Optimizing Zone Temperature Setpoint Excitation to Minimize Training Data for Data-Driven Dynamic Building Models*

Jie Cai[1], Donghun Kim[1], James E. Braun[1] and Jianghai Hu[2]

Abstract— **Poorly designed excitation signals could lead to inaccurate or, even worse, highly correlated parameter estimates in a data-driven model, so it is critical to have an informative training data set in order to obtain an accurate model in a cost effective manner. This paper investigates a sequential optimal design of experiments (DOE) approach to generate an optimal training data set for varying zone temperature setpoints that maximizes the accuracy of parameter estimates for an intended building model structure. This method was applied to a whole building case study in which a simplified thermal network modeling approach was adopted and the thermal parameters were estimated. The obtained optimal trajectory was always under a bang-bang type which enables an exhaustive search scheme to start the sequential design procedure. The designed excitation signals led to significant improvements in model accuracy compared to night-setup/back control strategies that are typically used to vary setpoints.**

I. INTRODUCTION

To model existing buildings for control purposes, data-driven modeling approaches are advantageous compared to physically-based modeling. When trained using on-site measurements, a data-driven model is better able to capture actual system behavior. Most often, a data-driven model adopts a specific simplified model structure and requires relatively little effort for model setup compared to a physical model. In addition, the computational requirements for a data-driven model are significantly less than those for a detailed model, which enables it to be applied for real-time control or fault diagnosis.

Lumped resistance-capacitance models are commonly used for data-driven dynamic building models and there is extensive research focusing on application of this type of modeling approach (e.g., [1]-[5]). However, most of the previous model training approaches have used passive identification where the training data is collected under normal operating conditions that can lead to inaccurate model parameter estimates.

Optimal experimental design is an effective approach to generate the optimal excitation inputs for identification of a dynamic system. Good references can be found in the monograph [6] and survey papers [7][8]. The reference in [9] provides a review of research and industrial applications of the optimal experimental design technique in different fields. However, the application of experimental design in the building dynamic model identification is still lacking but the importance of obtaining an informative training data set is being recognized in the building modeling community, because building dynamics are typically very slow and inappropriately designed experiments could require a long duration of data collection to obtain a meaningful model ([5]).

In reference [4], a forced-response experiment was designed that led to improved parameter estimates compared to the use of normal operation data. However, the experiment was based on a trial and error method. This is not a scalable approach since it would require expensive experiments to achieve good parameter estimates. In reference [10], a conceptual active identification scheme was proposed for multi-zone buildings. They studied the effect of different excitation signals on model training results for different model structures. Although they provided some useful criteria for experimental design, the work did not tackle the design problem explicitly. In reference [11], the optimal sensor location and sampling time were investigated under an optimal experimental design framework for a slab heat transfer problem. But no input design was involved and a whole building heat transfer problem is more complicated than that for a single slab.

This paper presents a method to excite a building in a systematic manner so that the resulting training data leads to efficient estimates of the key parameters in an intended model structure. The method relies on a sequential optimal experimental design procedure which is formulated as a model predictive control (MPC) problem. The MPC problem maximizes a functional scalar of the Fisher Information Matrix (FIM) under constraints from the system dynamics, space comfort and actuation capacity. Before carrying out the design procedure, a sensitivity-based approach is used to identify the non-influential parameters to avoid wasting energy to excite those structually non-identifiable parameters. By fixing the non-influential parameters to some nominal values, the estimation parameter set can be reduced leading to less computational requirements. Lastly, this sequential experimental design procedure was applied to a whole building case study, where optimal zone air temperature setpoints were determined to maximize the Fisher information matrix.

II. DESIGN CRITERIA

For a predefined model structure, the estimate of a general single-output dynamic system can be assumed to take the

*Research supported by the National Science Foundation under Grant No. 1329875 and US Department of Energy through the Consortium of Building Energy Innovation

[1]J. Cai, D. Kim and J. E. Braun are with School of Mechanical Engineering, West Lafayette, IN 47906 USA (cai40@purdue.edu).

[2]J. Hu are with School of Electrical and Computer Engineering, West Lafayette, IN 47906 USA.

form:

$$y(t,\boldsymbol{\theta}) = \hat{y}(t,\mathbf{x}_0,\mathbf{u}_t,\boldsymbol{\theta}) + \varepsilon(t,\boldsymbol{\theta}) \tag{1}$$

where $\boldsymbol{\theta}$ is the set of parameter values with dimension n_p, $\mathbf{x}_0$ is the initial state vector with dimension n, $\mathbf{u}_{t-} = [u(1),\ldots,u(t)]^T$, ε is the model residual, y and $\hat{y}$ are the measured and predicted outputs, respectively. If the actual system can be recovered under the proposed model structure with $\boldsymbol{\theta}_0$ being the "true" parameters, then a common assumption that can be made is that $\varepsilon(t,\boldsymbol{\theta}_0) = \varepsilon(t)$ is a white Gaussian noise.

Assuming the estimator is unbiased, the Cramer-Rao inequality says the Fisher Information Matrix (FIM) defined below

$$\mathbf{M} = E\left\{ \left[\frac{d}{d\boldsymbol{\theta}} \log f_y(\boldsymbol{\theta},\mathbf{y}_N)\right] \left[\frac{d}{d\boldsymbol{\theta}} \log f_y(\boldsymbol{\theta},\mathbf{y}_N)\right]^T_{|\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right\}, \tag{2}$$

where f_y is the probability distribution function of $\mathbf{y}_N = [y(1),\ldots,y(N)]^T$, serves as the inverse of the lower bound for the estimate covariance matrix [12]. N is the number of sample points. For efficient estimators, e.g., maximum likelihood estimator or least-square estimator with white Gaussian noise, the inverse of FIM is close to, and thus could be used as an approximation of, the estimated covariance matrix, which is an indicator of the accuracy of parameter estimates.

With the white Gaussian noise assumption, a least square estimator is equivalent to a maximum likelihood estimator and it is straightforward to show that

$$\begin{aligned} \tfrac{d}{d\boldsymbol{\theta}} \log f_y(\boldsymbol{\theta},\mathbf{y}_N) &= \tfrac{d}{d\boldsymbol{\theta}} \log\left(\textstyle\prod_{t=1}^N f_e(\varepsilon(t,\boldsymbol{\theta}))\right) \\ &= -\textstyle\sum_{t=1}^N \frac{f_e'(\varepsilon(t,\boldsymbol{\theta}))}{f_e(\varepsilon(t,\boldsymbol{\theta}))} \cdot \psi(t,\boldsymbol{\theta}) \end{aligned}$$

where

$$\psi(t,\boldsymbol{\theta}) = \frac{d}{d\boldsymbol{\theta}}\hat{y}(t,\boldsymbol{\theta}) = -\frac{d}{d\boldsymbol{\theta}}\varepsilon(t,\boldsymbol{\theta}), \tag{3}$$

f_e is the probability distribution function of ε and f_e' is the derivative of f_e with respect to the residual ε. For an actual system f_e is fixed so the FIM can be expressed as

$$\mathbf{M} = \tau \sum_{t=1}^{N} E\left(\psi(t,\boldsymbol{\theta})\psi^T(t,\boldsymbol{\theta})\right)_{|\boldsymbol{\theta}=\boldsymbol{\theta}_0} \tag{4}$$

where the constant τ is a functional of f_e. The optimal experimental design involves trying to optimize some functional of the FIM. The scalar τ in front of the summation in Equation (4) is assumed be to a constant and have no impact on the optimal solution in the design process. For ease of notation, the FIM is used to denote only the summation part for the rest of this paper. Hence, an estimate of the FIM could be written as

$$\hat{\mathbf{M}}(\hat{\boldsymbol{\theta}}) = \sum_{t=1}^{N} \psi(t,\hat{\boldsymbol{\theta}})\psi^T(t,\hat{\boldsymbol{\theta}}) = \Psi(\hat{\boldsymbol{\theta}})\Psi^T(\hat{\boldsymbol{\theta}}) \tag{5}$$

where $\Psi(\hat{\boldsymbol{\theta}}) = \left[\psi(1,\hat{\boldsymbol{\theta}}),\ldots,\psi(N,\hat{\boldsymbol{\theta}})\right] \in \mathbb{R}^{n_p \times N}$ is called the sensitivity matrix. For systems with different scales it is critical to normalize the FIM to avoid an ill-conditioned matrix caused by different parameter scales. The normalized FIM is given below,

$$\hat{\mathbf{M}}_{norm}(\hat{\boldsymbol{\theta}}) = Diag(\hat{\boldsymbol{\theta}})\left(\Psi(\hat{\boldsymbol{\theta}})\Psi^T(\hat{\boldsymbol{\theta}})\right)Diag(\hat{\boldsymbol{\theta}}) \tag{6}$$

and should be used to correctly rank parameter significance and design the experiments. Since the rest of the paper only deals with estimated parameters and FIM, the 'hat' symbol will be omitted for ease of notation.

III. Identifying non-influential parameters

Most gray-box models are over-parameterized so that some non-influential parameters are present or some of the parameters are correlated in the intended model structure. The non-influential parameters may not be neglected due to their physical presence in an actual process. However, if they were estimated together with other parameters, the estimation would be highly correlated and the corresponding confidence region could explode. In the experimental design procedure, a waste of excitation energy could result from an optimal solution that tries to excite those non-influential parameters. Thus, it is a critical step to identify the non-influential parameters before performing experimental design.

This study proposes a principal-component-analysis (PCA)-based approach to order parameter significance from a pre-estimation data set. The pre-estimation data set could be collected under a conventional control logic for a short period of time prior to the design procedure. This data set is used to pre-estimate the parameters which are used as initial parameter values for the experimental design process. In addition, the pre-estimation data set is used to order parameter significance with the following steps ([13]):

1) Calculate the FIM based on the pre-estimation data set using Equation (6);
2) Find the minimum eigenvalue and the corresponding eigenvector $\mathbf{v}$ of the FIM;
3) Find the maximum entry and the corresponding position index k of the eigenvector found in the previous step, i.e., $k = \underset{i}{\text{argmax}}|v_i|$ where $\mathbf{v} = [v_1,\ldots,v_{n_p}]$;
4) The kth parameter is the least influential parameter; remove the kth row and kth column of the FIM;
5) Check the termination condition and, if not satisfied, repeat steps 2 to 4 for the reduced FIM.

Let $\phi_i \in \mathbb{R}^{1\times N}$ denote the ith row of the sensitivity matrix $\Psi(\boldsymbol{\theta})$ defined in Equation (5). So ϕ_i carries the sensitivity information for θ_i. Then the element in the ith row and jth column of the FIM is $\mathbf{M}(i,j) = \phi_i\phi_j^T$. Note that the 4-th step above removes all the information with respect to the kth parameter and the remaining FIM is the same as the FIM with estimation parameter set excluding the kth parameter.

With the aforementioned steps, a list of indices could be obtained corresponding to the non-influential parameters. The termination condition used in the last step concerns whether the condition number (*cond*) of the current FIM is above some threshold.

Note that the presented approach only concerns the data-dependent identifiability of the parameters, which poses a

potential risk that some influential parameters might be mistakenly identified as non-influential if the pre-estimation data set does not provide a good excitation in some directions of the parameter space. In this regard, a data-independent approach seems to be more suitable. A structural identifiability index was introduced for the estimation parameters in [14] based on sensitivity information of the Markov parameters to the estimation parameters. This approach can detect the parameter identifiability regardless of the training data set. However, the data-independent approach assumes uncorrelated inputs with perfect excitation which might not be reasonable for a building dynamic model. The building response strongly depends on the weather related disturbances which are highly correlated with each other. For example, the solar radiation absorbed on the external wall and the transmitted solar radiation, as will be shown in the case study section, are two different inputs but are highly correlated. Thus, the solar absorptance of external walls and the window transmittance are intrinsically correlated which cannot be detected by the data-independent approach. This is why the data-dependent method is adopted in the present study.

IV. Sequential experimental design

A. Building model formulation

Building envelope models have a discrete-time state-space representation given by

$$\begin{aligned} \mathbf{x}[k+1] &= \mathbf{A}(\boldsymbol{\theta})\mathbf{x}[k]+\mathbf{B}_w(\boldsymbol{\theta})\mathbf{w}[k]+\mathbf{B}_u(\boldsymbol{\theta})u[k] \\ y[k] &= \mathbf{C}(\boldsymbol{\theta})\mathbf{x}[k]+\mathbf{D}_w(\boldsymbol{\theta})\mathbf{w}[k]+\mathbf{D}_u(\boldsymbol{\theta})u[k], \end{aligned} \tag{7}$$

where $\mathbf{w}$ represents disturbances such as weather conditions and internal heat gains, u is the controllable input which is the desired zone air temperature and the output y is the cooling/heating load. Let n_w and n_u denote the number of disturbance inputs and controllable inputs, respectively ($n_u = 1$ for the present study). Matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B}_w \in \mathbb{R}^{n\times n_w}$, $\mathbf{B}_u \in \mathbb{R}^{n\times n_u}$, $\mathbf{C} \in \mathbb{R}^{1\times n}$, $\mathbf{D}_w \in \mathbb{R}^{1\times n_w}$ and $\mathbf{D}_u \in \mathbb{R}^{1\times n_u}$ are dependent on the parameter set $\boldsymbol{\theta}$ but this dependence will be omitted in the rest of the paper for ease of notation. The parameter values that are being pursued in the training process are assumed to be time-invariant so the target model is linear time-invariant (LTI). The relationships between the system matrices and model parameters can be found in [1][2].

B. Sequential design formulation

The goal of the optimal experimental design is to find the trajectory of the controllable input that maximizes information in the resulting training data given predicted disturbances. Since weather prediction is increasingly uncertain with increasing look-ahead horizons and parameter estimates could vary as more training data is collected, experimental design should be carried out sequentially with the most recent weather predictions and parameter updates. Define $\mathbf{u}_{t+} = \left[u[t],...,u[t+Np-1]\right]^T$ where Np is the prediction horizon. Similarly define $\mathbf{x}_{t++} = \left[\mathbf{x}[t+1],...,\mathbf{x}[t+Np]\right]^T$ and $\mathbf{y}_{t++} = \left[y[t+1],...,y[t+Np]\right]^T$. For each decision step, a model predictive control (MPC) problem is solved with information richness being the objective function:

$$\begin{aligned} &\max_{[\mathbf{u}_{t+},\mathbf{x}_{t++},\mathbf{y}_{t++}]} \quad \mathscr{E}\left\{\mathbf{M}(\mathbf{u}_{t+},\mathbf{W}_{t+},\boldsymbol{\theta}_t)+\mathbf{M}_{-}(\boldsymbol{\theta}_t)\right\} \\ &\text{s.t.} \\ &\left.\begin{aligned} \mathbf{x}[k+1] &= \mathbf{A}\mathbf{x}[k]+\mathbf{B}_w\mathbf{w}[k]+\mathbf{B}_u u[k] \\ LB_u[k] &\le u[k] \le UB_u[k] \end{aligned}\right\} \text{ for } k=t, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad ...,t+Np-1 \\ &\left.\begin{aligned} y[k] &= \mathbf{C}\mathbf{x}[k]+\mathbf{D}_w\mathbf{w}[k]+\mathbf{D}_u u[k] \\ LB_y[k] &\le y[k] \le UB_y[k] \end{aligned}\right\} \text{ for } k=t+1, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad ...,t+Np \end{aligned} \tag{8}$$

where $\mathbf{W}_{t+}$ contains all the predicted disturbances from step t to step $t+Np-1$ and $\boldsymbol{\theta}_t$ is the most recent parameter estimate. $\mathbf{M}_{-}$ is the information matrix contribution from the historical data prior to the current time step. Note that $\mathbf{M}_{-}$ is a function of historical data as well as the current parameter estimate and thus, it has to be updated at each decision step. But in the optimization process, it is a constant matrix. $\mathscr{E}$ is a scalar functional of the information matrix, which could be the trace (T-optimality), determinant (D-optimality) or condition number (modified E-optimality). The information matrix determinant (D-optimality) characterizes the hyper-rectangular volume that contains the elliptical confidence region ball. The trace (T-optimality) carries information in regard to the total length of the intersection segments of the confidence region with different axes and the reciprocal of the condition number (modified E-optimality) provides a measure of how skinny the confidence ball is. The case study used the determinant as the cost function since it is a direct indicator of the size of confidence region, while other criteria were also monitored.

The incremental information matrix $\mathbf{M}$ in the cost function shown in Equation (8) is an implicit function of the design variables. It would be better to transform it into an explicit form for the convenience of analysis. Define $(\cdot)_{*i} = \partial(\cdot)/\partial\theta_i$, e.g., $\mathbf{x}[k]_{*i} = \partial\mathbf{x}[k]/\partial\theta_i$, for $i = 1,...,n_p$ where θ_i is the ith parameter in $\boldsymbol{\theta}$. Then

$$\begin{aligned} \mathbf{x}_{*i}[k+1] &= \tfrac{\partial}{\partial\theta_i}\left(\mathbf{A}\mathbf{x}[k]+\mathbf{B}_w\mathbf{w}[k]+\mathbf{B}_u u[k]\right) \\ &= \mathbf{A}\mathbf{x}_{*i}[k]+\mathbf{A}_{*i}\mathbf{x}[k]+\mathbf{B}_{w,*i}\mathbf{w}[k] \\ &\quad +\mathbf{B}_{u,*i}u[k] \end{aligned} \tag{9}$$

Define $\overline{\mathbf{x}}[k] = [\mathbf{x}[k];\mathbf{x}_{*1}[k];...;\mathbf{x}_{*np}[k]] \in \mathbb{R}^{(n_p+1)n\times 1}$ where ';' represents concatenation in the column direction. Then

$$\begin{aligned} \overline{\mathbf{x}}[k+1] &= \overline{\mathbf{A}}\overline{\mathbf{x}}[k]+\overline{\mathbf{B}}_w\mathbf{w}[k]+\overline{\mathbf{B}}_u u[k] \\ y[k] &= \tilde{\mathbf{C}}\overline{\mathbf{x}}[k]+\mathbf{D}_w\mathbf{w}[k]+\mathbf{D}_u u[k] \end{aligned} \tag{10}$$

where

$$\begin{aligned} \overline{\mathbf{A}} &= \left[\begin{array}{c|ccc} \mathbf{A} & \mathbf{0} & \dots & \mathbf{0} \\ \hline \mathbf{A}_{*1} & \mathbf{A} & \dots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{A}_{*n_p} & \mathbf{0} & \dots & \mathbf{A} \end{array}\right] \in \mathbb{R}^{((n_p+1)\cdot n)\times((n_p+1)\cdot n)} \\ \overline{\mathbf{B}_w} &= \left[\mathbf{B}_w;\mathbf{B}_{w,*1};\dots;\mathbf{B}_{w,*n_p}\right] \in \mathbb{R}^{((n_p+1)\cdot n)\times n_w} \\ \overline{\mathbf{B}_u} &= \left[\mathbf{B}_u;\mathbf{B}_{u,*1};\dots;\mathbf{B}_{u,*n_p}\right] \in \mathbb{R}^{((n_p+1)\cdot n)\times n_u} \\ \tilde{\mathbf{C}} &= [\mathbf{C},\mathbf{0},\dots,\mathbf{0}] \in \mathbb{R}^{1\times((n_p+1)\cdot n)} \end{aligned}$$

The output sensitivity with respect to θ_i at time k is

$$\begin{aligned} y_{*i}[k] &= \frac{\partial y[k]}{\partial \theta_i} = \frac{\partial}{\partial \theta_i}\left(\mathbf{C}\mathbf{x}[k] + \mathbf{D}_w\mathbf{w}[k] + \mathbf{D}_u u[k]\right) \\ &= \mathbf{C}\mathbf{x}_{*i}[k] + \mathbf{C}_{*i}\mathbf{x}[k] + \mathbf{D}_{w,*i}\mathbf{w}[k] + \mathbf{D}_{u,*i}u[k]. \end{aligned} \tag{11}$$

Then the output sensitivity vector at time step k is

$$\begin{aligned} \boldsymbol{\psi}[k] &= \frac{\partial y[k]}{\partial \boldsymbol{\theta}} = \left[y_{*1}[k], \ldots, y_{*n_p}[k]\right]^T \\ &= \overline{\mathbf{C}}\overline{\mathbf{x}}[k] + \overline{\mathbf{D}}_w\mathbf{w}[k] + \overline{\mathbf{D}}_u u[k] \in \mathbb{R}^{n_p \times 1} \end{aligned} \tag{12}$$

where

$$\begin{aligned} \overline{\mathbf{C}} &= \left[\begin{array}{c|ccc} \mathbf{C}_{*1} & \mathbf{C} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{C}_{*2} & \mathbf{0} & \mathbf{C} & \ldots & \mathbf{0} \\ \vdots & \vdots & & \ddots & \vdots \\ \mathbf{C}_{*n_p} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{C} \end{array}\right] \in \mathbb{R}^{n_p \times (n\cdot(n_p+1))} \\ \overline{\mathbf{D}}_w &= \left[\mathbf{D}_{w,*1}; \ldots; \mathbf{D}_{w,*n_p}\right] \in \mathbb{R}^{n_p \times n_w} \\ \overline{\mathbf{D}}_u &= \left[\mathbf{D}_{u,*1}; \ldots; \mathbf{D}_{u,*n_p}\right] \in \mathbb{R}^{n_p \times n_u} \end{aligned}$$

Then the MPC problem in (8) can be reformulated as

$$\max_{[\mathbf{u}_{t+}, \overline{\mathbf{x}}_{t++}, \boldsymbol{\psi}_{t++}, \mathbf{y}_{t++}]} \mathscr{E}\left\{\sum_{k=t+1}^{t+Np}\left(\boldsymbol{\psi}[k]\boldsymbol{\psi}^T[k]\right) + \mathbf{M}_-(\boldsymbol{\theta}_t)\right\}$$

s.t.

$$\left.\begin{aligned} \overline{\mathbf{x}}[k+1] &= \overline{\mathbf{A}}\overline{\mathbf{x}}[k] + \overline{\mathbf{B}}_w\mathbf{w}[k] + \overline{\mathbf{B}}_u u[k] \\ LB_u[k] &\le u[k] \le UB_u[k] \end{aligned}\right\} \text{ for } k = t, \ldots, t+Np-1$$

$$\left.\begin{aligned} \boldsymbol{\psi}[k] &= \overline{\mathbf{C}}\overline{\mathbf{x}}[k] + \overline{\mathbf{D}}_w\mathbf{w}[k] + \overline{\mathbf{D}}_u u[k] \\ y[k] &= \tilde{\mathbf{C}}\overline{\mathbf{x}}[k] + \mathbf{D}_w\mathbf{w}[k] + \mathbf{D}_u u[k] \\ LB_y[k] &\le y[k] \le UB_y[k] \end{aligned}\right\} \text{ for } k = t+1, \ldots, t+Np \tag{13}$$

where $\overline{\mathbf{x}}_{t++}$ and $\boldsymbol{\psi}_{t++}$ are defined similarly as for $\mathbf{y}_{t++}$. Note that the involved system matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are all dependent on $\boldsymbol{\theta}_t$.

Remark: When the T-optimality criterion is used, the cost function is of a least square form and thus, is convex. So the optimal solution is always bang-bang with the optimal control taking values at either the lower bound or the upper bound of the feasible range. Although convexity cannot be established for the D-optimality or the modified E-optimality criterion, the optimal inputs fall on the bounds most of the time to create maximal sensitivity curves [9]. This bang-bang behavior could also be observed in the optimization results obtained in the case study where the D-optimality criterion was used.

C. Steps for the optimal experimental design

Figure (1) shows the flow chart of the proposed design methodology. In the initialization step, some rough information, e.g, wall areas and materials, is collected to construct an initial guess for parameter values which are primarily used to specify the search region. The information does not need to be accurate but the search region should be assigned properly to cover the true parameter values. For uncertain parameters, a relatively large band can be used to allow the estimation to search in a larger range while for parameters with good confidence, a small region should be assigned to reduce computational burden.

In the pre-estimation step, data is collected for a short period of time under conventional control logic. If a building management system (BMS) is in place, the most recent historical data set can be downloaded and used directly. This pre-estimation step provides an initial parameter estimate that can be used in the design procedure and more importantly, the data set can be used to identify and eliminate the non-influential parameters. If the adopted model structure takes uncorrelated inputs, a data-independent approach elaborated in [14] is preferred for identifying the non-influential parameters.

Sequential optimal experimental design is then started with the pre-estimated parameters and for the reduced parameter set. Weather predictions are obtained from external weather forecast services, e.g., the National Oceanic and Atmospheric Administration (NOAA) website, and an optimal input trajectory can be determined by solving the MPC problem in Equation (13). The first Nd decisions are applied to the building control system where Nd is the decision interval. In this study, zone air temperature setpoints are designed inputs so the optimal setpoint profile for the first Nd steps are sent back to the BMS so that local PI controllers will adjust the heating/cooling actuation to maintain the temperature setpoints. The design process idles until the next decision time comes, and the parameters are re-estimated based on the updated historical data set. Then the experimental design step is repeated with the updated parameters and the iterations move forward until a threshold is reached for the parameter accuracy.

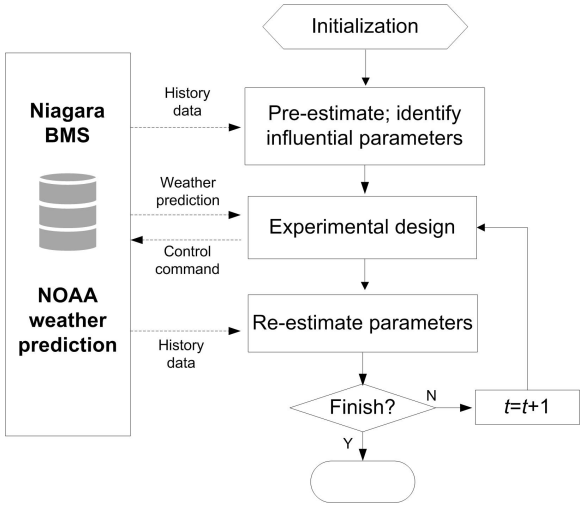

Fig. 1. Flow chart for the experimental design procedure.

V. Case Study: A Whole Building

This section looks at a 20-zone office building located in Philadelphia, PA. It is served by one air handling unit (AHU) and ten VAV boxes. Thus only ten of the zones are controlled which correspond to the main office areas, while the other ten zones (stairwells, hallways, etc.) have floating temperatures.

A. Baseline model

A detailed simulation model was constructed from the building's physical properties to serve as a baseline model,

and the modeling approach was elaborated in [14]. It was assumed that all the zone air temperatures share an identical setpoint which is the sole controllable input to the model. Sensible heating/cooling was calculated as the model output.

B. Data-driven model

The inverse model adopts a simplified thermal network shown in Figure 2 which consists of an external wall, a window and an internal thermal storage element. The model is formulated under the discrete-time state-space form in Equation (7) with the disturbance vector defined as

$$\mathbf{w} = \left[T_{amb}, Q_{irrad,ext}, Q_{irrad,wind}, Q_{gain}\right]^T$$

where T_{amb} is the ambient temperature, $Q_{irrad,ext}$ is the solar irradiation that hits the external wall surface (W), $Q_{irrad,wind}$ is the solar irradiation that hits the windows (W) and Q_{gain} is the electrical heat gain (W). Note that $Q_{irrad,ext}$ and $Q_{irrad,wind}$ were calculated in a pre-process step by considering all the wall and window areas with different orientations. The heat gain effects for the absorbed solar radiation on external wall surfaces and the transmitted solar through windows are determined by

$$\begin{aligned} Q_{sol,ext} &= \theta_{ab} \cdot Q_{irrad,ext} \\ Q_{sol,trans} &= \theta_{tr} \cdot Q_{irrad,wind} \end{aligned}$$

where θ_{ab} and θ_{tr} are to-be-estimate parameters corresponding to the solar absorptance on external walls and window transmittance. θ_{conv} is another parameter that needs to be estimated which represents the ratio of the convective heat gain to the total electrical heat gain:

$$\begin{aligned} Q_{gain,conv} &= \theta_{conv} \cdot Q_{gain} \\ Q_{gain,rad} &= (1 - \theta_{conv}) \cdot Q_{gain} \end{aligned}$$

The estimated parameter set includes the aforementioned heat gain related parameters and the thermal resistances and capacitances shown in Figure (2).

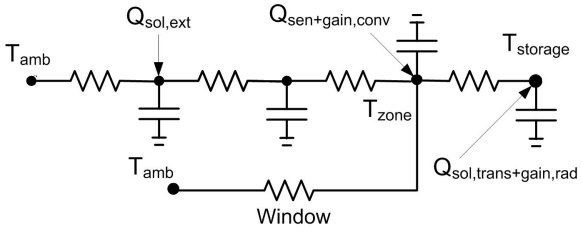

Fig. 2. Thermal network for the case study building.

C. Optimal experimental design settings

The optimal design problem formulated in Equations (13) includes bound constraints for the zone air temperature and cooling/heating rate. The zone air temperature is bounded to a narrow range during occupied periods to avoid discomfort and those bounds are relaxed during unoccupied hours for potentially better excitation (dash lines in Figure (3)). The bounds for the output are imposed due to the capacity of the air conditioning system and these bounds were set to $LB_y = -120$ (kW) and $UB_y = 120$ (kW). However, these bounds were never reached and thus, were always inactive in simulations within this study.

A one hour time step was used in the discrete-time state-space representation in Equation (7) and a 50-hr data set under a conventional night setback control strategy was used in the pre-estimation step. In the parameter significance ranking step, a condition number of 1500 was used for the threshold and two parameters were identified as non-influential (θ_{ab} and the inner resistance of the external wall). The choice of this threshold was conservative in order to reduce the chance of mistakenly identifying a significant parameter as a non-influential one. The T-optimality criterion was used to start the design procedure. Since the optimal solution is bang-bang, an exhaustive search method was implemented to find the optimal trajectory in the first design step. In the subsequent decision steps, D-optimality was used and the optimization initial guess was constructed from the overlapped portion of the optimal solution obtained in the previous decision step. Interior-point method ([16]) was used to solve the MPC problem in Equation (13) for the subsequent design steps. The Levenberg-Marquardt method ([17]) was used to estimate the parameters and the most recent parameter estimates were used as the initial guesses for the current estimation step.

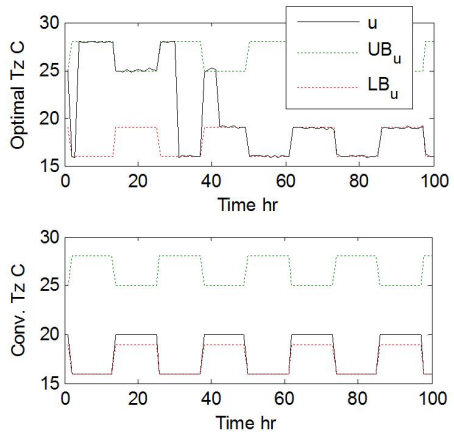

Fig. 3. Optimal and conventional temperature setpoint profiles.

The optimal temperature setpoint profile is shown in Figure (3) along with the conventional setpoints, which are fixed to 20 C during occupied periods and 16 C during unoccupied periods. Note that the D-optimality criterion was used, but the optimal trajectory is still a bang-bang type with several switches between the upper and lower bounds. Performance of the optimal input trajectory is compared to that of the conventional trajectory in Figure (4). As explained before, the different criteria characterize the confidence region in different aspects, thus the three criteria exhibit different behaviors within the designed experiment. The determinant (D-optimality) exhibits the fastest increase as it was considered explicitly in the optimization cost function. The 'DET' plot shows that, with the optimally designed setpoints, the size of the confidence region is reduced by approximately 30 times.

So to achieve a certain accuracy level, the training data size can be dramatically reduced. Taking the demonstrated results as an example, the optimal design only requires 18-hr data to achieve the same parameter accuracy that requires 100-hr training data using the conventional strategy (plus the 50-hr pre-estimation data in both cases). The other two criteria also have significant improvements compared to the night setback strategy.

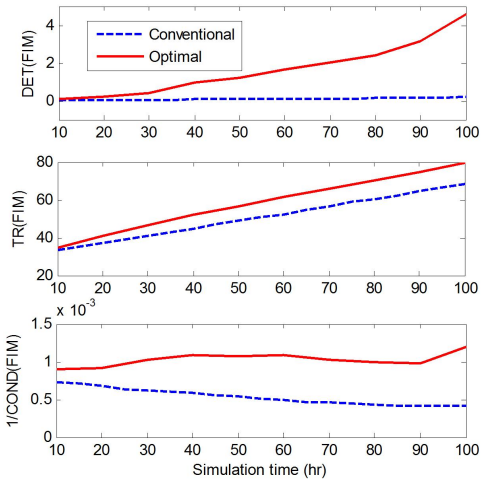

Fig. 4. Performance comparison of optimal and conventional input trajectories.

D. Validation results

Figure (5) compares the sensible load predictions from the models trained by conventional and optimal data sets, respectively, with the baseline sensible load profile. Both models are able to accurately predict sensible load since both training data sets drive the estimated parameters to the optimal parameter values for this case study (both data sets are persistently exciting for the model structure). But with the optimally designed inputs, the parameter estimates converge with significantly improved accuracy, i.e., with faster shrinking confidence region. This is critical in a practical point of view because the optimal parameter estimates are not known a priori and the only way to determine if an experiment is adequate for identifying a satisfactory model is to check the confidence region for the estimated parameters. With optimally designed excitation, the parameter confidence region shrinks much faster and thus, shorter experiments are needed to reach an accuracy requirement.

VI. Conclusion & Future work

This paper develops a methodology to optimally design an experiment for the use of data-driven dynamic modeling of buildings. This technique was applied to a whole building case study and significant improvements were achieved in parameter precision. In another perspective, the required size of the training data could be dramatically reduced to achieve a certain accuracy level, which leads to more cost-effective experiments.

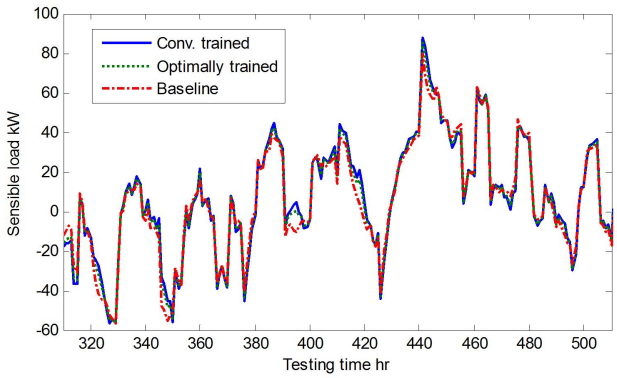

Fig. 5. Performance comparison of models trained with conventional and optimal training data sets.

This study only considered a single-zone case study to demonstrate the efficacy of the proposed approach where only a single controllable input and a single output are involved. However, the overall method is still applicable for multi-zone buildings and most of the analysis results will carry over into multi-zone cases with simple modifications. For example, within a multi-zone case, the sensitivity function ψ defined in Equation (3) would become a matrix with the number of columns equal to the number of zones or output channels (n_u). The design criteria and problem formulations would remain the same except for dimension changes of the corresponding matrices. The step of identifying non-influential parameters has already been applied in a multi-zone data-driven modeling study in [18]. In addition, the application of the experimental design techniques in multi-zone buildings could well address the inter-zonal coupling estimation issues raised by [4] where the neighboring zone temperature setpoints could be optimally perturbed by the optimal design technique to provide a good excitation in estimating the inter-zonal coupling properties.

A preliminary field test has been carried out within an office space in the Center for High Performance Buildings at Purdue University. Parameter precision was improved significantly although more tests will need to be performed to have an overall performance assessment of the proposed methodology.

References

[1] Braun, J.E. and Chaturvedi, N., "An Inverse Gray-box Model for Transient Building Load Prediction", *HVAC&R Research*, Vol. 8, No. 1, pp. 73-99, 2002.

[2] Cai, J. and Braun, J.E., "Efficient and robust training methodology for inverse building modeling", *Fifth National Conference of IBPSA-USA*, 2012.

[3] O'Neil, Z, Narayanan, S. and Brahme, R., "Model-based thermal load estimation in buildings", *Fourth National Conference of IBPSA-USA*, 2010.

[4] Lin, Y., Middelkoop, T. and Barooah, P., "Issues in identification of control-oriented thermal models of zones in multi-zone buildings", *IEEE 51st Annual Conference on Decision and Control*, 2012.

[5] Privara, S., etc., "Modeling and Identification of a Large Multi-Zone Office Building", *IEEE International Conference on Control Applications*, 2011.

[6] Goodwin, G.C. and Payne, R.L., Dynamic system identification: experiment design and data analysis, New York: Academic Press, 1977.
[7] Mehra, R.K. "Optimal input signals for parameter estimation in dynamic systems–Survey and new results." *Automatic Control, IEEE Transactions on*, 1974.
[8] Gevers, M., "Identification for Control: From the Early Achievements to the Revival of Experiment Design." *European journal of control*, 2005.
[9] Franceschini, G. and Macchietto, S., "Model-based design of experiments for parameter precision: State of the art." *Chemical Engineering Science*, 2008.
[10] Agbi, C., Song, Z. and Krogh B., "Parameter identifiability for multi-zone building models", *IEEE 51st Annual Conference on Decision and Control*, 2012.
[11] Emery, A.F., "Using the concept of information to optimally design experiments with uncertain parameters." *Journal of heat transfer*, 2001.
[12] Ljung, L., System Identification: Theory for the User, Prentice Hall, 1999.
[13] Posten, C. and Munack, A., "On-line Application of Parameter Estimation Accuacy to Biotechnical Processes", *American Control Conference*, pp. 2181-2186, 1990.
[14] Doren, V, eta., "Determining identifiable parameterizations for large-scale physical models in reservoir engineering", *Proceedings of the 17th IFAC World Congress*, 2008.
[15] Kim, D. and Braun, J.E., "Reduced-order modeling for application to model-based predictive control", *Fifth National Conference of IBPSA-USA*, 2012.
[16] Waltz, R. A., J. L. Morales, J. Nocedal, and D. Orban. "An interior algorithm for nonlinear optimization that combines line search and trust region steps." *Mathematical Programming*, 2006.
[17] K. Madsen, H.B. Nielsen and O. Tingleff, "Methods for Nonlinear Least Squares Problems", Technical University of Denmark, 2004.
[18] Cai, J. and Braun, J.E., "An inverse hygrothermal model for multi-zone buildings", *Journal of Building Performance Simulation*, 2015.