Inverse Modeling of Portfolio Energy Data for Effective Use with Energy Managers

Honey Berk, Marco Ascazubi, Michael Bobker
CIUS Building Performance Lab, City University of New York
New York, New York, USA

Abstract

Development, features, and use of a tool for inverse energy modeling for a portfolio of municipal buildings is described. The tool is software-automated to enable batch-processing. Statistical fitness is automatically evaluated to provide a baseline for energy retrofit program measurement and verification. Particular attention is paid to tool outputs that support use by an end user, such as an energy manager, for initial facility diagnostic purposes to guide further investigation.

Introduction

This paper examines how large quantities of property portfolio energy data can be effectively modeled and presented for use by energy management professionals. Driven by greenhouse gas emissions reduction initiatives, cities have advanced in their collection of municipal energy use data. Implementation programs in municipal property portfolios require measurement and verification to validate the prudent use of public funds. This paper reports on work conducted over four years to develop an automated system for generating and using inverse-modeled energy data in a large municipal setting.

State of the Art and Deficiencies. While access to 15-minute electricity interval data is increasing, for larger cities, the volume of even monthly data can be a hindrance to its effective use.

The most widely used platform for energy data aggregation is ENERGY STAR Portfolio Manager (PM). While useful for comparative purposes, PM does not provide statistically-validated models for measurement and verification (M&V) of change programs such as energy retrofitting. For this purpose, inverse modeling at the Whole Facility level (IPMVP Option C) is the state of the art. Such models are commonly generated on a building-by-building basis.

Municipal Portfolio Programming. New York City has an aggressive energy management program for its approximately 4,000 municipal buildings, which includes energy standards for new construction and major rehabilitations, retrofitting, retro-commissioning, and training. Tasked with establishing an industry-standard approach to M&V for the portfolio's energy reduction initiatives, we quickly discovered the need to automate the inverse modeling process in order to reach scale in an acceptable timeframe.

Objectives and Modeling Approach

The fundamental objective of the work was to provide an industry-standard basis for evaluating NYC's municipal energy retrofit program. This process required a statistically-validated linear regression model for whole facility energy use, per the requirements of the International Performance Monitoring and Verification Protocol (IPMVP) and ASHRAE Guideline 14. The methodology set forth in the ASHRAE Inverse Modeling Toolkit (Kissock, 2003) was applied. It should be noted that, for retrofits that are not projected to achieve savings of at least 10% of total metered usage, Retrofit Isolation with parameter measurement (IPMVP Options A & B) is generally recommended, as whole facility savings cannot be reliably differentiated from the normal year-to-year fluctuations in building energy use. Future work will attempt to enhance the utility of the whole facility method through the use of interval data and associated methods.

Our objectives did not include improvements to the ASHRAE Inverse Modeling procedure. Rather, our primary goal was to facilitate batch processing of large sets of building energy data using this methodology. We purposely maintained the integrity of those calculations without modification so as to preserve the claim to "industry standard" method, as prescribed by our sponsor. To introduce calculation changes, even if improvements, would subject the entire effort to extensive scrutiny, review and delay.

During the course of the work, we realized that, in addition to serving as the basis for M&V, the method could also be used to provide diagnostic insights into building performance and potential systems improvements (Kissock 2004). This became an important objective of the work as the client agency Energy Managers are, at present, much more involved in energy efficiency project identification than M&V. This balance is expected to change over the next few years as more projects are implemented and the M&V program progresses.

A third, less developed, objective was to be able to use inverse models in conjunction with physics-based forward models. In work with other teams (Eicker, 2016; Schumacher, 2017) we discovered that the partitioning of baseload versus weather-sensitive loads in change-point models can be useful to validation and calibration of forward models. Simulated monthly consumption data generated through forward modeling can be run through the inverse modeling method, then compared to results based on actual data. Match of modeled results for baseload and weather-sensitive loads can then point to distinct variables to be tuned.

For example, poor calibration of modeled electricity may be due to air-conditioning (weather-sensitive) or lighting power density (baseload), or a combination of the two; use of the change-point models can help demonstrate when an appropriate calibration fit is achieved. We suggest that this procedure may be significant to improving the accuracy of forward modeling conducted on large numbers of buildings at an urban scale (Pang, 2013; Eicker, 2017) and may be useful as part of energy model auto-tuning techniques (Chaudhary and New 2017). .

Methods

After an initial two years of manual processing using industry-standard tools, including Energy Explorer (Kissock, 2000), a semi-automated batch-processing procedure was developed. Visual Basic (VB) scripts were used to facilitate import of CSV data files from the NYC utility database into a spreadsheet environment (MS Excel) and to clean and prepare data. Cleaned data files were batch-processed using Energy Explorer, and the output was used to recreate change-point linear regression models in Excel. A scripted algorithm used a series of three tests -- a shape test, a t-test and a data population test -- to select the best-fitting model (2-, 3-,4- or 5-parameter) (Paulus, 2014), and the model coefficients and statistical metrics were displayed along with the model visualization and compared to identify best and worst performers.

Batch-Process Automation. In Year 4, the modeling process was fully automated using a custom-built Python application, PyBEMA, linked to a database management system (MS Access). Scripts in Access prepare the data and perform several initial calculations, and simple queries allow for selected processing of data by date range, agency or other filters. Change-point models are generated in PyBEMA using a piecewise least-squares linear regression method, and processed data is then pushed to an enhanced Excel dashboard that runs the model selection algorithm, chooses the best-fitting model, combines regression results with building metadata, and creates a series of data visualizations (Figures 1 and 2) and tables with associated metrics for both electricity and thermal energy (natural gas/steam/fuel oil[1]) usage. Model outputs were systematically checked against independent runs of the same test data using Energy Explorer, to ensure accuracy.

Timeframes. Our early modeling efforts used monthly energy consumption data for a 12-month period, with approximate billing period start and end dates. The lack of availability of actual billing period dates was a significant source of uncertainty in these early models, as it was impossible to accurately match up outdoor air temperature (OAT), our independent variable, with energy consumption; this effect was especially pronounced during shoulder months, when OAT is most variable. This issue was addressed in Year 3, when we were provided with more granular utility data that included actual billing period dates. Although this improved the model output, another confounding challenge was the all-too-common presence of estimated meter readings in the electricity and natural gas data. Even one or two estimated data points in a 12-month period was enough to skew a model and negatively affect goodness-of-fit metrics.

While researching methods to address timeframe issues, we learned of the “sliding NAC analysis” technique (Lammers, 2011). This methodology uses typical meteorological year (TMY2) data to generate normalized annual consumption (NAC) data; the goal is to filter out the “noise” of variable annual weather patterns so that the “true” energy signatures of facilities can be discerned. The NAC data is then used to generate change-point models for sets of sequential 12-month periods, and changes in the model parameters (i.e., baseload, change-point(s) and cooling/heating sensitivity) from one model to the next are used as indicators of the factors that are driving energy consumption patterns.

We ran sliding NAC analysis on 24- and 36-month datasets for courthouse facilities and analyzed the changing parameters. Initial trials indicated the potential to contribute valuable diagnostic insights into building performance; so, the technique was added to the PyBEMA development roadmap for incorporation into future versions. One immediate result of this work, however, was our adoption of a longer usage period (24 months) for baseline modeling for performance analysis; we continue to use 12 months of pre- and post-installation data for Whole Facility M&V following IPMVP guidelines.

Data Visualization and Usability Testing. PyBEMA generates a data visualization “dashboard”, a sample of which is provided as Appendix A. The dashboard displays visualizations and metrics for one facility at a time, as selected via a dropdown list. Currently implemented in Excel, this user interface provides another important area for testing, beyond the quantifiable accuracy of results. To date, anecdotal user feedback has been collected from energy management personnel during one-on-one portfolio analysis review

[1] Metered fuel oil consumption data is not available at this time; as such, fuel oil data is only modeled as pertains to interruptible natural gas service.

meetings and is discussed further below. Formal usability tests of the dashboard are scheduled to be conducted later this year with NYC client agency Energy Managers; results are anticipated to be available for presentation in summer 2017.

We are currently working to migrate the dashboard user interface to a Python interactive visualization library that presents very large datasets in a (modern) web browser environment; a prototype will be completed in spring 2017.

Results

The PyBEMA application readily processes hundreds of facilities at a time, producing individual change-point models and associated diagnostics that can then be accessed in the standalone Excel dashboard (shown in Appendix A). The dashboard is designed to present the information that is most relevant for a high-level assessment of facility energy consumption, for both electricity and available thermal fuel data, such as site and source EUIs, time series graphs, change-point models, energy use breakdowns and CO_2e emissions.

Statistical Metrics. Figure 1 shows typical output, in this case a 3-parameter cooling (3PC) model for electricity, with data points and best-fit line.

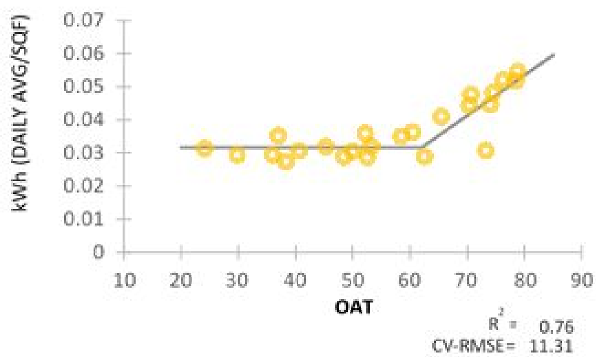

Figure 1. Sample visualization: electricity change-point model for a NYC municipal office building.

Figure 2 shows further elements, based on combined electricity and thermal energy, using LEAN methodology (that are automatically displayed as part of the dashboard interface.

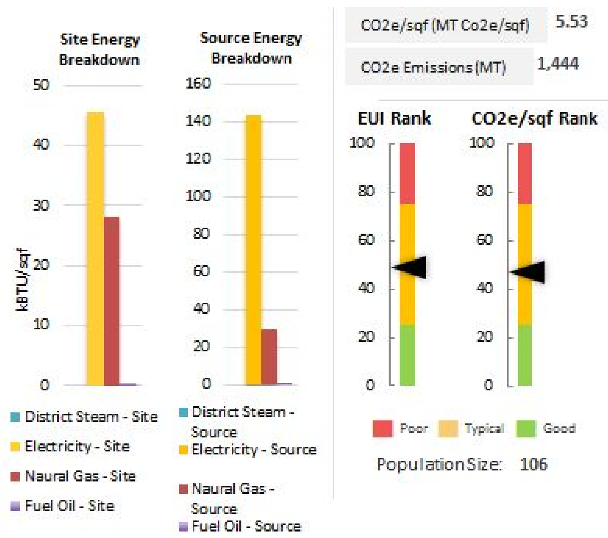

Figure 2. Sample visualization: combined energy metrics for a NYC municipal office building.

Note the statistical measures of goodness-of-fit in the lower right of Figure 1; these are used to evaluate the baseline models, in line with industry standards. First, the coefficient of determination (R^2), represents the proportion of variation in the dependent variable that can be explained by the independent variable. Here, R^2 indicates how well the regression model can predict future energy consumption based upon OAT. Our work requires that $R^2 \geq 0.75$ for the model to be valid; lower values signify that factors other than OAT are driving energy consumption.

The second statistical metric used is the coefficient of variation of the root mean square error (CV-RMSE), defined as the root mean square error (RMSE) divided by the mean of the measured data. CV-RMSE describes how well the model fits the data; the higher its value, the more scattered the data points appear around the best-fit regression line and vice versa. Our work requires CV-RMSE $\leq 20\%$ for the model to be valid; however, we note that we frequently find these values to be higher in natural gas/steam models. Similar findings have been attributed by others to high variations in the amount of outdoor air and related factors during shoulder season months (Matutinovic, 2014); our research into this phenomenon is ongoing.

Our current work in this area involves residual analysis and testing of additional statistical metrics that may be used to fine-tune model selection and the identification and elimination of data outliers.

Lean Energy Analysis (LEA). Our work incorporates the statistical Lean Energy Analysis (LEA) technique (Kissock, 2004; Abels, 2011), which uses regression coefficients from the change-point models to compare performance across portfolios of buildings. LEA is applied to peer facilities based upon Commercial Buildings Energy Consumption Survey (CBECS) building categories (e.g., office buildings, hospitals or schools). Facilities are ranked based upon the value of model parameter coefficients: baseload, change-point(s), and heating and cooling slopes. Quartiles are used to flag best and worst performers for each parameter within a peer group, and results are visualized using a graphic rating scale (Figure 3).

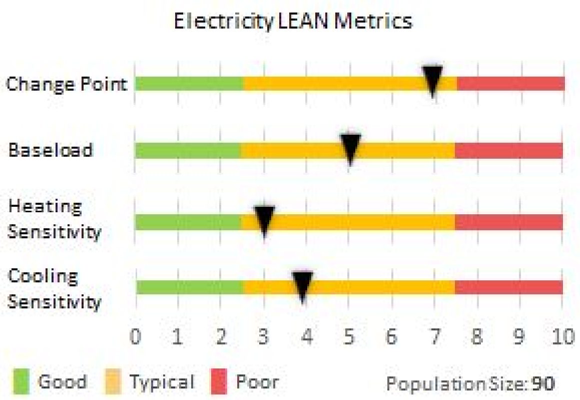

Figure 3. Sample visualization: electricity graphic rating scale used to display facility rankings.

LEA is also used to generate analysis of NYC client agency portfolios, such as all facilities belonging to the NYC Sanitation Department or the NYC Fire Department (FDNY). In this use case, dashboard metrics are not used to gauge comparative performance between facilities, as there is no basis for comparison across building typologies (e.g., FDNY vehicle repair shop vs. FDNY firehouse). Rather, the dashboard affords an agency Energy Manager a comprehensive look at the entire portfolio, and helps pinpoint the largest energy consumers and identify potential efficiency opportunities.

Note that we have observed that agency Energy Managers are more interested in the ranking of a given facility against its NYC agency peers, rather than against a wider population of peers nationally, as indicated by PM scores.

Diagnostics. The change-point models, as discussed further below, can be used for building performance diagnostics to identify areas of likely performance deficiencies for further investigation. For example, a high heating or low cooling change-point may indicate a need to investigate a building for infiltration, ventilation rates, outside air damper operations and/or building insulation conditions; a relatively high baseload might, instead, suggest investigation of equipment scheduling and continuous loads.

The ability to characterize and filter buildings by change-point, slope and baseload parameters enables the creation of quad charts that sort and help identify best and worst performing facilities along specified dimensions. Figure 4 shows the comparison of multiple facilites with regard to cooling change-point and cooling senstivity; the quadrants are delineated by the median value for each of these parameters. As an example, facilities with a high cooling change-point and low cooling sensitivity (those in the upper left quadrant) are considered best performers; while, those in the lower right quadrant are considerd worst, as they exhibit a low cooling change-point and high cooling sensitivity.

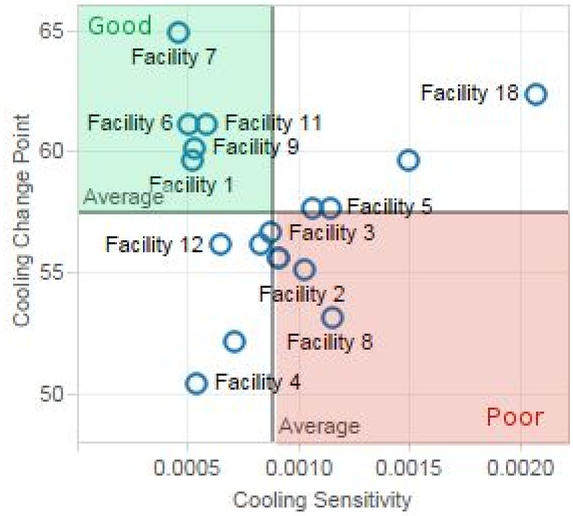

Figure 4. Sample visualization: quad chart displaying multiple facilities by change-point vs. cooling sensitivity (slope).

This portfolio-view quad chart format has yielded very positive response in initial informal trials with energy managers. Further work in the use of multiple parameters for performance diagnosis is ongoing (Ascazubi, 2017).

Discussion and Analysis of Results

PyBEMA will attempt to model all facility data included in a batch run, but there are usually a number of models that are not modeled due to a poor R^2 value (< 0.75). As mentioned previously, a low R^2 indicates that the independent variable (OAT) does not adequately explain the energy usage, so the model is invalid. In our earlier work, we also used the CV-RMSE value as a validity threshold; however, the higher values we tend to see in thermal energy models caused us to reconsider the affected models. Facilities that do not pass are listed, with associated metrics, under separate tabs (i.e., Poor Electricity Models, Poor Fuel Models).

Interpreting Poor Fit. There are some common causes for poor models, some of which result from data quality issues, and others that may indicate something significant about building operations.

We do not usually see facilities with missing data, but we do see estimated utility meter readings quite often. Estimated readings show a characteristic pattern readily identified in the time series chart that contributes to poor fit statistics. Interruptible gas with a lack of fuel oil consumption data is another fairly common situation that can be addressed: fuel oil delivery data for a 12-month period is aggregated and allocated to specific (known) days and hours of interruption, then modeled together with the natural gas data. Fuel oil-only sites represent a more difficult problem, as in most cases only delivery history is available; experiments with oil metering schemes have revealed the complexities in reaching a uniformly applicable solution for digitized metering.

Once data problems are ruled out, operational issues must be considered. With public school facilities, for example, vacation and summer closures commonly result in poor electricity data fit with OAT alone; a second independent variable, such as occupancy or operating hours, needs to be introduced. Multivariate regression analysis is enabled in the PyBEMA tool, but forward-looking work will involve testing its use with the most appropriate variables and applicable situations.

Interpreting Baseload. Baseload is identified as a corollary of change-point, and is quantified as the value of that point at the y-axis (i.e., y-intercept). In order to compare baseloads across facilities, usage is normalized by gross square footage. For certain kinds of facilities, normalization by building volume may be preferable; this is a construct that will be tested in future work.

If there is a natural gas or steam baseload, it is most often a year-round service hot water load, often relatively small in commercial/institutional facilities. A high baseload of this type can usually be interpreted to represent one of a relatively small set of services: domestic hot water, pool heating, cooking or laundry. Thermally-driven cooling (i.e., steam turbine chillers, gas engine-driven chillers, absorption units) may produce what looks like a high fuel baseload (and would most likely be modeled using a 5-parameter change-point model); techniques are being developed to diagnose and separate this case from high thermal baseload.

Electricity baseload usually represents the predominant portion of electricity use, typically comprised of lighting, plug loads and ventilation (i.e., fans), so high baseloads can usually be traced back to unusually high usage in one or more of these functions; a facility with a large data center or other 24/7 operation (e.g., emergency services call centers) will typically have a high baseload.

The absolute value of baseload (per square foot) is a useful indicator for the energy manager. The relationship of baseload to heating and cooling loads can also be used to interpret building equipment usage and operations. The total relative usages of baseload, heating and cooling can help guide the energy manager on where to prioritize efforts.

Interpreting Change-points. Change-point is analogous to (but not exactly representative of) a facility's thermal balance point, above which energy is used for cooling and below which energy is used for heating.[2] A change-point is considered poor when it is too low for cooling (say, 55°F) and too high for heating (say, 65°F), indicating unusually extended hours of heating and/or cooling. With cooling, at lower temperatures the facility should be taking advantage of free cooling from economizer operation at air handling units and/or chiller plant cooling tower; a water-side economizer cycle uses incremental energy, but much less than mechanical cooling, and an air-side economizer even less so. In looking at combined energy, the dashboard displays both electricity and natural gas/steam models on the same plot, so that instances of simultaneous heating and cooling can be identified (i.e., when heating change-point is higher than cooling change-point).

Interpreting Slopes. In a model with cooling energy, the slope of the line indicates how much additional energy is consumed as the outdoor air temperature increases; conversely, in a model with thermal energy, the slope indicates how much additional energy is consumed as the outdoor air temperature decreases. The steepness of the slope can thus provide an indication of a building's overall heating or cooling efficiency, but it is important to consider other factors, like the total number of months of heating or cooling.

It is tempting to interpret this slope as the mechanical system efficiency (Donnelly, 2013), but this is incorrect, as the building ventilation and conduction conditions are also implicated. Some of our work seems to suggest that the change-point can be a dominant factor in the magnitude of the slope. This is to say that, given an energy use for the maximum load condition, the slope is then established by the temperature location of the change-point. We are currently testing this concept with a combination of forward and inverse modeling. It has been suggested that, in order to compare overall building efficiencies by use of the temperature-sensitivity slope, we must first normalize the respective change-points. Future publications will address this analysis.

Initial User Experience Testing Results. As described above, we are in an early stage of testing with users. Initial experience indicates that, when presented with change-point model visualizations, users initially focus most strongly on the value of the heating/cooling change-point and the magnitude of heating/cooling slopes. Change-points tend to be interpreted as directly indicative of thermostat set-points for initiating heating or cooling; effort is necessary to ensure a proper understanding as representing something more like a building balance point, indicative of mass and insulation in building construction.

---

[2] For electricity, this is true for 3-parameter cooling, 4-parameter and 5-parameter models; for natural gas/steam, this is true for 3-parameter heating, 4-parameter and 5-parameter models.

Similarly, users tend to interpret slopes as direct proxies for heating/cooling system efficiency, especially when peer group buildings have significantly contrasting slopes. We have found in our work that this characterization is not entirely accurate, as the slope may be affected by the change-point temperature, the total number of months with seasonal energy consumption, and other model characteristics.

Finally, we have observed that this primary focus on change-point and slopes often causes users to overlook the magnitude of the baseload, visualized as the value at which the change-point intercepts the y-axis. This is of concern, since the baseload often represents the largest area of energy use, especially for electricity.

Conclusion

A programmed solution is developed for usability of energy data sets in large building portfolios. An automated batch-processing tool has been developed and tested to produce statistically-validated baseline models from monthly energy consumption data. The models provide the basis for industry-standard whole facility measurement & verification (M&V) for a portfolio-wide energy retrofit program. Beyond this primary objective, the energy data tool also provides energy managers with evaluative indicators for individual building and portfolio-wide comparative energy use, to support and improve decision-making and focus for investigations. Testing is in process with energy managers and further versioning is anticipated.

Acknowledgements

We gratefully acknowledge funding, supply of data, staff time and general support for this work from NYC DCAS Energy Management.

References

Abels, B., Sever, F., Kissock, K., Ayele, D. (2011). Understanding Industrial Energy Use Through Lean Energy Analysis*, SAE Int. J. Mater. Manuf. 4, 495e504.*

Ascazubi, M. (2017). Combinatorial Parameter Logics for Building System Performance Diagnosis. CUNY BPL unpublished working paper.

Chaudhary, Gaurav, Joshua New et. al. (2017) Evaluation of "Autotune" Calibration Against Manual Calibration of Building Energy Models. Applied Energy v.182

Donnelly, M. (2013). LEAN Energy Analysis: Using Regression Analysis to Assess Building Energy Performance. Johnson Controls Institute for Energy Efficiency.

Efficiency Valuation Organization (2016). International Performance Measurement and Verification Protocol (IPMVP).

Eicker, Ursula (2016). Personal communication.

Eicker, Ursula (2016). Urban-scale Energy Modeling Workshop, CUNY Advanced Science Research Center.

Eicker, Ursula et.al. (2017) Thinking Local, Acting Global: Urban-scale Modeling for Global City Governance SimBuild2017 (accepted)

Kissock, J.K. (2000). Energy Explorer Data Analysis Software, Version 1.0. University of Dayton, Dayton, OH.

Kissock, J.K., Haberl. J. and Claridge, D.E. (2003). Inverse Modeling Toolkit (1050RP): Numerical Algorithms. *ASHRAE Transactions, Vol. 109, Part 2*.

Kissock, J.K. and Seryak, J. (2004) Lean Energy Analysis: Identifying, Discovering and Tracking Energy Savings Potential, presented at the Proc. SME, Livonia, MI, USA, 2004, no. October, pp. 1–11.

Lammers, N., Kissock, J.K., Abels, B. and Sever, F. (2011). Progress with Normalized Energy Intensity. *Mechanical and Aerospace Engineering Faculty Publications.* Paper 163.

Matutinovic, L. (2014). Using Calibrated Energy Models to Help Understand, Manage and Improve Existing Building Energy Performance, presented at eSIM Conference, Ottawa, Ontario, Canada, 2014: IBPSA-Canada.

Pang, Xuiufeng,T. Hong and M.Piette (2013). Improving Building Performance at Urban Scale with a Framework for Real-time Data Sharing LBNL-6303E

Schumacher, Jurgen et. al. Forward And Inverse Modelling Of New York Buildings In An Urban Scale Simulation SimBuild2017 (acceptance pending)

Sever, Frank, K .Kissock, D. Brown, S. Mulqueen (2011). Estimating Industrial Building Energy Savings Using Inverse Simulation. *ASHRAE 2011 – 86073*.

Appendix A: Dashboard User Interface