RESEARCH PERFORMANCE PROGRESS REPORT

Project Title:	Intra-Hour Dispatch and Automatic Generator Control Demonstration with Solar Forecasting
Project Period:	09/30/2013 – 09/30/2015
Reporting Period:	09/30/2013 – 09/30/2015
Submission Date:	02/25/2016
Recipient:	University of California, San Diego
Recipient DUNS #:	804-355790
Address:	9500 Gilman Drive La Jolla, CA 92093
Award Number:	DE-EE0006330
Awarding Agency:	DOE EERE SunShot SUNRISE subprogram
Working Partners:	Clean Power Research PNNL SMUD
Cost-Sharing Partners:	California Public Utility Commission
Principal Investigator:	Carlos F. M. Coimbra Professor Phone: 858-534-4285 Email: ccoimbra@ucsd.edu
HQ Tech Manager:	Rebecca Hott
HQ Project Officer:	Christine Bing
GO Grant Specialist:	Jeremey Mikrut
GO Contracting Officer:	Diana Bobo

Executive Summary

In this project we address multiple resource integration challenges associated with increasing levels of solar penetration that arise from the variability and uncertainty in solar irradiance. We will model the SMUD service region as its own balancing region, and develop an integrated, real-time operational tool that takes solar-load forecast uncertainties into consideration and commits optimal energy resources and reserves for intra-hour and intra-day decisions. The primary objectives of this effort are to reduce power system operation cost by committing appropriate amount of energy resources and reserves, as well as to provide operators a prediction of the generation fleet's behavior in real time for realistic PV penetration scenarios. The proposed methodology includes the following steps: clustering analysis on the expected solar variability per region for the SMUD system, Day-ahead (DA) and real-time (RT) load forecasts for the entire service areas, 1-year of intra-hour CPR forecasts for cluster centers, 1-year of smart re-forecasting CPR forecasts in real-time for determination of irreducible errors, and uncertainty quantification for integrated solar-load for both distributed and central stations (selected locations within service region) PV generation.

Table of Contents

Background **4**

Introduction **7**

Task 1.0: Detailed Clustering Analysis of the Solar Resources in SMUD Territory **9**
Task 1.0 Deliverables 9
Task 1.0 Milestone 9
Task 2.0: Clean Power Research Forecasts for All Sectors of SMUD's Territory **10**
Task 2.0 Subtasks 10
Task 2.0 Deliverables 11
Task 2.0 Milestones 11
Task 3.0 Integrated Solar-Load Forecasts for SMUD **12**
Task 3.0 Subtasks 12
Task 3.0 Milestones 12
Task 4.0: Provide DOE with Publicly Accessible Data Created from this Project **12**
Task 4.0 Deliverables 13
Task 5.0 Optimization of CPR's forecasts using UCSD's methodologies **13**
Task 5.0 Subtasks 13
Task 5.0 Milestones 13
Task 5 Deliverables 14
Task 6.0 Solar-Load Uncertainty Implementation into ESIOS and PROMOD **14**
Task 6.0 Subtasks 14
Task 6 Milestones 15
Task 6 Deliverables 15
Task 7: Provide DOE with Public Accessible Data Created from this Project **16**
Task 7.0 Deliverables 16

Project Results and Discussion **17**
Task 1.0: Detailed Clustering Analysis of the Solar Resources in SMUD Territory **17**
Data and Preprocessing 17
Experimental Analysis 27
Clustering Results 30
Task 2.0: Clean Power Research Forecasts for All Sectors of SMUD's Territory **33**
Data for the benchmarking 33
Methods 34
Results 35
Task 3.0: Integrated Solar-Load Forecasts for SMUD **37**
Methods 37
HAM Load Re-Forecasting Results 44
DAM Load Re-Forecasting Results 46
Task 5.0 Optimization of CPR's forecasts using UCSD's methodologies **50**
Re-Forecast methodology 50
Optimization Algorithm 51
Error metrics 52
Results 53
Task 6.0 Solar-Load Uncertainty Implementation into ESIOS and PROM **61**
Methodology for Reserve Requirements 62
Day-ahead Planning Reserve Calculation 62
Regulation Reserve Calculation 63

Reserve Requirements for SMUD Study Cases 65
DA Planning Reserve 65
Regulation Reserve 67
Statistics of PV Forecast Performance 68
Remarks on PV Forecast Impacts on Reserve Requirements 69
Intra-hour Dispatch and AGC Simulations 69
Task 7: Dissemination of the work supported by this Project 71
Journal Publications 74
Technical Presentations 74
Works cited 75

Background

The heightened awareness surrounding climate change on a global scale underlines the significance of the technological and economic issues associated with increased levels of renewable penetration into the power grid. In particular, solar power generation has recently seen strong increases in market share and corresponding growth in grid penetration rates, which has led to issues concerning the variability of the solar resource at ground level and associated ramp-rates in solar power production. These issues arise from the coupling of cloud dynamics and resource availability that are radically unlike traditional power generation technologies (such as fossil and nuclear power). These were designed to run in stable output modes and have resulted in the majority of power grid variability originating from demand fluctuations (Lew et al., 2013). Consequently, utilities must develop new mitigating measures to offset weather-dependent solar power fluctuations in order to effectively integrate increased levels of solar penetration into the grid while maintaining grid reliability (Boyle, 2009; Helman et al., 2010; Rodriguez, 2010; Inman et al., 2013).

A number of methods have been proposed for mitigating such variability effects through the improvement of irradiance forecasting techniques, see e.g., Marquez and Coimbra (2011); Marquez et al. (2013); Chow et al. (2011). Quality forecasting models allow for the development of tools aimed at lowering operational costs through improved grid regulation, load-following production, power scheduling, unit commitment, automatic generation control, and a diminished need to dispatch ancillary fossil fuel generators (Inman et al., 2013). To date, however, high fidelity, robust solar forecast systems that work for widely different microclimates remain evasive. The problem is of great complexity due to the non-linear and chaotic effect of cloud motion on solar irradiance at the ground level. However, a number of promising approaches have been developed in the past years, and the incipient research field of solar meteorology for renewable generation has grown considerably by aggregating diverse areas of knowledge such as atmospheric physics, solar instrumentation, machine learning, forecasting theory and remote sensing.

Likewise, nascent methods for producing physical mappings of regions that exhibit similar climatological attributes have also been proposed in the literature. Understanding of attributes, such as solar resource variability, associated with different regions would enable system operators to improve decisions on unit dispatch by increasing the confidence of unit commitment operations, predicting intra-hour dispatch and reducing automatic generation control errors. Diabate et al. (2004) performed a clustering analysis of the monthly means of the daily clearness indexes of 62 stations and defined 20

coherent solar radiation regions in Africa. Kabalci (2011) developed an intelligent analysis tool based on an agglomerative hierarchical clustering approach to provide a mapping of solar radiation parameters of the Central Anatolian Region of Turkey. Similarly, Zagouras et al. (2013) proposed a novel method for the optimal placement of radiometric stations over the country of Greece, where a cluster analysis of data representing the biannual fluctuation of cloud modification factors was performed to conclude that 22 ground-based instruments could efficiently track the surface solar irradiance over the observed area.

More recently, Zagouras et al. (2014a) presented a cluster analysis for the determination of coherent microclimates of Global Horizontal Irradiance (GHI) for a utility scale territory in California, which is serviced by San Diego Gas & Electric. These methods were also applied on Lanai, Hawaii, in order to explore grids that operate in island mode (Zagouras et al., 2014b). In both cases, the identification of coherent solar microclimates would allow utilities and power plants to realize cost savings through regional planning and operation activities; e.g., mitigation of solar power variability through the intelligent placement of solar farms and increases in forecast quality through the optimal placement of radiometric stations.

In addition to offsetting weather-dependent solar resource availability, short-term load forecast (STLF) also plays a key role in the operation, control, and management of utility scale electrical grids. Depending on the time horizon, each forecast has a specific application for grid balancing and scheduling. Day-ahead market (DAM) load forecasts are required by Independent System Operators (ISOs), utilities, and electricity market participants for operation planning and unit commitment of generating plants. Hour-ahead market (HAM) load forecasts are needed for real-time control and load following. To compensate for the uncertainty in load forecasts, ISO maintains an operating reserve (WECC, 2010), increasing costs for utilities and customers (Ortega-Vazquez & Kirschen, 2006). Continuous research is being conducted to lower the uncertainty and increase the accuracy of load forecasts. Various methodologies based on time-series, regression, fuzzy logic, Artificial Neural Networks (ANNs), expert systems, hybrid models, etc., have been proposed (Suganthi & Samuel, 2012). Most of these methods rely on external inputs such as meteorological forecasts, temperature, dew point, etc., (Matthewmann & Nicholosn, 1968; Hippert et al., 2001). Since no forecast is completely accurate, the errors in meteorological forecasts are propagated through the load forecasts (Matthewmann & Nicholosn, 1968). As a result, studying the forecast error distribution is essential to reveal the error structure, which, if systematic and non-random, can easily be corrected leading to a better prediction.

All three aspects discussed above are being fully addressed to accelerate the timeline and improve the quality of the current project through: (1) the use of a novel clustering algorithm for the determination of coherent zones of GHI for utility scale service territories; (2) the production of HAM/DAM solar

irradiance forecasts for the cluster centers using current state-of-the-art; and (3) the demonstration that with the proper re-forecasting and ensemble techniques it is possible to substantially improve utility load forecasts. In this project we implement an operational tool to quantify the combined Variability and Uncertainty (V&U) of photovoltaic (PV) generation and load, and determine system capability to meet reliability requirements and optimal resource dispatch strategies for systems with high PV penetration. Objectives of the effort are to reduce power system operation costs by committing appropriate amounts of energy resources and reserves, and to provide operators a prediction of the generation fleet's behavior in real time for possible scenarios.

Introduction

This project addresses one of the remaining technological bottlenecks for widespread integration of variable resources into the grid. Specifically, our team will demonstrate, in a fully replicable way, how a utility company can benefit from accurate solar resourcing and forecasting in order to reduce operational costs associated with balancing area (BA) intra-hour dispatch and automatic generator control (AGC).

This project brings together, for the first time, proven methodologies for generating state-of-the-art solar and load forecasts to the operational planning, dispatching and trading desks of a utility company, with the objective of reducing operational costs associated with high PV penetration. The participating utility, Sacramento Municipal Utility District (SMUD), provides electricity to a population of 1.4 million spread over a 900 square mile service area, which makes it the nation's sixth-largest community-owned electric service provider. This project is transformational as it implements smart decision desk tools based on integrated solar-load forecast uncertainties for entire utility service regions, and for multiple PV penetration scenarios, while employing the most advanced forecasting tools available. The operational tools will be applied to and tested in the SMUD service region.

Anticipated direct benefits of this project for SMUD operations include:

- Improved dispatch of generation sources, resulting in operational cost savings;
- Improved commitment of generation sources, resulting in operational cost savings;
- Improved system reliability through improved resource adequacy in the operation time scales;
- Maintenance of the current North American Electrical Reliability Corporation (NERC) reliability standards for increasing penetration levels of variable generation as measured by current levels of area control error; and
- Accommodation of increased levels of renewable resources.

The methodology for evaluating each of these benefits will be determined as follows:

I. Improved dispatch and unit commitment of generation sources, resulting in operational economic cost savings

Scenarios will be studied based on actual solar conditions and system loading conditions using the software "Electric System Intra-Hour Operation Simulator" (ESIOS) together with existing operation tools and including solar forecasting input. Generation unit commitment and dispatch for these simulated

scenarios will be compared to the actual generation dispatch, unit commitment, and corresponding production costs. The difference between the simulation production costs and the actual operational costs will be the estimated cost benefits.

II. Improved system reliability through improved resource adequacy for operations

The ESIOS model will test scenarios where actual net load (i.e. load minus solar PV generation) is higher than expected, due to less PV generation than forecasted. With advanced solar forecasting input, ESIOS integrated with existing operation tools, will simulate the same scenario and determine the required unit commitment and dispatch. This will be done for several simulation scenarios to identify any conditions where inadequate generation reserve margin may have occurred, or a lower cost alternative to ensuring generation adequacy could have been achieved. This cost differential will be the identified benefit.

III. Maintenance of Compliance Levels of Area Control Error (ACE)

Scenarios will be studied based on actual solar conditions and system loading conditions using the ESIOS model together with existing operation tools and including solar forecasting input. The amount of regulation required to maintain the currently low ACE levels will be calculated by ESIOS and compared to the actual regulation required to maintain ACE for those specific scenarios. The differential in cost between actual cost of regulation and the results of simulations yield the identified benefit. Maintaining NERC reliability levels as measured by current ACE values at different penetration of variable generation (VG) is likely to be the most effective way to demonstrate the value of this project to SMUD and other utilities. Many utilities, such as SMUD, have a high interest in AGC algorithm development to move from a ”responsive” to a ”predictive” control strategy. Currently, SMUD maintains low ACE by responding to changes in system dynamics; however, as the penetration of PV increases, it will become critical that system operators and AGC make decisions for dispatching units in advance of moments of high variability. This project introduces solar-load forecasting along with the uncertainty into AGC, which is of high importance to SMUD operators. In order to evaluate the performance of ESIOS with solar forecasting, the system ACE will be compared with the current Energy Management System (EMS). Good performance can be measured by observing ACE behavior. It is expected that if the ESIOS with solar forecasting can improve system reliability, ACE ramps will be minimized, and ACE will be at or near zero more frequently than with existing EMSs.

IV. Accommodation of increased levels of renewable resources

The ESIOS model will simulate various levels of solar PV penetration with realistic solar forecasting uncertainties, representing future growth of PV in the SMUD control area. These same conditions will be modeled using conventional operational generation dispatch methods as currently used by SMUD. A comparison of the amount of solar PV that can be accommodated reliably (measured by same ACE levels with cases without PV) using both dispatch methods will be determined in order to assess this benefit. The SMUD implementation plan will be determined accordingly based on the performance achieved in items I-IV, and by the progress shown in Tasks 1-6 outlined below. All intermediate results from this project will be shared in real-time with SMUD operations, which will assess the benefits of incorporating the real-time information into their operational desks at different stages of the project.

A summary of the tasks within the Statement of Project Objectives (SOPO) for the project's lifetime, including relevant milestones and deliverables, follows next.

Task 1.0: Detailed Clustering Analysis of the Solar Resources in SMUD Territory

The objective of this task is to identify the number and location of regions of coherent solar and load variability within the SMUD service territory. Knowledge of the number and location of these distinct regions will help define simple algorithms for the forecast uncertainty quantification and the effect of solar variability, solar irradiance, and load variability for different scenarios of distributed and central stations penetration by region.

Task 1.0 Deliverables

D1.1 Summary report on coherent regions of solar resourcing, solar contribution to the power grid, solar variability, and load variability for SMUD territory.

D1.2 Updated cluster maps of solar variability for SMUD territory.

Task 1.0 Milestone

M1.1 Produce and make available to the public detailed clustering maps of solar resource, solar resource variability, and solar power output variability.

Task 2.0: Clean Power Research Forecasts for All Sectors of SMUD's Territory

The purpose of this task is to develop regional forecasts for the SMUD service territory (all sector plus aggregate) to be integrated with the ESIOS tool developed by Pacific Northwest National Laboratory (PNNL). Clean Power Research (CPR) will provide the PV fleet simulation software SolarAnywhere® FleetView® as well as dedicate engineering resources to help design and build an integrated solution with PNNL ESIOS. FleetView® provides utilities and PV planners the necessary simulation tools to accurately model and forecast the impacts of PV on the grid. Within this software, individual system specifications are aggregated to create fleet groupings, which can then be simulated using historical data or forecast. PV production is generated by combining satellite-derived irradiance with regionally corrected weather inputs, along with the system specifications and then simulating production using established PV output models. CPR proposes the following tasks to be completed across multiple budgeting periods. These tasks are designed to support project management throughout the whole scope of work and certain tasks may, in practice, become part of separate budgeting periods. This will depend on the rate and success of each project task.

Task 2.0 Subtasks

Subtask 2.1: CPR will provide license access to the fleet simulation software SolarAnywhere® FleetView® for historical fleet simulations. This access is supported by product and software developer application programming interface (API) documentation. The API will be the primary interface by which the PNNL ESIOS tool will access FleetView®.

Subtask 2.2: CPR will conduct a webinar or in-person tutorial for all stakeholders in the project. This presentation will walk through functional aspects of the FleetView® web user interface and API, along with a technical description of how FleetView® models and forecasts PV output.

Subtask 2.3: CPR to engage PNNL to understand how the ESIOS tool calls and processes PV simulation data, including inputs of both fleet power and variability. This task will involve the CPR engineering team understanding the product requirements of the ESIOS tool and then designing any product modifications into the FleetView® tool to best deliver information.

Subtask 2.4: CPR will take input from SMUD to clean all PV system data using metered output and define the appropriate fleet groupings.

Subtask 2.5: CPR will deploy FleetView® for the SMUD region including forecasts creation using the existing state-of-the-art. Presently, forecasts for short-term include 1-minute power forecasts out to 30-

minutes, followed by 30-minute increments. These forecasts are generated anew every 30-minutes. In the near-term, the cloud vector model approach is applied to recent satellite cloud images and is accurate out to five-hours. Beyond the five-hour time horizon, forecasts are generated using the locally tuned numerical weather prediction output provided by the National Digital Forecast Database for the United States. CPR will generate forecasts based on fleet inputs from SMUD and customize output to ESIOS based on the desired time horizons (i.e. 5-min for RTD or 24-hour plus for DAM).

Subtask 2.6: To continue to provide the most accurate forecast generation, CPR will devote engineering resources to linking PowerClerk (system of record for all SMUD systems) to FleetView®. This task will automate the addition of systems to the fleet forecast from FleetView® once they have achieved a certain installation status through PowerClerk. This task will ensure that SMUD is not receiving a static forecast based on dynamic fleet configuration. Forecast uncertainties will be generated in real-time.

Subtask 2.7: CPR already has a relationship with SMUD and UCSD, and they anticipate that a number of teleconferences and personal meetings with SMUD operations group and with UCSD researchers will be needed to clarify the most useful format of the data to be exchanged for the UCSD re-forecasts, which have the objective of reducing CPR forecast errors. CPR will work closely with UCSD to implement advanced forecasting methods into FleetView®, including real-time smart re-forecasting and cloud-seeded Numerical Weather Prediction (NWP) models developed specifically to the SMUD service region. This process will involve as many meetings as necessary to understand the technology and design a product specification that can be implemented operationally. CPR will devote engineering resources for the incorporation of UCSD advanced forecasting techniques, specific to the SMUD area.

Task 2.0 Deliverables

D2.1 One year of intra-hour forecasts up to 30 minutes for the SMUD region, including the cluster centers determined by Task 1.0, and made available to the public free of charge.

D2.2 Ramp-rate forecast uncertainty maps and statistics for whole service region.

D2.3 Day-ahead forecast up to 24-hour plus for the SMUD region, including the cluster centers determined by Task 1.0.

Task 2.0 Milestones

M2.1 Expected intra-hour forecasts skill over persistence forecast: 0.15 ($s \sim 1 - RMSE_f / RMSE_p \geq 0.15$).

M2.2 Expected day-ahead forecasts skill over persistence forecast: 0.20 ($s \sim 1 - RMSE_f / RMSE_p \geq 0.20$).

Task 3.0 Integrated Solar-Load Forecasts for SMUD

UCSD will implement, post-process, optimize and integrate operational solar-load forecasts using an ensemble methodology that has been validated both at the local (campus community, utility feeder, utility service region) and at regional California Independent System Operator (CAISO) levels. This methodology allows the level of irreducible load forecast errors, minimum Mean Average Percent Error (MAPE), and minimum Root Mean Square Error (RMSE) through ensemble forecasts that ingest minimum amounts of weather data together with load time-series. The ensemble method uses a combination of k-Nearest-Neighbor (kNN), Autoregressive Exogeneous (ARX), Autoregressive Moving Average Exogeneous (ARMAX), Box-Jenkins model (BJ), Polynomial, Nonlinear Autoregressive Exogeneous model (NARX), and Hammerstein-Wierner Non-Linear (NL) models to generate highly robust RTD/HAM and DAM load forecasts.

Task 3.0 Subtasks

Subtask 3.1 - Hour-ahead load forecasts: Implement Real Time Dispatch/Hour Ahead Market (RTD/HAM) ensemble forecast for SMUD region and for each cluster of interest given the solar penetration growth scenarios generated by SMUD.

Subtask 3.2 - Day-ahead load forecasts: Implement kNN DAM forecasts for SMUD region for each cluster of interest given the scenarios and needs stipulated by SMUD.

Task 3.0 Milestones

M3.1 Reduce SMUD RTD/HAM forecasts errors by 50% in terms of Mean Average Percent Error (MAPE).

M3.2 Reduce SMUD DAM forecasts errors by 25% in terms of MAPE.

Task 4.0: Provide DOE with Publicly Accessible Data Created from this Project

This task will collect and make available free of charge in the public domain to be easily viewed and downloaded all data that was generated as part of the scope of this project, including data generated using federal and cost-share funds, such as, but not limited to: solar irradiance and power forecasts, name of a city as a minimum location, and actual values, values of the measurements taken to assess specific

system impacts due to improved forecast accuracy, electricity grid/feeder models developed or used in this project, and/or strategic plans developed for solar integration. All data will be thoroughly checked for any errors and other quality issues corrected before making it available for download. Exceptions for redaction of proprietary information may be made on a case-by-case basis before posting the data in public domain for compliance to this task.

Task 4.0 Deliverables

D4.1: All data generated as part of the scope of this project in Budget Period 1 will be delivered in an easily viewed and downloadable format no later than 30 days after the conclusion of Budget Period 1.

Task 5.0 Optimization of CPR's forecasts using UCSD's methodologies

The purpose of this task is to incorporate real-time smart re-forecast models to the intra-hour and day-ahead CPR forecasts and to implement more accurate NWP models for day-ahead forecasting in order to reduce forecast errors. Accuracy is determined by using both standard error metrics between forecasted and measured values (MBE, RSME, MAPE) and a forecasting skill metric developed by UCSD. Overall improvement in all error metrics is the goal. This information will be reported to DOE in quarterly progress reports.

Task 5.0 Subtasks

Subtask 5.1 - Real-time re-forecast intra-hour model implementation based on the first year historical intra-hour forecasts: UCSD will implement real-time forecast engines based on ensemble stochastic learning to optimize the error reduction process.

Subtask 5.2 - Real-time re-forecast day-ahead model implementation based on the first year of historical day ahead forecasts: Coimbra's group at UCSD will implement real-time forecast engines based on ensemble stochastic learning to optimize the error reduction process. A comparative analysis of these methods will be reported to DOE in the regular quarterly progress report.

Task 5.0 Milestones

M5.1 Improvement of forecasting skill compared to original CPR intra-hour forecasts for cluster centers by 0.15 (25% over CPR, 35% over persistence).

M5.2 Improvement of forecasting skill compared to original CPR intra-day forecasts for cluster centers by 0.25 (25% over CPR, 50% over persistence).

Task 5 Deliverables

D5.1 Real-time smart forecasts for cluster centers.

D5.2 Report on smart re-forecast improvements over intra-hour and intra-day/NWP models providing a comparative analysis of different methods (Pros/Cons) conditions of result reliability.

Task 6.0 Solar-Load Uncertainty Implementation into ESIOS and PROMOD

PNNL, working closely with UCSD, CPR and SMUD will implement an operation tool to determine system capability to meet reliability requirements and optimal resource dispatch strategies for systems with high PV penetration. Objectives of the effort are to reduce power system operation cost by committing appropriate amounts of energy resources and reserves, as well as to provide operators a prediction of the generation fleet's behavior in real time for possible scenarios. This operational tool will be tested in the SMUD service region.

Task 6.0 Subtasks

Subtask 6.1: Data collection and building SMUD system model in ESIOS, including resources to meet SMUD load, operation practice for balancing generation and load, automatic generation control (AGC) algorithm, and control performance criteria. Operation characteristics of resources including base-load and AGC generators, applicable market contracts, demand response resources and storage devices (if applicable) will be modeled. (Transmission network model generally is not required; however, specific congestion areas can be monitored and managed by coordinating ESIOS with an existing power flow solver if so desired.)

Subtask 6.2: Incorporation of V&U quantifications and forecasts in ESIOS: V&U quantifications and forecasts will create stochastic simulation cases for the next few operation hours to be simulated by ESIOS. An individual module will be developed to implement this function to work with the existing ESIOS model.

Subtask 6.3: Study the impact of V&U on day-ahead unit commitment (UC): UC runs with and without PV forecast data will be performed and compared to determine the difference in generation cost on

conventional generators. This task will need to leverage a parallel effort at SMUD to study PV integration cost and UC model for SMUD is assumed available.

Subtask 6.4: Integrate UC results with ESIOS to perform intra-hour dispatch and AGC simulation for SMUD, and evaluate the impact of PV generation V&U on BA control performance and AGC units operation.

Subtask 6.5: Work with SMUD to develop and test the use of ESIOS to perform stochastic simulations on intra-hour dispatch and AGC for the next few operation hours based on combined PV and load forecasts, verifying in real time the generation fleet's capability to handle most of the possible scenarios.

Subtask 6.6 Work with SMUD to evaluate the value of real-time ESIOS simulation in reducing operator anxiety and improving optimality of operator dispatch decisions.

Task 6 Milestones

Milestone M6.1 Complete the offline ESIOS model for SMUD, which is capable of performing intra-hour dispatch and AGC simulations for a certain future period.

Milestone M6.2 Complete an analysis on PV impacts on SMUD intra-hour dispatch and control performance for a projected PV penetration scenario using ESIOS to simulate different penetration levels for each cluster region determined by Task 1. The metric to analyze impact is maintenance of NERC reliability standard as measured by ACE values as compared to current values at low VG levels.

Milestone M6.3 Complete the ESIOS model capable of online stochastic simulations on intra- hour dispatch and AGC if links to SMUD energy management system (EMS) are established to provide real-time generator dispatch data and combined PV and load forecast. Report on the cost benefit analysis of the forecasting service for intra-hour dispatch.

Task 6 Deliverables

D6.1 An ESIOS model for SMUD system to perform offline studies on future scenarios that allow SMUD to maintain NERC reliability at different levels of VG penetration as measured by ACE values (the goal is to keep current levels of ACE, which are low due to low penetration).

D6.2 An ESIOS model for SMUD system that is capable of online stochastic simulations on intra-hour dispatch and AGC, once links to SMUD EMS are established.

D6.3 To allow for replication of these results to other utilities, the uncertainty quantification values (forecast error bands for different levels of penetration and variability) will be programmed into the ESIOS (and possibly PROMOD) software that are used by the majority of utilities nationwide. Note that this project will use CPR's FleetView® and SolarAnywhere® products, but alternative intra-hour and 30-

300 minute forecast services based on satellite image processing could be used for the same purposes by other utilities without the need to license CPR's products (see, e.g., references 6, 10-15 in the technical volume).

Task 7: Provide DOE with Public Accessible Data Created from this Project

This task will collect and make available in the public domain free of charge to be easily viewed and downloaded all data that was generated as part of the scope of this project, including data generated using federal and cost-share funds, such as, but not limited to, solar irradiance and power forecasts, name of city as a minimum location, and actual values, values of the measurements taken to assess specific system impacts due to improved forecast accuracy, electricity grid/feeder models developed or used in this project, strategic plans developed for solar integration. All data will be thoroughly checked for any errors and other quality issues corrected before making it available for download. Exceptions for redaction of proprietary information may be made on a case-by-case basis before posting the data in public domain for compliance to this task.

Task 7.0 Deliverables

Deliverable D7.1: All data that was generated as part of the scope of this project in Budget Period 1 in easily viewable and downloadable format no later than 30 days after the conclusion of Budget Period 2.

Dissemination, Replication and Impact of EERE funding

Lessons learned from data sharing and the simulations of high-renewable penetration scenarios will be shared with the broader power system operators' stakeholder group by coordinating with DOE's other parallel activities as appropriate and through SEPA and UVIG. The principal investigators and utility partners are regular attendees and speakers at SEPA webinars and UVIG forecasting workshops. Through this EERE funded project, power system operators will become familiar with and anticipate the various issues that arise under high solar penetration, and be able to test ways to handle them, and develop and incorporate mitigating tools into the utility power system operations. We anticipate these outcomes to be fully replicable to other utilities. All real-time forecasts will be available for viewing and down-loading free of charge to any interested stakeholders through a public web portal (http://coimbra.ucsd.edu/projects/sunrise-ucsd/index.html) managed and maintained by UCSD. Lessons learned and best practices will be published in one or more peer review journal(s) during each budget period.

Project Results and Discussion

Task 1.0: Detailed Clustering Analysis of the Solar Resources in SMUD Territory

A novel cluster analysis for the determination of coherent zones of GHI for a utility scale territory as developed by UCSD (Zagouras et al., 2014a) was employed in this project. The clustering method was improved with the addition of a variability index and an original method to evaluate the inter/intra-cluster correlations for each partition by satisfying two criteria: the maximization of the intra-cluster correlation agreements and the intra-cluster disagreements. An efficient feature extraction technique was applied to satellite-derived solar data yielding a set of vectors representing the average daily clear-sky index and average daily variability over the gridded domain of interest. In order to obtain different partitions for a varying number of clusters, the k-means clustering algorithm as developed by MacQueen (1967) was used in this project. A deterministic initialization method of seed centers was utilized to address the fundamental drawback of k-means, which is its solution instability.

Clustering validation was performed by the computation of three internal validity indices in order to investigate the number of clusters that best captured the cohesion and separation of the clustering partition with respect to the parameterization of the variability distribution problem. The appropriate number of clusters was estimated by a simple but efficient graphical method, known as the L-method (Salvador and Chan, 2004). This work is summarized in the following sections.

Data and Preprocessing

Satellite data

Irradiance data derived from satellite images was used to investigate coherent clusters of similar broadband GHI while maintaining a uniform spatial discretization over the SMUD service area. In particular, we used GHI data from the SolarAnywhere® Enhanced Resolution dataset for 2009 through 2013. This dataset consists of GHI derived from the semi-empirical SUNY model that extracts global and direct irradiances from the visible channel of geostationary weather satellites (Perez et al., 2002). The spatial and temporal resolutions of the dataset are 0.01°× 0.01° (~1 km × 1 km) and 30 minutes, respectively. The spatial domain of interests covers the SMUD service territory between 38.2°-38.8°N and 121.0°-121.7°W (Figure 1).

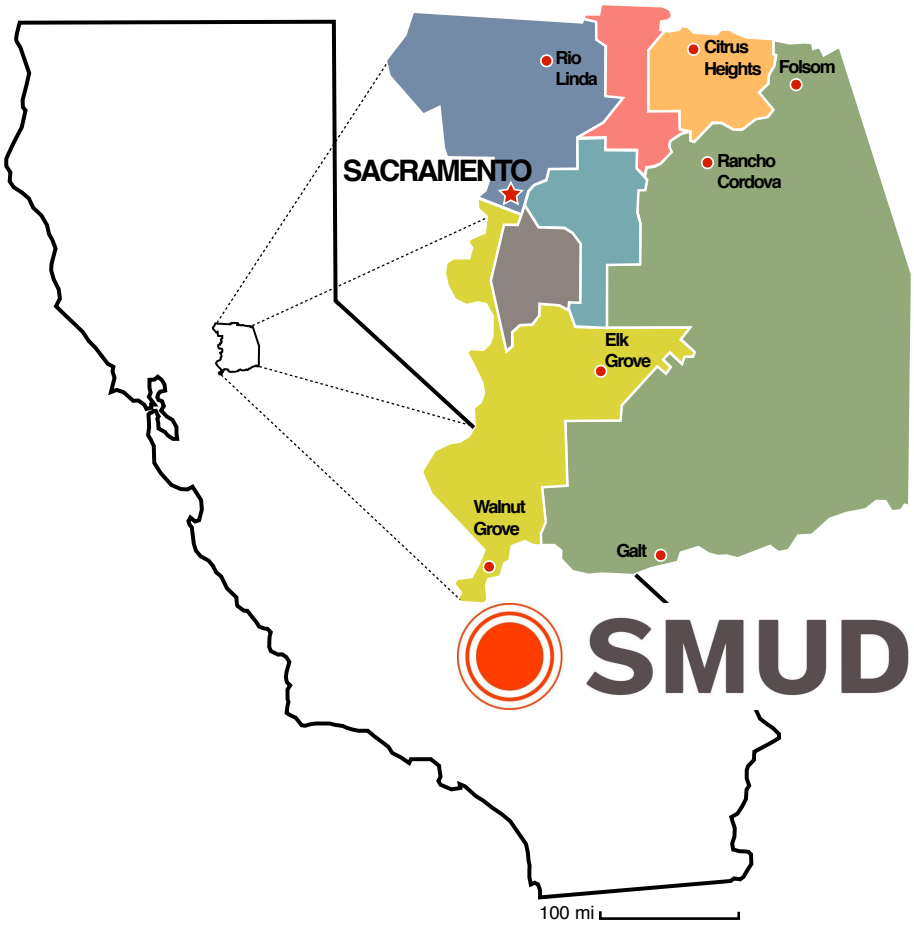

Figure 1: Map showing the state of California as well as the SMUD service territory between 38.2°-38.8°N and 121.0°-121.7°W, which supplies power to a population of 1.4 million spread over a 900 square mile service area making it the nation's sixth largest community owned electric service provider.

Clear Sky Model

Clear Sky Models (CSMs), which estimate ground level irradiance in the absence of clouds, were employed in several areas of solar engineering including forecasts, establishing forecasting skill metrics, normalizing satellite data, and defining clear sky indices (Inman et al., 2013). While a number of broadband GHI CSMs exist in the literature, it has been shown that it is not the model itself, but the quality of the input parameters (most notably turbidity) that have the highest influence on GHI CSM accuracy (Ineichen, 2006; Reno et al., 2012). Rather than employ a multi-parameter CSM, which can require up to eight atmospheric inputs, in this project we used the bulk-parameter CSM developed by Ineichen and Perez (2002) that requires only the Linke turbidity coefficient as an input. This CSM is based on an airmass independent formulation of the Linke turbidity factor that has the advantages of being independent of zenith angle and matching the original Linke turbidity factor at an airmass of 2. In

addition, it is well known that the local Linke turbidity factor varies in space and time with fluctuating concentrations of aerosol particles suspended in the atmosphere (Eltbaakh et al., 2012). To address this, the CSM references Linke turbidity maps of the world for each month, developed by Remund et al. (2003) using a combination of ground measurements and satellite data.

As discussed in Zagouras et al. (2014a), it is important to note that this project investigated the step changes of the average daily clear-sky index at each node in the domain in order to create coherent clusters. In other words, the vectors that were being compared are temporal vectors at a fixed location. To this end, the uncertainties in the Linke turbidity factor at neighboring locations do not influence the nature of the vector elements at a given location. In addition, any bias introduced through uncertainties in the Linke turbidity at a specific location was removed through the examination of the step changes of the variability index during the construction of the temporal vectors.

Clear Sky Index

In order to remove variability associated with deterministic diurnal/seasonal solar cycles, the GHI time-series was first normalized with respect to the CSM described in the previous section. The normalized GHI, or clear-sky index $k_c(t)$ as it is more commonly known, is defined as:

$$k_c(t) = \frac{G_h(t)}{G_{hc}(t)} \quad (1)$$

where $G_h(t)$ is the modeled GHI at time t, $G_{hc}(t)$ is the modeled clear-sky GHI at time t, and the dimensionless quantity $k_c(t)$ varies between 0 and 1.

Variability of the Solar Resource

As a first step toward coherent GHI clustering, daily solar resource was considered in this project. The 30-min data derived from SolarAnywhere® was preprocessed and compiled into daily parameters at each pixel. In order to accomplish this, $k_c(t)$ frames corresponding to each day were averaged at each location yielding vectors of average daily clearness. Note that prior to the normalization, integration of daily GHI values over a day is equivalent to the total daily energy per pixel area (Jm^{-2}), which reflects the node's relative potential for energy production. After normalization the integral represents the daily average clear-sky index ρ for a given pixel, which is a dimensionless parameter that varies between 0 and 1. The extracted feature, denoted as ρ, for a daily time course of $k_c(t)$ defined on a closed interval $[0, T_\gamma]$ was approximated by the trapezoidal rule:

$$\rho(D) = \frac{1}{T_\gamma}\int_{t=0}^{T_\gamma} k_c(t)dt \approx \frac{1}{\gamma}\left[k_c(0) + 2\sum_{i=1}^{\gamma-1} k_c(i) + k_c(\gamma)\right] \quad (2)$$

where T_γ is the total daylight time (s) of the γ frames for each day, which varies with the day/season (more in summer, fewer in winter). Likewise, an average hourly clear sky index was developed in order to investigate intra-daily variability as:

$$\rho_h(H) = \frac{1}{2}\int_{t=0}^{T_h} k_c(t)dt \approx T_h k_c(0) + \frac{1}{2}T_h\left[k_c(T_h) - j_c(0)\right] \quad (3)$$

where T_h represents two consecutive non-overlapping 30-minute time-stamps in order to form hourly features. These definitions of $\rho(D)$ and $\rho_h(H)$ have several benefits including, being dimensionless, eliminating inconsistencies in the length of days, and being independent of deterministic fluctuations.

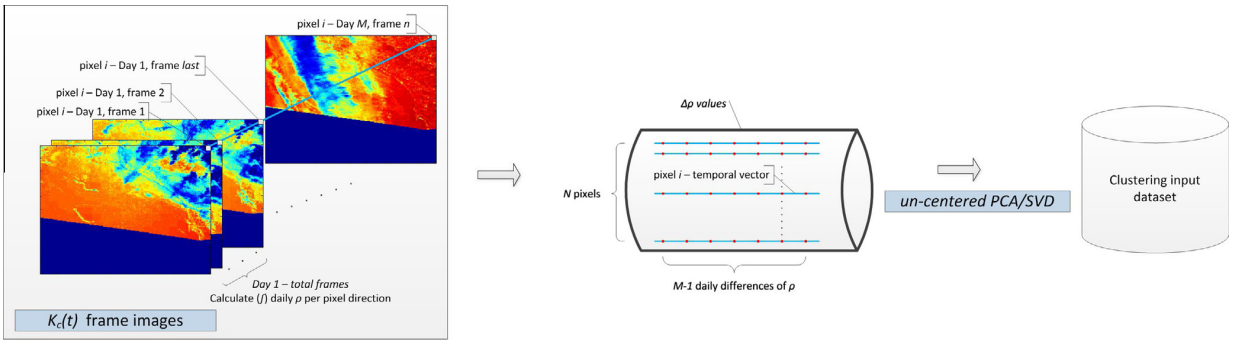

Figure 2: An example of how the concatenated pixels over a specific location are used to calculate the daily $\rho(D)$ value is illustrated (left). All the temporal vectors (blue lines) that represent the M-1 daily differences of $\rho(D)$ (red dots) for each of the N pixel locations, which are then stacked to create the V features dataset (middle). A dimensionality reduction method (un-centered PCA/SVD) proceeds to provide the clustering input dataset (right).

Rather than using the temporal vectors of $\rho(D)$ or $\rho_h(H)$ in the cluster analysis, a final post processing was applied to investigate the variability of clearness at each pixel. To this end, we defined a new variability measure of the daily average clear sky index as:

$$V(D) = \frac{1}{N-1}\sum_{N}|\Delta k_c(t)| \quad (4)$$

and,

$$V_h(H) = |\Delta k_c(t)| \quad (5)$$

where $|\Delta k_c(t)| = k_c(t+H) - k_c(t)$. We expect to have low values of $V(D)$ for consecutive similar cloud-conditions and high values for a sequence of days that vary between clear and cloudy conditions. Similarly, we expect to have low values of $V_h(H)$ for days that are consistently clear or overcast and high

values when a day transitions from clear to cloudy. With this preprocessing technique the original dataset was transformed such that each element of a vector represents the consecutive daily or hourly absolute step-changes of $\rho(D)$ and $\rho_h(H)$ for each node in the domain of interest. Figure 2 depicts the described feature extraction stage, where the N pixels were used to construct N temporal vectors that represent the course of $V(D)$ and $V_h(H)$ at each node's location. After this process, and before the clustering analysis began, we also applied a dimensionality reduction algorithm, which is described in the following section.

Dimensionality Reduction

The high dimensionality of the temporal vectors requires a dimensionality reduction in order to lower the computational complexity and restrict the noise effect of the data. For that purpose we employ the most widely used linear dimensionality reduction method, the Principal Component Analysis (PCA) (Jolliffe, 1986). Ding and He (2004) indicated that the dimensional reduction via PCA is directly related to unsupervised clustering as a result of the continuous solutions of the discrete cluster membership indicators in k-means clustering actually being the principal eigenvectors given by the PCA. Although the centering of the data matrix to the origin (i.e. column mean subtracting) is essential before applying the classical PCA, many researchers debate this demand and investigate the differences between the centered and un-centered PCA (Cadima and Jolliffe, 2009).

Centered PCA subtracts the mean from each column, which is equivalent to translating the original high-dimensional axes to the mean of the coordinates. In this case the PCA relates the variability of the differences from means about the incipient origin. Additionally, the centering deprives the physical interpretation of the original data by representing unsigned physical data with negative values. On the contrary, the un-centered PCA provides a low-dimensional coordinate system by the Principal Components (PCs), which is explained by the variance about the center of gravity of the dataset (Cadima and Jolliffe, 2009). Considering the advantages resulting from the un-centered scheme such as (i) the untransformed representation of the data structure in the feature space in terms of discontinuous data (Noy-Meir, 1973) and (ii) the purposeful understanding of the time dependencies, we choose not to apply a column-based mean subtracting prior to the PCA. In this project, the PCA via Singular Value Decomposition (SVD) projects the initial high dimensional un-centered data into the best low-dimensional linear approximation in such a manner that 99% of the initial variance of the data is preserved.

In addition to the approach described above, we examined the impact of variable dimensionality of the feature space to the clustering partitions. That is accomplished by applying the dimensionality

reduction algorithm to various dimensions in the range 98% ± 2%. Thus, the clustering is performed at each of the multiple reduced feature spaces and the L-method finally determines the appropriate number of classes (NC) for each dimensionality.

Clustering Algorithm

The k-means clustering algorithm is one of the most widely used and simple methods of unsupervised learning (Jain, 2010; Wu et al., 2008). The reasons for its popularity rely primarily on its scalability and simplicity. On the other hand, the algorithm also suffers from a number of limitations. Primarily, it considers the underlying structure of the data as hyper-spherical, owing to the typical selection of the Euclidean distance as the primary clustering criteria. For this reason, k-means partitions may be fallacious for dataset structures composed by non-hyper-spherical shapes.

In addition, the requirement to define a priori the number of *k* clusters can also be considered as a primal handicap. In order to address this issue, the default iterative refinement algorithm (Lloyd, 1982) of *k*-means uniformly chooses a random number of *k* points as the initial centers of the desired *k* clusters, where each point of the dataset is assigned to its closest center. Subsequently, the position of the *k* centers is iteratively optimized by the minimization of the distance criteria between the points of a cluster to its center. The algorithm stops either after a predefined number of iterations or when a convergence threshold value of a criterion function is reached. Hence, it is obvious that the effectiveness of *k*-means depends on how close the initial centers are to the final partition. The initialization of different seed centers generates divergent final clustering solutions. In addition, the risk of convergence to local minima of the criterion distance is high.

Initialization method

In order to achieve a stable solution, several heuristic algorithms have been proposed. Celebi et al. (2013) presented a comprehensive survey along with a comparative study of *k*-means initialization methods. These methods are mainly distributed by their complexity and their deterministic or non-deterministic heuristic approach to select the initial centers. In this project, we present a deterministic initialization scheme that provides stable seed centers with respect to a structural parameter, as described below. A pseudo-code for the presented initialization scheme is shown in the Appendix of Zagouras et al. (2014a).

The first step is to initialize a set of candidate initial centers from the entire dataset based on the reverse nearest neighbor (RNN) search as proposed by Xu et al. (2009). According to this scheme, the nearest neighbor (NN) of each point is calculated using the Euclidean distance and then the points are sorted in descending order according to how many times each one was selected as another points NN, i.e.

according to the number of RNNs. This concept is based on the assumption that a point assigned as a NN multiple times will have a dense distribution of points around it, thus it is reasonable to consider it as a cluster center. Based on this method it is likely that close neighbor centers occur, especially in regions with high density. To avoid extremely close candidate centers a threshold distance is applied upon which the first *m* points are selected as the set of candidate centers *S*. An empirical rule to define the threshold value based on a "reasonable portion" of the maximum distance that occurs in the dataset is employed (i.e. not finer than the spatial resolution, and not greater than the order of magnitude of the distance between the farthest points).

After the determination of the initial set of potential centroids $S_{(m)}$, the algorithm extracts the final cluster centers. If we let the number of candidate centroids be *m*, the remaining *N*-*m* points are assigned to them according to the minimum Euclidean distance, thus creating clusters of data. We select the first *n* clusters with the maximum number of members as initial centroids. This approach is based on the *maxmin* algorithm (Theodoridis and Koutroumbas, 2001). It is expected that this approach provides more reasonable clusters than the first rough selection thus yielding more robust and dense clusters. In this study the number of the first *m* centers is empirically chosen as the threefold of the desired number of *n* initial cluster centers. An example of how the initial cluster centers are posed on a 3D projection of the dataset is illustrated in Figure 3.

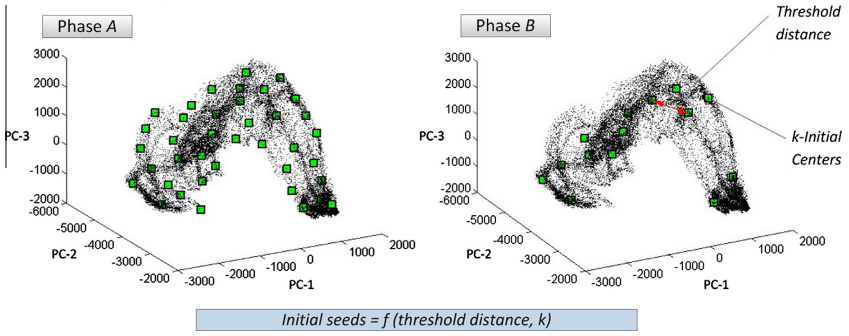

Figure 3: Example of the three-dimensional plots of the dataset (black dots) used in the experiments for the Principal Component (PC) axis. Each of the black dots represents the 3-component feature vector at every spatial location (pixel) of the service area as lying in the direction/subspace that corresponds to the three first PCs. On the left, the first $m=3k$ selected centers at the first step of the initialization method are shown as green square symbols. On the right, a selection of *k* centers spaced at least by a predefined distance that are finally used as the cluster centers. Different initial seeds are acquired by applying different threshold distances and *k* number of centers.

Internal Clustering Validity Indices

The criteria used to estimate how well a proposed clustering fits the structure underlying the partitioned data are called cluster validity indices. In the case that no correct or known partition is

available, the clustering validation is achieved by estimating internal measures of the data such as the compactness and the inter separation of the clusters. These types of criteria are known as internal cluster validity indices. Milligan and Cooper (1985) compared 30 validity indices that existed by that time and that work constitutes an important reference in cluster analysis. Recently, Arbelaitz et al. (2013) published an extensive comparative study of popular cluster validity indices in different experimental configurations and suggested guidelines to select the most suitable for any particular environment. Typically these indices are based on the computation of the intra-cluster cohesion and the inter-cluster separation. The indices can then estimate the partitions quality in terms of different variations of ratio-type and summation-type factors. These factors are commonly related to the geometrical or statistical properties of the clusters, the similarity or dissimilarity between the data, the number of partitioned data, and/or the number of clusters.

Indeed, a good clustering is equivalent to close quantifiable distances among the member points of a cluster and, at the same time, high distinction from the points of other clusters. Most of these methods tend to consider the clusters as hyper-spherical shapes providing the cluster centers as a benchmark for the measurement of compactness and separation. The performance of the internal validity indices is related to factors such as the type of datasets and which clustering algorithms are used (Arbelaitz et al., 2013), and their validation properties depend on aspects such as the monotonicity, the noise and the density of the data, the impact of sub-clusters or the unequal size of clusters (Liu et al., 2010). Taking into account the properties of the most frequently cited internal validity measures the Caliński and Harabasz (1974), the Silhouette (Rousseeuw, 1987), and the Davies and Bouldin (1979) indices are employed as the most suitable for this study among others. The following subsections provide outlines of the above validity indices.

Caliński-Harabasz

A commonly used validity index is provided by the Caliński-Harabasz (CH) method. The index is based on the cluster center and likewise the overall data centroid to estimate separation and compactness. The distance from all the members of any cluster to the respective cluster centroid $\bar{x}_k$ defines the cohesion of the clusters. The separation is calculated as the distance between each cluster centroid to the centroid $\bar{x}$ of all the data. The ratio-type relationship is described by the following equation:

$$\mathrm{CH} = \frac{\mathrm{Tr}(\mathbf{B})}{\mathrm{Tr}(\mathbf{W})} \times \frac{N-n}{n-1} \tag{6}$$

where Tr(·), **B** and **W** denote the trace, “between cluster scatter” matrices, and “within cluster scatter” matrices respectively. They are defined as:

$$\mathbf{B} = \sum_{\mathbf{c_k} \in \mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c_n}} \mathbf{N_k} (\overline{\mathbf{x}}_\mathbf{k} - \overline{\mathbf{x}})(\overline{\mathbf{x}}_\mathbf{k} - \overline{\mathbf{x}})^\mathbf{T} \tag{7}$$

where $\mathbf{N_k}$ is the number of points grouped to cluster c_k, and

$$\mathbf{W} = \sum_{\mathbf{c_k} \in \mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c_n}} \sum_{\mathbf{x} \in \mathbf{c_k}} (\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^\mathbf{T} \tag{8}$$

Thus, maximization of this index indicates the highest quality partitioning. The term (N−n)/(n−1) serves as a normalizing factor to optimize the CH score with respect to the increase of the number of *n* clusters.

Silhouette

The Silhouette index (SIL) for each point provides a score based on the relation of the point to all the points in its respective cluster compared to points in other clusters. Unlike other methods the SIL is a normalized summation-type index where the compactness is estimated by the average distance a_{x_i} between each point x_i of a cluster c_k to every other point x_i in the cluster. The separation is measured based on the minimum value b_{x_i} of the average distances between that point and the points belonging to all other clusters $\{c_1, c_2, \ldots, c_n/c_k\}$. The difference of these terms, normalized by their maximum value, results in a Silhouette value of the *i*th point that ranges from -1 to 1. The overall Silhouette index for the clustering of *N* points of a dataset is given by:

$$\mathrm{SIL} = \sum_{c_k \in c_1, c_2, \cdots, c_n} \sum_{x_i \in c_k} \frac{b_{x_i} - a_{x_i}}{\max(b_{x_i}, a_{x_i})} \tag{9}$$

A Silhouette value that approaches unity signifies that the *i*th point belongs to a coherent cluster (small a_{x_i}) and is located at a maximal distance from its nearest cluster (large b_{x_i}), resulting in tightly clustered data in well-separated groups.

Davies-Bouldin

The Davies-Bouldin (DB) index is based on the ratio of inter-group and intra-group distances. Assuming that a dataset is classified into *n* clusters, where c_i is defined as the cluster center of each of the clusters. The DB index is defined as:

$$\mathrm{DB} = \frac{1}{k} \sum_{i=1}^{k} \max_{i=1,\cdots,k, i \neq j} \left(\frac{d_i + d_j}{d(c_i, c_j)} \right) \tag{10}$$

where d_i is the average intra-cluster distance between all data classified into class *i* to the cluster center c_i, and $d(c_i,c_j)$ is defined as the inter-cluster distance between cluster centers c_i and c_j. According to the definition above, the DB index defines the average of similarity between each cluster and separated clusters, since it minimizes the potential similarity between each other.

L-Method

Estimating the appropriate number of clusters is one of the most ambiguous steps in cluster analysis. Increasing the numbers of desired clusters in a partitioning problem usually implies a monotonic increase or decrease of the validity index values due to the inherent trend to provide better clustering quality with the increase of clusters. The higher the number of clusters the more degrees of freedom are available for the system to assign the data points to many small and compact groups. However, this assumption is inconsistent with a clear determination of the correct number of clusters. On the other hand, a global extreme of the validity index would be ambiguous as this conflicts with the above consideration about monotonicity. Particularly for the clustering of large datasets with many parameters and an unknown number of clusters, the abundance of local minima renders the correctness of each solution subjective. Therefore, we seek the knee point of a validity index curve that corresponds to the number of clusters after which no significant change in value of the considered index occurs. A heuristic procedure to estimate the number of clusters via the gap statistic was proposed by Tibshirani et al. (2001), whereas Zhao et al. (2008) presented a graphical knee point detection method based on the Bayesian Information Criterion (BIC) curves in model-based clustering. In this project we adopt an efficient knee point detection method, called the L-method.

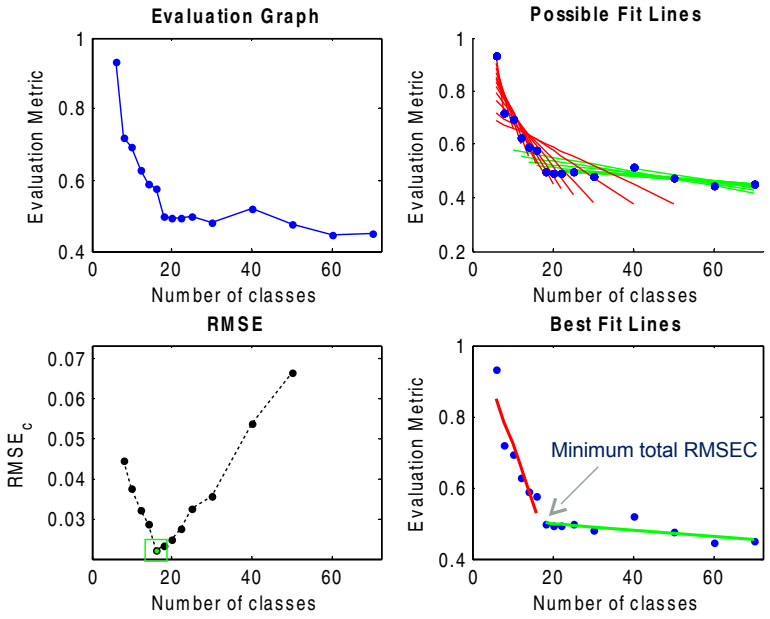

Figure 4: Determination of the appropriate number of clusters by the implementation of the L-method over the evaluation graphs of the DB cluster validity indices. This plot shows the values of the considered index (ordinate) as a function of the number of clusters (abscissa), the plot of all possible pairs of fit lines (red and green), the RMSE curve with respect to each candidate knee point (the minimum RMSE, i.e. the knee point, is marked with a green square) and the best fit lines where the point of discontinuity defines the knee point.

The *L*-method (Figure 4) is used for the location of the crucial point on the evaluation graph of any validity index. The main advantage of the *L*-method is that it does not require the execution of the clustering algorithm itself. It performs a rapid standalone procedure on the curve of an already implemented validity index graph. Unlike the model-based methods, the *L*-method detects the boundary between the pair of straight lines that best fit either side of the curve. For example, the ideal shape of the curve forms an '*L',* which implies a sharp change of the considered index values to a uniform segment. In this case, two lines are fit to the approximately linear left and right parts of the curve indicating the crucial point of discontinuity of the two lines. It should be noted that no significant change of the validity index prevails at the curve segment following the knee point indicating that the clusters are no longer discrete and they do not contribute to an appropriate partitioning. Regarding less ideal curves where a monotonically smooth decrease occurs, the point of discontinuity of the pairs of lines that best fit the underlying shape of the curve locates the point after which the curve continues more smoothly than at any other point.

The L-method can be implemented by defining an evaluation graph where the values of a validity index are on the y-axis and the number of clusters on the x-axis – Figure 4. By selecting iterative sequences of points left and right of every possible point that can be considered as a knee point, we create all the possible pairs of fitted lines on either side. The first 2 points of the curve must necessarily comprise the first left sequence of points whereas the right part contains the remaining and so forth. This method covers every possible pair of lines. A first-degree polynomial P approximates the given points of every line segment linearly. The L-method determines the appropriate pair of lines that best fit the monotonicity of the curve by minimizing the total root-mean-square error (RMSET) calculated as:

$$\mathrm{RMSE}_T = \frac{c-1}{b-1}\mathrm{RMSE}_L + \frac{b-c}{b-1}\mathrm{RMSE}_R \tag{11}$$

where c corresponds to the vertical projection of the x-axis on the point of discontinuity of the left and right lines, RMSE_L and RMSE_R are their root-mean-square error, respectively, and b is defined as the maximum number of clusters. The crucial point $\theta \in [3, b-2]$ is defined as:

$$\theta = \mathrm{argmin}_c\,[\mathrm{RMSE}_T] \tag{12}$$

Experimental Analysis

Each location's k_c time-series is transformed to the row-oriented data matrix of the variability feature $V(D)$ and $V_h(H)$, based on Eq. (4) and Eq. (5) respectively. The ith row of the data matrix represents the variability feature values among all the M days (i.e. columns) for the pixel indicated by the row index. It should be noted that each element of a feature vector is calculated by either the absolute

daily step changes of the clearness feature $\rho(D)$ or the absolute hourly step changes of $\rho_h(H)$ over a specific pixel. Therefore, the initial dimensionality is equivalent to the corresponding time period of this dataset in days or daylight hours, respectively. The variability feature space is then reduced to a lower dimensional space by performing un-centered PCA via SVD, after which, the presented stable initialization scheme for cluster centroids is carried out on the generated feature space. As previously mentioned, the initialization method requires a pre-defined parameter; the closest allowable distance between the candidate centroids. The derived initial centers obtained by the presented method tend to vary as a function of threshold distance. For the current dataset it was experimentally found that threshold distances beyond one third of the maximum length scale in the domain causes inadequate selection of initial seed points. In addition, the spatial resolution of the satellite data, combined with the scale of the spatial domain in this study, results in a reasonable upper limit of 80 clusters. Consequently, the number of clusters in this study was allowed to vary between 4 and 80. The number of clusters is identical with the number of initial centroids applied in each run. The block diagram of the proposed methodology is displayed in Figure 5.

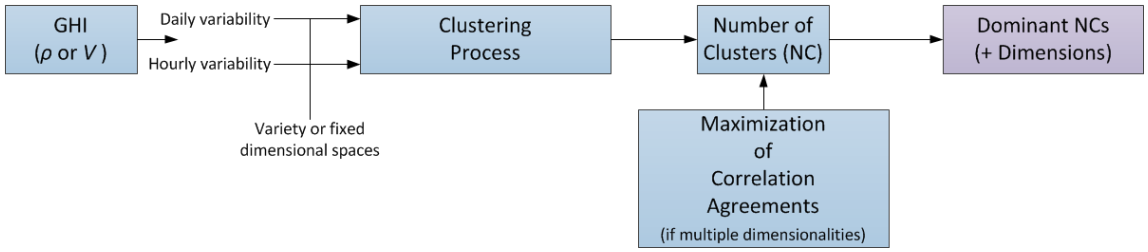

Figure 5: Block diagram of the clustering methodology.

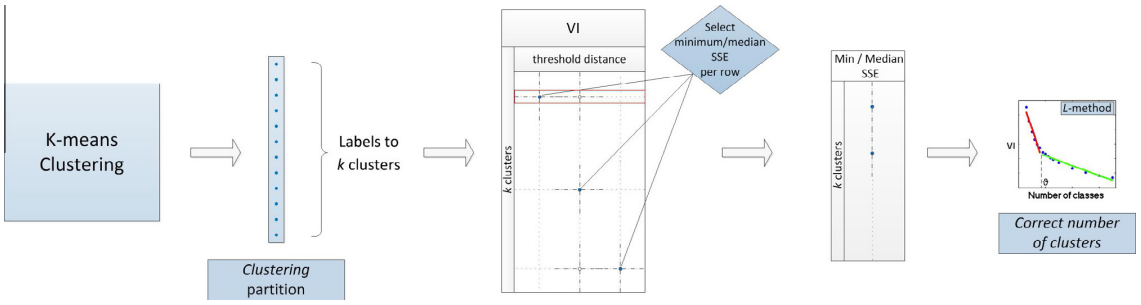

Figure 6: The *k*-means clustering algorithm provides column-based clustering partitions of *N* cluster labels for each predefined clustering parameterization, i.e. threshold distance of the initial seeds and *k* number of clusters. The VIs (CH, SIL, and DB) are calculated correspondingly and for each *k* we select the VI value from the partition that corresponds to the minimum of the SSE (blue dots) among different threshold distances (red strip row). The L-method is applied on the evaluation graph of the selected VI values with respect to *k* and determines the appropriate θ number of clusters.

Several partitions were built for each number of clusters with respect to the various threshold distances. The default criteria function that *k*-means uses to estimate the cohesion of a cluster is the sum of squared error (SSE) (Celebi et al., 2013). The minimization of the SSE determines the convergence of the *k*-means to the desired number of clusters. Although the various partitions were built to cover all the possible scenarios of threshold, no statistically significant differences occur among the runs of each *k*, which demonstrates the solution stability of *k*-means for the seed centers initialization. Furthermore, for a given initialization set it is not required to execute *k*-means multiple times since the initialization method is considered deterministic and produces an identical partition. The three validity indices CH, SIL, and DB were computed for each single partition and the L-method employed to identify the reasonable number of clusters for the overall experimentation. For more detail see Figure 6, which illustrates the process that leads to the final number of clusters.

Finally, a maximum correlation distinction analysis is conducted among the k_c time-series of the N pixels in the SMUD territory. A correlation matrix **R** is composed such that:

$$R_{ij}, \quad \text{where} \quad 1 \leq i,j \leq N \tag{13}$$

For each clustering partition we calculated the intra-correlation agreement as:

$$E_{\text{intra}} = \sum_{i=1}^{N} \frac{\sum_{j=1}^{N_{c\in i}} R_{ij}}{N_{c\in i}} \tag{14}$$

if the pixel j belongs to the same cluster as pixel i, and the inter-correlation disagreement as:

$$E_{\text{inter}} = \sum_{i=1}^{N} \frac{\sum_{k=1}^{N_{c\notin i}} R_{ik}}{N_{c\notin i}} \tag{15}$$

if pixel k belongs to a different cluster than pixel i. Finally, we selected the most appropriate clustering by choosing the clustering that maximizes the distinction:

$$\Delta E = E_{\text{intra}} - E_{\text{inter}}. \tag{16}$$

Table 1: Optimal number of clusters for a variety of cluster features with respect to the variability indices.

Feature	VI		
	CH	SIL	DB
ρ(D)	12	11	6
V(D)	8	8	6
ρ_h(H)	8	9	10
V_h(H)	8	12	10

Clustering Results

Regarding the result of the approach which used a SVD via PCA that conserved a fixed 98% of the variance of the initial dataset, the *L*-method determines the optimal number of clusters for a variety of cluster features with respect to the validity indices as shown in Table 1. The resulting clusters for the daily features are also shown in Figure 7. In general, two major conclusions arise here: a) there is a narrow range of 6-12 clusters that can be considered as the spectrum of appropriate clustering schemes for the SMUD service area; b) there is a consistent shape of the clusters regardless of the clustering scheme. Comparing the validity indices, there appears to be no significant difference between the numbers of clusters obtained.

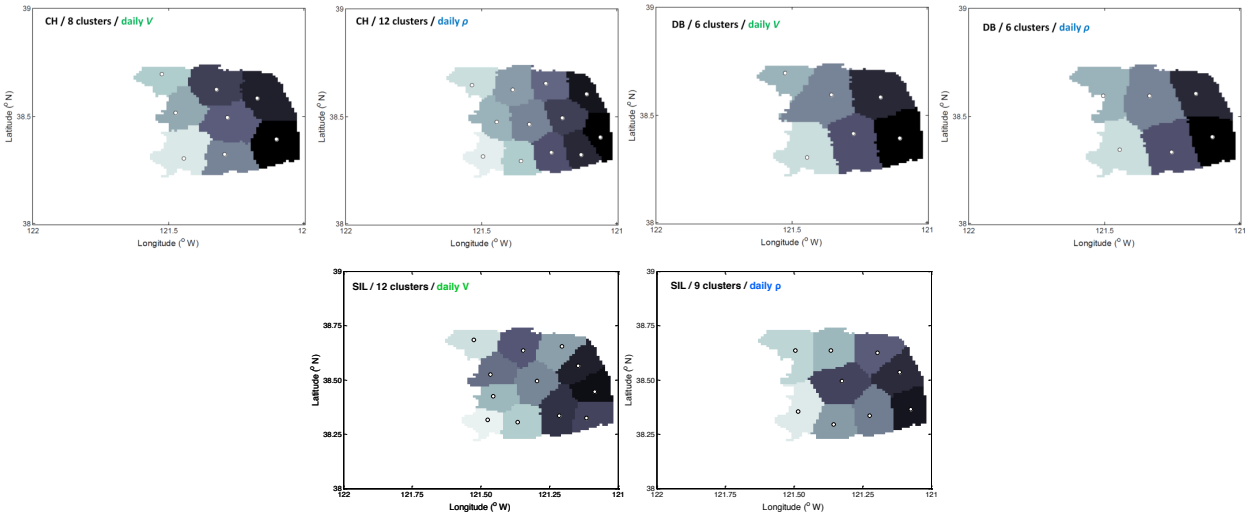

Figure 7: The resulting clusters for the daily features ($\rho(D)$ and $V(D)$) for the fixed dimensionality corresponding to 98% of the initial variance being conserved.

While the approach above is based on the assumption that the dimensionality of the feature space is reduced to those dimensions that preserve 98% of the variance of the initial data, we also performed an analysis to examine the impact of variable dimensionality of the feature space to the clustering partitions. The dimensionality was reduced to a variety of dimensions within 98% ± 2%. Thus, the clustering was performed at each of the multiple reduced feature spaces and the *L*-method determined the appropriate number of clusters for each dimensionality. While a number of results for a clustering feature and validity index occur due to the multiple dimensionalities, we selected the dimensionality that performed best based on the maximization of the distinction between the intra-inter correlation among the clusters, see Eq. (16). Table 2 shows the number of clusters that correspond to the most appropriate dimension (blue

font) with respect to the clustering features and the validity indices. This approach determined an even narrower range of clusters varying from 8 to 12. The cluster shapes confirm the previous results of consistent topology of the coherent clusters. In addition, the resulting clusters for the daily features are also shown in Figure 8. Again, a narrow range of clusters as well as a consistent shape of clusters between methods is observed.

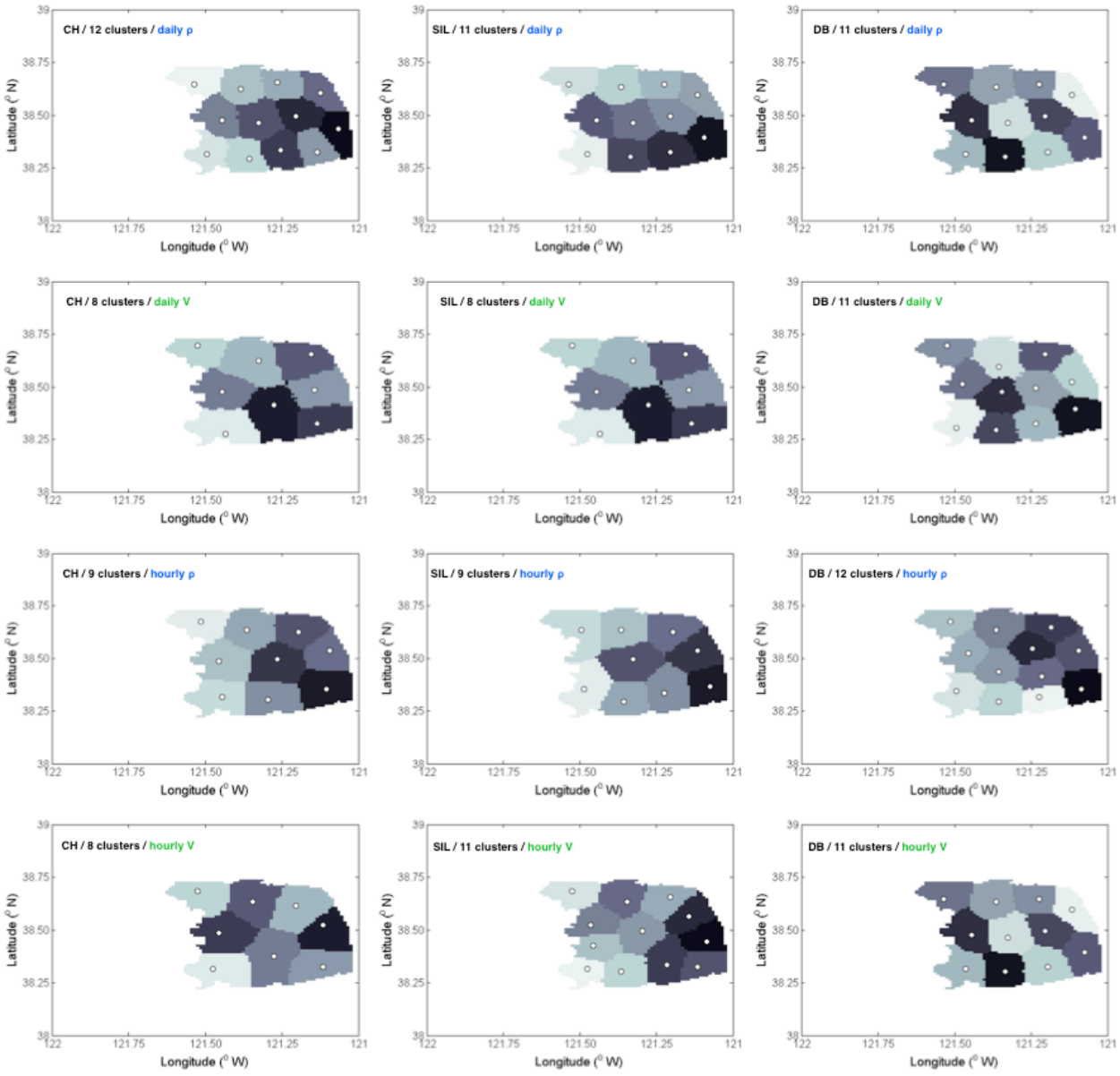

Figure 8: The resulting clusters for the daily features ($\rho(D)$ and $V(D)$) for the fixed dimensionality corresponding to 98% of the initial variance being conserved.

Table 2: Optimal number of clusters for a variety of cluster features with respect to the validity indices for a varying number of dimensionalities.

Feature	VI, Dimension					
	CH		SIL		DB	
ρ(D)	12	246	11	246	11	90
V(D)	8	98	8	98	11	267
ρ_h(H)	9	779	9	413	12	901
V_h(H)	8	779	12	1145	11	47

Task 2.0: Clean Power Research Forecasts for All Sectors of SMUD's Territory

The purpose of this task is to benchmark regional forecasts for the SMUD service territory to be integrated with the ESIOS tool developed by PNNL. CPR provided one year of historical forecasts of GHI for the 12 cluster centers determined in Task 1.0 (Detailed Clustering Analysis of the Solar Resource in SMUD Territory). These forecasts were issued in two distinct time-horizons and resolutions: Intra-Hour (IH: 30 minute horizon) and Day-Ahead (DA: 24 hour horizon), which are resolved at 1 minute and 30 minute centered averages respectively. It should be noted that both of these forecast were issued every half hour and, as a result, there is no layover in the IH forecast and 23.5 hours of layover for the DA forecasts, see Figure 9.

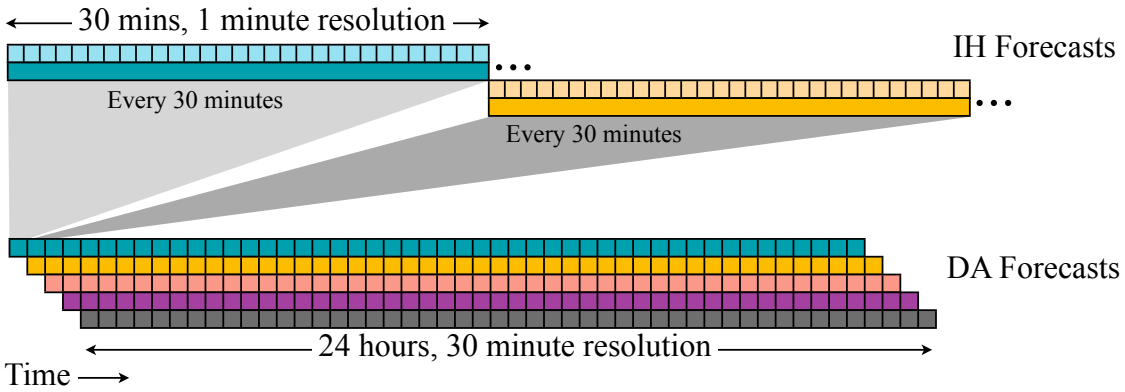

Figure 9: Illustration of the frequency with which the Intra-Hour (IH) and Day-Ahead (DA) forecasts are issued as well as an illustration of their horizons and resolutions. It should be noted that because the IH forecast horizons matched the frequency with which it is issued, there is no overlap. However, the DA forecast horizon is much longer than the frequency with which it is issued and results in nearly 98% overlap.

Data for the benchmarking

In order to benchmark CPR's irradiance forecasts, which were generated for the center of each cluster as determined by Task 1.0, ground truth was required at each of the cluster centers. This was achieved through the use of SMUD's ground based sensor network, which is deployed on the same 5 km grid as the NOAA National Digital Forecast Database (NDFD). The sensor that was located closest to the calculated cluster center from Task 1.0 was used in the benchmarking and CPR's forecasts were generated for this same location. The location (lat, lon) of the calculated cluster centers from Task 1.0, the location of the sensor that is closest to the calculated cluster center, the distance from each other, as well as the cluster number and SMUD sensor number are shown in Figure 10. It is clear from the figure that

all of the sensors were typically located within 2 km of the calculated cluster center, which was a result of the ~5 km grid on which they were deployed.

No.	Cluster Center Lat (Deg)	Cluster Center Lon (Deg)	Sensor Lat (Deg)	Sensor Lon (Deg)	Distance (km)	SMUD Sensor No.
1	38.685	-121.525	38.684796	-121.505844	1.663	71
2	38.525	-121.465	38.510128	-121.462761	1.035	37
3	38.425	-121.455	38.425040	-121.438402	0.899	22
4	38.315	-121.475	38.327528	-121.472831	0.874	7
5	38.635	-121.345	38.622223	-121.316835	1.759	61
6	38.495	-121.295	38.496066	-121.276611	0.998	49
7	38.305	-121.365	38.299355	-121.347410	1.031	20
8	38.655	-121.205	38.633020	-121.205389	2.445	62
9	38.565	-121.145	38.554294	-121.137628	1.352	47
10	38.445	-121.085	38.474108	-121.106113	3.723	24
11	38.335	-121.215	38.326817	-121.188386	2.494	11
12	38.325	-121.115	38.341042	-121.117222	1.795	74

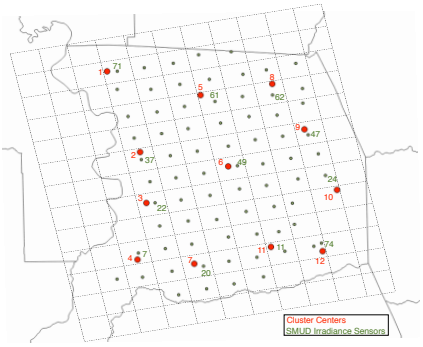

Figure 10: Illustration of the location of the calculated cluster centers from Task 1.0, the location of the sensor that is closest to the calculated cluster center, the distance from each other, as well as the cluster number and SMUD sensor number.

Methods

The benchmarking of CPR's IH and DA forecast was performed via the typical forecasting skill metric s, which is defined as the difference between solar variability and forecast uncertainty normalized with respect to the variability

$$s = \frac{V - U}{V} = 1 - \frac{U}{V} \tag{17}$$

where the variability of solar irradiance is defined as the standard deviation of the step-changes of the clear sky index

$$V = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{I_t}{I_{\mathrm{clr},t}} - \frac{I_{t-1}}{I_{\mathrm{clr},t-1}}\right)^2} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\Delta k_t)^2} \tag{18}$$

and the forecast uncertainty is defined as a version of the RMSE normalized with respect to the clear sky irradiance rather than the average irradiance

$$U = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{\hat{I}_t - I_t}{I_{\mathrm{clr},t}}\right)^2} \tag{19}$$

It is well known that the forecasting skill metric is closely approximated by the difference between the RMSE of the persistence forecast and the RMSE of the forecast being benchmarked normalized with respect to the RMSE of the persistence forecast (Inman et al., 2013),

$$s \simeq \frac{\mathrm{RMSE}_P - \mathrm{RMSE}_F}{\mathrm{RMSE}_P} = 1 - \frac{\mathrm{RMSE}_F}{\mathrm{RMSE}_P}, \tag{20}$$

which is the metric used in this work. It is clear from this definition that the forecasting skill metric s can be viewed as an improvement over the persistence model.

It is worth pausing here to note that the definition of a persistence forecast is slightly different for IH and DA forecasts. Specifically, persistence performs particularly well for short horizons and begins to

deteriorate with increasing extrapolation. As a result, for IH forecasts, persistence, as the name implies, is defined as having the clearness persist for the following forecast horizon

$$k_{t+\Delta t} = k_t = \frac{I_t}{I_{\text{clr},t}} \tag{21}$$

where Δt is the forecast horizon. Thus, the IH persistence forecast is computed as

$$I_{\text{pers},t+\Delta t} = k_t I_{\text{clr},t+\Delta t} \tag{22}$$

Day-ahead forecasts, on the other hand, use a persistence of the daily clearness. Specifically, when issuing a DA persistence forecast at a specific time of the day, the previous day's clearness is assumed to persist, which is illustrated in Figure 11 for a specific instance of the forecast. This type of DA persistence results in a more uniform performance of the persistence forecast, see Figure 11.

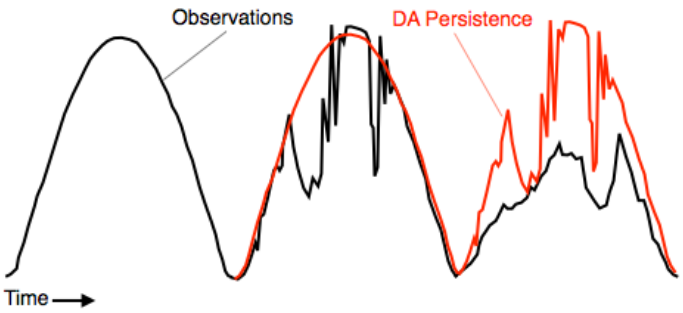

Figure 11: Illustration of the day-ahead persistence model used in the benchmarking of the CPR DA forecasts.

Results

Given the forecast at the twelve cluster centers, the ground truth at these locations, a definition of a forecasting metric and persistence forecasts, we were able to calculate the skill as a function of time horizon for both the IH and DA forecasts from CPR, which are summarized in Figure 12. Each plot shows the mean forecast skill relative to persistence as well as the RMSE from which they were calculated. All curves also include the 95% confidence interval for the twelve locations benchmarked here.

A few points should be made about these results. First of all, as expected, the IH persistence forecast exhibits a low error for short horizons after which error increases logarithmically with time. The IH forecast from CPR, on the other hand, exhibits a relatively uniform RMSE of ~80 Wm^{-2} for all time horizons. For horizons less than 10 minutes the RMSE of the persistence is less than 80 Wm^{-2} and, consequently, the forecast skill is negative for such short horizons. However, for horizons longer than 10 minutes, the persistence forecast's RMSE exceeds 80 Wm^{-2} and the CPR's forecasts begin to offer some improvement over persistence, which can be up to 15% for horizons longer than 20 minutes. Also as one would expect, rather than exhibiting the typical logarithmic growth of error associated with short term

persistence forecasts, the DA persistence forecast exhibits a strongly uniform RMSE as a function of forecast horizon. This is a result of the sliding overlap of the forecast from CPR (see Figure 9) that are issued and subsequently benchmarked and is a reflection of the inter-daily variability of the irradiance for the SMUD service area. It should be noted that the error of the DA forecast from CPR performs well for short horizons and tends to deteriorate with longer horizons, similar to the IH persistence forecast. For all time horizons the CPR search forecasts errors are less than persistence resulting in a positive improvement over persistence of up to 50% for short horizons and 30% for longer horizons.

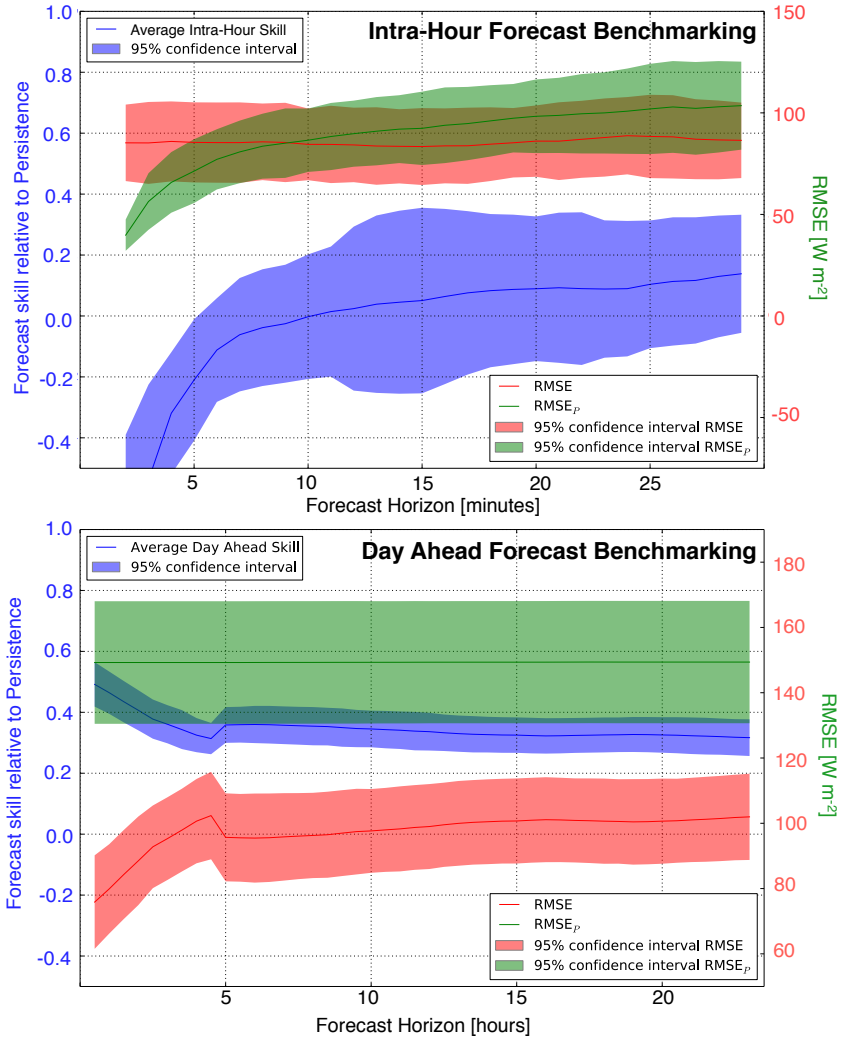

Figure 12: Summary of benchmarking results for solar irradiance forecasts from CPR at all twelve cluster centers. Each plot shows the forecast skill relative to persistence (blue; left axis), which is calculated from the ratios of the RMSEs (red, green; right axis). Top: Intra-hour forecasts. Bottom: Day-ahead forecasts.

Task 3.0: Integrated Solar-Load Forecasts for SMUD

In this task UCSD implemented, post-processed, optimized and integrated operational load forecasts using an ensemble methodology that has been validated both at the local (campus community, utility feeder, utility service region) and at regional California Independent System Operator (CAISO) levels.

Methods

The schematic for the proposed methodology to improve the accuracy of the SMUD load forecasts and generate forecasts for shorter forecast horizons with no exogenous input is shown in Figure 13. A base forecast with given forecast horizon k' A base forecast with given forecalogy to produce k steps ahead forecast such that $k \leq k'$ ≤ eps $k = k'$ = eps ahead forecast such that ecalogy to produce ve the accuracy of the SMUD load(no improvement is possible) otherwise re-forecast and ensemble techniques are applied until the residuals are white. Similarly, for $k < k'$ < erwise re-forecast and ensemble techniques are applied until the residuals are white. Similarly, for te forecasts for shorter forecast horizons with no th at the locawas applied for $k = 1$ and $k \in \{1, 2, \ldots, 24\}$ forecast horizons and the details are explained in the following sections.

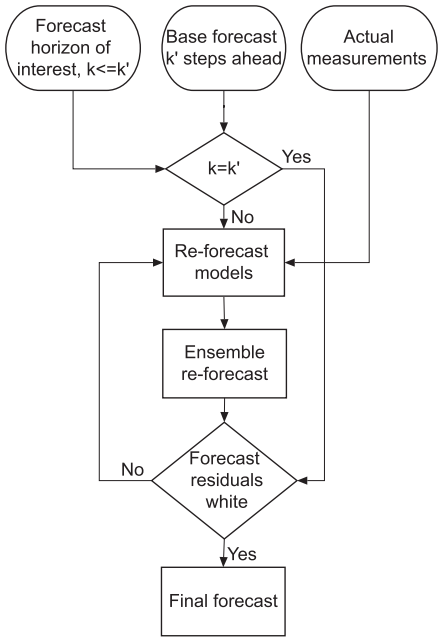

Figure 13: Schematic of the methodology for ensemble re-forecasting. A base SMUD forecast up to k' steps ahead forecast horizon is required to produce k steps ahead forecast such that $k \leq k'$. If the forecast residuals are white then it is considered the final forecast else re-forecast and ensemble techniques are applied unless the residuals are white.

Hour-Ahead Re-Forecast

Using real-time system measurements, a given input $u(t)$ was processed every hour by ensemble re-forecast models (m) to produce Hour-Ahead Market (HAM) point forecasts $\hat{y}(t+k)$,

$$\hat{y}(t+k) = m\left(y(t), y(t-1), \cdots, u(t), u(t-1), u(t-2), \cdots\right), \tag{23}$$

where t is time and k=1 is the forecast horizon. We used 24-hourly Day-Ahead Market (DAM) load estimates provided by SMUD as an exogenous input, filter it every hour and produce HAM load forecasts for the integrated system load. For example, SMUD DAM forecast produced on August 1, 2014 for the following 24 hours is corrected hourly by our re-forecast in order to produce HAM point forecasts.

Day-Ahead Re-Forecast

The above idea was extended for DAM load prediction with forecast horizons $k \in \{1, 2, \ldots, 24\}$. In this case, hourly forecasts were produced at midnight for each hour of the day that just started $h \in \{1, 2, \ldots, 24\}$, i.e.,

$$\hat{Y} = [\hat{y}(t+1), \hat{y}(t+2), ..., \hat{y}(t+24)]^T \tag{24}$$

The time-series models $m(h)$ were built specific for each hour of the day with an exogenous input consisting of estimates for that hour $u(t\text{-}24n)|_{n=0,1,\ldots}$ and corresponding measurements $y(t\text{-}24n)|_{n=1,2,\ldots}$. In summary, we used the twenty-four hourly DAM estimated load provided by SMUD as an input, refined it by applying the models developed for each hour, combined the forecast from all the models and produced a vector of 24 load forecasts corresponding to all hours of the next day.

De-trending

In this project we applied a simple polynomial fit that captures the daily trend for the electric load. Using a simple least squares method (from the Curve Fitting toolbox available in MATLAB) we determined that daily trends are best fitted by a 6th order polynomial as a function of time of the day t, specific for each weekday $w \in \{1,2, \ldots ,7\}$ where Sunday is represented as 1, Monday as 2, and so on. The actual load data can be expressed as a sum of the polynomial fit and the de-trended load, i.e.,

$$y(t, w) = \psi(t, w) + \delta(t, w), \tag{25}$$

where y represents the actual load and ψ is the corresponding de-trended load using a polynomial fit for each day of the week $\delta(t, w)$. Similarly, the input for the models, i.e., a given estimate u is de-trended,

$$u(t, w) = \mu(t, w) + \delta(t, w), \tag{26}$$

where μ is the de-trended input using the polynomial fit $\delta(t, w)$.

Re-Forecast Models

This project focused on extracting the information from non-random noise to produce better forecasts using time-series models with the assumption that the model's residual is white Gaussian noise. The filtered output produced by any of the models is called a re-forecast.

Various time-series models were applied at the re-forecast stage. The linear and non-linear time-series models used in this project are well defined in the literature (Matthewmann and Nicholosn, 1968; Willis and Northcote-Green, 1983; Hagan and Behr, 1987; Ljung, 1999). Only a brief description is provided here.

- **Linear Models**

A generalized model (GM) that linearly combines current and past values of the input $\mu(t)$ and past values of the output $\psi(t)$ to model current output can be defined as (Ljung, 1999),

$$A(q)\psi(t) = \frac{B(q)}{F(q)}(q)\mu(t-\eta_k) + \frac{C(q)}{D(q)}(q)\varepsilon(t), \tag{27}$$

where t represents the time, η_k is the input-output delay parameter, ε is assumed to be white noise, q is a shift operator, and $A(q)$, $B(q)$, $C(q)$, $D(q)$, and $F(q)$ are the polynomials of order η_a, η_b, η_c, η_d, and η_f, i.e.,

$$q^{\pm N}\psi(t) = \psi(t \pm N), \tag{28}$$

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{\eta_a} q^{-\eta_a}, \tag{29}$$

$$B(q) = b_1 + \cdots + b_{\eta_b} q^{-\eta_b+1}, \tag{30}$$

$$C(q) = 1 + c_1 q^{-1} + \cdots + c_{\eta_c} q^{-\eta_c}, \tag{31}$$

$$D(q) = 1 + d_1 q^{-1} + \cdots + d_{\eta_d} q^{-\eta_d}, \tag{32}$$

$$F(q) = 1 + f_1 q^{-1} + \cdots + f_{\eta_f} q^{-\eta_f}. \tag{33}$$

In this project we also applied other forms of GM such as an Autoregressive model with exogenous inputs (ARX), Autoregressive moving average model with exogenous inputs (ARMAX), and Box-Jenkins models (BJ). Table 3 lists the polynomials used in each one of these models.

Table 3: Polynomials used for various model structures.

Model parameters	$A(q)$	$B(q)$	$C(q)$	$D(q)$	$F(q)$
ARX	✓	✓			
ARMAX	✓	✓	✓		
BJ		✓	✓	✓	✓
GM	✓	✓	✓	✓	✓

- **Non-Linear Models**

To model the non-linearities in the load time-series we used the Non-linear Autoregressive model (NARX). In this model the current output is expressed as a non-linear combination of the input and the past values of the output,

$$\psi(t) = f(\psi(t-1), \psi(t-2), ..., \mu(t), \mu(t-1), \mu(t-2), ...), \tag{34}$$

where $f(\cdot)$ can be computed using any non-linear estimator. In this study, we used a wavelet network non-linear estimator. The prediction error method (PEM) was applied to derive the model parameters that minimize the weighted norm of the prediction error for a training dataset.

Ensemble Re-Forecasting

Ensemble re-forecast was produced by a weighted average of the predictions from the previous models. The trend $\delta(t, w)$ was added to the re-forecast $\hat{\psi}(t)$ to produce a final re-forecast,

$$\hat{y}(t, w) = \hat{\psi}(t, w) + \delta(t, w). \tag{35}$$

We investigated three ways of forming an ensemble by optimizing the averaging weights based on global, hourly and weekly performance of the models. The mathematical details regarding the ensembles are provided next.

- **Least Squares Model Ensemble**

An ensemble forecast $\hat{\phi}$ for any given time t is produced by linearly combining forecasts from n forecasting models for k forecast horizons,

$$\begin{pmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} & \cdots & \hat{y}_{1,n} \\ \hat{y}_{2,1} & \hat{y}_{2,2} & \cdots & \hat{y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{k,1} & \hat{y}_{k,2} & \cdots & \hat{y}_{k,n} \end{pmatrix} \tag{36}$$

using the weights $W=[\omega_1, \omega_2, \ldots, \omega_n]^T$, i.e.,

$$\hat{\varphi} = \hat{Y}W \tag{37}$$

For HAM load prediction k is 1 and for DAM k is {1, 2, ..., 24}. The weights ω_i were optimized based on the global performance of the models applied to the T and S datasets, which consist of p actual load measurements $Y = [y_1, y_2, ..., y_p]^T$. Each measurement can be uniquely identified by a timestamp $T = [t_1, t_2, ..., t_p]^T$. Correspondingly, there are p hourly forecasted data points from n forecasting models $\hat{y}_{i,j}$ for $i = 1, 2, ..., p$ and $j = 1, 2, ..., n$,

$$\begin{pmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} & \cdots & \hat{y}_{1,n} \\ \hat{y}_{2,1} & \hat{y}_{2,2} & \cdots & \hat{y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{p,1} & \hat{y}_{p,2} & \cdots & \hat{y}_{p,n} \end{pmatrix} \tag{38}$$

The forecasted data points $\hat{Y}$ are combined using $W = [\omega_1, \omega_2, ..., \omega_n]^T$ to produce a final forecast

$$\hat{\varphi} = \left[\hat{\varphi}_1, \hat{\varphi}_2, ..., \hat{\varphi}_p\right]^T \tag{39}$$

with an error $E=[e_1,e_2,...,e_p]^T$,

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} & \cdots & \hat{y}_{1,n} \\ \hat{y}_{2,1} & \hat{y}_{2,2} & \cdots & \hat{y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{p,1} & \hat{y}_{p,2} & \cdots & \hat{y}_{p,n} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{pmatrix} = \begin{pmatrix} \hat{\varphi}_1 \\ \hat{\varphi}_2 \\ \vdots \\ \hat{\varphi}_p \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{pmatrix} \tag{40}$$

such that the $E = Y - \hat{Y}W$, and W is approximated using linear least squares with an objective function of minimizing the sum of squares of errors, i.e., $G(T, W) = \frac{1}{2}E^TE$, whose global solution is $W = (\hat{Y}^T\hat{Y})^{-1}\hat{Y}^T\hat{Y} \in R_n$.

- **Least Squares Hourly Ensemble (LS-HE)**

In this model an ensemble forecast $\hat{\varphi}(h)$ depends on the hour of the day h, which can be easily retrieved from the timestamp, i.e., $h \subset t$. The ensemble forecast $\hat{\varphi}(h)$ was produced by linearly combining forecasts for the hth hour from n forecasting models with k forecast horizons $\hat{y}_{i,j} \in \{1,2,...,k\}$ and $j \in \{1, 2, ..., n\}$ using the weights $\omega_j(h)$,

$$\hat{\varphi}(h) = \sum_{i=1}^{k}\sum_{j=1}^{n}\hat{y}_{i,j}(h)\omega_j(h). \tag{41}$$

Using the timestamps T, only q data points with $q \leq p$ corresponding to hth hour were considered to compute the weights $\omega_i(h)$ and the same procedure as followed, i.e.,

$$\begin{pmatrix} y_1(h) \\ y_2(h) \\ \vdots \\ y_q(h) \end{pmatrix} = \begin{pmatrix} \hat{y}_{1,1}(h) & \hat{y}_{1,2}(h) & \cdots & \hat{y}_{1,n}(h) \\ \hat{y}_{2,1}(h) & \hat{y}_{2,2}(h) & \cdots & \hat{y}_{2,n}(h) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{q,1}(h) & \hat{y}_{q,2}(h) & \cdots & \hat{y}_{q,n}(h) \end{pmatrix} \begin{pmatrix} \omega_1(h) \\ \omega_2(h) \\ \vdots \\ \omega_n(h) \end{pmatrix} + \begin{pmatrix} e_1(h) \\ e_2(h) \\ \vdots \\ e_q(h) \end{pmatrix} = \begin{pmatrix} \hat{\varphi}_1(h) \\ \hat{\varphi}_2(h) \\ \vdots \\ \hat{\varphi}_q(h) \end{pmatrix} + \begin{pmatrix} e_1(h) \\ e_2(h) \\ \vdots \\ e_q(h) \end{pmatrix} \quad (42)$$

The above algorithm was repeated twenty-four times $\forall h \in \{1, 2, ..., 24\}$ to compute individual weights for each hour of the day for all n models. In this case, $W \in \mathrm{R}^{n,h}$.

- **Least Squares Weekday Ensemble (LS-WE)**

This model produceed ensemble forecast $\hat{\varphi}(d)$ depending on the day of the week $d \in \{1, 2, ..., 7\}$, which was computed from the timestamp t, i.e., $d \subset t$. The ensemble forecast $\hat{\varphi}(d)$ was produced by linearly combining forecasts from n forecasting models with k forecast horizons $\hat{y}_{i,j}(h)$, $i \in \{1,2,...,k\}$ and $j \in \{1, 2, ..., n\}$ using the weights $\omega_i(d)$ specific for each weekday,

$$\hat{\varphi}(d) = \sum_{i=1}^{k} \sum_{j=1}^{n} \hat{y}_{i,j}(d)\omega_j(d). \quad (43)$$

In this case, only r data points with $r \leq p$ corresponding to the d were considered to compute the weights $\omega_i(d)$,

$$\begin{pmatrix} y_1(d) \\ y_2(d) \\ \vdots \\ y_r(d) \end{pmatrix} = \begin{pmatrix} \hat{y}_{1,1}(d) & \hat{y}_{1,2}(d) & \cdots & \hat{y}_{1,n}(d) \\ \hat{y}_{2,1}(d) & \hat{y}_{2,2}(d) & \cdots & \hat{y}_{2,n}(d) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{r,1}(d) & \hat{y}_{r,2}(d) & \cdots & \hat{y}_{r,n}(d) \end{pmatrix} \begin{pmatrix} \omega_1(d) \\ \omega_2(d) \\ \vdots \\ \omega_n(d) \end{pmatrix} + \begin{pmatrix} e_1(d) \\ e_2(d) \\ \vdots \\ e_r(d) \end{pmatrix} = \begin{pmatrix} \hat{\varphi}_1(d) \\ \hat{\varphi}_2(d) \\ \vdots \\ \hat{\varphi}_r(d) \end{pmatrix} + \begin{pmatrix} e_1(d) \\ e_2(d) \\ \vdots \\ e_r(d) \end{pmatrix} \quad (44)$$

The above algorithm was repeated seven times $\forall d \in \{1,2,...,7\}$ to compute individual weights for each weekday and forecasting model. For this ensemble $W \in \mathrm{R}^{n,d}$.

- **Recursive Least Squares Model Ensemble (RLS-ME)**

This method only applies to point forecasts with a single forecast horizon. For real-time applications, this model produces an hourly ensemble forecast $\hat{\varphi}(t)$ depending on the time of the forecast

t. The ensemble forecast $\hat{\varphi}(t)$ was produced by linearly combining forecasts from n forecasting models, $\hat{y}_i(d)$, $i \in \{1, 2, ..., n\}$ using the weights $\omega_i(t)$ updated based on the last measurements,

$$\hat{\varphi}(t) = \sum_{i=1}^{n} \hat{y}_i(t)\omega_i(t). \tag{45}$$

The model was initialized using the weights computed by the LS-ME method as they represent the best estimate of weights for the individual model. The weights $\omega_i(t)$ were then updated using the following RLS algorithm (Passino, 2005).

$$\omega(t) = \omega(t-1) + K(t)(y(t) - \hat{y}(t)')\omega(t-1), \tag{46}$$

$$K(t) = \frac{P(t-1)\hat{y}(t)}{1 + \hat{y}(t)'P(t-1)\hat{y}(t)}, \tag{47}$$

and,

$$P(t) = (I - K(t)\hat{y}(t)')P(t-1). \tag{48}$$

Load Data

For (DAM) forecast, the forecast horizon ranges from 1-24 h; i.e. the forecast is issued at midnight for the entire day. For accurate day-ahead load forecast it is essential to anticipate the type of daily load profile. The load profile from SMUD is highly influenced by seasonal cycles. Therefore, the data was divided into five periods: period 1 (Jan - Apr), period 2 (May), period 3 (Jun - July), period 4 (Aug - Sep) and period 5 (Oct - Dec) as shown in Figure 16. For the first period, the maximum load is 1.75 GW and the load profile remains uniform throughout the 4-month period. The second period marks the beginning of the summer season. During this period, the load profile begins to change due to increasing demand caused by high cooling and air-conditioning loads. Furthermore, in period 3 during the summer season the load profile is very chaotic and demand can be as high as 3 GW, twice the demand for period 1. For these reasons, this is the most challenging period of the year to forecast. By the 4th period, the load demand starts to decrease with a maximum load of 2.75 GW. After the month of September, during the 5th period, again load profile becomes very smooth with maximum demand of 1.5 GW. For period 1 and 5, the load demand shape for the whole day is very persistent day-to-day, whereas for the other periods (2-4) high demand can be observed between the 10th to 20th hours. By dividing the time-series into these 5 periods, we trained independent local forecasting models for each type of load profile. The data was further divided into 2 disjointed data sets, the first to train/estimate the free parameters in the various forecast models and the second to test the models performance (Table 4).

Table 4: Division of the load data from SMUD for training and testing.

Dataset	**Time period**
Training set	Jan 2012 – Dec 2012
Testing set	Jan 2013 – Dec 2013

HAM Load Re-Forecasting Results

Error Statistics

The results and discussion presented here are for the re-forecasts and ensemble re-forecasts produced for the testing set. For the RLS algorithm, the HAM re-forecast started two days before the validation set to discard the initial effect of sudden unstable changes in weights. For HAM, the parameters of the SMUD time-series model were selected using minimum description length (MDL) criterion.

Standard error metrics commonly used to compare the performance of the forecast models are: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Bias Error (MBE), and Root Mean Square Error (RMSE). MAPE measures the accuracy of a method in terms of percentage error. A MAPE of zero implies a perfect fit and there is no upper bound on its value. The mean of absolute errors is represented by MAE and bias in the forecast is represented by MBE. RMSE measures the root mean squared distance between the measured and forecasted load. The results after applying these metrics for the proposed models are presented in Table 5. It is clear that the HAM re-forecast methodology used here results in a reduction in all of the error metrics.

Table 5: Error metrics for the SMUD and UCSD HAM forecasts. The last column shows the improvement when re-forecasting tools are applied to SMUD's predictions.

	MAPE [%]	MAE [kW]	MBE [kW]	RMSE [kW]	Imp [%]
SMUD	4.66	0.03	60.69	78.88	--
UCSD	1.20	0.01	15.17	21.58	74

Correlation Analysis

The assumption in time-series forecasting is that a good forecast model should have white noise as a residual. Otherwise, the information in the residuals should be used to enhance model performance.

Our analysis showed that the residuals from SMUD forecasts have a periodic correlation in the autocorrelation function (ACF), thus justifying the implementation of re-forecast. The ACF of the forecast residuals produced for SMUD and the re-forecast model for HAM are shown in Figure 14. The residuals from the re-forecast are white and performance of the models shows improvement, which validates our approach. One drawback of this analysis is that it does not provide any information about the time of the errors, which is very important in load forecasts as the errors during peak times have higher implications on the stability of the grid than off-peak times. Consequently, a temporal analysis of errors is discussed below.

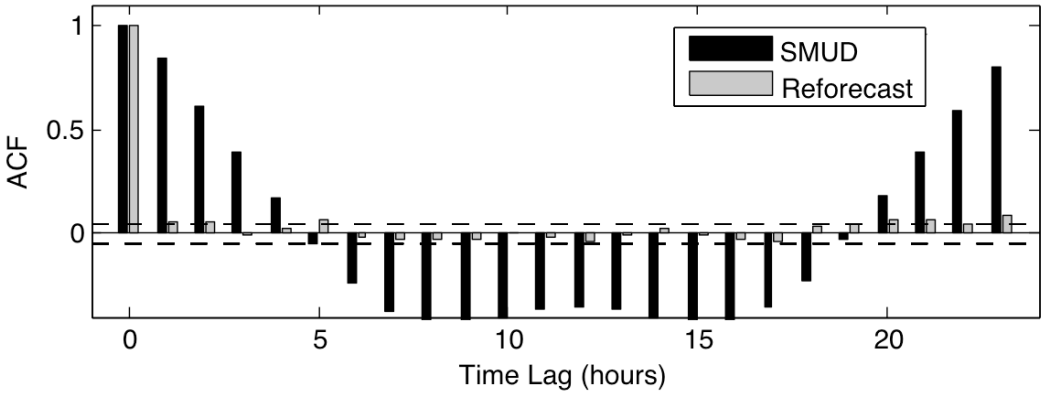

Figure 14: Error ACF for the HAM load forecast.

Temporal Analysis

For SMUD load demand, midnight to 5:00 am PDT in the morning Time Lag (hours) is considered super off-peak time, 6:00 am to 12:00 noon and 8:00 pm to midnight PDT is considered off-peak time, and noon to 8:00 pm PDT is considered an on-peak time. The grid is very sensitive during on-peak times because of high demand. Figure 15 shows the mean and standard deviation in HAM load forecast errors from SMUD and re-forecast model. This analysis shows that the SMUD forecast accuracy varies throughout the day, while the performance of the re-forecast model is consistent throughout the day. The re-forecast model significantly improves forecast quality for all times of the day. Similar results were found for SMUD DAM load forecasts (see Figure 17).

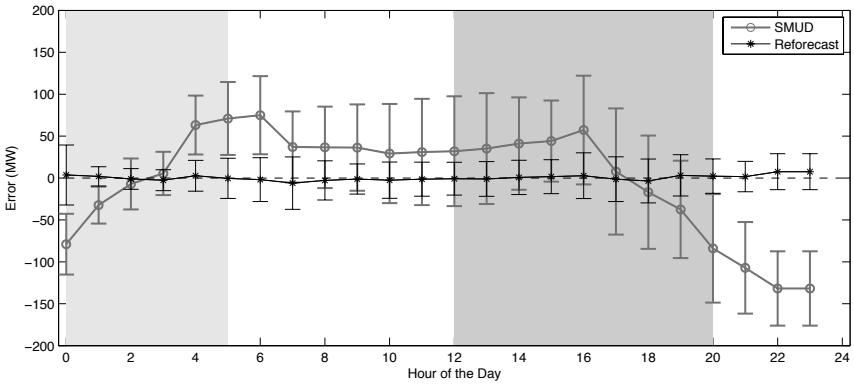

Figure 15: Mean error and error standard deviation versus the hour of the day for the HAM forecast.

DAM Load Re-Forecasting Results

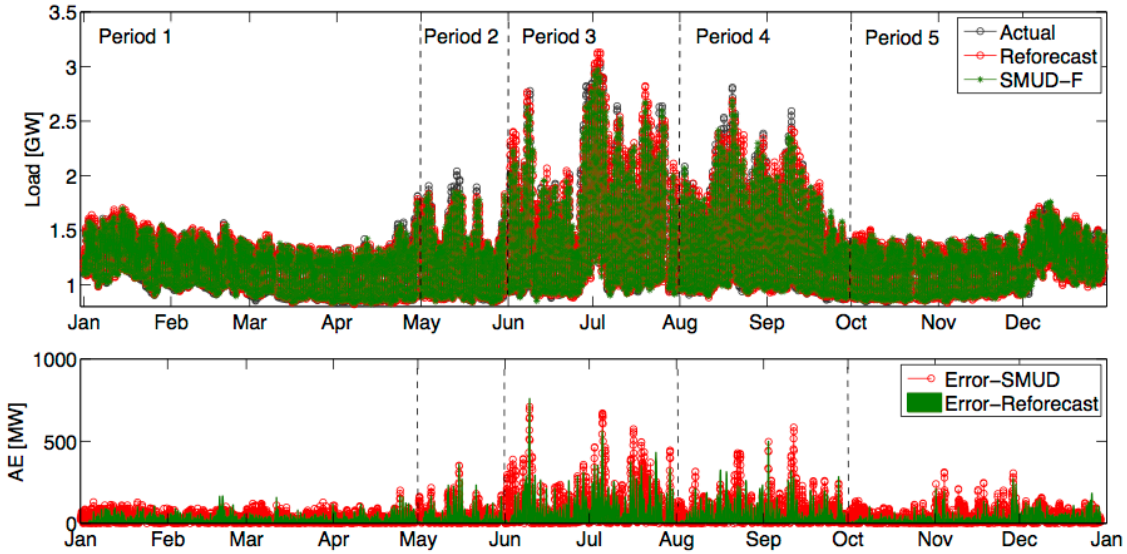

Figure 16: Load forecast and actual time-series for DAM. Load profile is influenced by the seasonal cycles. During summer time there is a sudden increase in load demand and correspondingly it is the most challenging period to forecast for day-ahead market. As expected, the error for period 1 and 5 are lower as compared to the errors in other periods.

The forecasting models were tested for the whole year using the test set ranging from January 2013 to December 2013. The overall time-series with the forecasting results are shown in Figure 16 and the error statistics are listed in Table 6. Period 1 and 5 have low absolute forecasting error as compared to period 2, 3, and 4. Load demand is sensitive to heating and cooling loads, therefore sudden changes in load demand can be observed during the summer season, which results in increased error during the 2nd, 3rd, and 4th period.

Table 6: Error metrics for the SMUD and UCSD DAM forecasts versus the seasonal period. The last column shows the improvement when re-forecasting tools are applied to the SMUD's predictions.

Period		MAPE [%]	MAE [kW]	MBE [kW]	RMSE [kW]	R^2	Imp [%]
1	SMUD	2.54	-17.38	30.51	39.31	0.95	--
	UCSD	2.04	-3.36	24.77	34.59	0.96	**19.7**
2	SMUD	3.38	-7.31	44.10	66.59	0.93	--
	UCSD	3.27	2.60	42.60	66.73	0.93	**3.3**
3	SMUD	6.61	-51.08	1.5.88	157.50	0.89	--
	UCSD	5.39	-3.44	85.44	131.84	0.92	**18.5**
4	SMUD	4.60	-22.69	68.08	103.46	0.92	--
	UCSD	3.71	-0.66	55.85	85.24	0.95	**19.3**
5	SMUD	3.20	-5.69	38.65	55.32	0.90	--
	UCSD	2.43	-0.15	28.37	42.82	0.94	**27.0**
Net Improvement							**20**

Error Statistics

Similar to the HAM re-forecasting, the DAM re-forecast showed improvements in all error statistics for all periods of the year (see **Table 6**), which again is a validation of the assumption that valuable information remained in the residuals of the original forecasts. As discussed above, the month of May (period 2) is a transition period and was the most difficult to improve with a reduction in terms of MAPE of only 3.3%. However, for all of the other periods (11 months of the year) the performance of the re- forecast was over 5 times greater, with an average improvement in terms of MAPE of over 20%.

Temporal Analysis

The temporal error analysis for the DAM is also shown for all five of the periods used in this project. For periods 1 and 5, relatively small error results from the smooth load profile. On the other hand, other periods (especially period 3) show higher errors. However, throughout the year, the magnitude of errors averages out and the mean of the re-forecast error is always nearly zero regardless of the hour of the day. On the contrary, SMUD tends to overestimate for hours 12-20, see **Figure 17**.

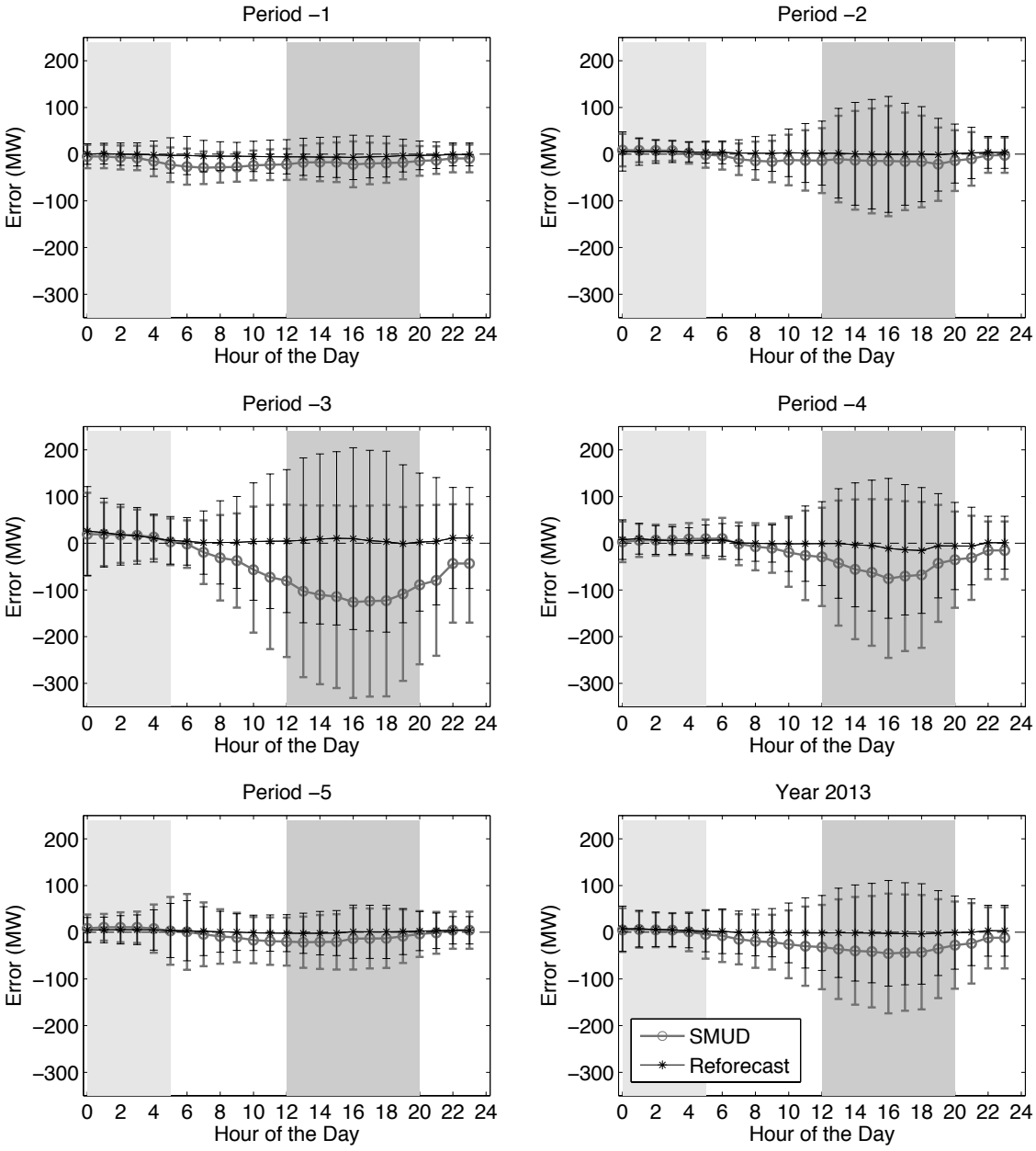

Figure 17: Temporal error analysis for DAM forecast for all the periods and for year 2013 (test set). Throughout the year the magnitude of errors averages out and mean of re-forecast error is always close to zero irrespective of the hour of the day and SMUD overestimates for 12- 20 hours.

Correlation Analysis

An ACF between day-ahead errors for all 24 hours of the day was also performed. As discussed above, for day ahead forecasting, understanding of daily load shape is critical. If load demand shape

changes due to heating or cooling, it takes the algorithm at least one day to adapt to the new profile and learn from the errors (see **Figure 18**). As a result, if the load profile is predicted incorrectly, the errors over the hourly lags are highly correlated. Therefore, in order to check that the method is learning from past errors and that there are no structural error residuals to learn from, we examine the ACF at a 24-hour lag i.e., one day. It can be seen that for hours 10-23 there is a high correlation in the SMUD re-forecast at a one day lag and it is less than the 5% confidence level for the re-forecasting method, see **Figure 18**. Hence, this is the best forecast that can be produced using the given time-series with no exogenous inputs such as weather forecasts.

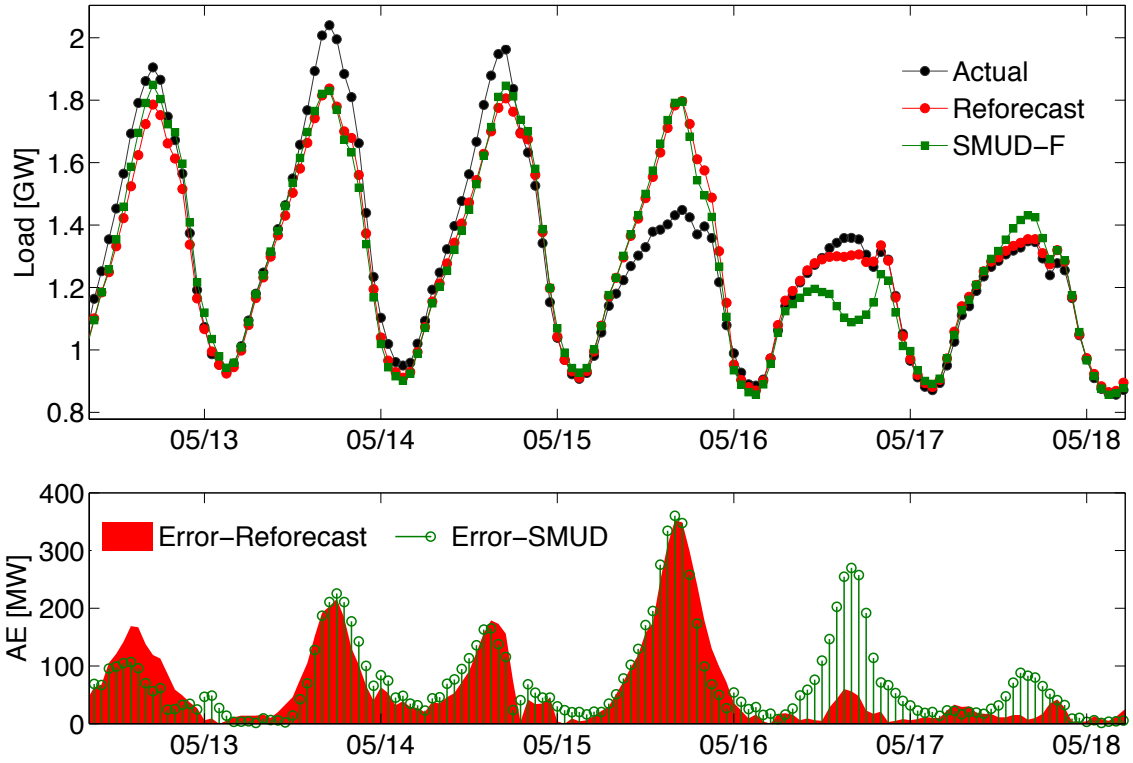

Figure 18: Actual and forecasted load time-series for SMUD from 05/13/2013 to 05/18/2013 (Summer). It can be seen that the shape of daily load profile changes suddenly on 05/15/2013 and causes a large magnitude error in DAM forecast for both SMUD forecasts and SMUD forecasts models. Whereas the following day, re-forecast methods adapt to the changing profile and as a result produce a more accurate load forecast as compared to forecasts produced by SMUD.

Task 5.0 Optimization of CPR's forecasts using UCSD's methodologies

CPR has provided one year of historical forecasts of GHI for the 12 cluster centers determined in Task 1.0. These forecasts were issued in two distinct time-horizons and resolutions: Intra-Hour (IH: 30 minute horizon) and Day-Ahead (DA: 24 hour horizon), which are resolved at 1 minute and 30 minute centered averages respectfully. Both of these forecasts were issued every half hour and, as a result, there is no layover in the IH forecast and 23.5 hours of layover for the DA forecasts as illustrated in Figure 9.

CPR's irradiance forecasts were benchmarked against ground truth data at each of the cluster centers. This was achieved through the use of SMUD's ground based sensor network, which is deployed on the same 5 km grid as the NOAA National Digital Forecast Database (NDFD), see Figure 10.

In order to benchmark the IH and DA forecasts we also consider two simpler models as baseline forecast techniques. For the IH forecasts the baseline is the persistence model that assumes the persistence of the clear-sky index, see Eqn. (21).

For the DA forecasts, we use the GHI forecasts provided by NOAA from the North American Mesoscale Forecast System[1] (NAM). These forecasts are produced for the CONUS region every 6-hours for the next 0-36h at hourly resolution. In order to compare the NAM forecast with the CPR reforecast we extracted the GHI from the NAM files for the 12 nodes and resample the data in 30 min intervals (via a simple linear interpolation).

Re-Forecast methodology

In this task we used Machine Learning (ML) tools to reforecast the IH and DA predictions produced by CPR. The idea behind this task is that the satellite-based forecast models used by CPR can be improved by adding information about the ground data and other local variables. In this task we used ML tools such as Artificial Neural Networks, k-nearest-neighbors (kNN) or non-linear Least Squares (nLS). These tools are well suited to improve the CPR forecasts due to their overall flexibility and nonlinear pattern recognition abilities. However, the forecast skill of such methods depends on (i) the inputs variables and their preprocessing; (ii) the architecture of the ML methods (e.g. the number of layers, numbers of neurons per layer in an ANN); and (iii) the fraction and distribution between training and testing data. In a data rich scenario such as this one where CPR forecasts, NAM forecasts and ground data are available it is not always evident which variables should be included in the reforecast model.

[1] http://www.emc.ncep.noaa.gov/index.php?branch=NAM
[2] L_{10} is the bound for ACE 10-minute averages in CPS2 calculations, above which a violation is counted for a

There are no "recipes" or theorems to guide us in these decisions and the forecast ability of the ML methods is often sub-optimally exploited. Thus in this task we applied an optimization algorithm to filter the best set of input variables and determine the ML model that most improves on the CPR forecasts for the 12 nodes.

The general form for the models used to reforecast the DA and IH GHI predictions for the 12 nodes is:

$$\hat{G}_i(t) = f(G_{0,i}(t), G_{0,j}(t), \ldots, G_{0,k}(t), M_i(t-\tau_1), \ldots, M_i(t-\tau_n), \vec{u}(t)) \qquad (49)$$

where $\hat{G}_i(t)$, is the reforecasted GHI at time *t*. The forecast function can be: a linear Least-Squares (LS) curve fit, a non-linear LS curve fit, a non-linear ANN, a k-nearest-neighbor model, etc. The final function is determined via the optimization procedure explained below. The inputs used in the reforecast can be separated in several groups:

- $G_{0,i}(t)$ is the CPR forecast for node *i*.
- $G_{0,j}(t)$, …, $G_{0,k}(t)$ are the CPR forecasts for nodes other than node *i*.
- $G_{0,i}(t-\tau_1)$, …,$G_{0,i}(t-\tau_n)$ are the time-delayed GHI ground data for node *i*. The delays $\tau_1,\ldots,\tau_n$ must be such that the ground data is available at the time of the reforecast.
- $\vec{u}(t)$ includes any other relevant variables. Here we consider GHI forecasts obtained from NAM for the 12 nodes. Other variables include the time of the day, the solar elevation angle, the solar azimuth angle and the GHI clear-sky index.

The initial data is split 50/50 into two disjointed datasets: the training and validation/testing sets. The training data is used in the input selection, model selection and model optimization stage. The validation set is used to assess the models' performance. All the results presented below were obtained for the validation set since they are indicative of how the reforecast models perform in real-time.

Optimization Algorithm

The optimization algorithm used in this work follows a two-step procedure. In the first step we use a linear model to select the best set of inputs. This approach allows filtering the best combination of inputs out of the millions of possible combinations. In the second step several non-linear machine-learning algorithms are tested to find the best reforecast model using the variables determined in the first step. **Figure 19** illustrates this procedure graphically.

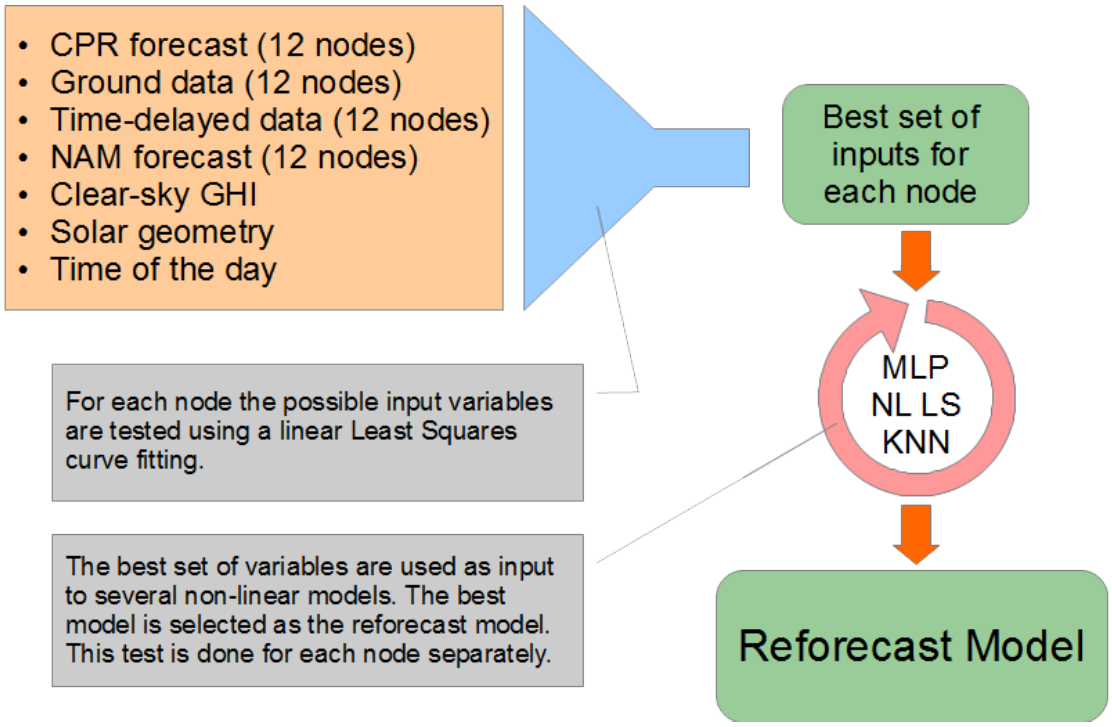

Figure 19: Schematic for the methodology used to determine the best reforecast model for each node.

Error metrics

Once the best reforecast models are determined we analyzed their performance and compared them against the other forecasts (CPR and Persistence for the IH forecasts and CPR and NAM for the DA forecasts). The following bulk metrics were used in this analysis:

- Mean Bias Error

$$\mathrm{MBE} = \frac{1}{N}\sum_{i}^{N} G_i - \hat{G}_i \tag{50}$$

- Mean Absolute Error

$$\mathrm{MAE} = \frac{1}{N}\sum_{i}^{N} \left| G_i - \hat{G}_i \right| \tag{51}$$

- Root Mean Square Error

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i}^{N} \left(G_i - \hat{G}_i \right)^2} \tag{52}$$

- Coefficient of Determination

$$\mathrm{R}^2 = 1 - \frac{SS_{\mathrm{res}}}{SS_{\mathrm{tot}}}, \quad \text{with} \quad SS_{\mathrm{res}} = \sum_{i}^{N} (G_i - \hat{G}_i)^2, \quad \text{and} \quad SS_{\mathrm{tot}} = \sum_{i}^{N} (G_i - \bar{G})^2. \tag{53}$$

In these equations $\hat{G}$ is the forecasted value for a given model, G is the ground data, and $\bar{G}$ is the averaged ground data, all for the validation set.

Results

Intra-hour reforecast

Once the best models for the IH forecasts were determined using the training data we applied them to the validation dataset. Table 7 lists the error metrics for each node. The cells highlighted in green identify the best model for the respective metric. The table shows that the UCSD reforecast achieves the lowest errors and highest coefficient of determination almost always. The only two exceptions happen in the MBE metric. Although, given that MBE values are very small, such occurrences do not impact the improvement of the UCSD reforecasts significantly as the improvements on the other metrics are substantial. The values in the tables show uniformity for all the nodes, indicating that there is very little geographical variation in forecast performance.

Table 7: Intra-hour forecast error metrics for the 12 nodes

	MBE [Wm^{-2}]			MAE [W/m^{-2}]			RMSE [W/m^{-2}]			R^2 [-]		
node	PER[1]	CPR[2]	RFT[3]	PER	CPR	RFT	PER	CPR	RFT	PER	CPR	RFT
1	2.0	0.1	0.0	30.8	38.9	30.5	78.5	73.3	66.4	0.92	0.93	0.94
2	5.1	-3.4	0.5	30.4	42.8	28.3	75.5	73.2	62.8	0.93	0.93	0.95
3	1.9	2.7	0.7	33.5	43.2	32.5	78.2	78.2	66.2	0.92	0.92	0.95
4	1.6	7.3	-1.9	30.7	41.9	29.4	75.2	73.5	63.8	0.93	0.93	0.95
5	3.2	0.2	-1.9	34.3	43.8	30.9	86.0	78.7	69.4	0.91	0.92	0.94
6	2.2	7.9	1.0	30.9	39.4	28.3	77.6	72.3	62.8	0.92	0.93	0.95
7	2.3	-6.7	-0.5	30.0	39.2	28.4	75.1	73.2	61.6	0.92	0.93	0.95
8	3.5	12.6	0.2	33.1	45.0	30.5	81.9	76.5	68.4	0.92	0.93	0.94
9	2.6	11.0	1.3	31.1	45.3	29.2	78.5	75.0	63.0	0.92	0.93	0.95
10	2.1	19.8	1.1	31.7	45.2	30.1	78.9	75.5	66.3	0.92	0.93	0.95
11	1.4	7.3	0.0	31.7	42.5	30.8	77.2	75.6	65.4	0.92	0.93	0.94
12	2.7	10.1	-1.2	31.2	47.0	30.8	77.9	82.9	65.9	0.92	0.91	0.95

[1]Persistence forecast, [2]CPR forecast, [3]UCSD reforecast

Table 8 quantifies the improvement of the UCSD reforecast relative to the Persistence and the CPR forecasts. Values highlighted in red indicate a decrease in performance, and values highlighted in green indicate instances for which the UCSD predictions are at least 15% better than the other ones for a given error metric.

Figure 20 shows the RMSE for the three forecast models (Figure 20 (a-c)) and the forecasts skill of the UCSD reforecast relative to the CPR predictions overlaid to the map of clusters (Figure 20(d)). This figure allows observing the reduction in RMSE from the basic persistence model to the CPR forecast and the UCSD reforecast. This figure shows that there is little variation in the forecast performance from node to node.

Table 8: Intra-hour reforecast improvement relative to the persistence and the CPR forecasts. MBE improvement calculated as the difference of absolute MBE for the two models compared. All the other improvements percentage values relative to the model the UCSD reforecast is being compared to.

	MBE		MAE		RMSE		R^2	
node	RFT/PER	RFT/CPR	RFT/PER	RFT/CPR	RFT/PER	RFT/CPR	RFT/PER	RFT/CPR
1	2.0	0.04	0.8	21.5	15.3	9.4	2.5	1.3
2	4.6	2.9	7.0	33.9	16.8	14.2	2.3	1.9
3	1.2	1.9	2.8	24.7	15.4	15.4	2.3	2.3
4	-0.3	5.4	4.2	29.8	15.1	13.1	2.2	1.9
5	1.3	-1.8	9.9	29.6	19.3	11.8	3.5	1.8
6	1.3	6.9	8.5	28.3	19.0	13.1	2.8	1.7
7	1.8	6.2	5.4	27.5	18.1	16.0	2.8	2.3
8	3.3	12.5	7.8	32.3	16.4	10.5	2.7	1.5
9	1.3	9.7	6.3	35.5	19.8	16.1	3.0	2.3
10	1.0	18.7	5.1	33.5	16.1	12.2	2.5	1.7
11	1.4	7.3	3.0	27.5	15.4	13.5	2.4	2.0
12	1.5	8.9	1.3	34.4	15.3	20.5	2.3	3.5

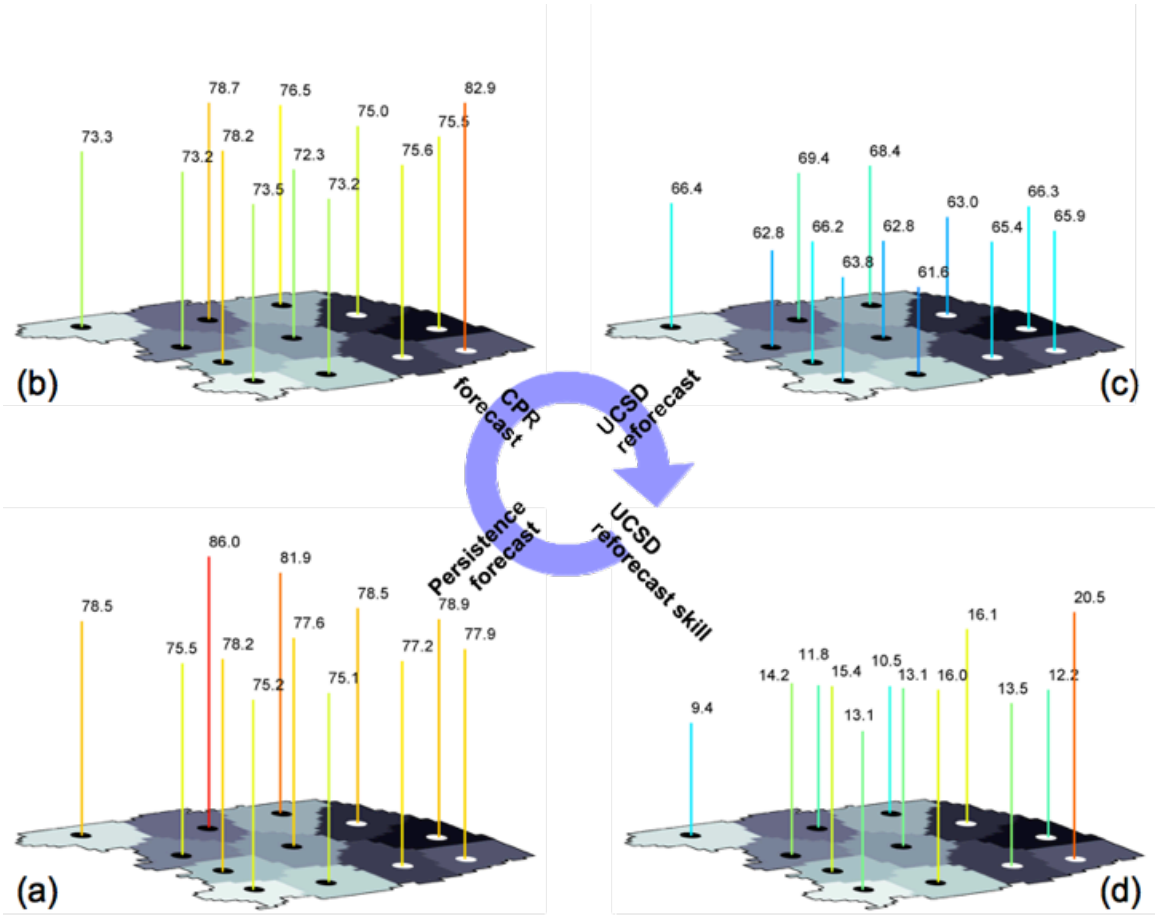

Figure 20: Four maps that illustrate the different forecast and how their performance varies from node to node. (a) Persistence forecast RMSE, (b) CPR forecast RMSE, (c) UCSD reforecast RMSE and (d) UCSD reforecast percentage improvement over CPR forecast. The height and color lines indicate the magnitude of RMSE (a-c) or forecast skill (d). RMSE values are in Wm^{-2}.

The previous analysis does not factor the forecast horizon; it considers all the validation data in bulk. In Figure 21 we assess the forecast performance as a function of the forecast horizon (1 to 30 min in the case of the IH forecasts). Figure 21 (top) shows the RMSE for the 12 nodes and the three forecasts. First of all, as expected, the IH persistence forecast exhibits a low error for short horizons after which error increases. The IH forecast from CPR, on the other hand, exhibits a relatively uniform RMSE for all time horizons since it does not benefit from using the latest ground data. The UCSD reforecast by incorporating the CPR forecasts and ground data is able to reduce the RMSE for all horizons. Figure 21 (bottom) shows the forecast skill of the UCSD reforecast over CPR and Persistence and the forecast skill of the CPR forecast over the persistence model.

Finally Figure 22 shows seven days in the validation dataset that illustrate some of the improvements achieved by the UCSD reforecast. The annotations in the figure highlight instances for which the CPR bias is corrected and the benefit of using the latest ground data to obtain a better prediction for GHI.

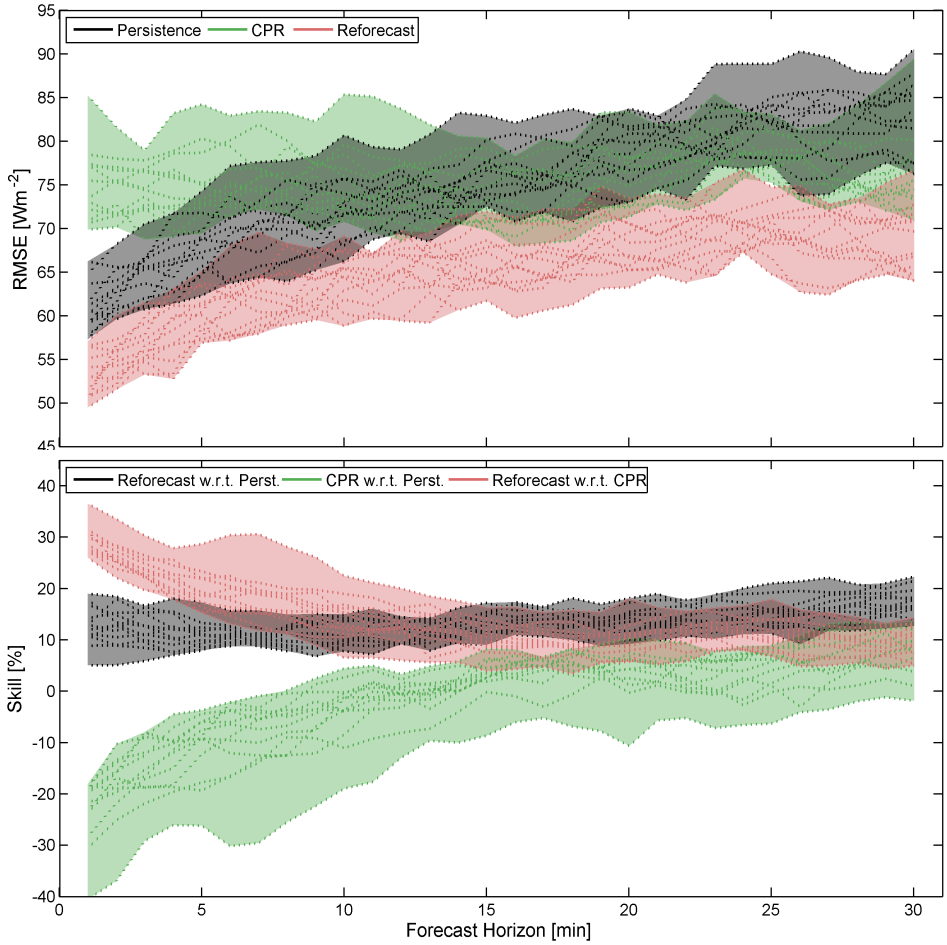

Figure 21: Forecast performance versus the forecast horizon. Top: RMSE for the Persistence model, CPR model and UCSD reforecasts. Bottom: CPR improvement with respect to Persistence and UCSD reforecast improvement over CPR reforecast.

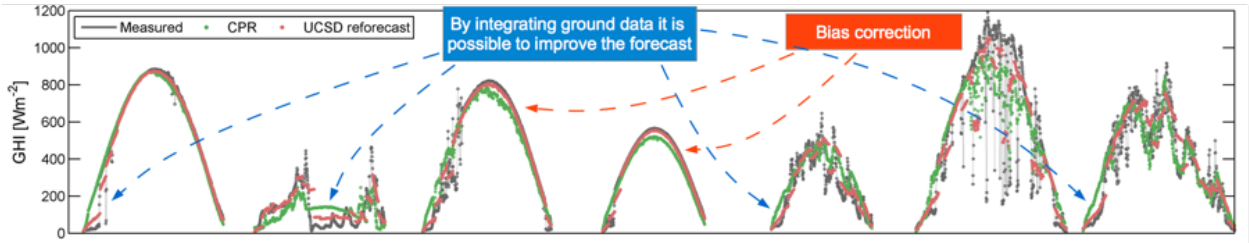

Figure 22: Measured ground data, CPR forecast and UCSD reforecast for seven different days in the validation dataset. The annotations highlight some of the improvements achieved by the UCSD reforecast.

Day-ahead reforecast

The error metrics for the day-ahead forecast are shown in Table 9. Three forecasts are reported: NAM, CPR and UCSD reforecast. The values in green highlight the best forecast for the respective error metric. The table shows that the UCSD reforecast is the best forecast in terms of all metrics. In four instances the CPR forecast shows a smaller MBE than the UCSD reforecast, but as in the IH case the improvements in the other metrics are substantial. Table 10 compares the improvements of the UCSD reforecast with respect to the predictions from CPR and NAM. Values highlighted in green indicate an improvement above 15%, and values highlighted in red indicate a decrease in performance. Figure 23 shows the RMSE for the three forecast models (Figure 23 (a-c)) and the forecasts skill of the UCSD reforecast relative to the CPR predictions overlaid to the map of clusters (Figure 23 (d)).

Table 9: Day-ahead forecast error metrics for the 12 nodes.

	MBE [Wm^{-2}]			MAE [W/m^{-2}]			RMSE [W/m^{-2}]			R^2 [-]		
node	NAM[1]	CPR[2]	RFT[3]	NAM	CPR	RFT	NAM	CPR	RFT	NAM	CPR	RFT
1	-72.2	-3.8	-9.8	82.3	55.0	42.9	109.7	90.0	77.4	0.84	0.89	0.92
2	-74.7	-6.3	-8.9	83.7	53.5	46.3	111.1	88.2	77.2	0.84	0.90	0.92
3	-68.4	0.0	-7.8	78.7	56.7	48.0	108.4	90.5	79.7	0.85	0.90	0.92
4	-62.1	6.4	-5.3	74.0	58.9	45.7	101.8	90.8	76.6	0.86	0.89	0.92
5	-71.7	-1.6	-9.1	79.8	56.6	47.6	110.0	88.8	78.0	0.85	0.90	0.92
6	-63.0	6.3	-7.7	73.7	55.9	42.7	99.1	88.5	72.7	0.87	0.90	0.93
7	-77.8	-10.3	-4.2	86.9	61.3	41.8	109.8	90.6	71.8	0.83	0.88	0.93
8	-60.3	9.3	-8.3	70.4	57.6	47.2	101.8	89.8	77.5	0.87	0.90	0.92
9	-58.4	10.8	-4.8	69.9	60.6	48.3	97.9	91.0	75.5	0.88	0.89	0.93
10	-49.7	19.7	-2.7	64.2	58.7	44.8	92.2	92.9	74.3	0.89	0.89	0.93
11	-62.1	6.7	-4.4	73.4	58.6	45.0	100.6	90.8	75.2	0.87	0.89	0.92
12	-60.3	10.6	-7.1	72.5	58.1	46.9	101.8	90.8	77.3	0.87	0.89	0.92

[1]NAM forecast, [2]CPR forecast, [3]UCSD reforecast

Table 10: Day-ahead reforecast improvement relative to the NAM and the CPR forecasts. MBE improvement calculated as the absolute difference in the two models compared. All the other improvements are relative to the model the reforecast is being compared to.

	MBE		MAE		RMSE		R^2	
node	RFT/NAM	RFT/CPR	RFT/NAM	RFT/CPR	RFT/NAM	RFT/CPR	RFT/NAM	RFT/CPR
1	62.4	-6.0	47.9	22.0	29.4	14.0	9.6	3.1
2	65.8	-2.6	44.7	13.5	30.5	12.4	9.7	2.6
3	60.6	-7.8	39.0	15.3	26.4	11.9	8.0	2.6
4	56.8	1.1	38.2	22.4	24.7	15.6	6.9	3.5
5	62.6	-7.5	40.4	16.0	29.1	12.2	9.0	2.5
6	55.3	-1.4	42.0	23.5	26.7	17.9	6.7	3.7
7	73.6	6.1	51.9	31.8	34.6	20.8	11.8	4.9
8	52.0	1.0	32.9	18.1	23.9	13.6	6.3	2.9
9	53.6	6.0	30.8	20.3	22.9	17.1	5.7	3.7
10	47.1	17.0	30.2	23.6	19.4	20.1	4.2	4.4
11	57.8	2.3	38.6	23.1	25.2	17.2	6.8	3.9
12	53.2	3.5	35.3	19.2	24.0	14.8	6.5	3.3

Figure 24 (top) shows the forecast performance as a function of the forecast horizon. For the DA forecasts there are 48 values issued for the next 24 h. The shaded areas envelop the forecasts for the 12 nodes for a given methodology. The large width of the NAM RMSE envelope indicates that the forecast performance varies significantly from node to node, the other two forecasts show much more uniformity. The figure also shows that the NAM and the UCSD forecasts performances do not vary much with the horizon. Conversely the CPR forecast shows larger variations. Figure 24 (bottom) shows the improvement of CPR forecast over NAM and UCSD reforecast over NAM and CPR.

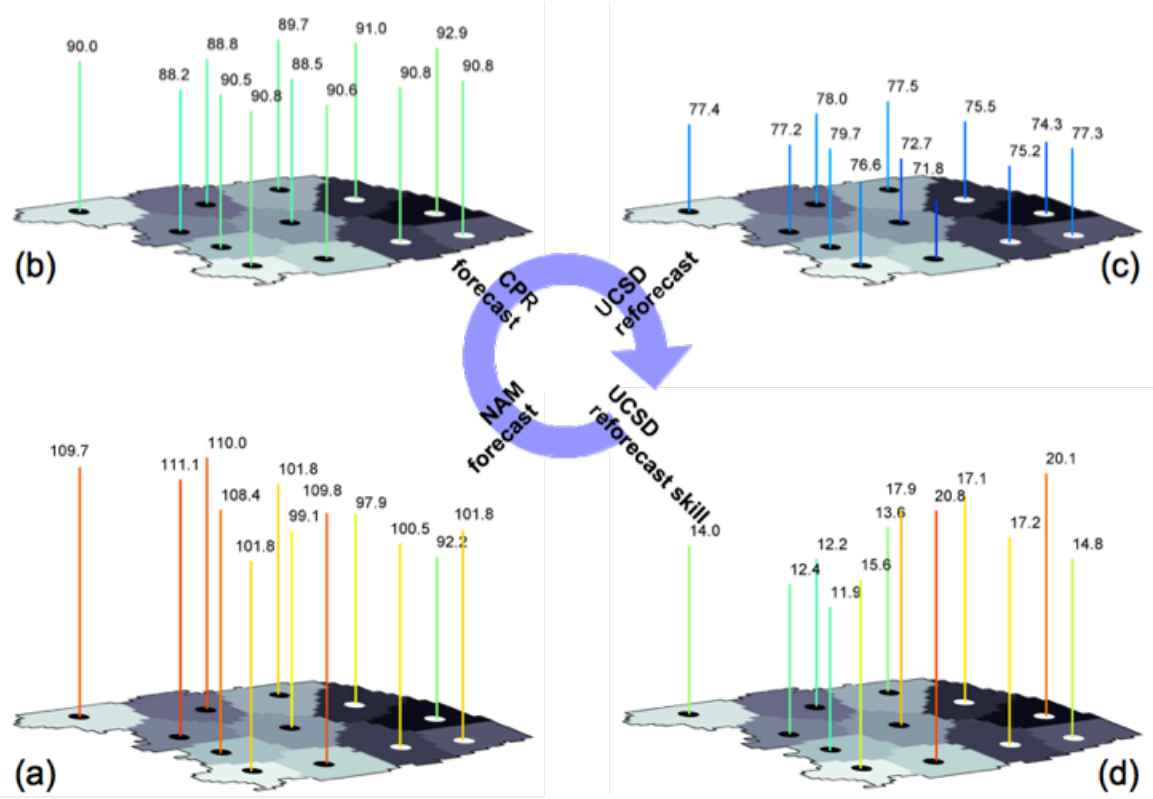

Figure 23: Four maps that illustrate the different forecasts and how their performance varies from node to node. (a) NAM forecast RMSE, (b) CPR forecast RMSE, (c) UCSD reforecast RMSE and (d) UCSD reforecast improvement over CPR forecast. The height and color lines indicate the magnitude of RMSE (a-c) or forecast skill (d). RMSE values are in Wm^{-2}.

Finally Figure 25 shows eight days in the validation dataset that illustrate some of the improvements achieved by the UCSD reforecast. The annotations in the figure highlight instances for which the CPR bias is corrected and the benefit of using the latest ground data to obtain a better prediction for GHI. In some cases we also observe that the reforecast can have a negative impact in the forecast performance. However in view of the overall forecast improvement reported in the tables 9 and 10 it is possible to conclude that those cases are few and far between.

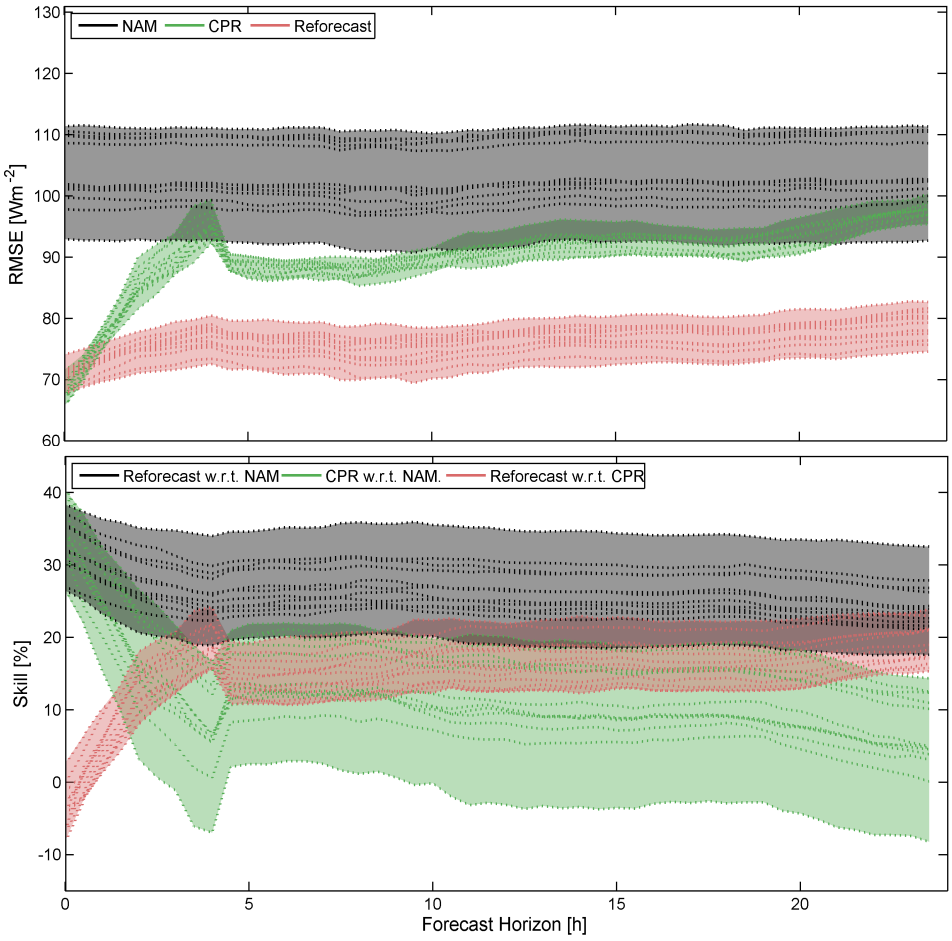

Figure 24: Forecast performance versus the forecast horizon. Top: RMSE for the Persistence, CPR and UCSD reforecasts RMSE. Bottom: CPR improvement with respect to Persistence and UCSD reforecast improvement over CPR reforecast.

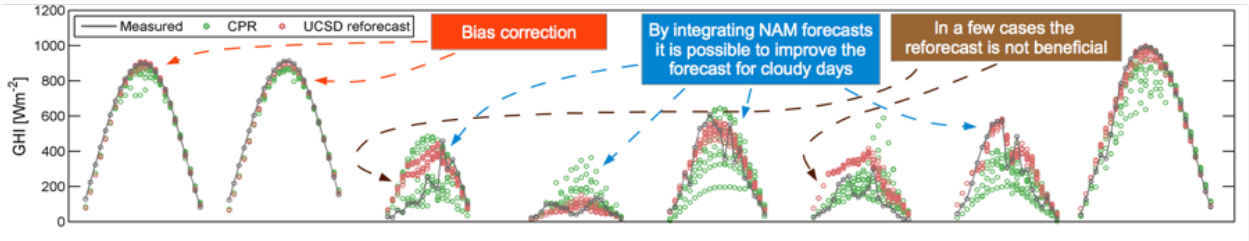

Figure 25: Measured ground data, CPR forecast and UCSD reforecast for eight different days in the validation dataset. The annotations highlight some of the improvements achieved by the UCSD reforecast.

Task 6.0 Solar-Load Uncertainty Implementation into ESIOS and PROM

PV forecast improvements have been recognized as one of the approaches to reducing PV integration cost and reliability risks to the power grid. In this study, the value of improved PV forecast performance was evaluated through operating reserve requirements and power system operation cost. Operating reserve requirements are affected by the inherent variability of load and PV, as well as the forecast uncertainty of them. Reserve requirements are then used as inputs to the production simulations involving hourly and real-time dispatch generation resources. Two study cases based on the SMUD system were developed for this purpose: 1) the base case, which has a PV forecast from one-day-ahead persistence model, i.e., using today's PV production as the forecast for tomorrow; 2) the change case, which uses PV forecast provided by UCSD and CPR through the state-of-the-art forecast techniques for the projected SMUD PV fleet. Results are presented in the following sections.

Types of Reserves

In the SMUD system study, the following reserves are used when simulating generation operations:

- **Day-ahead planning reserve (DA PR):** Flexible generation capacity required in DA planning that are intended to mitigate forecast errors (uncertainty), within-hour variability and loss of generation. However, it is assumed that the current contingency reserve requirement does not increase regardless of the amount of PV integration. With respect to this study, DA PR refers only to the components associated with forecast uncertainty and variability with system generation and load.
- **Regulation reserve (RR):** RR is the reserve needed to cover variations of load and PV at the one to ten-minute time scale. RR requires fast response and therefore must be provided by units on AGC. The amount of RR can be provided by a resource needs to be attainable within 10 minutes. RR is also a component of DA PR.

Dynamic operating reserves that vary with hour of the day were applied in the unit commitment process in the study. In general, response time of up to half an hour is acceptable for DA PR. Therefore, DA PR can be fulfilled by AGC, other online units, offline units, or storage and demand response

resources that can reach desired output within 30 minutes. Table 11 summarizes the types of reserves and the resources that can provide them.

Table 11. Types of Reserves and Qualified Resources

Reserves	Qualified Resources
DA PR Up	Online base and AGC generation, online and offline peaking units, storage and demand response resources that can respond within 30 minutes
DA PR Down[(a)]	Online base and AGC generation, online peaking units, storage and demand response that can respond within 30 minutes
RR Up	Online AGC generation and storage that can respond within 10 minutes
RR Down	Online AGC generation and storage that can respond within 10 minutes

[(a)] The current operation practice usually does not involve a DA PR down requirement, other than having a plan as to what units will be taken off and in what order if there is minimum load or over generation issues. The reason is that without the presence of significant variable generation, the issue of generation "tripping on" (generation showing up that was not expected) does not exist. Over generation does not happen very often. With high penetration of solar, generation can actually "trip on," which requires system operators to manage differently.

Methodology for Reserve Requirements

Day-ahead Planning Reserve Calculation

DA PR is calculated as the difference between the actual load and DA load forecast. In the case with PV generation, it is the difference between the actual *net load* and the DA forecast of *net load*, where *net load* = load – PV. When forecast data is not available, which is the case for future study years, statistical models can be constructed based on historical forecast performance to simulate forecast errors (ref). In this study, DA load forecast was provided by SMUD. As to DA PV forecast, a one-day-ahead persistence PV forecast was used in the base case, and a forecast produced by UCSD and CPR was used in the change case. Reserve requirements were produced for each of these two cases. DA PR requirements were calculated separately for each month for each operating hour, using the following procedure:

1. Calculate the DA PR requirements time series (load or net load minus its DA forecast) for the entire month, with 1-minute resolution.

2. For each operating hour, group all the reserve requirements data within this hour during the entire month.
3. Rank all the data within this hour and truncate 5 percent of the points. This will give us 95 percent of confidence level to meet control performance requirements within this hour.
4. Find the maximum and minimum values of the truncated time series as the DA PR requirements for this operating hour.
5. Twenty series of Monte Carlo simulations are used to provide a more robust evaluation. The final DA PR values are the average of the 20 runs.

Regulation Reserve Calculation

Regulation reserve is calculated as the difference between actual net load and real-time schedule of net load. There are multiple approaches to create real-time forecasts. A persistence model is used for both the base case and the change case in this study. The naive persistence model assumes that the load is the same as what it was a certain number of minutes ago (10 minutes forecast lag time is used in this study). Then a 10-minute ramp length between two real-time forecasts was added.

A similar procedure to DA PR calculation is applied to find the regulation requirements for each operating hour in a month. The L_{10}[2] limit of CPS2 is deducted from RR requirement because ACE values within the L_{10} limit are allowed. The L_{10} is assumed to be 35.84 MW. ACE is simulated as the imbalance between generation and load.

Figure 26 illustrates the day-ahead and real-time load forecasts, as well as the amount of uncertainty and variability that needs to be covered by DA planning reserve and regulation reserve, respectively.

The same persistence model for the real-time load forecast is used to calculate the real-time PV forecast. Day-ahead and real-time PV forecasts, as well as the type of reserves needed to cover associated uncertainty are illustrated in Figure 27. Because in both base and change cases, the same persistence model was used for real-time PV forecast, the regulation requirements are expected to be the same for both cases.

[2] L_{10} is the bound for ACE 10-minute averages in CPS2 calculations, above which a violation is counted for a 10-minute period.

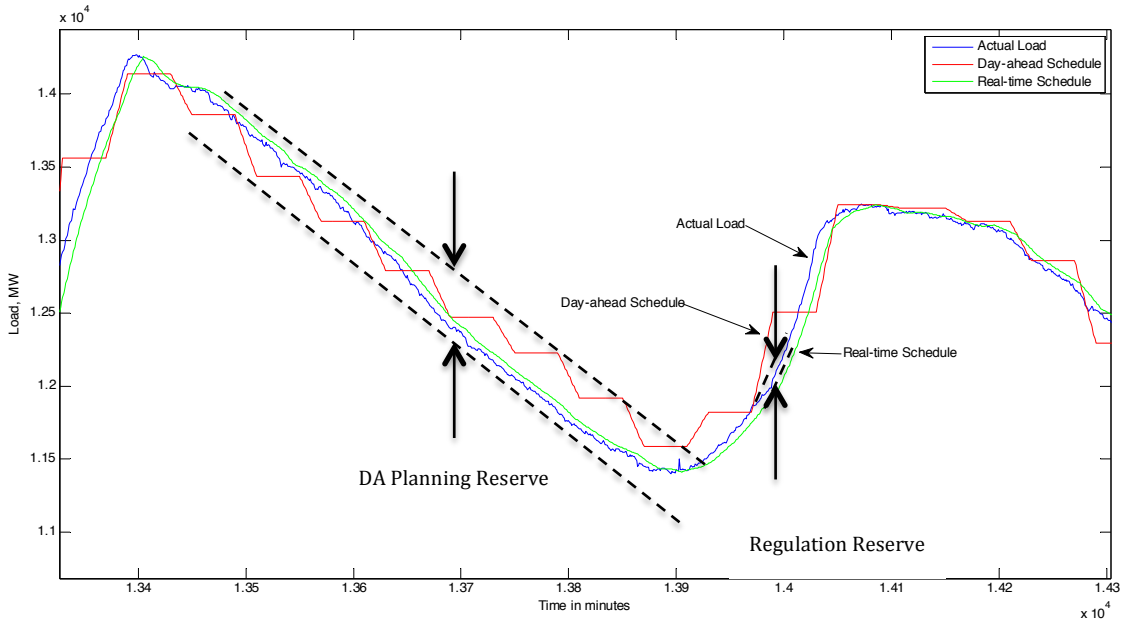

Figure 26: Illustration of DA and Real-Time Load Schedules and Associated Reserve Requirements

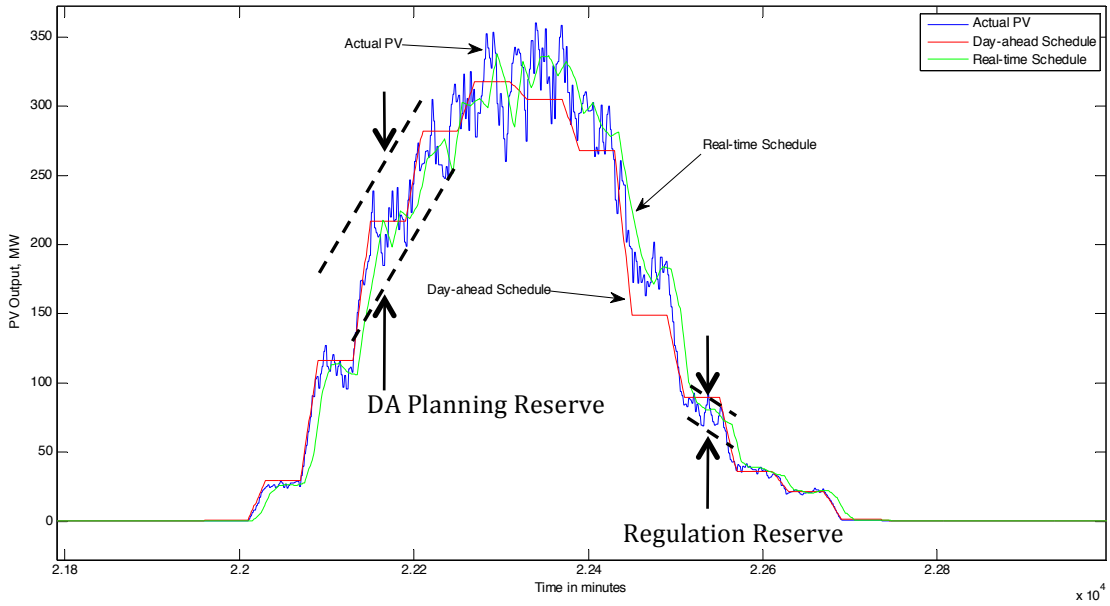

Figure 27: Illustration of DA and Real-Time PV Schedules and Associated Reserve Requirements

Reserve Requirements for SMUD Study Cases

DA Planning Reserve

This section shows DA PR requirements in January (winter), August (summer), April (spring) and October (fall) as examples.

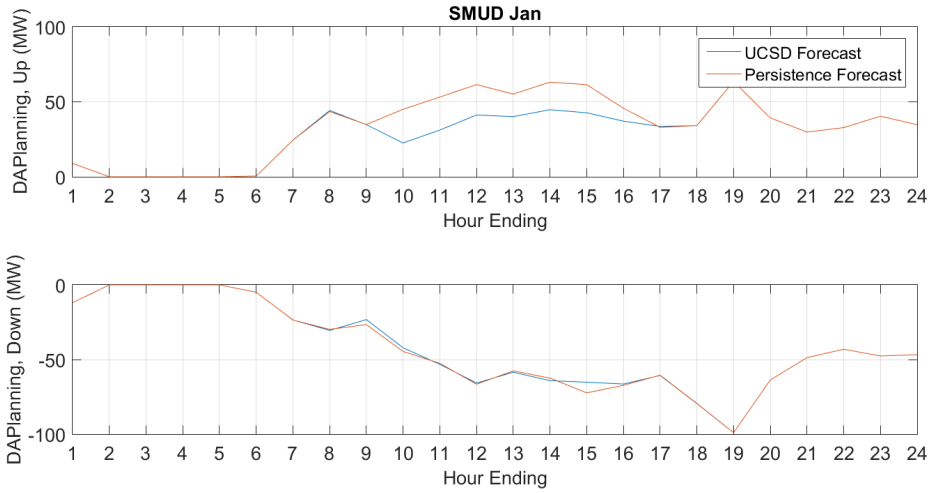

Figure 28: DA Planning Reserve Requirements in January

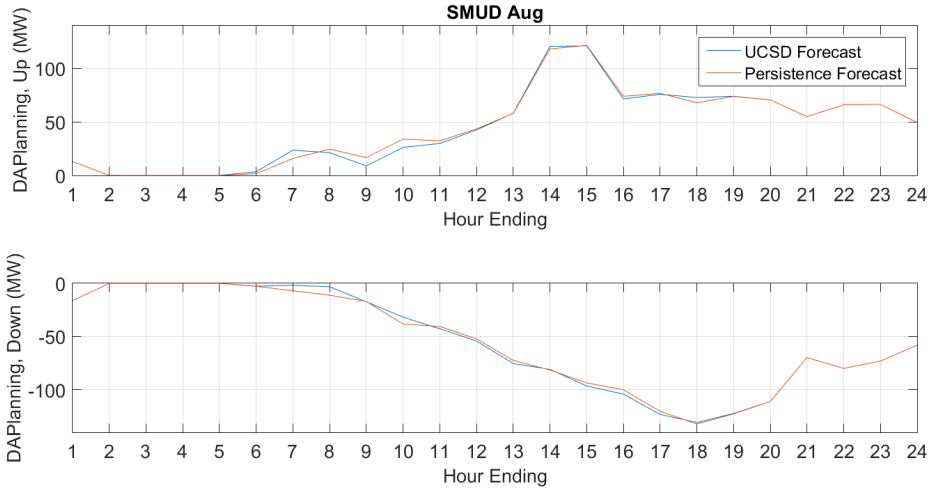

Figure 29: DA Planning Reserve Requirements in August

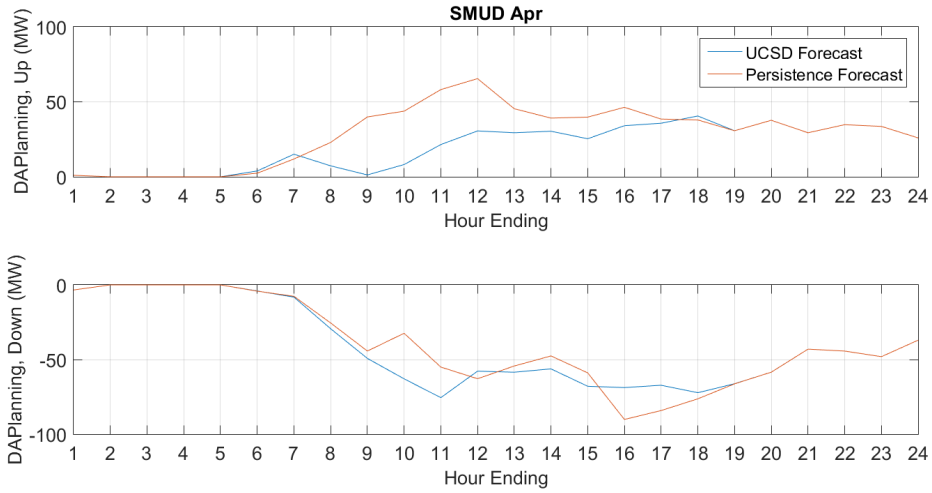

Figure 30: DA Planning Reserve Requirements in April

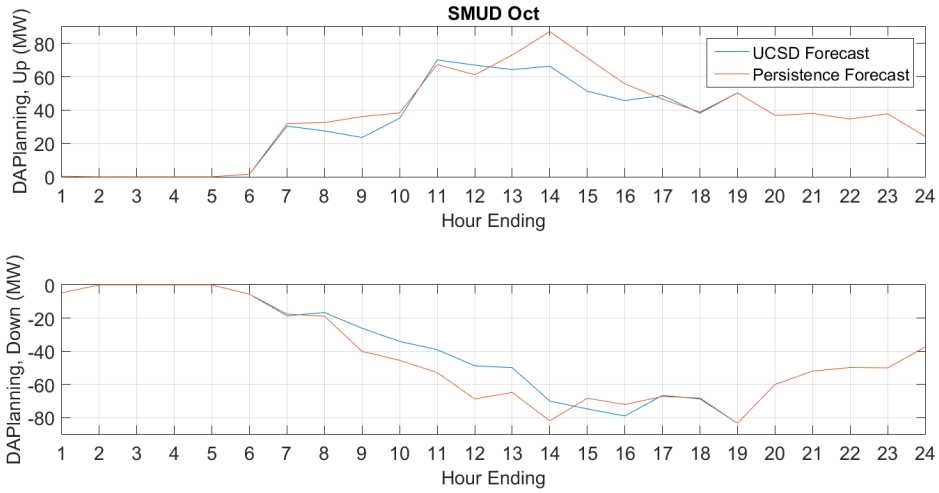

Figure 31: DA Planning Reserve Requirements in October

By comparing the DA PR requirements between the cases with UCSD forecast and persistence forecast, we can observe the following:

1. In both winter and shoulder seasons, UCSD PV forecast results in smaller DA PR requirements.

2. In the summer, DA PR requirements are almost the same for both cases, which may indicate weather in the summer is very stable from day to day and therefore, persistence model can get comparable performance with the UCSD model.

Overall effect of different forecasts on DA PR requirements throughout the year can be viewed from Figure 33. On average, UCSD PV forecast reduces DA PR Up requirements by 20 MW, and DA PR Down requirements by 10 MW in the daylight hours. Savings of DA PR Up requirements amount to 1663 MW-hour, and DA PR Down requirements 899 MW-hour for the entire year.

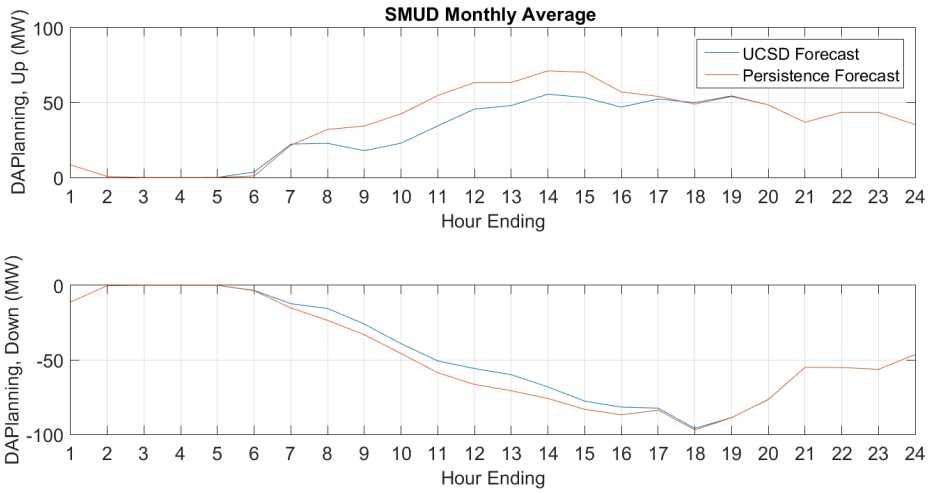

Figure 32: Monthly Average DA Planning Reserve Requirements

Regulation Reserve

This section provides example plots in January and August to illustrate regulation reserve requirements. Because in both base and change cases, the same persistence model was used for real-time PV forecast, the regulation requirements are the same for both cases. Note that when reserve requirements are smaller than L_{10} of SMUD system, they are shown as zero.

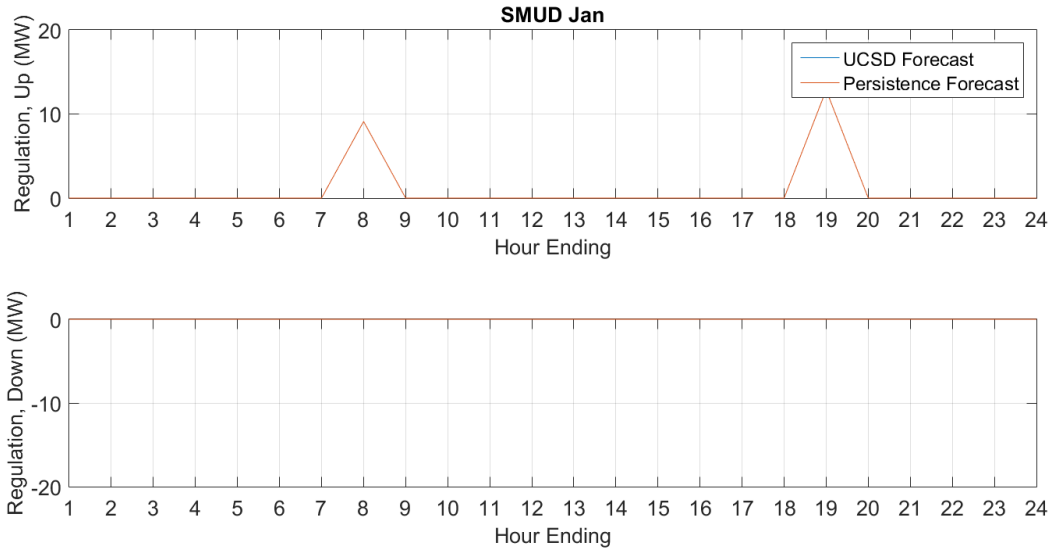

Figure 33: Regulation Reserve Requirements in January

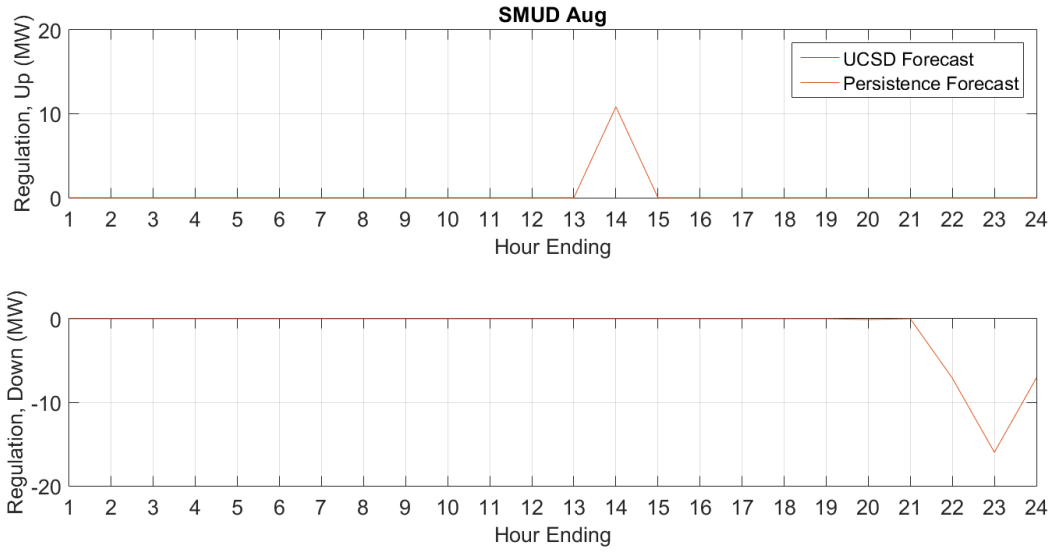

Figure 34: Regulation Reserve Requirements in August

Statistics of PV Forecast Performance

PV forecast errors were analyzed for both the PV forecast produced using the UCSD approach and using DA persistence model. The forecast errors were normalized by the maximum of output of each hour and then categorized by the clearness index (CI) of that hour. Table 12 shows the statistics derived from the persistence model and UCSD PV forecast. It can be observed from the table that when CI is

between 20% to 50%, persistence model is slightly better than UCSD forecast. In the other three CI categories, UCSD forecast is better than persistence model.

Table 12. PV Forecast Performance Statistics

SMUD	**DA Persistence Model**		**UCSD Forecast**	
Clearness Index	Mean	Std	Mean	Std
CI<=0.2	-0.10	0.20	-0.05	0.14
0.2<CI<=0.5	-0.06	0.24	-0.05	0.28
0.5<CI<=0.8	0.03	0.18	0.05	0.17
0.8<CI	0.04	0.14	0.04	0.11

Remarks on PV Forecast Impacts on Reserve Requirements

Compared to the reserve requirements derived from persistence forecast, on average, UCSD PV forecast reduces DA PR Up requirements by 20 MW, and DA PR Down requirements by 10 MW in the daylight hours. Savings of DA PR Up requirements amount to 1663 MW-hour, and DA PR Down requirements 899 MW-hour for the entire year.

Comparison of the performance statistics of the two forecasts shows that UCSD PV forecast has a few percentages improvement over DA persistence forecast.

Intra-hour Dispatch and AGC Simulations

Production cost simulations in the study involve two steps:

1. Build the SMUD system model in PLEXOS with “actual” load and PV data and create the hourly schedule of generation resources; reserve requirements calculated in Section 2.0 were used as inputs to the PLEXOS model;
2. Build the SMUD system model in ESIOS to perform real-time dispatch and AGC simulations; the hourly schedule from PLEXOS is used as input to the ESIOS model and ESIOS redispatch generation resources to follow real-time variations of load and PV.

If the generation schedules in the base case and change case are different because of different reserve requirements, impact on system control performance and production cost could be observed from EISOS runs.

In this study, however, in the step of PLEXOS run, it was found that the base case and the change case have exactly the same generation schedule, even though the reserve requirements in these two cases have 10-20 MW differences in the daylight hours. Because of this, the real-time dispatch and AGC performance in ESIOS would also be the same since PV and load data are the same in the two cases. No difference in production cost would be observed.

The reason for this result can be attributed to the small amount of reserve requirements for balancing the system, compared to the available resources in the SMUD system. Reserve requirements are not binding constraints in any case; therefore, they did not result in any difference in generation schedule.

Task 7: Dissemination of the work supported by this Project

In this task (same as task 4 but for the whole project) we collected and made available in the public domain, free of charge, all data that was generated as part of the scope of this project. The web portal can be accessed at: http://coimbra.ucsd.edu/projects/sunrise-ucsd/. In this website the public has access to the data used and created in this project, the list of publications and presentations that were generated with the support of this project and the real-time day-ahead GHI forecast for the 12 clusters in SMUD. The web page is organized in a simple and intuitive manner with three separate pages for each type of product.

Data

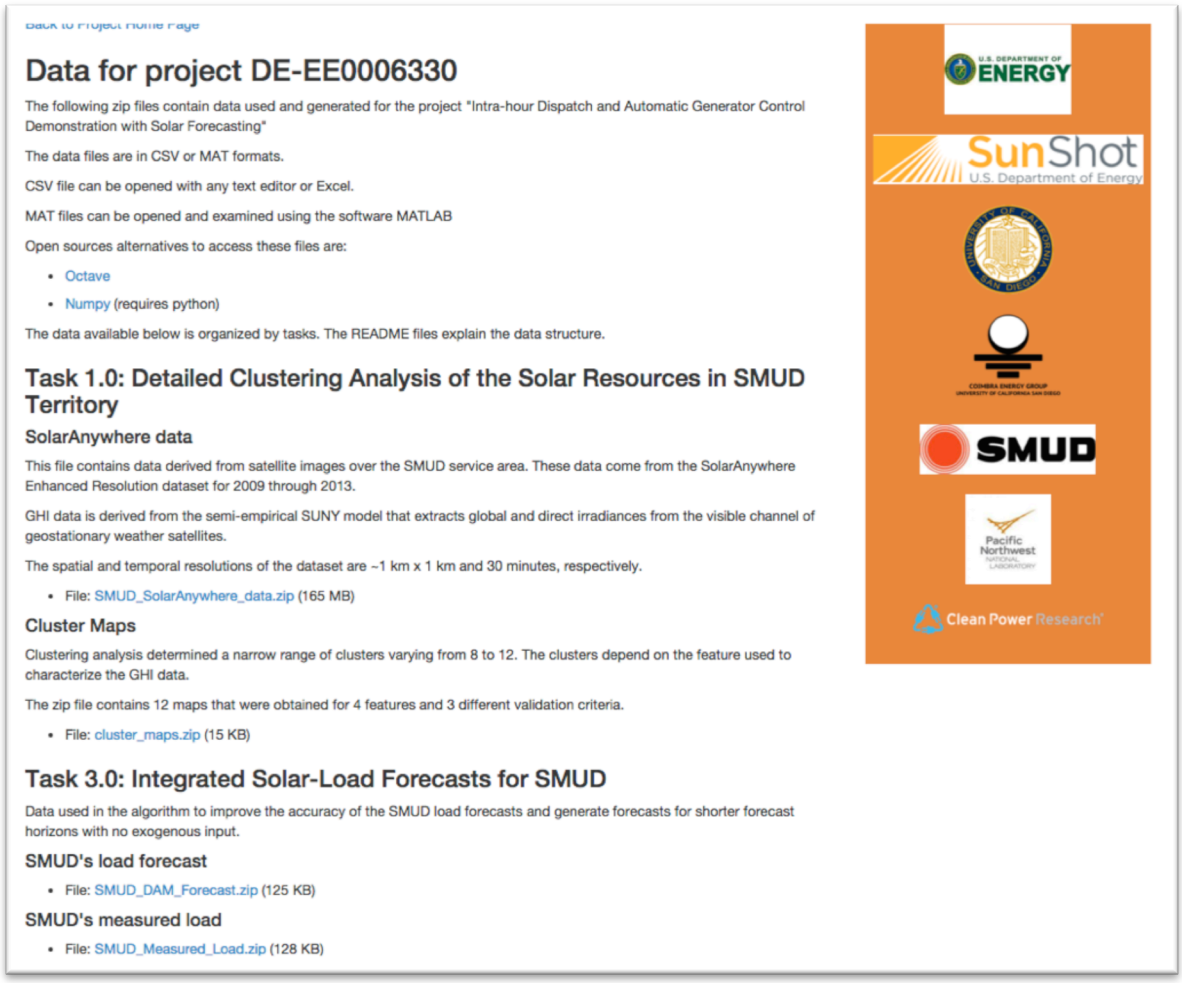

Figure 35: Screenshot of web portal that gives access to the data used and generated by this project.
URL: http://coimbra.ucsd.edu/projects/sunrise-ucsd/sunrise_data.html

Publications

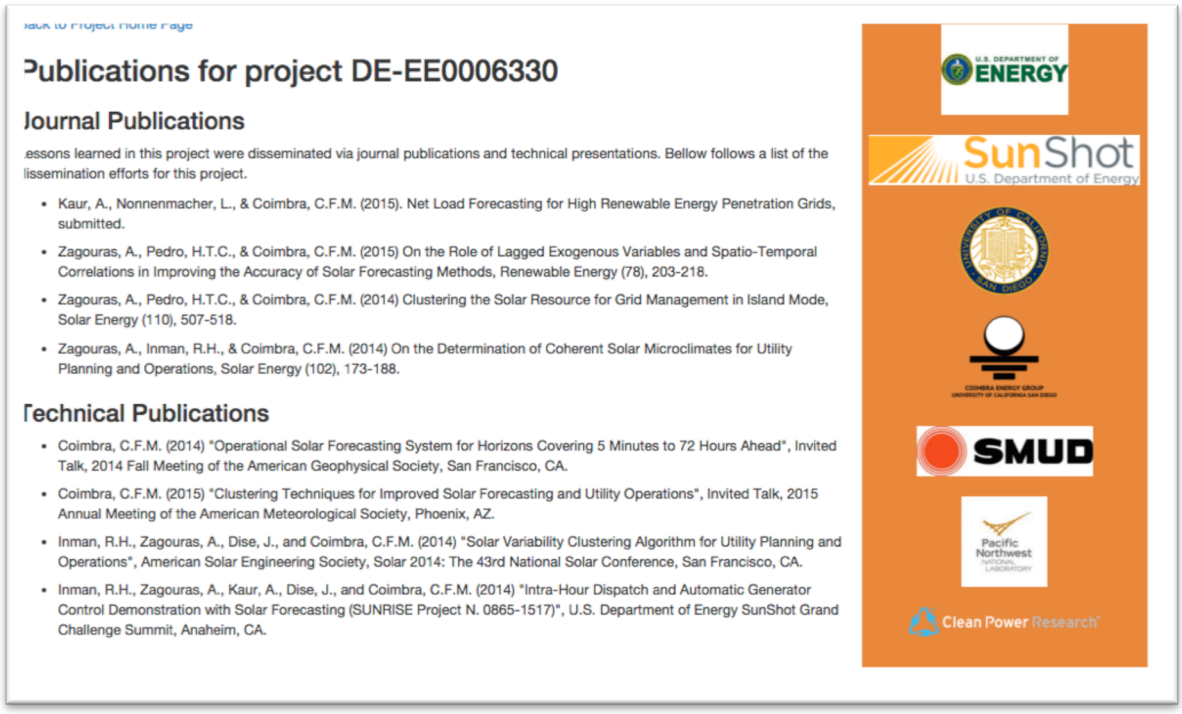

Figure 36 Screenshot of web portal that lists the presentations and journal publications supported by this project. URL: http://coimbra.ucsd.edu/projects/sunrise-ucsd/publications.html

Real-time GHI forecasting

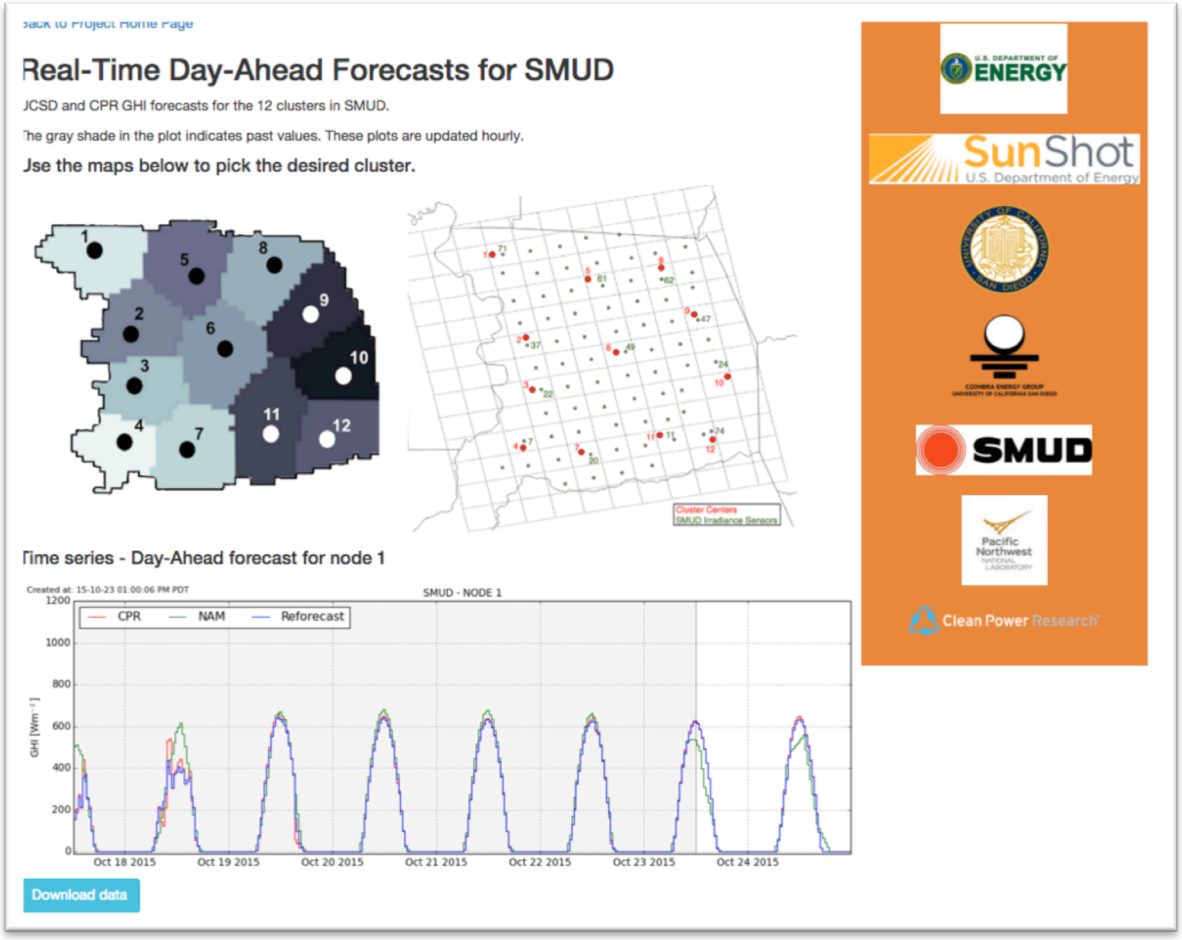

Figure 37: Screenshot of the real-time day-ahead GHI forecast for the 12 nodes in the SMUD region.
URL: http://coimbra.ucsd.edu/projects/sunrise-ucsd/sunrise_forecast.html

The optimized DA models obtained in task 5 were implemented in real-time in our servers at UCSD. The latest GHI forecasts for NAM. CPR and UCSD reforecast models are available at http://coimbra.ucsd.edu/projects/sunrise-ucsd/sunrise_forecast.html. The predictions for a given cluster are accessed by picking the corresponding node in the clustering map. The "Download data" button at the bottom downloads the data in *csv* format.

Other activities that disseminate the findings from this project consist of technical presentations and publication of peer-reviewed articles in the most impactful solar energy journals. Below follows the list of such efforts.

Journal Publications

- Kaur, A., Nonnenmacher, L., & Coimbra, C.F.M. (2015). Net Load Forecasting for High Renewable Energy Penetration Grids, submitted.
- Zagouras, A., Pedro, H.T.C., & Coimbra, C.F.M. (2015) On the Role of Lagged Exogenous Variables and Spatio-Temporal Correlations in Improving the Accuracy of Solar Forecasting Methods, *Renewable Energy* (78), 203-218.
- Zagouras, A., Pedro, H.T.C., & Coimbra, C.F.M. (2014) Clustering the Solar Resource for Grid Management in Island Mode, *Solar Energy* (110), 507-518.
- Zagouras, A., Inman, R.H., & Coimbra, C.F.M. (2014) On the Determination of Coherent Solar Microclimates for Utility Planning and Operations, *Solar Energy* (102), 173-188.

Technical Presentations

- Coimbra, C.F.M. (2014) "Operational Solar Forecasting System for Horizons Covering 5 Minutes to 72 Hours Ahead", Invited Talk, 2014 Fall Meeting of the American Geophysical Society, San Francisco, CA.
- Coimbra, C.F.M. (2015) "Clustering Techniques for Improved Solar Forecasting and Utility Operations", Invited Talk, 2015 Annual Meeting of the American Meteorological Society, Phoenix, AZ.
- Inman, R.H., Zagouras, A., Dise, J., and Coimbra, C.F.M. (2014) “Solar Variability Clustering Algorithm for Utility Planning and Operations”, American Solar Engineering Society, Solar 2014: The 43rd National Solar Conference, San Francisco, CA.
- Inman, R.H., Zagouras, A., Kaur, A., Dise, J., and Coimbra, C.F.M. (2014) “Intra-Hour Dispatch and Automatic Generator Control Demonstration with Solar Forecasting (SUNRISE Project N. 0865-1517)”, U.S. Department of Energy SunShot Grand Challenge Summit, Anaheim, CA.

Works cited

Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*,*46*(1), 243-256.

Boyle, G. (2012). *Renewable electricity and the grid: the challenge of variability*. Earthscan.

Cadima, J., & Jolliffe, I. (2009). On relationships between uncentred and column-centred principal component analysis. *Pak J Statist*, *25*(4), 473-503.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis.*Communications in Statistics-theory and Methods*, *3*(1), 1-27.

Celebi, M. E., Kingravi, H. A., & Vela, P. A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*, *40*(1), 200-210.

Chow, C. W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., & Washom, B. (2011). Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy*, *85*(11), 2881-2893.

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2), 224-227.

Diabate, L., Blanc, P., & Wald, L. (2004). Solar radiation climate in Africa. *Solar Energy*, *76*(6), 733-744.

Ding, C., & He, X. (2004, July). K-means clustering via principal component analysis. In *Proceedings of the twenty-first international conference on Machine learning* (p. 29). ACM.

Eltbaakh, Y. A., Ruslan, M. H., Alghoul, M. A., Othman, M. Y., & Sopian, K. (2012). Issues concerning atmospheric turbidity indices. *Renewable and Sustainable Energy Reviews*, *16*(8), 6285-6294.

Hagan, M. T., & Behr, S. M. (1987). The time series approach to short term load forecasting. *Power Systems, IEEE Transactions on*, *2*(3), 785-791.

Helman, U., Loutan, C., Rosenblum, G., Rothleder, M., Xie, J., & Zhou, H. (2010). Integration of renewable resources: Operational requirements and generation fleet capability at 20% rps. *California Independent System Operator*.

Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *Power Systems, IEEE Transactions on*, *16*(1), 44-55.

Ineichen, P. (2006). Comparison of eight clear sky broadband models against 16 independent data banks. *Solar Energy*, *80*(4), 468-478.

Ineichen, P., & Perez, R. (2002). A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy*, *73*(3), 151-157.

Inman, R. H., Pedro, H. T. C., & Coimbra, C. F. M. (2013). Solar forecasting methods for renewable energy integration. *Progress in energy and combustion science*, *39*(6), 535-576.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, *31*(8), 651-666.

Jolliffe, I. T. (1986). Principal Component Analysis, vol. 487Springer-Verlag.*New York*.

Kabalci, E. (2011). Development of a feasibility prediction tool for solar power plant installation analyses. *Applied Energy*, *88*(11), 4078-4086.

Lew, D., Brinkman, G., Ibanez, E., Hodge, B. M., & King, J. (2013). The western wind and solar integration study phase 2. *Contract*, *303*, 275-3000.

Liu, Y., Li, Z., Xiong, H., Gao, X., & Wu, J. (2010, December). Understanding of internal clustering validation measures. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on* (pp. 911-916). IEEE.

Ljung, L. (1999). System Identification: Theory for the user. *PTR Prentice Hall Information and System Sciences Series*, *198*.

Lloyd, S. P. (1982). Least squares quantization in PCM. *Information Theory, IEEE Transactions on*, *28*(2), 129-137.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, No. 14, pp. 281-297).

Marquez, R., & Coimbra, C. F. M. (2011). Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database. *Solar Energy*, *85*(5), 746-756.

Marquez, R., Pedro, H. T. C., & Coimbra, C. F. M. (2013). Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs. *Solar Energy*, *92*, 176-188.

Matthewman, P. D., & Nicholson, H. (1968). Techniques for load prediction in the electricity-supply industry. In *Proceedings of the Institution of Electrical Engineers* (Vol. 115, No. 10, pp. 1451-1457). IET Digital Library.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, *50*(2), 159-179.

Noy-Meir, I. (1973). Data transformations in ecological ordination: I. Some advantages of non-centering. *The Journal of Ecology*, 329-341.

Ortega-Vazquez, M. A., & Kirschen, D. S. (2006, June). Economic impact assessment of load forecast errors considering the cost of interruptions. In: *Power Engineering Society General Meeting, 2006. IEEE* (pp. 8-pp). IEEE.

Passino, K. M. (2005). *Biomimicry for optimization, control, and automation*. Springer Science & Business Media.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., & Vignola, F. (2002). A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, *73*(5), 307-317.

Remund, J., Wald, L., Lefèvre, M., Ranchin, T., & Page, J. (2003). Worldwide Linke turbidity information. In *ISES Solar World Congress 2003* (Vol. 400, pp. 13-p). International Solar Energy Society (ISES).

Reno, M. J., Hansen, C. W., & Stein, J. S. (2012). Global horizontal irradiance clear sky models: Implementation and analysis. *SANDIA report SAND2012-2389*.

Rodriguez, G. D. (2010, July). A utility perspective of the role of energy storage in the smart grid. In *Power and Energy Society General Meeting, 2010 IEEE*(pp. 1-2). IEEE.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, *20*, 53-65.

Salvador, S., & Chan, P. (2004, November). Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In *Tools with Artificial Intelligence, 2004. ICTAI 2004. 16th IEEE International Conference on* (pp. 576-584). IEEE.

Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. *Renewable and sustainable energy reviews*, *16*(2), 1223-1240.

Theodoridis, S., & Koutroumbas, K. (2001). Pattern recognition and neural networks. In *Machine Learning and Its Applications* (pp. 169-195). Springer Berlin Heidelberg.

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 411-423.

WECC (2007). WECC Standard BAL-STD-002-0–Operating Reserves.

Willis, H. L., & Northcote-Green, J. E. (1983). Spatial electric load forecasting: a tutorial review. *Proceedings of the IEEE*, *71*(2), 232-253.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1-37.

Xu, J., Xu, B., Zhang, W., Zhang, W., & Hou, J. (2009). Stable initialization scheme for k-means clustering. *Wuhan University Journal of Natural Sciences*,*14*(1), 24-28.

Zagouras, A., Kazantzidis, A., Nikitidou, E., & Argiriou, A. A. (2013). Determination of measuring sites for solar irradiance, based on cluster analysis of satellite-derived cloud estimations. Solar Energy, 97, 1-11.

Zagouras, A., Inman, R. H., & Coimbra, C. F. M. (2014a). On the determination of coherent solar microclimates for utility planning and operations. *Solar Energy*,*102*, 173-188.

Zagouras, A., Pedro, H. T. C., & Coimbra, C. F. M. (2014b). Clustering the solar resource for grid management in island mode. *Solar Energy*, *110*, 507-518.

Zhao, Q., Xu, M., & Franti, P. (2008, November). Knee point detection on Bayesian information criterion. In *Tools with Artificial Intelligence, 2008. ICTAI'08. 20th IEEE International Conference on* (Vol. 2, pp. 431-438). IEEE.