Predicting Energy Savings Using Inverse Modelling Simulation

Honey Berk
CUNY Building Performance Lab

James Hamski
CUNY School of Professional Studies

Dan Fanelli
CUNY School of Professional Studies

Sonya Hong
CUNY School of Professional Studies

Abstract

Inverse modelling was used to calculate energy consumption for a small sample of NYC office buildings. Model coefficients were then used to populate thermal energy balance equations for electricity-related end uses, and simulations were run to investigate retrofit and retro-commissioning scenarios designed to increase energy efficiency and lower consumption. Total energy consumption, cooling energy demand, cooling system efficiency and internal loads were calculated before and after simulations, and changes in these parameters were examined. Simulation using inverse modelling was found to correctly simulate real-world conditions, and could prove to be valuable for "back-of-the-envelope" calculations prior to complex physical energy modelling.

Keywords energy efficiency, inverse modelling, thermal energy balance, simulation

1. Keywords

energy efficiency, inverse modelling, thermal energy balance, simulation

2. Introduction

There were an estimated 5.6 million commercial buildings in the United States in 2012, comprising 87 billion square feet of floorspace. These buildings consumed a total of 6,963 trillion Btu in site energy, with electricity representing 61% of that total and natural gas, 32%. Heating, ventilation & cooling account for 44% of total energy used (U.S. EIA, 2012). Nationally, buildings account for about 40% of greenhouse gas (GHG) emissions; while, in New York City (NYC), buildings account for 73% of total GHG emissions through the use of natural gas, electricity, heating oil, steam, and biofuel (City of New York, 2014).

In 2007, then-NYC Mayor Bloomberg announced PlaNYC, with a bold proclamation that the City would cut GHG emissions 30% from 2005 levels by 2017. In 2014, newly-elected Mayor Bill de Blasio upped this goal to an 80% reduction in emissions by 2050 (80x50) in his 'One City, Built to Last' plan. Since then, various studies have been commissioned to figure out exactly how New York, the largest city in the country, with about 8.5 million residents, will be able accomplish such a drastic reduction in GHG emissions. One of the major targets for reduction is NYC's 4,600-building municipal portfolio, which includes about 120 office buildings, 25 courthouses, 23 hospitals, 300 firehouses, 1,200 public school facilities, hundreds of garage and repair service facilities, and many more facilities of multiple typologies, across more than 40 agencies. In light of the sheer number of municipal buildings, there is a tremendous potential for GHG mitigation in the NYC portfolio through capital equipment replacement and changes in operations and maintenance practices, commonly known as retrofit and retro-commissioning projects.

For the last five years, the CUNY Institute for Urban Systems Building Performance Lab (BPL) has been working with the NYC Department of Citywide Administrative Services Energy Management division (DCAS DEM) to create baseline energy models for municipal facilities; to date, nearly 1,000 facilities have been modelled. BPL uses an industry-standard methodology that involves generating linear change-point regression models with monthly energy consumption data as the dependent variable and monthly average outdoor air temperature as the independent variable.

The current work replicates the change-point modelling procedure in the R computing environment, then uses the resulting model coefficients and equations for a simplified thermal energy balance model to simulate energy efficiency measures and project associated energy savings. The thermal balance principle considers the whole building as an entity with a number of sources and sinks. The thermal balance is the net amount of all gains and losses in the building, and the building's energy efficiency is defined through its thermal balance. (Anecdotally, it was noted that the use of sources and sinks corresponds to the design of discrete event simulations.)

This work builds on a methodology described in a 2011 paper co-authored by one of the pre-eminent researchers in this area, John Kelly Kissock, Ph.D, P.E., Professor and Chair of the Department of Mechanical and Aerospace Engineering, University of Dayton. The paper in question, "Estimating Industrial Building Energy Savings using Inverse Simulation" (Sever and Kissock, 2011) will be discussed, along with other influential works, in the Literature Review, below.

3. Literature Review

Linear change-point regression modelling has been used in the analysis of energy consumption data for more than two decades. The method was officially introduced to the industry in the early 2000s when ASHRAE, the leading industry association for heating, ventilating, air conditioning and refrigeration professionals, initiated a research project on the topic, which culminated in the development of The Inverse Modeling Toolkit (IMT) (Kissock et al, 2003). In the IMT and in subsequent research, a range of techniques has been used to predict energy consumption in commercial and industrial facilities using historical weather data (Zhang, 2013; Carpenter, 2010), as energy consumption in buildings can be estimated as a function of outdoor air temperature (Sever and Kissock, 2011).

Change-point models are also used in measurement and verification (M&V) of savings from energy efficiency projects, in accordance with industry-standard M&V protocols (EVO, 2016; ASHRAE, 2014). The M&V process involves creating baseline models of pre-retrofit energy usage, then estimating savings by comparing actual post-retrofit usage with adjusted baseline usage; the adjusted baseline is usage that would have occurred had the retrofit not been implemented, but normalized to the actual weather of the post-retrofit period.

Retrofit and large retro-commissioning projects are typically planned following recommendations from an energy audit report. These recommendations commonly include a series of energy conservation measures (ECMs), such as a lighting upgrade or boiler replacement, for which energy savings have been projected using engineering calculations and/or simulations generated using energy modelling software packages (e.g., eQUEST, EnergyPlus). Estimating energy savings in this way is usually a time intensive process, requiring a building simulation model to be developed then run with actual weather data; further, the model must be calibrated to actual historical energy consumption data, and then modified to include changes before it is run again. The process involves a significant number of assumptions and is fraught with calibration errors, as a result (Sever and Kissock, 2011).

The current work is based on an inverse simulation methodology that uses actual historical energy data (Sever and Kissock, 2011). The authors first created baseline models using linear change-point regression analysis. Next, they altered model parameters to simulate changes in the building due to ECMs (e.g., installing highly-insulated triple pane windows). To estimate resulting energy savings, the authors then compared the baseline model with the model generated after the parameters were changed. In comparing results from inverse simulation to energy simulation for the same projects, the researchers found that normalized annual savings were predicted within $\pm 1\%$ of simulated savings.

The inverse simulation methodology was found to be as robust as energy modelling and, notably, allowed for the elimination of the time-intensive model calibration because the model uses actual historical energy data. There was, however, associated time-scale error and phase-shift error found that reduced precision (Sever and Kissock, 2011).

Building Name	Year Built	Gross Floor Area (sq. ft.)
EXC	1911	59,000
LIC	1874	59,300
MFC	1975	481,000
MSC	1901	212,500
SUN	1845	242,062

Table 1. Basic information on five sample buildings

4. Methodology

4.1. Data Sources and Preparation

4.1.1. Building Energy Consumption Data

Raw data were obtained from DCAS DEM for five NYC municipal office buildings. Each data file contained meter-level monthly energy consumption data for electricity and thermal energy consumption; thermal energy consisted of either natural gas, steam or both. Consumption for electricity was presented in kilowatt hours (kWh), while thermal energy consumption was converted into British Thermal Units (Btu) from natural gas therms and/or thousand pounds (kLbs) of steam.

4.1.2. Weather Data

Daily average temperature data was obtained from the Average Daily Temperature Archive at The University of Dayton. The file contained data for the LaGuardia Airport weather station in Queens, NY for January 1, 1995 to present; source data are from the National Climatic Data Center. In lieu of monthly average outdoor air temperatures, similar research has used heating and cooling degree days (HDD, CDD) as the independent variable; these are measures of the number of days in the period multiplied by the sum of the number of degrees reached above (for CDD, or below, for HDD) a set balance point temperature (typically, 65°F) for those days. There is some evidence to suggest that imposing an artificially selected balance point temperature introduces bias into the change-point model; so outdoor air temperature was selected for this work, not degree days.

4.1.3. Data Preparation

The building data obtained from DCAS DEM contained meter-level energy consumption data for the last four fiscal years (July 2012-June 2016). Data preparation for each facility included selecting the proper 24-month time period (July 2014-June 2016), aggregating multiple meters for each energy type, and converting natural gas and steam data to Btu. A CSV file was then prepared for each facility, containing the start and end dates for each billing period, and a column each for electricity and thermal energy consumption data. To simplify the analysis for this work, the billing periods were assumed to be an entire month; in reality, this is often not the case, as utility meter reading dates vary from month-to-month. Future work will include actual billing period start and end dates, for increased precision.

4.2. Linear Change-Point Regression Models

4.2.1. Background

When analyzing building energy consumption, two major patterns of usage emerge: weather-independent (baseload) and weather-dependent. With electricity, baseload energy represents end-uses that are relatively constant, such as

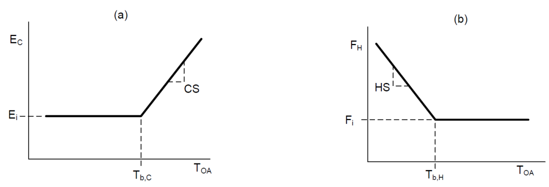

Figure 1. 3-parameter heating (3PH) (l.) and cooling (3PC) (r.) models (Kissock, 2011)

plug loads from computers and office equipment, lighting, and fans used for ventilation; with natural gas or steam, a baseload usually represents energy used for domestic hot water production, or services such as cooking or laundry. In contrast, weather-dependent (also known as weather-sensitive) energy represents the amount consumed for cooling and heating end uses, and so it follows that this consumption varies with outdoor air temperature. It is important to note here that weather-sensitive energy consumption is seen only in climates where there is a need for mechanical heating and/or cooling; this is an intuitive concept that is sometimes overlooked.

Based upon this natural dichotomy of building energy consumption in climates such as that of NYC, linear change-point regression is used as an industry standard for modelling. In a building energy model, the change-point represents the balance point temperature, which is the temperature above which energy is used for cooling and below which energy is used for heating. In this way, change-point models are much better representations of a building than standard straight-line regression models, as they are illustrative of the physical conditions of the building.

To wit, in addition to the change-point, the model's y-intercept represents the building's baseload – the energy consumed by weather-independent end uses; and the slope of the weather-sensitive line segment is an indication of the efficiency of the cooling or heating system. These three factors represent a 3-parameter change-point model – either heating (3PH) or cooling (3PC), as seen in the figure below. Four- and 5-parameter change-point models add a second sloped line segment and second change-point, respectively; those models types will not be addressed in this paper.

4.2.2. Modelling

Using the prepared data files, a change-point analysis was performed on both electricity and thermal energy consumption data using the R package changepoint (Killick and Eckley, 2014). The change-point analysis utilized the AMOC (At Most One Change) method to specify a single change-point for each dataset, with no loss function. For electricity, the change-point was determined using the mean of the electricity load data. Given the wide range of electricity loads over the observed temperature range, the change in mean at warmer temperatures leads to an identifiable change-point. For fuel, both mean and variance changes were incorporated to determine the change-point, as at warmer temperatures there is very little change in the fuel load (i.e., the baseline load variance is minimal, but the heating load variance is significant). It should be noted that the changepoint package did not allow for selection based upon other metrics, such as root mean square error (RMSE), which is used more often in building energy analysis. This issue will need to be addressed in future work, likely with manual piecewise linear regression in lieu of the use of an R package.

After the change-points were determined, a linear regression model was fit to the electricity data points at temperatures greater than or equal to the change-point; for thermal load, the linear regression model was fit to data points at less than or equal to the change-point. These two linear models represent the observed response rates of energy consumption to changes in average monthly outside air temperature. That is to say, they are equal to the amount of energy used (kWh or Btu) or cooling or heating per degree Fahrenheit. Change-point and baseload values for the electricity and fuel models for the five buildings can be found in the Results section, below.

After some exploratory analysis of the electricity and thermal energy models for the five buildings, the electricity model for the LIC facility was selected for further investigation, as it looked to be the best 3-parameter cooling model out of the sample.

4.2.3. Regression and Thermal Energy Balance Equations

As per the inverse model simulation methodology described earlier, the coefficients for the electricity and fuel regression equations were used, along with a number of assumptions and calculated values, to build out thermal energy balance equations. In effect, these equations disaggregate energy consumption into physical building parameters. These equations, presented below, were solved and then used as the basis for a series of energy efficiency scenario simulations. A table of definitions of terms follows the equations.

3PC Change-point Model:

$$E = E_i + CS * (T_{OA} - T_{b,C})^+ \tag{1}$$

3PH Change-point Model:

$$F = F_i + HS * (T_{b,H} - T_{OA})^+ \tag{2}$$

Cooling Coefficient:

$$CC = UA + (\dot{V} * n * \rho * cp) \tag{3}$$

Heating Coefficient:

$$HC = UA + (\dot{V} * n * \rho * cp) \tag{4}$$

Cooling Slope:

$$CS = \frac{CC}{\eta_C} \tag{5}$$

Heating Slope:

$$HS = \frac{HC}{\eta_H} \tag{6}$$

Cooling Change-point:

$$T_{b,C} = T_{set} - \frac{Qi}{CC} \tag{7}$$

Heating Change-point:

$$T_{b,H} = T_{set} - \frac{Qi}{HC} \tag{8}$$

Note that cooling coefficient and heating coefficient terms represent cooling and heating demand in the building, which is essentially the amount of space conditioning needed to keep the interior temperature at the thermostat setpoint temperature.

Total energy, baseload, change-point and slope were determined by the regression equation; density and specific heat of air are constant terms; surface area and volume were

Term	Definition
E	Total Energy Consumption
F	Total Fuel Consumption
E_i	Electric Baseload
F_i	Fuel Baseload
CS	Cooling Slope
HS	Heating Slope
T_{OA}	Outdoor Air Temperature
$T_{b,C}$	Cooling Change-point
$T_{b,H}$	Heating Change-point
CC	Cooling Coefficient
HC	Heating Coefficient
U	U-Value
A	Surface Area
$\dot{V}$	Volume
n	Air Changes Per Hour
ρ	Density of Air
cp	Specific Heat of Air
η_C	Cooling System Efficiency
η_H	Heating System Efficiency
T_{set}	Thermostat Setpoint
Q_i	Internal Loads

Table 2. Defintion of model terms

determined by the geometry of the building; and the remaining terms were either assumptions, or were calculated based upon assumptions. It should be noted that the equations used in the current work differed from those of the work being replicated in that the "n" term was added to represent the number of air changes per hour in a building, which was then multiplied by the total volume of conditioned space ($\dot{V}$). The reference article seems to have considered that term to be the net product of those terms (Sever and Kissock, 2011).

In the course of this work, research into the thermal energy balance equation methodology revealed that the concept is the subject of an international standard, ISO 13790:2008 (ISO, 2008), which provides extremely detailed calculation methods for assessment of the annual energy use for space heating and cooling of a building. This method includes the calculation of factors such as: heat transfer by transmission and ventilation; contribution of internal and solar heat gains to the building heat balance; the annual energy needs for heating and cooling required to maintain specified setpoint; and the annual energy used for heating and cooling end uses.

A cursory review of the ISO standard revealed that many of the terms used in the energy balance equation in the current work are highly simplified representations of what are included in much greater detail in the Standard. As an example, Q_i represents internal loads in the simplified equation, but is really composed of the loads from multiple sources, such as solar gain (Q_{solar}) and gains from occupants (Q_{people}). Further, envelope conductance (U-value) can be broken down into conductance of walls, windows, roof, ground and thermal bridges. Future work will involve expanding the simplified equations for a more detailed and precise analysis.

4.3. Simulation

Once the change-point, y-intercept and slope were determined using the change-point and linear regression procedures, and the constants were set, each of the remaining equation terms (as described above) were either set via assumption or calculated using the energy balance equations. It took a number of trials to determine the best assumptions so that projected total energy most closely matched regression equation results. In addition, there were calculations required to correct the units to be used for each term, as many standard building terms are in SI units (e.g., u-value is measured in watts per square meter kelvin $[W/m_2K]$); ultimately, there was a mix of SI and imperial units used, with a conversion of results back to imperial units where necessary.

With the basic terms set for each equation, a series of four simulations was performed, each of which represented a change in a physical building parameter or other related factor. As mentioned earlier, the simulations were run for one specific facility, LIC, and for electricity data only. The simulations performed were as follows:

4.3.1. Simulation 1: Outdoor Air Temperature

Method: Used change-point model equation, simulated outdoor air temperature from 10-100°F in 5°F increments. All other parameters were kept constant.

Purpose: This simulation was used to predict total electricity usage at a full range of outdoor air temperatures. This exercise served as a verification of the change-point model results, with predicted temperatures that fit with the regression equation.

4.3.2. Simulation 2: Thermostat Setpoint

Method: Substituted different values for thermostat setpoint term (T_{set}) in cooling change-point equation. Simulated setpoint temperatures ranged from 50°F to 75°F, in 5°F increments. All other parameters were kept constant.

Purpose: This simulation was used to predict changes in building parameters when the thermostat setpoint temperature is raised or lowered, thereby allowing for insight into the efficacy of setpoint changes as an energy efficiency measure.

4.3.3. Simulation 3: Thermal Conductance

Method: Substituted different values for u-value term (U) in cooling coefficient equation. Simulated u-values ranged from 0.6 to 2, in 0.2 increments. All other parameters were kept constant.

Purpose: This simulation was used to predict changes in building parameters when thermal conductance is increased or decreased. An example of this in an actual building would be tightening up the building envelope by adding more insulation into the walls, or by installing highly-insulated triple pane windows.

4.3.4. Simulation 4: Air Changes

Method: Substituted different values for the air changes per hour term (n) in cooling coefficient equation. Simulated number of air changes ranged from 1 to 3, in 0.5 increments. All other parameters were kept constant.

Purpose: This simulation was used to predict changes in building parameters when the number of air changes per hour were increased or decreased. This term is a measure of the air volume added to or removed from a space. Typically, air in a space is neither uniform or perfectly mixed, and the actual percentage of a space's air which is exchanged in a period depends on the airflow efficiency of the space and the way in which it is ventilated. As such, changing the number of air changes per hour would represent a change in the way the space was ventilated.

Findings for each of the building simulations run can be found in the Results section, below.

4.3.5. Simulation 5: Weather Sampling

In addition to building simulations, a number of weather sampling simulations were conducted, with the goal of simulating a typical meteorological year (TMY), similar to those used for industry standard building simulation. TMY datasets are a collation of selected weather data for a specific location, generated from a data bank much longer than a year in duration. TMY data files represent a full range of weather phenomena for the location in question, while still giving annual averages that are consistent with the long-term averages for the location in question. The National Renewable Energy Laboratory maintains the National Solar Radiation Database, which contains solar and meteorological data for 1,454 U.S. locations and territories (NSRDB, 2016).

For the current work, a single temperature per month was required, but that single temperature can be chosen in a variety of ways. Strategies might include the following:

- Choose mean temperature for the specified month and year
- Choose maximum temperature for the specified month and year
- Choose minimum temperature for the specified month and year
- Choose maximum temperature for the specified month for ALL years

Anecdotally, it is interesting to note that there has been discussion in the industry regarding the use of more recent TMY data (i.e., 10 years vs. 30 years), since the climate has changed so dramatically that 30-year TMY data is likely less longer applicable to today's changed climate conditions.

One additional weather simulation generated random sampled temperatures by month, using a strategy of sampling a random day per month for the specified year. Sample output can be found in the Results section, below.

5. Results

5.1. Change-point Analysis and Linear Regression

Following are the change-point and baseload values for the electricity and fuel models for the five buildings:

	EXC	LIC	MFC	MSC	SUN
ELECTRICITY					
Baseload (MWh)	47.6	25.2	430.4	215.3	257.6
Change-point (°F)	62.76	62.76	66.14	66.14	66.14
FUEL					
Baseload (MMBtu)	876.4	185.9	992.2	578.4	4.776
Change-point (°F)	62.76	66.14	52.58	58.06	70.53

5.2. Change-point Model Visualization

The ggplot2 plotting system for R was used to visualize the change-point model equations generated in R by the changepoint package, as follows:

5.2.1. Building Simulations

Simulation 1: Outdoor Air Temperature Total electricity results were as expected, and served to validate the original model. As can be seen in the table below, total electricity is equal to the baseload value (25,200 kWh) until the outdoor air temperature (OAT)exceeds the change-point temperature; at that point, weather-dependent energy consumption begins and total electricity increases as OAT increases (positive correlation). Cooling demand (cooling coefficient), cooling efficiency and internal loads stay constant.

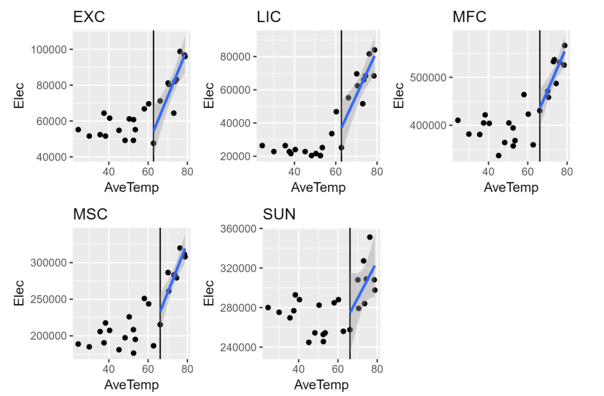

Figure 2. Electricity models for the five sample buildings

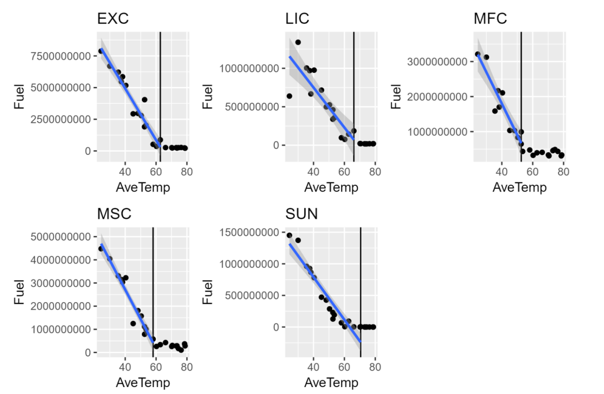

Figure 3. Fuel models for the five sample buildings

sim.temps	total.e	cool.coef	cool.eff	i.loads
10	25200.0	9406.399	3.490315	124540.7
15	25200.0	9406.399	3.490315	124540.7
20	25200.0	9406.399	3.490315	124540.7
25	25200.0	9406.399	3.490315	124540.7
30	25200.0	9406.399	3.490315	124540.7
35	25200.0	9406.399	3.490315	124540.7
40	25200.0	9406.399	3.490315	124540.7
45	25200.0	9406.399	3.490315	124540.7
50	25200.0	9406.399	3.490315	124540.7
55	25200.0	9406.399	3.490315	124540.7
60	25200.0	9406.399	3.490315	124540.7
65	31236.8	9406.399	3.490315	124540.7
70	44711.8	9406.399	3.490315	124540.7
75	58186.8	9406.399	3.490315	124540.7
80	71661.8	9406.399	3.490315	124540.7
85	85136.8	9406.399	3.490315	124540.7
90	98611.8	9406.399	3.490315	124540.7
95	112086.8	9406.399	3.490315	124540.7
100	125561.8	9406.399	3.490315	124540.7

Table 3. Results for Simulation 1: Outdoor Air Temperature

sim.Tset	total.e	cool.coef	cool.eff	i.loads
50	60881.8	9406.399	3.490315	-120025.65
55	60881.8	9406.399	3.490315	-72993.66
60	60881.8	9406.399	3.490315	-25961.66
65	60881.8	9406.399	3.490315	21070.33
70	60881.8	9406.399	3.490315	68102.33
75	60881.8	9406.399	3.490315	115134.33

Table 4. Results for Simulation 2: Thermostat Setpoint

sim.insulation	total.e	cool.coef	cool.eff	i.loads
0.6	60881.8	6704.239	2.487658	88764.13
0.8	60881.8	7304.719	2.710471	96714.48
1.0	60881.8	7905.199	2.933284	104664.84
1.2	60881.8	8505.679	3.156096	112615.19
1.4	60881.8	9106.159	3.378909	120565.55
1.6	60881.8	9706.639	3.601721	128515.90
1.8	60881.8	10307.119	3.824534	136466.26
2.0	60881.8	10907.599	4.047347	144416.61

Table 5. Results for Simulation 3: Thermal Conductance

Simulation 2: Thermostat Setpoint An analysis of the results showed that lowering setpoint temperature by just 5°F decreases internal loads significantly. This is as designed, since the simplified internal loads term includes internal gain from electricity, so increased electricity usage for cooling would increase internal loads. For the current work, the total energy term is kept constant during this simulation. Future work would disaggregate the simulation equation function so that the internal loads and total energy would be variable and therefore the balance between the two might better be discerned.

Simulation 3: Thermal Conductance An analysis of the results showed that lowering thermal conductance (i.e., lowering u-values) has an effect on the internal loads, cooling efficiency and cooling demand (cooling coefficient). This is as designed, since a building envelope that is tighter (i.e., more insulated) will require less cooling energy to satisfy the resulting lower cooling demand, which means the system can work at a less intense efficiency and therefore the internal loads from electricity are deceased. For the current work, the total energy term is kept constant during this simulation. Future work would disaggregate the simulation equation function so that all related parameters would be variable and therefore the balance between them might better be discerned.

Simulation 4: Air Changes An analysis of the results showed that increasing the number of air changes per hour has an effect on the internal loads, cooling efficiency and cooling demand (cooling coefficient). This is as designed, since more air changes require increased usage of ventilation fans, which causes an increase in electricity usage and therefore in internal loads; this then reduces cooling system efficiency and results in an increase in the amount of electricity needed to satisfy the building's increased cooling demand. For the current work, the total energy term is kept constant during this simulation. Future work would disaggregate the simulation equation function so that all related parameters would be variable and therefore the balance between them might better be discerned.

5.2.2. Weather Simulations

Weather simulations were run, according to parameters detailed in the methodology section. Following is a visualization of sample results:

sim.V	total.e	cool.coef	cool.eff	i.loads
1.0	60881.8	9406.399	3.490315	124540.7
1.5	60881.8	11857.799	4.399925	156997.3
2.0	60881.8	14309.198	5.309536	189453.8
2.5	60881.8	16760.598	6.219146	221910.3
3.0	60881.8	19211.997	7.128756	254366.8

Table 6. Results for Simulation 4: Air Changes

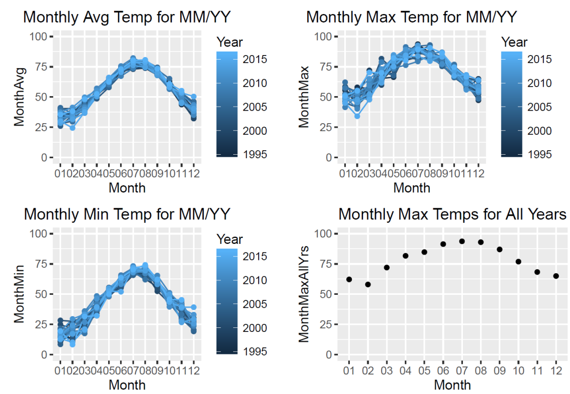

Figure 4. Weather simulation results by type of strategy

1995 Month	Sampled Temperature (°F)
01	56.9
02	31.3
03	47.6
04	52.2
05	50.7
06	76.8
07	81.2
08	73.2
09	68.3
10	61.7
11	46.3
12	41.5

Table 7. Output from monthly random sampled temperature simulation for the year 1995

6. Conclusion

The current work was successful in replicating the methodology used in the reference paper, "Estimating Industrial Building Energy Savings using Inverse Simulation" (Sever and Kissock, 2011), but extended the authors' work from industrial facilities in a small midwestern city to commercial/institutional office buildings in New York City.

Using the primary model equation function, predicted results from simulation 1 were verified against actual historical electricity consumption data, and found to be within a 10% margin; while, results from simulations 2, 3 and 4 were found to be consistent with real-world work with energy efficiency in buildings.

During the course of this work, the authors identified certain enhancements that would likely provide for greater precision in future work, in addition to the inclusion of the “air change per hour” term that was included herein. The most influential change to the current methodology would likely be the disaggregation of the energy balance equations into more detailed parameters – not enough to be burdensome and overly complicate the process, but at a level that would enhance calculability without onerous requirements.

In light of the sizable potential for greenhouse gas emissions reduction in New York City’s municipal portfolio, inverse simulation looks to be a valuable addition to an energy manager’s toolkit, to perform simple “back-of-the-envelope” calculations for viability of proposed energy conservation measures.

References

[1] 2012 Commercial Buildings Energy Consumption Survey: Energy Usage Summary. CBECS 2012: Energy Usage Summary. U.S. Energy Information Administration, 18 Mar. 2016. Web.

[2] ASHRAE Guideline 14-2014: Measurement of Energy and Demand Savings ASHRAE, Atlanta, Georgia (2014)

[3] Carpenter, Kevin et al. "Profiling And Forecasting Daily Energy Use With Monthly Utility-Data Regression Models". ASHRAE Transactions 116.Part 2 (2010): 639-651. Print.

[4] The City of New York. Mayor's Office of Sustainability. Inventory of New York City Greenhouse Gas Emissions in 2014. New York: n.p., 2014. Print.

[5] Efficiency Valuation Organisation (2016). International Performance Measurement and Verification Protocol (IPMVP). Core Concepts 10000-1:2016.

[6] ISO. Energy performance of buildings - Calculation of energy use for space heating and cooling. ISO 13790:2008. Geneva, Switzerland: InternationalOrganization for Standardization, 2008.

[7] Killick R and Eckley IA (2014). "changepoint: An R Package for Changepoint Analysis." *Journal of Statistical Software*, *58*(3), pp. 1-19.

[8] Killick R, Haynes K and Eckley IA (2016). *changepoint: An R package for changepoint analysis*. R package version 2.2.2.

[9] John Kelly Kissock, Jeff S. Haberl, and David E. Claridge. "Inverse Modeling Toolkit: Numerical Algorithms". ASHRAE Transactions - Research 109. Part 2 (2003), pp. 425–545.

[10] "NSRDB: 1991- 2010 Update". Rredc.nrel.gov. N.p., 2016. Web. 21 Dec. 2016.

[11] Franc Sever, J. Kelly Kissock et al. "Estimating Industrial Building Energy Savings using Inverse Simulation". Proceedings of the 2011 ASHRAE Winter Conference. ASHRAE, 2011.

[12] Temperature Data Archive. url: http://academic.udayton.edu/ kissock/http/Weather/default.htm.

[13] Hadley Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009. isbn: 978-0-387-98140-6. url:http://ggplot2.org.

[14] Yuna Zhang et al. "Inverse Model with Uncertainty Quantification to Estimate the Energy Performance of an Office Building". Proceedings of the BS2013: 13th Conference of International Building Performance Simulation Association. Chambery, France: Building Performance Simulation Association, 2013.

A. Appendix: LIC Model Parameters and Equations

```
 # Assumptions
  baseload <- 25200
  #u.value <- 1.5  # watts per sq. meter/K building envelope conductance  SI
  area <- 7500  # sq. meter envelope area  SI
  volume <- 37800  # cubic meter volume (conditioned space) (assumption) SI
  #air.chg <- 1  # air change per hr.
  rho <- 1.2  # kg/m cubed density of air  SI
  cp <- 0.27  # watt hrs./cubic meter in K specific heat of air  SI
  cool.slope <- 2695  # cooling slope (from lm)
  cool.cp <- 62.76  # cooling CP (from CP model)
  #setpoint <- 76  # cooling setpoint (assumption)
  #oa.temp

  # Cooling Coefficient
  cool.coef <- ((((u.value * area) +
  (volume * air.chg * rho * cp)) / 1000) * 0.556) * (30*24)
  # convert from watts to kw, = kwh per F (multiply by 0.556)

  # Cooling Efficiency
  cool.eff <- cool.coef / cool.slope

  # Internal Loads
  i.loads <- -cool.coef * (cool.cp - setpoint)

  # Total Electricity
  total.e <- ifelse(oa.temp - cool.cp > 0,
    baseload + (cool.slope * (oa.temp - cool.cp)),
    baseload) # E = expected kWh at Toa - CP MODEL EQUATION

      #))
}
```

B. Appendix: R Code

Data Preparation, Exploration, Visualization

```
files <- c("EXC","LIC","MFC","MSC","SUN")

df_exc <- read.csv("data/EXC.csv", stringsAsFactors = FALSE)
df_exc$Facility <- "EXC"

df_lic <- read.csv("data/LIC.csv", stringsAsFactors = FALSE)
df_lic$Facility <- "LIC"

df_mfc <- read.csv("data/MFC.csv", stringsAsFactors = FALSE)
df_mfc$Facility <- "MFC"

df_msc <- read.csv("data/MSC.csv", stringsAsFactors = FALSE)
df_msc$Facility <- "MSC"

df_sun <- read.csv("data/SUN.csv", stringsAsFactors = FALSE)
df_sun$Facility <- "SUN"

df <- df_exc
df <- rbind(df, df_lic)
df <- rbind(df, df_mfc)
df <- rbind(df, df_msc)
df <- rbind(df, df_sun)

df$Start_Date <- as.Date(df$Start_Date, format = "%d-%b-%y")
df$End_Date <- as.Date(df$End_Date, format = "%d-%b-%y")

df$Facility <- as.factor(df$Facility)

df$MonthYear <- format(as.Date(df$End_Date), "%Y-%m")

# Floor area

Facility <- c("MSC", "MFC", "EXC", "SUN", "LIC")
Gross.floor.area <- c(212500, 481000, 59000, 242062, 59300)
Floor.area <- cbind(Facility, Gross.floor.area) %>% as.data.frame()
Floor.area$Facility <- as.factor(Floor.area$Facility)

df <- left_join(df, Floor.area, by = "Facility")

# Get temp data
```

```
NYC_weather <- read.csv("data/NYNEWYOR.csv", header=FALSE)
colnames(NYC_weather) <- c("Month", "Day", "Year", "AveTemp")
NYC_weather$Month <- str_pad(NYC_weather$Month, 2, pad = "0")
NYC_month_ave <-aggregate(AveTemp ~ Month + Year, NYC_weather, mean)
NYC_month_ave$MonthYear = paste(NYC_month_ave$Year, "-", NYC_month_ave$Month, sep = "")
df <- left_join(df, NYC_month_ave, by = "MonthYear") %>%
  select(MonthYear, Elec, Fuel, Facility, Gross.floor.area, AveTemp)
Save the Dataframe

save(df, file = "report_dataframe.Rda")

# Visualization

ggplot(df, aes(x = AveTemp, y = Elec)) + geom_point() + stat_smooth()
ggplot(df, aes(x = AveTemp, y = Elec, color = Facility)) + geom_point() + stat_smooth()
ggplot(df, aes(x = AveTemp, y = Fuel)) + geom_point() + stat_smooth()
ggplot(df, aes(x = AveTemp, y = Fuel, color = Facility)) + geom_point() + stat_smooth()

# Changepoint Analysis

elec.bcp <- bcp(y = df$Elec, x = df$AveTemp)
plot(elec.bcp)
fuel.bcp <- bcp(y = df$Fuel, x = df$AveTemp)
plot(fuel.bcp)
```

Electric Change-point Modelling

```
get.elec.changepoint <- function(building, data){
  cp.elec.data <- data %>%
    filter(Facility == building) %>%
    select(Elec, AveTemp) %>%
    arrange(AveTemp)

  cp.elec <- cpt.mean(cp.elec.data[,1], method = 'AMOC', penalty = 'None')
  return(cp.elec.data[cpts(cp.elec),])
}

elec.changepoints <- sapply(unique(df$Facility), FUN = get.elec.changepoint, data = df)
colnames(elec.changepoints) <- unique(df$Facility)

plot.changepoint.elec <- function(data, series, facility, changepoint){
  df <- filter(data, Facility == facility)

  g <- ggplot(df, aes_string(x = "AveTemp", y = series)) +
    geom_point() +
    geom_vline(xintercept = as.numeric(changepoint)) +
```

```
    ggtitle(facility) +
    stat_smooth(method = lm, data = df %>% filter(AveTemp >= changepoint), mapping = aes_str

  return(g)
}
kable(elec.changepoints)

grid.arrange(plot.changepoint.elec(df, "Elec", "EXC", elec.changepoints[2,1]),
             plot.changepoint.elec(df, "Elec", "LIC", elec.changepoints[2,2]),
             plot.changepoint.elec(df, "Elec", "MFC", elec.changepoints[2,3]),
             plot.changepoint.elec(df, "Elec", "MSC", elec.changepoints[2,4]),
             plot.changepoint.elec(df, "Elec", "SUN", elec.changepoints[2,5]),
             ncol = 3, nrow = 2)
get.fuel.changepoint <- function(bui
```

Fuel Change-point Modelling

```
get.fuel.changepoint <- function(building, data){
  cp.fuel.data <- data %>%
    filter(Facility == building) %>%
    select(Fuel, AveTemp) %>%
    arrange(AveTemp)

  cp.fuel <- cpt.meanvar(cp.fuel.data[,1], method = 'AMOC')

  return(cp.fuel.data[cpts(cp.fuel),])
}

fuel.changepoints <- sapply(unique(df$Facility), FUN = get.fuel.changepoint, data = df)
colnames(fuel.changepoints) <- unique(df$Facility)

plot.changepoint.fuel <- function(data, series, facility, changepoint){
  df <- filter(data, Facility == facility)

  g <- ggplot(df, aes_string(x = "AveTemp", y = series)) +
    geom_point() +
    geom_vline(xintercept = as.numeric(changepoint)) +
    ggtitle(facility) +
    stat_smooth(method = lm, data = df %>% filter(AveTemp <= changepoint), mapping = aes_str

  return(g)
}

kable(fuel.changepoints)

grid.arrange(plot.changepoint.fuel(df, "Fuel", "EXC", fuel.changepoints[2,1]),
```

```
             plot.changepoint.fuel(df, "Fuel", "LIC", fuel.changepoints[2,2]),
             plot.changepoint.fuel(df, "Fuel", "MFC", fuel.changepoints[2,3]),
             plot.changepoint.fuel(df, "Fuel", "MSC", fuel.changepoints[2,4]),
             plot.changepoint.fuel(df, "Fuel", "SUN", fuel.changepoints[2,5]),
             ncol = 3, nrow = 2)
```

LIC Facility Linear Regression

```
  LIC.elec.lm <- lm(Elec ~ AveTemp,
                  data = df %>%
                    filter(Facility == "LIC") %>%
                    filter(AveTemp >= elec.changepoints[2,2]))


LIC.fuel.lm <- lm(Fuel ~ AveTemp,
                  data = df %>%
                    filter(Facility == "LIC") %>%
                    filter(AveTemp <= fuel.changepoints[2,2]))

LIC.temp.model <- function(t){
  ifelse(t <= elec.changepoints[2,2],
         return(elec.changepoints[1,2]),
         return(predict(LIC.elec.lm, newdata = data.frame(AveTemp = t))))
}
```

LIC Equations: Parameters

```
U <- 1.5  # sq. meter/K building envelope conductance (assumption - 12" brick)  SI
A <- 8000  # sq. meter envelope area (assumption based on Pluto + other NYC DCAS data)  SI
V <- 37800  # cubic meter ventilation/inflitration flow rate (assumption) (1-3, step 0.5)  S
n <- 1  # air change per hr.
rho <- 1.2  # kg/m cubed 0.00234 density of air (calc. based on IAT 68F)  SI
cp <- 0.27  # watt hrs./cubic meter in K specific heat of air (calc. based on IAT 68F)  SI
CS <- 2695  # cooling slope (from lm)
TcpC <- 62.76  # cooling CP (from CP model)
Tset <- 76  # cooling setpoint (assumption)

# 3PC Model
Ei <- 25200  # baseload (non-weather sensitive usage)
Toa <- 80  # sample temp. (VARIABLE)

E <- Ei + (CS * (Toa - TcpC))  # E = expected kWh at Toa - CP MODEL EQUATION
# cat(E, "=", Ei, "+ (", CS, "* (", Toa, "-", TcpC, "))" )
cat("Expected kWh at Toa:", E)

# Cooling Coefficient
```

```
CC <- ((((U * A) + (V * n * rho * cp)) / 1000) * 0.556) * (30*24) # convert from watts to kw
cat("Cooling Coefficient:", CC)

# Efficency
Effc <- CC / (CS)
cat("Cooling efficiency:", Effc)

# Qi - sum of internal loads from electricity use, solar gain and occupants
#TcpC <- Tset - Qi / CC
Qi <- -CC * (TcpC - Tset) # Just kWh (not per degree)
cat("Internal loads:", Qi)
```

Simulation Equations (LIC)

```
# Function: CP Model Equation

bldg.sim <- function(oa.temp, setpoint, u.value, air.chg) {

  # Assumptions
  baseload <- 25200
  #u.value <- 1.5  # watts per sq. meter/K building envelope conductance  SI
  area <- 7500  # sq. meter envelope area  SI
  volume <- 37800  # cubic meter volume (conditioned space) (assumption) SI
  #air.chg <- 1  # air change per hr.
  rho <- 1.2  # kg/m cubed density of air  SI
  cp <- 0.27  # watt hrs./cubic meter in K specific heat of air  SI
  cool.slope <- 2695  # cooling slope (from lm)
  cool.cp <- 62.76  # cooling CP (from CP model)
  #setpoint <- 76  # cooling setpoint (assumption)
  #oa.temp

  # Cooling Coefficient
  cool.coef <- ((((u.value * area) + (volume * air.chg * rho * cp)) / 1000) * 0.556) * (30*2
  # convert from watts to kw, = kwh per F (multiply by 0.556)

  # Cooling Efficiency
  cool.eff <- cool.coef / cool.slope

  # Internal Loads
  i.loads <- -cool.coef * (cool.cp - setpoint)

  # Total Electricity
  total.e <- ifelse(oa.temp - cool.cp > 0,
    baseload + (cool.slope * (oa.temp - cool.cp)),
    baseload) # E = expected kWh at Toa - CP MODEL EQUATIO
```

```
  # Parameters
  parameters <- c(u.value, area, volume, air.chg, rho, cp, cool.slope, cool.cp,
                  cool.eff, setpoint, baseload, oa.temp, total.e, i.loads)
  # as.data.frame(parameters)
  names(parameters) <- c('u-value', 'surface area', 'volume', 'air changes', 'density of air
                         'specific heat of air', 'cooling slope', 'cooling change-point',
                         'cooling efficiency', 'setpoint', 'baseload', 'outdoor air temp.',
                         'total electricity', 'internal loads')

  return(cbind(total.e, cool.coef, cool.eff, i.loads))
  # Return
  #return(
  #cat("Expected kWh at Toa:", total.e,
      #"\nCooling Coefficient:", cool.coef,
      #"\nCooling Efficiency:", cool.eff,
      #"\nInternal Loads:", i.loads,
      #"\nParameters\n",
      #parameters

      #))
}

bldg.sim(50, 76, 1.5, 1)
```

Simulations (LIC)

```
# Sim #1: Toa -- use CP Model Equation and simulate Toa from 10-100 degrees F in steps of 5

library(devtools)
scipen=9999

sim.temps <- seq(from = 10, to = 100, by = 5)

temp.range.results <- lapply(sim.temps, FUN = bldg.sim,  setpoint = 76, u.value = 1.5, air.c

temp.range.results <- as.data.frame(do.call(rbind, temp.range.results))
temp.range.results <- cbind(sim.temps, temp.range.results)
print(temp.range.results)

# Sim #2: Tset -- substitute other values from 50 to 75, in steps of 5 degrees -- this simul

slibrary(devtools)
scipen=9999

sim.Tset <- seq(from = 50, to = 75, by = 5)
```

```
Tset.range.results <- lapply(sim.Tset, FUN = bldg.sim, oa.temp = 76, u.value = 1.5, air.chg

Tset.range.results <- as.data.frame(do.call(rbind, Tset.range.results))
Tset.range.results <- cbind(sim.Tset, Tset.range.results)
print(Tset.range.results)


# Sim #3: U -- substitute other values -- this simulates adding insulation, etc. to tighten

sim.insulation <- seq(from = 0.6, to = 2, by = 0.2)

insulation.range.results <- lapply(sim.insulation, FUN = bldg.sim, oa.temp = 76, setpoint =

insulation.range.results <- as.data.frame(do.call(rbind, insulation.range.results))
insulation.range.results <- cbind(sim.insulation, insulation.range.results)
print(insulation.range.results)

#Sim #4: V -- substitute other values: 1 to 3, in steps of 0.5 -- this simulates improved/wo

sim.V <- seq(from = 1, to = 3, by = 0.5)

V.range.results <- lapply(sim.V, FUN = bldg.sim,  oa.temp = 76, setpoint = 76, u.value = 1.5

V.range.results <- as.data.frame(do.call(rbind, V.range.results))
V.range.results <- cbind(sim.V, V.range.results)
print(V.range.results)
```

Weather Simulations

```
NYC_weather <- read.csv("data/NYNEWYOR.csv", header=FALSE)
colnames(NYC_weather) <- c("Month", "Day", "Year", "AveTemp")
NYC_weather$Month <- str_pad(NYC_weather$Month, 2, pad = "0")
NYC_weather <- NYC_weather[NYC_weather$AveTemp >= -80, ]
#kable(head(NYC_weather))


month_avgs <- sqldf("select Year, Month, avg(AveTemp) as MonthAvg from NYC_weather group by

p1 <- ggplot(data=month_avgs, aes(x=Month, y=MonthAvg, group=Year, color=Year)) + ggtitle("M

month_maxes <- sqldf("select Year, Month, max(AveTemp) as MonthMax from NYC_weather group by

p2 <- ggplot(data=month_maxes, aes(x=Month, y=MonthMax, group=Year, color=Year)) + ggtitle("

month_mins <- sqldf("select Year, Month, min(AveTemp) as MonthMin from NYC_weather group by
```

```
p3 <- ggplot(data=month_mins, aes(x=Month, y=MonthMin, group=Year, color=Year)) + ggtitle("M

month_maxes_all_years <- sqldf("select Month, max(AveTemp) as MonthMaxAllYrs from NYC_weathe

p4 <- ggplot(data=month_maxes_all_years, aes(x=Month, y=MonthMaxAllYrs)) + ggtitle("Monthly 

grid.arrange(p1, p2, p3, p4, ncol = 2)
```