



Design of Experiments and Big Data Analytics for Energy Efficient Buildings

Pragnesh Patel, Kwai Wong, Joshua New, Jibonananda Sanyal, George
Ostrouchov and Mark Adams
NICS, UTK, ORNL

Outline

- **Background**
- **Simulation & Experimental Design**
- **Statistical techniques for sensitivity analysis**
- **Analysis results**
- **Workflow system**
- **Conclusion**
- **Future work**

Energy is the Defining challenge of our time

- **Building in U.S.**
 - 40% of Primary energy/carbon
 - 73% of electricity, 34% of gas
- **Building in China.**
 - 60% of urban building floor space in 2030 has yet to be built
- **Building in India.**
 - 67% of all building floor space in 2030 has yet to be built

**Global energy
consumption will
increase 50% by
2030**

Goals of ECSS project

- **Port EnergyPlus simulation code on XSEDE resources**
- **Integration of EnergyPlus into openDIEL to allow to manage and run multiple instances of EnergyPlus simulation.**
- **Implementation of experimental design framework**
- **Implementation of sensitivity analysis tool**
- **Fast and portable workflow solution that leverages existing workflow mechanisms**

EnergyPlus

- It is a whole building energy simulation program that engineers, architects, and researchers use to model both energy consumption—for heating, cooling, ventilation, lighting, and plug and process loads—and water use in buildings.
- It contains many of the recent building dynamic simulation algorithms.
- It has been transitioning to C(~750,000 lines) from FORTRAN.
- It provides large number of built-in HVAC and lighting control strategies and an extensible runtime scripting system for user-defined control.
- Standard summary and detailed output reports as well as user definable reports with selectable time-resolution from annual to sub-hourly, all with energy source multipliers.
- <https://energyplus.net>

Inputs

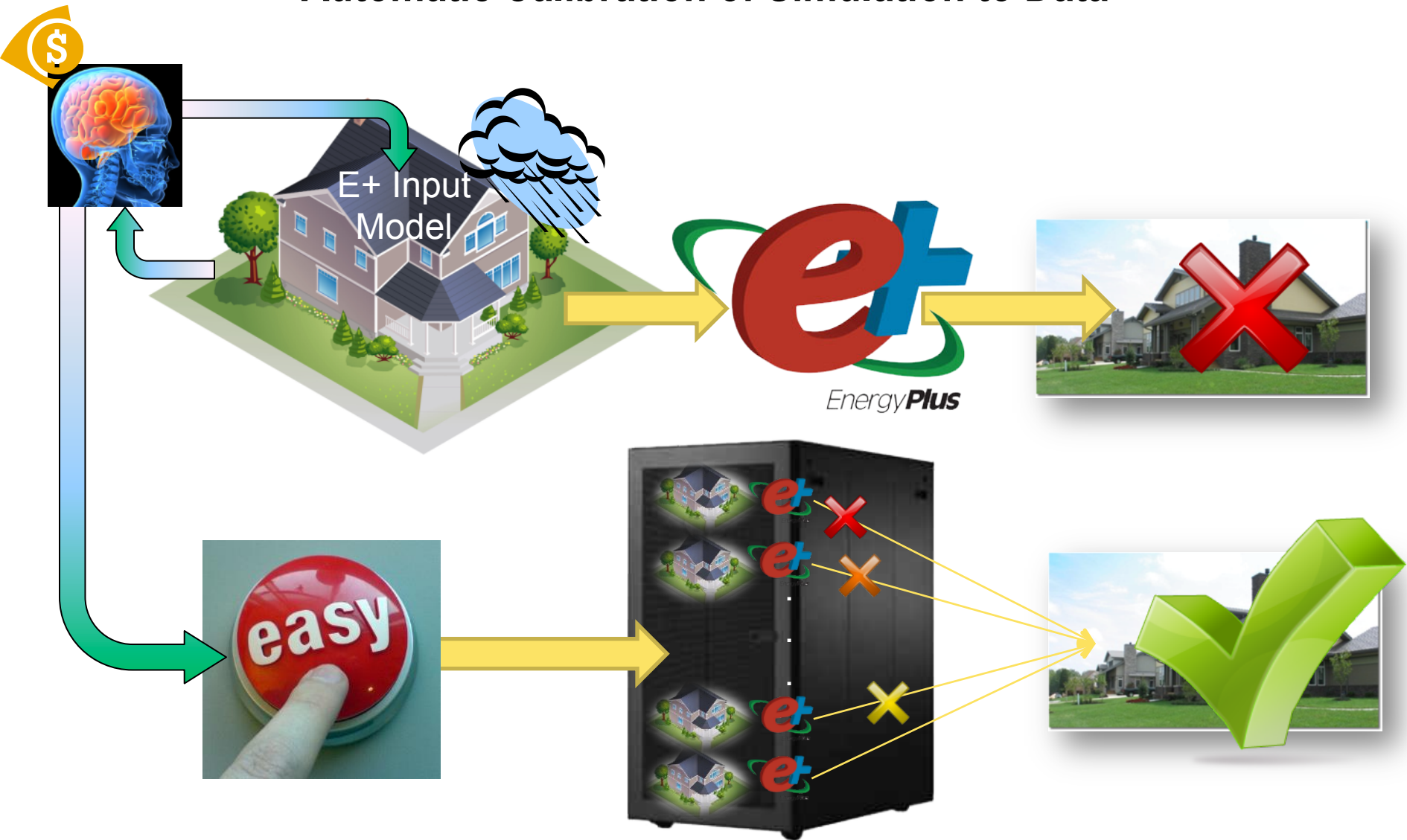
- **Standard input file to describe a building has on the order of 3,000 lines.**
- **Domain experts pick out the most useful parameters that could affect building consumption and meaningfully discretize values for those parameters.**
- **The most useful group consisted of 156 parameters.**
- **Need to turn on/off nobs (or Min/Max values)**
- **This resulted in GBs of data for 1 building type and 1 climate zone (e.g. Atlanta)**

Our motivation

- **Simulation plays a big role in understanding the behavior of building envelopes**
- **Parametric analysis determinants**
 - DOE 2.1 and EnergyPlus
 - OpenStudio's parametric tool
- **Uncertainty in two types of buildings**
 - Building envelope's thermal conductivity properties
 - 156 parameters
 - Sensitivity analysis of stand-alone retail building
 - 20 parameters
 - Fractional factorial design
- **16 different building types and 16 types of climate zones.**
- **Datasize depends on selection of Hourly, Daily, Weekly, Monthly parameters**

Autotune

Automatic Calibration of Simulation to Data



Experimental setup 1

- **ZEBRAIliane homes and Markov Order Design**
- **Home has energy efficient technologies**
 - Standing seam metal roof with infrared reflective pigments to boost solar reflectance
 - ENERGY STAR appliances
 - Triple-pane low emittance Argon-filled windows
 - Compact fluorescent lighting
 - Horizontal ground loop installation that leverages foundation and utility excavations
 - High-efficiency water-to-air heat pump for space conditioning
 - High-efficiency water-to-water heat pump for hot water heating
 - Emulated occupancy

Markov sampling

- ZEBRAAlliane homes and Markov Order Design
- 156 parameters determined by experts
- Brute force
 - All possible combinations of min and max sizes (~12,000 simulations, 143 GB of data)
- Markov Order 1
 - 299 Simulations
 - 3.9 GB
- Markov Order 2
 - 28,000 simulations approximately
 - 450 GB

Experimental setup 2

- DOE's reference Stand-Alone Retail building
- Fractional Factorial Design
- Factors and responses
 - Input parameters and output variables
- Effect sizes of factors
- Fractional design expression

$$L_{k-p}$$

- L: number of factor (parameter)
- k: number of factors (parameters) investigated
- p: size of fraction of full factorial used (degree of confoundedness)

Fractional factorial designs

- Sparsity of effects principle
- Resolution of fractional factorial designs
 - Ability to separate main effects and low-order interactions from one another
 - Most common are III, IV and V
 - Resolution IV: Estimate main effects unconfounded by two-factor interactions, two-factor interaction effects may be confounded with other two-factor interactions
- Regular design, power of 2
- Non-regular designs are design where run size is a multiple of 4, which introduces partial aliasing

Fractional factorial designs

- **20 factors and resolution VI design**
 - 1024 simulation runs
 - Resolution VI designs alias main effects (single factors) with order five or higher interactions and alias two-factor interactions with order four or higher interactions
- **Generation: FrF2 package**
- **For more than 2 levels, response surface methodology is useful**
 - 2nd degree polynomial
 - Identify factors affecting response variable
 - Follow-up with more complicated designs

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	0.00023 1	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	0.00023 1	0.000429
	Point_Of_Sale_Infil	ZF_POS	Flow per Ext Surface Area	0.00033	0.00023 1	0.000429
DsgnSpec: OutdrAir	SZ DSOA	DS_BaSp	Outdr Airflow per Area	0.0008	0.00056	0.00104
	Back_Space					
	SZ DSOA Core_Retail	DS_CoRt	Outdr Airflow per Area	0.00175	0.00122 5	0.002275
Sizing: Parameters	Sizing:Parameters	Sz_Heat	Heating Sizing Factor	1.25	0.875	1.625

Variables

Reference: http://web.eecs.utk.edu/~new/presentations/2014_Purdue_Uncertainty.pdf

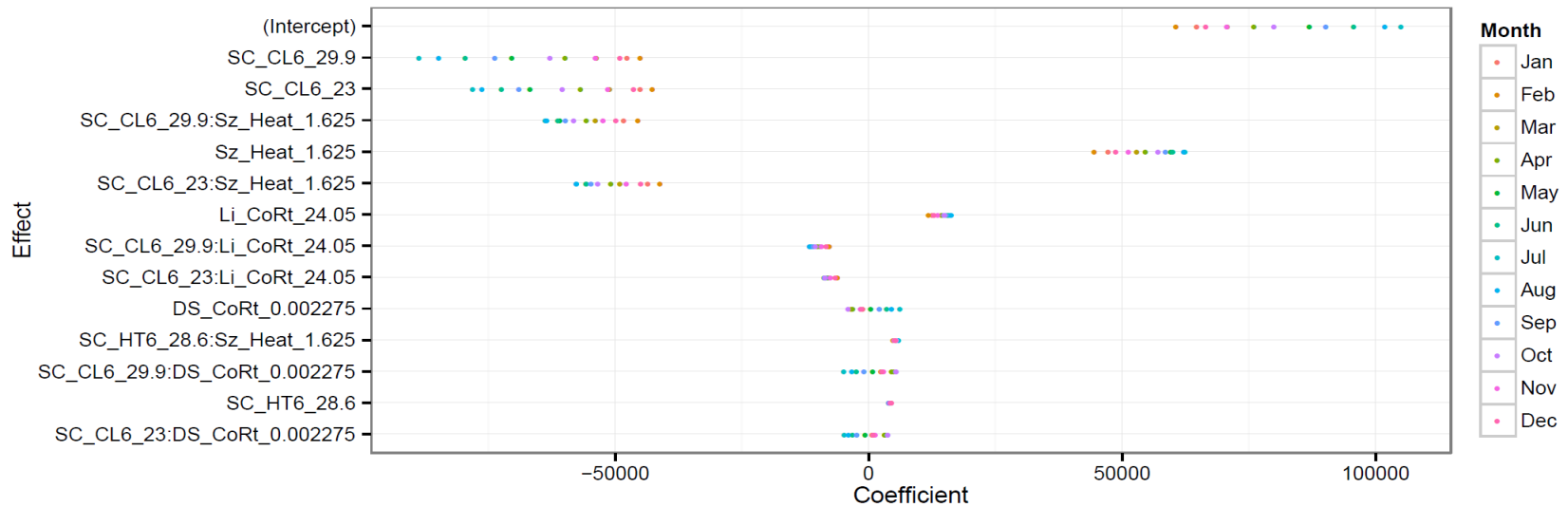
Sensitivity analysis

- **Analysis of variance**
 - Model containing main effects and all two factor interactions
- **Simulation time of 1 year and use of monthly totals in the analysis**
 - 12 measurements used in a repeated measures multivariate analysis
 - More powerful than treating each month as separate analysis of variance.

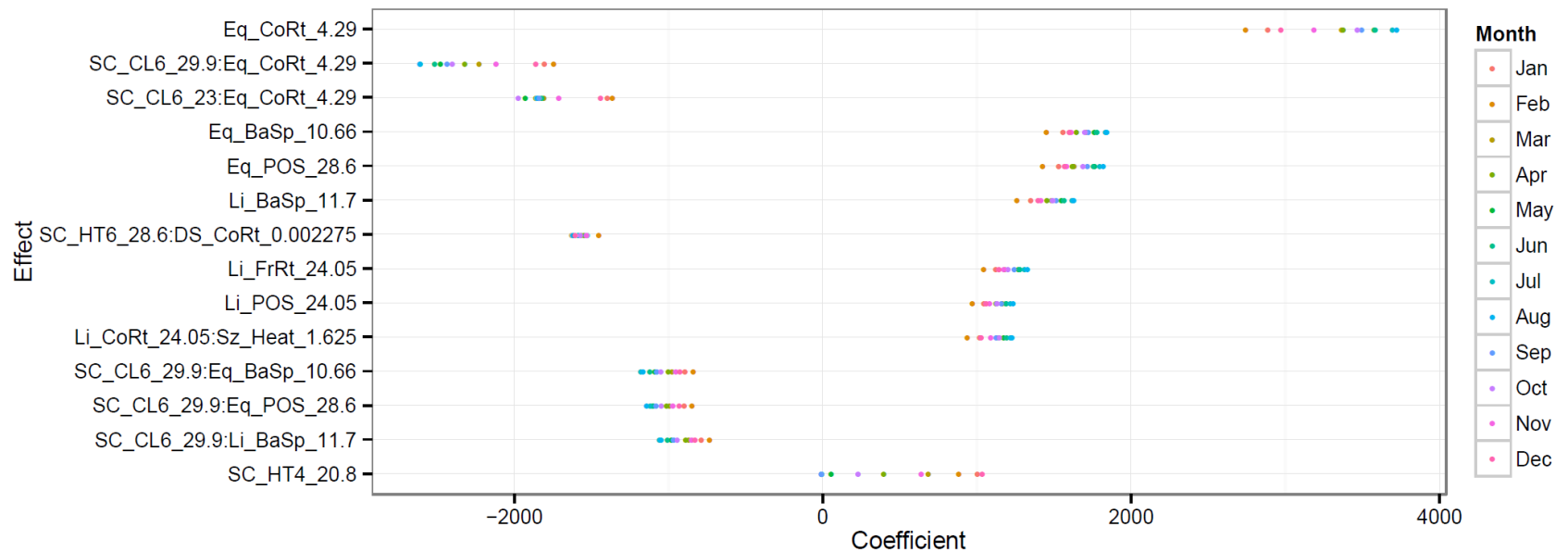
Sensitivity analysis

- 20 factors, all 2-way interactions model fit to 1024 observations, we still have 778 degrees of freedom left for the error term
- Strong significance with Pillai statistic (below 0.0001)
 - All factors except
 - LightsFront Entry
 - Zone infiltration flow rate: Air changes per hour, Constant term coefficient
- 380 possible 2 way interactions
 - 58 were significant at <0.01
 - Effect size is important
 - Charts terms with over 1,000 KWh effect

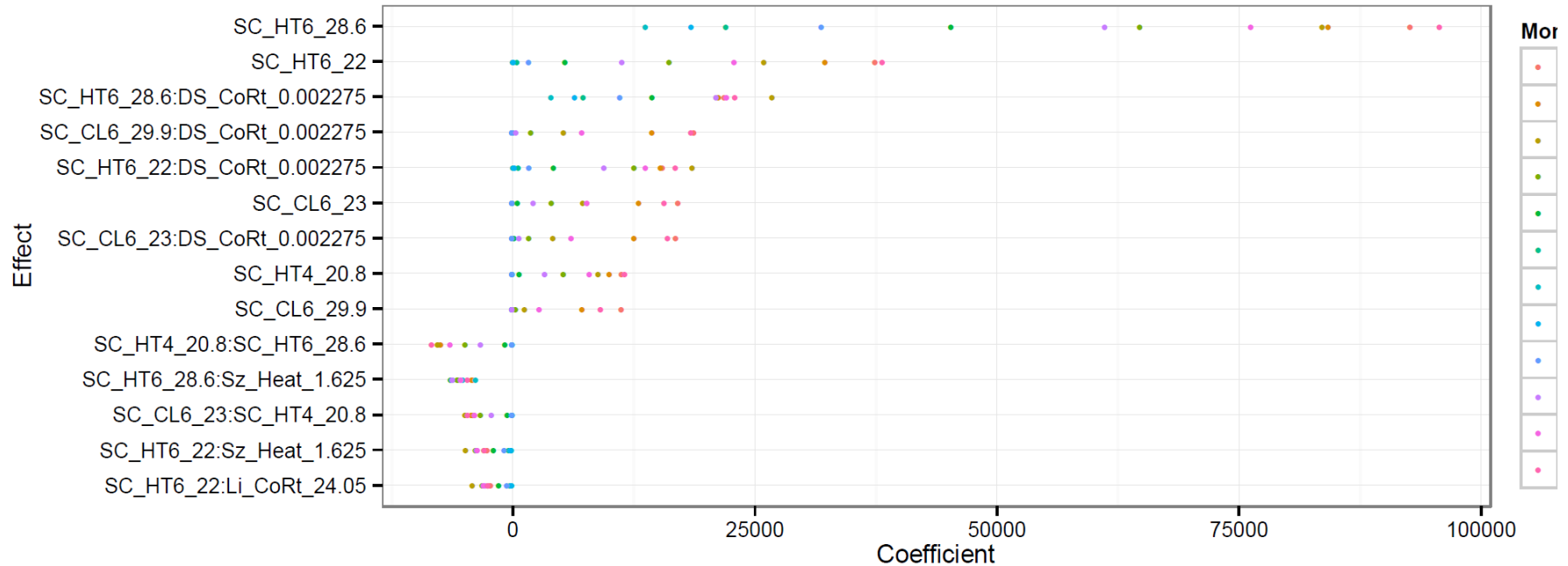
Factors with largest effects on electricity use



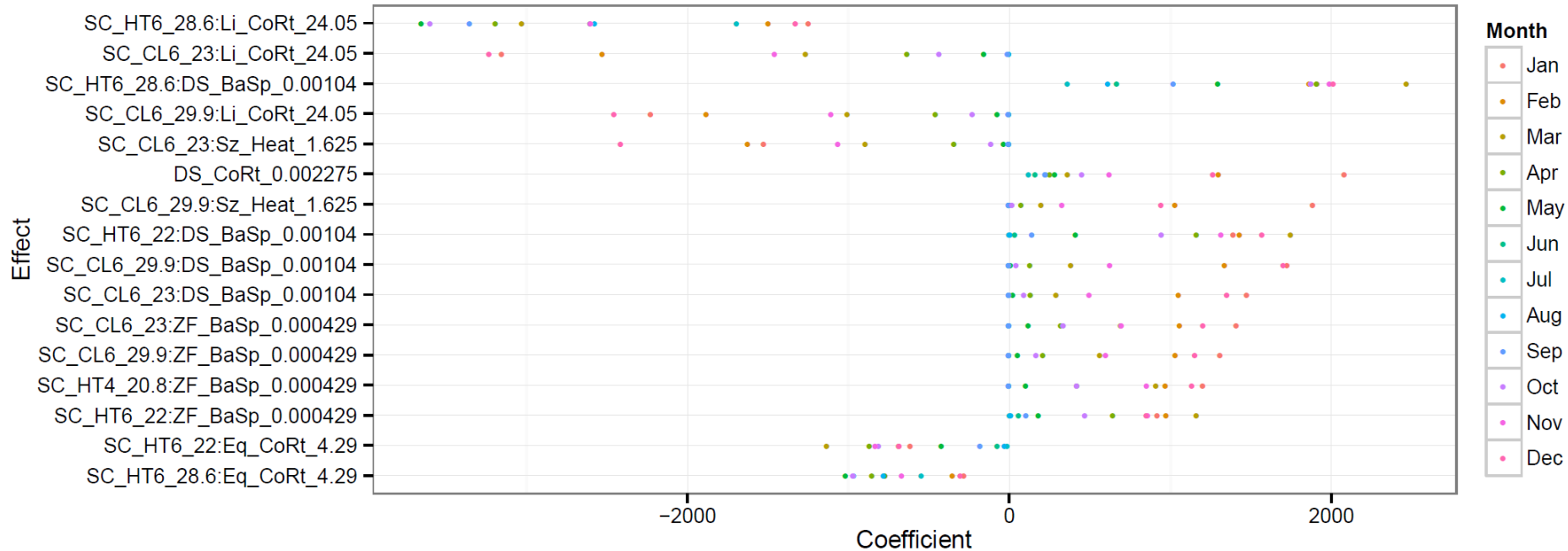
Factors with largest effects on electricity use(2)



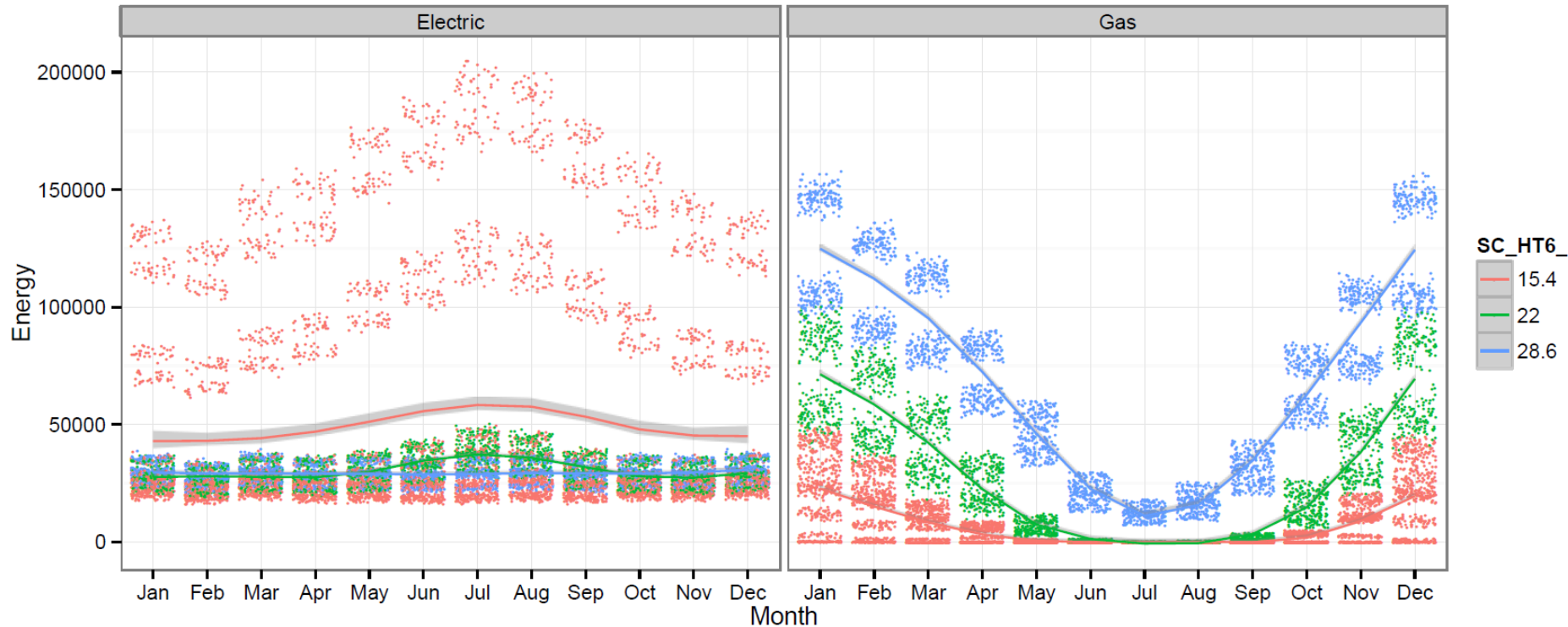
Factors with largest effects on gas use



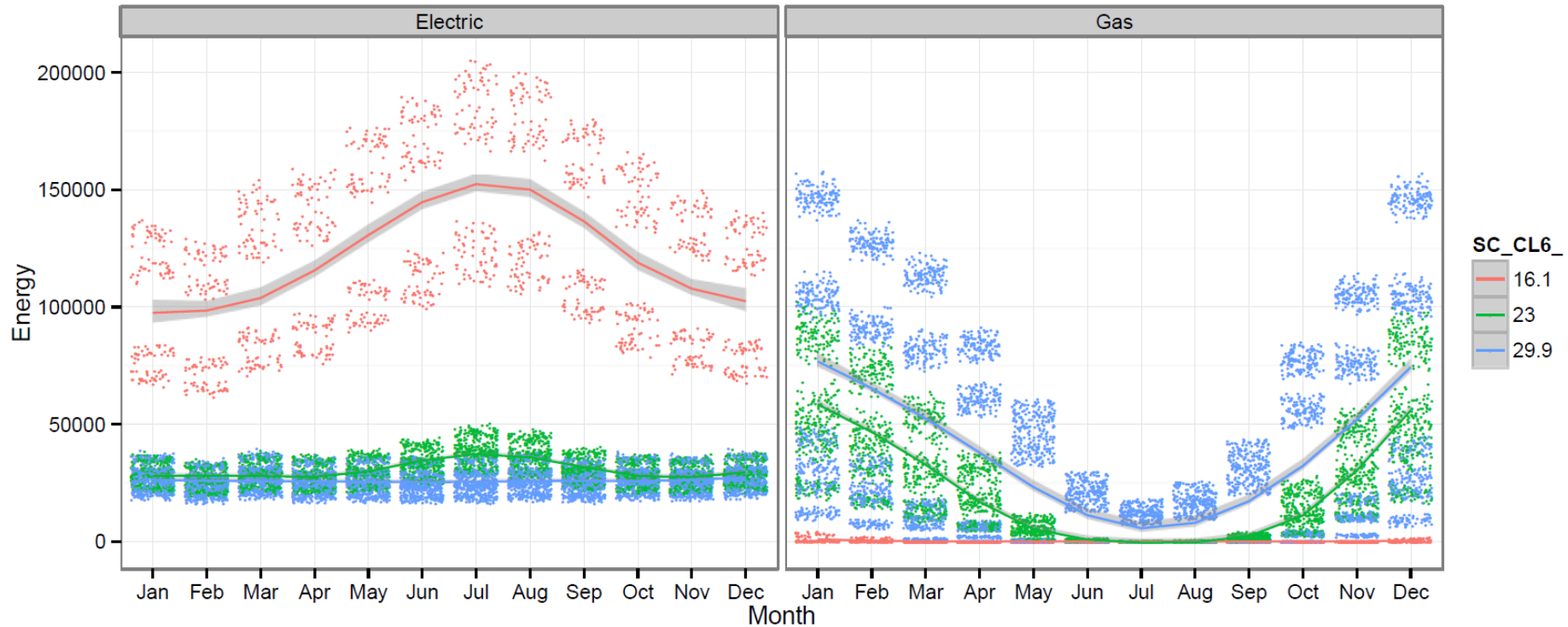
Factors with largest effects on gas use



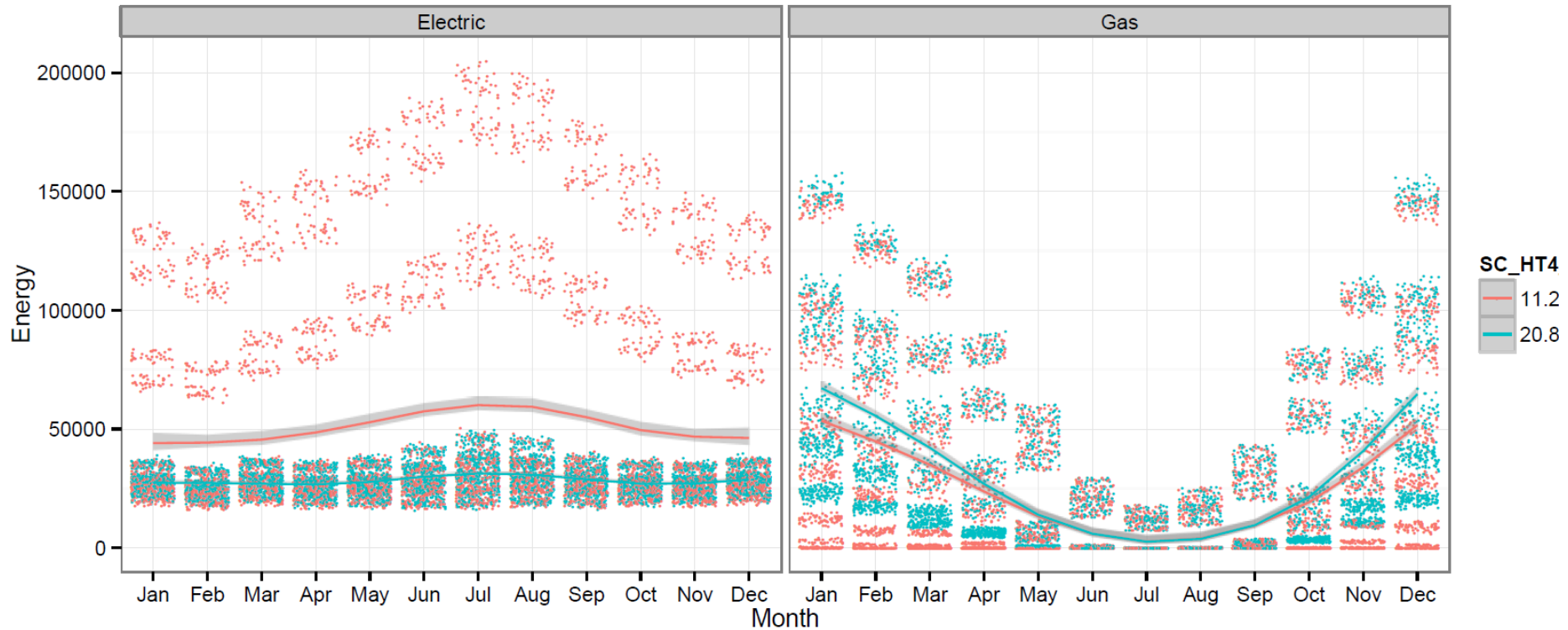
Monthly Electric and Gas energy use



Monthly Electric and Gas energy use



Monthly Electric and Gas energy use



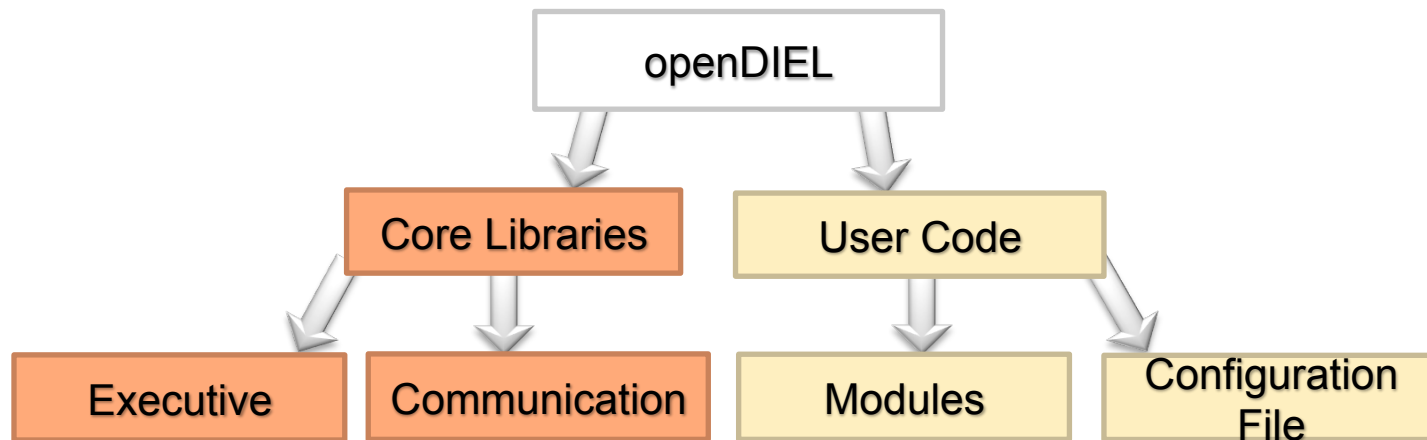
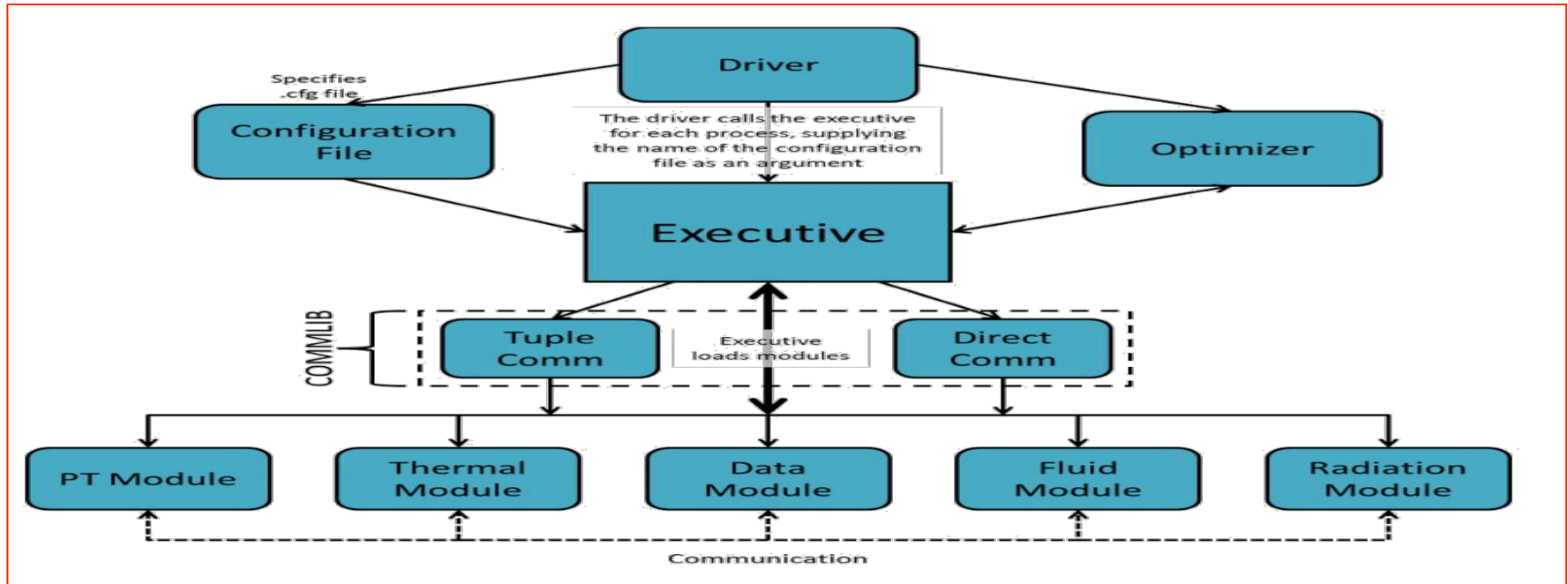
Conclusion

- **Sampling of 156 parameters for residential building**
 - More time and data
- **Sampling of 20 parameters on stand-alone retail**
 - Fractional factorial
 - Analysis of variance
- **Insight**
 - Practical issue in experimental design
 - Analytical approaches in understanding thermal properties of building envelopes
- **Future work**
 - Parallel analysis using Programming with Big Data in *R* (*pbdr*)
 - For All 16 types of building and 16 climate zones
 - *In Situ* Analysis

openDIEL

- **openDIEL (open Distributive Interoperable Executive Library) aims to facilitate communication between user-created loosely coupled simulations.**
- **Loosely coupled simulations are mostly serial programs that rely on data points from other simulations; these simulations get their input from and send output to other simulations.**
- **The openDIEL communicates through a tuple server-based method, and now is able to communicate large chunks of contiguous memory more efficiently through direct communication.**
- **The organization of these simulations is organized and facilitated by the workflow implementation of the openDIEL.**

openDIEL organization



Direct Communication

- **Wrappers around MPI to communicate large chunks of data**
- **Functions IEL_send (nonblocking), IEL_recv (blocking) and IEL_move (blocking)**
- **Uses a set of shared boundary conditions in a conceptual grid, each process having access to a different section of this grid**
- **Movement parameters specify point-to-point transfer**

Direct Communication

```
modules = (
{
    function = "jacobi";
    args = ();
    libtype = "static";
    library = "libjacobi.a";
    size = 4;
    points = (
        (([0,16]), ([0,16]), ([0,0]), ([0,0]), ([0,0]),
         ([0,0])),
        (([0,16]), ([0,16]), ([0,16]), ([0,16]), ([0,0]),
         ([0,0])),
        (([0,0]), ([0,0]), ([0,16]), ([0,16]), ([0,16]),
         ([0,16])),
        (([0,0]), ([0,0]), ([0,0]), ([0,0]), ([0,16]),
         ([0,16]))
    );
}
)

movement = (
    (0, ([0, 16]), (0), (1)),
    (1, ([0, 16]), (1), (0)),
    (2, ([0, 16]), (1), (2)),
    (3, ([0, 16]), (2), (1)),
    (4, ([0, 16]), (2), (3)),
    (5, ([0, 16]), (3), (2))
)
```

Configuration file using direct communication

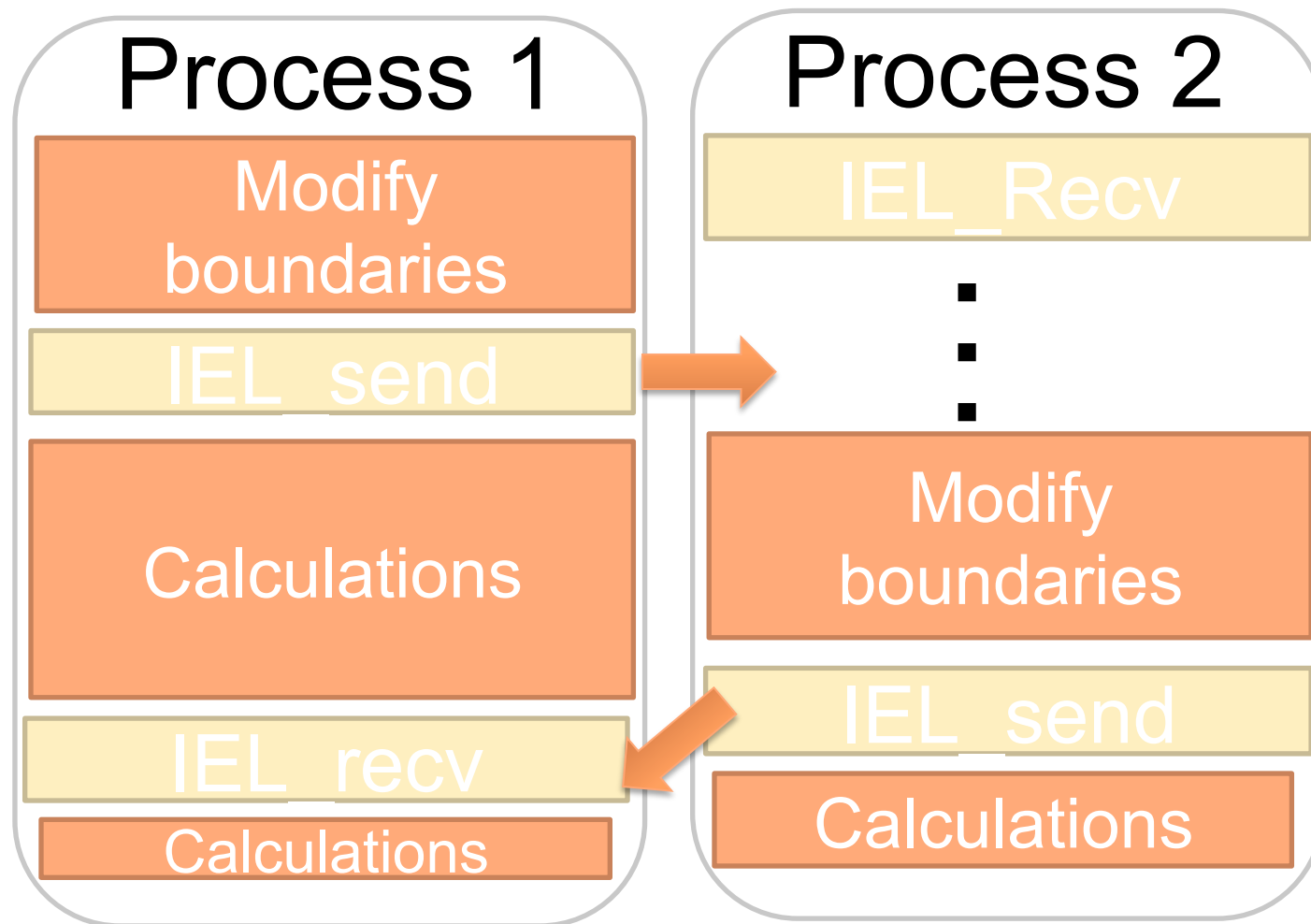
```
if( mype < npes-1 )
/* Send my data down to the processor below me; Only npes-1 do this */
    MPI_Send(&t[NRL][1], NC, MPI_FLOAT, mype+1, DOWN, MPI_COMM_WORLD);
if( mype != 0 )
/* Sending my data up to the processor above me ; Only npes-1 do this */
    MPI_Send(&t[1][1], NC, MPI_FLOAT, mype-1, UP, MPI_COMM_WORLD);
if( mype != 0 )
/* Receive new data from UP processor of any source */
    MPI_Recv(&t[0][1], NC, MPI_FLOAT, MPI_ANY_SOURCE, DOWN,
             MPI_COMM_WORLD, &status);
if( mype != npes-1 )
/* Receive new data from DOWN processor of any source */
    MPI_Recv(&t[NRL+1][1], NC, MPI_FLOAT, MPI_ANY_SOURCE, UP,
             MPI_COMM_WORLD, &status);
```

Example: MPI Communication Code

```
// Exchange data with the process below
if (myRank != NPES - 1)
    IEL_move(exec_info, myRank + 1);
// Exchange data with the process above
if (myRank != 0)
    IEL_move(exec_info, myRank - 1);
```

Example: IEL Communication Code

Direct Communication



openDIEL and Energyplus with Analysis

EnergyPlus – (Single Program Multiple Data)

EnergyPlus is a program developed by the DOE to model the power consumption of buildings. The openDIEL workflow engine is used to run EP7, Readvars, and an R analysis script(along with preprocessing code of simulation output) to perform statistical analysis on the dataset as a whole.

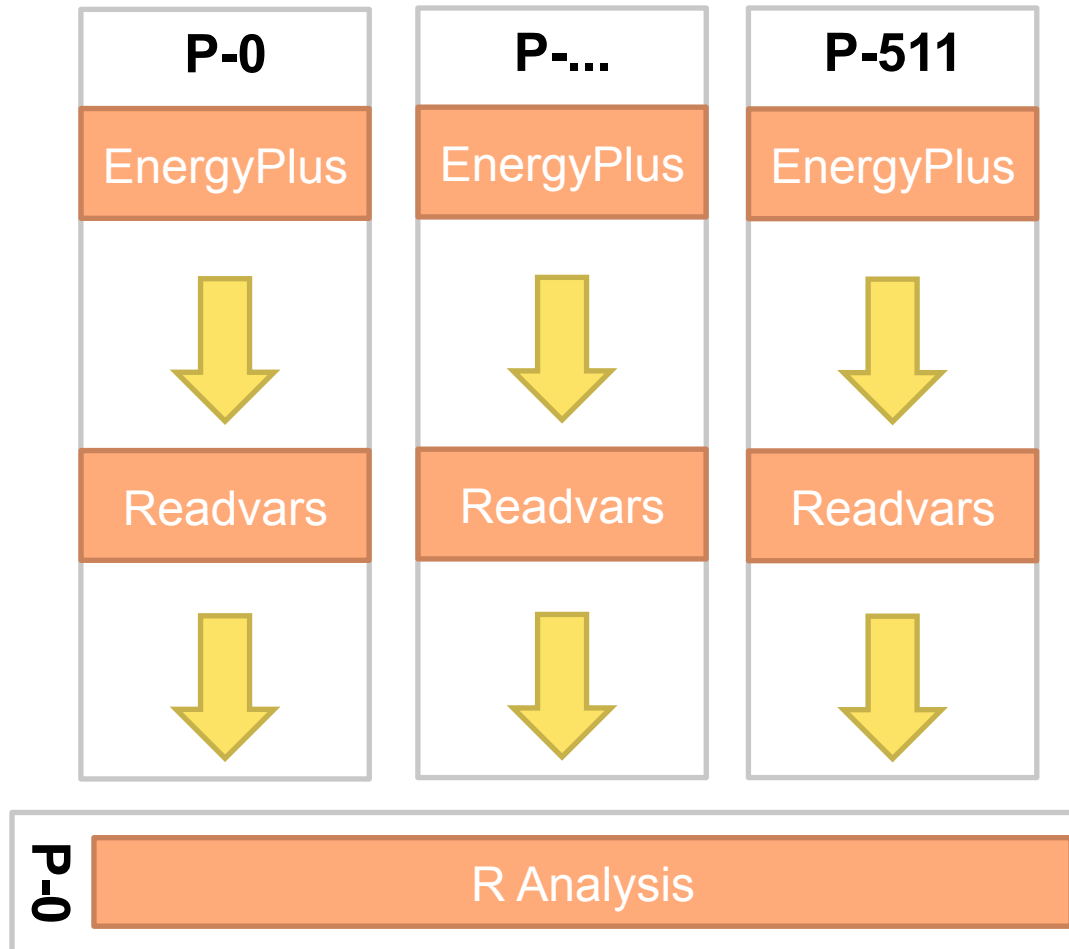
openDIEL and Energyplus with Analysis

One group is used for this case, containing EnergyPlus, Readvars, and R Analysis (a module that runs arbitrary R scripts).

Each EP7/Readvars module executes in a separate directory on a different set of input files.

Once EP7 is complete, Readvars extracts the relevant information from the output files.

After this data is collected, an R script is used to perform statistical analysis on the produced data.



openDIEL and Energyplus with Analysis

```
workflow:  
{
```

```
  groups:  
  {
```

```
    ep7-readvars:  
    {
```

```
      order=("modep7", "modreadvars")  
    }
```

```
    RAnalysis:  
    {
```

```
      order=("RAnalysis")  
    }
```

```
  }
```



openDIEL : Work in progress

Workflow/Tuple Server Control Module

The tuple server will be expanded to read a makefile-like list of dependencies and dynamically schedule tasks using the available processes.

ModMaker

The ModMaker toolset will be ported from bash to python and made available on the Python Package Index (PyPi) for an easy, one line installation.

Tuple Server Scaling

The tuple server will be expanded to allow multiple processes to operate as a single server (currently each processes operates separately).

openDIEL GUI

The openDIEL GUI will be extended to expose openDIEL features in a more user friendly package.

References

- http://web.eecs.utk.edu/~new/presentations/2014_Purdue_Uncertainty.pdf
- EnergyPlus: <https://energyplus.net>
- FrF2: <https://cran.r-project.org/web/packages/FrF2>
- OpenDIEL: <http://cfdlab.utk.edu/opendiel>
- R: <https://www.r-project.org>
- pbdR: <http://r-pbd.org>

**Thanks to Joshua New, Jibonanada
Sanyal, Kwai Wong, George
Ostrouchov and Mark Adams**