10-19-2013

Uncertainty Propagation from Sensing to Modeling and Control in Buildings - Technical Report

Madhur Behl
University of Pennsylvania, mbehl@seas.upenn.edu

Truong Nghiem
University of Pennsylvania, nghiem@seas.upenn.edu

Rahul Mangharam
University of Pennsylvania, rahulm@seas.upenn.edu

Uncertainty Propagation from Sensing to Modeling and Control in Buildings - Technical Report

Abstract
A fundamental problem in the design of closed-loop Cyber-Physical Systems (CPS) is in accurately capturing the dynamics of the underlying physical system. To provide optimal control for such closed-loop systems, model-based controls require accurate physical plant models. It is hard to analytically establish (a) how data quality from sensors affects model accuracy, and consequently, (b) the effect of model accuracy on the operational cost of model-based controllers. We present the Model-IQ toolbox which, given a plant model and real input data, automatically evaluates the effect of this uncertainty propagation from sensor data to model accuracy to controller performance. We apply the Model-IQ uncertainty analysis for model-based controls in buildings to demonstrate the cost-benefit of adding temporary sensors to capture a building model. Model-IQ's automated process lowers the cost of sensor deployment, model training and evaluation of advanced controls for small and medium sized buildings. Model-IQ provides recommendation of sensor placement and density to trade-off the cost of additional sensors with energy savings by the improved controller performance. Such end-to-end analysis of uncertainty propagation has the potential to lower the cost for CPS with closed-loop model based control. We demonstrate this with real building data in the Department of Energy's HUB.

Keywords
energy efficient buildings, control systems, uncertainty analysis

Disciplines
Computer Engineering | Controls and Control Theory | Electrical and Computer Engineering

Comments
@TechReport{uncertinty2013, author = {Behl, M. and Nghiem, T. X. and Mangharam, R.}, title = "{Uncertainty Propagation from Sensing to Modeling and Control in Buildings - Technical Report}", institution = {University of Pennsylvania}, year = 2013, month = "Oct"}

Model-IQ: Technical Report
Uncertainty Propagation from Sensing to Modeling and Control in Buildings

Madhur Behl, Truong X. Nghiem and Rahul Mangharam
Dept. of Electrical and Systems Engineering
University of Pennsylvania
{mbehl, nghiem, rahulm}@seas.upenn.edu

***Abstract*— A fundamental problem in the design of closed-loop Cyber-Physical Systems is in accurately capturing the dynamics of the underlying physical system.**

I. Introduction

One of the biggest challenges in the domain of cyber physical energy systems is in accurately capturing the dynamics of the underlying physical system. In the context of buildings, the modeling difficulty arises due to the fact that each building is designed and used in a different way and therefore, it has to be uniquely modeled. Furthermore, each building system is a collection of a large number of interconnected subsystems which interact in a complex manner and are subjected to time varying environmental conditions.

Controls-oriented models are needed to enable optimal control in buildings. Learning mathematical models of buildings from sensor data has a fundamental property that the model can only be as accurate and reliable as the data that it was learned from. Any measurement exhibits some difference between the measured value and the true value and, therefore, has an associated uncertainty. Non-uniform measurement conditions (e.g. sensor placement and density), less accurate sensors and limited sensor calibration make the measurements in the field vulnerable to errors. With appropriate understanding of the source of error and their effect on the operating cost of the controller, the error associated with some of these sources can be minimized. In the case of using sensor data for training inverse models (e.g. grey box or black box), the goal is to provide maximum benefit, in terms of model accuracy, for the least sensor cost.

One major challenge to the use of models for buildings controls lies in understanding the impact of uncertainty in the model structure, the estimation algorithm, and the quality of the training data. It is known that the quality of the training data, characterized by uncertainty, depends on factors such as the accuracy of sensors, sensor placement and density, and the assumption that air is well mixed. It is intuitive to assume that installing additional sensors to obtain higher quality training data should result in more accurate models, which will further result in better performance of a model based controller (e.g. Model Predictive Control (MPC)). However an understanding of the cost-benefit associated with adding additional sensors to a building is either limited or missing altogether. The reason for this is twofold:

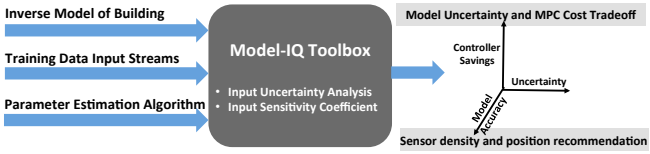


Fig. 1. Model-IQ Toolbox uncertainty tradeoff analysis for energy-efficient buildings

a) It is not clear how the data quality from sensors affects the accuracy of a building model. This is because data quality is only one of several factors that may affect model accuracy.
b) It is hard to analytically establish the effect of model accuracy on the performance of a model based closed-loop controller for buildings.

One of the goals of this paper is to study the cost-benefit effect of adding temporary low cost wireless sensors to a zone to enable implementation of advanced control schemes such as Model Predictive Control (MPC). We first perform an input uncertainty analysis to classify the effect of the quality of training data on the accuracy of a "grey-box" building inverse model. We show how additional sensors can affect the training data quality. We empirically evaluate the tradeoff between model accuracy and MPC performance, i.e. all things being the same, for different model accuracies how MPC performance varies. This paper has the following contributions:

1) ModelIQ, a methodology for offline assessment of training data quality versus model accuracy is presented. This scheme can be used to rank training inputs which affect the model accuracy the most.
2) A simulation based approach to study the influence of model accuracy on the performance of MPC for buildings is also presented.
3) The inferences from (1) and (2) allow us to formulate the trade-off between additional sensor cost and model accuracy for real buildings based on measurement data.

A. Understanding Sources for Uncertainty in Modeling

Uncertainty in modeling the dynamics of the underlying physical system is largely due to (a) the model structure, (b) the performance of the parameter estimation algorithm and (c) the uncertainty in the input-output data. In this

effort, we assume the first two are fixed as there are well-established inverse models for buildings. Our focus is thus on understanding the effect of uncertainty of the training data from the building and environment sensors on the overall cost of operating a model-based controller for the buildings HVAC systems.

The uncertainty in training data can be characterized in two ways: bias error or measurement noise (i.e. random error). Biases are essentially offsets in the observations form the true value. Bias error can also be referred to as the systematic error, precision or fixed error. The bias in the sensor measurement is due to a combination of two reasons. The first reason is the sensor precision. The best corrective action in this case is to ascertain the extent of the bias (using the data-sheet or by re-calibration) and to correct the observations accordingly. The sensor may also exhibit bias due to its placement, especially if it is measuring a physical quantity which has a spatial distribution, e.g. air temperature in a zone. In this case, it is hard to detect or estimate the bias unless additional spatially distributed measurements are obtained. Random error is an error due to the unpredictable and unknown extraneous conditions that can cause the sensor reading to take some random values distributed about a mean. Furthermore, random errors can be additive or multiplicative. Additive errors are independent of the magnitude of the observation while multiplicative error depend on the magnitude of the observation.

In buildings, the density and location of sensors in a zone effects the deviation of the measured temperature value from the true temperature value. For instance, a temperature sensor placed too close to the wall, window, supply or return air duct can introduce a bias in the sensor measurement. As we will see in the next section, a bias in the zone temperature value can lead to wastage of energy and discomfort with simple zone air control schemes like On-Off and PID control.

B. A Simple Example of Uncertainty Propagation

In this example we examine how underestimating or over-estimating the temperature of a zone can directly affect the energy consumption and comfort levels in a zone. Consider the case of controlling a heater in a single zone house as shown in Fig 2. The set point of the temperature inside the zone is fixed at $21°$C. Due to the low temperature (average value $8°$C) outside the house the heater has to constantly work to maintain the zone temperature at the set-point.

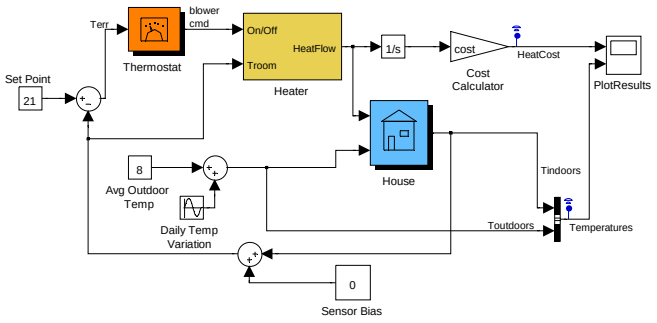


Fig. 2. Thermal model of a house with on-off control

We consider two different control strategies for the heater: On-Off control and PID control, which are widely use din existing buildings. In On-Off control the thermostat switches the heater on or off while ensuring that the temperature in the zone always remains within $\pm 2°$ of the set-point temperature. This simple control scheme is also referred to as 2-set thermostat control. In the second case, the PID controller directly controls the supply air temperature and flow into the zone in order to always maintain the set point temperature. The simulation period is run for 48h and the cost of energy is fixed at $15c$ per kW h. The baseline scenario is the case when there is no sensor bias in the zone temperature value. In this case the mean zone temperature for the duration of the simulation is always the same as the set point temperature of $21°$C for both On-Off and PID control. For the baseline case, the total energy cost for On-Off control is $\$36.41$ and $\$36.33$ for the PID case. The costs are nearly identical since the mean value of the zone temperatures are nearly the same in both cases.

In both the cases we deliberately introduce a fixed sensor bias and evaluate the zone output in terms of comfort and cost of energy. Let us assume that the bias occurs due to incorrect placement of the temperature sensor. Figure 3 shows the comparison of the zone temperature for different bias values for the two different control schemes. It also shows the change in the energy cost of the zone from the baseline cost for different values of the sensor bias. When the sensor underestimates the zone temperature by some bias value (say $-3°$C) the mean temperature of the zone increases by the same bias and causes the zone to overheat i.e. the mean zone temperature goes up to $24°$C from $21°$C. This causes the heating system to consume more energy as it needs to compensate for the underestimated zone temperature. On the other hand, if there is a positive sensor bias (say $+3°$C) then it causes the mean value to decrease by the amount of bias and results in a lower and much cooler zone temperature i.e. mean zone temperature drops from $21°$C to $18°$C. Although, the energy consumption for this case is less, it is only at a the cost of zone comfort. The change in energy cost is the same for both the control schemes. This is because the mean temperature of the zone for all values of the sensor bias is very similar for both the cases and hence they consume almost the same amount of total energy.

Therefore, a bias in the temperature measurement of the zone (due to the location and density of temperature sensors) affects the zone comfort and energy consumption. However, note that both the controllers are *model-independent* i.e. they only used the measurement of the *process variable* (which is the room temperature in this case) to compute the control signal that will either track the set-point (PID) or keep the temperature bounded around the set-point (On-Off). As we proceed to apply *model based control* schemes for building retrofits, the bias or the uncertainty in the measured data will also influence the accuracy of the model itself which in turn affects the performance of the model based controller, which is the focus of this paper.

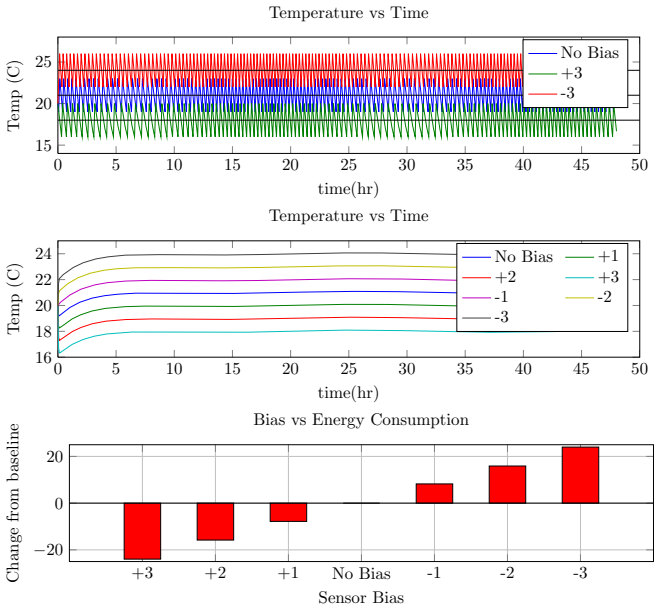


Fig. 3. The zone temperature for the simulation period is plotted for different temperature sensor bias values for: (a) On-Off control, and (b) PID control. (c) shows the change in the energy cost of the zone from its baseline value for different bias values

Organization: This paper is structured as follows. We begin with a short primer on the inverse modeling process for buildings in Section II. The Model-IQ approach for input uncertainty analysis for inverse models is presented in Section III. In Section IV, we quantify the effect of the model accuracy on the performance of a model predictive controller. Section V then presents a data-driven case study in which we demonstrate our approach on sensor data obtained from a real building. Section VI presents the result of an experiment to understand the relationship between the quality of data and the location of the sensor. Section IX follows the related work and concludes the paper with a discussion on the use of the free and open-source Model-IQ toolbox.

II. Inverse Modeling

The main objective of an HVAC system for air temperature control is to reject disturbances due to outside weather condition and internal heat gain caused by occupants, lighting and plug-in appliances. Therefore, the building model must accurately capture the thermal response of the building to the different disturbances.

Building models can be broadly classified into three categories:

1) *White-box* models are based on the laws of physics and permit accurate and microscopic modeling of the building system. High fidelity building simulation programs like EnergyPlus and TRNSYS ♡[1] fall into this category. Although such models provide a high degree of accuracy they are unsuitable for control design due to their high level of complexity. These models use a large number of parameters which must be obtained from a detailed description of the building. Furthermore, the process of constructing the model and tuning the parameters with limited data is very time consuming and not cost effective.
2) *Black-box models* are not based on physical behaviors of the system but rely on the available data to identify the model structure. These models are often purely statistical and have a simple structure (e.g. linear regression). However, they provide little insight into the dynamics dictating the system behavior.
3) *Grey-box models* fall in between the two above categories. A simplified model structure is chosen loosely based on the physics of the underlying system and the available data is used to estimate the values of the model parameters. These models are suitable for control design and still respect the physics of the system.

[1]RAHUL: Need citations

A. Model Structure

While there are several methods available for modeling the dynamics of a building, in this paper we will focus on the analysis of only grey-box RC models. A commonly used grey-box representation of the thermal response of a building due to heat disturbances uses a lumped parameter Resistive-Capacitative (RC) network. This approach for modeling buildings has been used widely, e.g. in [1], [2], [3]. Figure 4 shows an example of such a model for a single zone, as used in [1]. In this representation, the central node of the RC network represents the zone temperature $T_z(°\text{C})$. The geometry of the zone is divided into different kinds of surfaces, each of which is modeled using a 'lumped-parameter' branch of the network. For instance, all the external walls of the zone are lumped into a single wall with 3R2C (3 resistances and 2 capacitance) parameters. The same process is applied to the ceiling, the floor and the internal (or adjacent) walls of the zone. The zone is subject to several (heat) disturbances which are applied at different nodes in the network in the following manner: (a) solar irradiation on the external wall $\dot{Q}_{sol,e}(\text{W})$ and the ceiling $\dot{Q}_{sol,c}(\text{W})$ is applied on the exterior node of the lumped wall. (b) incident solar radiation transmitted through the windows $\dot{Q}_{solt}(\text{W})$ is assumed to be absorbed by the internal and adjacent walls, (c) radiative internal heat gain $\dot{Q}_{rad}(\text{W})$ which is distributed with an even flux to the walls and the ceiling, (d) the convective internal heat gain $\dot{Q}_{conv}(\text{W})$ and the sensible cooling rate $\dot{Q}_{sens}(\text{W})$ is applied directly to the zone air, (e) The zone is also subject to heat gains due to the ambient temperature $T_{amb}(°\text{C})$, ground temperature $T_g(°\text{C})$ and temperatures in other zones which are accounted for by adding boundary condition nodes to each branch of the network. The list of all parameters in the model and their descriptions is given in Table I.

Given this model, the nodal equations for the lumped

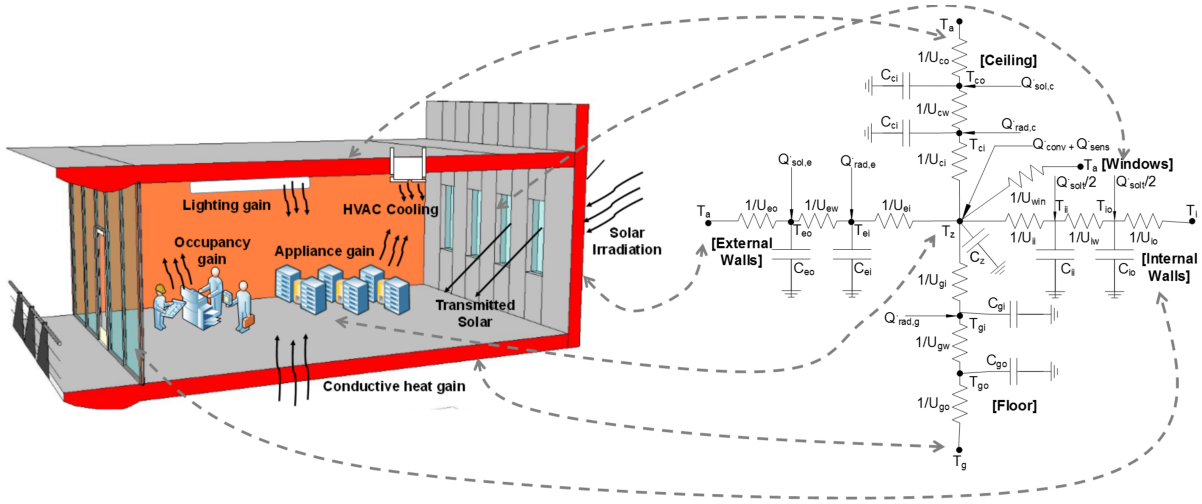


Fig. 4. RC lumped-parameter model representation for a thermal zone

external wall and the ceiling network are:

$$C_{eo}\dot{T}_{eo}(t) = U_{eo}(T_a - T_{eo}(t)) + U_{ew}(T_{ei}(t) - T_{eo}(t)) + \dot{Q}_{sol,e} \tag{1a}$$

$$C_{ei}\dot{T}_{ei}(t) = U_{ew}(T_{eo}(t) - T_{ei}(t)) + U_{ei}(T_z(t) - T_{ei}(t)) + \dot{Q}_{rad,e} \tag{1b}$$

$$C_{co}\dot{T}_{co}(t) = U_{co}(T_a - T_{co}(t)) + U_{cw}(T_{ci}(t) - T_{co}(t)) + \dot{Q}_{sol,c} \tag{1c}$$

$$C_{ci}\dot{T}_{ci}(t) = U_{cw}(T_{co}(t) - T_{ci}(t)) + U_{ci}(T_z(t) - T_{ci}(t)) + \dot{Q}_{rad,c} \tag{1d}$$

TABLE I
LIST OF PARAMETERS

$U_{\star o}$	convection coefficient between the wall and outside air
$U_{\star w}$	conduction coefficient of the wall
$U_{\star i}$	convection coefficient between the wall and zone air
U_{win}	conduction coefficient of the window
$C_{\star\star}$	thermal capacitance of the wall
C_z	thermal capacity of zone z_i
g	floor
e	external wall
c	ceiling
i	internal wall

Similarly, one can write the equations for the dynamics of the nodes of the floor and internal wall network. The law of conservation of energy gives us the following heat balance equation for zone

$$\begin{aligned} C_z\dot{T}_z(t) = {} & U_{ei}(T_{ei}(t) - T_z(t)) + U_{ci}(T_{ci}(t) - T_z(t)) \\ & + U_{ii}(T_{ii}(t) - T_z(t)) + U_{gi}(T_{gi}(t) - T_z(t)) \\ & + U_{win}(T_a(t) - T_z(t)) + \dot{Q}_{conv} + \dot{Q}_{sens} \end{aligned} \tag{2}$$

Differential equations (1a) to (1d) and (2) can be combined to give the state space model of the system:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{3}$$

where the state x is a vector of all node temperatures of the model, i.e. $x = [T_{eo}, T_{ei}, T_{co}, T_{ci}, T_{go}, T_{gi}, T_{io}, T_{ii}, T_z]^T$. The input u is a vector of all the inputs to the systems, i.e. $u = [T_a, T_g, T_i, \dot{Q}_{sol,e}, \dot{Q}_{sol,c}, \dot{Q}_{rad,e}, \dot{Q}_{rad,c}, \dot{Q}_{rad,g}, \dot{Q}_{solt}, \dot{Q}_{conv}, \dot{Q}_{sens}]^T$. The control input to the zone is the sensible cooling rate $\dot{Q}_{sens}$. The cooling rate can be controlled by changing the mass flow rate of cold air which enters the zone (in case of cooling) or by changing the set point of the supply air temperature. The rest of the inputs are disturbances to the zone. The elements of the state matrix A and the input matrix B depend non-linearly on the U and C parameters of the model. Let us consider $\theta = [U_{eo}, U_{ew}, U_{ei}, \dots C_{io}, C_{ii}]^T$, as a vector of all the parameters of the model Then the state space equations have the following representation which emphasizes the parameterization of the A, B, C and D matrices.

$$\begin{aligned} \dot{x}(t) &= A_\theta x(t) + B_\theta u(t) \\ y(t) &= C_\theta x(t) + D_\theta u(t) \end{aligned} \tag{4}$$

The entries of output matrix C and the feed-forward matrix D depend on the output of interest. For instance, if the output of the model is the zone temperature then $C = [0, 0, \dots, 0, 1]$, which is a row vector with all entries equal to zero except the last entry corresponding to the zone temperature T_z equal to one. In this case $D = \mathbf{0}$, the null matrix. The model structure is based on the underlying assumption that the air inside the zone is well mixed and thus, it can be represented by a single node. Furthermore, only one dimensional heat transfer is assumed for the walls and there is no lateral temperature difference. The parameters of the model are assumed to be time invariant.

B. Parameter Estimation (Model Training)

We first consider the discrete-time state space representation of the dynamical system of Eq. (4)

$$\begin{aligned} x(k+1) &= A_\theta x(k) + B_\theta u(k) \\ y(k) &= C_\theta x(k) + D_\theta u(k) \end{aligned} \tag{5}$$

The goal of parameter estimation is to obtain estimates of the parameter vector θ of the model from input-output time series measurement data. The parameter search space is constrained both above and below by $\theta_l \leq \theta \leq \theta_u$. For a given parameter set θ, the model, given by Eq. (5), can be

used to generate a time series of the zone air temperature T_{z_θ} using the measured time series data for the inputs $u(t)$. The subscript θ denotes that the temperature value T_{z_θ} is the predicted value using the model with parameters θ and the inputs u. This model generated time series T_{z_θ} may then be compared with the corresponding observed values of the zone temperature T_{z_m}, and the difference between the two is quantified by a statistical metric. The metric usually chosen is the sum of the squares of the differences between the two time series. The parameter estimation problem is to find the parameters θ^*, subject to $\theta_l \leq \theta \leq \theta_u$, which result in the least square error between the predicted and the measured temperature values, i.e.

$$\min_{\theta^*} J = \sum_{i=1}^{N} (T_{z_m}(i) - T_{z_\theta}(i))^2 \tag{6}$$

where the summation is over the N data points of the input-output time series under investigation.

The least square optimization of Eq. (6) is a constrained minimization of a non-linear objective. It is numerically solved using a trust region reflective algorithm [4] such as the Levenberg-Marquardt [5] algorithm. A well known problem with non-linear search algorithms is the problem of the solution getting stuck at a local minima. For this reason, it is required that the initial parameter estimates θ_0 should be as close as practicable to their (unknown) optimal (true) values. Generally speaking, the further the initial guess about the parameters values is away from the true values, the bigger the search region is for the optimization. As search region grows, it becomes more likely that the estimation process will converge to a local optimum.

III. ModelIQ

In this section, we describe the ModelIQ approach for analyzing uncertainty propagation for building inverse models. We describe the methodology through an elaborated case study.

The accuracy of the building inverse model depends primarily on the following three factors:

(a) **The structure of the model** which depends on the extent to which the model respects the physics of the underlying physical system,

(b) **The performance of the estimation algorithm**. As discussed previously, in the case of non-linear estimation the performance of the algorithm depends heavily on the nominal values of the parameters, and

(c) **The quality of the training data**, which can be characterized by its uncertainty.

The main premise of the input uncertainty propagation is that once the model structure and the parameter algorithm are fixed, one can study the influence of the uncertainty in the training data on the accuracy of the model using virtual simulations which utilize artificial data-sets. For the remaining part of the section we describe the results of conducting the input uncertainty propagation analysis for a virtual building modeled in TRNSYS.

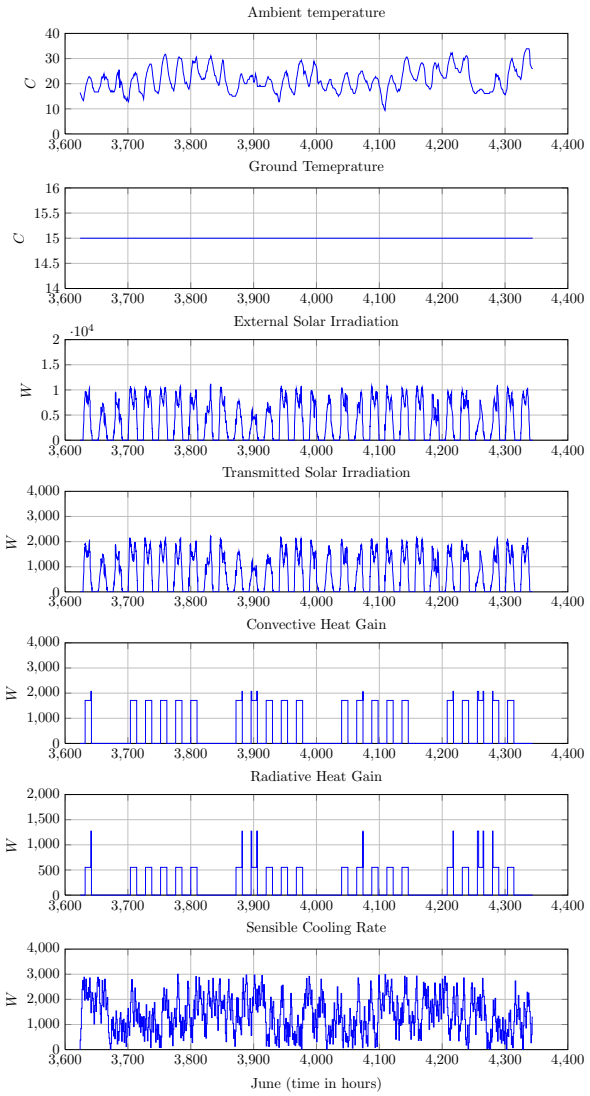
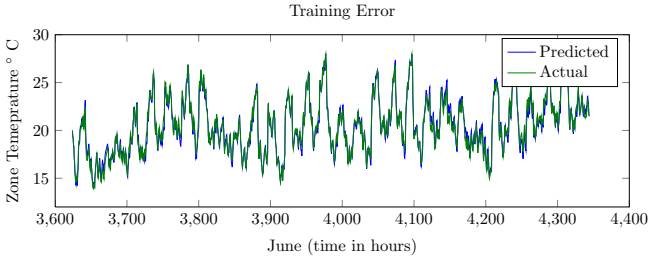


Fig. 5. Inputs sampled at 2 minutes during the month of June (hour 3624 to 4344) were used for training the single zone model.

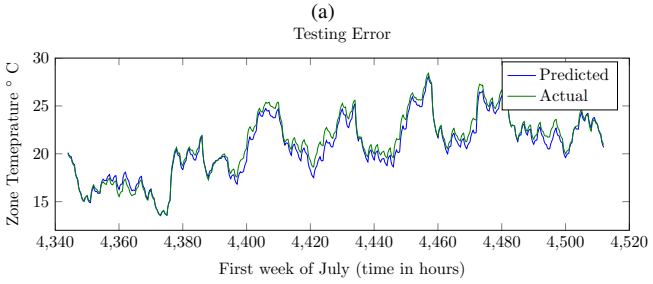
A. Inverse Model

The test-bed used for the input uncertainty analysis is a single zone building modeled in TRNSYS. The building is north facing, has 4 external brick walls each of which contains a large window, a concrete ceiling and a floor. For the simulation we use the `Philadelphia – TMY2` weather file which provides the ambient temperature and solar irradiation data for modeling. The building is assumed to be equipped with a HVAC system with a maximum cooling power of 3.5kW. In addition to the heat gains due to outside temperature and incident irradiation, the building is also subject to internal heat gains from occupants, appliances and lighting fixtures. Without lack of generality we only consider the case when the building is being cooled. The operation of the heating system would be similar.

The objective is to construct an inverse model for the thermal response of the building which can be used for model based control. A lumped parameter RC model was constructed for the building with a structure similar to the model explained in Section II. The model contains 12 RC



(a)



(b)

Fig. 6. The fit between the predicted and actual values of the zone temperature for the training period (June) is shown in (a) while, (b) shows the fit between the predicted and actual zone temperature values for the testing period which was the first week of July

parameters which need to be estimated. The inverse model contains a total of seven inputs and an output. The six disturbance inputs are: the ambient temperature (T_a), the ground temperature (T_g), the external incident solar irradiation (Q_{sole}), the solar irradiation transmitted through the windows (Q_{soltr}), radiative heat gain (Q_{grad}) and convective internal heat gain (Q_{conv}). The output of the inverse model is the temperature of the zone while the control input is the sensible cooling rate (Q_{sen}). The training data for the model is in the form of time-series data for each of the inputs and the output. The training period is the month of June (hours 3624-4344 in TRNSYS). The input-output time-series data is generated at a sampling rate of 2 minutes for the entire training period. All the different training inputs used for inverse modeling are shown in Fig. 5.

For this case study, the nominal values of the RC parameters of the model were estimated from the construction details of the building, obtained from TRNSYS. Fig. 6(a) shows the result of non-linear parameter estimation problem (6) for the training period. The result of the inverse model training are estimates of the RC parameters, and hence the matrices A,B,C and D of the state-space model. The comparison between the predicted zone temperature values from the model and the actual zone temperature is shown. The root mean square error (RMSE) of the fit was 0.187 and the R^2 value is 0.971. The R^2 coefficient of determination is a statistical measure of the goodness of fit of a model. Its value lies between $[0, 1]$ with a value of 1 indicating that the model perfectly fits the data. The R^2 coefficient also indicates how much of the variance of the data can be described by the model. However, measuring the fit statistics on the training period alone is never sufficient, since the model may be over-fitting the data. Therefore, the accuracy of the inverse model was also tested on a test data set. The test data set corresponds to the first week of July (hours 4344-4512 in TRNSYS). The time-series of inputs for the test period were used with the learned model and the results of the comparison between the predicted model output and the actual zone temperature is shown in Fig. 6(b). The RMSE for the testing period was 0.292 with a R^2 value of 0.961. These stats are a better indicator of the accuracy of the model since during the testing period the model is subject to an input data-set that it was not trained on.

B. Input Uncertainty Analysis

The aim of this analysis is to determine the influence of the uncertainty (bias) in the training data inputs on the accuracy of the inverse model and then, to quantify the relative importance of the inputs. First, some notation is introduced for brevity. We consider a model with $m > 0$ training input data sets denoted by $U = \{u_1, \cdots, u_m\}$. Note that these are inputs for model training, not the inputs for the model itself, e.g. even though zone temperature is a model output, it is still a required data-set (hence, an input) for model training. $U_{i,\delta} = \{u_i = u_i + \delta, u_j = u_j | i, j \leq m, j \neq i\}$ denotes the artificial data-set obtained by perturbing input u_i by an amount δ while keeping all other inputs data sets unperturbed. U_0 denotes the data-set in which all the inputs are unperturbed. Now, $\hat{M}_{U_{i,\delta}}$ is the inverse model with obtained by training on the data-set $U_{i,\delta}$ and $\hat{M}_{U_0}$ is the model obtained by training on a completely unperturbed data-set. We denote the RMSE of the model $\hat{M}_{U_{i,\delta_k}}$ by $r(\hat{M}_{U_{i,\delta_k}})$. The ModelIQ approach for conducting an input uncertainty analysis consists of the following steps:

(a) Establish a baseline (reference) model: The baseline model, $\hat{M}_{U_0}$, is the inverse model obtained by training on the unperturbed data set U_0, which is considered as the ground truth.
(b) Determine which model outputs will be investigated for their accuracy and what are their practical implications.
(c) Each of the input data streams are then perturbed within some bounds. There are a total of N perturbations $\delta_1, \cdots, \delta_N$ for each input stream $u_i, i \leq m$. This results in N artificial data-sets $U_{i,\delta_1}, \cdots, U_{i,\delta_N}$ for each input stream i.
(d) Corresponding to every perturbation, the inverse modeling process is run again and a new model $\hat{M}_{U_{i,\delta_k}}$ is obtained.
(e) The prediction accuracy of each of the trained model is evaluated on a common input data stream U_T. The accuracy of the model $\hat{M}_{U_{i,\delta_k}}$ is measured by the RMSE $r(\hat{M}_{U_{i,\delta_k}})$ between the predicted and the actual model output values for the common input stream U_T.
(f) Using the RMSE of the fit and the magnitude of the perturbation, determine the sensitivity coefficient (or influence coefficient) for each input training stream.

An overview of the steps for the input uncertainty analysis is shown in Fig. 8.

For our case study, the baseline model is the model trained on unperturbed training data-set i.e. the original input data-set corresponding to the month of June. The RMSE for the

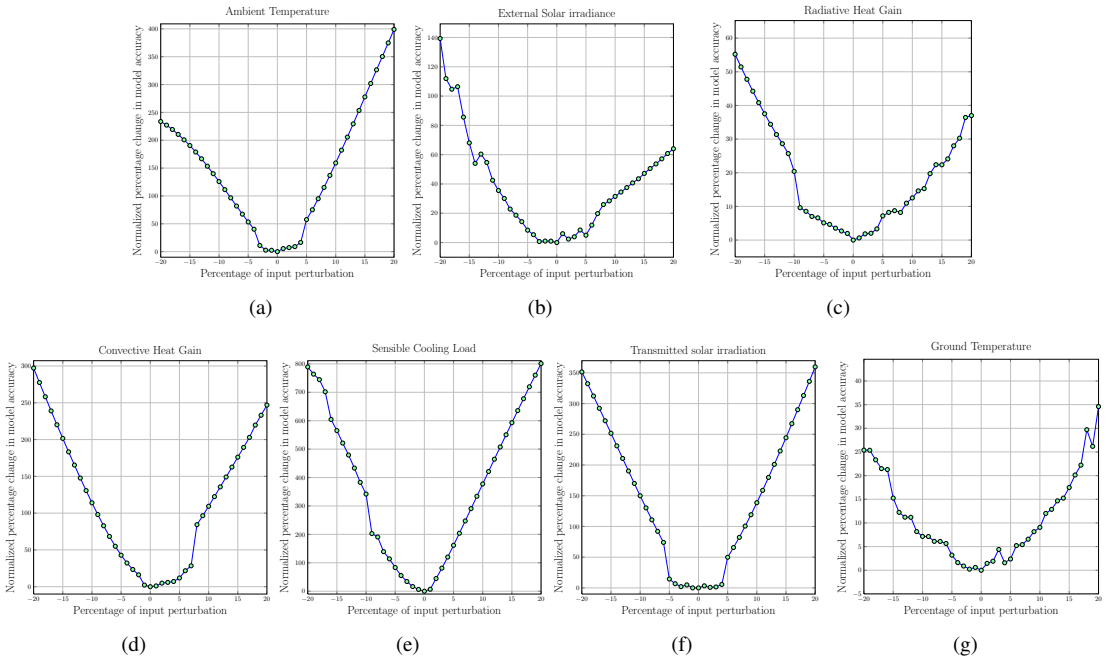


Fig. 7. Input uncertainty analysis results for a single zone TRNSYS model. The x axis shows the magnitude of the perturbation in percent change from the unperturbed data while the y axis is the percent change in the model accuracy compared to the RMSE for the model trained on unperturbed data. The following inputs are shown: (a) ambient temperature (°C); (b) incident solar irradiation on the external walls (W); (c) radiative internal heat gain (W); (d) convective internal heat gain (W); (e) sensible cooling rate (W); (f) solar irradiation transmitted through the windows (W); and, (g) floor (ground) temperature (°C).

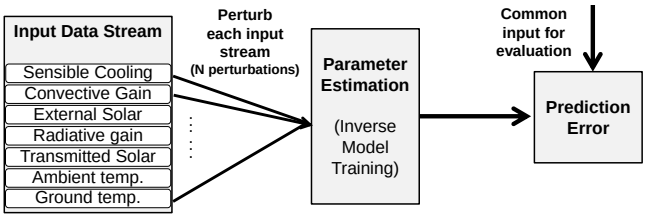


Fig. 8. Overview of the ModelIQ input uncertainty analysis methodology, an offline method to confirm the influence of each training input on the accuracy of the model.

baseline model is denoted by $r(M_{U_0})$. The artificial data-sets are created in a normalized manner by adding (and subtracting) a bounded and fixed bias to the unperturbed data in the form of the per-cent change from the unperturbed (baseline) value i.e. the perturbations δ_k's are in form of per-cent changes around the unperturbed data point. Therefore, each data-point x_i belonging to the unperturbed input u_i gets perturbed to a new value of $\tilde{x}_i = x_i(1 + \delta_k/100)$. This is done so that every input is treated in the same manner regardless of the scale of the input. One can relate the per-cent change to the absolute value of the change, simply through the mean of the data-set. For e.g. if the mean of unperturbed ambient temperature was 20°C, then the mean of data which was perturbed $\delta = +10\%$ would be 22°C which is equivalent to a mean absolute bias of 2° degrees in the ambient temperature. Each of the 7 training input data streams (Fig. 8) are perturbed one at a time within $[-20\%, 20\%]$ around the unperturbed nominal value with increments of 1%. Every perturbation for each of the inputs creates an artificial training set for the inverse model. Therefore for each of the 7 input streams, $N = 40$ additional artificial data-sets $U_{i,\delta_1}, \cdot, U_{i,\delta_40}$ were created resulting in a total 280 different training data-sets. The inverse model for the single zone building was trained on each of the artificial data-set and the accuracy of the model was evaluated in terms of the RMSE on the test data-set. The use of a common test data-set for evaluating the accuracy of the model ensures a fairness in the comparison of the influence of the uncertainty among different inputs on the model accuracy.

Finally the model accuracy sensitivity coefficient is calculated as follows:

$$\gamma_i = \text{Mean}_{(k=1,\cdots,N)} \left(\frac{r(\hat{M}_{U_{i,\delta_k}}) - r(M_{U_0})/r(M_{U_0})}{|\delta_k|} \right) \tag{7}$$

It is the mean of the ratio of the normalized change in the model accuracy to that of the normalized change in the magnitude of the input data stream. Both normalization's are with respect to the baseline case. The magnitude of the sensitivity coefficient γ_i can be interpreted as the mean value of the change in the RMSE of the model due to 1% bias uncertainty in the training data stream i. The sensitivity coefficient is sometimes also referred to as the influence coefficient or point elasticity. The results for the input uncertainty analysis for the TRNSYS building are shown in Fig. 7. These results align well with the intuition that as the magnitude of the uncertainty bias increases in the input data stream the inverse model becomes worse and its prediction error increases. This is the case for all the input data streams and it results in the parabolic trend. The shape of the curve varies from input to input, due to a different sensitivity coefficient value, and is an

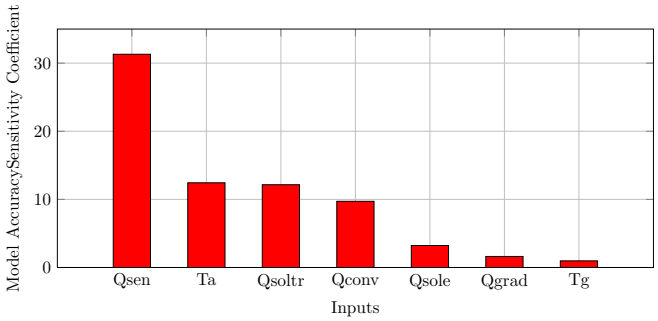


Fig. 9. Comparison of model sensitivity coefficients for the different input training data streams.

indicator of the extend to which a particular input influences the model accuracy.

The model accuracy sensitivity coefficients were calculated for every input data stream and their comparison is shown in Fig. 9. For the building under consideration, it is clear that the sensible cooling rate, ambient temperature, transmitted solar gain and convective heat gains should be measured accurately in order to learn a good inverse model. Note that, although the results presented in this paper assume a particular building and a fixed model structure, the ModelIQ approach itself is a general approach which can be used to identify the inputs which should be measured more accurately in order to obtain accurate building models.

IV. Model accuracy vs MPC performance

While the input uncertainty analysis reveals important insights about the relationship between data quality and model accuracy, it is not clear how much of an impact does model accuracy have on the performance of a model based building controller. Therefore, it is also necessary to examine if the model accuracy can have any direct economic impact. Especially when energy-efficient control algorithms rely on the accuracy of the underlying mathematical model of the building in order to figure out optimal control inputs. Installation of additional sensors in a building can yield better quality of data for model training but there is a trade-off between how good can an inverse model can perform versus how much cost is spent on obtaining the inverse model. These trade-offs can be better understood if an end-to-end relationship between the data uncertainty, model accuracy and control cost are known. There is significant value in knowing how much the cost of a model predictive controller changes with changes in the model accuracy. This information can be used to provide "target" accuracy levels for the inverse model, which in turn specify the degree of accuracy required on the sensing.

However, it is a hard problem to analytically determine the impact of the model accuracy on the MPC performance. The problem arises due to the complexity of the model structure and the MPC formulation itself. For this reason, to quantify the effect of model accuracy on the performance of a MPC controller we make use of an empirical analysis with the same single zone TRNSYS model used in section III. First, a model predictive controller was designed for the zone. The MPC simulation is then ran for models of different

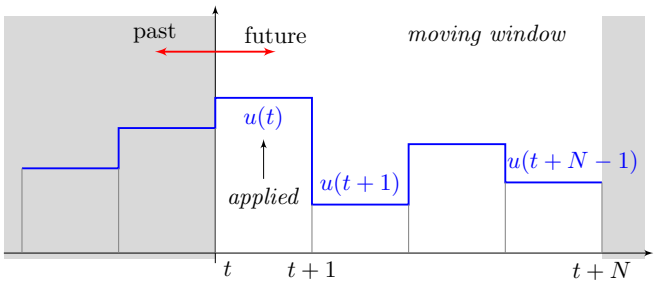


Fig. 10. Finite-horizon moving window of MPC: at time t, the MPC optimization problem is solved for a finite length window of N steps and the first control input $u(t)$ is applied; the window then recedes one step forward and the process is repeated at time $t+1$.

accuracy and the outputs are compared to reveal the trend of model accuracy vs MPC cost. We now describe the MPC formulation followed by the results of this analysis.

A. MPC formualtion

The MPC problem involves optimizing a cost function subject to the dynamics of the system, over a finite horizon of time. Based on the cost function and the constraints on the zone temperature, the MPC controller calculates a sequence of control inputs which minimize the cost for the length of the finite horizon. The first computed input is applied, and at the next step the optimization is solved again as shown in Figure 10.

The state-space model zone model in eq (5) can also be written as follows:

$$x(k+1) = Ax(k) + Bu(k) + Ed(k) \tag{8a}$$

$$y(k) = Cx(k) \tag{8b}$$

Where the cooling rate to the zone is the control input $u(k)$ and $d(k)$ is the vector of all the disturbances into the zone (ambient temperature, heat gains etc).

To reduce the number of control variables we use the move-blocking technique. During each move-blocking window of length l, the control u is held constant. So $u(0) = u(1) = \cdots = u(l-1)$, $u(l) = u(l+1) = \cdots = u(2l-1)$ and in general $u(il) = u(il+1) = \cdots = u((i+1)l-1)$ i.e. MPC re-optimizes at integral multiples of the window length il only.

Let us consider a control horizon H in terms of move-blocking windows, so the number of time steps is Hl. At time $t = il$, the MPC problem is to minimize

$$\sum_{k=0}^{H-1} \sum_{\sigma=t+kl}^{t+(k+1)l-1} \left(P_U(\sigma)u(k) + P_T(\sigma)\left(y(\sigma - y_{sp}(\sigma)\right)^2 \right)$$

subject to

$$x(t) = x_0$$
$$\begin{bmatrix} cx(t+kl+1) \\ \vdots \\ x(t+(k+1)l) \end{bmatrix} = \mathrm{diag}(A) \begin{bmatrix} cx(t+kl) \\ \vdots \\ x(t+(k+1)l-1) \end{bmatrix} + \mathrm{col}(B)u(k) + \mathrm{diag}(E) \begin{bmatrix} cd(t+kl) \\ \vdots \\ d(t+(k+1)l-1) \end{bmatrix}$$
$$u_{min}(\sigma) \leq u(k) \leq u_{max}(\sigma)$$

where the last two constraints hold for all $k = 0, \ldots, H-1$ and $\sigma = t+kl, \ldots, t+(k+1)l-1$, $\mathrm{diag}(\cdot)$ represents a block diagonal matrix of appropriate dimensions, and $\mathrm{col}(B)$ is the column vector constructed by stacking the columns of matrix B. $P_U(\sigma)$ is the price of electricity at time σ and $P_T(\sigma)$ is the penalty for errors in tracking the desired zone temperature trajectory $y_{sp}(\sigma)$. Both the cost and penalty functions vary throughout the day, for example the price of electricity can be high during the peak hours of the day as compared to the off peak hours. Similarly, the set-point temperature can change during the day depending on the zone occupancy. Note that in our MPC formulation we only consider soft constraints on the zone temperature. The initial state of the system is x_0, while $u_{min}(\sigma)$ and $u_{max}(\sigma)$ are the lower and upper bounds on the cooling rate $u(k)$ which can vary during the day to account for equipment schedules.

B. State Observer

In order for us to use the state space model (5) for model predictive control, we also need to design a state observer. The state observer is designed to provide the estimates of $\hat{x}(k|k)$, the state of the plant model at every MPC time step. The estimates are computed from the measured output $y_m(k)$ by the linear state observer. The reason for estimating the states of the plant is that the states $x_1, \cdots, x_{n-1}$ are lumped parameter temperatures which are hard (and almost impossible) to measure.

Given the discrete plant model

$$x(k+1) = Ax(k) + Bu(k) + Gw(k)$$
$$y(k) = Cx(k) + Du(k) + Hw(k) + v(k)$$

where, $w(k)$ is white process noise and $v(k)$ is white measurement noise satisfying $E(w(k)) = E(v) = 0$, $E(ww^T) = Q$, $E(vv^T) = R$, $E(wv^T) = N$. The estimator has the following state equation:

$$\hat{x}(k+1|k) = A\hat{x}(n|n-1) + Bu(k) + L(y(k) - C\hat{x}(k|k-1)) \quad (9)$$

The gain matrix L is derived by solving a discrete Riccati equation:

$$L = (APC^T + \bar{N})(CPC^T + \bar{R})^{-1} \quad (10)$$

where,

$$\bar{R} = R + HN + N^T H^T + HQH^T \quad (11)$$
$$\bar{N} = G(QH^T + N) \quad (12)$$

The prediction $\hat{x}(k|k-1)$ is updated using the new measurement $y(k)$ as:

$$\hat{x}(k|k) = \hat{x}(k|k-1) + M(y(k) - C\hat{x}(k|k-1) - Du(k))$$

where the innovation gain M is defined as:

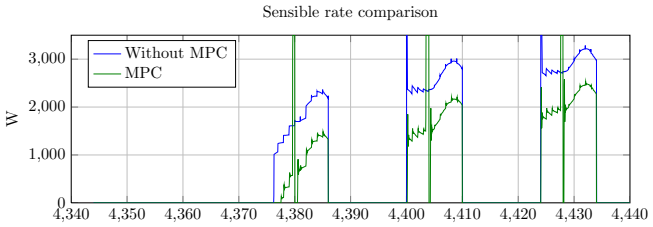
$$M = PC^T(CPC^T + \bar{R})^-1$$

For the simple case, $E(wv^T) = N = 0$, $D_d = D_u = 0$, $H = 0$, and $G = B$.

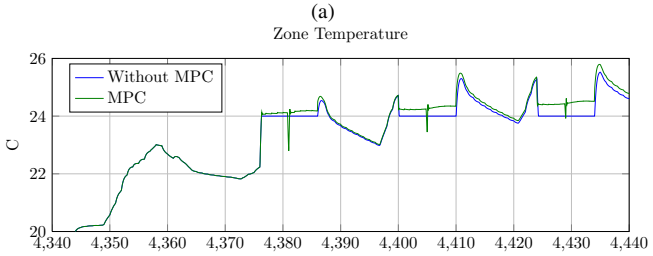
C. Single zone example

The MPC control described above was implemented for the single zone TRNSYS model. The cooling system of the building is always switched on during the occupancy period from 8 a.m. in the morning to 6 p.m. in the evening on weekdays and remains off during the weekend. The maximum and minimum constraints on the cooling rate were $u_{max} = 3500(\mathrm{W})$ and $u_{min} = 0$. The set point temperature of the zone was kept at 24°C for the occupancy period. The zone temperature is allowed to float during the weekend. The simulation was ran for a part of the first week of July from hour 4344-4400 in TRNSYS. The building is also subject to peak demand pricing, i.e. the price of electricity is 10 time the nominal price during the on-peak hours , which are from 1 p.m. to 5 p.m.. We compare two different cases. The first is the comparison of the building operation with and without an MPC controller. Without MPC control, the cooling switches on at 8 am and then tries to supply exactly the amount of cooling energy required to keep the temperature at 24 degrees for the occupancy period. The total energy consumption for the simulation period is $93.71(\mathrm{kW\,h})$. In this case, the power consumption of the cooling system remains high even during the peak pricing hours resulting in a total cost of 511.83 units.

The baseline model for MPC is the model with the best RMSE (0.187) for the testing data. This is the model which was trained on unperturbed data and was also used as the baseline for the input uncertainty analysis. The move-blocking step of MPC is 5 minutes and the MPC horizon is 2 hours. The performance of the MPC controller with the case without MPC is shown in figure 11. The MPC controller rapidly pre-cools the zone just before the peak power pricing period begins at 1 p.m. This can be seen in both the cooling rate and the zone temperature plots. As a result of this, the energy consumption of the building during the peak hours is reduced and it results in an overall lower energy cost. The total energy consumption for this case was $87.29(\mathrm{kW\,h})$ and the total energy cost was reduced to 442.06 units. So there is a 13.63% reduction in the energy cost and a 6.85% reduction in the total energy consumption. While the primary reason for the reduction in cost is the pre-cooling of the zone, another reason for the lower energy consumption is

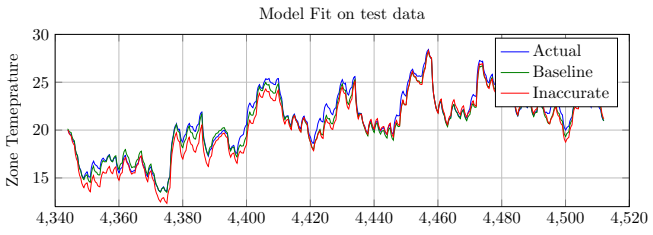


(a)

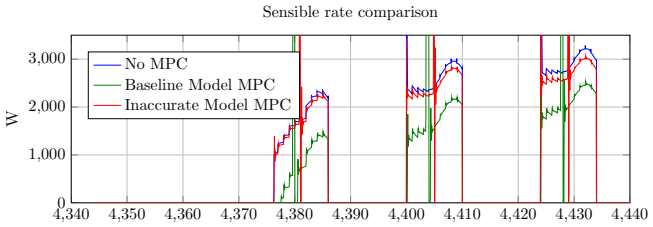


(b)

Fig. 11. (a) shows the comparison between the performance of an MPC controller with the default case without MPC. (b) shows the zone temperature values for both the controllers.



(a)



(b)

Fig. 12. (a) shows the comparison of the fit on the test data between the baseline model RMSE 0.187) and an inaccurate model (RMSE 0.538) (b) shows the comparison between the MPC performance of the models.

that because the MPC has soft temperature constraints, the zone temperature is slightly above the set-point temperature requiring it to use less cooling energy.

Having implemented MPC for the baseline model (trained on unperturbed data), we next use models trained on perturbed data and compare their performance with the baseline case. This allows us to observe the trend between MPC performance and model accuracy. An example of such a simulation run is shown in figure 12. The MPC performance of the baseline model has been compared with the performance of a relatively inaccurate model, with a much higher RMSE (0.538) value than the baseline model with RMSE (0.187). It can be seen that an inaccurate model performs poorly compared with the "good" baseline case. The total energy consumption for the inaccurate model was $91.68(\mathrm{kW\,h})$ which is about 2.2% less than the baseline case. The total energy cost was 492.53 units, which is a reduction of only

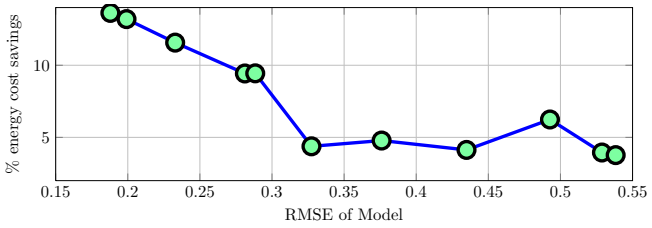
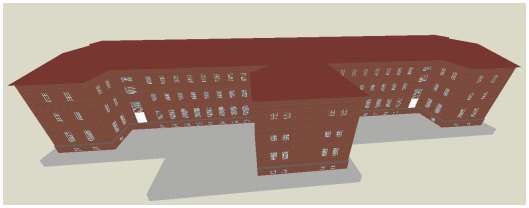
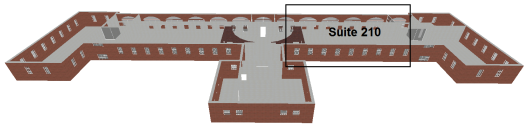


Fig. 13. Change in the performance of the model predictive controller for models of different accurateness.



(a)



(b)

Fig. 14. 3D view of Building 101, the site chosen for the case study and the location of suite 210 in the north-wing of the building.

3.77% as compared with the cost reduction of 13.63% for the baseline model based MPC. Several inverse models with different accurateness, in terms of their testing RMSE, were ran with the MPC controller. The savings achieved by MPC for the different models is shown in figure 13. The savings are measured with respect to the case when no MPC was used. The trend of the plot aligns with intuition and shows that MPC performance deteriorates as the underlying model becomes worse.

We have now seen that models can lose their predictive performance if they are trained on uncertain (biased) data. The input uncertainty analysis revealed the extent to which different inputs are responsible for the accuracy of the inverse model. By empirically, establishing a relationship between model accuracy and MPC performance, one can take informed decisions about the investment on additional sensor requirements to improve the data quality. In the next section we apply the ModelIQ tool-chain on a model for a real building using real sensor data.

V. Case Study

The ModelIQ approach described in Section III was applied to real sensor data. The site chosen for analysis is called Building 101. Building 101, located in the Navy Yard in Philadelphia, is the temporary headquarters of the U.S. Department of Energy's Energy Efficient Building Hub [6]. It is a highly instrumented commercial building and the acquired data is continuously stored and is made available to Hub researchers. The building is in the shape of a "T" with three wings (Figure 14(a)), and is comprised of offices, a lunchroom, mechanical spaces, and miscellaneous support

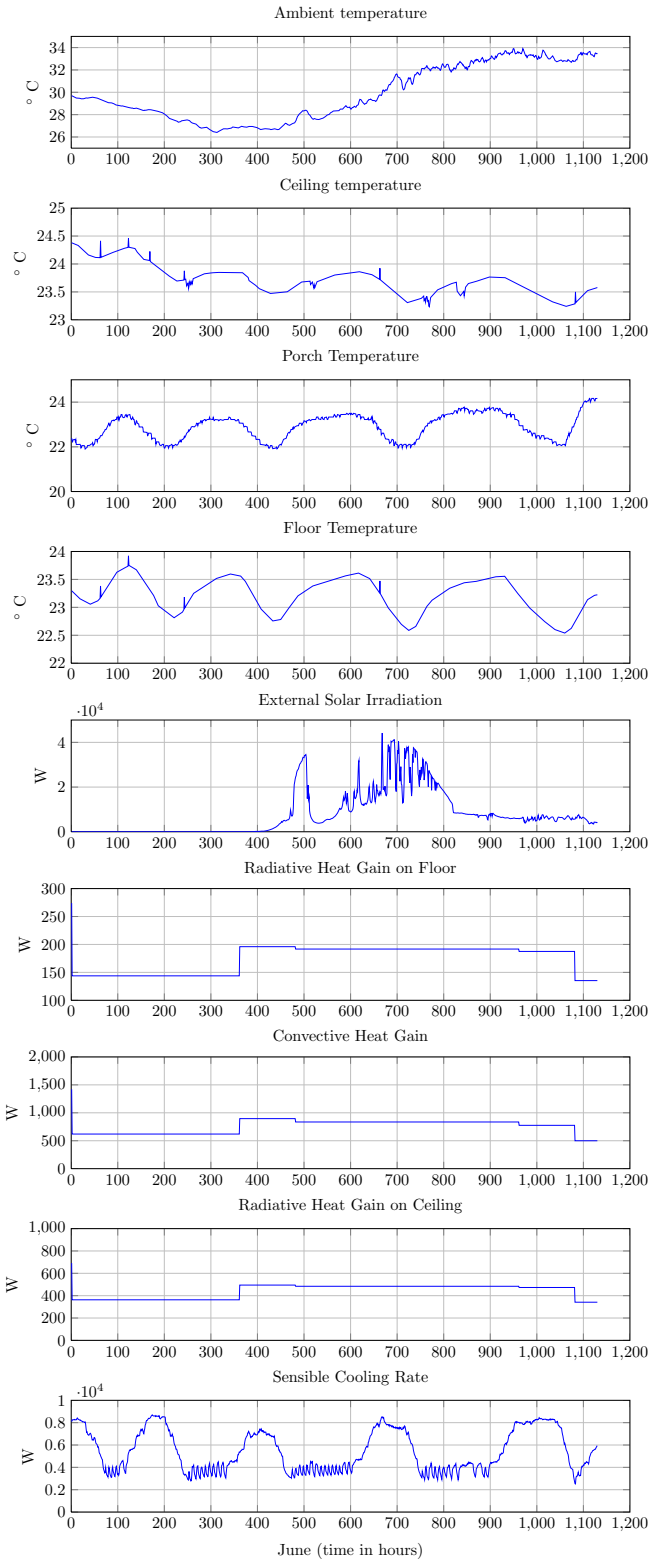


Fig. 15. Training data for the inverse model for Suite 210 of Building 101. The data obtained by running a functional test on the zone's air handling unit from 19-07-2013 18:07 to 20-07-2013 22:29 is partitioned into the training (80%) and test (20%) set for calculating the training and test fit error

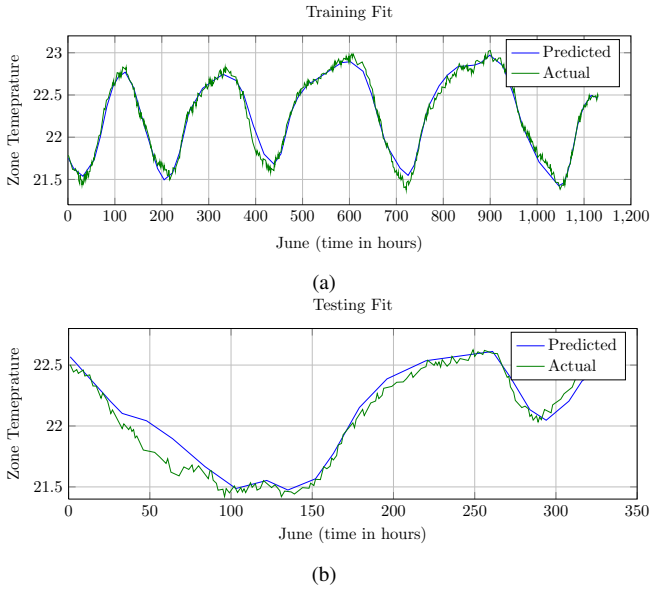


Fig. 16. The fit between the predicted and actual values of the zone temperature in suite 210 of building 101 for the training period is shown in (a). It has a RMSE of 0.062 and a R^2 value of 0.983; (b) shows the fit between the predicted and actual zone temperature values for the testing period with a RMSE of 0.091 and a R^2 value of 0.948

spaces, as well as a lobby/atrium located in the center of the building.

For the case study, we only focus on suite 210, a large office space on the second floor of the north-wing of the building as shown in Figure 14(b). This zone has a single external wall on the east side with 8 windows, a large interior wall on the west side which is adjacent to the porch area on the north-wing and two more adjacent walls on the north and the south side. In July 2013, functional tests were ran from 18:07, 19-07-2013 to 22:29, 20-07-2013, on the air handling unit serving suite 210 as a part of an ongoing Hub project. During a functional test, the supply air temperature is changed rapidly so that there is enough thermal excitation in the zone to generate a rich data-set for learning its dynamical model.

We created the lumped parameter RC-network model for suite 210 using the principles described in Section II. The model has 9 states, 9 inputs and 1 output. There were a total of 22 RC parameters in the model structure for this zone. The temperature inputs to the model were the ambient temperature T_a(°C), boundary condition for the floor T_f(°C) given by the temperature of the zone on the first floor underneath suite 210, boundary condition for the ceiling T_c(°C) given by the temperature of the zone on the third floor above suite 210 and temperature of the adjacent porch area T_p(°C). The external solar irradiation Q_{sole} incident on the east wall is logged by a pyranometer. For the internal heat gain calculation, we consider 3 different heat sources: occupants, lighting and appliances. The number of people in the zone at different times during the functional test period was estimated using data from the people counter. We assume, using ISO standard 7730, that in a typical office environment the occupants are seated, involved in light activity and emit 75 (W) of total heat gain, 30% of

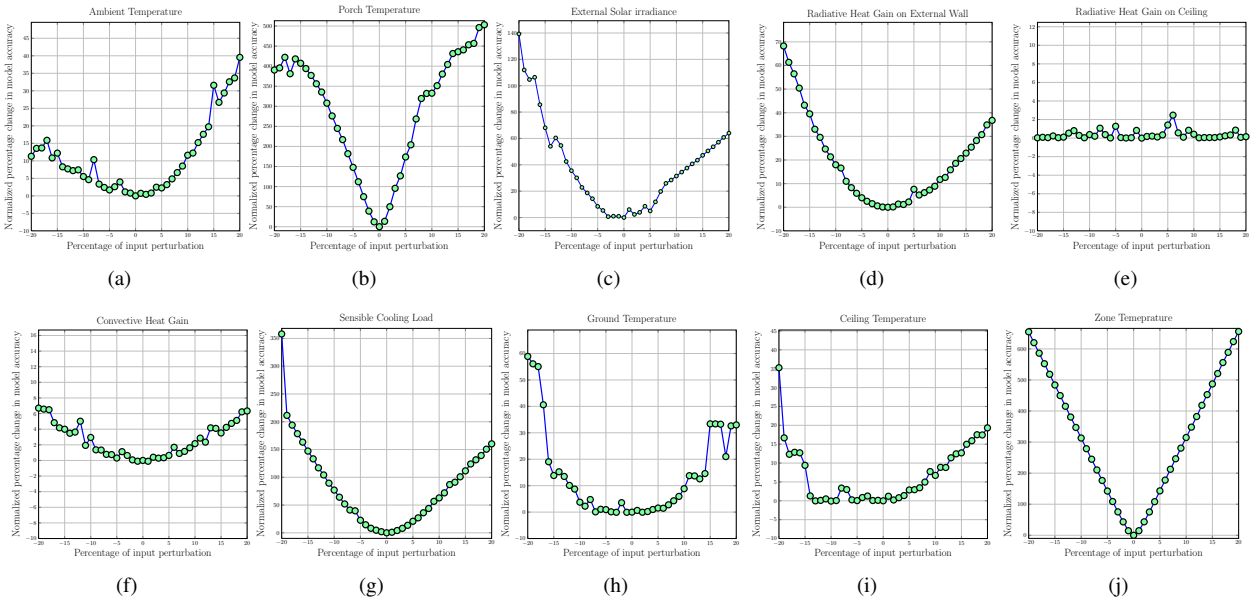


Fig. 17. Input uncertainty analysis results for Building 101 inverse model. The x axis shows the magnitude of the perturbation in percent change from the unperturbed data while the y axis is the percent change in the model accuracy compared to the RMSE for the model trained on unperturbed data. The following inputs are shown: (a) ambient temperature (°C); (b) porch temperature (°C); (c) incident solar irradiation on the external walls (W); (d) and (e) radiative internal heat gain on external wall and ceiling (W); (f) convective internal heat gain (W); (g) sensible cooling rate (W); **??** solar irradiation transmitted through the windows (W); (h) floor temperature (°C); (i) ceiling temperature (°C), and (j) zone temperature (°C)

which is convective and 70% is radiative gain. Using the power rating of the lighting fixtures and their efficiency, one can calculate the heat gain due to lighting. In this zone, lights contribute about 13 (W/m^2) with a $40\% - 60\%$ split between the convective and the radiative part. A constant heat gain due to the electrical appliances and computers is also assumed. The total internal convective heat gain Q_{conv} was obtained by adding the convective gain contributions from the three different heat gain sources. The total internal radiative heat gain was obtained in a similar way. The total internal radiative gain is further split into the radiative gain on the external wall Q_{qgrade} and the radiative gain on the ceiling Q_{qgradc} and applied as two separate inputs. The sensible cooling rate Q_{sen} was calculated using the temperature and mass flow rate measurements for the supply and the return air.

The sampling rate of the data was 1 minute. The total available data was split into a training set (80% of the data) and a test set (remaining 20% data). All the inputs for training the inverse model are shown in Figure 15. The output of the inverse model is the zone temperature T_z. After completion of the training process, the zone temperature predicted by the model is compared with the actual value of the zone temperature for both the training and the test period. The results of the inverse model training are shown in Figure 16. The RMSE for the training data-set was 0.062 with R^2 equal to 0.983 (Figure 16(a)) while the RMSE and R^2 values for the test set were 0.091 and 0.948 respectively (Figure 16(b)).

After successfully training the inverse model, we conduct an input uncertainty analysis on the input-output training data-set as described in Section III-B. The model trained

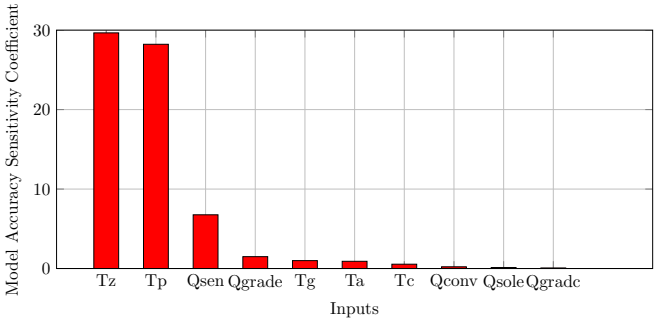


Fig. 18. Model accuracy sensitivity coefficients for suite 210 in building 101

on unperturbed data serves as the baseline model for the uncertainty analysis. Similar to the case of the single-zone TRNSYS model, we created artificial data-set form the training data by perturbing each input data stream within $[-20\%, 20\%]$ of the unperturbed values in increments of 1%. For this case study, we also wanted to characterize the influence of uncertainty in the output of the model, the zone temperature, on the accuracy of the model. Therefore, in addition to the 9 aforementioned model inputs, perturbations were also introduced in the output training data-set i.e. in T_z. With 40 additional data-sets each, there were a total of 400 artificial data-sets. Each of these data-sets were used again for model training and the resulting model was evaluated for its accuracy in terms of the RMSE on the test-set.

The results of the input uncertainty analysis for suite 210 in building 101 are shown in Figure 17. Yet, again we see the parabolic trend obtained as a result of 'artificial' uncertainty in the training data for each of the training data-sets. The sensitivity coefficients for the different training

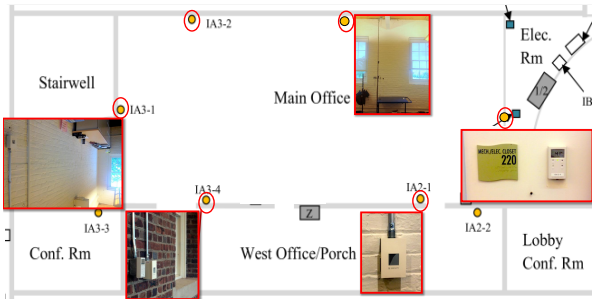


Fig. 19. Temperature sensor locations for suite 210. The thermostat is located on the right wall. The location of 4 IAQ temperature data loggers and the portable temperature sensor cart is also shown.

inputs were calculated. Figure 18 shows the comparison of the model accuracy sensitivity coefficients for the inverse model for suite 210. It is seen that the zone temperature has the largest model accuracy sensitivity coefficient suggesting that the accuracy of the model is quite sensitive to the zone temperature measurement. This suggests that getting the correct measurement for the zone temperature will result in a better model.

VI. Sensor Placement and Data Quality

So far, we have shown the adverse effects of having uncertainty in the training data on the accuracy of the building inverse model which in turn influences the performance of a model predictive controller. In this section, we show how the location of the sensor effects the quality of measured data. Specifically, we compared the thermostat measurement of suite 210 in building 101 with the mean of several temperature measurements made in the same zone but at different locations. A single point temperature measurement of a zone is based on the assumption that the air inside the zone is well mixed. The aim of our experiment was to analyze data from suite 210 and determine if there is any location bias in the thermostat reading. The true value of the temperature of a zone (air volume) is extremely hard to determine. A better approximation of the true zone temperature is mean of temperature measurements taken from sensors which are uniformly located in the zone. Suite 210 at building 101 contains several sensors which log air temperature at different locations int he zone. The layout of the zone and the location of the temperature sensors is shown in Figure 19. There are a total of six different locations in the zone where air temperature is logged. The zone thermostat is placed on the south wall. There are 4 indoor air quality (IAQ) sensors which also measure zone temperature placed on the west, north and the east wall. An additional source of temperature measurement is a portable cart which measures temperatures at different height levels. The location of the cart was not changed for almost an entire year therefore its data can be treated as data measured form the same location. Since the different temperature sensors are located around the zone in a uniform manner, the mean of all the temperature measurements (including the thermostat) is a better representation of the zone temperature.

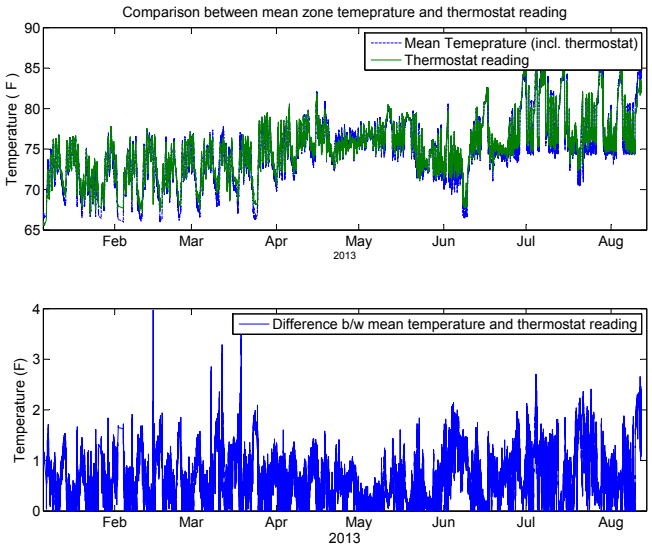


Fig. 20. The comparison of thermostat reading and the mean temperature reading for suite 210 is shown here. The bottom figure is the plot of residuals between the two data-sets.

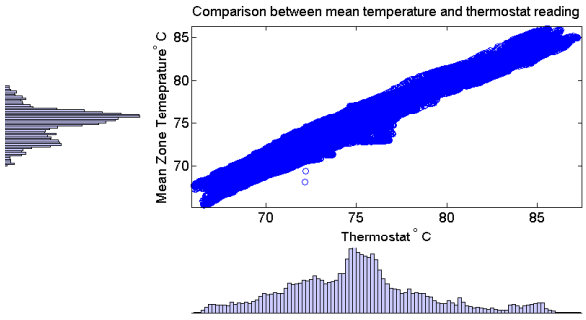


Fig. 21. Scatter plot between the mean temperature measurement (y-axis) and the thermostat reading (x-axis)

The mean temperature value is compared with the thermostat measurement in Figure 20. The values of the residuals have also been plotted. It can be seen that the difference between the thermostat reading and the mean temperature reading can be upto 4 degrees under certain conditions. The trend in the data suggests that the reading of the thermostat may be biased due to its location. The mean deviation in the temperature value is 0.71° which is about 1% bias in the data. A better way to compare the two data-sets is through a scatter plot between the mean temperature and the thermostat reading. Figure 21 shows such a comparison along with a histogram plot for each axis. Two main inferences can be drawn form this plot. First, the width (or the spread) of the data reveals how much the thermostat reading deviates from the mean temperature. A lower spread indicates that the two measurements are in agreement and that the well mixed assumption holds well for the zone. Second, the histogram of the data-sets reveals that the thermostat data has a much larger variance that the mean temperature measurement.

From the results of section V, we have seen that the bias in the zone temperature measurement has the maximum influence on the accuracy of the inverse model for suite 210. This suggests that for this zone, it would be better to deploy additional low-cost wireless sensors just during the

model training phase and get a better estimate of the zone temperature for training the inverse model. Also, the mean value obtained by adding more sensors could be used to re-calibrate or correct the thermostat reading for location bias, resulting in data which can yield an inverse model which can better represent the dynamics of the zone.

VII. Related work

A brief summary of related work in different areas is presented next.

A. MPC related

The treatment and analysis of the implementation of model based control schemes like MPC and optimal control for buildings has been very thorough. Several papers [7], [8], [9] describe the implementation of model predictive control for energy efficient operation of buildings, supported by strong case studies. In [10] authors consider uncertainty in the prediction of disturbances and propose a stochastic version of MPC. In [11], a reduced order model has been used for model based predictive control. [12] advocates the use simpler models for buildings based on the physical description of the bundling. The authors highlight the building modeling process as a crucial part for building predictive control.

B. Sensitivity analysis related

Parametric sensitivity analysis of a model reveals the important parameters of the model which effect the model output the most. [13], [14] among several others are devoted to sensitivity analysis for building modeling. In [13], important input design parameters are identified and analyzed from points of view of annual building energy consumption, peak design loads and building load profiles. In [14] the authors extend traditional sensitivity analysis and increase the size of analysis by studying the influence of about 1000 parameters.

C. Uncertainty related

It is only recently [15], [16], [17] that researchers have analyzed the uncertainty in modeling for close loop control. In [15], the authors acknowledge that the performance of advanced control algorithms depends on the estimation accuracy of the parameters of the model. They design an MPC algorithm using a control model that is structurally identical to the plant model but has perturbed parameters. The closed loop system is simulated and the impact of the parameter perturbations on the energy cost is evaluated. Although, this methodology bears some similarity with the ModelIQ approach, there are some key differences. Firstly, for a fixed model structure, the model parameters can change either due to the estimation process or due to the quality of data. The cause of the parameter change has not been addressed in their work. So although one can identify which parameters should be estimated well, it is not clear how can one go about in getting a good estimate for that parameter. Secondly, the use of the same model as the control and the plant model is questionable. In reality once can never learn the exact plant model and the control model can only be an approximation of the plant dynamics. Which is why we used the TRNSYS building as the plant model in our MPC simulation to make it more realistic. In [16], the authors discuss the development of a control oriented simplified modeling strategy for MPC in buildings using virtual simulations [17] presents a methodology to automate building model calibration and uncertainty quantification using large scale parallel simulation runs. The method considers global sensitivity analysis using probabilistic data while we consider a fixed bias error.

VIII. discussion and limitations

Firstly, although the ModelIQ approach has been presented for the case of a single zone, it can be easily extended for a multi-zone scenario in which zones interact with each other. One method of dealing with this case is to treat the neighboring zone as a boundary condition (temperature node) for the zone of interest. We saw this in the example of the input uncertainty analysis for suite 210, in the case study in Section V, where the porch area was an adjacent zone and its temperature was a boundary condition for our zone model. Secondly, we assume that there exists a sensor deployment which can provide a data-set for training an inverse model. If a building has been retro-fitted for advanced control (like MPC) then it is likely that such a sensor configuration exists. Lastly, the accuracy of the model also depends on the measurement process itself. For instance, the accuracy of the model will be effected by the sampling rate and total number of data-points. Moreover, often it is necessary that the model is tuned or re-trained as the operating conditions of the building change or due to seasonal weather changes. Problems of optimal experimentation design for building inverse models, minimum frequency of model re-tuning and minimum duration of training period are of interest to us and will be investigated as part of future work.

IX. conclusion

In this paper we have introduced ModelIQ, a methodology and a tool-chain for analysis of uncertainty propagation for building inverse modeling and controls. ModelIQ enables the modeling framework to incorporate uncertainty to a level that enables end-users understanding of the limits of their models and controls. Through analysis with a high fidelity virtual building modeled in TRNSYS and then through a case study using real data measurements from an office building, we have shown:

(a) Uncertainty in the form of bias, if present in the training input can adversely effect the accuracy of the building inverse model estimated from that data. This effect can be quantified through an input uncertainty analysis and the extent of the influence of uncertainty in each training data stream on the accuracy of the model can be measured.

(b) We evaluate the relationship between model accuracy and performance of a MPC controller. This was done for a building modeled in TRNSYS. We believe our empirical treatment of this analytically hard problem is both new and realistic compared to related work. We

observe that an accurate building inverse model can result in a MPC cost reduction of more than 13% while a bad model will barely reduce the cost (3%).

(c) We run the ModelIQ tool-chain using data obtained form a real building. Also using real sensor data, we show that the density and placement of sensors (temperature sensors in this case) can be responsible for introducing a location based bias in the measured data from the true value. For the case study, we saw that it can influence the model accuracy of the zone in excess of 20%.

We are continuing our efforts to develop ModelIQ into a an open source toolbox to automate the input uncertainty analysis for building inverse models. Results from this study has also motivated us to address the problem of optimal sensor placement and density for learning building models. This paper is a first step towards having an automated tool to determine the minimum number of sensors, with their appropriate placement in the building, required to capture an adequate building model for model-based control strategies.

REFERENCES

[1] J. E. Braun and N. Chaturvedi, "An inverse gray-box model for transient building load prediction," *HVAC and R Research*, vol. 8, no. 1, pp. 73–99, 2002.

[2] T. L. McKinley and A. G. Alleyne, "Identification of building model parameters and loads using on-site data logs," *governing*, vol. 10, p. 3, 2008.

[3] T. Dewson, B. Day, and A. Irving, "Least squares parameter estimation of a reduced order thermal model of an experimental building," *Building and Environment*, vol. 28, no. 2, pp. 127 – 137, 1993, ¡ce:title¿Special Issue Thermal Experiments in Simplified Buildings¡/ce:title¿.

[4] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on optimization*, vol. 6, no. 2, pp. 418–445, 1996.

[5] J. J. Moré, "The levenberg-marquardt algorithm: implementation and theory," pp. 105–116, 1978.

[6] U. D. of Energy. (2013, Oct.) Eeb hub: Energy efficient buildings hub. [Online]. Available: http://www.eebhub.org/

[7] D. Gyalistras, A. Fischlin, M. Morari, C. Jones, F. Oldewurtel, A. Parisio, F. Ullmann, C. Sagerschnig, and A. Gruner, "Use of weather and occupancy forecasts for optimal building climate control," Technical report, ETH Zürich, Tech. Rep., 2010.

[8] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, and P. Haves, "Model predictive control for the operation of building cooling systems," *Control Systems Technology, IEEE Transactions on*, vol. 20, no. 3, pp. 796–803, 2012.

[9] F. Oldewurtel, A. Parisio, C. N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari, "Use of model predictive control and weather forecasts for energy efficient building climate control," *Energy and Buildings*, vol. 45, pp. 15–27, 2012.

[10] F. Oldewurtel, A. Parisio, C. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth, "Energy efficient building climate control using stochastic model predictive control and weather predictions," in *American Control Conference (ACC), 2010*, 2010, pp. 5100–5105.

[11] D. Kim and J. E. Braun, "Reduced-order building modeling for application to model-based predictive control 2," *gen*, vol. 1, p. 2, 2012.

[12] S. Prívara, J. Cigler, Z. Váňa, F. Oldewurtel, C. Sagerschnig, and E. Žáčeková, "Building modeling as a crucial part for building predictive control," *Energy and Buildings*, 2012.

[13] J. C. Lam and S. Hui, "Sensitivity analysis of energy performance of office buildings," *Building and Environment*, vol. 31, no. 1, pp. 27–39, 1996.

[14] B. Eisenhower, Z. O'Neill, V. A. Fonoberov, and I. Mezić, "Uncertainty and sensitivity decomposition of building energy models," *Journal of Building Performance Simulation*, vol. 5, no. 3, pp. 171–184, 2012.

[15] S. Bengea, V. Adetola, K. Kang, M. J. Liba, D. Vrabie, R. Bitmead, and S. Narayanan, "Parameter estimation of a building system model and impact of estimation error on closed-loop performance," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, 2011, pp. 5137–5143.

[16] J. A. Candanedo, V. R. Dehkordi, and P. Lopez, "A control-oriented simplified building modelling strategy," *13th Conference of international Building Performance Simulation Association*.

[17] S. N. Slaven Peles, Sunil Ahuja, "Uncertainty quantification in energy efficient building performance simualtions," in *International High Performance Buildings Conference*, 2012.