RC25337 (WAT1211-117) November 29, 2012
Computer Science

IBM Research Report

Building Energy Performance Analytics on Cloud as a Service

Young M. Lee, Lianjun An, Fei Liu, Raya Horesh, Young Tae Chae, Rui Zhang, Estepan Meliksetian, Pawan Chowdhary, Paul Nevill, Jane Snowdon
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 208
Yorktown Heights, NY 10598
USA

Building Energy Performance Analytics on Cloud as a Service

Young M. Lee*, Lianjun An, Fei Liu, Raya Horesh, Young Tae Chae, Rui Zhang, Estepan Meliksetian,

Pawan Chowdhary, Paul Nevill, Jane Snowdon

IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598, U.S.A.,

Tel: (914) 945-3562, Fax: (914) 945-3434, Email: ymlee@us.ibm.com

*Corresponding author

Submitted to the Special Issue of Service Science on Cloud Computing as a Service

Building Energy Performance Analytics on Cloud as a Service

<u>Abstract</u>

Keywords: Cloud, Energy Performance, Energy Simulation, Building Energy Analytics, Visualization

Saving energy, improving energy efficiency and reducing greenhouse gas (GHG) emissions are among the most important initiatives in today's world. Occupied buildings consume a substantial amount of energy; mounting to about 40% of overall energy consumption in most countries. The majority of the world's population either lives or works in buildings; therefore, everybody has a responsibility and a role to play in reducing energy consumption, controlling GHG emissions, and mitigating climate change and its potential impact. We developed an analytical tool which can assist building owners, facility managers, operators and tenants of buildings in assessing, benchmarking, diagnosing, tracking, forecasting, simulating and optimizing energy consumption in building portfolios. Furthermore, for greater dissemination, we have made this analytic service available on demand in a flexible cloud environment. Cloud is an efficient and effective medium to provide building energy analytics capability to various functions and people in a variety of roles in buildings without investing a substantial amount of money in hardware, software and IT infrastructure. We present results of the building energy analytics developed for New York City's K-12 public school buildings and a commercial office building in Korea.

1 Introduction

Energy consumption by humans causes a gradual increase in concentrations of greenhouse gas (GHG) in the Earth's atmosphere, and is considered to be the main source of global warming (Silver & DeFries, 1991). In the United States, 40% of the nation's total energy consumption is due to commercial and residential buildings (U.S. Dept of Energy, 2006), and this figure is constantly increasing. Such buildings also contribute 45% of the country's GHG emissions (EPA ESPM Greenhouse Gas Inventory 2009). The majority of the world's population either lives or works in buildings; therefore, everybody has a responsibility and a role to play in reducing energy consumption, controlling GHG emissions, and confronting climate change and its potential impacts. End-use energy efficiency can contribute to more than 50% of total global energy conservation (World Energy Outlook 2009).

Saving energy, improving energy efficiency and reducing greenhouse gas (GHG) emissions are key initiatives for many cities, municipalities, building owners and operators. For example, New York City (NYC)'s government spends over $1 billion a year on energy for their approximately 4,000 buildings (e.g. public schools, prisons, court houses, administrative buildings, waste water treatment plants, etc.), and is committed to reducing the City government's energy consumption and CO_2 emissions by 30% by 2017 from their 2005 levels through an initiative called PlaNYC (PlaNYC 2007). NYC plans to invest, each year, an amount equal to 10% of its energy expenses in energy-saving measures over the next 10 years. The largest segment of NYC government buildings is the 1,200 K-12 public schools serving 1.1 million students and covering about 150 million square feet. The New York City Department of Education was interested in understanding how energy efficient their buildings are, what factors contribute to inefficiencies, what the opportunities for improvement given budget constraints are, and how much they can contribute in energy savings and GHG emission reductions toward NYC's PlaNYC initiative, and wanted to have the information readily available to the department, school custodians and even to students.

In order to reduce energy consumption in buildings, however, one needs to understand patterns of energy usage and heat transfer as well as characteristics of building structures, operations and occupants' behavior that influences energy consumption. This understanding can be aided through development of scientific models which are based on physics, mathematics and statistics. The models can then be used to simulate the impact of possible changes that can be made to buildings on energy consumption, energy costs and GHG emissions, and provide decision support for making buildings more energy efficient. The changes can be structural changes such as retrofits (e.g. new boiler, insulation, windows or roof), operational changes such as operating hours, behavioral changes such as running appliances at different times of the day when electricity prices are lower, and external changes such as weather factors. The models can also be used for optimizing the changes that can be made to the buildings given an energy conservation target. Developed along this effort is the IBM Building Energy and Emissions Analytics (called *i-BEE^TM^*) Toolset, an analytical tool that assesses, benchmarks, diagnoses, tracks, forecasts, simulates and optimizes energy consumption in building portfolios. As initial efforts of this initiative, IBM collaborated with the City University of New York (IBM, 2011a) to develop the tool for the portfolio of K-12 public school buildings in New York City to identify energy saving opportunities. The *i-BEE^TM^* tool for NYC K-12 schools was deployed on the cloud so that all the building energy analytics were readily accessible by building owners, facility managers, operators and tenants of the buildings through web browsers without the need to install any special hardware, software, data interface or mathematical toolbox. Additionally, we are also developing the building energy analytics for a commercial office complex in Korea using near real time sensor/meter data collected in the building management system (BMS).

Cloud computing is an efficient and effective way to deliver the techniques, platform, software as services, and integration of multiple services into an enterprise solution. We deploy our building analytic modules into the cloud platform as services, make building and energy data retrieved from the BMS, weather data and other collected data available on the cloud, provide interactive reports and customizable visualization for a portfolio of buildings on the cloud, and enable collaboration among building operators owners as well as occupants.

The rest of the paper is organized as follows. In section 2, the data flow and analytics of the building energy analytics toolset are presented. In section 3, the heat transfer model and inverse modeling approach for estimating heat transfer coefficients are described. Then, section 4 shows statistical models that integrate multiple regression models and a time series model of building energy. Section 5 explains the deployment of the tool on cloud. Finally, conclusions are provided in section 6.

2 Data flow and analytics of building energy Analytics toolset

The building energy models in *i-BEE*^{TM} incorporate data integrated from multiple sources. Figure 1 describes how various data are collected and assembled into a common database and used by models and analytics.

One type of data used by models is data coming from the building management system (BMS) such as temperature and flow rate of supply air and return air from/to air handling units (AHU), which provide warm or cool air into various rooms inside buildings. Data collected from various sensors in spaces in buildings such as temperature, humidity and occupancy are also used. The sensor data can be recorded and collected through the BMS or individually. Data can also come from meters and sub-meters that measure energy use such as electricity, natural gas, steam and chilled water for the whole building or parts of the building or equipment. Other data elements include: energy bills, e.g. monthly bills for electricity, natural gas and steam consumption; building characteristics such as gross floor area (GFA), age of building, number of occupants, wall area, window area, roof area, number of floor, percentage of building cooled and heated, operating hours, number of computers and other equipment such as refrigerators, freezers etc; thermal loads and plug loads; weather data for historic and current conditions such as temperature, humidity, solar radiation and wind; and real-time electricity prices from the local utility. Data from the Environmental Protection Agency (EPA) can also be used for calculating *source* *energy* and greenhouse gas (GHG) emissions for which a building is responsible (EPA ESPM Source Energy, 2009). *Source energy* is the amount of energy required to generate and transport the energy requirements of the building. *Site energy* is the amount of energy consumed by a building for heating, cooling, lighting, plug loads, and so on.

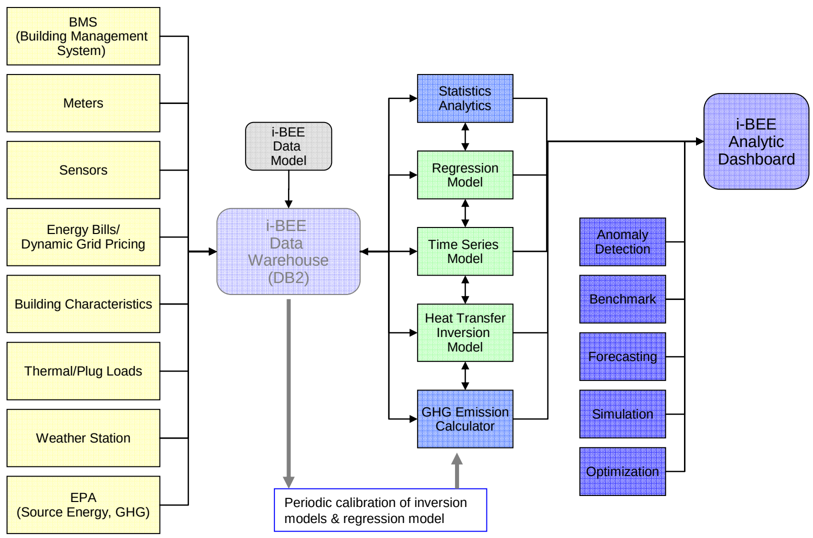

Figure 1: Data Flow and Analytics

The *i-BEE*^{TM} data described above is organized into a database in a data warehouse, which is designed through a data model that defines the relationship between data elements. The data from the database is used in populating three fundamental base models: heat transfer inverse model, multivariate regression model, and time series model. The models are described below in the following sections. These three base models are being calibrated periodically as new data are brought into the database, such that the model stays accurate even if the building energy performance degrades over time. In addition to the model, basic statistical analyses including temporal and spatial analysis, and GHG emissions calculations are done for the buildings in the portfolio. Using the three base models in an integrated way, various analytical tasks can be performed. One task is anomaly detection for energy use for each building. This capability identifies abnormal energy consumption for a building or a part of the building by comparing actual energy consumption with predicted energy consumption and calculating upper and lower control bounds using statistical process control methods. Another task is benchmarking, which computes energy performance indicators of each building for each energy type and for each type of energy load (e.g., base load, heating load and cooling load), and identifies under-performing buildings, which can be candidates for energy efficiency improvements. A third task is forecasting future energy consumption in the short-term (i.e., for the next 24 hours) and longer-term (e.g., for the next few months). A fourth task is simulation (what-if analysis) of changes that could be made to building structures, operations and occupant incentives on energy consumption and GHG emissions. Finally, several types optimization can also be done as explained in the sections below. All the analytics and reports are accessible securely to users through an analytics dashboard deployed on the cloud.

3 Heat Transfer Modeling and Inversion

In this section, we describe how thermal energy is used to provide a comfortable climate, e.g., temperature and humidity, inside a building using thermal physics principles. Figure 2 shows a simplified view of heat transfer in a building. Building occupants desire comfortable temperature and humidity inside the building. However, since buildings are neither perfectly insulated nor blocked from sunlight, warm and humid climate conditions outside a building come into the building during the summer season, and cold and dry air comes into the building during the winter season making the inside climate uncomfortable for the occupants. In order to compensate for the influence of the outside climate, the heating ventilation and air conditioning (HVAC) system and plant equipment (e.g., chillers and boilers etc.) are operated to provide thermal energy into the building to maintain occupants' thermal comfort. Typically heat transfer from outdoor conditions into the space involves heat transfer through a building component like external or internal walls, windows, roof, and the ground (foundation). The heat transfer includes heat conduction (e.g., heat flows through wall materials), heat convection (e.g., heat flows through the air from interior walls into the space), solar radiation (e.g., solar energy on the exterior wall or through openings such as a window onto an interior wall or object like a piece of furniture), infiltration of outside air into the space through cracks around windows, doors and opened windows and doors, and internal heat gain from light, equipment, and occupants. The plant and its systems produce thermal energy sources such as steam, hot water or chilled water, which is then transferred to the space through the system equipment. In the case of an all-air based system, heat exchangers convert the source energy into warm and cold air with a certain supply temperature, humidity and flow rate, and blow them to each zone of the building using AHUs and other fan systems.

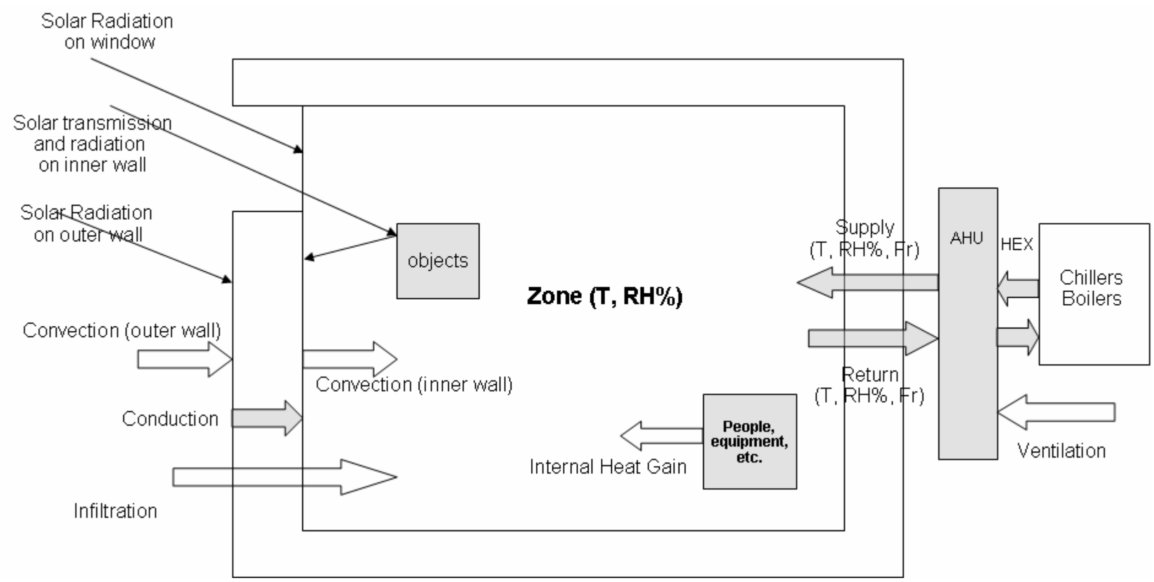

Figure 2: Schematic View of Heat Transfer in a Building

The heat transfer can be simplified as the following set of equations:

$$\frac{\partial}{\partial t}[(\rho_e C_{p,e})T_e] = \frac{\partial}{\partial x}(k_e \frac{\partial T_e}{\partial x})$$

$$\forall e \in \{wall_E, wall_W, wall_N, wall_S, window, roof, ground...\} \quad (1)$$

$$q_x = -k_e \frac{\partial T_e}{\partial x}|_I = h_I(T_{I,e} - T_z) + \dot{Q}_{sol}\lambda_{w,I} \quad (2)$$

$$q_x = -k_e \frac{\partial T_e}{\partial x}|_O = h_O(T_{amb} - T_{O,e}) + \dot{Q}_{sol}\lambda_{w,O} \quad (3)$$

$$\rho_z V_z C_p \frac{dT_z}{dt} = \dot{m}_{\inf} C_p (T_{amb} - T_z) + m_{sys} C_p (T_{sys} - T_z) + \sum_e h_{I,e} A_e (T_{I,e} - T_z) + \dot{Q}_{sol}\lambda_{shgc} A_{win} + \dot{Q}_{\text{int}} \quad (4)$$

Equation (1) is energy balance on the building envelope, e, which can consist of walls in different directions, windows, roof and ground. At the outer and inner envelope surfaces, heat transfer is controlled by a convective heat transfer from ambient air and zone air, and solar radiation. At the outer and inner envelope surfaces, heat transfer is controlled by a convective heat transfer from ambient air and zone air, and solar radiation. The boundary conditions at the outer and inner envelope surfaces are expressed by equation (2) and (3). Equation (4) is energy balance inside of the building (zone). Here,

$T_e, T_{I,e}, T_{O,e}, T_z, T_{amb}, T_{sys}$ are temperatures of envelope, inner envelope surface, outer envelope surface, zone, outside air (ambient) and the HVAC system,

ρ_e, ρ_z are densities of the building envelope and the air inside the zone,

V_z, A_e, A_{win}, are volumes of zone, wall, areas of the envelope and windows,

$C_{p,e}, C_p$ are specific heat of the building envelope and air,

k_e, h_I, h_O are coefficients of heat conduction of the envelope and heat convection coefficients at the inner and outer surfaces of the building envelope,

$\dot{m}_{\inf}, \dot{m}_{sys}$ are flow rates of infiltration and the system (HVAC),

$\dot{Q}_{sol}, \lambda_{w,I}, \lambda_{w,O}, \lambda_{shgc}$ are rates of total solar radiation, solar absorption constants at the inner and outer walls, and the solar heat gain constant of window,

$\dot{Q}_{\text{int}}$ is the internal heat gain from people, lighting and other equipment inside the building.

For heat transfer inverse modeling of buildings, often there isn't enough sensor data or meter data to allow for the estimation of all the physical parameters that support the complexity of the heat transfer model (e.g., equations). Therefore, the heat transfer equations above can be simplified further.

Heat transfer coefficients are often expressed in terms of R-values and U-values, i.e., $R_{wall}, R_{roof}, U_{win}$, which are wall and roof heat resistances and the window heat transfer coefficient.

In equations (1) - (4), the parameters that are not known are $k_e, h_I, h_O, C_{p,e}, \dot{m}_{\inf}$,which represent heat conduction coefficients of the envelope, heat convection coefficients of the inner and outer building envelope surfaces, heat capacity of envelope and infiltration coefficient of the building envelope. The parameters can be estimated by forming the problem as an inverse problem (Beck and Woodbury 1998).

Depending on the availability of data for the building, the observed data might not allow the estimation of all parameters, moreover, some parameters are correlated. For the portfolio of K-12 public school buildings analyzed, we generated an innovative procedure to address this issue. First, we derive a static heat transfer model by integrating the differential equations described above over heating seasons and cooling seasons. Then, using the energy consumption data, we estimate the overall heat transfer coefficient and solar contribution for each building. Lastly, a clustering method is used to separate the overall heat transfer into thermal resistance of the wall, roof and window for a collection of similar buildings. The details of this approach are published in a separate paper (An et al. 2012).

We present here an early result of the heat transfer inverse model applied to a 5-story commercial office building located in Korea. We calibrated the heat transfer model using the past 5 days of historical sensor data, and predicted the energy profile for the next 5 days using hourly resolution. Figure 3 below shows predicted and actual zone temperature profiles for a modeled zone in the building on July 2-6. The mean absolute percentage error (MAPE), with respect to the actual and peak energy consumption is 8.9% and 8.2%, respectively. The occurrence of the peak energy differs by 3.5 hours on average. We are in the process of improving the calibration procedure of the heat transfer inverse model in order to improve the accuracy of the prediction. The heat transfer model is planned to be used to simulate the impact of operational changes such as a set point change, pre-cooling and pre-heating.

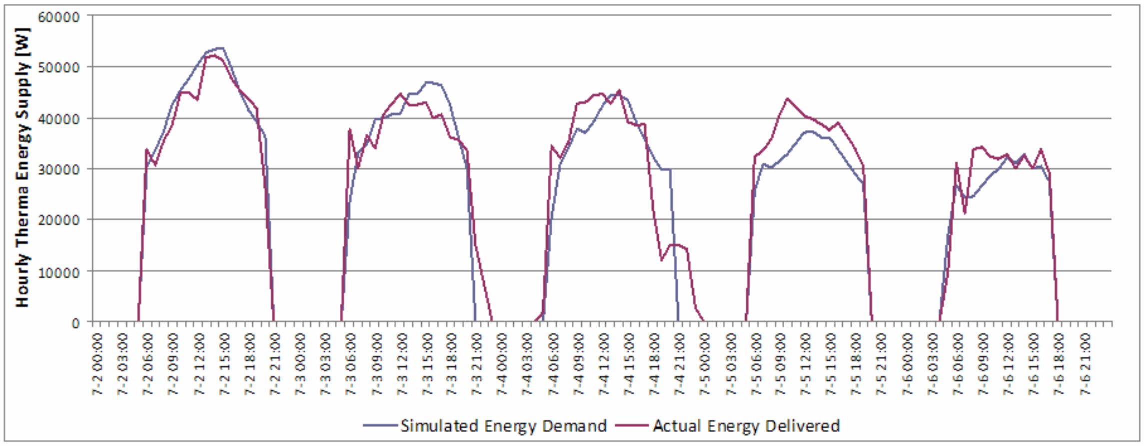

Figure 3: Predicted and Actual Energy Consumption from Heat Transfer Model

4 Statistical Models

We developed statistical methods to help understand the energy usage patterns for portfolios buildings as well as individual buildings. The Variable Base Degree Day (VBDD) regression model (ASHRAE 2009) is a popular approach to analyze energy consumption, which assumes an independent error structure for the regression model. The assumption may not be realistic in practice, especially for our application with a large portfolio of buildings. A new method, which incorporates building heterogeneity and the dependent error structure, is thus developed.

We developed a multi-step statistical analysis procedure, which combines the multivariate regression model, the VBDD regression model and the Auto Regressive Integrated Moving Average (ARIMA) model (Brockwell and Davis 2006), to assess energy efficiency, predict energy consumption and detect anomaly of energy use. In the first step, we build a regression model that correlates energy consumption with building characteristics. The energy related building characteristics are then identified through the stepwise variable selection technique. The results are valuable in providing building energy performance scores for the whole portfolio. Additionally, this method can predict the energy consumption of buildings of similar characteristics. In the second step, to accommodate building heterogeneity, we build VBDD regression models separately for each building. For a VBDD model, the total monthly energy usage data for building i, in period t is denoted y_{it}, and is modeled as:

$$y_{it} = b_i + c_i CDD_t(T_i^{(b)}) + h_i HDD_t(T_i^{(b)}) + \varepsilon_{it} \tag{5}$$

where b_i is the base load usage, c_i is the cooling coefficient, h_i is the heating coefficient, and ε_{it} are the error terms reflecting the month-to-month variations that can not be explained by base, heating or cooling usage. We further restrict that $b_i > 0$, $c_i > 0$ and $h_i > 0$. The heating degree day (HDD) and the cooling degree day (CDD) for building i in month t are defined as

$$HDD_t(T_i^{(b)}) = \sum_{d=1}^{d_t}(T_i^{(b)} - T_{itd})^+, \text{ and } CDD_t(T_i^{(b)}) = \sum_{d=1}^{d_t}(T_{itd} - T_i^{(b)})^+ \tag{6}$$

Here, T_{itd} is the outdoor temperature for building i on day d of month t, $i \in \{1,\ldots,n\}$, $t \in \{1,\ldots,m\}$, $d \in \{1,\ldots,d_t\}$, and $T_i^{(b)}$ is the balance-point temperature for building *i*. $HDD_t(T_i^{(b)})$ and $CDD_t(T_i^{(b)})$ are the cumulative heating and cooling energy usage for month t when the balance point temperature is set to be $T_i^{(b)}$. The VBDD model uses daily energy and temperature data to compute the monthly energy consumption. With the availability of the data in finer time resolution, i.e., hourly, a variation of the VBDD model, Variable Base Degree Hours (VBDH), can also be developed. The VBDH model uses hourly data to compute the daily energy consumption.

These models are used to separate the base load energy consumption from the weather dependent usage. The results of this step consist of the base temperature estimates, as well as the estimated coefficients for HDD and CDD for all buildings. In the third step, we further analyze energy efficiency by energy consumption type (i.e., heating, cooling and base load) by building the multivariate regression models for the results from the VBDD model, from which the performance scores can be derived for base load, heating, and cooling. The multivariate regression model takes the form

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \ldots + x_{ip}\beta_p + \varepsilon_i \quad (7)$$

where y_i is the quantity of interest, typically referred to as the response variable, and $x_{i1},\ldots,x_{ip}$ are the p predictor variables, $\beta_1,\ldots\beta_p$ are the regression coefficients, and ε_i is the error term.

Finally, in the last step, we model the dependent error structure through the ARIMA model. Recall that $\hat{\varepsilon}_{it}$, $t = 1,....,m$, are the estimated error terms from the VBDD models in equation (5). The time series modeling is conducted for each individual building independently. Firstly, we remove the seasonal patterns by using a regression model, with $\hat{\varepsilon}_{it}$ being the response variable and the 12 monthly seasonal factors being the predictor variables. To avoid the colinearity issue, we set the monthly seasonal factor for December equal to 0. We denote the error terms from equation (5) after removing the seasonal patterns by $\hat{\varepsilon}_{it}$. Then, the dynamic structure of $\hat{\varepsilon}_{it}$ is modeled using the ARIMA model to capture time series data in order to better understandthe present data and accurately forecast future data points (Brockwell and Davis 2006). Despite its popularity in statistical literature, the ARIMA model has been rarely used in the context of building energy, partly because of its complex modeling schemes. Nevertheless, the ARIMA model provides a more flexible, possibly non-stationary structure to model building energy patterns, which is essential for simultaneously modeling a large number of buildings. The ARIMA model takes a form

$$(1-\sum_{\ell=1}^{p}\phi_{i\ell}L^{\ell})(1-L)^{\ell}\tilde{\varepsilon}_{it} = (1+\sum_{\ell=1}^{q}\theta_{i\ell}L^{\ell})\eta_{it} \quad (8)$$

where L is the lag operator, $L\tilde{\varepsilon}_{it} = \tilde{\varepsilon}_{i,t-1}$; p, r, q are non-negative integers and are the orders of auto-regressive, integrated, and moving average parts of the model; $\{\phi_{i\ell}, \ell = 1,\ldots,p\}$ and $\{\theta_{i\ell}, \ell = 1,\ldots,q\}$ are the parameters associated with the auto-regressive and moving average parts of the model; and η_{it} are mutually independent standard normal random variables. The ARIMA models are the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing. In fact, the order of the integrated part r reflects the trend of the data (e.g., $r = 0$ no trend, $r = 1$ linear trend, $r = 2$ quadratic trend, etc), while p and q control how fast the auto-correlation decays. The details of the technique are described in Liu et al. (2011).

The ARIMA model, combined with either the VBDD or VBDH regression models, can be used to predict and forecast energy consumption for each individual building. In addition, the anomaly dectection with respect to the historical energy usage may be conducted by comparing the observed energy useage with the resulting upper and lower control limits for each building. We present here sample results from the VBDH/ARIMA model on foreasting and anomaly detection. Figure 4 shows the predicted energy consumption (solid line), actual energy consumption (square dots), detected energy consumption anomalies (square dots circled), and a 3-day forecast (line breaks in circle) in daily scale for an electricty meter for a 5 story commercial building. The figure is for the scenario of using one month of historical energy consunption data (March 1st - March 31s) to calibrate the VBDH/ARIMA model, and then forecast energy consumption for 3 future days (April 1st, 2nd and 3rd). On March 29th and 30th, the energy consumption observed by this particular meter is higher than the control bound, thus generating alerts that anomalies occurred. Figure 5 shows the forecast accuracy for energy consumption recorded through all the meters (4 different types, i.e., electrcity meters, gas meters, total meters and thermal meters) in the commercial building in the scatter diagram (log scaled). Here, the total meters are aggregated values of several meters for the purpose of tracking and reporting, and the thermal meters are computed energy consumption values based on metered data to represent the thermal energy for certain zones and equipment. The forecast error in MAPE for all the meters is about 10% for the building of this study.

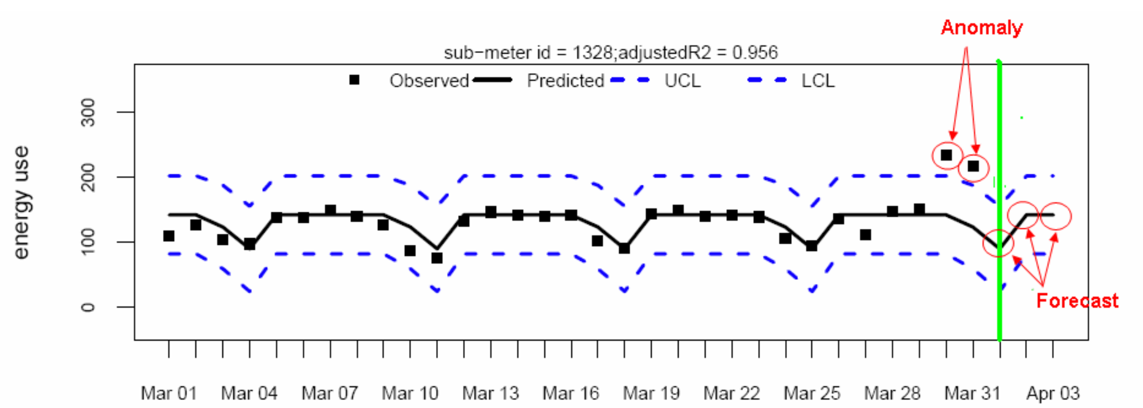

Figure 4: Energy Forecast and Anomaly Detection

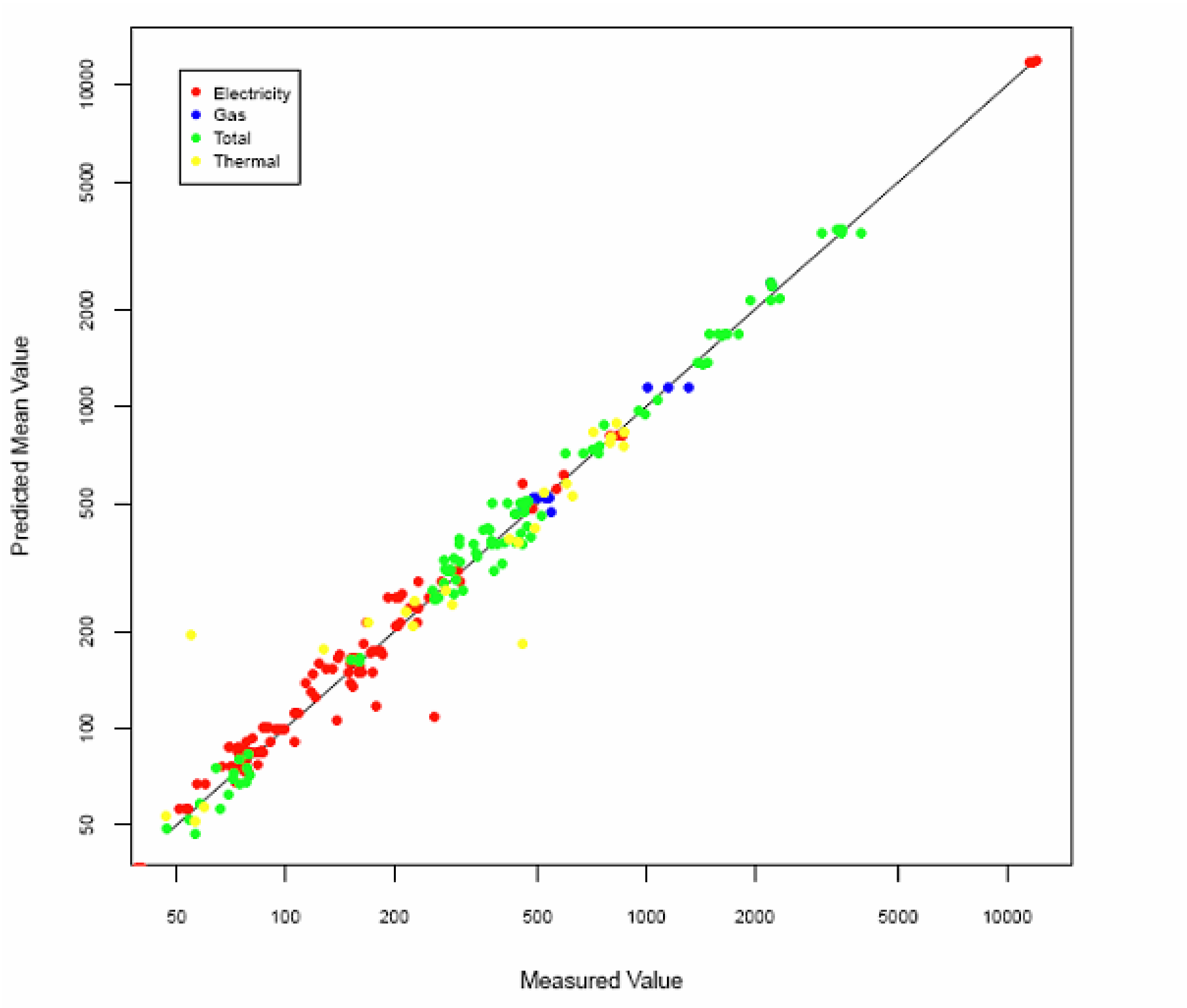

Figure 5: Scatter Plot of Measured Meter Values and Forecasted Values

In addition to the anomaly detection by control bounds, the VBDH/ARIMA model is also used for detecting drifting trend of energy consumption through the analysis of Cumulative Sum of Differences, CUSUM (Duncan, 1986), which computes the cumulative difference between predicted and actual energy consumption and identifies the trend that goes out of the statistical boundary. Let $\hat{y}_{it}$ be the target energy consumption for meter i at time t, i.e., the energy consumption predicted based on the VBDD regression model. Then the lower cusum C_{it}^{-} and the upper cusum C_{it}^{+} are defined as

$$C_{it}^{-} = \max(0, (\hat{y}_{it} - k) - y_{it} + C_{i,t-1}^{-}) \quad (9)$$

$$C_{it}^{+} = \max(0, y_{it} - (\hat{y}_{it} + k) + C_{i,t-1}^{+}), \quad (10)$$

where k is an user specified value that reflects the drift level. Comparing the lower cusum and the upper cusum with the alert level H (typically chosen as 5), we can generate an alert for mean shift whenever the upper cusum or the lower cusum is outside the control range. Figure 6 shows an example of a drifting anomaly from CUSUM. Although the top graph of Figure 6 indicates that all the observed energy consumption data are within the control bounds, the bottom graph (CUSUM) indicates that on March 28th - 30th, there is a substantial increasing trend of actual energy consumption with respect to the model predicted energy consumption. This CUSUM analysis is an effective way to detect the drifting energy consumption which cannot be easily detected by control bound calculation.

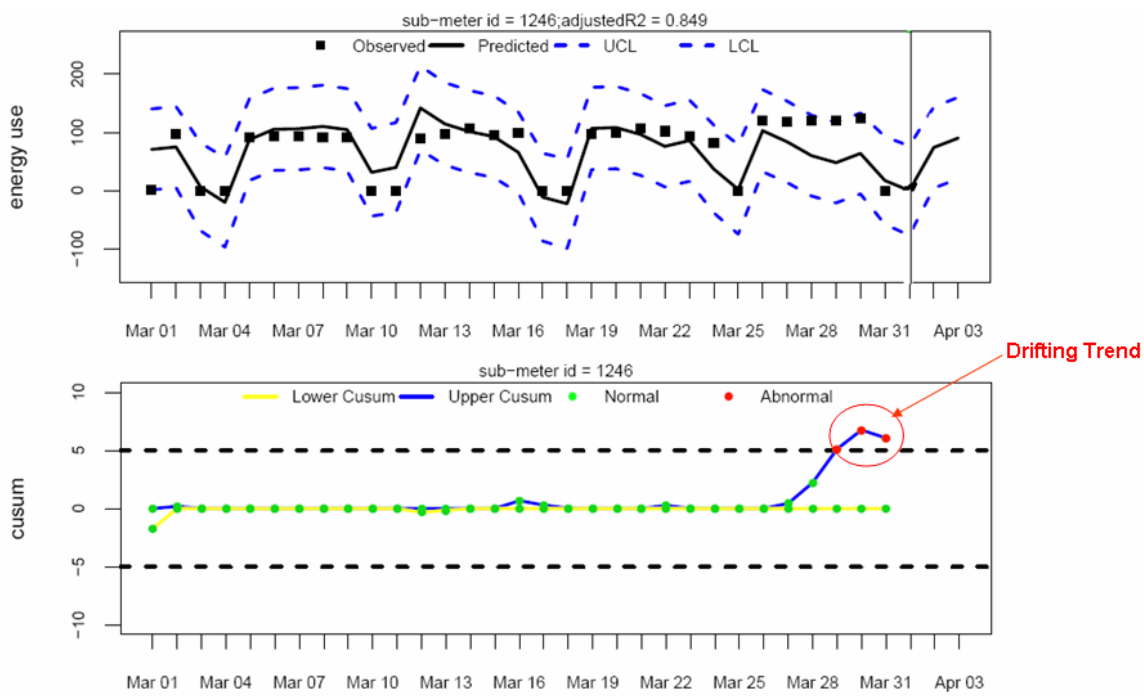

Figure 6: Anomaly Detection by Control Bounds and CUSUM

The statistical technique provides an integrated analysis for building heterogeneity, the weather dependent patterns and the temporal dependent patterns. It has wide applicability in anomaly detection, forecasting, energy efficiency analysis by energy type for building portfolios. We remark that, in contrast to the common practice of utilizing physics-based models to perform the Fault Detection and Diagnosis (FDD) in the building operation industry, the statistical modeling approach avoids the complexity of model calibration and provides a more flexible tool for such purpose. However, a statistical model is appropriate for bigger time resolutions, e.g., monthly, weekly or daily energy consumption while the physics-based model is effective for computing a dynamic profile (i.e., finer time resolution such as 1 minute, 15 minutes and hourly) of energy consumption.

5 Cloud Deployment

All the data, models, analytics, visualization and other useful information for energy performance of buildings can be very valuable for many people including those who own and manage the buildings and tenants of the buildings. However, the building energy models and analytics require substantially configured computing hardware and software including data interface programs, database, data integration and transformation, data warehouse, statistical engine, optimization engine, business intelligence/dashboard tool and visualization. Also depending upon the size of the data and amount of analysis there is a need to scale up and scale down the install base. Typical users, especially tenants of buildings such as students and public school custodians cannot afford such resources individually. As suggested, the cloud is a secure, flexible and scalable way of delivering such analytic services to many users by sharing all the resources in a cost effective and secure way among many users. Running this kind of energy service on cloud for multiple clients using the shared computing resource is actually more energy-efficient than providing the same service to individual clients with a traditional computing medium without the cloud. Depending on the number and roles of the users, an appropriate level of resource can be easily configured virtually on the cloud, and the level of access to resources can be adjusted on demand. Using cloud, all the building energy analytics can be readily accessible by building owners, facility managers, operators and tenants of buildings through web browsers without installing any special hardware, software, data interface and mathematical tools.

The cloud-based deployment is based on the Software as a Service (SaaS) model for software deployment. The i-BEE tool is hosted as a service provided to customers across the Internet. The decision to deploy the i-BEE tool on the cloud was to eliminate the need to install and run the tool on the customer's own computers along with alleviating any burden to the customer with regard to software maintenance, ongoing operation and support. The i-BEE cloud deployment is a hybrid of both public and private clouds that is designed to provide rapid access to security-rich, enterprise-class virtual server environments, well suited for dynamic workloads. The i-BEE cloud offers the capabilities to control access and configure security. The control access capability creates accounts, manages permissions, modifies user information, and more. The configure security capability manages the security of the virtual environment such as managing keys and passwords for the encrypted connections, and assigning instances to virtual private networks and virtual local area networks.

Even though the i-BEE tool resides on a public cloud platform where control of the cloud services to operate, access and secure is maintained through a separated authority, the IT infrastructure acts as a shared pool of computing resources i.e.; servers, networks, and storage that is rapidly provisioned for the benefit of a single organization. Furthermore, acting as a private cloud the ability to obtain automatic failover between hardware platforms is simpler than providing disaster recovery services that can be extended to i-BEE customers.

Managing data from multiple disparate sources requires security, privacy, policy, and governance. Policies vary by government, geography, and industry. The right mix of technology is important for enforcing different policy choices. Figure 7 below shows how the smarter building energy analytics were deployed on the cloud. There are two types of data sources that provide the data that is required for the models and analytics. One type of data is coming directly from the internet through web-service type of applications that include weather data from the local weather stations and grid pricing data, which can change dynamically over time based on energy demand and supply. The other type of data is data that describes building configuration, energy consumption, energy load, data from meters/submeters and sensors for each building. These data can come directly from a server in a building through the BMS or other central data depository, or manually in papers. All these data are brought into a server on the cloud and are cleaned and aggregated in a data staging area, and then stored in a data warehouse. The data in the data warehouse is used by various models and analytics described in the earlier sections of the paper including data mining, statistical model, spatial analyzer, heat transfer model, simulation and optimization, all of which reside on the cloud. The data, energy performance information, analysis results and recommendations are communicated to users through the analytics dashboard, which can be accessed via a PC web browser. or hand held device such as smart phone and tablet computer.

For this application of smarter building energy analytics, the data model is built such that a single warehouse can handle data from multiple clients. All building and energy related data is identified to belong to a particular client. Weather data can be shared across clients if multiple clients would be referring to the same location. The dashboard can easily identify data specific to a particular client and display just that data and information. Since some clients may not have large amounts of data, this allows many clients to share resources, thereby reducing individual costs. If there are security or privacy concerns, the data can be separated out into a separate instance. Scaling on the cloud is a fairly straightforward process. Once the base software, models, and analytics are installed, one can make an image of this "empty" system. This image can then be copied and deployed on another cloud instance, as long as they share the same operating system. This will require some slight configuration changes to allow it to work properly (IP Address/Hostname updates on various software, licenses and path updates if necessary). With this method, you can quickly get another empty instance up and running that is ready to accept and process additional client data once a previous resource has filled up. We have not run into a situation where one particular client has a large enough set of a data that we are not able to accommodate them on a single piece of hardware. Even a "large" amount of data for a particular BMS system, is quite small as far as being able to be processed and stored. If a case like this were to arise a partitioned database may be needed to accommodate the large data size. The rest of the software and processing would still be accommodated by a single instance.

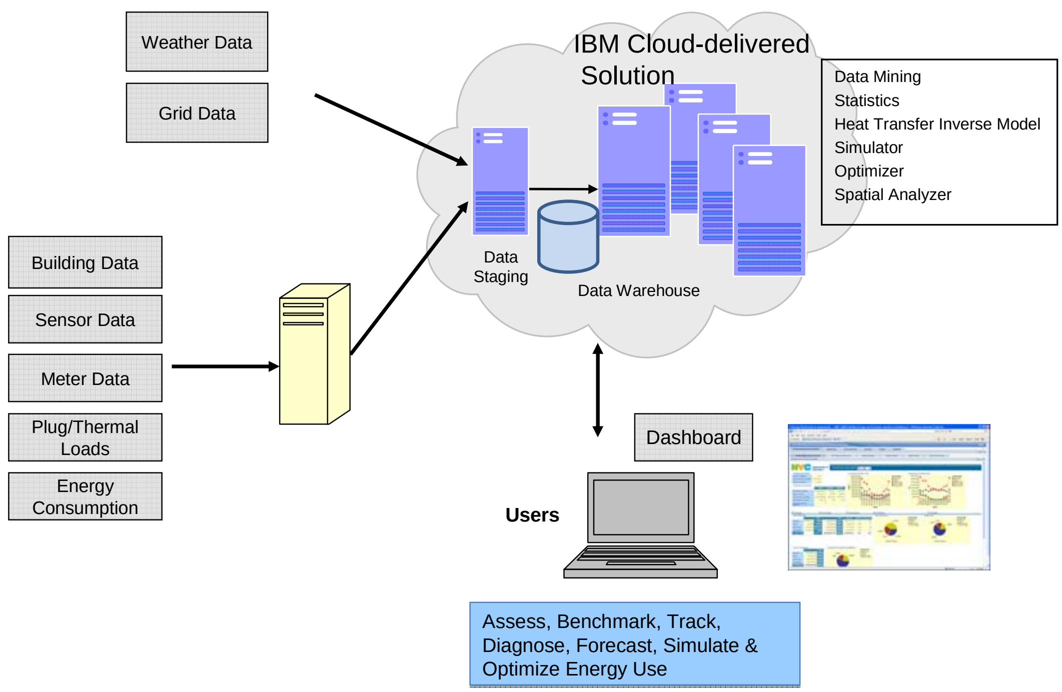

Figure 7: Cloud-Delivered Smarter Building Energy Analytics

6 Conclusion

The cloud is an efficient and effective way to deliver smarter building energy analytics to building owners, facility managers, operators and tenants of buildings by sizing, sharing and managing computational resources. We developed a smarter building analytics tool which provides useful information in assessing, benchmarking, diagnosing, tracking, forecasting, simulating and optimizing energy consumption in building portfolios, and identifying energy savings opportunities. The cloud-based tool has been deployed on the cloud for the K-12 public school buildings in New York City. The tool is also being used for NYC building operator certification (BOC) training classes (Bobker et al. 2011) and has received very positive feedback from both the instructors and the students. The tool is also being extended to analyze energy efficiency and simulate and optimize the operational alternatives of highly instrumented commercial buildings. We also plan to expand the use of the tool to more public school systems in the U.S. and other public buildings portfolios. Energy savings, energy efficiency and GHG emissions from buildings have become critical issues, and therefore there are many opportunities to be explored to make an impact, and cloud can facilitate the delivery of smarter building energy analytics to a wider spectrum of users. Let's build smarter buildings on cloud together.

References

Anderson, E. G., D. J. Morrice. 1999. A Simulation Model To Study The Dynamics in a Service-Oriented Supply Chain, In Proceedings of the 1999 Winter Simulation Conference.

An, L., R. Horesh, Y. Chae, Y.M. Lee. 2012. Estimation of Thermal Parameters of Buildings Through Inverse Modeling and Clustering, In Proceedings of the 2012 SimBuild Conference.

ASHRAE. 2009. ASHRAE Handbook: Fundamentals (SI Edition) Atlanta, GA, American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Beck, James V, and K.A. Woodbury. 1998. Inverse Problems and Parameter Estimation: Integration of Measurement and Analysis, Measurement Science Technology, 9, 839-847.

Bobker, M., J. Belfast, D. Leifer, J. Snowdon, Y. Lee, P. Dail, P. Weisner. 2011. Energy-data Dashboard: Designing for Use in NYC School, 2011. ICEBO 2011.

Brockwell, P. J. and R.A. 2006. Time Series: Theory and Methods. Springer, 2 edition.

Duncan, A. 1986. Quality Control and Industrial Statistics, Fifth Edition. Richard D. Irwin, INC.

eGRID, 2010. U.S. EPA's Emissions & Generation Resource Integrated Database. eGRID2007 Version 1.1. , Accessed March, www.epa.gov/cleanenergy/egrid.

EPA ESPM (EnergyStar Portfolio Manager) 2010. Greenhouse Gas Inventory and Tracking in Portfolio Manager, August 2009. Accessed February. http://www.energystar.gov/ia/business/evaluate_performance/Emissions_Supporting_Doc.pdf.

EPA ESPM Source Energy, Energy Star Performance Ratings Methodology for Incorporating Source Energy Use, 2009. Accessed September 2010, http://www.energystar.gov/ia/business/evaluate_performance/site_source.pdf.

EPA Greenhouse Gas Equivalencies Calculator, March 2010. Accessed September 2010. http://www.epa.gov/cleanenergy/energy-resources/calculator.html#results 9.

Lee, Y.M., F. Liu, E. Meliksetian, J.Snowdon, and M. Bobker, Building Performance Metrics Calculation and Visualization for K-12 School Buildings in New York City, In Proceedings of the 2012 SimBuild Conference.

Liu, F., H. Jiang, Y.M. Lee, J. Snowdon. 2011. Statistical Modeling for Anomaly Detection, Forecasting and Root Cause Analysis of Energy Consumption for a Portfolio of Buildings, IBPSA2011.

IBM 2011(a). City University of New York and IBM to Reduce Energy Consumption in Public School Buildings. http://www.ibm.com/press/us/en/pressrelease/34080.wss.

IBM 2011(b). IBM Works With McMaster University to Create First Canadian Energy-Smart Buildings on Campus, http://www-03.ibm.com/press/us/en/pressrelease/33838.wss.

Kissock, J. K., J. Haberl, and D.E. Claridge. 2003. Inverse modeling toolkit (1050rp): Numerical Algorithms. ASHRAE Transactions, 109:425–434.

PLANYC 2007. PlaNYC 2030: A Greener, Greater New York. http://www.nyc.gov/html/ planyc2030/html/home/home.shtml.

Silver, Cheryl Simon, R.D. DeFries. 1991. One Earth One Future, National Academy of Sciences.

U.S. Department of Energy. Buildings Energy Data Book. 2006. Accessed September, 2010. http://buildingsdatabook.eren.doe.gov/TableView.aspx?table=1.1.1.

U.S. Environmental Protection Agency, Climate Leaders Program, Direct Emissions from Stationary Combustion Sources, Appendix B, May 2008. Accessed March, 2010. http://www.epa.gov/climateleaders/documents/resources/stationarycombustionguidance.pdf.

Vison 2010, 2008. City Housing Hamilton Demonstrates Commitment to Energy Conservation, http://www.hamilton.ca/NewsandPublications/NewsReleases/2008News/01-17-08ka.htm.

Weather Warehouse, 2010, Accessed September, 2010, https://weather-warehouse.com/index.html.

World Energy Outlook 2009, Accessed September 2010. International Energy Agency.