2014 ASHRAE/IBPSA-USA
Building Simulation Conference
Atlanta, GA
September 10-12, 2014

SPREADSHEET TOOL DEVELOPMENT FOR VISUALIZING BUILDING PERFORMANCE AND SIMULATION DATA TO HELP CALIBRATING MODELS

Simon Sansregret[1], Karine Lavigne[1]
[1] Laboratory of Energy Technologies, Hydro-Québec Research Institute,
Shawinigan, Qc, Canada
sansregret.simon@lte.ireq.ca, lavigne.karine@lte.ireq.ca

ABSTRACT

One of the most important challenges in calibrating a building energy model is to visualize and analyze a significant amount of timeseries data in a practical and meaningful way. The paper presents a spreadsheet tool under development that addresses this challenge. The tool is based on new features of Excel™ 2010 and later version that use efficient data filtering capabilities. The tool populates different worksheets that address specific calibration tasks such as analyzing time domain profile differences, selecting short periods to calibrate, comparing start-stop schedules on daily profiles and quantifying the fit of the model with statistical indicators. The paper will discuss the functionalities of the tool by presenting a case study of the calibration of an EnergyPlus model using 15-minute total electric consumption measurements.

INTRODUCTION

Calibrated Building Energy Models (BEM) can provide numerous benefits to on-going building operations and inform retrofit measures and their feasibility. However, calibrating a BEM, especially in hourly and sub-hourly measurements, is not a trivial task. One of the most important challenges is to visualize and analyze a significant amount of timeseries data in a practical and meaningful way in order to help modellers improve the precision of their model. Most of the time, the task of comparing simulation results and measurements has the side effect of finding building operations problems.

Visualization techniques to compare measured and simulated data has been documented in several publications. Guideline 14-2002 (ASHRAE, 2002) provides specific graphical comparison techniques such as weather day type 24-hour profile plots or three dimensional surfaces. Reddy (2006) also reported a number of data visualization techniques such as timeseries plots, simple bar charts, scatter plots and comparative 3D surface plots. More recently, for the comparison of two large data sets, Raftery & Keane (2011) proposed the use of daily carpet-contour plots on the relative errors between measured and simulated data. Also, an innovative approach for classifying measured and simulated daily profiles based on a clustering technique was proposed by Fournier & Lavigne (2010).

On the other hand, some initiatives have been undertaken in recent years to develop free tools to manage and plot specific graph formats for measured and simulated building performance data. All of the following tools have interesting capabilities and their own advantages. Universal Translator (PG&E, 2008) stores imported data from loggers into an SQL server service. It allows to manipulate data (synchronization, interpolation due to missing data, data filtering), proposes graphs (XY and scatter plots) and analysis tools (economizers, lighting, plug loads). ECAM (PNNL, 2013) is an Excel add-in that allows to plot specific graphs related to building performance data with filtering capabilities and the calculation of normalized metrics. VizTool (LBNL, 2007) was intended to manage and plot building performance information from different data sources in a number of charting formats. Also, Dview (NREL, 2010) plots hourly timeseries from a text file in different time intervals and graph formats.

However, the authors faced some challenges when visualizing data for the calibration of specific building models using 15-minute total electric consumption measurements. None of above tools are designed to support both simulated and measured data and none of them are able to reproduce, with the desired flexibility, the visualization techniques described earlier. O’Donnell et al. (2013) also reported the lack of a user-friendly data visualization tool for comparing simulated and measured data.

To overcome those limitations, the development of a spreadsheet tool was initiated. The idea was to develop a flexible viewer where the user can easily adapt it to different projects. The tool, which is planned to be freely distributed to the community, populates different worksheets that address visualization techniques especially for calibration tasks.

TOOL DESCRIPTION

The tool, called VizBEM, is based on a new feature in Excel™ 2010 and later versions that use efficient data filtering capabilities. It was inspired by the ECAM tool that presents very interesting views of building performance data using pivot table functionalities. The use of the Excel™ pivot table functionality with its new filtering capabilities, called slicers, provides a powerful way to manipulate data in an efficient and flexible way. Excel™ has pro and cons, but it undoubtedly provides the user with a lot of flexibility. Because it is a well-known application, its users can easily customize their graph formats or create new worksheet reports. However, the storage of large datasets, the limited graphical capabilities and dealing with different updates and language versions of Excel™ can be cumbersome.

VizBEM is an Excel™ template file with macros (VizBEM.xlmt) allowing to rapidly create new project files with custom timeseries data. It contains a *Data* worksheet that includes all the timeseries data to which pivot tables of report worksheets are connected. At the present time, the timeseries data are simply copied/pasted on the *Data* worksheet by the user. The *Configuration* worksheet allows to update all pivot tables with a single command and to configure a group of settings, such as identifying variables to plot on second y-axis graphs.

Each of the remaining worksheets proposes a specific report to analyze the data in a different way. Figure 1 illustrates an example of a report worksheet. These report worksheets usually contain a graph with a checkbox list at the left to which the user may want to add or remove timeseries data. They also contain a number of customizable slicers at the top of the graph that enable powerful filtering capabilities. This setup allows the user to dig very rapidly into requested information. The following report worksheets are included in the VizBEM template file. They are presented in the case study section through Figures 2 to 7.

***TimeFrame* worksheet**: plots selectable timeseries data with advanced filtering options and based on different time intervals (timesteps, hourly, daily, weekly, monthly or annually). The graph format automatically switches from a line chart to a bar chart according to the selected time interval. This report provides a quick view for a preliminary analysis.

***WeeklyProfile* worksheet**: plots selectable timeseries data using weeks as a series. Two graphs on the same report allow to analyze other selected timeseries at the same time. This report is particularly useful for selecting week periods to calibrate.

***DailyProfile* worksheet**: plots selectable timeseries data on daily profile data using customized fields as the series such as day of the month, day of the week or daily bin temperature. This view is interesting in order to compare the operational schedules between the measured and simulated data.

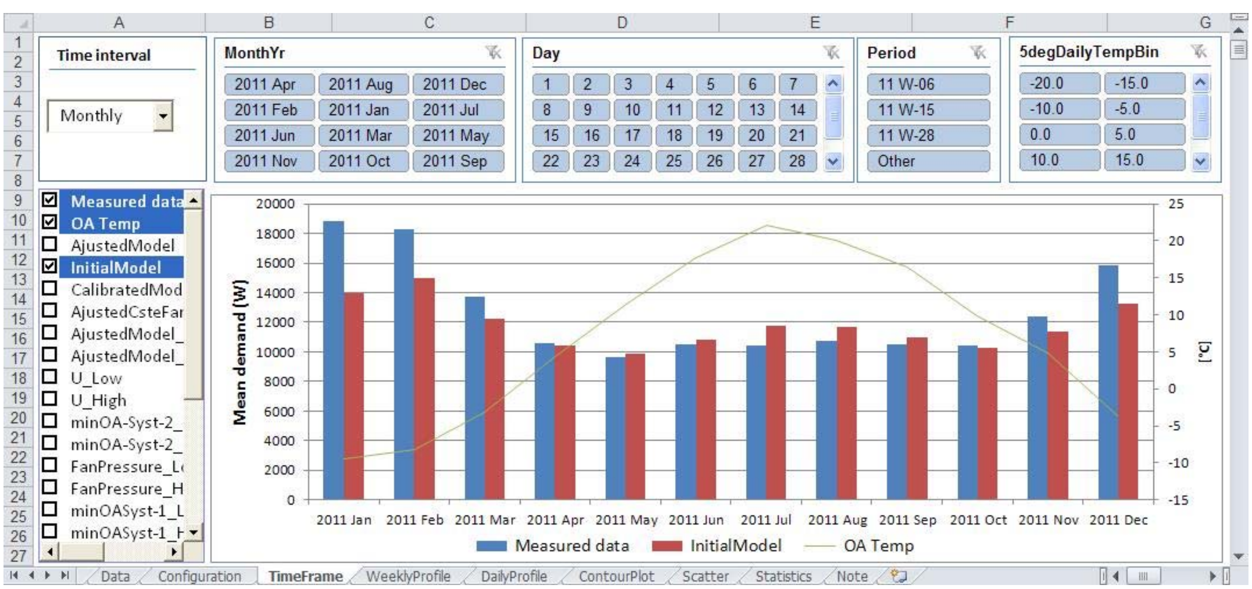

Figure 1 Screen capture of *TimeFrame report* worksheet –*Monthly* Time interval

.***Scatter* worksheet**: plots selectable timeseries on a scatter chart against an independent timeseries selected by the user. Also, the graph can show data through customized series such as occupied or non-occupied periods and data can be aggregated using the time interval selection.

***Statistics* worksheet**: computes and plots the normal mean bias error (NMBE) and coefficient of variation of the root mean square error (CVRMSE) between two selected timeseries. The evaluation of these statistical indicators takes into account the tool's filtering capability (for example, according to specific time periods and the selected time interval). These indicators are plotted on bar graph and compared to targets that can be modified by the user. Also, the report plots the bias errors between the selected timeseries according to the period and time interval selected.

Flexibility

VizBEM was developed to be flexible enough so that the users can modify and create their own report worksheets based on their project specificities. The users can easily copy, rename and modify the majority of report worksheets. When the data source is updated, all the report worksheets, including the customized ones, are refreshed with proper list fields and timeseries data. An effort was made to limit the use of VBA codes behind the interface in order to enable the user to easily make modifications.

The filtering capabilities can be easily adapted. The *Data* worksheet contains specific fields used expressly for filtering purposes. The users can easily add or modify these types of fields according to their project and add new slicers to the report worksheets. For example, an "Operation Schedule" field can be added with the "ON" or "OFF" value at each timestep. It is then easy to add a slicer or a series to the graph of a report worksheet to analyze specific timeseries according to the operation schedule. Here are other examples: nights/days, day of the week, pre/post retrofit and daily bin temperature. The filtering function can also be applied to remove or to highlight time period; for example, shutdown period, peaking hour or specific selected week for calibration purposes. The case study shown below exposes some examples of the use of these fields for filtering purposes.

The timeseries on the *Data* worksheet can be as low as 1-minute intervals. Tests were performed with three years of data on 15-minute intervals (more than 100,000 lines) with more than 50 custom fields. Even if the file size exploded (> 100 Mo), the navigation and filtering capabilities did not significantly depreciate on a standard PC with 3 Go of random access memory. To deal with a higher amount of data, a direct connection to a database should be considered. One irritant was the duration of the update process for all pivot tables, which could be significant in the case of a larger file.

The *Configuration* worksheet lets the user easily define calculated fields to add to the available field list. The calculated fields, which are treated directly in the cache memory, are obtained by an equation involving other fields. For example, a calculated field defined as the difference or the relative error between two timeseries could be defined and easily analyzed on different report worksheets. This functionality allows to efficiently analyze the difference between timeseries.

However, the user must be proficient with Excel™ spreadsheets in order to take advantage of all the tool's flexibility. Even if the tool is relatively simple to operate, a certain learning period is required, as is the case with all software applications.

CASE STUDY

In order to describe the capabilities of VizBEM, a case study of a calibration process is presented. The objective of the case study was to obtain a detailed calibrated energy model of a specific building in order to assess the impact of different demand response strategies. The building selected for the study was a small single storey bank branch of 426 m² built in 2009. It is located in Quebec, Canada, where the climate is cold, which explains the high power consumption of about 260 kWh/m² per year. Three electric rooftop units cool and heat this all-electric building. Only the total electricity consumption measurements in 15-minute intervals were used to calibrate the model, no sub-metering or data from the building's automated system were available. A detailed on-site survey was conducted and an initial EnergyPlus model was created. The following section describes the specific use of the VizBEM worksheet reports in different phases of the calibration process.

Analyzing Measured Data

Before attempting to get a highly detailed energy model, an analysis of the measured data was conducted using several report worksheets, such as the worksheets illustrated in Figure 1 and Figure 2. It was important to get a reliable and predictable set of measured data in order to calibrate the model. Figure 2 shows the *ContourPlot report* worksheet. A weekly surface chart of the selected "Measured data" field for all the weeks of 2011 (w-01 to w-52) is depicted in a single view. The user can easily switch from a weekly to a daily surface chart and use slicers to filter the data to plot (ex. season, weekday, etc.). The graph shows, at a glance, that there were no activities on Sundays and a high consumption level at the start-up of the systems in the winter, especially on Monday mornings. Also, the building's hours of operation were extended during the month of February (w-05 to w-09). This information was taken into account in the input schedules of the energy model.

Comparing Measured and Simulated Data

After a first energy model had been obtained, an analysis to compare measured and simulated data was conducted. First of all, operation schedules were compared using the *DailyProfile* report worksheet, as shown in Figure 3. This report allows to plot the mean daily profiles of the selected fields according to customized filters. For example, on Figure 3, the "Measured data", "Initial model" and the "CalibrationModel" profiles were selected and filtered with the "2011 Jan" and "Friday" items. Indeed, the mean profiles of the 4 Fridays of the month, January 7th, 14th, 21st and 28th of 2011 were depicted for each field. This view shows an offset of 30 minutes between the start/stop of the operation schedule (meaning, in this particular case, the HVAC system) between the measured data and the initial model. This offset was corrected in the calibrated model, as shown on Figure 3.

The *Scatter* report worksheet was used to plot the difference between the initial model and the measurements against the outside air temperature. The time interval was set to *daily* and the data was filtered by selecting *Not occupied* under the Schedule slicer. Figure 4 shows the results. The resulting graph clearly indicates that the initial model underpredicted the consumption as the temperature became colder. This information confirmed that the parameters affecting the temperature difference loads, such as the envelope U value and outside air infiltration rate, should be adjusted. This report worksheet report is very interesting in that it shows an interaction with an independent variable established by the user with custom filtering elements.

Selecting Calibration Periods

Selecting short calibration periods is beneficial. It reduces the simulation running time when performing parametric/optimization runs. It also allows to select periods with a predictable pattern and to only focus the analysis on specific periods, which is important when there is a submersion of data. The *WeeklyProfile* report worksheet was set up to facilitate the selection of weekly periods. In the calibration of the bank model, three specific weeks in the winter (week 06 of 2011: Jan. 30th to Feb. 5th), summer (week 28 of 2011: July 3rd to 9th) and shoulder season (week 15 of 2011: Apr. 3rd to 9th) were chosen, as shown on Figure 5. After this selection, a new field called "Period" was updated in the data source with an indication of the selected week as a value for each timestep. This provides the user with the flexibility to add a "Period" slicer at the top of each worksheets in order to easily focus the data analysis on these selected periods. This is a simple example of customizing the filtering options in VizBEM.

Analyzing Parametric Runs

A large set of parametric runs was performed in order to see the impact of varying predefined parameter values on the bank branch model results. This task is very important for the selection of the most influent group of parameters before launching optimization calculations and for the detection of any model definition problems. Figure 6 shows the results of two parametric runs related to the minimum outside air of the main rooftop unit with high (50%) and low (20%) values using the *TimeFrame* report worksheet. The use of the "Period" slicer allowed to simultaneously analyze the influence for the three selected calibration week periods. It was observed that the outside air influence had a huge impact during HVAC operations for the selected winter week (11 w-06), but the effect was negligible for the selected summer week (11 w-28).

Evaluating Statistical Indices

The calibration process requires many iterations before obtaining a model with acceptable accuracy. The limit proposed by ASHRAE Guideline 14 [ASHRAE, 2002] was selected for this project, where the NMBE should be inside +/- 10% and the CVRMSE lower than 30% when evaluated on hourly time intervals. The *Statistics* report worksheet assesses and plots these statistical indices as depicted on Figure 7. It shows that the initial model was not declared calibrated because the CVRMSE computed on hourly intervals for the three selected calibration weeks was higher than the imposed limit of 30%. The bars are automatically coloured in red on the graph when their value are higher than the specified limit. The bar graph indicates that "CalibratedModel1" was well below the target criteria of Guideline 14 for hourly interval assessments of the NMBE and CVRMSE. This report is interesting to qualify improvements of the calibrated model.

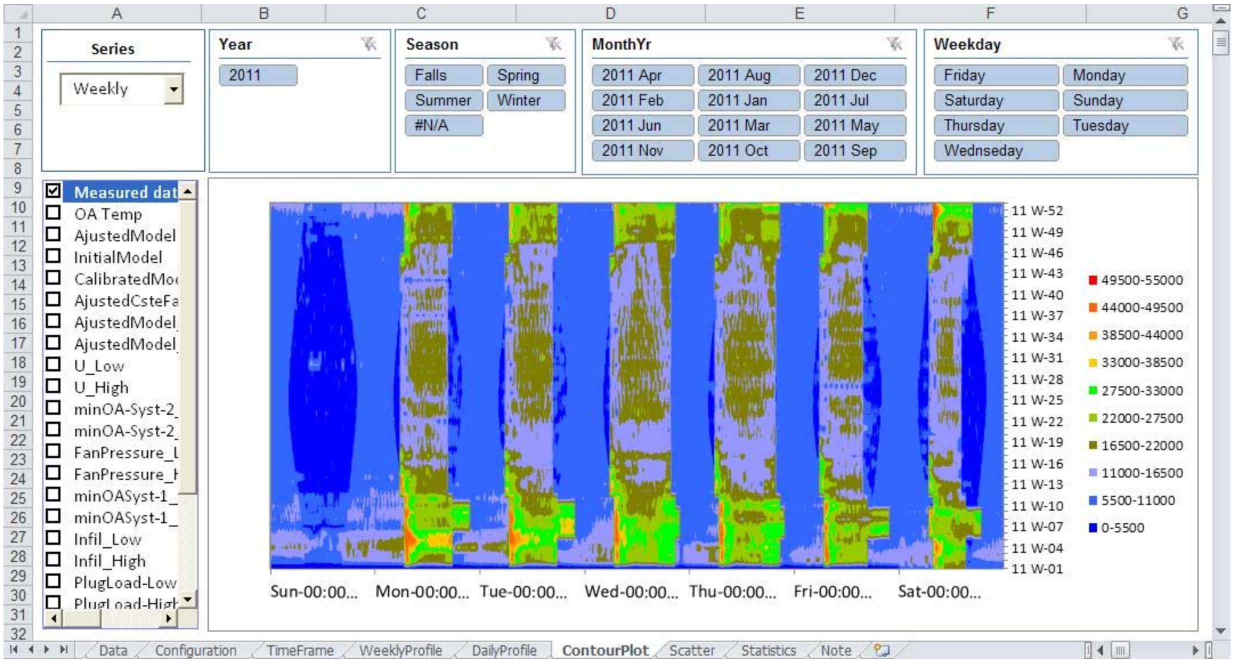

Figure 2 Screen capture of *ContourPlot* report worksheet

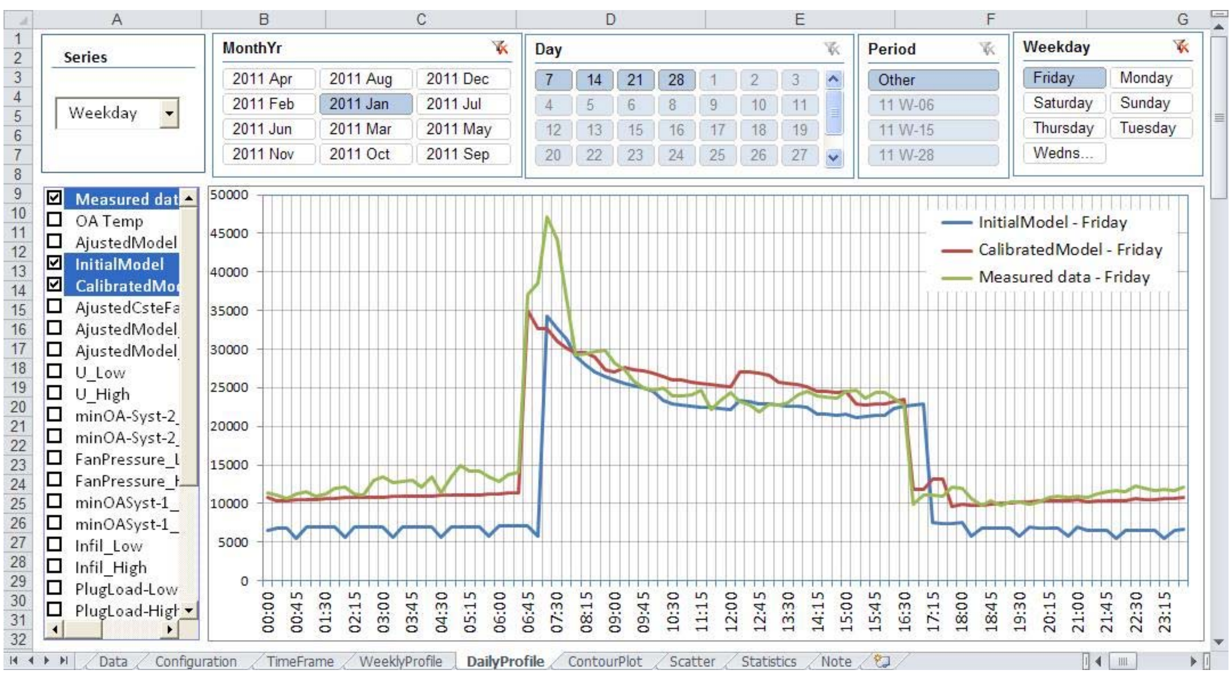

Figure 3 Screen capture of *DailyProfile* report worksheet

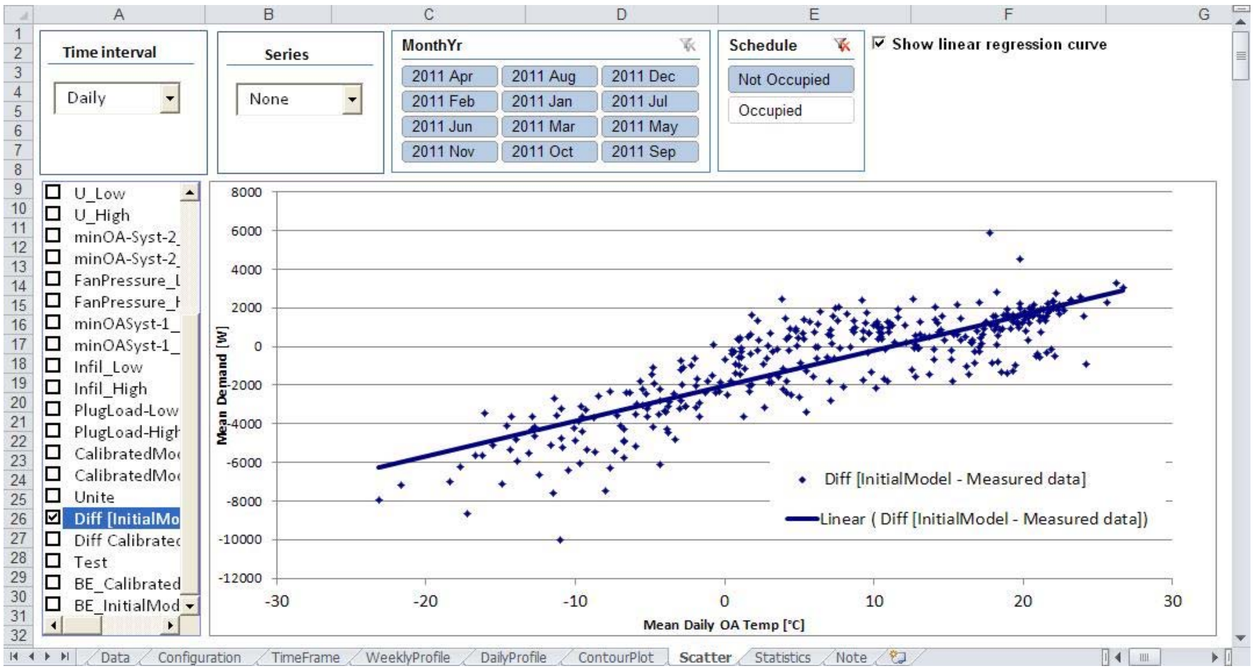

Figure 4 Screen capture of *Scatter* report worksheet

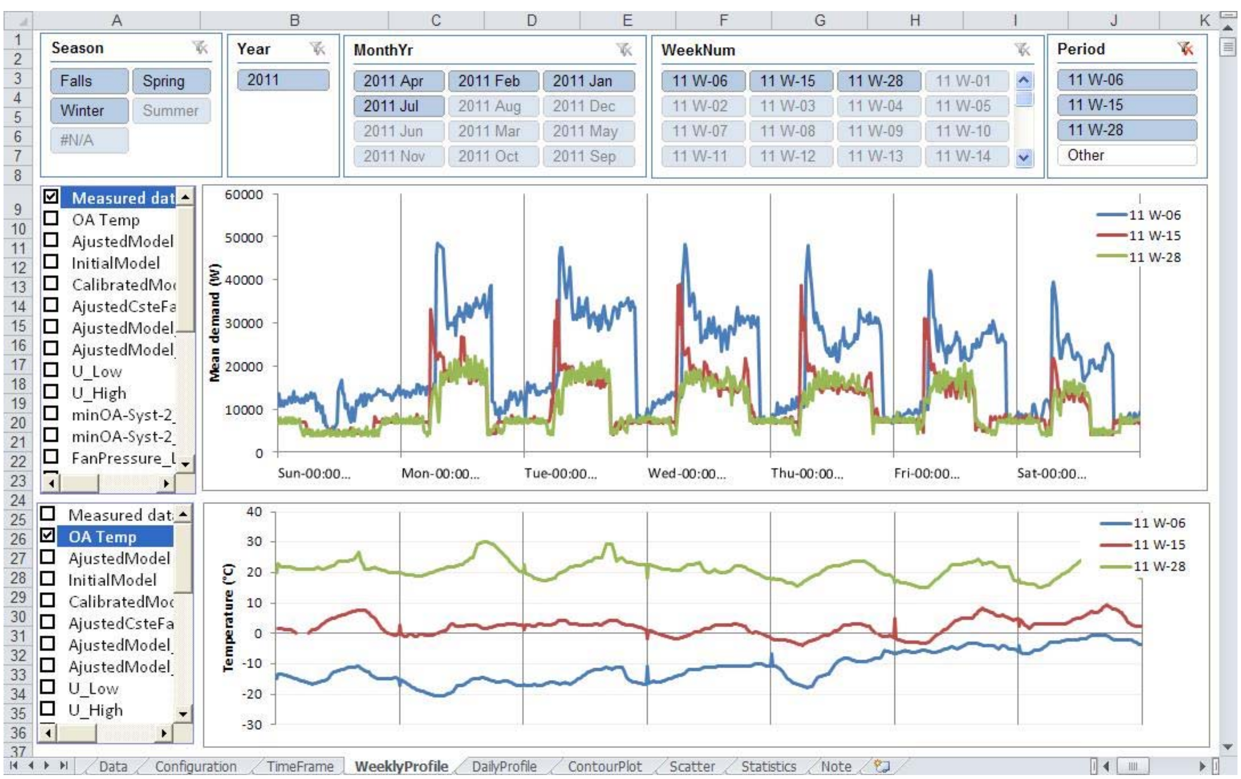

Figure 5 Screen capture of *WeeklyProfile* report worksheet

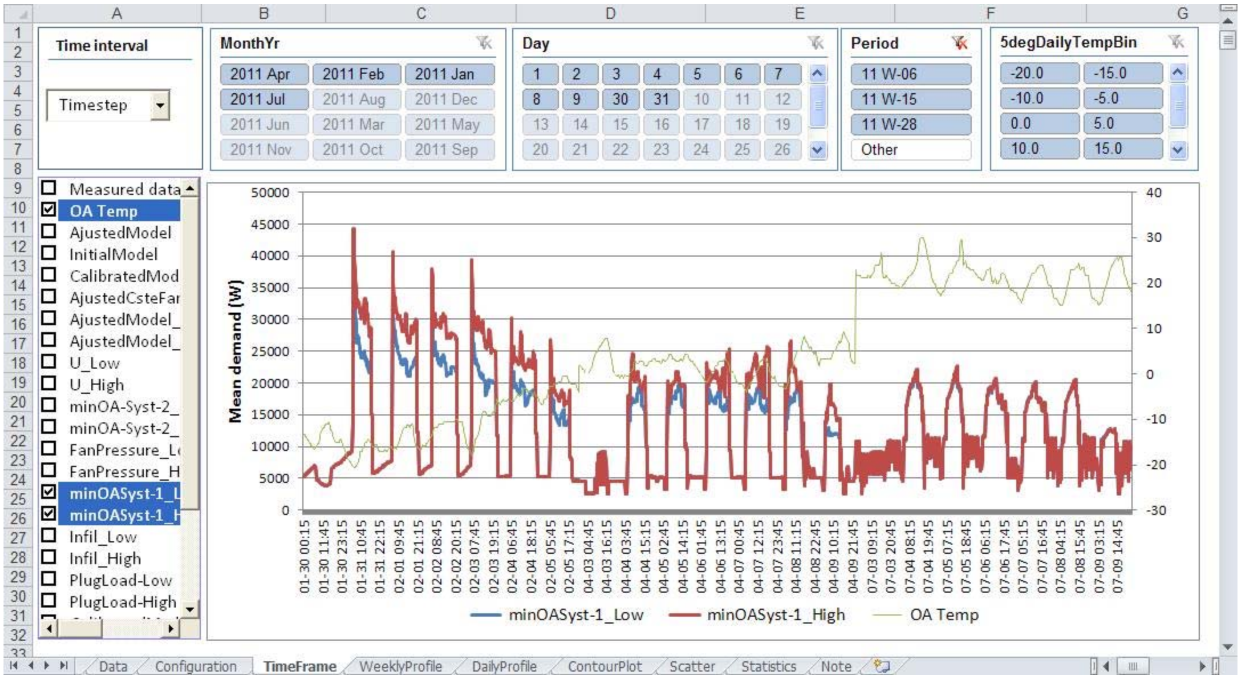

Figure 6 Screen capture of *Timeframe* report worksheet – *Timestep* Time intervals

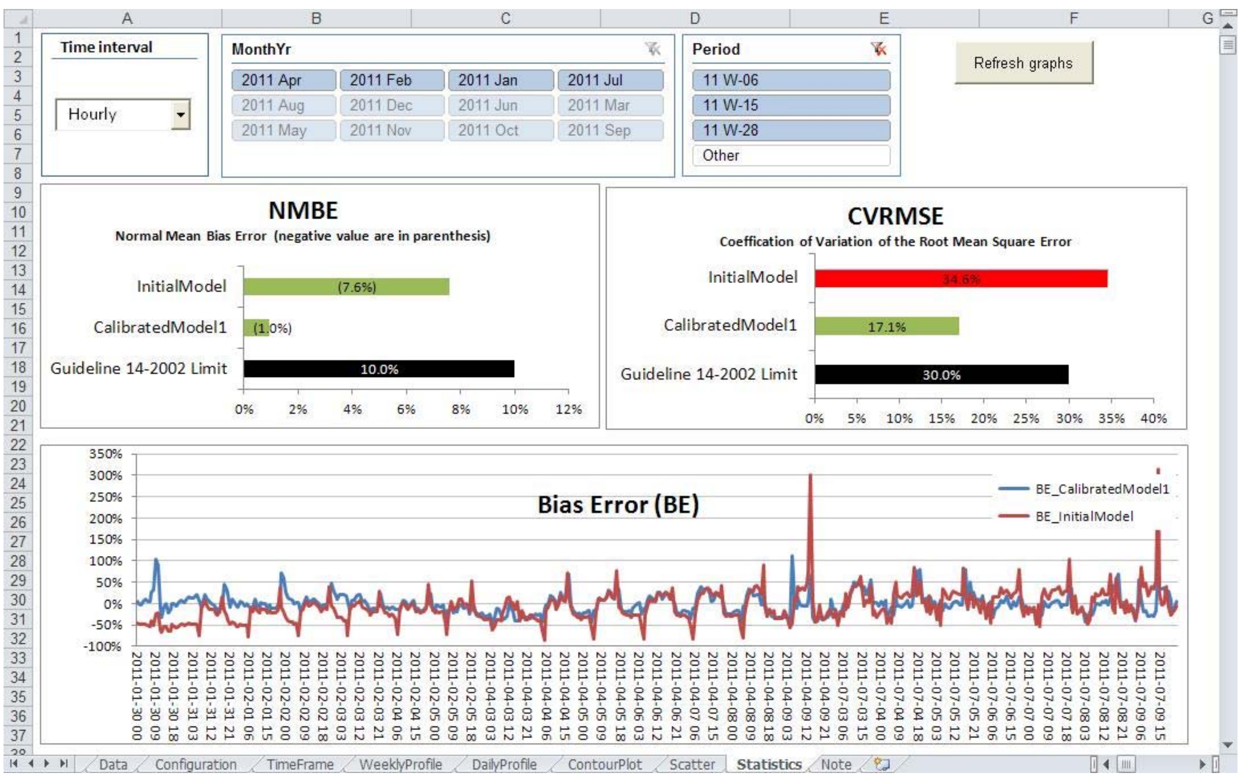

Figure 7 Screen capture of *Statistics* worksheet

FUTURE DEVELOPMENTS

Besides improving the interface's usability and robustness, interesting developments are in the radar for this visualization tool. A new report worksheet that classifies daily profiles based on clustering techniques (Fournier & Lavigne 2010) is considered. The integration of this approach would allow to easily classify the daily profiles of the relative errors between two timeseries. Also a new report with binned interquartile daily profiles using whisker plots, as initially reported by Bou-Saada & Haberl (1995) could be eventually integrated.

The import process of external data should also be improved. A command that easily imports data from EnergyPlus CSV result files into VizBEM will be integrated. This functionality will facilitate the comparison between different simulated alternatives. With this improvement, VizBEM could be used as a powerful viewer to EnergyPlus, especially for sub-hourly report frequencies.

A proof of concept will be considered as well to link external data from a database to VizBEM report worksheet capabilities. This database will gather information from building automated system (BAS) data points, sub-utility metering and simulated results. A robust structure that maps together this information should be developed in order to facilitate the comparison between different BAS data points and simulated object results. Calibration of models using multiple measured data, such as zone temperatures, zone airflows or sub-utility metering brings significant challenges.

CONCLUSION

This paper presented a tool developed for visualizing building performance and simulation data to help the calibration of models. This tool is based on Excel™ pivot table functionalities with their efficient filtering capabilities. It addresses a need for easily plot graphs adapted for timeseries data visualization and comparison purposes. Some of the limitations of using the widely known Excel™ spreadsheet software are compensated by all the flexibility it provides to adapt the viewer to specific needs.

Data visualization is an essential process in the calibration of a model. More adapted tools should be developed in order to manage and analyze measured and simulated building performance data. By simplifying and accelerating the visualization and analysis of generated data, energy modellers and researchers may concentrate on exploring new sets of techniques to create better calibrated models.

REFERENCES

ASHRAE. ASHRAE Guideline 14-2002, Measurement of Energy and Demand Savings, 2002

Bou-Saada, T. E., Haberl, J. S. 1995. An improved procedure for developing calibrated hourly simulation models, Proceedings of the 5th IBPSA Building Simulation Conference, Madison, Wisconsin, USA.

Fournier, M., Lavigne, K., 2010. Daily load profile clustering: a tool for simulation calibration, ESim 2010: The 6th IPBSA-Canada Conference, Winnipeg, May 19-20, 2010.

LBNL. 2007, VizTool, Available from <http://www.viztool.org/> [Accessed 12 July 2012]

NREL. 2010, DView, Available from < https://beopt.nrel.gov/downloadDView/> [Accessed 27 February 2014]

O'Donnell, J., Maile, T., Settlemyre, K., Haves, P. 2013, A Visualization Environment for Analysis of Measured and Simulated Building Performance Data, Proceedings of 13th Conference of International Building Performance Simulation Association, Chambéry, France, August 26-28, 2013

PG&E. 2008, Universal Translator (UT). Available from <http://www.utonline.org/> [Accessed 27 February 2014]

PNNL. 2013, ECAM : Energy Charting and Metric Tool. [Accessed 27 February 2014] Available from < http://buildingretuning.pnnl.gov/ecam.stm/>

Raftery, P., Keane, M. 2011, Visualizing Patterns in Building Performance Data, Proceedings of 12th Conference of International Building Performance Simulation Association, Sydney, November 14-16, 2011 .

Reddy, T. A., 2006. Literature Review on Calibration of Building Energy Simulation Programs: Uses, Problem, Procedures, Uncertainty and Tools, ASHRAE Transactions, 2006, Vol. 112 Issue 1, p. 22