Title

Distributed Energy Resource Optimization Using a Software as Service (SaaS) Approach at the University of California, Davis Campus

Author
Michael, Stadler

Publication Date
2011-02-07

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Distributed Energy Resource Optimization Using a Software as Service (SaaS) Approach at the University of California, Davis Campus

Stadler Michael, Chris Marnay, Jon Donadee, Judy Lai, Olivier Mégel, Prajesh Bhattacharya, Afzal Siddiqui

Environmental Energy Technologies Division

February 6 2011

http://eetd.lbl.gov/EA/EMP/emp-pubs.html

This work was funded by the Berkeley Lab FY09 Technology Commercialization Fund within the Office of Energy Efficiency and Renewable Energy (EERE) at the U.S. Department of Energy in addition to partnering and cofunding under a Cooperative Research and Development Agreement (CRADA) with OSIsoft LLC.
The Berkeley Lab is managed and operated by the University of California for the U.S. Department of Energy under contract DE-AC02-05CH11231.

ACKNOWLEDGEMENTS

This work was funded by the Berkeley Lab FY09 Technology Commercialization Fund within the Office of Energy Efficiency and Renewable Energy (EERE) at the U.S. Department of Energy in addition to partnering and cofunding under a Cooperative Research and Development Agreement (CRADA) with OSIsoft LLC.

We are very appreciative of support and assistance from J. Patrick Kennedy, Robert J Eisele, and Andy Feiyan Wu of OSIsoft, as well as Chris Cioni, Alan Ebler, and Joshua Morejohn of UC Davis.

The Berkeley Lab is managed and operated by the University of California for the U.S. Department of Energy under contract DE-AC02-05CH11231.

ACRONYMS and ABBREVIATIONS

BAS	Building Automation System
CAISO	California Independent System Operator
CIMIS	California's Irrigation Management Information Systems
DER	Distributed Energy Resources
DER-CAM	Distributed Energy Resources Customer Adoption Model
DG	Distributed Generation
GAMS®	General Algebraic Modeling System
GHG	Greenhouse Gas
kW	kilowatt
kWh	kilowatt hour
LMP	Locational Marginal Price
NOAA	National Oceanic and Atmospheric Administration
PG&E	Pacific Gas and Electric Company (utility company)
PV	Photovoltic
SaaS	Software as a Service
TCF	Technology Commercialization Fund
TOU	Time of Use (tariffs)
UCD	University of California, Davis
USDOE	United States Department of Energy
WAPA	Western Area Power Administration
WECC	Western Electricity Coordinating Council

ABSTRACT

Together with OSIsoft LLC as its private sector partner and matching sponsor, the Lawrence Berkeley National Laboratory (Berkeley Lab) won an FY09 Technology Commercialization Fund (TCF) grant from the U.S. Department of Energy. The goal of the project is to commercialize Berkeley Lab's optimizing program, the Distributed Energy Resources Customer Adoption Model (DER-CAM) using a software as a service (SaaS) model with OSIsoft as its first non-scientific user. OSIsoft could in turn provide optimization capability to its software clients. In this way, energy efficiency and/or carbon minimizing strategies could be made readily available to commercial and industrial facilities. Specialized versions of DER-CAM dedicated to solving OSIsoft's customer problems have been set up on a server at Berkeley Lab. The objective of DER-CAM is to minimize the cost of technology adoption and operation or carbon emissions, or combinations thereof. DER-CAM determines which technologies should be installed and operated based on specific site load, price information, and performance data for available equipment options. An established user of OSIsoft's PI software suite, the University of California, Davis (UCD), was selected as a demonstration site for this project. UCD's participation in the project is driven by its motivation to reduce its carbon emissions. The campus currently buys electricity economically through the Western Area Power Administration (WAPA). The campus does not therefore face compelling cost incentives to improve the efficiency of its operations, but is nonetheless motivated to lower the carbon footprint of its buildings. Berkeley Lab attempted to demonstrate a scenario wherein UCD is forced to purchase electricity on a standard time-of-use tariff from Pacific Gas and Electric (PG&E), which is a concern to Facilities staff. Additionally, DER-CAM has been set up to consider the variability of carbon emissions throughout the day and seasons. Two distinct analyses of value to UCD are possible using this approach. First, optimal investment choices for buildings under the two alternative objectives can be derived. Second, a week-ahead building operations forecaster has been written that executes DER-CAM to find an optimal operating schedule for buildings given their expected building energy services requirements, electricity prices, and local weather. As part of its matching contribution, OSIsoft provided a full implementation of PI and a server to install it on at Berkeley Lab. Using the PItoPI protocol, this gives Berkeley Lab researchers direct access to UCD's PI data base. However, this arrangement is in itself inadequate for performing optimizations. Additional data not included in UCD's PI database would be needed and the campus was not able to provide this information. This report details the process, results, and lessons learned of this commercialization project.

Keywords: building optimization, forecasting, distributed energy resources, DER-CAM

Please use the following citation for this report:

Stadler Michael, Chris Marnay, Jon Donadee, Judy Lai, Olivier Mégel, Prajesh Bhattacharya, Afzal Siddiqui, (2011): "Distributed Energy Resource Optimization Using a Software as Service (SaaS) Approach at the University of California, Davis Campus" Lawrence Berkeley National Laboratory.

TABLE OF CONTENTS

PROJECT HIGHLIGHTS

Motivation and Background

The goal of this TCF project is to set up and commercialize Berkeley Lab's optimizing program, the Distributed Energy Resources Customer Adoption Model (DER-CAM) as a web-based software service. The objective of the model is to minimize the cost of DER technology adoption and operation or carbon emissions, or combinations thereof. It determines which technology should be operated based on specific site load, price information, and performance data for available equipment options.

The site used is the University of California, Davis (UCD), specifically, the Segundo Dining Commons. Because UCD utilizes OSIsoft's PI system for gathering and storing sub-metering data, Berkeley Lab has been able to access the historical and real-time electricity, natural gas, steam, and chilled water usage information for Segundo through OSIsoft's PItoPI service.

The campus currently buys electricity through the Western Area Power Administration (WAPA) at an essentially flat, favorable tariff of $0.085/kWh. Berkeley Lab set out to investigate the hypothetical scenario wherein UCD has to purchase electricity on a time-of-use (TOU) tariff from Pacific Gas and Electric (PG&E), and therefore, has to consider the variability of carbon emissions throughout the day and seasons.

Accomplishments

The following are some of the key accomplishments of this project:

- A PI server at Berkeley Lab has been installed and PItoPI link between UCD and Berkeley Lab servers has been established.
- Historic Segundo electricity, natural gas, chilled water, and steam usage data are directly downloadable on the fly.
- Fully functional investment & planning optimization that takes into account historic metered data from the Segundo Dining Commons building.
- A download tool for automatic querying of the California Independent System Operator (CAISO, for system load and locational marginal price data) has been written.
- A time series database for storing and accessing CAISO and National Oceanic and Atmospheric Administration (NOAA) weather data has been created.
- A "forecaster" for Segundo Dining Commons building week-ahead demand depending on temperature data has been implemented.
- Week-ahead optimization has been partially implemented (user interface is missing and some load data regarding load shifting could not be modeled due to missing data points and support at Segundo Dining Commons building).

Example Results

WebOpt was used to conduct a series of distributed generation (DG) investment analyses for the Segundo Dining Commons. Experiments were also conducted with the week-ahead DER-CAM to assess the CO_2 emissions reductions and energy cost savings from rescheduling electrical loads. Figure 1 shows an example result of the investment and planning optimization running in WebOpt. Instead of a flat WAPA tariff, the PGE E-19 TOU tariff was used (see also PG&E E-19). In this cost minimization example, 44.5 kWh of lead acid batteries are adopted. As can be seen from Figure 1, DER-CAM is scheduling battery discharging around noon to reduce the on-peak related TOU costs and demand charges. The investment results from the investment and planning optimization were used as input for the week-ahead optimization. The week-ahead optimization result based on Figure 1 can be found in the main body of this report (Figure 17).

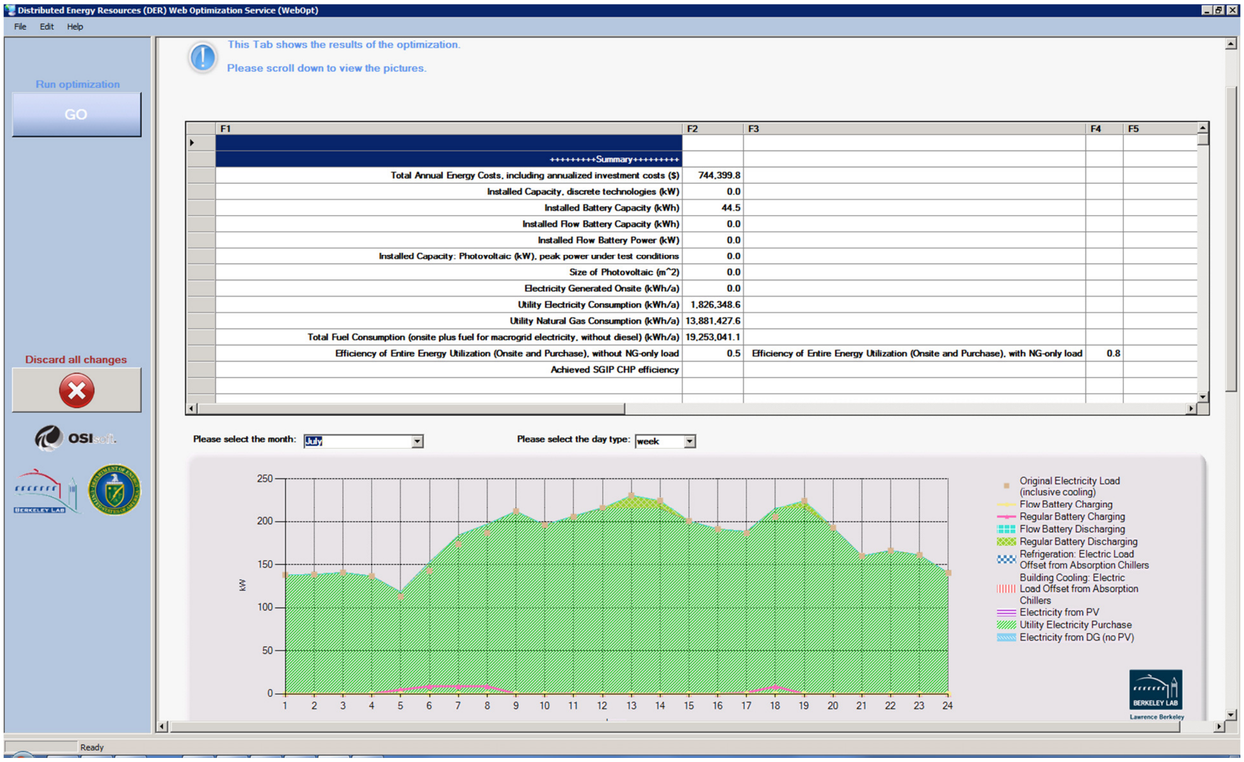


Figure 1. Investment & Planning WebOpt Result, Cost Minimization using PGE E-19 TOU Tariff

Figure 2 shows the example result for the week-ahead optimization based on a CO_2 minimization strategy.

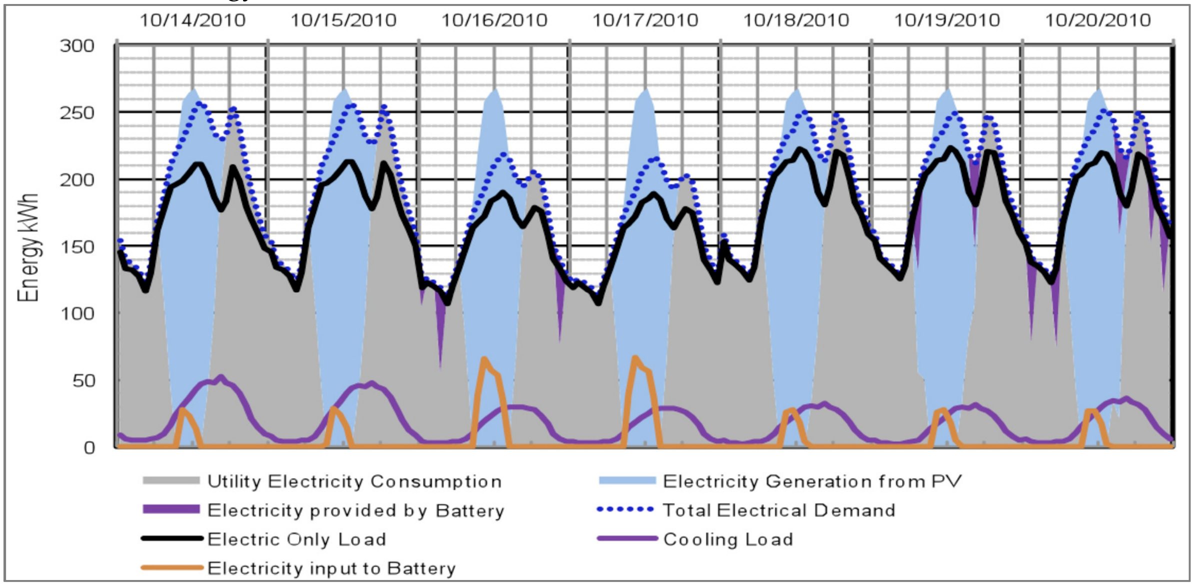


Figure 2: The Forecasted Energy Needs and Operating Schedule for the CO_2 Minimizing Combination using the Week-Ahead DER-CAM[1]

Conclusions, Lessons Learned

Berkeley Lab has demonstrated that, if accurate building use data and other required inputs were available, it is possible for a customer to perform investment & planning as well as week-ahead optimizations in an automated manner by using a Software as Service (SaaS) approach over the internet with WebOpt and without specialized hardware or software. Having established this capability, Berkeley Lab has made great strides towards offering near real-time forecasts that will impact the site's operations.

Having access to historical and real-time building electricity, natural gas, and steam data via OSIsoft's PItoPI server was an important aspect of this project. However, several problems arose with regards to the carbon emissions data and data availability. For example, Berkeley Lab had limited possibilities to model demand response since no detailed enduse data were available at UCD, e.g. data for refrigeration. Also, UCD is very interested in minimizing carbon emissions, but the lack of accurate marginal CO_2 emission data from the macrogrid made it difficult to estimate possible CO_2 reduction potential.

Also, the issue of data feedback to the existing Energy Management System (EMS) is not solved at this point since it is not supported by OSIsoft. The final fully automated SaaS should feature the possibility to feedback near real-time optimization results to reduce the human interaction and should also be able to consider uncertainty in, for example, weather data (Siddiqui et al. 2010).

---

[1] Electricity Only Loads are electric loads for services that only can use electricity. Cooling is not an electricity only load since waste heat / absorption cooling could be also used for cooling. Typical electricity only loads are lighting or computing.

Finally, Berkeley Lab received 40% less TCF funding than originally requested, and setting up a fully functional version of DER-CAM proved more difficult and significantly more time consuming than anticipated. Consequently, it is now clear that a larger, more substantial effort is needed to set up a fully functioning SaaS implementation.

Technical Details

Project Overview

Together with OSIsoft LLC as its private sector partner and matching sponsor, the Lawrence Berkeley National Laboratory (Berkeley Lab) won an FY09 Technology Commercialization Fund (TCF) grant from the U.S. Department of Energy. The goal of this TCF project is to commercialize Berkeley Lab's optimizing program, the Distributed Energy Resources Customer Adoption Model (DER-CAM) using a software as a service (SaaS) model, with OSIsoft as its first non-scientific client. OSIsoft could in turn offer optimization capabilities to users of its PI software suite (see OSIsoft). DER-CAM is an economic-engineering model of customer DER adoption and operation. DER-CAM has been implemented on the General Algebraic Modeling System (GAMS®) optimization software platform, and has been in development at Berkeley Lab since 2000. The objective of the model is to minimize the cost of technology adoption and operation or its carbon footprint, or combinations thereof. DER-CAM determines which technologies should be installed and operated based on specific site energy service requirements typically in the form of hourly load shapes, price information, and performance data for available equipment options.

A central objective for the TCF is to choose a pilot PI user and to demonstrate that optimization results of value to the site can be executed via an interconnection between the site's PI installation and one installed at Berkeley Lab. Because UCD utilizes OSIsoft's PI system for gathering and storing sub-metering data, Berkeley Lab has been able to access the historical and real-time electricity, natural gas, steam, and chilled water usage information for multiple UCD buildings through OSIsoft's PItoPI protocol. A test building on the UCD campus was needed for the pilot, and after due consideration, the Segundo Dining Commons was chosen.

UCD's participation in this project is driven by its motivation to the reduce carbon footprint of its operations. The main campus currently buys its electricity through the Western Area Power Administration (WAPA) at essentially a flat tariff of $0.085/kWh, and some of its auxiliary sites buy from PG&E and Sacramento Municipal Utility District (SMUD). Throughout the mid 2000's, the main campus purchased a mix of PG&E and WAPA power, so the possibility of having to revert to a PG&E tariff exists. Berkeley Lab therefore investigated the hypothetical scenario wherein UCD has to purchase electricity on a standard commercial time-of-use (TOU) tariff from Pacific Gas and Electric (PG&E). There are two analyses of interest to UCD: optimal equipment choice and operation under a cost minimizing objective using the applicable PG&E tariff, and carbon minimizing operation.

For each scenario, there are two distinct types of optimization:

a) **an investment and planning optimization.** The first is an investment choice based on a test year of historic operations. This analysis suggests what equipment and operating schedule achieves either of the two objectives, or a combination.
b) **a week-ahead optimization.** Considering temperature forecasts and changing load profiles, the second optimization assumes that the installed equipment is fixed, and given expected requirements and conditions, delivers a week-ahead forecast of building operations that satisfies either of the two objectives, or a combination. To minimize carbon footprint, the optimization has to consider the variability of carbon emissions throughout the coming week.

There are several key benefits to having DER-CAM's optimization capabilities accessible via a SaaS implementation:

- incorporating it into a graphical interface makes DER-CAM more usable
- running optimizations on Berkeley Lab's secure server and returning results to the client means that there is no need for specialized hardware or software on the user end
- Berkeley Lab bears the burden of expensive licensing and maintenance costs for GAMS® and related mathematical solvers
- neither the software clients nor their customers are required to enter into licensing agreements with U.S. DOE, the California Energy Commission, and the University of California Berkeley Lab for accessing DER-CAM
- Easy central maintenance of DER-CAM
- Simple user management and
- User tailored DER-CAM version management.

The following graphic shows the completed work (enclosed in dashed lines) and broader ultimate goals of this commercialization effort. As represented, Berkeley Lab has accomplished the first steps and realized the goal of making DER-CAM optimizations accessible as a SaaS.

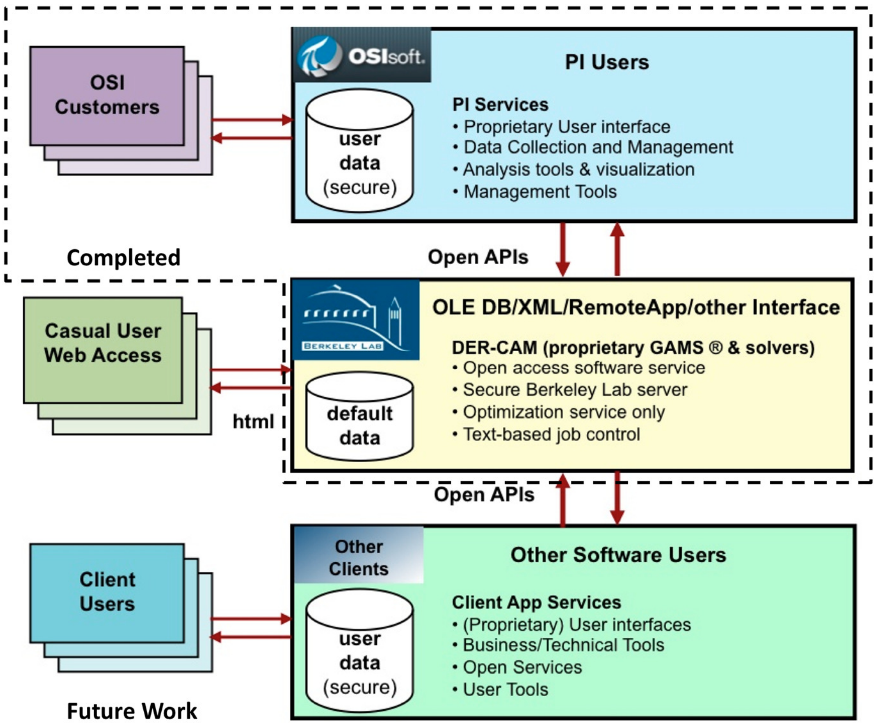
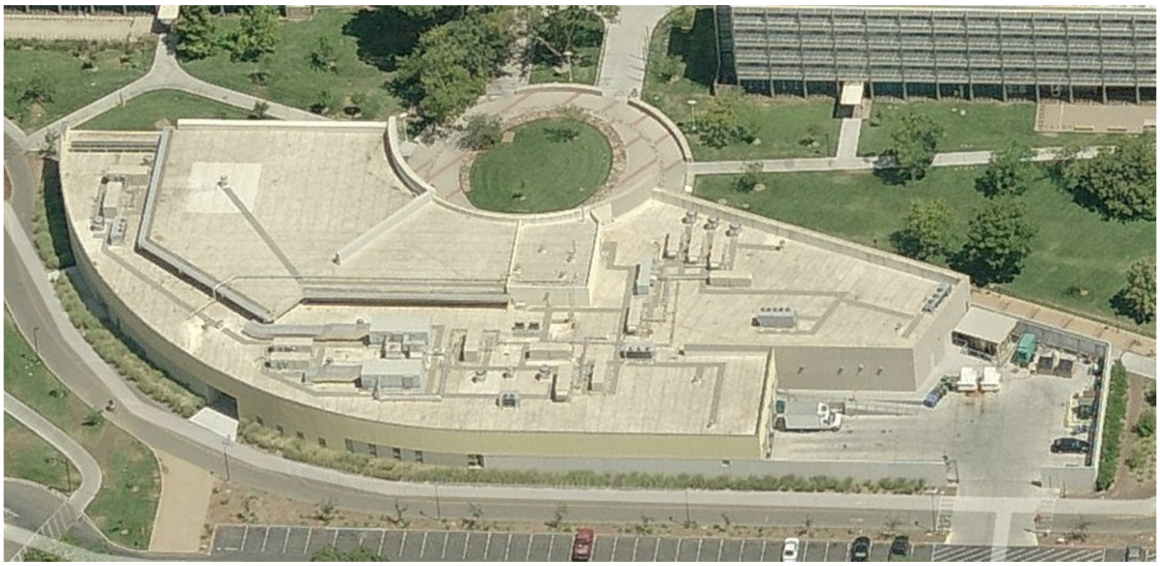


Figure 3. Final Software as Service Structure

The final product will be a web based optimization service (WebOpt) that embeds DER-CAM results into existing energy management software tools, beginning with OSIsoft's PI data management suite. OSIsoft clients, and in the future others, can be provided with sophisticated financial/environmental/engineering analysis results, cost-environmental trade-offs, and deployment paths without using DER-CAM directly. Marrying Berkeley Lab's technical building energy analyses and simulation capabilities with existing tools will make complex optimization accessible quickly to a large installed customer base. This combined skill set can quickly deliver effective products to a large primed market, and yield dramatic energy efficiency deployment and renewable adoption gains.

The Demonstration Site - Segundo Dining Commons

Segundo Dining Commons, see Figure 4, is a ~50,000 ft^2 (~4650 m^2) dining hall that serves three meals a day to students on weekdays and two meals on weekends, and its kitchen is occasionally used for catering other campus events and meetings.



Birdseye View (from the south)

Exterior

Courtyard

Rooftop

Interior

Interior

Interior

source: Birdseye view from http://www.bing.com/maps/, all others by Berkeley Lab

Figure 4. Photos of Segundo Dining Commons

Approximately two years of historical sub-metered data (electricity, natural gas, steam, and chilled water) for Segundo have been made available to Berkeley Lab for this project via OSIsoft's PItoPI interface.

WebOpt

The structure of the Saas setup is shown in Figure 5. The blue area to the right shows the tools installed on Berkeley Lab's server. The yellow area in the upper left shows the data archived in UCD's PI server. The grayed out data sets are needed to complete the WebOpt capability. The Berkeley Lab PI server directly accesses UCD data using the proprietary PItoPI protocol. The data flow at the bottom shows the data collection being conducted with other data sources. A Java tool downloads CAISO and weather data and together with estimated marginal CO_2 emissions a demand forecast for the week-ahead optimization is performed. The role of WebOpt then is to establish a GAMS job to be executed by the appropriate DER-CAM version. There is currently no automated way of returning results to UCD, so they would need to be sent by alternative means such as email or can be shown directly in WebOpt.

The final user interface, which handles the data management and DER-CAM and runs on a Berkeley server behind firewalls, is called WebOpt. The user executes WebOpt through a secure Remote Desktop Connection (Terminal Services Client 6.0) and does not need to have any specialized software installed or run any other program.

WebOpt collects data from the Berkely Lab PI server using DataLink, a standard OSIsoft product, and also calls the format changer macro, which converts the raw PI data to a format DER-CAM can use. WebOpt will handle two distinct versions of DER-CAM:

a) investment & planning optimization and
b) week-ahead optimization

WebOpt currently performs the investment & planning optimization which utilizes historic Segundo load data. However, since OSIsoft's PI system does not currently support data feed-back, the optimization results cannot be sent back to the building directly, and therefore, must currently be shown in the WebOpt interface.

The week-ahead optimization capabilities have been developed, but are not currently implemented in WebOpt because strategies such as load shift measures cannot be considered due to missing data streams for the Segundo buildings. The week-ahead optimization runs in principle without any user interface on Excel and individual Java applications.

The important forecaster (at the bottom of Figure 5) forecasts the Segundo load profiles depending on weather data. These forecasted loads will be sent to the week-ahead optimization and DER-CAM is executed.

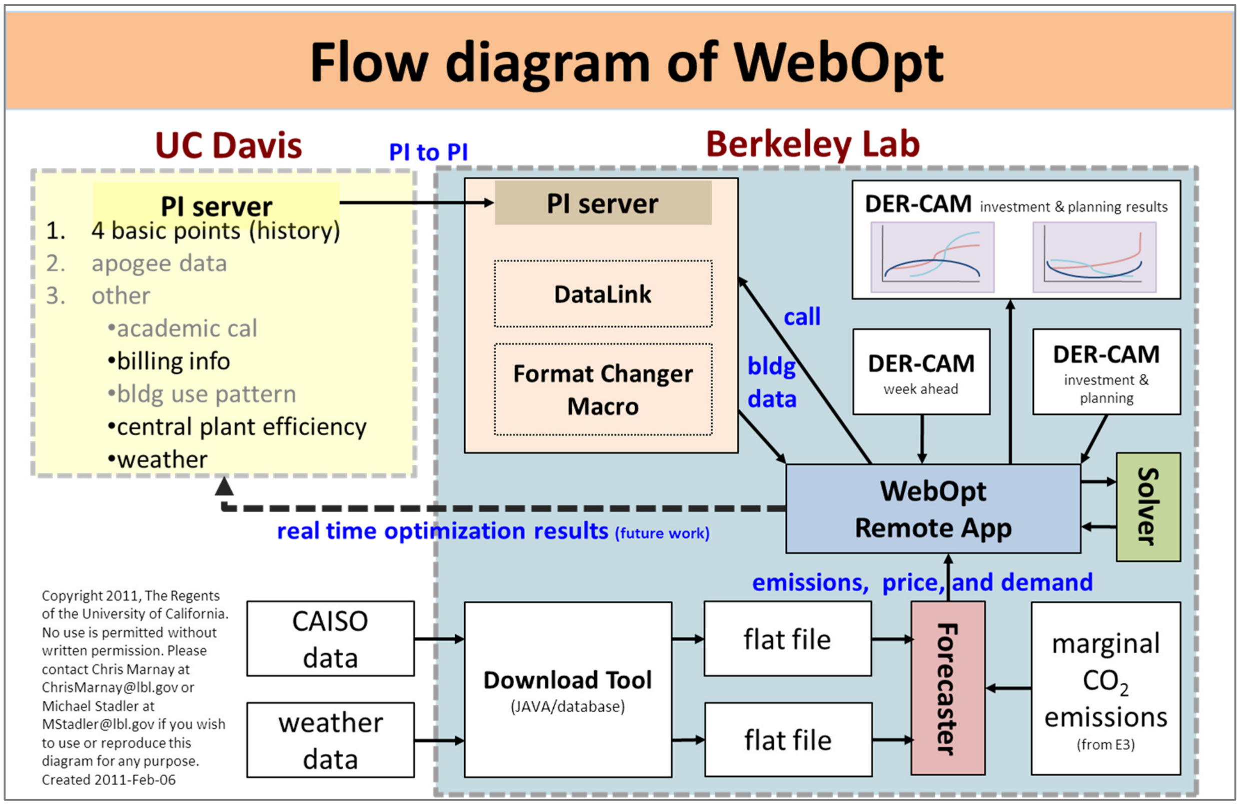


Figure 5. Flow Diagram of WebOpt

Distributed Energy Resources Customer Adoption Model (DER-CAM)

DER-CAM finds optimal combinations of building equipment and an operating schedule, given the facility energy service requirement, the prevailing economic conditions, and a menu of supply and passive technologies to choose from. It has been developed at Berkeley Lab over several years, has been rigorously reported in more than 30 peer-reviewed papers as well as in many reports and conference proceedings, and has been used for many example building studies and other analysis by a worldwide user group of researchers in Australia, Austria, Belgium, Canada, China, Denmark, Germany, Ireland, Italy, Japan, New Zealand, Poland, Portugal, Spain, Taiwan, and USA. To date, it has been primarily applied to commercial buildings with peak electrical loads of 100 kW – 5 MW. For more information on DER-CAM, please refer to APPENDIX B: Background on DER-CAM (Marnay et al. 2008, Stadler et al. 2008, Stadler et al. 2009).

DER-CAM Optimization Engines

This project utilizes two different versions of DER-CAM:

a) investment & planning optimization and
b) week-ahead optimization.

The investment & planning optimization delivers an optimal investment portfolio for DER technologies and planned operational schedule for a test year. The test year is based on historic UCD Segundo load profile data, collected through PItoPI and Datalink. This investment &

planning optimization assumes that all loads, tariffs, macrogrid emissions data, etc. are fixed and known. This version determines the equipment that would minimize cost or CO_2 emissions.

The week-ahead version of DER-CAM is used to create a rolling optimal week ahead operation schedule given a fixed set of DER technologies, weather, and load. The structure for real-time pricing is implemented in the week-ahead DER-CAM version, but was not utilized in this demonstration.

This week-ahead optimization represents a big step towards a full real-time optimization based on user-defined timestamps. It would be also capable of real-time demand response and load shifting when supplied with the needed data points e.g. lighting and refrigeration loads. The motivation for making day-ahead scheduling decisions on a rolling week-ahead forecast is twofold. First, operations tomorrow naturally depend heavily on conditions tomorrow, but also on subsequent days and that should be recognized and considered. Second, for reliability purposes, it is useful to have future days' schedules established should communications fail on any given day.

PI Server and PI System

OSIsoft's PI Server can accept data from disparate sources, e.g., enterprise systems, databases, operational data sources, etc, and is capable of real-time data gathering, archiving, and distribution, all key capabilities for this optimization project. The PI Server strives to optimize the storage and retrieval of vast amounts of data so as to provide users with a comprehensive, real-time view into operational, IT infrastructure, and business activities. UC Davis utilizes a PI System to collect data from its utility meters (electricity, gas, water, steam and chilled water) in conjunction with Siemens' Apogee Building Automation System (BAS).

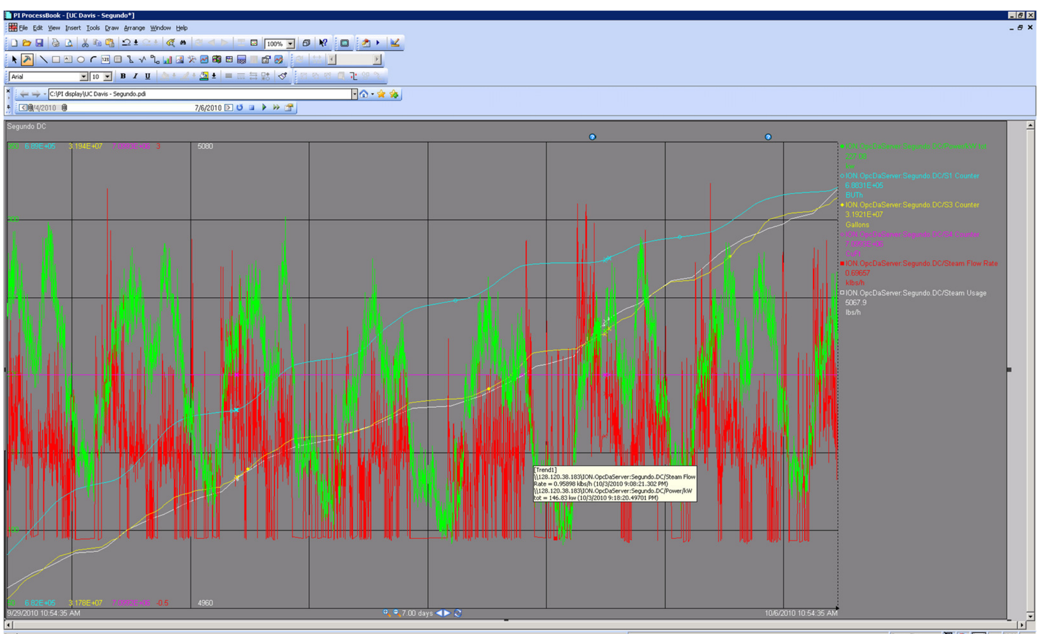


Figure 6. View of PI Processbook (Realtime Data View). Legends from the top are 1) total kW (green), 2) cumulative counter for BTUh (blue), 3) cumulative counter for gallons of water (yellow), 4) klbs/h of steam flow rate (red), 5) lbs/h of steam usage (white)

Figure 6 shows a screencapture of PI Processbook, which allows viewing the metered data. The utility data is collected for every building and the BAS data is collected for a few specific buildings. A PI System was installed at Berkeley Lab to gather some of the data archived on the UC Davis PI System, using a PItoPI.

Data Collection

The following subsections discuss in detail the various boxes shown in Figure 5.

Data from UC Davis

Historical building data for electricity, natural gas, chilled water, and steam for the Segundo Dining Commons are based on actual PI data collected from April 1st 2009 to present. Berkeley Lab has been able to access the above four points via PItoPI, as shown in the WebOpt flow diagram, but has not been granted access to Siemen's APOGEE building monitoring data, information regarding building use (for example, when Segundo dining commins is responsible for extra cooking/catering for campus events) or the academic calendar (for example, shutdowns over holidays and term breaks). This APOGEE data would be needed for modeling demand response and load shifting in the week-ahead optimization version of DER-CAM. Historical weather as measured at UC Davis was only made available to Berkeley Lab on the afternoon of the last day of the project, by which time alternative sources of weather data, discussed in later sections, had been found and used for the demand forecaster.

California Independent System Operator (CAISO) Data

The following California Independent System Operator (CAISO) data dating back to April 1st 2009 (the start of CAISO's Market Redesign and Technology Upgrade, MRTU) have been downloaded and are stored on Berkeley Lab's server:

- Hourly actual CAISO load
- 2-day ahead CAISO load forecast
- Day-ahead CAISO load forecast
- Hourly Davis area Locational Marginal Price (LMP) node information

The following forecasts are downloaded periodically to Berkeley Lab's server:

- Hourly actual CAISO load (download daily)
- Hourly Davis area Locational Marginal Price (LMP) node information (download daily)
- 15-minute Davis area (LMP) node information (download daily)

The LMP could be used in the week-ahead optimization considering real-time prices.

Weather Data

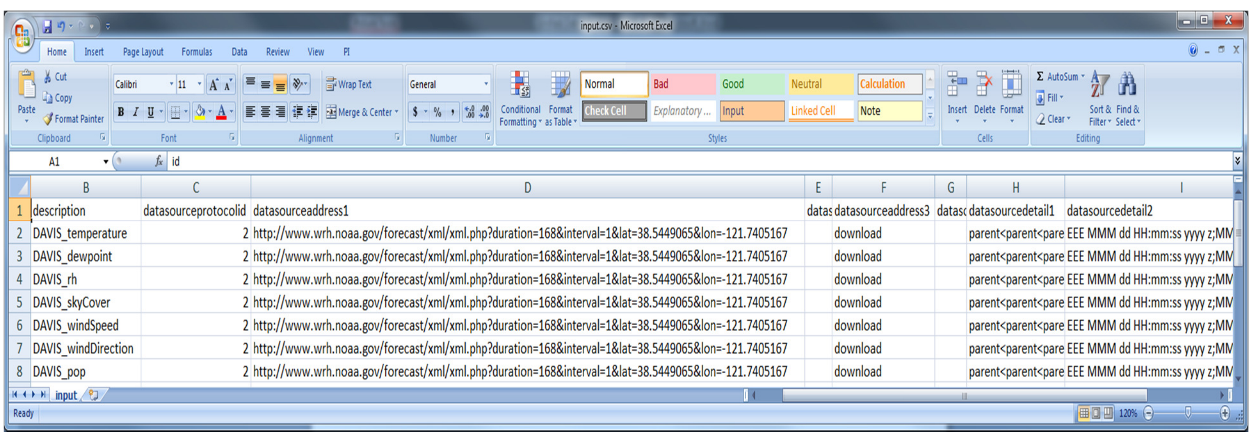
Historical weather data for the Davis area are downloaded from the California's Irrigation Management Information Systems (CIMIS) website. The temperature forecasts are taken from National Oceanic and Atmospheric Administration (NOAA) website (see NOAA). Historical daily high/low temperatures are stored in Berkeley Lab's time-series database and used as input into the demand forecaster regression model. Hourly high/low 7-day ahead forecasts are also downloaded from the above link and used as input into the forecaster.

Automated Data Download Tool

In order to automate the downloading of system load, pricing, and weather data from CAISO and NOAA servers, specialized Java applications have been written and are being run continuously. The data is saved in a database and passed as flat files when needed by the forecaster.

System Setup

1. Since all the applications are Java-based, the first step was to install Java Runtime Environment (JRE) on the host computer.
2. The next was to install the database program – postgreSQL, for the current version of the system.
3. A setup tool has been created which includes a few text files with the database configuration information and a Java package which performs the actual task of setting up the initial database structure.
4. The steps described above constitute the generic part of the setup process. After that the site-specific setup has to be performed. Essentially, it involves setting up every time-series point in the system with the individual details. A spreadsheet was created using specific headers. Each row of that spreadsheet contains information for a single point. The spreadsheet is then converted to a csv file which is used as the input to the system administration tool. Below is a screenshot of how the spreadsheet looks when it is populated with all the information. The column headers of the spreadsheet are essentially mapped to the columns in the Point table described in APPENDIX D: Unified Data Manager (UDM).



	B	C	D	E	F	G	H	I
1	description	datasourceprotocolid	datasourceaddress1	datas	datasourceaddress3	datasc	datasourcedetail1	datasourcedetail2
2	DAVIS_temperature	2	http://www.wrh.noaa.gov/forecast/xml/xml.php?duration=168&interval=1&lat=38.5449065&lon=-121.7405167		download		parent<parent<pare	EEE MMM dd HH:mm:ss yyyy z;MM
3	DAVIS_dewpoint	2	http://www.wrh.noaa.gov/forecast/xml/xml.php?duration=168&interval=1&lat=38.5449065&lon=-121.7405167		download		parent<parent<pare	EEE MMM dd HH:mm:ss yyyy z;MM
4	DAVIS_rh	2	http://www.wrh.noaa.gov/forecast/xml/xml.php?duration=168&interval=1&lat=38.5449065&lon=-121.7405167		download		parent<parent<pare	EEE MMM dd HH:mm:ss yyyy z;MM
5	DAVIS_skyCover	2	http://www.wrh.noaa.gov/forecast/xml/xml.php?duration=168&interval=1&lat=38.5449065&lon=-121.7405167		download		parent<parent<pare	EEE MMM dd HH:mm:ss yyyy z;MM
6	DAVIS_windSpeed	2	http://www.wrh.noaa.gov/forecast/xml/xml.php?duration=168&interval=1&lat=38.5449065&lon=-121.7405167		download		parent<parent<pare	EEE MMM dd HH:mm:ss yyyy z;MM
7	DAVIS_windDirection	2	http://www.wrh.noaa.gov/forecast/xml/xml.php?duration=168&interval=1&lat=38.5449065&lon=-121.7405167		download		parent<parent<pare	EEE MMM dd HH:mm:ss yyyy z;MM
8	DAVIS_pop	2	http://www.wrh.noaa.gov/forecast/xml/xml.php?duration=168&interval=1&lat=38.5449065&lon=-121.7405167		download		parent<parent<pare	EEE MMM dd HH:mm:ss yyyy z;MM

Figure 7. Specific Header Spreadsheet

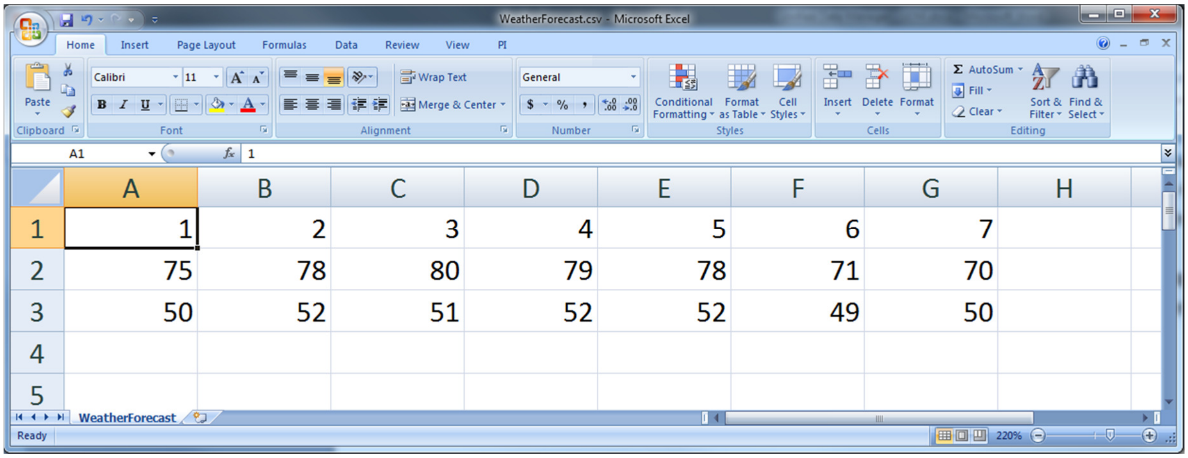
5. After the point creation, the data collector application (app) needs to be started. Ideally the data collector apps should run as a service. However, at this time it is manually started as a batch job.

Operation

Every data collector app has a unique identifier associated with it. At the time of start-up, a data collector app queries the database for points that match the app's unique identifier. The details of the points are loaded into the memory. Then it runs in an infinite loop. It "sleeps" until the next poll time arrives and then performs the work of data fetching, parsing and storing and then again goes back to sleep until the next poll time.

Current Features

1. File download capability. For example, if the data is published in a csv file or an XML on a website or on any remote host, the file can be fetched to the local host.
2. Unzip capability. For example, California ISO publishes its demand and price data in a zipped file format. One then has to unzip the file to get the actual csv or XML file. Unzipping is required for locally residing files as well.
3. Generic XML file parsing. For example, 7 day hourly weather forecast data from the NOAA website is collected, and parsing the file for values and timestamps for different quantities is done.
4. Structured csv or tab separated file parsing.
5. Capability to store forecast data. Forecast data is different from actual historic data in terms of timestamp. In historic data, there is only one timestamp associated with a value. In forecast data, there is a timestamp for when the forecast was made and a second timestamp for the time for which the forecast was made. For example, weather forecast for Sunday, is made on Thursday, Friday and Saturday.
6. Timestamp is stored in UTC millisecond. This ensures transparent operation across time zones.
7. Automatic application of offset on the poll time. For example, if the poll interval for a point is 1 hour, the polling takes place slightly after the top of the hour. This helps in smoothing the load on the local host. But more importantly, it helps in avoiding time-outs and lack-of-response from a remote server where several people send request at the top of the hour (such as CA ISO oasis server).
8. Backfill mode for the text file data collector app.
9. Multi-threaded operation of the data collector apps enables better performance and scalability. A thread can have one or more points for which it performs data collection. In the current version, a separate thread is created for every different combination of address of the data-source (IP address or web URL) and the poll interval.
10. Because of the non-flat nature of the database structure, writing plain SQL queries to extract data can be cumbersome. Hence, a query has been built to extract the 7 day forecast of daily high and low temperatures. Following is a screenshot of the csv file that is generated out of this query, to be consumed by the WebOpt load forecaster for the UC Davis project.



	A	B	C	D	E	F	G	H
1	1	2	3	4	5	6	7	
2	75	78	80	79	78	71	70	
3	50	52	51	52	52	49	50	
4								
5								

Figure 8. Daily High and Low Temperatures for WebOpt

More such generic queries can be built pretty easily.

11. Backfill does not require any manual intervention, such as making manual changes in datatable start time (archiving start time).

For more information on that download tool see APPENDIX D: Unified Data Manager (UDM).

Marginal CO_2 Emissions Data

The hourly marginal emissions data are based on the Greenhouse Gas (GHG) Calculator developed by E3 (see E3). E3 derived the emissions by simulating the generation capacity and mix in the Western Electricity Coordinating Council (WECC), and the emissions represent the carbon contribution from energy consumed in California regardless of whether the energy was generated instate or out of state. Emissions were calculated for 2008 and 2020 and intermediate years were linearly extrapolated.

For this project, the emission rates in the CAISO region have been assumed to be interchangeable with emission rates reported for the state. Berkeley Lab analyzed the CO_2 data and found extremely high emissions volatility. Figure 9 below shows a winter/January comparison of CAISO load (MW, right y-axis) with average and marginal CO_2 emissions (g/kWh, left y-axis). The month is represented by five weeks, with values for each week in color and the average of all weeks' values shown in black. Weeks start from Monday; refer to calendar in lower right hand corner for week counts and special holiday conditions. As can be seen from the graphic, the hourly CAISO loads display a predictable pattern throughout the days and week and the hourly average carbon emissions rates remain fairly steady, around 400 g/kWh. The behavior of the marginal emissions, however, varies wildly from hour to hour, day to day, and shows no discernable pattern or relationship to demand/time of day.

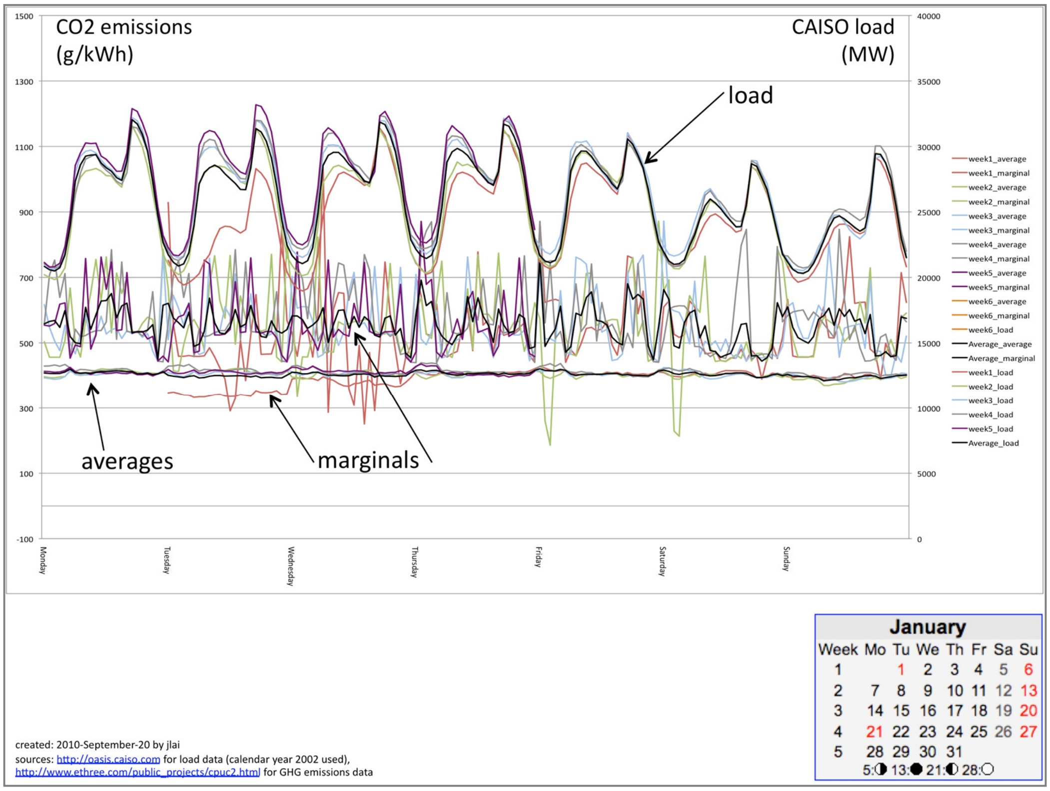


Figure 9. January CO_2 Emissions and Load Comparison

Figure 10 below shows a summer/July comparison. The CAISO loads and average emissions are again predictable, and the behavior of the marginals completely unpredictable. Taking week 4 as an example, there seems to no relationship between daily peak demand and marginal emissions; some days the peak demand aligns with peak emissions (Tue), other days peak demand yields minimum emissions (Wed, Thu, Fri), and for the remainder of the week the marginal emissions are unremarkable (Sun and Mon). Additionally, when comparing the same day of the week, for example Tuesday, the peak demand seemingly correlates to an extremely high marginal emission rate (week 4), an extremely low marginal emissions rate (week 2 or 3), or falls somewhere in the middle (week 1 and 5).

Due to such unexplained high volatility in the hourly marginal emission values, Berkeley Lab has elected to use *averaged* hourly marginal emissions based on the GHG Calculator developed by E3.

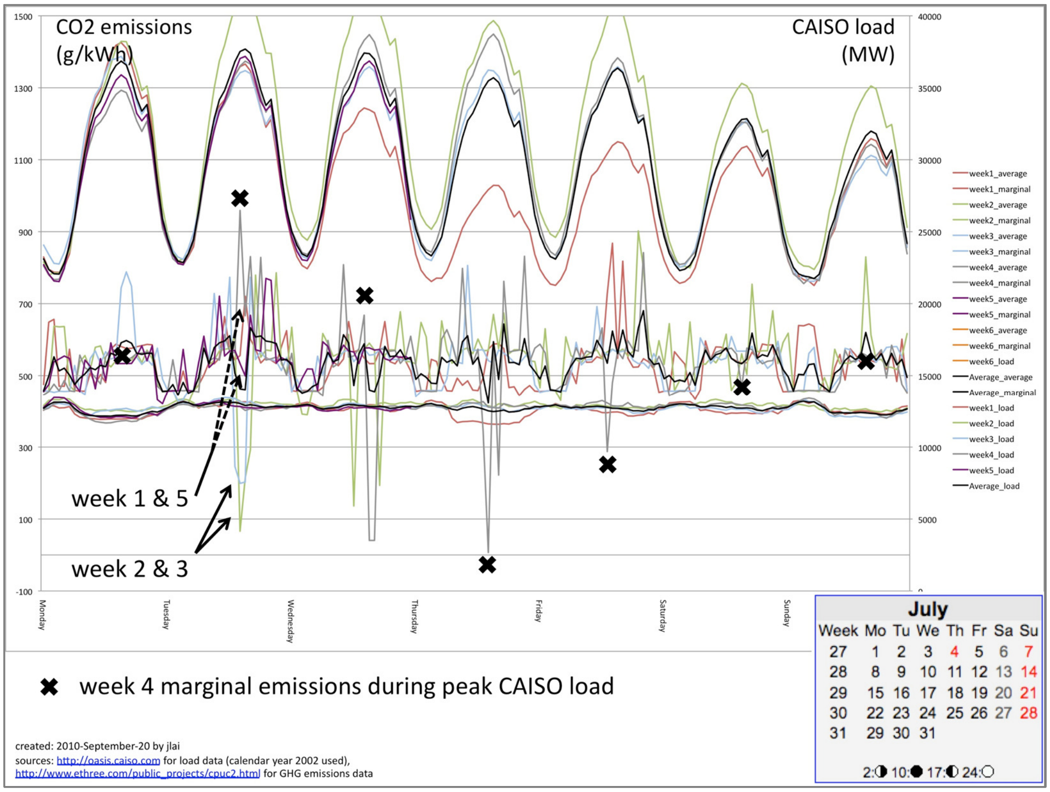


Figure 10. July CO_2 Emissions and Load Comparison

More information regarding emissions and comparison graphics for other months of the year can be found in APPENDIX A: CAISO System Load and California CO_2 Emission Rates Comparison.

Input Data Forecast

In order for DER-CAM to create an optimal DER operating schedule for the next seven days, Berkeley Lab sought to predict the heating, cooling, electricity, and natural gas demands that must be met. Given the predicted demands, the week-ahead DER-CAM will create an operation schedule that minimizes cost, emissions, or a weighted combination of the two.

Please note that this prediction is not needed for the investment & planning DER-CAM since it uses historic load date stored on the PI system.

Energy Demand Forecast for Week-Ahead Optimization

Hourly energy demands are forecasted using a multiple linear regression on the following factors:

- hour of the day
- daily high temperature
- daily low temperature

- school day (binary yes or no)

Interactions of these factors are also considered for the regression, excluding daily low interacting with daily high. To perform this regression, historical energy use data from UC Davis's PI server and historical temperature from the California's Irrigation Management Information Systems (CIMIS) website, are used. The temperature forecasts are taken from the National Oceanic and Atmospheric Administration (NOAA) by using the automated data download tool described above. Some attempts have been made to remove corrupted data before performing the regression, but this was not a comprehensive filtering. For example, some energy use data points were zero for long durations. UC Davis staff confirmed that there were data outages and that this data should be discarded. Regression coefficients were calculated using the open source statistical software *R* (see R-software). These coefficients, forecasted weather data, and estimated[2] calendar of school days are then used to compute predicted energy needs for Segundo Dining Commons. Figure 11 and Figure 12 below are example plots showing the actual energy needs and forecasted energy needs for an example week in October 2009. Note that the night of October 5th and morning of October 6th were missing energy use data and are replaced by a straight line with no markers. Better prediction results would likely be possible if more information on building occupancy or cooking schedules were available. There could also be non-linear relationships that our simple regression model does not account for.

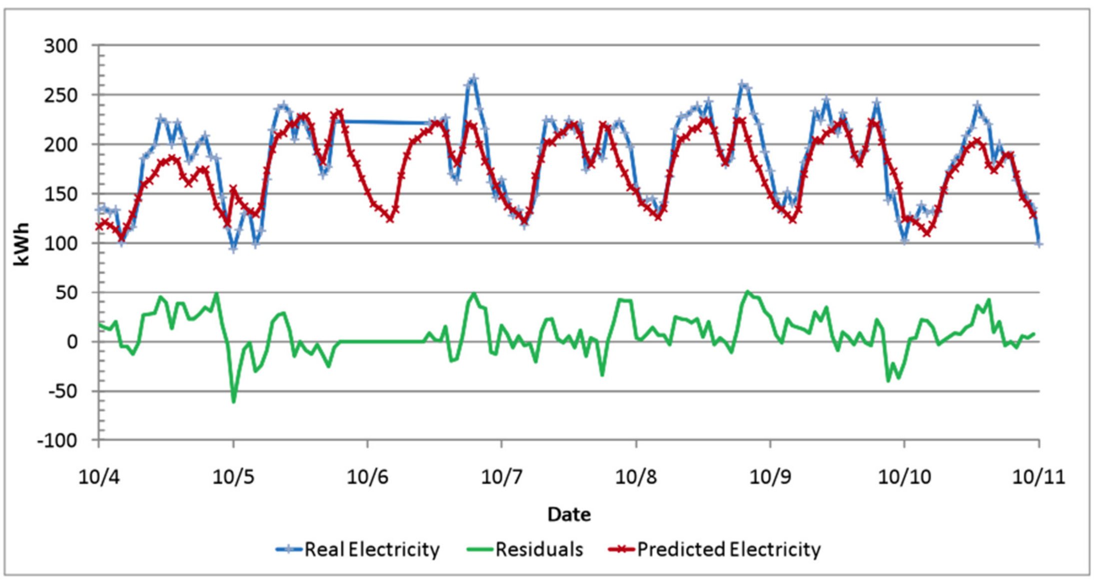


Figure 11. Actual and Forecasted Electricity Demand and Errors for One Week in October 2009

[2] Please note that the Berkeley Lab team did not receive the UCD calendar, and therefore, Berkeley Lab had to estimate the calendar.

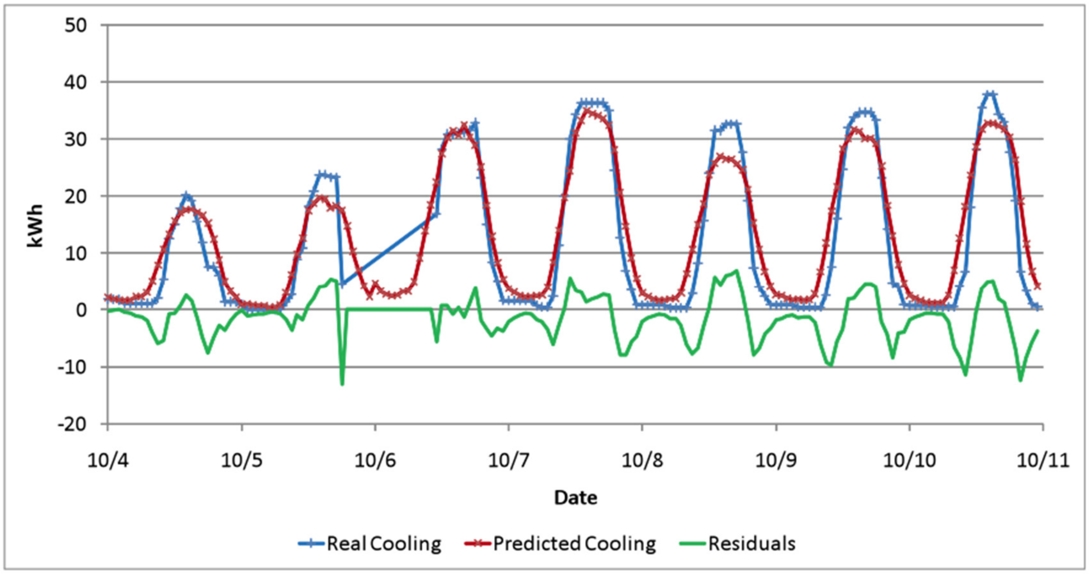


Figure 12. Actual and Forecasted Cooling Demand and Errors for One Week in October 2009

CO_2 Emissions Forecast

In order to help UCD to reduce its carbon footprint, Berkeley Lab sought to forecast CO_2 emissions from utility provided electricity on an hourly basis. Although UCD purchases its energy through WAPA, Berkeley Lab substituted readily available information on carbon emissions of energy on the CAISO market as a proxy. In other words, electricity in WECC is assumed fully fungible such that a kWh saved by UCD is equivalent to a marginal kWh not traded on CAISO. Unexpectedly, review of this highly volatile data did not yield a discernable relationship between CAISO market demand and carbon emissions. Instead of a regression analysis, a table of averaged marginal carbon emissions for each hour in each month was created. Simulated marginal emissions data was available for the years 2008 and 2020. Data is interpolated to the current year before use in Operations DER-CAM. This project highlights the need for more access to real time carbon intensity information for grid supplied electricity. If real carbon intensity information were available, energy use patterns could be changed to minimize carbon emissions.

More information regarding marginal emissions can be found in APPENDIX A: CAISO System Load and California CO_2 Emission Rates Comparison.

Results

WebOpt was used to conduct a series of trial distributed generation (DG) investment analyses for the Segundo Dining Commons. Experiments were also conducted with the week-ahead DER-CAM to assess the CO_2 emissions reductions and energy cost savings from rescheduling electrical loads.

Investment Analyses

For the investment analyses, photovoltaics (PV), lead acid batteries, and Zinc-Bromide flow batteries were considered. 23 000 ft^2 of roof area, equal to half the square footage of Segundo, was assumed to be available for installation of PV. It was assumed that energy generated by PV onsite or released from energy storage could be used to supply electrical or cooling needs and that cooling is supplied at the time of use. Figure 13 shows the WebOpt interface for the investment & planning DER-CAM. As already mentioned, this interface also handles data management and allows the user to specify the time frame for the historic Segundo load data. By clicking on "Update PI data" the interface executes Datalink as well as the format changer macro (see Figure 5) and updates the load data for the investment optimization. Figure 14 shows the technology data used for the investment & planning runs. Example results for the investment decision and ideal planning are shown in Figure 15.

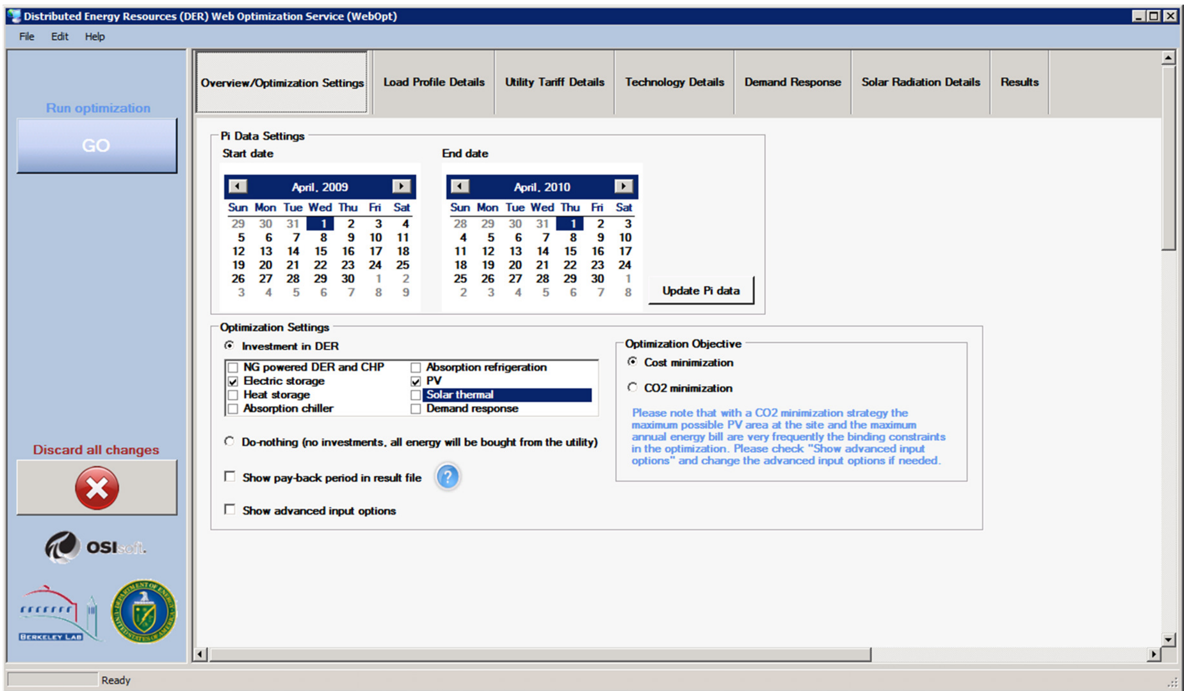


Figure 13. Investment & Planning DER-CAM using PV and Electric Storage as Possible Options

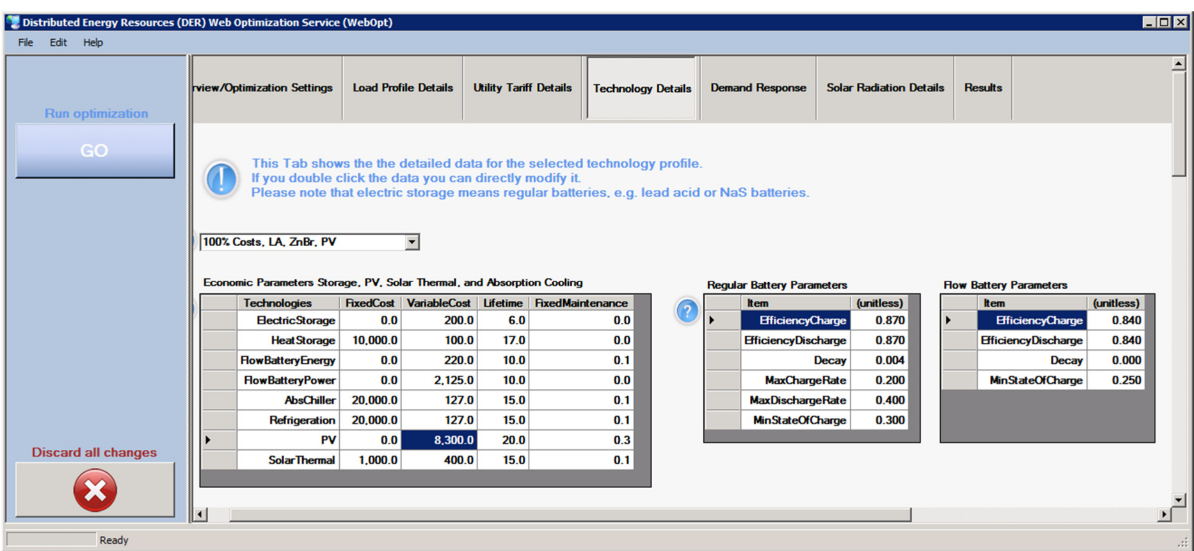


Figure 14. Technology Parameters used for the Investment & Planning DER-CAM Runs

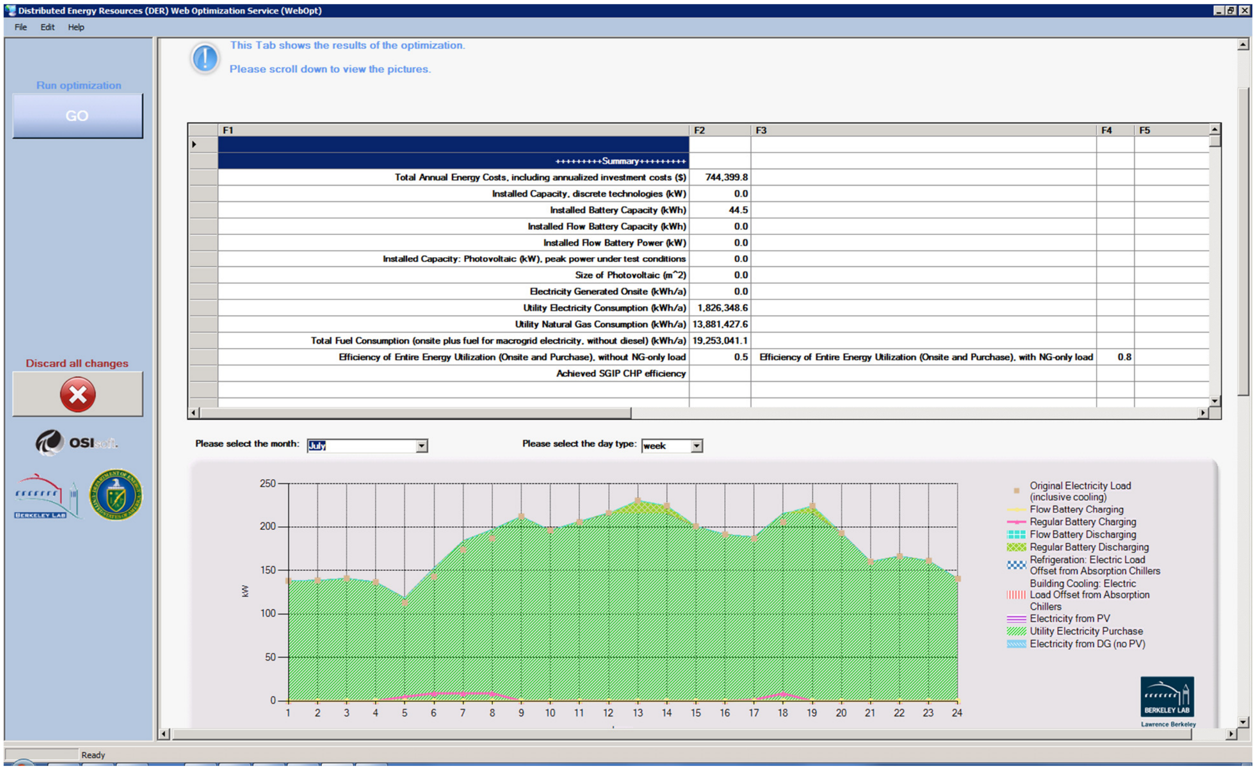


Figure 15. Example Investment & Planning WebOpt Result

Results of the four DER-CAM scenarios are shown below in Table 1. The four different scenarios shown in Table 1 are the cost and CO_2 minimization with the current flat WAPA tariff and the possible E-19 TOU tariff from PG&E (PG&E E-19).

	Cost Minimizing Flat Tariff (Base Case)	Cost Minimizing PG&E E-19 Tariff	CO_2 Minimizing Flat Tariff	CO_2 Minimizing PG&E E-19 Tariff
Reduction in CO_2 Emissions (kg/yr)	0	-1949	334,604	334,604
Reduction in Cost From Base ($/yr)	0	-104,901	-226,839	-304,632
Battery Capacity Installed (kWh)	0	44.5	0	0
Flow Battery Power Installed (kW)	0	0	76.5	76.5
Flow Battery Energy Installed (kWh)	0	0	573.8	573.8
PV Installed (kW)	0	0	326.7	326.7

*** NOTE: negatives are increases**

Table 1: Optimal Investments and Resulting Costs based on the Investment & Planning DER-CAMA cost minimizing run of DER-CAM found that no DG investments are needed to achieve

minimum energy costs. In other words, the low $0.085/kWh flat WAPA rate makes the purchase and installation of equipment uneconomic. This trial did, however, provide a baseline energy cost and carbon emissions level for comparison to other cases. Berkeley Lab used the investment & planning DER-CAM to identify the CO_2 minimizing combination of PV and electrical energy storage. DER-CAM suggests installing a PV array with a rated peak power of 326.7 kW as well as a flow battery with a rated peak power of 76.5 kW and 573.8 kWh of energy storage.

Operational Analyses

The set of equipment from Table 1 was then used in the week-ahead DER-CAM and an example forecast and operation schedule were generated as shown in Figure 16.

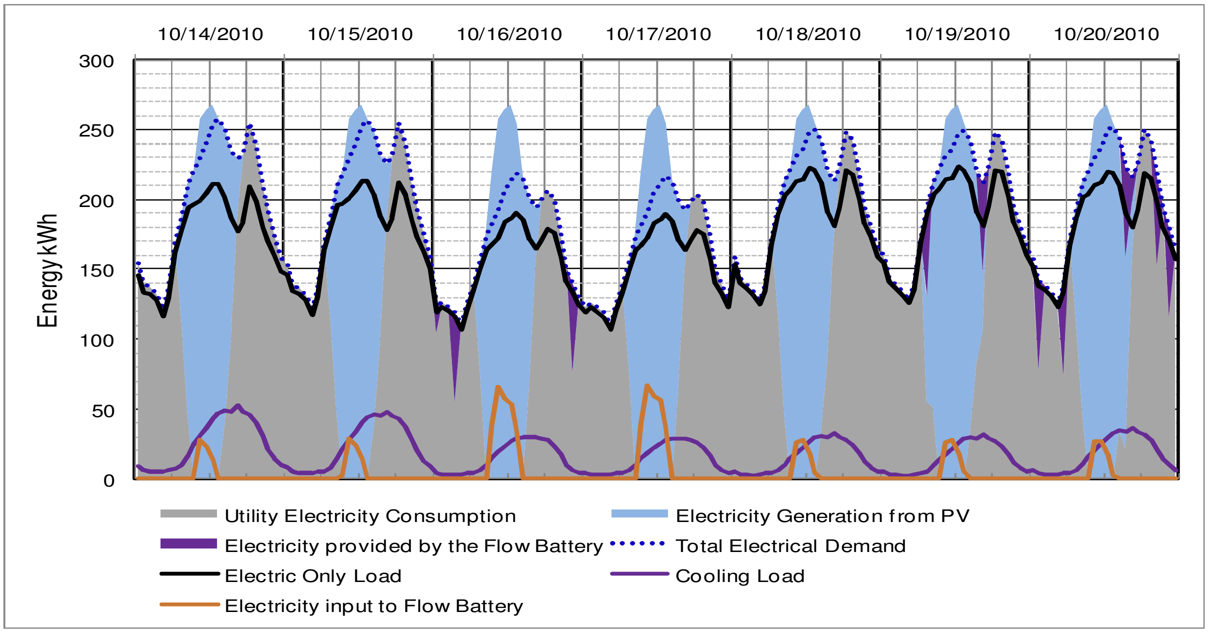


Figure 16: The Forecasted Energy Needs and DG Operating Schedule for the CO_2 Minimizing Combination of DG by using the Week-Ahead DER-CAM[3]

DER-CAM was also used to determine the cost minimizing combination of PV and electrical energy storage if Segundo were subject to PG&E's E-19 tariff (for customers with peak demand between 500 and 1000 kW). The E-19 Tariff has varying TOU pricing as well as monthly TOU peak demand charges. In this case, DER-CAM suggests installing a lead-acid battery bankwith an energy storage capacity of 44.5 kWh. Again, this set of equipment was specified in the week-ahead DER-CAM and an example forecast and an operating schedule was generated as shown in Figure 17.

---

[3] Electricity Only Loads are electric loads for services that only can use electricity. Cooling is not an electricity only load since waste heat / absorption cooling could be also used for cooling. Typical electricity only loads are lighting or computing.

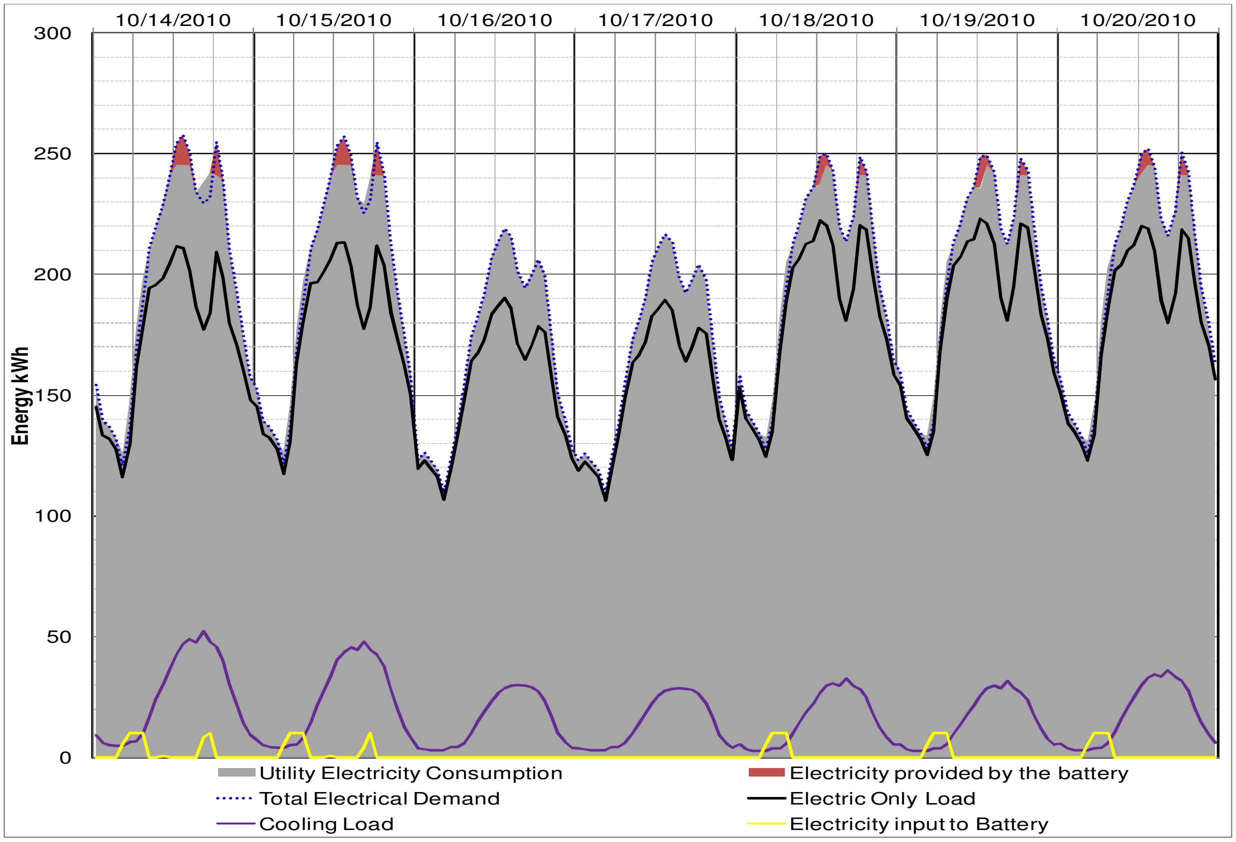


Figure 17: The Forecasted Energy Needs and Cost Minimizing DG Operation Schedule under PG&E E-19 Tariff

Load Rescheduling Experiments

The week-ahead DER-CAM optimization has the capability to suggest a load schedule that minimizes CO_2 emissions or energy costs. For these experiments, Berkeley Lab assumes that up to 15% of the electricity used in any hour at Segundo can be rescheduled to any other hour of the same day. An additional constraint is put in place such that no more than 100 kWh of electricity use can be reallocated to any one hour. No DG equipment is assumed to be in place for the following cases. Figure 18 and Figure 19 show example forecasts of energy needs as well the suggested rescheduling of energy use. In Figure 18, energy use is rescheduled to minimize cost when Segundo is charged under the E-19 Tariff. Compared to when operating without load shifting, volumetric electricity charges are reduced by $98, and monthly peak demand charges are reduced by $588. It should be noted that the monthly peak demand charge is based on the average of the highest 15 minute block of the month, and for the purposes of this project the usage in the week shown is assumed to represent the highest demand for the month.

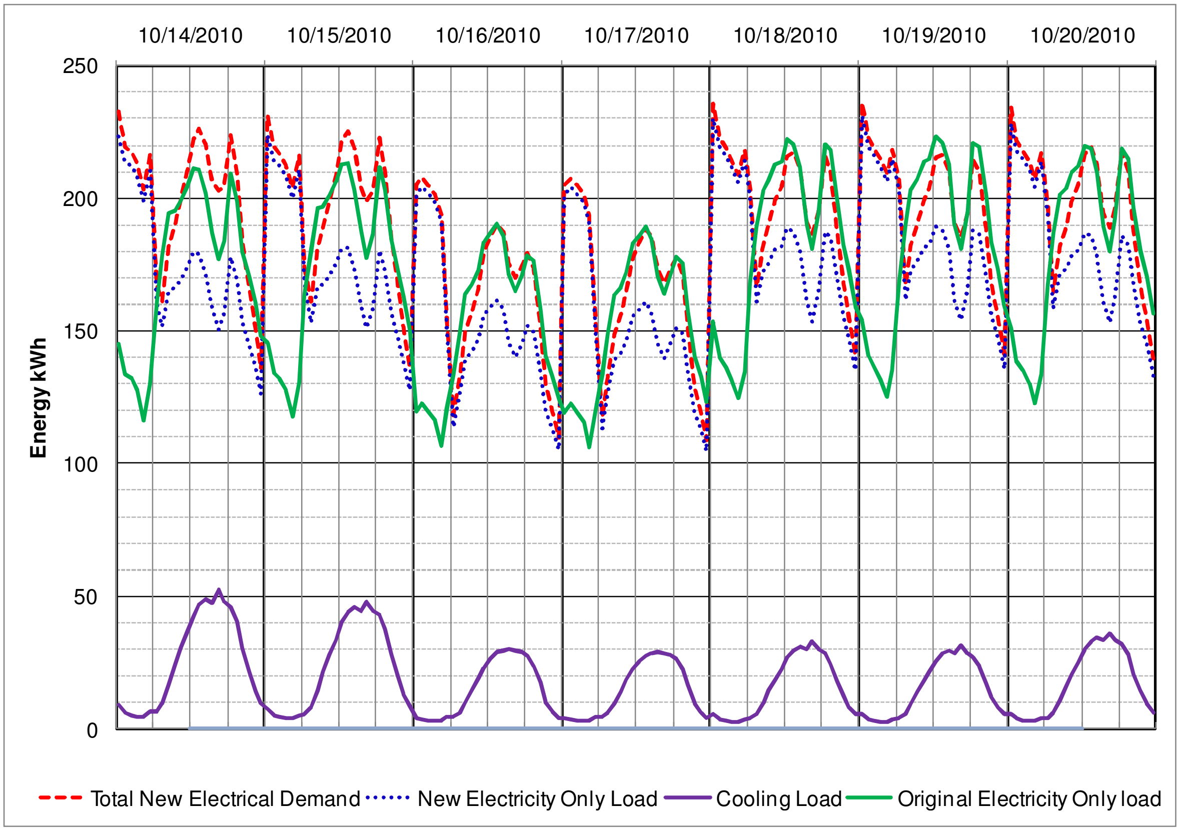


Figure 18 Forecast Energy Demand and Energy Demand Rescheduled to Minimize Cost under PG&E E-19 Tariff

Figure 19 shows how electricity consumption should be rescheduled to minimize CO_2 emissions if the emissions intensity for energy consumed at Segundo. Rescheduling saves 937 $kgCO_2$ for this optimized week. Please note that this run also assumes no adopted DG and all energy needs to be purchased from the utility.

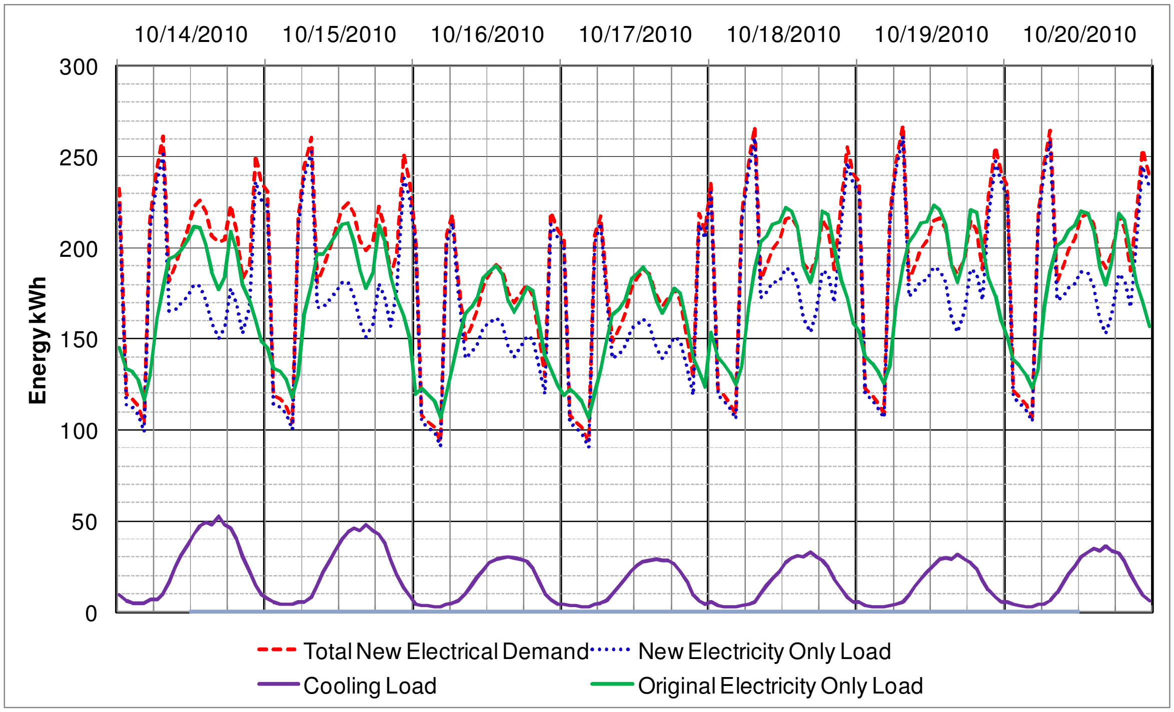


Figure 19: Forecast Energy Demand and Energy Demand Rescheduled to Minimize CO_2 Assuming no DG and All Energy From PG&E

References

CAISO (2010): "*Integration of Renewable Resources: Operational Requirements and Generation Fleet Capacity at 20% RPS*," August, 2010.

CIMIS, http://www.cimis.water.ca.gov/cimis/data.jsp

DER at Lawrence Berkeley National Laboratory, http://der.lbl.gov/

E3, http://www.ethree.com/public_projects/ghg.html

Marnay C., G. Venkataramanan, M. Stadler, A. Siddiqui, R. Firestone, B. Chandran (2008a): *"Optimal Technology Selection and Operation of Microgrids in Commercial Buildings,"* IEEE Transactions on Power Systems, Volume 23, Issue 3, page 975-982, August 2008, ISSN 0885-8950.

NOAA, http://www.wrh.noaa.gov/forecast/wxtables/index.php?lat=38.56&lon=-121.75&table=custom&duration=7&interval=1

OSIsoft, http://www.OSIsoft.com/

PG&E E-19, http://www.pge.com/tariffs/tm2/pdf/ELEC_SCHEDS_E-19.pdf

R-software, http://www.r-project.org/

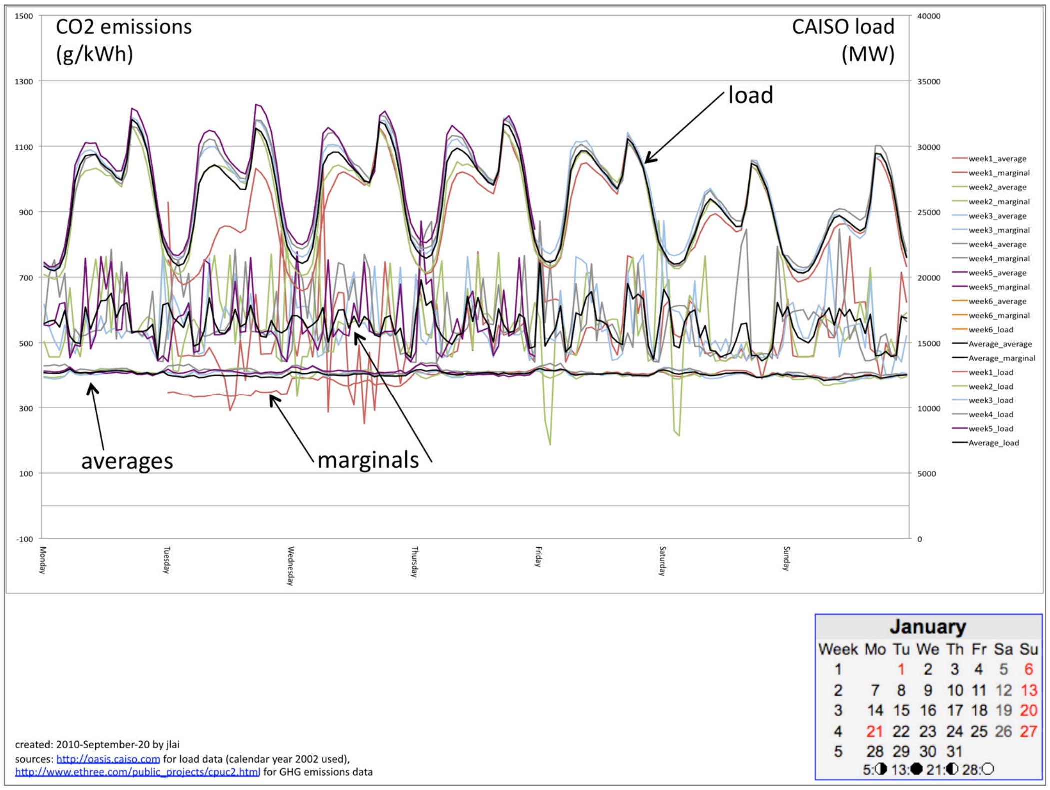
Siddiqui Afzal, Michael Stadler, Chris Marnay, and Judy Lai (2010): *"Optimal Control of Distributed Energy Resources and Demand Response under Uncertainty,"* IAEE's Rio 2010 International Conference, June 6-9, 2010, InterContinental Rio Hotel – Rio de Janeiro, Brazil, LBNL-3828E.

Stadler Michael, Afzal Siddiqui, Chris Marnay, Hirohisa Aki, Judy Lai (2009): *"Control of Greenhouse Gas Emissions by Optimal DER Technology Investment and Energy Management in Zero-Net-Energy Buildings,"* European Transactions on Electrical Power 2009, Special Issue on Microgrids and Energy Management, LBNL-2692E.

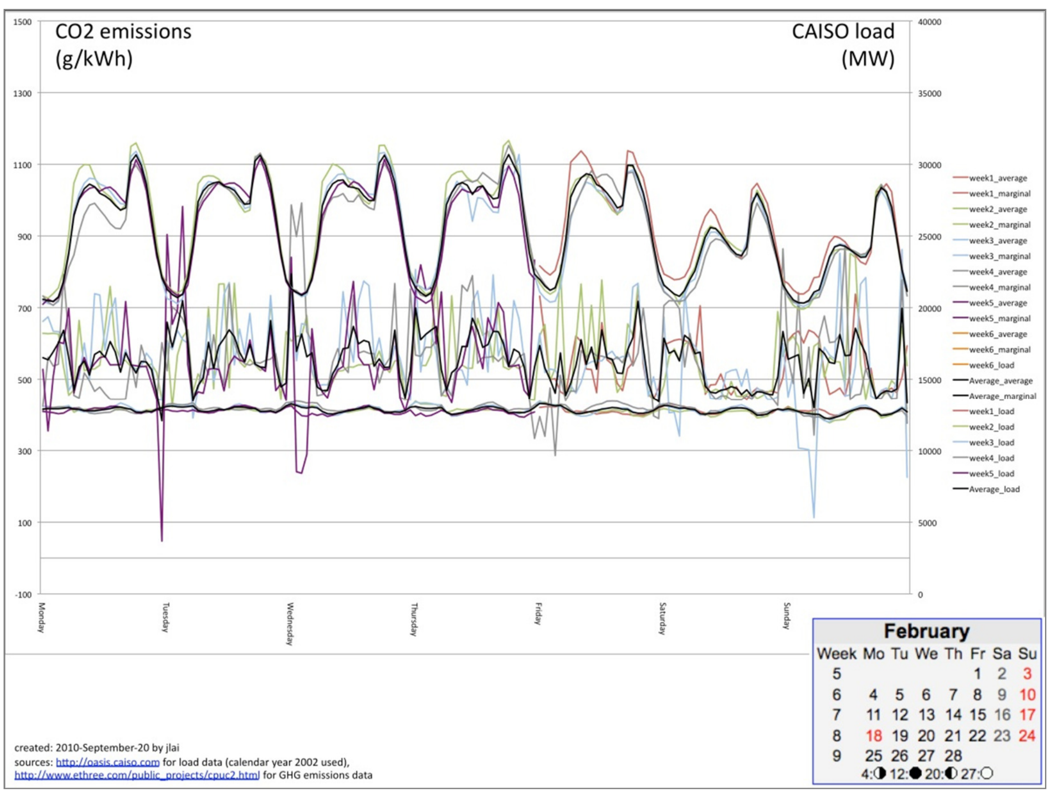
Stadler, Michael, Chris Marnay, Afzal Siddiqui, Judy Lai, Brian Coffey, and Hirohisa Aki (2008): *"Effect of Heat and Electricity Storage and Reliability on Microgrid Viability: A Study of Commercial Buildings in California and New York States."*, LBNL-1334E, December 2008.

APPENDIX A: CAISO System Load and California CO_2 Emission Rates Comparison

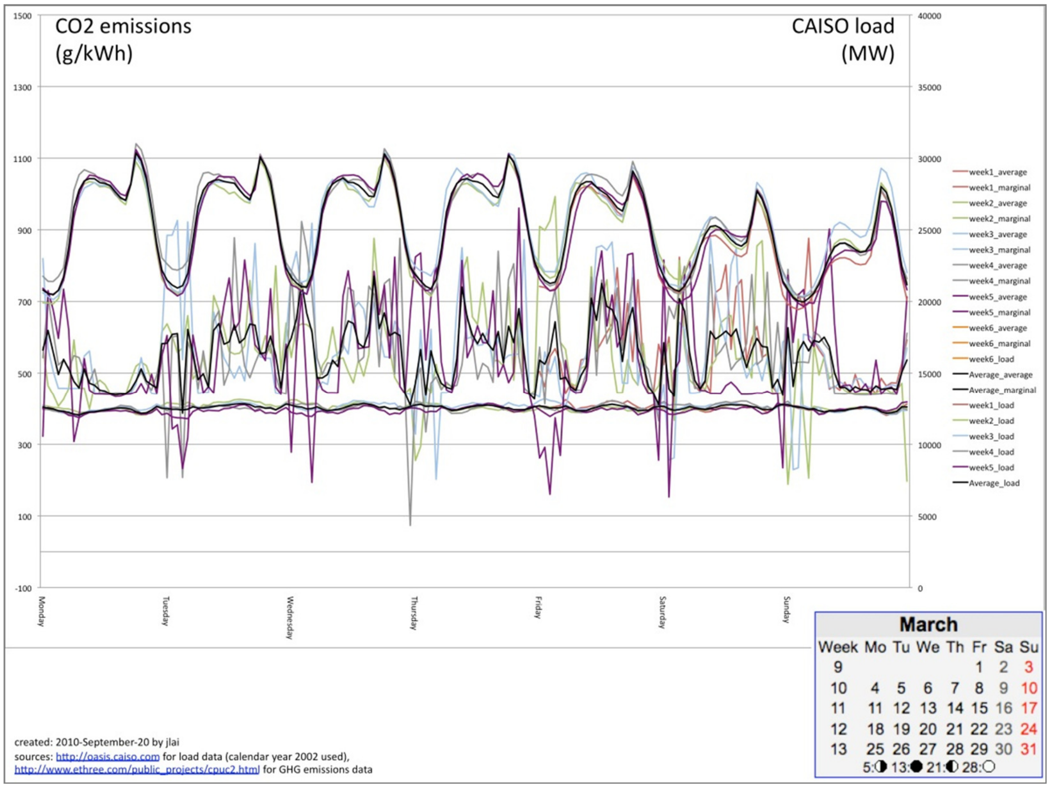
Using the 2002 calendar, the CAISO system load and California CO_2 emissions are estimated for 2008 (emissions estimated using E3's GHG calculator). Each plot below shows a month's worth of hourly data for CAISO load, average CO_2 emissions, and marginal CO_2 emissions. Each month is represented by 5 or 6 weeks (count starts from Monday, see calendar in lower right hand corner). Each week is graphed separately, and the average of the weeks' load and emission values for each hour is shown in black. As can be seen from the plots, there is great volatility from hour to hour for the marginal emissions. The *average* of the hourly marginal emissions are used by WebOpt to determine if or when carbon savings can be achieved.



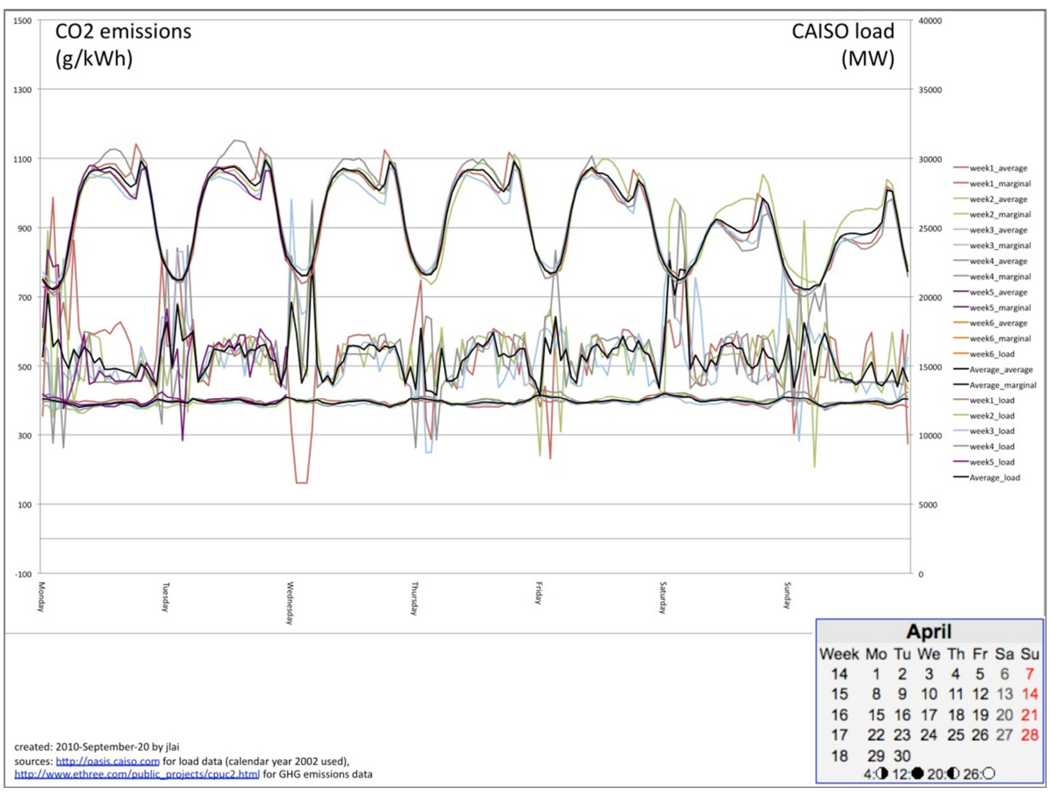
Appendix Figure 1. January CO_2 Emissions and Load Comparison



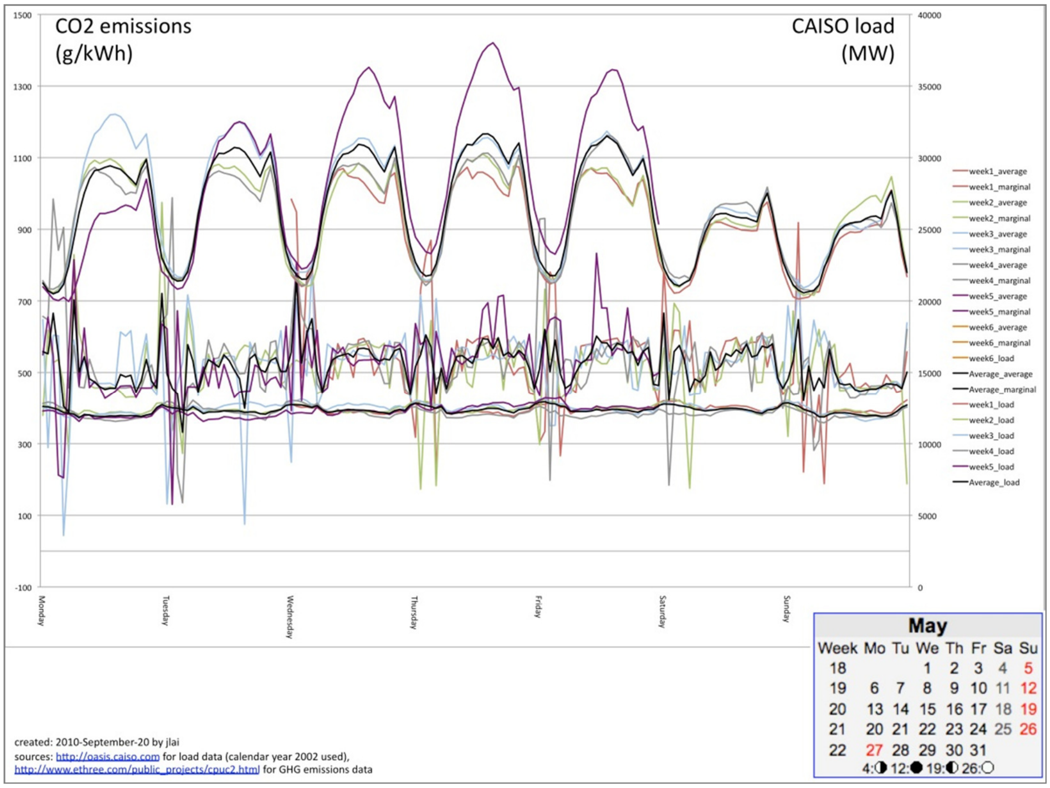
Appendix Figure 2. February CO_2 Emissions and Load Comparison



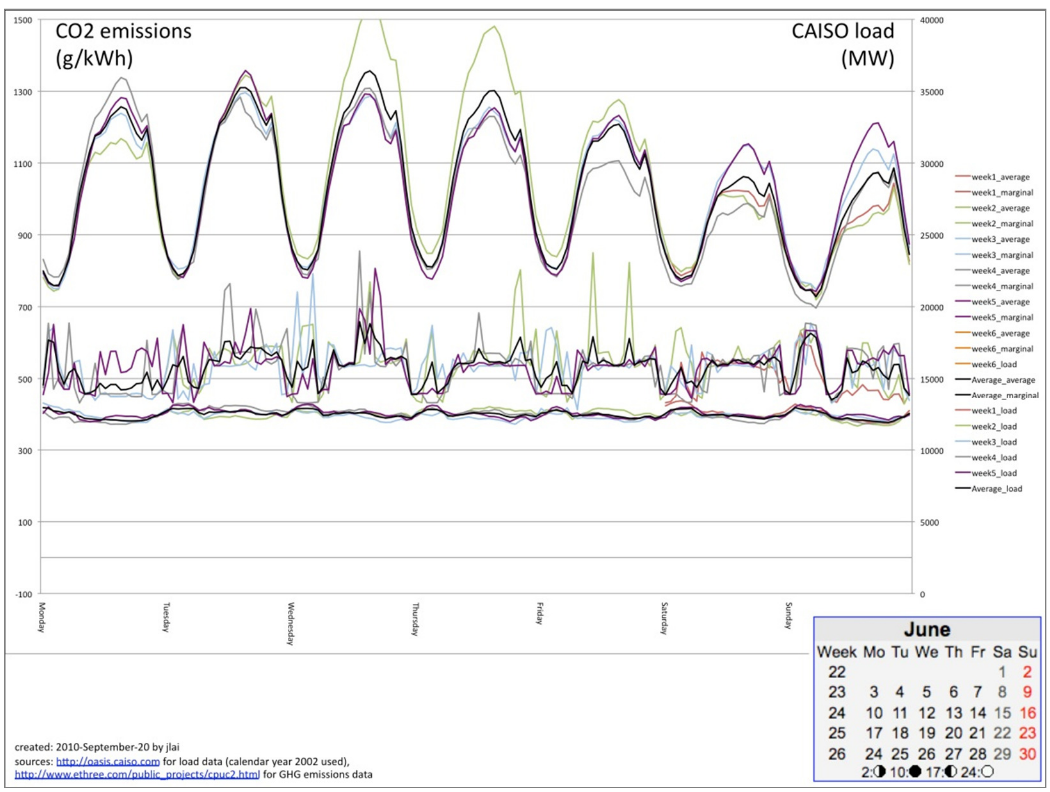
Appendix Figure 3. March CO_2 Emissions and Load Comparison



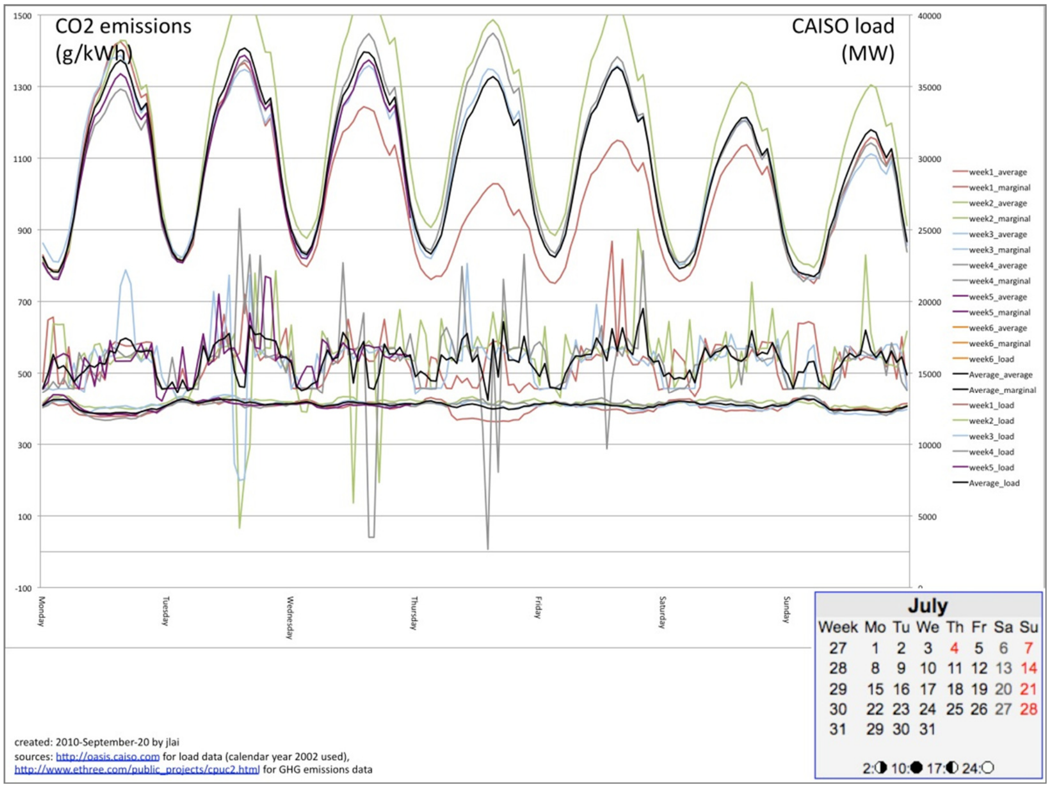
Appendix Figure 4. April CO_2 Emissions and Load Comparison



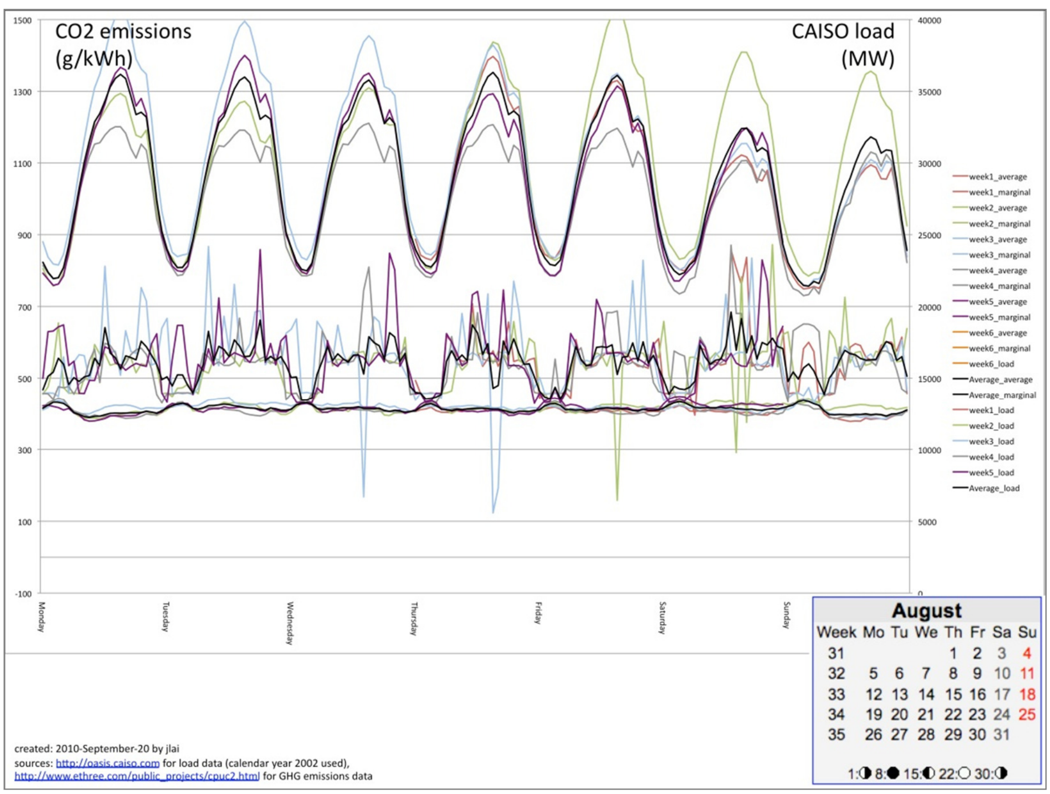
Appendix Figure 5. May CO_2 Emissions and Load Comparison



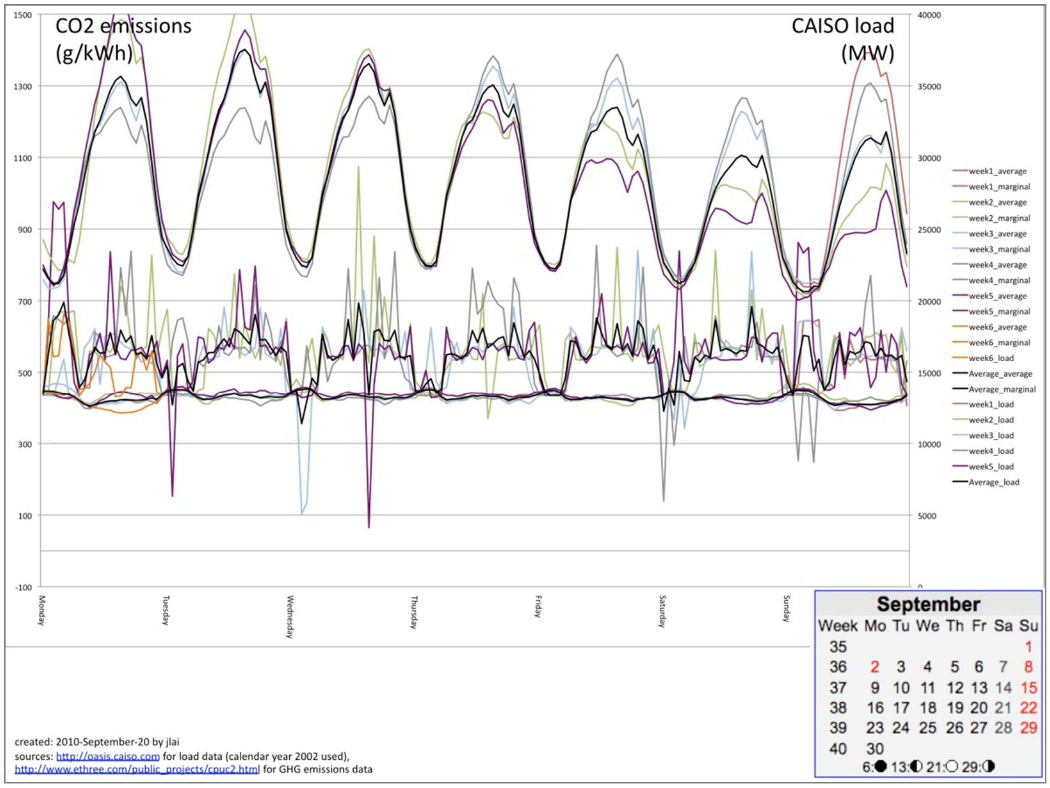
Appendix Figure 6. June CO_2 Emissions and Load Comparison



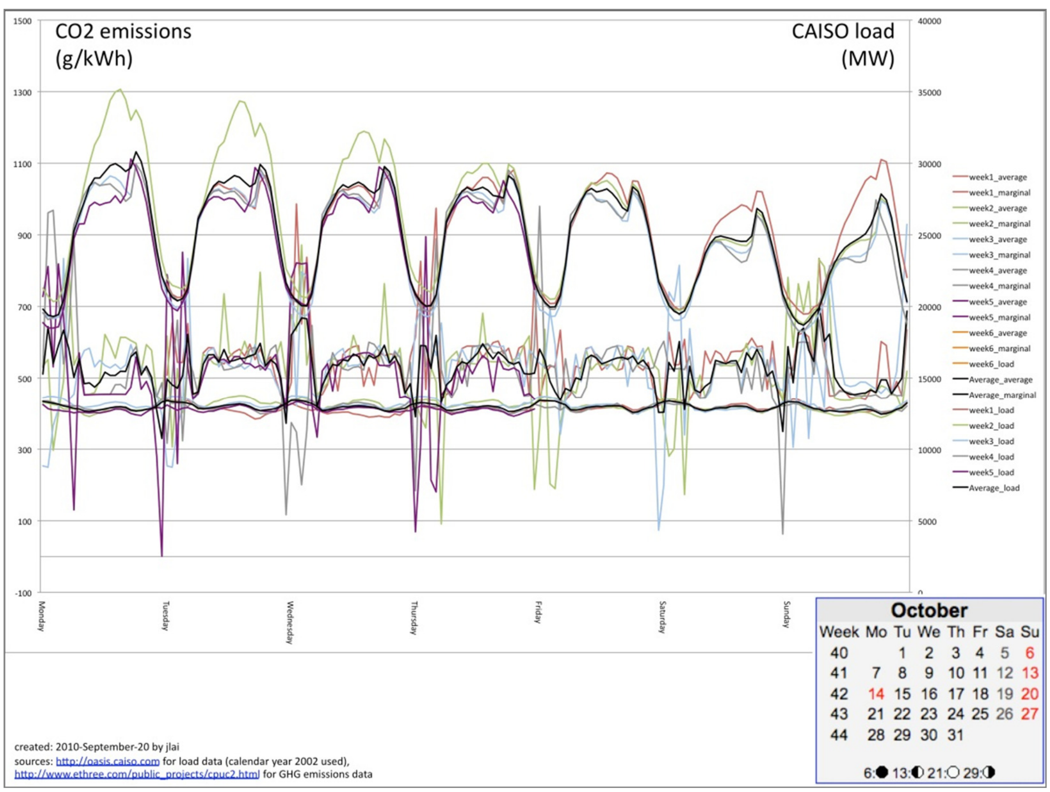
Appendix Figure 7. July CO_2 Emissions and Load Comparison



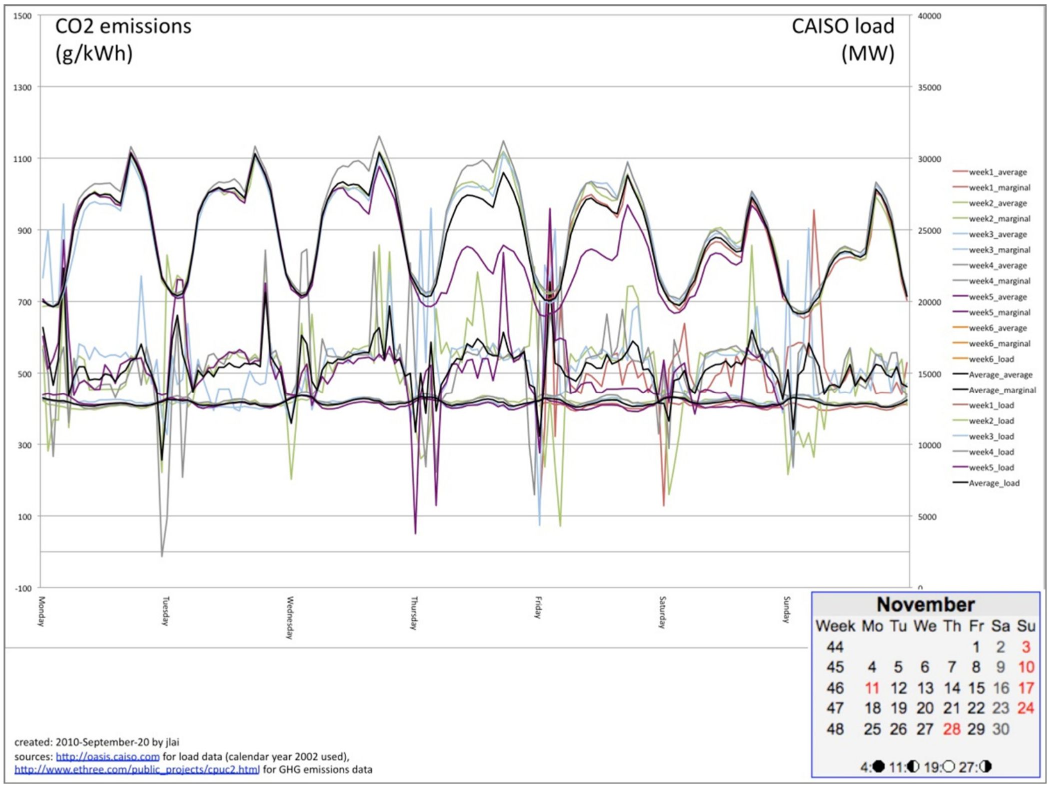
Appendix Figure 8. August CO_2 Emissions and Load Comparison



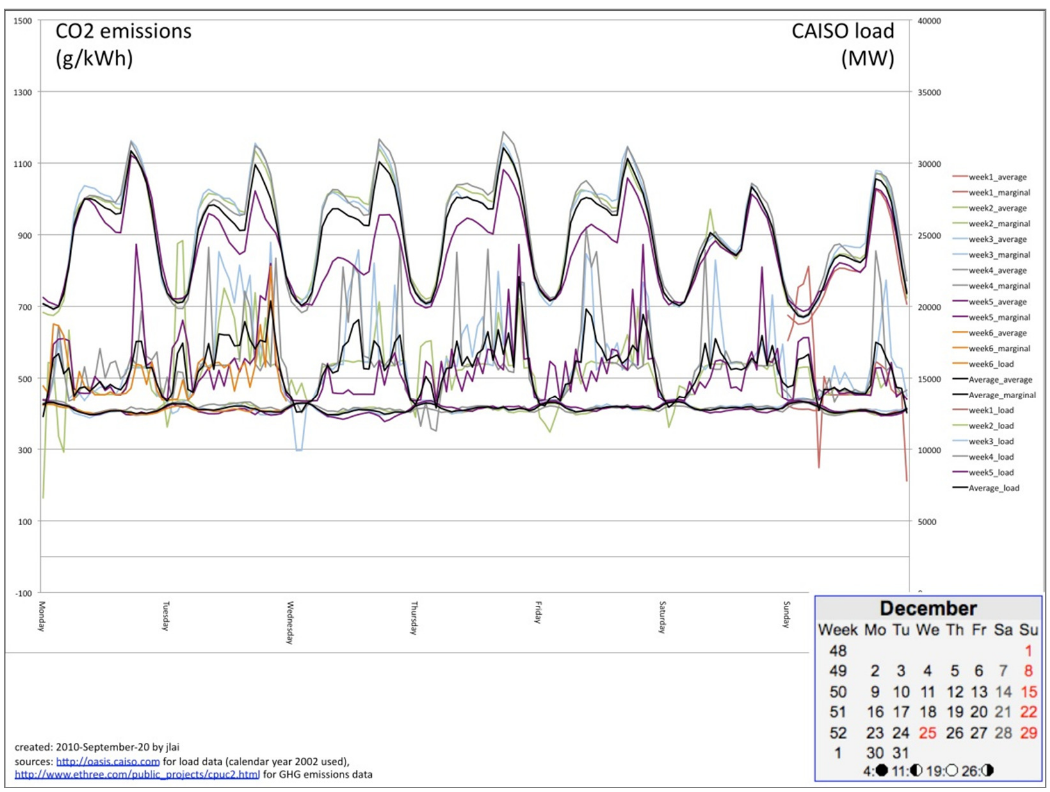
Appendix Figure 9. September CO_2 Emissions and Load Comparison



Appendix Figure 10. October CO_2 Emissions and Load Comparison



Appendix Figure 11. November CO_2 Emissions and Load Comparison



Appendix Figure 12. December CO_2 Emissions and Load Comparison

APPENDIX B: Background on DER-CAM

DER-CAM is an economic-engineering model of customer DER adoption implemented in the General Algebraic Modeling System (GAMS®) optimization software. This model has been in development at Berkeley Lab since 2000. The objective of the model is to minimize the cost or carbon emissions of operating on-site generation and combined heat and power (CHP) systems, either for individual customer sites or a μGrid. To achieve this objective, the following issues must be addressed:

- Which is the lowest-cost or lowest-carbon combination of distributed generation technologies that a specific customer can install?
- What is the appropriate level of installed capacity of these technologies that minimizes cost or carbon emissions?
- How should the installed capacity be operated so as to minimize the total customer energy bill or carbon emissions?

How does DER-CAM Work?

The DER-CAM model chooses which DG and/or CHP technologies a customer should adopt and how that technology should be operated based on specific site load and price information, and performance data for available equipment options. The inputs to and outputs from DER-CAM are illustrated below.

Key Inputs into the Model are:

1. Customer's end-use load profiles (typically for space heat, hot water, gas only, cooling, and electricity only)
2. Customer's default electricity tariff, natural gas prices, and other relevant price data
3. Capital, operating and maintenance (O&M), and fuel costs of the various available technologies, together with the interest rate on customer investment
4. Basic physical characteristics of alternative generating, heat recovery and cooling technologies, including the thermal-electric ratio that determines how much residual heat is available as a function of generator electric output.

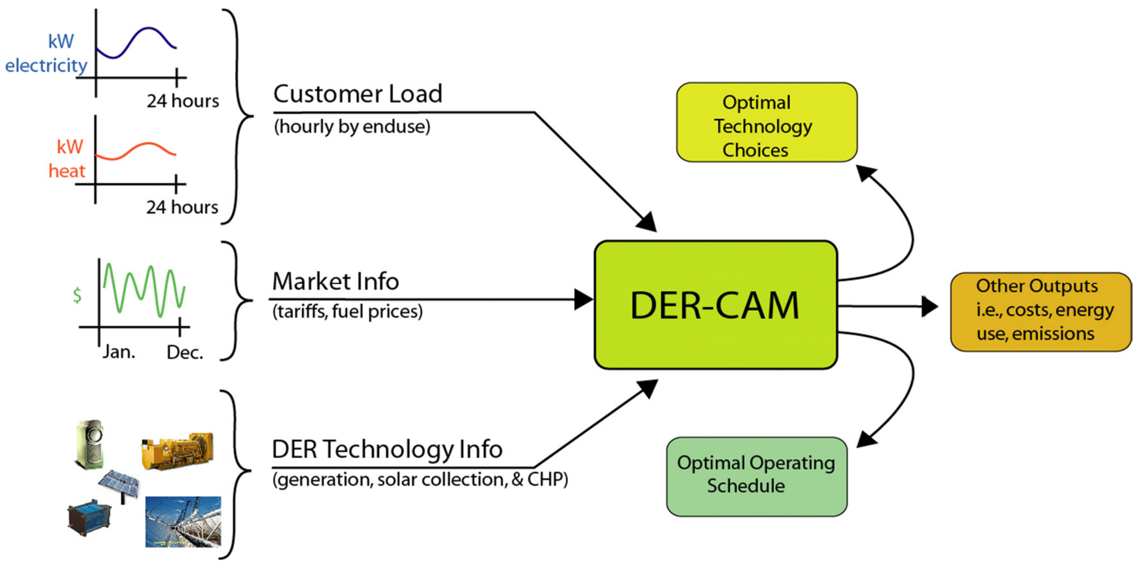
Outputs to be Determined by the Optimization Model are:

1. Capacities of DG and CHP technology or combination of technologies to be installed
2. When and how much of the capacity installed will be running
3. Total cost of supplying the electric and heat loads, and
4. CO_2 emissions.

The key Assumptions are:

1. Customer decisions are made based only on direct economic or environmental criteria. In other words, the only possible benefit is a reduction in the customer's electricity bill or CO_2 emission.

2. No deterioration in output or efficiency during the lifetime of the equipment is considered. Furthermore, start-up and other ramping constraints are not included.
3. Reliability and power quality benefits, as well as economies of scale in O&M costs for multiple units of the same technology are not directly taken into account.
4. Possible reliability or power quality improvements accruing to customers are not considered directly.

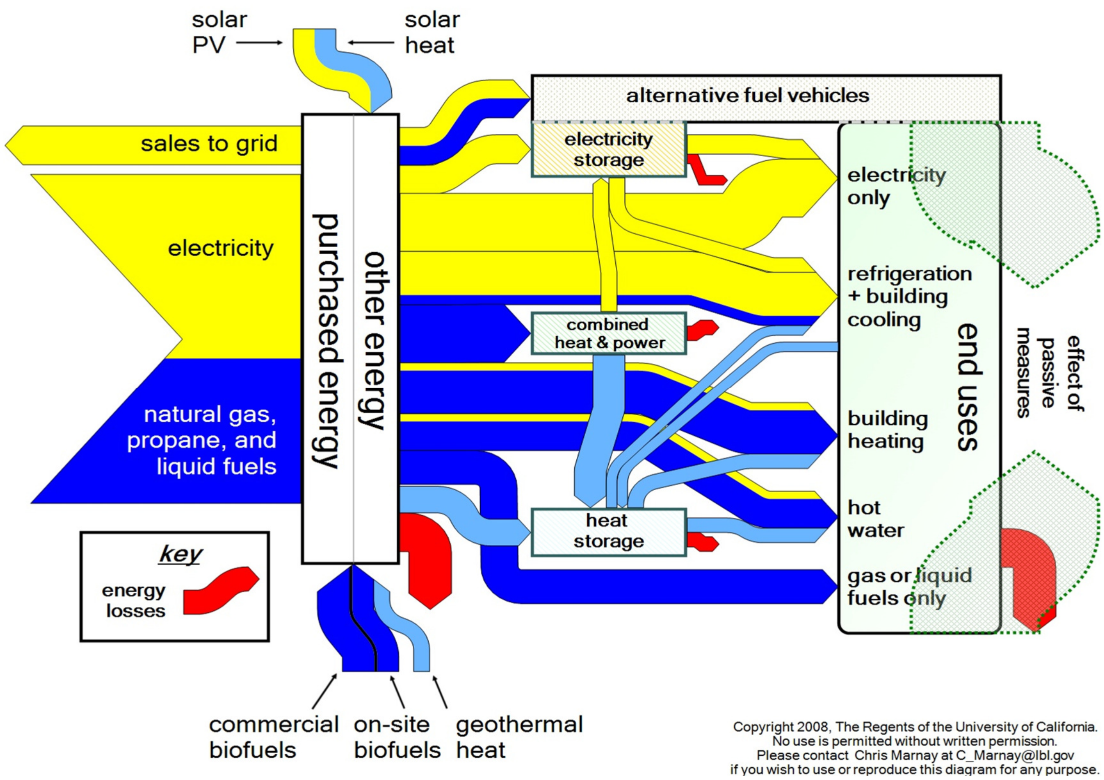


Appendix Figure 13. Structure of DER-CAM

Simultaneous Optimization Approach:

The next figure shows a high-level schematic of the energy flow modeled in DER-CAM. Possible energy inputs to the site are solar insulation, utility electricity and natural gas. For a given DG investment decision, DER-CAM selects the optimal combination of utility purchase and on-site generation required to meet the site's end-use loads at each time step.

a) Electricity-only loads (e.g. lighting and office equipment) can only be met by electricity
b) Cooling loads can be met either by electricity or by heat (via absorption chiller)
c) Hot water and space heating loads can be met either by recovered heat or by natural gas
d) Natural gas-only loads (e.g. mostly cooking) can only be met by natural gas.



Appendix Figure 14. Sankey Showing the Energy Flows Modeled by DER-CAM

For more information on DER-CAM please see DER at Lawrence Berkeley National Laboratory, Marnay et al. 2008, Stadler et al. 2009, and Stadler et al. 2008.

APPENDIX C: Background on PI Server and PI System

The following description of the PI System is taken directly from the OSIsoft website:

The PI System® brings all operational data into a single system that can deliver it to users at all levels of the company - from the plant floor to the enterprise level. This capability is often described as a Process Historian and the PI System is considered the standard of Process Historians. The PI System keeps business-critical data always online and available in a specialized time-series database by:

- *gathering event-driven data, in real-time, from multiple sources across the plant and/or enterprise*
- *applying advanced analytical calculations and business rules to contextualize and analyze this data*
- *configuring smart and thin client tools to distribute and visualize knowledge/ information to display critical operational metrics and integrate the user experience across different roles within the enterprise.*

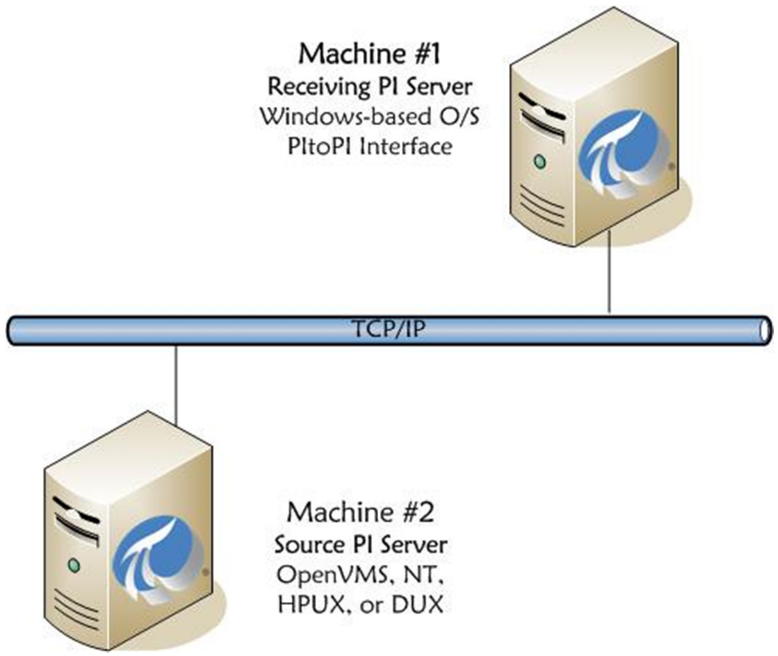
  The PI System functionality incorporates many features for analyzing, contextualizing, and visualizing real-time PI data.

The PItoPI interface copies tag data from one PI server to another. Data is moved in one direction, meaning data is copied from the source to the receiving PI server (also referred to as target PI server). The interface must run on a Windows Intel-based operating system (Windows 2000 SP4, XP, 2003, or higher). In this project, Windows Server 2008 is used.

Interface tags are created on the receiving PI server. Each interface tag is configured to receive data for a unique source tag. Tags receive either archive or exception data updates from the source tag. Exception data is data that has not yet been subjected to compression. The type of data collection, exception or archive, is configured through scan class assignment. By default, all tags belonging to the first scan class receive exception data. Tags assigned to any other defined scan class receive archive data.

The interface supports history recovery. History recovery enables users to recover data for time periods when the interface was not running or otherwise unable to collect data. The history recovery period is configurable; the default is 8 hours. Users have the option of performing time-range specific history recovery by specifying a start and end time. In this configuration the interface collects data for the specified time period then exits. We used this feature to backfill the LBNL PI server with data from the UC Davis PI server. The backfill data, in this case, dated as far back as March 01, 2009.

There are different ways in which the PItoPI interface can be configured. We used the following configuration described in the user manual of the interface.

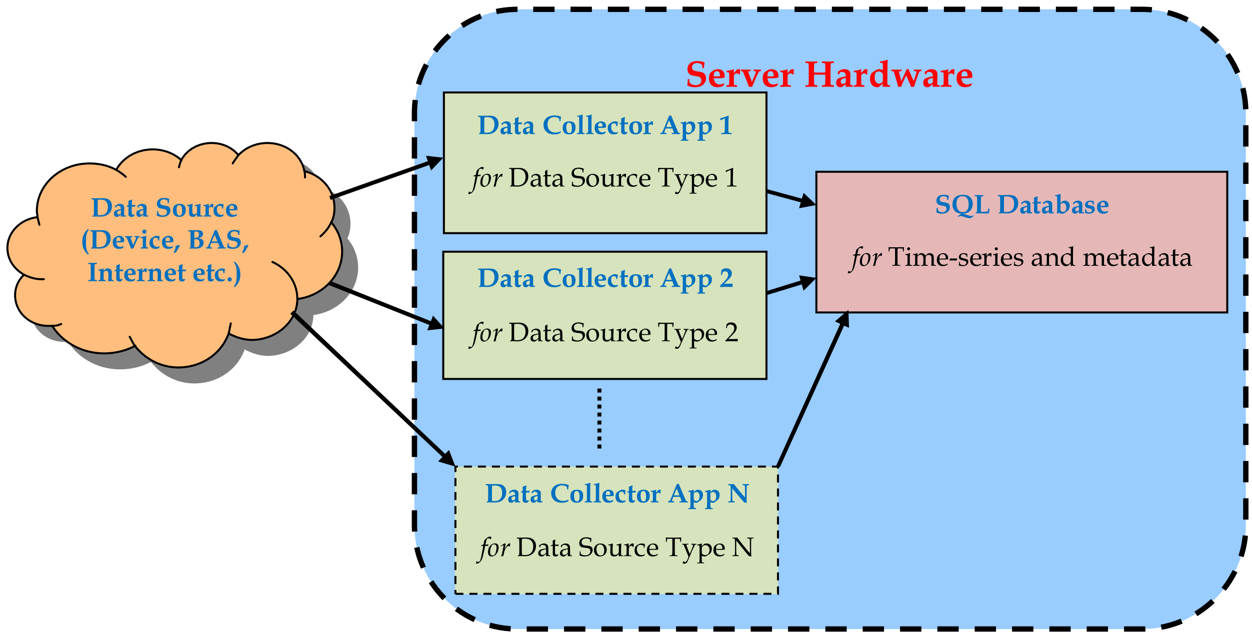


Appendix Figure 15. An Exceedingly Boring Image Showing two Computers Connected via TCP/IP

APPENDIX D: Unified Data Manager (UDM)

System Architecture

The Figure below shows a simplified schematic diagram of the system architecture used in Unified Data Manager (UDM). The dotted portions attempt to depict the extensibility of the system. A few examples of *Data Source Type* are xml files, structured csv files, external database, bacnet, modbus, SNMP, XML SOAP etc. Currently, *Data Collector Application* (app) for only two data source types have been implemented – the xml file and the csv file. The backend database and the data collector app can reside either on the same host or on different hosts. Berkeley Lab decided to use a database that conforms to the SQL standard because the support ecosystem for these databases is widely available. As a first attempt, the postgreSQL database program as the backend database was used, keeping SQL statements as standard-oriented as possible. But still there are places where they are specific to postgreSQL. It should be noted here that UDM's database structure is very different from the way usually Relational Databases are used for time-series data.



Appendix Figure 16. Simplified Schematic Diagram of the System Architecture used in Unified Data Manager (UDM)

Every time-series data stream is a different point in the system. In its current configuration, only a limited amount of metadata (data about data; in this case, information about the time-series data) can be stored in the database. Right now all that information is stored in the database in a table called *Point*. Each row in that table is dedicated to a single time-series point. The columns contain information about the unique identifier for a point, the human readable description of it, the unique identifier for the data source type, how to gather data for this point from the data source, how to interpret that data, whether the data is forecast data and how often the data source should be polled to collect the data for this point. The time-series data is stored in tables dedicated for individual points. The mapping between the points and their *DataTable*s is maintained in a table called *Point_DataTable_Map*. Every time a point is created in the system a new datatable, dedicated to that point, is created in the database. Similarly, when a point is deleted from the system, all datatables associated with it are dropped from the database.

All the applications were built on the Java platform to make the system compatible with both Windows and unix operating systems.

Since the database structure is not very flat, making changes to the system using plain SQL statements can prove to be cumbersome. Therefore a system administration utility has been developed which takes csv files as its input and outputs another csv file if there is supposed to be any output. The operations supported at present include point creation, point deletion and point import. The first two are probably self-explanatory; the point import action imports all the details about all the points into a csv file.

Known Problems

1. Although the timestamps on the data are perfect, polling takes place at GMT time. It is not perceivable for points with 1 hour or shorter poll interval, but is evident for points

with longer poll intervals. The work around has been to poll once every 8 hours for points that could have used a poll interval of 1 day.

2. In the specific case of the NOAA weather forecast, The XML file Data Collector might fetch faulty timestamp for the first 6 days of the year.

Proposed Near-Term Enhancements

1. Package all the files necessary for initial installation (jar libraries, UDM class files, database configuration text files) in a single container and then copy all the files into the respective directories during installation. This would simplify installation.
2. Implement logging.
3. Incorporate string datatype for header_alue_pair in method getPoints() of class PointUtility. This would help in creating better search for points.
4. Build generic queries.
5. Take care of the changes in the CLASSPATH so that user does not have to do it separately.
6. Right now, file download takes place in the bin directory under the UDM root directory. User should be able to specify this directory. It should be accepted as an input for the data collector app.
7. Link the starttime constraint of the datatable to the starttime stored in the Point_DataTable_Map as opposed to a hard number.
8. Portal-based visualization
9. RDBMS Data Collector
10. BACNet Data Collector
11. Modbus Data Collector
12. The method sendToDatabase of the DataCollector class, under realtime mode, creates the SQL query to INSERT all the data values for all the points in the thread in a single query. This enhances performance. However, if even one point has a duplicate key value existing in the database, the INSERT attempt would fail and none of the other values would get stored. In the catch block of this portion, a second attempt needs to be made where the rogue data point(s) will be identified and be excluded from the INSERT attempt.
13. Right now, the data collector apps query the database only at the time of start-up and cache that information for the rest of its life. It should check for changes in the system on a regular basis.
14. The timestamp finding technique in the Text File Data Collector should be made more generic.
15. Implement all the different scenarios left as TODO in the code (e.g. when timestamp is system generated).
16. Poll time offset should be calculated to distribute load more evenly. Right now it is 5% of the poll interval. It has to take the data throughput at a particular time under consideration.

17. In case of download or unzip, append the filenames with “backfill” so that backfill and realtime mode processes can run simultaneously without any conflict.
18. UPDATE and UPSERT option for backfill mode: right now, in the backfill mode, only those data can be inserted for which the timestamp does NOT already exist. However, under different situations (such as performing a rerun of a calculation after making changes to the model), one might be interested in overwriting existing data. Not only that, one might be interested in overwriting existing data as well as inserting new data.
19. Forecast interval, such as 2 day forecast or 7 day forecast, should be user defined.