ANL/MCS-TM-305

Weather Forecast-Based Optimization of Integrated Energy Systems

Mathematics and Computer Science Division

About Argonne National Laboratory
Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see www.anl.gov.

Availability of This Report
This report is available, at no cost, at http://www.osti.gov/bridge. It is also available on paper to the U.S. Department of Energy and its contractors, for a processing fee, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone (865) 576-8401
fax (865) 576-5728
reports@adonis.osti.gov

Disclaimer
This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor UChicago Argonne, LLC, nor any of their employees or officers, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of document authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, Argonne National Laboratory, or UChicago Argonne, LLC.

ANL/MCS-TM-305

Weather Forecast-Based Optimization of Integrated Energy Systems

by
V.M. Zavala,[1,2] E.M. Constantinescu,[1] T. Krause,[2] and M. Anitescu[1]
[1]Mathematics and Computer Science Division, Argonne National Laboratory
[2]Chemical Sciences and Engineering Division, Argonne National Laboratory

February 2009

Contents

Abstract **1**

1 Executive Summary **2**

2 Introduction **3**

3 Optimization Framework **6**
3.1 Stochastic Dynamic Optimization 6
3.2 Economic Impact of Forecasting 10
3.3 Gaussian Process Modeling 18

4 Uncertainty Quantification **23**
4.1 Numerical Weather Prediction Model – WRF 23
4.2 Ensemble Approach to Uncertainty Quantification 24
4.3 Uncertainty Model for Radiation 27

5 Discussions and Future Work **29**
5.1 Possible Deployment Strategy 30
5.2 Scientific and Technological Advances Needed 30

Acknowledgments **30**

References **31**

Weather Forecast-Based Optimization of Integrated Energy Systems

Victor M. Zavala,[†§] **Emil M. Constantinescu,**[†]
Theodore Krause,[§] **and Mihai Anitescu**[†]

[†]Mathematics and Computer Science Division

[§]Chemical Sciences and Engineering Division

Abstract

In this work, we establish an on-line optimization framework to exploit detailed weather forecast information in the operation of integrated energy systems, such as buildings and photovoltaic/wind hybrid systems. We first discuss how the use of traditional *reactive* operation strategies that neglect the future evolution of the ambient conditions can translate in high operating costs. To overcome this problem, we propose the use of a supervisory dynamic optimization strategy that can lead to more *proactive* and cost-effective operations. The strategy is based on the solution of a receding-horizon stochastic dynamic optimization problem. This permits the direct incorporation of economic objectives, statistical forecast information, and operational constraints. To obtain the weather forecast information, we employ a state-of-the-art forecasting model initialized with real meteorological data. The statistical ambient information is obtained from a set of realizations generated by the weather model executed in an operational setting. We present proof-of-concept simulation studies to demonstrate that the proposed framework can lead to significant savings (more than 18 % reduction) in operating costs.

1 Executive Summary

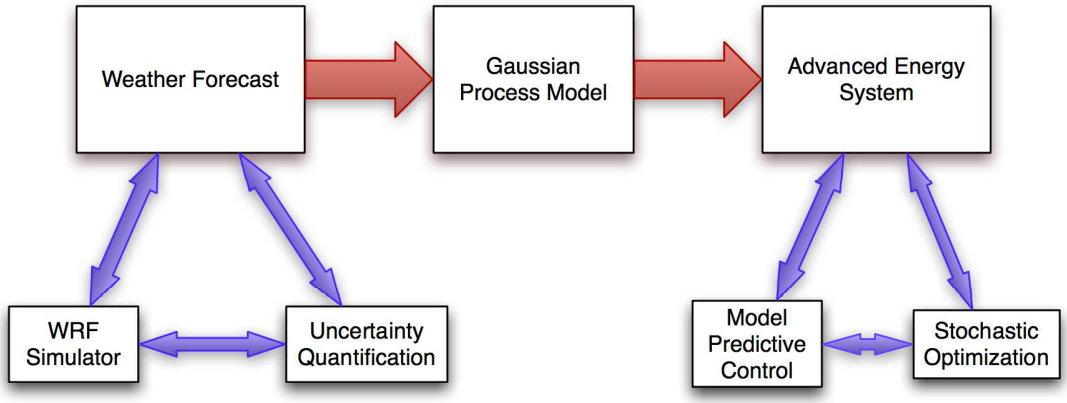


Figure 1: Scheme of our approach

Reduction of energy consumption costs while improving the standard of living, is an important national interest and thus a key research and development topic for all energy-related research. An important concern is that such a reduction may be too expensive or too difficult to deploy. Therefore, there is an increased interest in investigating "smart" technologies, those that efficiently incorporate available knowledge and information to reduce energy needs while providing the same or improved functionality.

In this paper, we demonstrate that incorporating forecasts and forecast uncertainty information in advanced integrated energy systems control techniques can result in significant cost and energy savings. Examples of such systems are hybrid energy systems for building climate control and systems which include photovoltaic and wind energy sources.

We use a nonlinear stochastic model predictive control technique that employs stochastic optimization to determine the optimal management schedule that preserves the functional objectives, minimizes the cost and accounts for the uncertainty in the environmental parameters, such as temperature and solar radiation.

Our approach has several layers, as shown in Figure 1. First, an uncertainty model, for the weather forecast, described in Section 4 is produced by enhancing the widely used Weather Research Forecast (WRF) code [30]. The model is then sampled in time at the coordinate points of the integrated energy system. The resulting samples are used to train a Gaussian process model of the uncertainty in the ambient parameters, as described in Section 3.3. Finally, a stochastic optimization–nonlinear model predictive control approach, described in Section 3, is solved by sample average approximation for determining the optimal functioning regime of the integrated energy system.

This initial study suggests that our approach can result in tenths of percents reduction in cost and/or energy consumption for the target integrated energy system, compared with some of the current control strategies (we have observed a 18% cost reduction in our

building model for standard insulation, as shown in Figure 12, and a 28 % cost reduction for improved insulation, as shown in Figure 13 and 75% reduction (Figure 7) in the operating costs of the photovoltaic hybrid energy system. Moreover, that implementation of our strategy requires minimal capital investment costs for the building model, beyond the computational effort (for example, a new climate control panel, which can be remotely controlled would be sufficient).

2 Introduction

During the past years, strong socio-economic pressures have forced diverse industrial sectors to reassess the efficiency of current energy production and consumption facilities. In particular, increasing fossil fuel prices and carbon emission taxes will require the consideration of more efficient designs able to accommodate multiple energy sources and operating strategies able to maximize their utilization in a cost-optimal manner. The design and operation of these integrated energy systems are complex tasks because of multiple physical phenomena arising in different units and their because of their strong dependency on *exogenous* disturbances such as the ambient conditions, time-varying demands, and time-varying fuel and electricity prices. In this context, the availability of powerful simulation technologies able to predict and assess the performance of these systems under a wide variety of operating environments will become increasingly important.

Rigorous simulation models for myriad energy systems have been developed over the past few years and have been used extensively for off-line design and retrofitting tasks. The availability of these models has led to more systematic practices and, consequently, to more cost-effective systems. Examples of available simulation software are: EnergyPlus and ADVISOR, developed at the NREL Laboratory to simulate the performance of building and hybrid vehicle systems [9; 22]; GCTool and PSAT, developed at Argonne National Laboratory to simulate hybrid vehicles and power train systems [13; 19]; and TRNSYS, developed at the University of Wisconsin to simulate a wide range of hybrid energy and building systems [18]. With this simulation technology at hand, some natural questions arise: Can we use these powerful models on-line to optimize the operation of energy systems? How can we integrate these simulators with available optimization technology? Can we handle highly uncertain and dynamic disturbances effectively?

The operation of industrial systems is normally decomposed in a hierarchical manner, as sketched in Figure 2. The high level is normally known as the *supervisory* or set-point optimization layer. At this level, the set-points are adjusted in order to optimize the system economic performance. In the context of energy systems, this level is also known as the centralized energy management system. The lower level is the *regulatory* control level that use available actuators to track the set-points dictated by the supervisory level. Most state-of-the-art energy simulation packages provide closed-loop simulation capabilities that can be used to design and test different operating strategies. These capabilities are based mostly on proportional-integral-derivative (PID) and *logic-based* controllers. PID controllers are limited to regulation tasks, while logic-based controllers can be used for both regulation and set-point optimization. Logic-based con-

trollers consist of a decision-making structure or tree designed to determine the controls as a function of the current outputs and exogenous disturbances. The decision-making structure and threshold values are tuned off-line using a simulation model in order to obtain a desired performance [34; 36]. The application of logic-based strategies is intuitive and can provide satisfactory results for regulatory control and basic optimization tasks. However, performing high-level supervisory optimization tasks can become cumbersome in large and tightly interconnected systems. The reason is that, as the amounts of information and number of decision variables grow, the logic structure becomes more and more complex and tuning the associated threshold values becomes time-consuming and impractical. In addition, once the logic-based controller is tuned by using the simulation model on a variety of scenarios, the model will no longer be used on-line. Consequently, it is difficult to guarantee adequate performance under unexpected conditions, and retuning might be necessary. Moreover, handling economic objectives and operational limits can become complicated under this framework.

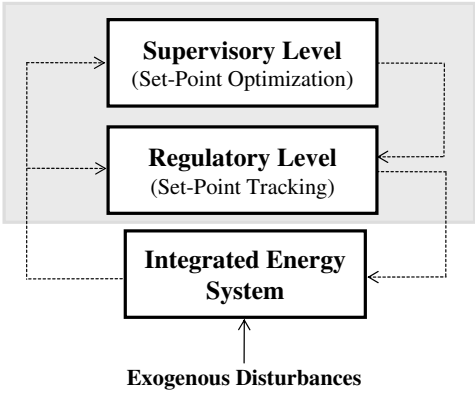


Figure 2: Schematic representation of operation hierarchy in industrial systems.

An alternative supervisory strategy is closed-loop real-time optimization (RTO) [23]. The idea is to use a *steady-state* rigorous model of the system and couple this to a large-scale optimization solver. The optimizer will determine the optimal set-points that maximize the system profit using the current information of the exogenous disturbances. Note that, since the set-points need to be satisfy the operational limits in the real system, a precise rigorous model is needed. The RTO output set-points are passed to a set of lower-level controllers that bring the system to the optimal steady-state. Once this is accomplished, the set-points are recomputed by RTO using the updated disturbance information. An advantage of this framework is that economic objectives and operational limits can be handled directly by the optimizer in a systematic manner. In addition, the rigorous model is always used and adapted on-line. Consequently, tuning tasks are significantly reduced. More important, it is always possible to guarantee that the system is at an optimal operating point. As can be seen, RTO offers significant advantages over logic-based strategies, specially in highly complex systems. Note also that, since

time-varying factors such as energy prices or weather conditions can be seen as dynamic disturbances that move the system away from the optimal operating point, RTO can be seen as a closed-loop optimizer that rejects these disturbances and keeps track of the maximum system performance. As a result of these desirable economic adaptation features, RTO capabilities are now widely used in conjunction with chemical process simulators such as AspenPlus© and ROMeo©. This technology has generated millions of dollars in annual savings in diverse sectors of the chemical industry [38]. Nevertheless, an important limitation of RTO and logic-based strategies is that they are entirely *reactive*, in the sense that only current disturbance information is used to make decisions. This feature can limit their ability to handle highly dynamic disturbances efficiently. For instance, if the set-points are updated at a higher frequency than the controller settling time, erratic performance and instabilities can arise. This represents an important limitation in the context of integrated energy systems where performance strongly depends on transient disturbances.

In this work, we propose to use a dynamic real-time optimization (D-RTO) strategy to perform economic supervisory decisions in integrated energy systems. The idea is in principle similar to RTO, but the key difference is that a rigorous *dynamic* model is used to compute future dynamic set-point trajectories [15]. We show that this strategy permits a consistent handling of highly dynamic disturbances and can directly incorporate forecast information. This adds *proactiveness* to the control actions and reduce costs. In particular, we show that incorporating weather forecast information can translate in significant savings in energy systems. Nevertheless, we demonstrate that using accurate forecasts and uncertainty information is critical to achieve a reliable system performance. To obtain this information, we first propose to construct data-based autoregressive models using a Gaussian process modeling technique. While this strategy is useful to obtain quick estimates of certain weather conditions and related uncertainty information, it is limited to short-term forecasts and can give rise to conservative uncertainty bounds. Therefore, we explore the potential of using detailed weather models. From an operational point of view, these models are attractive because they can provide comprehensive information such as spatio-temporal fields of ambient temperature, solar radiation and humidity. This information can be fully exploited by the rigorous model embedded within the D-RTO strategy. However, a limitation of current weather models is that the uncertainty information is limited and/or in forms that are inconsistent with existing optimization technology. To this end, we develop simplified uncertainty models for ambient variables that rely on model dynamics and only require a small amount of empirical assumptions. The weather model is driven in an operational setting with real data, and thus provides realistic and attainable estimates on the uncertainty found in the meteorological fields. We argue that connecting these powerful weather prediction models with modeling and optimization capabilities has the potential of achieving unprecedented energy utilization efficiencies and cost reductions in diverse industrial and residential sectors. We present simulation studies using a photovoltaic-hydrogen hybrid energy system and a building system to illustrate the concepts.

3 Optimization Framework

In this section, we derive the basic components of the proposed closed-loop dynamic optimization (D-RTO) framework, explain its advantages over steady-state optimization (RTO), and discuss extensions to consider stochastic disturbance information. We then illustrate the economic impact of folding weather forecast information in the operation of a photovoltaic-hydrogen hybrid energy system and a building system. Finally, we illustrate the use of Gaussian process modeling technique to obtain on-line forecast information, and we explain how this information can be connected to the optimization framework.

3.1 Stochastic Dynamic Optimization

To start the discussion, we consider a differential-algebraic equation (DAE) model of the form

$$\frac{dz}{d\tau} = \mathbf{f}\left(z(\tau), y(\tau), u(\tau), \chi(\tau)\right) \tag{1a}$$

$$0 = \mathbf{g}(z(\tau), y(\tau), u(\tau), \chi(\tau)) \tag{1b}$$

$$z(0) = x_k, \tag{1c}$$

where τ is the *model* time dimension and t_k is the current time in the *real* system. Variables $z(\tau)$ are differential states, $y(\tau)$ are algebraic states, $u(\tau)$ are the controls or manipulated variables, and $\chi(\tau)$ are the exogenous disturbances. In this context, the term *exogenous* refers to the fact that the disturbances are not affected by the system variables (e.g., energy prices). The differential equations (1a) represent conservation equations (energy, mass, and momentum), while the algebraic equations (1b) represent consistency conditions and expressions to calculate physico-chemical properties. The initial conditions at time t_k are given by the current state of the system x_k. Starting from this state and using a set of future control and disturbance trajectories, we can predict the evolution of the system. With these predictive capabilities, it is possible to formulate a dynamic optimization problem of the form:

$$\min_{u(\tau)} \int_{t_k}^{t_k+T} \varphi(z(\tau), y(\tau), u(\tau), \chi(\tau)) d\tau \tag{2a}$$

$$\frac{dz}{d\tau} = \mathbf{f}\left(z(\tau), y(\tau), u(\tau), \chi(\tau)\right) \tag{2b}$$

$$0 = \mathbf{g}(z(\tau), y(\tau), u(\tau), \chi(\tau)) \tag{2c}$$

$$0 \geq \mathbf{h}(z(\tau), y(\tau), u(\tau), \chi(\tau)) \tag{2d}$$

$$z(t_k) = x_k, \quad \tau \in [t_k, t_k + T], \tag{2e}$$

where T is the length of the prediction or forecast horizon. The objective function (2a) represents the system operational costs accumulated over the future horizon (e.g. heating/cooling utilities). The inequality constraints (2d) are used to represent operational limits (e.g. temperature, pressure, and voltage levels). The dynamic optimization problem is infinite-dimensional because it depends on time, which is a continuous parameter.

This problem can be approximated by a finite-dimensional nonlinear programming (NLP) problem through discretization techniques. Note that any partial differential equation (PDE) model can also be represented in DAE form through discretization along the space dimensions.

To formulate the optimization problem at time t_k, we know the value of the current disturbance $\chi(t_k)$ but the future disturbance trajectory is unknown. Nevertheless, we know that the disturbances belong to an uncertain space $\chi(\tau) \in \Omega_k, \tau \in [t_k, t_k + T]$ that we can approximate. To do so, we make use of *past* disturbance information $\chi(\tau), \tau \in [t_k - N, t_k]$ and a suitable forecast model. The forecast model can be either data-based or physics-based. In any case, we can assume that the model provides a predictive mean $\bar{\chi}(\tau)$ and that the prediction errors follow a normal or Gaussian distribution such that $\chi(\tau) = \mathcal{N}(\bar{\chi}(\tau), \mathbf{V}(\tau))$, where $\mathbf{V}(\tau)$ is the covariance matrix. With this, the uncertain space adopts an ellipsoidal form

$$\Omega_k := \{\, z \,|\, (z - \bar{\chi}(\tau))^T \mathbf{V}^{-1}(\tau)(z - \bar{\chi}(\tau)) \leq \alpha \}, \tag{3}$$

where α represents an appropriate confidence level. This uncertainty region is sketched in Figure 3. Under these assumptions, all that is needed to represent the uncertain space is the predictive mean and the covariance matrix. However, the proposed structure of the uncertainty space is a modeling assumption and hence might not be accurate. Nevertheless, from practical point of view, what we seek from the approximate space is that it can encapsulate the true disturbance realization and that it has a physically meaningful structure.

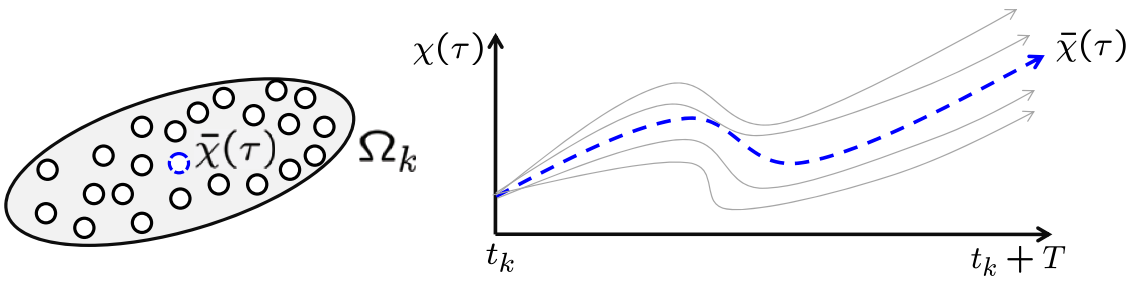


Figure 3: Schematic representation of ellipsoidal uncertainty region.

To exploit the entire statistical information at hand, we formulate an stochastic dynamic optimization problem of the form

$$\min_{u(\tau)} \mathop{\mathbf{E}}_{\chi(\tau)\in\Omega_k} \left[\int_{t_k}^{t_k+T} \varphi(z(\tau), y(\tau), u(\tau), \chi(\tau)) d\tau \right] \tag{4a}$$

$$\frac{dz}{d\tau} = \mathbf{f}\left(z(\tau), y(\tau), u(\tau), \chi(\tau)\right) \tag{4b}$$

$$0 = \mathbf{g}(z(\tau), y(\tau), u(\tau), \chi(\tau)) \tag{4c}$$

$$0 \geq \mathbf{h}(z(\tau), y(\tau), u(\tau), \chi(\tau)) \tag{4d}$$

$$z(t_k) = x_k, \quad \tau \in [t_k, t_k + T], \quad \chi(\tau) \in \Omega_k. \tag{4e}$$

where symbol $\mathbf{E}[\cdot]$ denotes the expectation operator. From the solution of this problem, we obtain the state and control trajectories $z^*(\tau), y^*(\tau), u^*(\tau), \tau \in [t_k, t_k + T]$ that we can send to a lower-level controller as set-points. The controller will try to keep the system at the recommended target. At the next time step t_{k+1}, we obtain the updated state of the system x_{k+1} and the updated forecast disturbance information Ω_{k+1} that we use to solve the next stochastic problem (4). In this way, feedback is introduced. The resulting *closed-loop* D-RTO strategy is as follows

1. **Current State and Forecast:** At time t_k, obtain current state x_k, disturbance mean $\bar{\chi}(\tau)$ and associated covariance matrix $\mathbf{V}(\tau), \tau \in [t_k, t_k + T]$.

2. **Compute Set-Points:** Solve stochastic optimization problem (4). Send optimal set-points $z^*(\tau), y^*(\tau), u^*(\tau), \tau \in [t_k, t_k + T]$ to low-level control layer.

3. **Update:** At $t_k + \Delta$, set $k \leftarrow k + 1$ and repeat process.

In Figure 4, we sketch this conceptual closed-loop optimization framework and its interaction with the low-level control layer and the forecasting capability. Note that the disturbance forecast information needs to be updated as frequently as possible in order avoid deterioration of the economic performance. In principle, we could use only the predictive mean $\bar{\chi}(\tau)$ without taking uncertainty into account and solve the deterministic problem (2). With this approach, however, we cannot guarantee satisfaction of the operational constraints. Note also that, in the presence of uncertainty, the cost function becomes a probability distribution because it depends on all the possible realizations of the disturbances. Therefore, optimizing a single instance of the cost function is meaningless. In the above formulation, we assumed that the mean of the objective distribution is an adequate measure of the performance of the system. However, this need not be the case. For instance, we could also choose the mean-risk approach of Markowitz where we seek to minimize simultaneously the mean and the variance of the cost distribution. In this stochastic optimization framework, the structure of the cost function becomes a design task because it is entirely problem dependent.

The stochastic dynamic formulation is significantly more computationally demanding than the deterministic dynamic optimization formulation. The reason is that the uncertain space Ω_k is also infinite dimensional. To solve the stochastic optimization problem, we propose to use a sample-average approximation (SAA) approach. The idea is to obtain independent samples from the disturbance distribution to obtain a N_s realizations $\{\chi_1(\tau), \chi_2(\tau), ..., \chi_{N_s}(\tau)\}$. The samples are illustrated in Figure 3. With this, the approx-

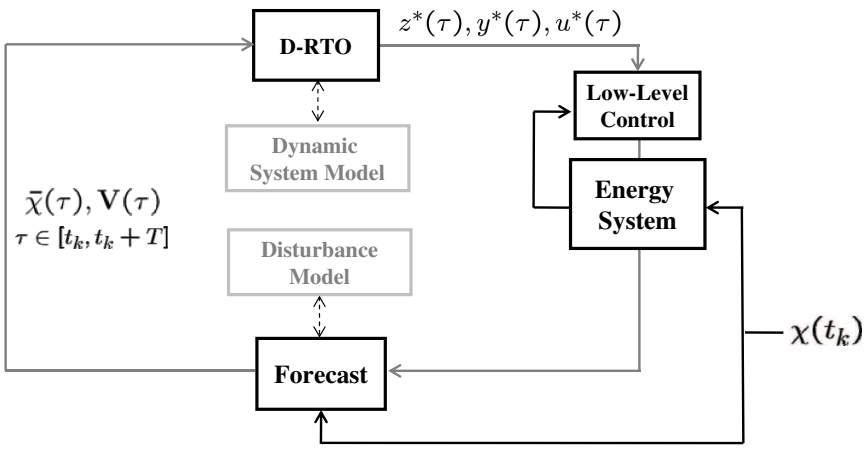


Figure 4: Structure of closed-loop stochastic optimization framework.

imate stochastic problem becomes

$$\min_{u(\tau)} \frac{1}{N_s} \sum_{j=1}^{N_s} \left[\int_{t_k}^{t_k+T} \varphi(z_j(\tau), y_j(\tau), u(\tau), \chi_j(\tau)) d\tau \right] \tag{5a}$$

$$\frac{dz_j}{d\tau} = \mathbf{f}\left(z_j(\tau), y_j(\tau), u(\tau), \chi_j(\tau)\right) \tag{5b}$$

$$0 = \mathbf{g}(z_j(\tau), y_j(\tau), u(\tau), \chi_j(\tau)) \tag{5c}$$

$$0 \geq \mathbf{h}(z_j(\tau), y_j(\tau), u(\tau), \chi_j(\tau)) \tag{5d}$$

$$z_j(t_k) = x_k, \quad \tau \in [t_k, t_k + T], \quad j = 1, ..., N_s. \tag{5e}$$

In this formulation, all the variables become a function of the particular disturbance realization except the controls, which are decision variables. One of the key advantages of the SAA approach is that it is straightforward to implement. Moreover, it is particularly suitable for large-scale systems, because it gives rise to highly structured problems [21]. This structure can be exploited in parallel computer architectures. The theoretical properties of the SAA approach have been widely studied in the context of nonlinear programming. For instance, it has been shown that solutions of the SAA problem converge at an exponential rate to the solution of the stochastic counterpart [29]. Although no formal convergence results exist in the context of infinite-dimensional dynamic optimization problems, we can expect that the available convergence guarantees can be used under some mild assumptions. Note that in the SSA approach the forecast capability can send the disturbance samples directly to the D-RTO component instead of the full covariance matrix. This can be useful in large-scale systems because the forecast capability can run in a centralized manner (consider a large weather model) and send the disturbance information to multiple D-RTO agents running on smaller dedicated machines.

If we shrink the prediction horizon of the stochastic dynamic optimization problem (2)

to zero, $T \to 0$, we recover the steady-state RTO problem:

$$\min_{u} \quad \varphi(z, y, u, \chi(t_k)) \tag{6a}$$
$$0 = \mathbf{f}(z, y, u, \chi(t_k)) \tag{6b}$$
$$0 = \mathbf{g}(z, y, u, \chi(t_k)) \tag{6c}$$
$$0 \geq \mathbf{h}(z, y, u, \chi(t_k)). \tag{6d}$$

This strategy finds the steady-state economic optimal operating point based only on the current disturbances $\chi(t_k)$, that are known. With this, we no longer rely on any forecasting mechanism, and the problem reduces to a finite-dimensional NLP problem that is significantly less computationally expensive. While all these seem to be practical advantages, in the next section we will show that strong economic penalties can be incurred by making these simplifications.

3.2 Economic Impact of Forecasting

In this section, we discuss some of the advantages of folding forecast information in operations. To do so, we present closed-loop D-RTO simulation studies on a photovoltaic-hydrogen (PV-H_2) hybrid energy system and on a building system. Our objective is to illustrate how the use of forecast information can add proactiveness to the D-RTO strategy and how this translates in lower operating costs.

Photovoltaic-Hydrogen Hybrid Energy System

PV-H_2 systems have been recently identified as a promising energy storage alternative. They are particularly attractive for remote areas where grid connections are expensive. In addition, they can be combined with wind power and diesel generators to provide back-up power in commercial buildings and industrial facilities. Economic studies have been performed in the U.S. by NREL [17; 24] and some prototypes have already been built in Germany, Switzerland and Norway [33]. The system consists of a set of photovoltaic arrays that generate electric power from the solar radiation. The radiation follows strong dynamic trends occurring at different time-scales (e.g. day-by-day and seasonal). In Figure 5, we present solar radiation data at position 41 59'N/87 54'W in the Chicago, IL, area for year 2004. The data are obtained from the National Solar Radiation Data Base [26].

The structure of the particular system under consideration has been obtained from [33] and is sketched in Figure 6. The available solar power from the arrays is used to satisfy a given user load. Any excess power can be stored in a battery system or can be used to produce hydrogen by water electrolysis. Hydrogen acts as an energy carrier that can be stored in pressurized tanks or in a solid material. When the solar power is not sufficient to satisfy the load, the deficit can be covered by the battery or from the stored hydrogen. If hydrogen is needed, this can be converted back to electric power through a fuel cell system. More details can be found in [39]. The dynamic model used in this study comprises a system-wide power balance motivated from [1; 20]. The power entering

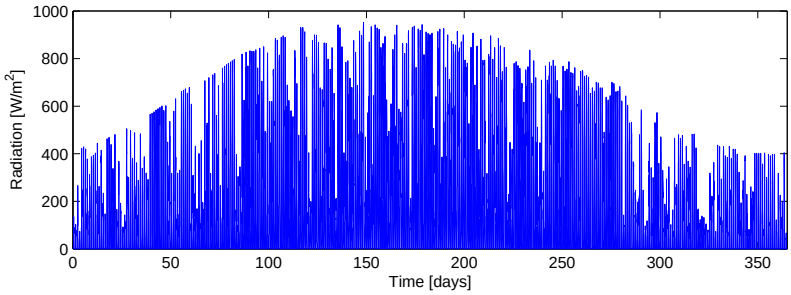


Figure 5: Total solar radiation in Chicago, IL, 2004.

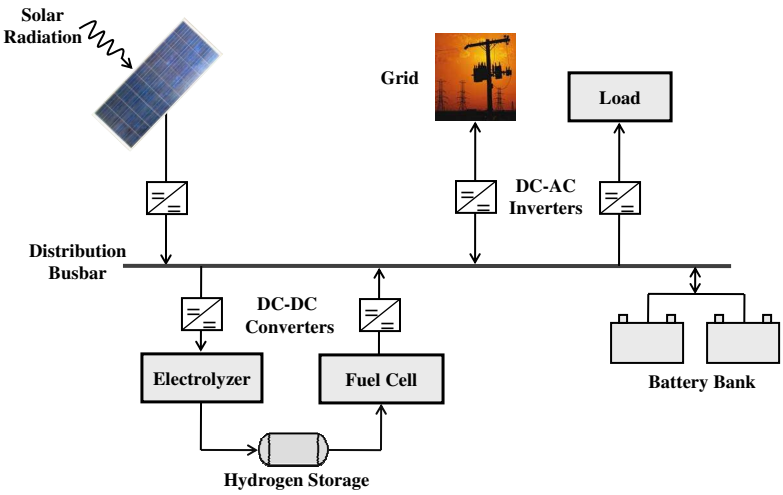


Figure 6: Schematic representation of photovoltaic-hydrogen hybrid system

through the solar module at a particular time τ is denoted by $P_{PV}(\tau)$ (kW). This can be calculated by using the measured radiation $G(\tau)$ (kW/m^2) and the module design characteristics. The electric current goes through a DC-DC converter that seeks to match the electric current voltage to the voltage of the distribution busbar. This conditioning process has an inherent efficiency θ_{PV} and generates power losses. The remaining power $\theta_{PV}P_{PV}(\tau)$ is sent to the busbar to satisfy the current load $P_{load}(\tau)$. The excess power can be used to produce hydrogen in the electrolyzer and/or to charge the battery. In order to run the electrolyzer, the power extracted $P_{EL}(\tau)$ passes through a buck DC-DC converter, which brings the current voltage down to the operating voltage of the electrolyzer. The efficiency of this step is θ_{BU}. The remaining power $\theta_{BU}P_{EL}(\tau)$ enters the electrolyzer. The conversion process to hydrogen has an efficiency θ_{EL}. The net price for each power unit produced by the electrolyzer is given by C_{EL}. Since hydrogen can be seen as an asset, the net price can be negative. The produced power $\theta_{BU}\theta_{EL}P_{EL}(\tau)$ in the form of hydrogen is

stored in a storage system modeled by the differential equation

$$\frac{dE_{H_2}}{d\tau} = \theta_{BU}\theta_{EL}P_{EL}(\tau) - P_{FC}(\tau), \tag{7}$$

where $E_{H_2}(\tau)$ is the total energy stored (kWh) at time τ. The hydrogen state of charge is defined as

$$SOC_{H_2}(\tau) = 100\frac{E_{H_2}(\tau)}{E_{H2}^{max}}, \tag{8}$$

where E_{H2}^{max} is the nominal maximum capacity (kWh). A certain amount of power $P_{FC}(\tau)$ can be withdrawn from the storage to feed a fuel cell and generate electric power. The conversion process has an efficiency θ_{FC}. The cost for each unit of power produced by the fuel cell is given by C_{FC}. The remaining power is then passed through a boost DC-DC converter that brings the voltage of the current up to the operating voltage of the busbar. The process has an efficiency θ_{BO}. The remaining power $\theta_{FC}\theta_{BO}P_{FC}(\tau)$ is sent to the distribution busbar. The system might be able to buy a given amount of power $P_G(\tau)$ from the grid in order to balance the system. This power will have a cost C_G that depends on the location and the degree of independence required by the application. In this study, we assume that the hybrid system is grid independent. Accordingly, we set $C_G = \infty$. In other words, if the demand cannot be met by the available resources, a high penalty cost will be paid. Excess power at the busbar can also be dumped to the grid or environment, which is modeled by variable $P_D(\tau)$. The cost of dumped power is C_D . If the power is dumped to the grid, this cost becomes an asset (set by net-metering rates). The power remaining at the busbar can be used to either charge or discharge the battery. The net battery power $P_B(\tau)$ is calculated by the balance at the busbar,

$$P_B(\tau) = \theta_{PV}P_{PV}(\tau) + P_G(\tau) + \theta_{FC}\theta_{BU}P_{FC}(\tau) - P_{EL}(\tau) - P_{load}(\tau) - P_D(\tau). \tag{9}$$

The stored energy in the battery is given by

$$\frac{dE_B}{d\tau} = P_B(\tau), \tag{10}$$

and the state-of-charge is,

$$SOC_B(\tau) = 100\frac{E_B(\tau)}{E_B^{max}}. \tag{11}$$

The fixed model inputs are $P_{PV}(\tau)$ and $P_{load}(\tau)$. The degrees of freedom (controls) are $P_{EL}(\tau)$, $P_{FC}(\tau)$, $P_G(\tau)$, and $P_D(\tau)$. Although this is a relatively simple hybrid, the operating decisions are complicated. For instance, the round-trip efficiency of the hydrogen storage system can be 70% or less, while that of the battery is around 90%. Therefore, larger amounts of power are lost if energy is decided to be stored as hydrogen. In addition, operating costs associated with the electrolyzer and the fuel cell must be taken into account (e.g., water and oxygen supplies). On the other hand, the battery capacity is usually small, giving a few days of autonomy. Consequently, it might not be immediately evident which component is the optimal one to store and provide energy at a particular time. The

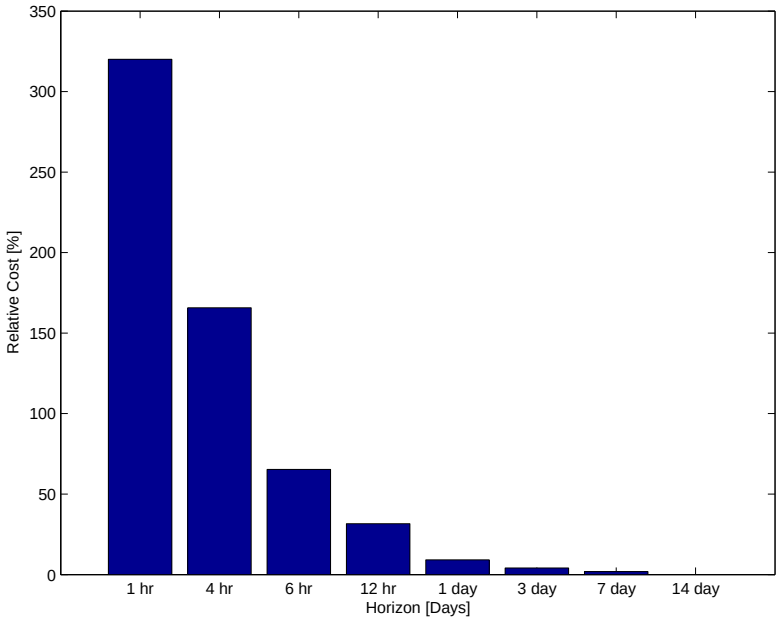


Figure 7: Impact of forecast horizon on economic performance of PV-H_2 system.

operational problems can become much more complicated if the system is connected to other devices with different efficiencies and operating costs, such as thermal solar plants, wind turbines, and diesel generators.

The operational objectives in these systems are to maximize hydrogen production, minimize operating costs, and satisfy the loads at all times. To obtain a reference for the best possible economic performance of the PV-H_2 system, we solve a dynamic optimization problem with *perfect* solar radiation information and a prediction horizon of 8760 hours (one year). The problem formulation is given in equation (12). In the following, we refer to this strategy as the open-loop strategy. The electrolyzer conversion efficiency was assumed to be 75%, while that of the fuel cell was assumed to be 50%. Both efficiencies were obtained from [36]. The unitary operating cost for the electrolyzer was assumed to be 7.4 \$/kg-H_2 and was obtained from [31]. For simplicity, we assumed equal fuel cell operating costs. A constant load of 1 kW and a maximum peak PV power of 5 kW were assumed. To solve this problem, we apply an implicit Euler discretization approach, implement the resulting NLP problem in AMPL [11], and solve it with the state-of-the-art optimization solver KNITRO© [5]. The solution of this problem provides the optimal plan for the electrolyzer, fuel cell, and battery powers that satisfy the load in a cost-optimal manner. The minimum operating costs for the electrolyzer and fuel cell are on the order of \$1,000. The total hydrogen produced is 380 Nm^3 which is equivalent to 1140 kWh. This is enough hydrogen to fulfill a load of 1 kW continuously for 45 days. Since the open-loop strategy uses all the available radiation information over one year, we use the minimum cost as a reference for the best achievable economic performance of the system. We then evaluate the performance of closed-loop D-RTO strategies with perfect forecast radiation

information with variable prediction horizons. This provides a measure of the economic impact of adding more and more forecast information to make the control decisions. The horizons considered are 1 hr, 3 hr, 6 hr, 9 hr, 12 hr, 1 day, 3 days, 7 days, and 14 days. From each scenario, we compute the relative costs using the 1-year horizon cost as reference. The scenario corresponding to 1 hr represents the economic performance of a purely reactive strategy (such as closed-loop RTO).

$$\min_{\substack{P_{EL}(\tau), P_{FC}(\tau),\\ P_G(\tau), P_D(\tau)}} \int_{t_k}^{t_k+T} \left[C_{EL}P_{EL}(\tau) + C_{FC}P_{FC}(\tau) + C_G P_G(\tau) + C_D P_D(\tau)\right] d\tau$$

$$\frac{dE_{H_2}}{d\tau} = \theta_{BU}\theta_{EL}P_{EL}(\tau) - P_{FC}(\tau) \tag{12a}$$

$$SOC_{H_2}(\tau) = 100\frac{E_{H_2}(\tau)}{E_{H2}^{max}} \tag{12b}$$

$$P_B(\tau) = \theta_{PV}P_{PV}(\tau) + P_G(\tau) + \theta_{FC}\theta_{BU}P_{FC}(\tau) - P_{EL}(\tau) - P_{load}(\tau) - P_D(\tau) \tag{12c}$$

$$\frac{dE_B}{d\tau} = P_B(\tau) \tag{12d}$$

$$SOC_B(\tau) = 100\frac{E_B(\tau)}{E_B^{max}} \tag{12e}$$

$$0 \leq SOC_{H_2}(\tau) \leq 100 \tag{12f}$$

$$0 \leq SOC_B(\tau) \leq 100 \tag{12g}$$

$$E_{H_2}(t_k) = E_{H2}^k \tag{12h}$$

$$E_B(t_k) = E_B^k, \tau \in [t_k, t_k + T] \tag{12i}$$

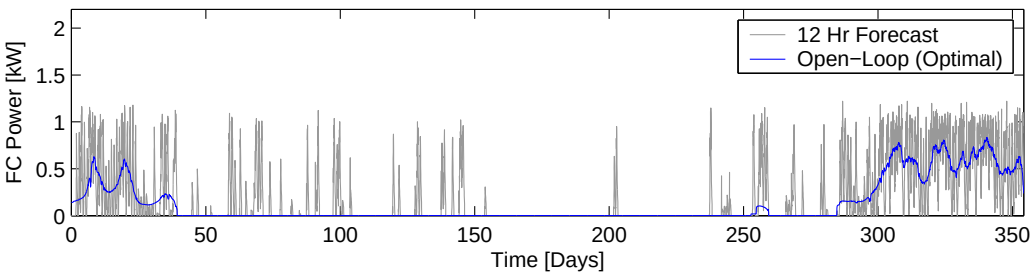


Figure 8: Impact of forecast horizon on fuel cell power profiles.

The first results of this economic study are presented in Figure 7. Several interesting and unexpected conclusions can be drawn: (1) the relative operating costs decay quickly to zero as the horizon is increased; (2) for a purely reactive RTO strategy (1 hr), the relative costs can go as high as 300%; and (3) the minimum cost can be obtained with a short forecast (14 days). This last result has important practical implications because it

implies that nearly optimal performance can be guaranteed for relatively short horizons. The economic penalty of using a forecast of 1 day is just an increase of 10% in relative costs, whereas the penalty for a forecast of 12 hr goes up to 31%. This implies that a practical forecast horizon should be sufficiently long to capture the periodicity of the daily radiation. The reason for these strong economic penalties becomes evident from Figure 8. Here, we present the power profiles for the fuel cell for both the optimal open-loop and the 12 hr closed-loop cases. Note that shorter forecasts induce much more aggressive control actions, which in turn affect the operating costs. As we increase the prediction horizon, the system is allowed to react more proactively and plan more carefully, resulting in smoother controls and lower costs. In Figure 9 we present the hydrogen storage profiles for increasing forecast horizons. The 14-day forecast produces the same amount of hydrogen as the optimal open-loop strategy. On the other hand, a 1-hr forecast strategy can lead to large hydrogen losses (around 50%). The main reason is that, as the horizon is increased, the operating strategy can exploit the battery capacity more efficiently and avoid power losses arising in the hydrogen storage loop.

We emphasize that steady-state RTO is the state of the art in industrial control technology and has brought significant economic benefits. The main reason is that it can manage multiple energy and material sources and identify the most cost-efficient conversion routes through the simulation model. In addition, it can adapt the operating point as the disturbances affecting costs change in time. Nevertheless, from the present study, one can see that the performance of RTO is far from the maximum potential performance. D-RTO inherits all the favorable properties of RTO, but it can also incorporate disturbance forecast information. As can be seen, adding this information directly reflects in significantly more cost-efficient operations.

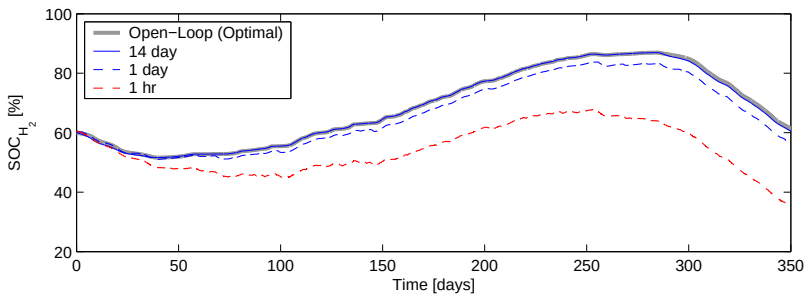


Figure 9: Impact of forecast horizon on hydrogen production.

Building System

Commercial buildings are energy-intensive facilities where considerable cost savings can be realized through optimal operating strategies. As an example, researchers have found that the thermal mass of a building can be used for temporal energy storage [3]. With this, it is possible to optimize the temperature set-points trajectories during the day to shift the heating and cooling electricity demands to off-peak hours and thus reduce costs.

For instance, a cooling strategy that has been used in commercial facilities consists in cooling down as much as possible the building at night when electricity is cheaper so as to reduce the amounts of cooling needed during the day when electricity is more expensive [4]. Since the thermal response of the building can be slow (order of hours), this can be exploited to reduce the on-peak electricity demand the next day. However, we point out that an issue might arise in the implementation of these peak-shifting strategies: namely, the optimal timing at which it is decided to start the cooling at night directly depends on the ambient temperature expected the next day. In addition, special care needs to be taken to stay within the thermal comfort zone at all times. Motivated by these factors, in this case study we analyze the effect of adding forecast temperature information in a D-RTO strategy. We demonstrate that the D-RTO strategy can be easily generalized to consider simultaneous peak-shifting for both cooling and heating during the entire year and to consider multiple energy sources.

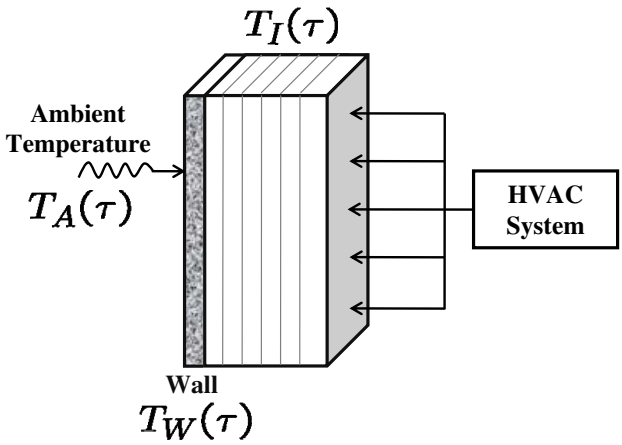


Figure 10: Schematic representation of building integration with heating, ventilation, and air-conditioning (HVAC) system.

The building system under consideration is sketched in Figure 10. We assume a total volume of 10,000 m^3 and a total surface area of 3,500 m^2. The building is equipped with a gas furnace, an electric heater, and an electric cooling system. Ambient temperature data at position 40 30'N/80 13'W in the Pittsburgh, PA, area for year 2006 were obtained from the National Weather Service Office [25]. The temperature profile is presented in Figure 11. As with solar radiation, strong stochastic effects can be observed at different scales. As opposed to solar radiation, however, accounting for variations at night is also necessary. The dynamic response of the building internal temperature is modeled by an ordinary differential equation; the building wall is modeled by a second-order PDE that accounts for conductive effects along the wall. The ambient temperature enters the model through a Neumann boundary condition at the wall external face. The basic heat-transfer model structure has been obtained from [6]. To analyze the effect of adding forecast information of the ambient temperature, we follow the approach described in the previous section. We

first solve an open-loop dynamic optimization problem with perfect forecast information and a prediction horizon of one year. The optimization problem has the following form:

$$\min_{\varphi^c(\tau),\varphi_h^{gas}(\tau),\varphi_h^{elec}(\tau)} \quad \int_{t_k}^{t_k+T} \left[C_{elec}(\tau)\varphi_c^{elec}(\tau) + C_{elec}(\tau)\varphi_h^{elec}(\tau) + C_{gas}\varphi_h^{gas}(\tau)\right] d\tau$$

$$C_I \cdot \frac{\partial T_I}{\partial \tau} = \varphi_h^{gas}(\tau) + \varphi_h^{elec}(\tau) - \varphi_c^{elec}(\tau) - S \cdot \alpha' \cdot (T_I(\tau) - T_W(\tau, 0)) \quad (13a)$$

$$\frac{\partial T_W}{\partial \tau} = \beta \cdot \frac{\partial^2 T_W}{\partial x^2} \quad (13b)$$

$$0 = \alpha' \left(T_I(\tau) - T_W(\tau, 0)\right) + \mathbf{k} \cdot \left.\frac{\partial T_W}{\partial x}\right|_{(\tau,0)} \quad (13c)$$

$$0 = \alpha'' \left(T_W(\tau, L) - T_A(\tau)\right) + \mathbf{k} \cdot \left.\frac{\partial T_W}{\partial x}\right|_{(\tau,L)} \quad (13d)$$

$$T_I^{min} \leq T_I(\tau) \leq T_I^{max} \quad (13e)$$

$$T_I(0) = T_I^k \quad (13f)$$

$$T_W(0, x) = T_W^k(x) \quad (13g)$$

where $T_A(\tau)$ is the ambient temperature, $T_I(\tau)$ is the internal temperature, and $T_W(\tau, x)$ is the wall temperature (all of them in ^{o}C). The controls are the gas heating power $\varphi_h^{gas}(\tau)$, the electric heating power $\varphi_h^{elect}(\tau)$, and the electric cooling power $\varphi_c^{elec}(\tau)$ (all of them in kcal/hr). The model parameters are summarized in Table 1. The base wall thickness was assumed to be 0.20 m. We assume an on-peak electricity price of 0.12 \$/kWh available from 9 a.m. to 10 p.m. The off-peak price is 0.04 \$/kWh. A demand rate of 16 \$/kW is charged for the monthly peak electricity demand. The natural gas price is fixed at 0.10 \$/kWh. Average prices were obtained from [35]. The thermal comfort zone is assumed to be 69-77^oF. The above PDE-constrained optimization problem was discretized using a central difference scheme in the axial dimension and an implicit Euler scheme in time. The resulting NLP was implemented in AMPL and solved with the solver IPOPT [37].

From the solution of the open-loop dynamic optimization problem, we obtain the optimal cost and use it as a reference for the best economic performance of the system. The resulting minimum annual cost is \$28,672 (demand cost is approximately 60% of total cost). We then solve closed-loop D-RTO problems over the entire year with prediction horizons of 1, 3, 6, 9, 12, 16, and 24 hr. An update time Δ of 1 hr was used. The relative costs (excluding demand costs) are presented in Figure 12. As can be seen, for a purely reactive strategy, the relative costs can go as high as 24% as a result of lack of proactiveness. In addition, we observe that an horizon of 24 hr is sufficient to achieve the minimum potential costs. This is explained by the fact that the thermal mass of the building cannot be used for a long time because there exists thermal losses through the wall. In fact, we found that as the building insulation is enhanced, the costs can be further reduced. In Figure 13, we present the relative costs with an increased wall thickness of 0.3 m. As can be seen, using a forecast of 24 hr can reduce costs by 45%. On the other hand, when the building is poorly insulated, increasing the forecast horizon does not reduces the costs. In other words, *the economic potential of adding forecast information is tightly*

related to the ability to store energy in the system, which is in turn influenced by the building characteristics.

In Figure 14 we present the temperature set-points for the 24-hr and 1-hr forecast cases during 10 days in the winter season. As can be seen, the 24 hr forecast strategy determines the optimal timing at which electric heating needs to be turned on at night. Note that the optimum timing and the peak temperature depend on the expected ambient temperature. On the other hand, the reactive strategy is not able to foresee the structure of the electricity prices. This strategy suggests that the optimal policy is to keep the temperature set-point always at the lowest possible value in order to reduce the overall heating costs. Although this strategy seems intuitive, it is clearly not optimal if the structure of the electricity rates and the thermal mass of the building can be exploited. From Figure 15, we observe that the optimal cooling policy during the summer follows a peak-shifting strategy. The resulting policy recommends to let the building cool down at night until the temperature gets close to the lower limit of the comfort zone. During the day, the building is allowed to heat up progressively until it reaches the highest limit of the comfort zone. Similar results have been obtained by Braun and coworkers [4]. The proposed closed-loop D-RTO framework can account for time variations and correct the policy automatically on-line. In this simplified study the cooling requirements are negligible since we only account for heat gains and losses through the wall. In addition, the day-night temperature difference at this location is quite large during summer, as seen in Figure 15. A more detailed study should also account for internal heat gains, radiation heating, air recycling, and humidity factors. Nevertheless, with these preliminary results, we would like to illustrate that the performance of operating strategies can benefit from anticipating the weather conditions.

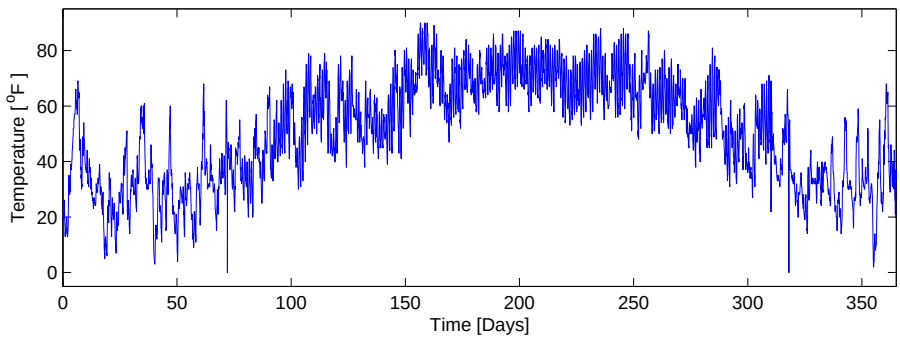


Figure 11: Ambient temperature in Pittsburgh PA, 2006.

3.3 Gaussian Process Modeling

In the previous sections, we have demonstrated that important economic benefits can be realized by using weather forecast information. But several questions arise: Can we get accurate forecast information? What techniques can be used? In which form is

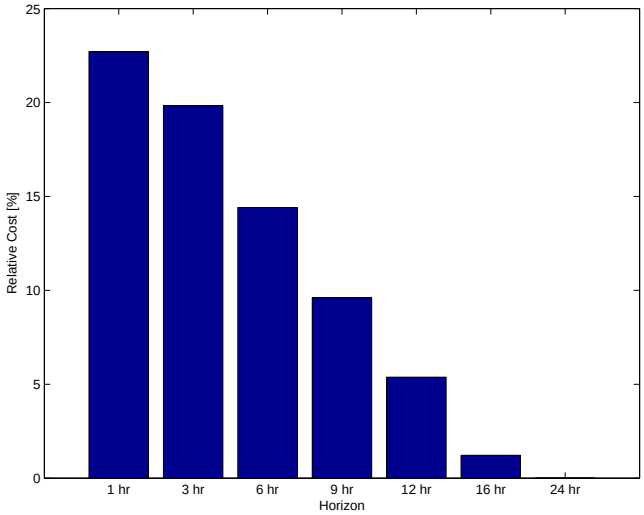


Figure 12: Impact of forecast horizon on economic performance of building system. Base insulation.

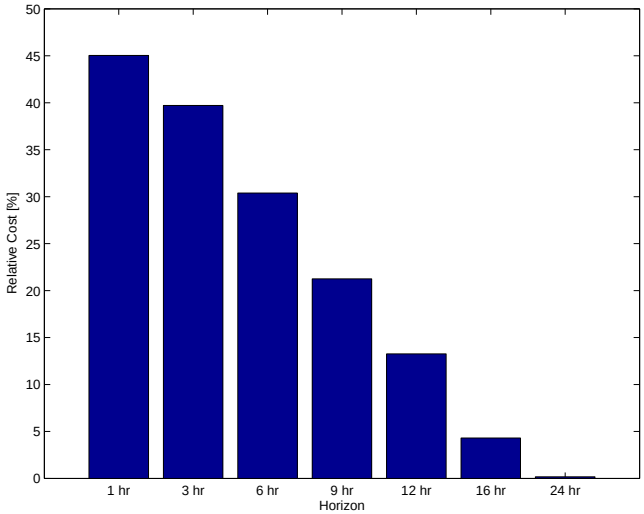


Figure 13: Impact of forecast horizon on economic performance of building system. Well-insulated building.

the information needed? What is the effect of neglecting uncertainty? In this section, we present a technique to derive empirical forecast and uncertainty information based solely on data. This will be used to explain how to connect the closed-loop optimization framework with the forecast capabilities. We show that, while the empirical strategy is practical and useful, it has important limitations. This motivates our interest in obtaining more consistent forecast information through a detailed weather model.

The most straightforward forecasting alternative is to use historical data to construct a

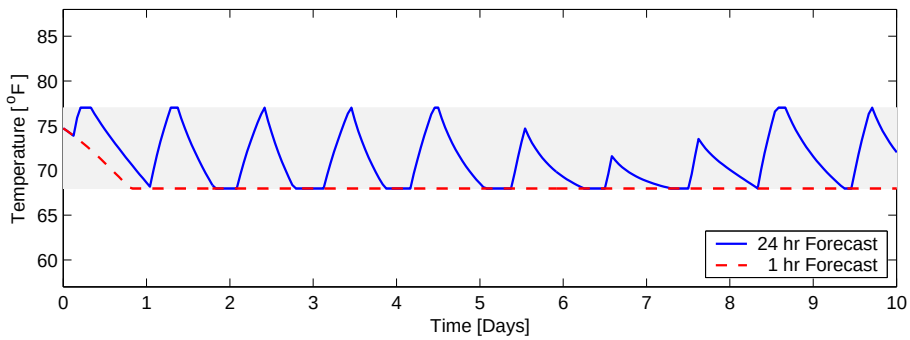


Figure 14: Optimal temperature set-points of closed-loop D-RTO with 1 hr and 24 hr forecasts. Comfort zone is highlighted in gray.

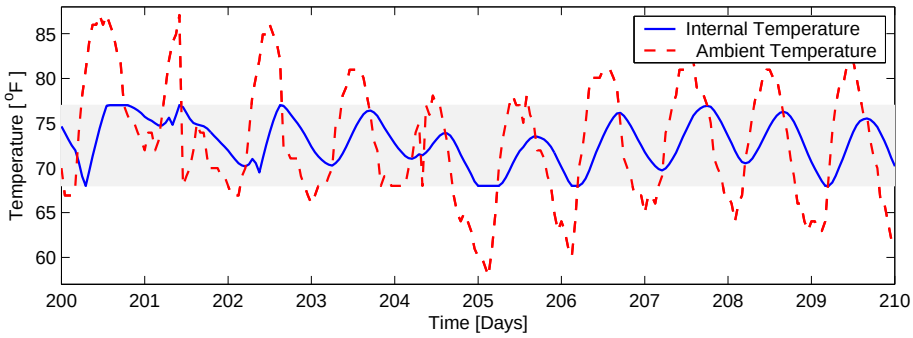


Figure 15: Internal temperature set-point and ambient temperature during 10 days in summer season. Closed-loop D-RTO with a forecast of 24 hr was used. Comfort zone is highlighted in gray.

time-series regression model. An approach that has recently received attention is Gaussian process (GP) modeling [28]. The idea is to construct the regressive model by specifying the structure of the covariance matrix rather than the structure of the dynamic model itself as in the Box-Jenkins approach [2]. We have found that this feature makes the GP approach more flexible. To illustrate the use of this technique, we construct a forecast model for the solar radiation by regressing the future radiation (output) χ_{k+1} to the current and previous radiation values $\chi_k, ..., \chi_{k-N}$ that can be obtained from available data bases [26]. In this case, N is selected long enough to capture the day-to-day periodicity of the radiation. With this, we define the model inputs as $\mathbf{X}_j = [\chi_k, ..., \chi_{k-N}]$ and the outputs as $\mathbf{Y}_j = \chi_{k+1}$. We collect a number of input-output pairs $j = 1, ..., N_T$ to obtain the training data sets. In this case, we assume that the inputs are correlated through an exponential covariance function of the form

$$\mathbf{V}(\mathbf{X}_j, \mathbf{X}_i, \eta) := \eta_0 + \eta_1 \cdot \exp\left(-\frac{1}{\eta_2}\|\mathbf{X}_j - \mathbf{X}_i\|^2\right), \tag{14}$$

where η_1, η_2 and η_3 are hyperparameters estimated by maximizing the log likelihood func-

Table 1: Building model parameters.

Parameter	Value	Units	Meaning
β	0.001	$\frac{m^2}{hr}$	thermal diffusivity of wall
C_I	34,800	$\frac{kcal}{^oC}$	internal heat capacity
$\mathbf{k}$	0.1	$\frac{kcal}{m \cdot hr \cdot ^oC}$	conductivity of wall
S	3,500	m^2	wall total surface area
A	1,000	m^2	usable total surface area
V	10,000	m^3	building total volume
α'	4	$\frac{kcal}{m^2 \cdot hr \cdot ^oC}$	convective heat transfer coefficient (wall inner side)
α''	10	$\frac{kcal}{m^2 \cdot hr \cdot ^oC}$	convective heat transfer coefficient (wall outer side)
L	0.20	m	wall thickness
C_{elec}	0.12	$\frac{\$}{kWh}$	off-peak electricity cost
C_{elec}	0.04	$\frac{\$}{kWh}$	on-peak electricity cost
C_{gas}	0.10	$\frac{\$}{kWh}$	natural gas cost

tion:

$$\log p(\mathbf{Y}|\eta) = -\frac{1}{2}\mathbf{Y}\mathbf{V}^{-1}(\mathbf{X},\mathbf{X},\eta)\mathbf{Y} - \frac{1}{2}\log\det(\mathbf{V}(\mathbf{X},\mathbf{X},\eta)). \tag{15}$$

Once the optimal hyperparameters η^* are obtained, we can compute mean predictions $\mathbf{Y}^P$ with associated covariance $\mathbf{V}^P$ at a set of test points $\mathbf{X}^P$. In our context, these are the time-varying radiation trends. The resulting GP posterior distribution is:

$$\mathbf{Y}^P = \mathbf{V}(\mathbf{X}^P,\mathbf{X},\eta^*)\mathbf{V}^{-1}(\mathbf{X},\mathbf{X},\eta^*)\mathbf{Y} \tag{16a}$$

$$\mathbf{V}^P = \mathbf{V}(\mathbf{X}^P,\mathbf{X}^P,\eta^*) - \mathbf{V}(\mathbf{X}^P,\mathbf{X},\eta^*)\mathbf{V}^{-1}(\mathbf{X},\mathbf{X},\eta^*)\mathbf{V}(\mathbf{X},\mathbf{X}^P,\eta^*) \tag{16b}$$

We use approximately 400 training data sets from the Chicago, IL, data presented in Figure 5. In Figure 16, we present the mean forecast and 100 samples drawn from the normal distribution $\mathcal{N}(\mathbf{Y}^P,\mathbf{V}^P)$ on a particular day. Note that the forecast distribution is able to capture the true radiation values. In addition, the variance is small at night, as suggested from physical intuition. This implies that the assumed covariance structure is reasonable.

To illustrate the importance of considering the forecast uncertainty explicitly in the D-RTO formulation, we revisit the case study for the photovoltaic-hydrogen system described in Section 3.2. A GP model is used to obtain the forecast information. We consider two closed-loop optimization strategies. The first one is a D-RTO deterministic strategy that uses only the mean forecast coming from the GP model. The second is a stochastic D-RTO strategy that samples the uncertain space predicted from the GP model. A forecast horizon of one day is used. In Figure 17, we present the hourly operating costs for both closed-loop strategies. As can be seen, the deterministic strategy is unable to satisfy the load, which is indicated by large penalty costs. On the other hand, the

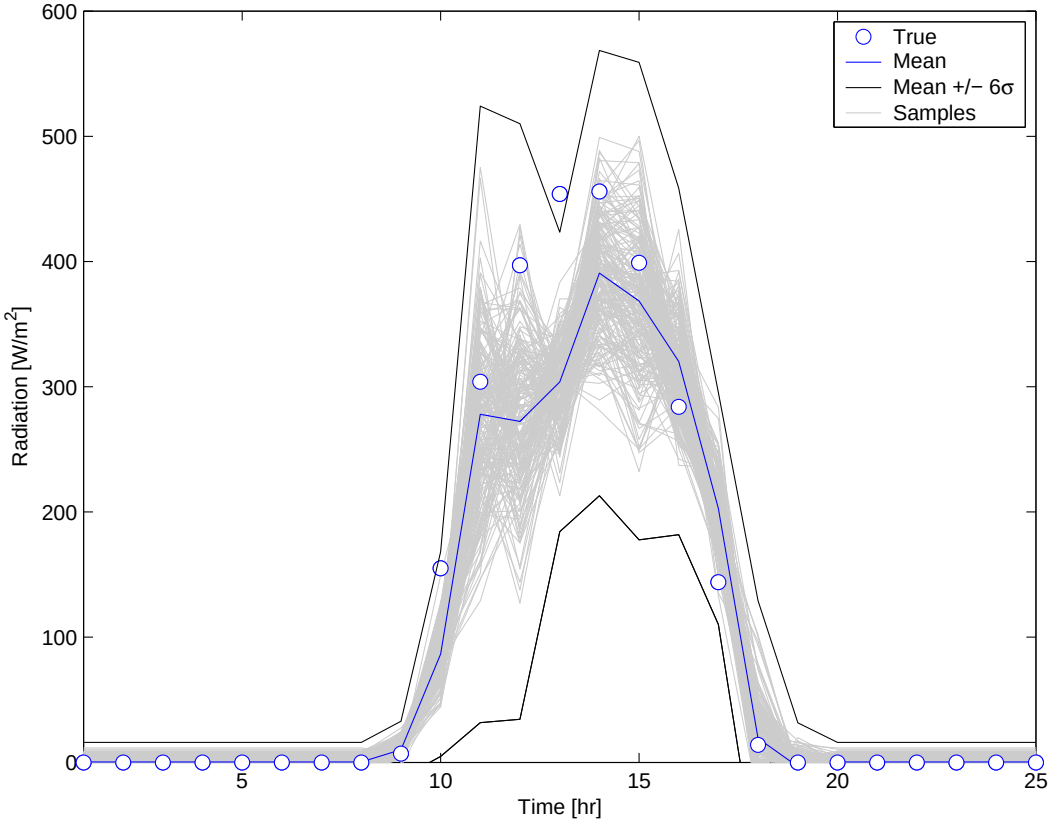


Figure 16: Solar radiation samples of predicted forecast distribution.

stochastic strategy is always able to satisfy the load. Note that at a particular time t_k the closed-loop optimizer dictates the future power production set-points that will satisfy the demand. Since the deterministic strategy does not take all possible realizations of the solar radiation into account, it can overestimate the actual solar radiation and produce less power than what is required. On the other hand, the stochastic strategy does take all possible realizations into account and ensures that no matter what the actual radiation will be, enough power will always be available. This enhanced robustness is particularly critical in off-grid or stand-alone systems. Nevertheless, we emphasize that this robustness can also lead to significant deterioration of economic performance if the uncertainty bounds are wide. From Figure 16, we see that the uncertainty bounds obtained from the Gaussian model are wide during the day. We have also found that the GP model is not able to capture long-term and spatial trends. Another issue that arises in this context is that obtaining measurement data is expensive and is usually not available on site. Motivated by these limitations, in the following section we explore the potential of using a detailed weather model to obtain more accurate and cheap forecast information. In addition, we present extensions to extract detailed uncertainty information in a systematic manner and in a form that is consistent with the proposed stochastic D-RTO strategy.

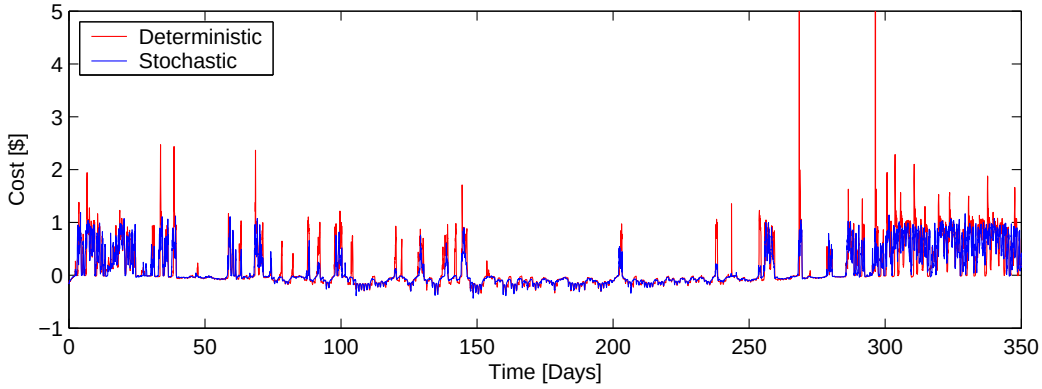


Figure 17: Performance of closed-loop deterministic and stochastic D-RTO strategies with forecast horizon of 24 hr.

4 Uncertainty Quantification

In this section we describe the development for an uncertainty model for the ambient temperature and ground solar radiation. The ground solar radiation uncertainty results from the incoming solar radiation and is driven through the complex and chaotic atmospheric processes. To our knowledge, there are no analytic or parametric models that describe this type of uncertainty. The reason is, in part, that solar radiation is included in the Earth energy budget feedback loop, and modeling it requires good knowledge of the state of the atmosphere. Furthermore, several physical processes involved in this system are not well understood or modeled. Parametric models calibrated to certain locations and time frames have been used to ascertain the radiation levels on the ground [32; 14]; however, the uncertainty in the radiation was never ascertained.

The major sources of uncertainty in solar radiation include cloud cover, temperature, and humidity [32]. Several models that rely on the minimum and maximum daily temperature have been devised in [32; 14]. This dependence implies a dynamic aspect of the uncertainty. For instance, the variation of the radiation is tightly connected with the probability of having or not having cloud cover and with the cloud type.

We employ a numerical weather prediction (NWP) model to assess the forecast uncertainty in the temperature and solar radiation. First we develop a model for a priori uncertainties in the temperature because of its tight connections with relative humidity and cloud cover. The uncertainties in the solar radiation as well as in the temperature field are then evolved though the NWP model dynamics.

4.1 Numerical Weather Prediction Model – WRF

Weather Research and Forecasting (WRF) model is a state-of-the-art mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs [30]. The current version of the model, WRF 3.1, with the default settings

is used for the forecast and uncertainty estimation in the temperature and radiation fields. The solar radiation model within WRF is based on the Dudhia scheme [10].

We briefly describe the experimental setting. The data used in the WRF model corresponds to North American Regional Reanalysis data set that covers 160E-20W, 10N-80N, with a resolution of 10 minutes of a degree. There are 29 pressure levels (1000-100 hPa, excluding the surface) and a three-hour output frequency. The time period under consideration ranges from August 1-30, 2006. This data set includes meteorological fields such as temperature, wind, and humidity, as well as geophysical forcings such as soil albedo and vegetation type.

The ground solar radiation field is predicted by the model physics based on the initial conditions and WRF forecast fields.

4.2 Ensemble Approach to Uncertainty Quantification

The uncertain model states are described through random variables. We consider random variables with a Gaussian probability density function $\mathcal{N}(\mu, \mathcal{B})$. This is a reasonable assumption for large-scale applications with many uncertain parameters such as WRF. By using the Gaussian distribution, the uncertainty is described completely by the random variable mean μ and covariance matrix B.

The covariance matrix is typically large (grows with the square of the number of uncertainty states) and hence in practice is approximated with a reduced model [8] or with an ensemble of realizations [27]. We next discuss the latter approach.

Uncertainty Representation via Ensemble of Realizations

The covariance matrix can be approximated from an ensemble of realizations in the following way. For n variables, the (unbiased) covariance matrix $B \in \mathbb{R}^{n\times n}$ can be approximated by an ensemble of m realizations, x_i, $1 \le i \le m$, with

$$B \approx \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \overline{x})(x_i - \overline{x})^T \ , \quad \overline{x} = \frac{1}{m}\sum_{i=1}^{m} x_i \, .$$

One also has

$$B = D^{\frac{1}{2}} C D^{\frac{1}{2}} \, , \quad C_{i,j} = \frac{B_{i,j}}{\sqrt{D_{i,i}}\sqrt{D_{j,j}}} = \frac{B_{i,j}}{\sigma_i \sigma_j} = \frac{\sigma_{i,j}^2}{\sigma_i \sigma_j} \, , \ 1 \le i, j \le n \, ,$$

where C is the full symmetric correlation matrix and D is a diagonal matrix holding the local variances ($D_{i,i} = \sigma_i^2$).

To explain how the uncertain variables are evolved by the model dynamics, we introduce the tangent linear and adjoint models and make some simplifying assumptions.

Tangent Linear and Adjoint Models

We denote by $\mathcal{M}$ the (non)linear model that evolves the state y from time index i to $i+1$:

$$y(t_{i+1}) = \mathcal{M}(y(t_i)) \, .$$

The tangent linear model (TLM), M, describes the propagation of small perturbations $\delta y(t_i)$ through the system $\mathcal{M}$:

$$\mathcal{M}\left(y(t_i)+\delta y(t_i)\right)=\mathcal{M}\left(y(t_i)\right)+\frac{d\mathcal{M}}{dy}\cdot\left(\delta y(t_i)\right)+\mathcal{O}(\delta y^2(t_i))\rightarrow$$
$$\delta y(t_{i+1})=M\cdot\delta y(t_i)+\mathcal{O}(\delta y^2(t_i))\,,\quad M=\frac{d\mathcal{M}}{dy}\,.$$

The transpose of the TLM is called the adjoint of the TLM model, M^*.

Error Propagation in the Presence of Model and Data Errors

Consider the true solution of the model at time t_i, $y_{true}(t_i)$. The numerical model is not perfect, and the true solution at t_{i+1} is given by

$$y_{true}(t_{i+1})=\mathcal{M}\left(y_{true}(t_i)\right)+\eta(t_i)\,,$$

where η represents the model errors that are considered unbiased with covariance Q, $\eta\in\mathcal{N}(0,Q)$.

The initial solution is not known exactly. Therefore, the numerical solution at time t_{i+1}, $y(t_{i+1})$, is obtained from the model evolution of the true solution perturbed with errors that are represented by a set of unbiased random variables $\varepsilon(t)$, $\varepsilon\in\mathcal{N}(0,B)$:

$$y(t_{i+1})=\mathcal{M}\left(y_{true}(t_i)+\varepsilon(t_i)\right)\,,$$

We remark that $y(t_{i+1})$ is a random variable, and we compute its covariance matrix P:

$$\begin{aligned}P=&\overline{\left(y(t_{i+1})-y_{true}(t_{i+1})\right)\left(y(t_{i+1})-y_{true}(t_{i+1})\right)^T}=\\&\overline{\left(\mathcal{M}\left(y_{true}(t_i)+\varepsilon(t_i)\right)-\mathcal{M}\left(y_{true}(t_i)\right)+\eta(t_i)\right)\cdot}\\&\overline{\cdot\left(\mathcal{M}\left(y_{true}(t_i)+\varepsilon(t_i)\right)-\mathcal{M}\left(y_{true}(t_i)\right)+\eta(t_i)\right)^T}\,.\end{aligned}\tag{17}$$

The following assumptions are made:

- Errors due to initial conditions modeled by ε and model errors (η) are uncorrelated, and
- the error growth is well approximated by the linearized model.

Therefore we have

$$\overline{\left(\mathcal{M}\left(y_{true}(t_i)+\varepsilon(t_i)\right)\right)\eta(t_i)^T}=0$$

and

$$\mathcal{M}\left(y_{true}(t_i)+\varepsilon(t_i)\right)-\mathcal{M}\left(y_{true}(t_i)\right)=M\cdot\varepsilon(t_i)\,.$$

It follows that (17) becomes

$$\begin{aligned} P &= \overline{\left(M \cdot \varepsilon(t_i) + \eta(t_i)\right)\left(M \cdot \varepsilon(t_i) + \eta(t_i)\right)^T}\,, \\ &= \overline{M\varepsilon(t_i)\varepsilon^T(t_i)M^T} + \overline{M \cdot \varepsilon(t_i)\eta(t_i)^T} + \overline{\eta(t_i)\varepsilon^T(t_i)M^T} + \overline{\eta(t_i)\eta(t_i)^T}\,, \\ &= MBM^* + Q\,. \end{aligned} \tag{18}$$

Equation (18) represents an approximation of the error covariance P obtained by the propagation of the error covariance B through the linearized model in the presence of model errors.

The WRF model propagates the uncertainties in the temperature field according to (18). However, in order to estimate P which characterizes the forecast uncertainty, we first need to specify the "background" covariance B. We next introduce a method that can be used to estimate B.

The NCEP Method for Covariance Estimation

The NCEP (former NMC) method [27; 12; 16] has been used to estimate the spatial uncertainty information in numerical weather prediction. An ensemble of realizations y is used to approximate the covariance matrix. Then differences among several perturbed forecasts verifying at the same time are used to approximate the covariance matrix [27; 7],

$$B \approx \overline{(y(t_{48\mathrm{h}}) - y(t_{24\mathrm{h}}))(y(t_{48\mathrm{h}}) - y(t_{24\mathrm{h}}))^T}\,,$$

where the time indices are not necessarily restricted to the indicated values and $\overline{(\ldots)}$ represents the ensemble average. This approach typically results in a low-rank noisy approximation; however, one can estimate characteristic correlation distances (e.g., average radius within one standard deviation). A new covariance matrix can be constructed by using a Gaussian distribution that fits the empirical correlation distance.

An Uncertainty Model for Temperature

The NCEP method is used to construct an uncertainty model for the temperature field. The inferred characteristic horizontal correlation distance for this experiment is approximated by $L_H = 2$ degrees and by $L_V = 500$ meters in the vertical direction. The spatial correlation function between two points $y(t, x_i, y_i, z_i)$ and $y(t, x_j, y_j, z_j)$ is then given by

$$C_{i,j} = \exp\left(-\frac{(x_j - x_i)^2 + (y_j - y_i)^2}{L_H^2} - \frac{(z_j - z_i)^2}{L_V^2}\right)\,. \tag{19}$$

The correlation function (19) is used to construct the covariance matrix from which the initial ensemble for the temperature field is drawn.

We consider that the true initial temperature field is not known exactly but is correctly represented in the probabilistic framework by an unbiased random vector $\varepsilon_T(t_0)$. The

initial temperature field, $T(t_0)$, is properly characterized by a random vector $T^B(t_0)$ with the following properties:

$$T^B(t_0) = T(t_0) + \varepsilon_T(t_0)\,,\ \overline{\varepsilon_T(t_0)\varepsilon_T(t_0)^T} = B_{TT}\,,\ \overline{\varepsilon_T(t_0)} = 0 \Rightarrow$$
$$\Rightarrow T^B(t_0) \in \mathcal{N}\left(T(t_0), B_{TT}\right)\,.$$

The initial temperature field is approximated by an m-member ensemble drawn from B_{TT}:

$$T^B_{[i]}(t_0) = T(t_0) + G\,(\sigma\, I)\, C\,(\sigma\, I)\, G^T \xi_i\,,\ 1 \le i \le m\,,\ \xi \in \mathcal{N}(0,1)\,,$$

where error variance is set to $\sigma = 1$, and G transforms the unbalanced variables into full quantity for temperature. Matrix G is chosen to be $\sigma_G(z_i)I$, where

$$\sigma_G(z_i) = \overline{T(t_0, z_i)} / \max_i \left(\overline{T(t_0, z_i)}\right)\,.$$

Therefore, the temperature perturbation depends on the average temperature on the corresponding layer, and most (99%) of the members have a variation of $\pm 5K$, which is a sensible value selected for this experiment.

Validation of the uncertainty temperature model. We next validate the forecast and uncertainty temperature model by using independent observations described in Sec. 3.2 for the Pittsburgh area [25]. In Figure 18 we show an ensemble of 30 members for temperature realizations, the expected temperature value, and an independent set of observations for five days (Aug. 1–6, 2006). The realization's envelope encloses the true (measured) solution for most time indices, except for day three, which indicates a possible sensor malfunction. Therefore the uncertainty model is well approximated at least for this test case.

Spatial error correlations for the temperature field. The ensemble forecast provides sufficient information to describe the spatial error distribution. As described in Sec. 4.2 we illustrate the horizontal correlation field for the temperature error in the Pittsburgh area [25]. In Figure 19 we show the error correlation that corresponds to 10 AM August 1 and August 2, 2006. The noise observed in the correlation field can be reduced by increasing the ensemble size.

4.3 Uncertainty Model for Radiation

It is known that the solar radiation is sensitive to variations in the temperature field [32]. To this end, we use the initial temperature ensemble developed in the previous section to characterize the uncertainty in both the forecast temperature and solar radiation. In other words, the weather model is used to evolve the uncertainty in the initial temperature field to uncertainty in the solar radiation. The resulting radiation and temperature ensembles are then used to construct approximations to the first two statistical moments of the radiation and temperature uncertainty variables.

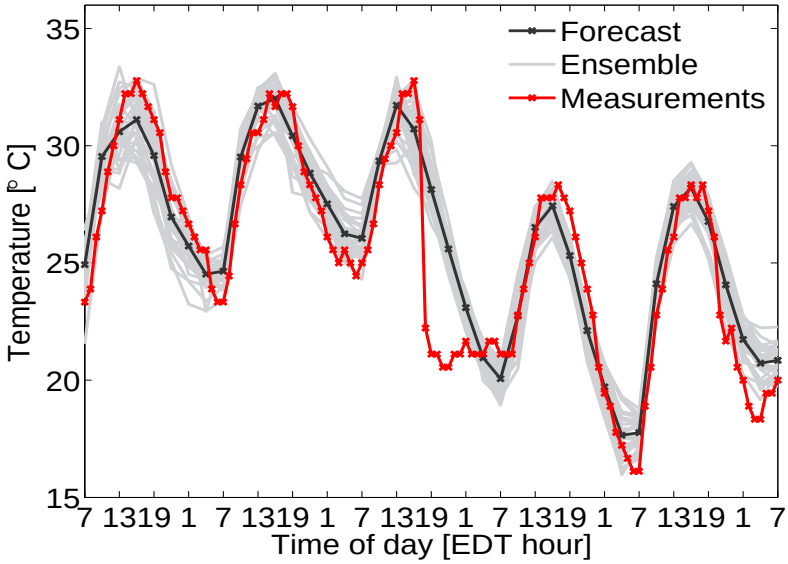
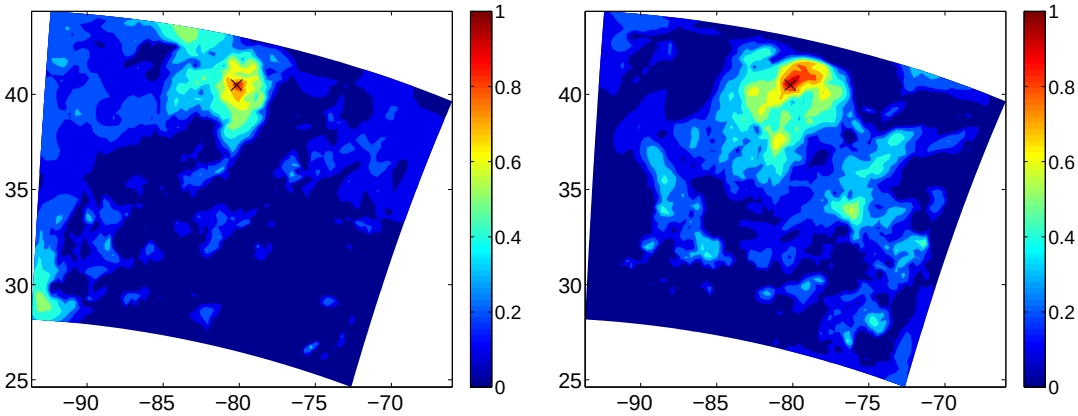


Figure 18: Forecast and measured temperature for the Pittsburgh airport area during August 1–6, 2006.



a) August 1st, 2006 b) August 2nd, 2006

Figure 19: Correlation field for the temperature errors in the Pittsburgh area at 10 AM on consecutive days.

Define $y = [T, \nu, \tilde{y}]^T$, where T represents the temperature, ν the ground solar radiation, and $\tilde{y}$ the rest of the state vector y. We remark that the ground solar radiation is not a model input and its prior distribution is not explicitly present in the weather model. The

temperature covariance, B_{TT}, is evolved in time as described in (18):

$$\left[\begin{array}{ccc} P_{T,T} & P_{T,\nu} & P_{T,\tilde{y}} \\ P_{T,\nu} & P_{\nu,\nu} & P_{\nu,\tilde{y}} \\ P_{T,\tilde{y}} & P_{\nu,\tilde{y}} & P_{\tilde{y},\tilde{y}} \end{array}\right] \Bigg| \varepsilon_T(t_0) = \mathcal{M}\left(T^B(t_0) = T(t_0) + \varepsilon_T(t_0), \tilde{y}\right) . \tag{20}$$

Further, the covariance matrix for the radiation can be approximated by

$$P_{\nu,\nu} = M_{T\to\nu} B_{T,T} M^*{}_{T\to\nu} + Q\,, \tag{21}$$

where $M_{T\to\nu}$ represents the TLM for $\nu(t_{i+1}) = \mathcal{M}_{T\to\nu}\left(T(t_i), \tilde{y}(t_i)\right)$; that is, $\mathcal{M}_{T\to\nu}$ is the subset of equations involved in obtaining the solar radiation from temperature.

Under these considerations, the uncertainty in the temperature field is completely described by the expected value on the model forecast and covariance matrix $P_{T,T}$. Similarly, the uncertainty in the ground radiation is also completely specified by the expected value of the ensemble forecast and covariance $P_{\nu,\nu}$.

The solar radiation forecast for the Pittsburgh area that corresponds to the temperature profile shown in Figure 18 is illustrated in Figure 20. The prediction for this case is relatively clustered, albeit a slight variation can be noticed during the day. Other regions, not shown in this study, present a different outcome in the daytime forecast.

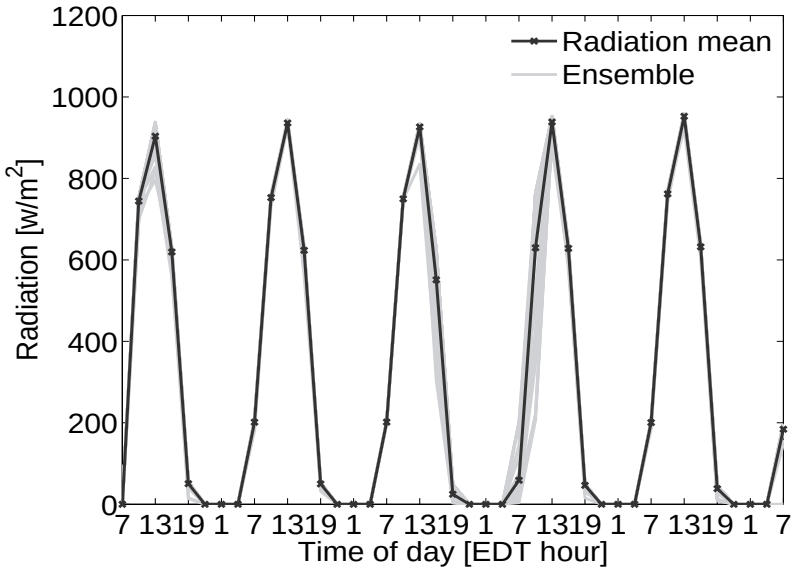


Figure 20: Forecast solar radiation for the Pittsburgh airport area during August 1–6 , 2006.

5 Discussions and Future Work

In this work, we have demonstrated that significant costs reductions, on the range of 18-75 %, can be achieved in integrated energy systems by using operation technology that

can anticipate the weather conditions in a systematic manner. In this section, we discuss scientific, engineering, and infrastructure obstacles that would need to be overcome.

5.1 Possible Deployment Strategy

In our calculations of the building model and the hybrid energy system, the largest computation time by far was taken by the calculation of the weather uncertainty model. On the other hand, should such a system be deployed on a regional or national scale, nearly spatially collocated customers could share virtually all of the information uncertainty calculation. This fact points out to a possible deployment strategy where a dedicated supercomputer is used to carry out the weather uncertainty portion of the calculation, regional or local Gaussian process modeling is done in a dedicated climate data center, and the control strategy is computed either in the customer facility or residence or by a vendor that runs one thread per customer and accesses the data center to obtain the uncertainty information at the required location.

5.2 Scientific and Technological Advances Needed

The assessment of the impact of our approach is based on a relatively limited number of detailed building and hybrid energy models. Therefore, one immediate need is an expansion of the database of models for which we can run our experiments, as well as carrying out such experiments at multiple locations (we point out, nonetheless, that our experiments were run in fairly high density areas, so we expect that our performance results will extrapolate to those cases as well). More detailed models may result in more savings as well as more computational needs. In addition, it is conceivable that advanced building technologies would result in the addition of sensors whose information should also be incorporated in the control strategy; a proper mathematical formulation for that case needs to be developed.

In terms of the further scientific advances needed, immediate needs include the extension of Gaussian process approach used in both space and time in a way that allows for scalable calculations when such a system is deployed on a regional scale.

In our examples we have assumed that the end user of our approach resides at one of the grid points of the mesh where the calculation was carried out; the computational and development requirements are expected to be much larger for off-grid uncertainty quantification.

Perhaps the area with the deepest need is the one of uncertainty quantification of weather models, where many problems that would result in superior models are not yet solved. These include appropriate uncertainty models that account for a multiscale approach (such as a coupled national-regional model), as well as improved ways of assessing the effects of initial choices of the uncertainty parameters on the end assessments.

Important advances are also needed in tools that are suitable for large-scale on-line stochastic optimization. These include adaptive sampling, which would result in a reduction of the number of scenarios that need to be considered for a given confidence level, as well as iterative methods for computing the optimization set-point.

Acknowledgments

This work was supported by the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract DE-AC02-06CH11357.

References

[1] A. Bilodeau and K. Agbossou. Control analysis of renewable energy system with hydrogen storage for residential applications. *Journal of Power Sources*, 162:757–764, 2006.

[2] G. E. P. Box, G. Jenkins, and G. Reinsel. *Time Series Analysis: Forecasting and Control*. Prentice-Hall, New Jersey, 1994.

[3] J. E. Braun. Reducing energy costs and peak electricty demand through optimal control of building thermal storage. *ASHRAE Transactions*, 96(2):876–888, 1990.

[4] J. E. Braun, K. W. Montgomery, and N. Chaturvedi. Evaluating the performance of building thermal mass control strategies. *HVAC&Research*, 7(4):403–428, 2001.

[5] R. H. Byrd, J. Ch. Gilbert, and J. Nocedal. A trust-region method based on interior-point techniques for nonlinear programming. *Math. Programm.*, 89:149–185, 2000.

[6] C. Carrara, L. F. Fantini, and A. Lorenzi. Influence of the outdoor temperature variation on the performance of a building heating system. In C. J. Hoogendoorn, editor, *Energy Conservation in heating, Cooling and Ventilating Buildings*. 1978.

[7] T. Chai, G.R. Carmichael, Y. Tang, A. Sandu, M. Hardesty, P. Pilewskie, S. Whitlow, E.V. Browell, M.A. Avery, P. Nedelec, J.T. Merrill, A. Thompson, and E. Williams. Four-dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements. *Journal of Geophysical Research*, 112, 2007.

[8] E.M. Constantinescu, T. Chai, A. Sandu, and G.R. Carmichael. Autoregressive models of background errors for chemical data assimilation. *Journal of Geophysical Research*, 112:D12309, 2007.

[9] D. B. Crawley, L. K. Lawrie, C. O. Pedersen, R. J. Liesen, D. E. Fisher, R. K. Strand, R. D. Taylor, R. C. Winkelmann, W. F. Buhl, Y. J. Huang, and A. E. Erdem. EneryPlus: A new-generation building energy simulation program. In *Proceedings of Building Simulation '99*, volume 1, pages 81–88, 1999.

[10] J. Dudhia. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric Sciences*, 46(20):3077–3107, 1989.

[11] R. Fourer, D. Gay, and B. Kernighan. *AMPL*. The Scientific Press, South San Francisco, 1993.

[12] R. Franke and E. Barker. Vertical correlation functions for temperature and relative humidity errors. *Monthly Weather Review*, 128(12):3962–3981, 2000.

[13] H. K Geyer and R. K Ahluwalia. GCtool for fuel cell systems design and analysis: User documentation. Argonne National Laboratory, 1998.

[14] L. A. Hunt, L. Kuchar, and C. J. Swanton. Estimation of solar radiation for use in crop modelling. *Agricultural and Forest Meteorology*, 91(3-4):293 – 300, 1998.

[15] J. Kadam, W. Marquardt, M. Schlegel, T. Backx, O. Bosgra, and P.J. Brouwer. Towards integrated dynamic real-time optimization and control of industrial processes. *Proceedings Foundations of Computer-Aided Process Operations (FOCAPO2003)*, pages 593–596, 2003.

[16] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, et al. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3):437–471, 1996.

[17] L. L. Kazmersky and K. Broussard. Solar photovoltaic hydrogen: The technologies and their place in our roadmaps and energy economics. *Technical Report NREL/CP-520-36401, National Renewable Energy Laboratory*, 2004.

[18] S. A. Klein, J. A. Duffie, and W. A. Beckman. TRNSYS: A transient simulation program. *ASHRAE Transactions*, 82:623–633, 1976.

[19] Argonne National Laboratory. PSAT (Powertrain Systems Analysis Toolkit), 2009. `http://www.transportation.anl.gov/`.

[20] J. Lagorse, M. G. Simoes, A. Miraoui, and P. Costberg. Energy cost analysis of a solar-hydrogen hybrid energy system for stand-alone applications. *International Journal of Hydrogen Energy*, 33:2871–2879, 2008.

[21] J. Linderoth, A. Shapiro, and S. Wright. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research*, 142(1):215–241, 2006.

[22] T. Markel, A. Brooker, T. Hendricks, V. Johnson, K. Kelly, B. Kramer, M. O'Keefe, S. Sprik, and K. Wipke. Advisor: a systems analysis tool for advanced vehicle modeling. *Journal of Power Sources*, 110(2):255–266, 2002.

[23] T.E. Marlin and A.N. Hrymak. Real-time operations optimization of continuous processes. In *Fifth International Conference on Chemical Process Control (CPC-5)*, 1996.

[24] A. Milbrandt and M. Mann. Potential for hydrogen production from key renewable resources in the United States. *Technical Report NREL/TP-640-41134, National Renewable Energy Laboratory*, 2007.

[25] National Oceanic and Atmospheric Association. National Weather Service Forecast Office, 2009. `http://www.erh.noaa.gov/pbz/hourlyclimate.htm`.

[26] National Renewable Energy Laboratory. National Solar Radiation Data Base, 2009. `http://rredc.nrel.gov/solar/old_data/nsrdb/`.

[27] D.F. Parrish and J.C. Derber. The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763, 1992.

[28] C. E. Rasmussen and C. K. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, 2006.

[29] A. Ruszczynski and A. Shapiro. *Stochastic Programming (Handbooks in Operations Research and Management Series)*. Elsevier Science BV, Amsterdam, 2003.

[30] W.C. Skamarock, J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, W. Wang, and J.G. Powers. A description of the Advanced Research WRF version 2. Technical Report Tech Notes-468+ STR, NCAR, 2005.

[31] R. E. Stoll and F. Von Linde. Hydrogen - what are the costs? *Hydrocarbon Processing*, 79(12):42–46, 2000.

[32] P. E. Thornton and S. W. Running. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agricultural and Forest Meteorology*, 93(4):211 – 228, 1999.

[33] O. Ulleberg. *Stand-Alone Power Systems for the Future: Optimal Design, Operation and Control of Solar-Hydrogen Energy Systems*. PhD thesis, Trondheim, Norway, 1998.

[34] O. Ulleberg. The importance of control strategies in PV-hydrogen systems. *Solar Energy*, 76:323–329, 2004.

[35] Kansas State University. Comparing fuel costs of heating and cooling systems. *Technical Report*, 2003.

[36] S. R. Vosen and J. O. Keller. Hybrid energy storage systems for stand-alone electric power systems: optimization of system performance and cost through control strategies. *International Journal of Hydrogen Storage*, 24(12):1139–1156, 1999.

[37] A. Wächter and L. T. Biegler. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Math. Program.*, 106:25–57, 2006.

[38] D. C. White. Online optimization: What have we learned? *Hydrocarbon Processing*, 77:55–69, 1998.

[39] V. M. Zavala, M. Anitescu, and T. Krause. On the Optimal On-Line Management of Photovoltaic-Hydrogen Hybrid Energy Systems. In R. M. de Brito Alves, C. A. Oller do Nascimento, and E. C. Biscaia Jr., editors, *Proceedings of 10th International Symposium on Process Systems Engineering*, 2009.

Mathematics and Computer Science Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 221
Argonne, IL 60439-4844

www.anl.gov

A U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC