July 2018

Practice-Oriented System Identification Strategies for MPC of Building Thermal and HVAC Dynamics

Anas Alanqar
Johnson Controls, Inc., United States of America, anas.alanqar@jci.com

Matthew J Ellis
Johnson Controls, Inc., United States of America, matthew.j.ellis@jci.com

Juan E Tapiero Bernal
Johnson Controls, Inc., United States of America, juan.esteban.tapiero.bernal@jci.com

Michael J Wenzel
Johnson Controls, Inc., United States of America, mike.wenzel@jci.com

Practice-Oriented System Identification Strategies for MPC of Building Thermal and HVAC Dynamics

Anas ALANQAR[1]*, Matthew J. ELLIS[1], Juan E. TAPIERO[1], Michael J. WENZEL[1]

[1] Johnson Controls Inc.; Technology and Advanced Development, Building Efficiency, Milwaukee, WI, United States of America
anas.alanqar@jci.com, matthew.j.ellis@jci.com, juan.esteban.tapiero.bernal@jci.com, mike.wenzel@jci.com,

* Corresponding Author

ABSTRACT

In this paper, we present two practice-oriented strategies to overcome difficulties in conducting system identification of building thermal and HVAC dynamics. Specifically, the first part of the paper presents a saturation detection and removal algorithm for buildings HVAC systems. This work focuses on identifying linear time invariant (LTI) models of building thermal and HVAC dynamics, and such models are not able to capture the nonlinear dynamics associated with controller saturation. For simplicity, data collected during periods where the controller was saturated is referred to as saturation data. The purpose of this algorithm is to detect and remove saturation data that are not suitable for system identification, and thereby utilize only the reliable excited data that produce more accurate models of buildings HVAC systems. The second part of this paper presents a systematic procedure that eliminates poor initial guesses when solving the non-convex system identification problem to identify building thermal and HVAC model parameters. The procedure presents guidelines to tests the quality of initial guesses and yields guesses that ultimately lead to models that accurately fit the data and captures the system dynamics. Although these strategies can be implemented to both black-box and grey-box system identification, we will present them in the context of grey-box. The proposed strategies were implemented to on a real building in cooling mode and resulted in very accurate models.

1. INTRODUCTION

The increasingly competitive HVAC market has made it necessary to develop technologies that exploit the economic potential in such systems to reduce energy consumption. Smart HVAC operation through optimization-based control strategies, such as model predictive control (MPC), serves as one method for achieving this goal. MPC has gained significant attention due to their ability to optimally operate HVAC system in order to minimize their energy consumption and/or energy cost while maintaining desired comfort temperatures (Mendoza-Serrano *et al.,* 2015, Arfam *et al.,* 2014, Rawlings *et al.,* 2017, in press). MPC requires the availability of a model of the process dynamics; therefore, several research work have studied the formulation and implementation of such control designs using empirical models (Alanqar *et al.,* 2015). Furthermore, grey-box system identification in order to model building thermal and HVAC dynamics has been attempted (Ellis *et al.,* 2016; Lin *et al.,* 2012). Nonetheless, empirically modeling HVAC systems in a robust and scalable manner is very challenging due to the non-convexity of the system identification optimization problem and the existence of complex actuator saturation limits. Therefore, this work develops grey-box system identification strategies that attack these challenges to enable the development of suitable empirical models of HVAC systems in practice.

With the building thermal and HVAC system we consider, saturation refers to periods whereby the HVAC system is operated at its minimum or maximum HVAC capacity for prolonged periods. The existence of significant saturation in the data collected is a common problem in industry that poses many challenges for identifying the dynamics of HVAC systems since they can affect the quality of the collected data and result in an inaccurate identified model. Classical approaches for dealing with saturation in system identification, such as using nonlinear functions to capture the saturation behavior (Wdanage *et al.,* 2004; Wills *et al.,* 2013), are not implementable due to the complex saturation behavior associated with HVAC systems. More specifically, saturation limits in commercial and residential HVAC systems may be both unknown and time varying. In addition, another challenge faced in identifying HVAC and building thermal dynamics is the existence of many roots in such non-convex system identification problems. Therefore, it is desirable in industry to avoid using initial guesses that lead to local optima which result in inaccurate models.

In the first part of this work, we develop an algorithm that is capable of detecting and removing saturation data from system identification experiment input-output data. This is done to extract the useful data sections that represent the excited HVAC system dynamics which are necessary to identify reliable models of the HVAC dynamics. Once the multiple useful data sections are extracted, system identification is conducted in order fit one model to all data sections. The second part of this work presents a strategy that avoids solving system identification problems all the way to local optimality using initial guesses that lead to local optima which result in poor models. The algorithm attempts to eliminate poor initial guesses and yield initial guesses that ultimately lead to great fits of the model to the data. The parameters of the grey-box models were identified via a two-step parameter estimation approach. In the first step, the model parameters were identified using simulation prediction error method (Ljung *et al.*, 1999). In the second step, the model was augmented with a disturbance model and the estimation gain (i.e., Kalman gain) was identified using the standard 1-step prediction error method (Ljung *et al.*, 1999). The proposed strategies were applied to data collected from real building HVAC systems and have shown to successfully work in practice.

2. PRELIMINARIES

2.1 Notation

The set of integers is denoted by $\mathbb{I}$ and the set of positive integers is denoted by $\mathbb{I}_{\geq 0}$. The notation $\hat{x}(i|k)$ is used to represent the estimated or predicted value of x at time step i given measurements up to time step k where $i \geq k$.

2.2 Dynamic System Model

The class of system models considered in this work is given by the following discrete-time linear time-invariant system model:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{aligned} \tag{1}$$

where $k \in \mathbb{I}_{\geq 0}$ is the time index, $x(k) \in \mathbb{R}^n$ is the state vector, $y(k) \in \mathbb{R}^p$ is the measured output vector, $u(k) \in \mathbb{R}^m$ is the input vector, $w(k)$ is process noise, and $v(k)$ is measurement noise. The matrix A is assumed to be stable; that is the real parts of the eigenvalues of A lie within the unit circle. The pair (A, B) is assumed to be controllable and the pair (A, C) is assume to be observable. Given that the full state is usually not measured in practice, state estimation is performed using a state observer. In addition, a disturbance model is added in order to remove the effects of correlated non-stationary disturbance (Pannocchia *et al.*, 2003). Therefore, closed-loop state estimation is described by the following estimator equations:

$$\begin{aligned} \begin{bmatrix} \hat{x}(k+1|k) \\ \hat{d}(k+1|k) \end{bmatrix} &= \begin{bmatrix} A & B_d \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{x}(k+1|k) \\ \hat{d}(k+1|k) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} K_x \\ K_d \end{bmatrix} (y(k) - \hat{y}(k|k-1)) \\ \hat{y}(k|k-1) &= \begin{bmatrix} C & C_d \end{bmatrix} \begin{bmatrix} \hat{x}(k|k-1) \\ \hat{d}(k|k-1) \end{bmatrix} + Du(k) \end{aligned} \tag{2}$$

where $\hat{x}(j|k)$,$\hat{d}(j|k)$ and $\hat{y}(j|k)$ are the estimated/predicted states, disturbance states and outputs vectors at time step j given the measurement at time step k. The disturbance model is characterized by the choice of the matrices B_d and C_d. The matrix K is the estimator gain. Under certain assumptions (i.e., system is linear, no plant-model mismatch, and the process and measurement noise is white with known covariance matrices), the Kalman filter is the optimal linear filter (Ljung *et al.*, 1999).

2.3 Grey-Box System Identification Background

System identification is the process of constructing mathematical models of dynamic systems (Ljung *et al.*, 1999). Grey box modeling is the class of system identification methods used in this work. It assumes that there is a known model structure M, often derived through simplified first-principals or semi-physical modeling. Subsequently, the model structure is parameterized with a parameter $\theta \in D_{\mathcal{M}} \subset \mathbb{R}^d$ where $D_{\mathcal{M}}$ is the set of admissible model parameter values. Then, parameter estimation techniques are used to estimate the parameters. The resulting possible models are given by the set:

$$M = \{\mathcal{M}(\theta) | \theta \in D_{\mathcal{M}}\} \tag{3}$$

To fit a proper parameter vector for a given system, input-output data points are collected, which are denoted by:

$$Z^N = [y(1), u(1), y(2), u(2), \dots, y(N), u(N)] \tag{4}$$

where $N > 0$ is the number of samples collected. Parameter estimation is the problem of selecting the best value $\hat{\theta}_N$ of the parameter vector with respect to an objective function given the input-output data Z^N (i.e., mapping: $Z^N \rightarrow \hat{\theta}_N \in D_{\mathcal{M}}$). The resulting parameterized state space system to be identified is given by the following:

$$\begin{aligned} \hat{x}(k+1|k) &= A(\theta)\hat{x}(k|k-1) + B(\theta)u(k) + K(\theta)(y(k) - \hat{y}(k|k-1)) \\ \hat{y}(k|k-1) &= C(\theta)\hat{x}(k|k-1) + D(\theta)u(k) \qquad \hat{x}(0|0) = x_0(\theta) \end{aligned} \tag{5}$$

Prediction error methods (PEM) refers to a particular family of parameter estimation methods that are of interest in the context of this work. Prediction error methods fits the parameter vector by minimizing some function of the difference between the predicted output and observed output. In this work, two versions of the prediction error methods are utilized in order to identify the plant model and the estimator gain.

2.3.1 One-step ahead PEM

The first version of the prediction error methods is the 1-step prediction method. At each time step, this method utilizes past input-output data in order to predict the output 1-step ahead. Specifically, let $\hat{y}(k|k-1,\theta)$ be the predicted output at time step k given the past input-output sequence Z^{k-1} and the model $\mathcal{M}(\theta)$ for $\theta \in D_{\mathcal{M}}$. Therefore, the prediction error at each time becomes:

$$\varepsilon(k,\theta) := y(k) - \hat{y}(k|k-1,\theta) \tag{6}$$

The 1-step prediction error method attempts to find the parameter estimates $\hat{\theta}_N$ that minimize a specific measure of this error. The general objective function of such optimization problem is:

$$\hat{\theta}_N(Z^N) = \arg\min_{\theta \in D_{\mathcal{M}}} \frac{1}{N}\sum_{k=1}^{N} \ell(\varepsilon_F(k,\theta)) \tag{7}$$

where $\ell(\cdot)$ is typically taken to be a positive definite function. In the context of the present work, we will use perhaps the most commonly encountered prediction error method, which we define as the *standard PEM*. The standard PEM uses a quadratic cost function that is the square of the prediction errors:

$$\big(y(k) - \hat{y}(k|k-1,\theta)\big) = \|y(k) - \hat{y}(k|k-1,\theta)\|_2^2 \tag{8}$$

where $\hat{y}(k|k-1,\theta)$ denotes the one-step ahead prediction of the output using the model $\mathcal{M}(\theta)$. When the prediction errors are independently and identically distributed random variables from a normal distribution (i.e., the process and measurement noise is Gaussian white noise) and the model being identified is linear, the cost function of (9) is optimal from a statistical point-of-view (Ljung *et al.*, 1999). Written in a compact way, the main objective is to find the optimal θ that minimizes the following cost function

$$\min_{\theta} \frac{1}{N}\sum_{k=1}^{N}\|y(k) - \hat{y}(k|k-1,\theta)\|_2^2 \tag{9}$$

2.3.2 Simulation PEM

The second version of prediction estimation method considered here is the simulation method. In this method, a simulation is conducted using the model $\mathcal{M}(\theta)$ and the input trajectory $[u(1), u(2), \dots, u(N)]$ starting from an initial state $x(0)$. This forward simulation produces all the predicted outputs in terms of θ (i.e. $[\ \hat{y}(1|0,\theta)\ \hat{y}(2|0,\theta) \dots \hat{y}(k|0,\theta) \dots \hat{y}(N|0,\theta)]$). The prediction error at time step k is given by

$$\varepsilon(k,\theta) := y(k) - \hat{y}(k|0,\theta) \tag{10}$$

In this problem, the main objective is to find the optimal θ that minimizes the following cost function

$$\min_{\theta} \frac{1}{N}\sum_{k=1}^{N}\|y(k) - \hat{y}(k|0,\theta)\|_2^2 \tag{11}$$

3. IDENTIFICATION OF BUILDING THERMAL AND HVAC DYNAMICS

In this work, we consider the grey-box system identification of building thermal and HVAC dynamics. The model structures considered capture the crucial dynamics for MPC and control-oriented applications. The thermal dynamics of a space within a building are assumed to be modeled by the following differential equations:

$$C_m \frac{dT_m}{dt} = \frac{1}{R_{mi}}(T_{ia} - T_m) \tag{12}$$

$$C_{ia} \frac{dT_{ia}}{dt} = \frac{1}{R_{oi}}(T_{oi} - T_{ia}) + \frac{1}{R_{mi}}(T_m - T_{ia}) + \dot{Q}_{HVAC} + \dot{Q}_{other} \tag{13}$$

where $T_{ia}, T_{oa}, T_{sp}, T_m$ are the indoor air temperature, outdoor air temperature, indoor set-point temperature, and lumped thermal mass temperature, respectively. The indoor air thermal capacitance, lumped mass thermal capacitance, indoor air-outdoor air thermal resistance, and indoor air-thermal mass thermal resistance are denoted as $C_{ia}\ C_m\ R_{oi}\ R_{mi}$. Sensible heat provided to/removed from the building space by the HVAC system is denoted as

$\dot{Q}_{\mathrm{HVAC}}$. The internal disturbance heat load, $\dot{Q}_{\mathrm{other}}$, models the heat generated by the building occupants, the heat transmitted via solar irradiance, and the heat generated by electrical equipment (e.g., light and computers). We assume that the relationship between the indoor air temperature and the temperature set-point follows a proportional-integral control law with saturation. Let

$$\begin{aligned} \dot{Q}_{\mathrm{HVAC}} &= K_p \varepsilon_{sp} + \frac{K_p}{\tau_{\mathrm{I}\,I}} \int_0^t \varepsilon_{sp}(s)\, ds \\ \varepsilon_{sp} &= T_{sp} - T_{ia} \end{aligned} \tag{14}$$

$\dot{Q}_{\mathrm{HVAC}}$ will have a positive value when heat is supplied to the building space and a negative value when cooling is supplied to the building space. If $\dot{Q}_{\mathrm{HVAC}}$, is within the constraints on $\dot{Q}_{\mathrm{HVAC}}$, then $\dot{Q}_{\mathrm{HVAC}} = \dot{Q}_{\mathrm{HVAC}}$. Otherwise, $\dot{Q}_{\mathrm{HVAC}}$,, is set to its upper or lower bound. Incorporating the thermal and the HVAC load models together and writing the system of equations as a linear system of differential equations gives the following state space representation:

$$\begin{aligned} \begin{bmatrix} \dot{T}_{ia} \\ \dot{T}_m \\ \dot{I} \end{bmatrix} &= \begin{bmatrix} \frac{1}{C_{ia}}\left(-K_p - \frac{1}{R_{mi}} - \frac{1}{R_{oi}}\right) & \frac{1}{C_{ia}R_{mi}} & \frac{K_p}{\tau_{\mathrm{I}} C_{ia}} \\ \frac{1}{C_m R_{mi}} & -\frac{1}{C_m R_{mi}} & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} T_{ia} \\ T_m \\ I \end{bmatrix} + \begin{bmatrix} \frac{K_p}{C_{ia}} & \frac{1}{C_{ia}R_{oi}} \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} T_{sp} \\ T_{oa} \end{bmatrix} + \begin{bmatrix} \frac{1}{C_{ia}} \\ 0 \\ 0 \end{bmatrix} \dot{Q}_{\mathrm{other}} \\ \begin{bmatrix} T_{ia} \\ \dot{Q}_{\mathrm{HVAC}} \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ -K_p & 0 & \frac{K_p}{\tau_{\mathrm{I}}} \end{bmatrix} \begin{bmatrix} T_{ia} \\ T_m \\ I \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ K_p & 0 \end{bmatrix} \begin{bmatrix} T_{sp,j} \\ T_{oa} \end{bmatrix} \end{aligned} \tag{15}$$

The combined building thermal and the HVAC model contains the following three inputs: 1) The zone temperature set-point T_{sp} is a controlled input 2) the outdoor air temperature T_{oa} which is a measured input and not under control 3) the $\dot{Q}_{\mathrm{other}}$ which is neither controlled nor measured. The outputs are the measured zone temperature T_{ia} and sensible heat provided to/removed $\dot{Q}_{\mathrm{HVAC}}$.

While the model above does not describe the detailed physical phenomena and dynamics of a building thermal mass and the HVAC load, it balances the trade-off between model accuracy and computational efficiency when used for online control computations. This balance is critical for practical implementation of model predictive control (MPC) since that MPC requires continuous online computation that incorporate feedback. This feedback coupled with online computation are necessary for the prediction accuracy and robustness of the MPC. The system identification procedure followed in this work consists of the following three steps:

1. An excitation signal is applied to the system and the corresponding outputs are recorded.
2. The thermal and HVAC load model parameters are identified first using simulation PEM.
3. The resulting state-space model is augmented with an integrating disturbance model and the Kalman filter gain is estimated using the one-step ahead PEM.

3.1 Step 1: Excitation Signal

Typically, to fit a dynamic model to input-output data, the inputs to a system are manipulated in an open-loop fashion as to sufficiently and persistently excite the system (Ljung *et al.*, 1999). Formally, a signal $\{s(k): k = 0,1,2, \ldots\}$ with spectrum $\phi_s(\omega)$ is persistently excitated (PE) if $\phi_s(\omega) > 0$ for almost all ω where:

$$\phi_s(\omega) = \sum_{m=-\infty}^{\infty} \left(\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} s(k)s(k-m) \right) \tag{16}$$

The approach selected for generating the excitation signals is to generate one from a pseudorandom binary sequence (PRBS). Thus, a PRBS temperature set-point T_{sp} signal is generated for system identification.

3.2 Step 2: Building Thermal Model Identification

In order to perform system identification, we first have to parameterize our model. The state space equations for the building thermal model can be represented compactly as follows:

$$\begin{aligned} \dot{x} &= A_c(\theta)x + B_c(\theta)u \\ y &= C_c(\theta)x + D_c(\theta)u \end{aligned} \tag{17}$$

where $x^T = [T_{ia}, T_m, I]$, $u^T = [T_{sp}, T_{oa}]$, $y^T = [T_{ia}, \dot{Q}_{\mathrm{HVAC}}]$, θ is the parameter vector to be identified. The parameterization followed in this work is presented in the following matrices:

$$\begin{aligned} A_c(\theta) &= \begin{bmatrix} -(\theta_1 + \theta_2 + \theta_3\theta_4) & \theta_2 & \theta_3\theta_4\theta_5 \\ \theta_6 & -\theta_6 & 0 \\ -1 & 0 & 0 \end{bmatrix}, B_c(\theta) = \begin{bmatrix} \theta_3\theta_4 & \theta_1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \\ C_c(\theta) &= \begin{bmatrix} 1 & 0 & 0 \\ -\theta_4 & 0 & \theta_5\theta_4 \end{bmatrix},\ D_c(\theta) = \begin{bmatrix} 0 & 0 \\ \theta_4 & 0 \end{bmatrix} \end{aligned} \tag{18}$$

Where: $\theta_1 = \frac{1}{C_{ia}R_{oi}}$ $\theta_2 = \frac{1}{C_{ia}R_{mi}}$ $\theta_3 = \frac{1}{C_{ia}}$ $\theta_4 = K_p$ $\theta_5 = \frac{1}{\tau}$ $\theta_6 = \frac{1}{C_m R_{mi}}$
Estimation of the model parameters will be conducted using simulation PEM.

3.3 Step 3: Estimation Gain Identification

The model in step 2 only tries to accounts for the main thermal dynamics of the system. In Step 3, the model from step 2 is converted into a discrete time system and it is augmented with another integrating disturbance model of the form described in Eq.2 where the disturbance is selected as follows $B_d = \frac{1}{C_{ia}}$ and $C_d = 0$. Then, a parameterized estimator gain matrix $K(\phi)$ is parameterized with the parameter vector ϕ as follows:

$$K_x(\phi) = \begin{bmatrix} \phi_1 & \phi_2 \\ 0 & 0 \\ \phi_3 & \phi_4 \\ \phi_5 & \phi_6 \end{bmatrix} \tag{19}$$

The identification of the estimator gain $K(\phi)$ will be performed using the one-step ahead PEM.

4. SATURATION DETECTION AND REMOVAL ALGORITHM

The first algorithm proposed is the saturation detection and removal algorithm. The purpose of this algorithm is to detect and remove data that are not useful for system identification purposes. This saturation detection algorithm can handle unknown and potentially, time varying saturation limits. The favorable remaining multiple data sections are extracted and system identification is performed in order fit one model to all data sections.

4.1 Algorithm

We consider a system identification experiment that contains a series of set-point changes. For an experiment that has N set-point changes, the indoor temperature absolute error for a given set-point i at a given time step k is:

$$e_{T_i}(k) = abs(T_{sp_i} - T_{ia}(k)) \text{ , } i = 1,2, \dots N, k = 0,1,2, \dots \tag{20}$$

Each set-pint starts at sampling period S_i and lasts for a certain period of time L_i. For each set-point change, the average absolute error is:

$$AvrgE_i = \frac{\sum_{k=S_i}^{k=S_i+L_i} e_{T_i}(k)}{L_i} \tag{21}$$

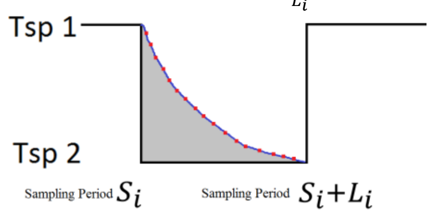

Figure 1. Illustration of temperature error accumulated during a set-point period.

Figure 1 illustrates an example of the temperature error during a set-point (in this case T_{sp2}) that lasts for the period of time L_i .When $AvrgE_i$ exceeds a predefined threshold, we check if the majority of the $\dot{Q}_{\text{HVAC}}$ data in that set are in the transient region of $\dot{Q}_{\text{HVAC}}$. If they are, we do not consider that set to contain significant saturation, and therefore we do not remove that set of data. If the majority of the $\dot{Q}_{\text{HVAC}}$ data in that set are not in the transient region of $\dot{Q}_{\text{HVAC}}$, we remove that set of data. While calculating the temperature absolute error may result in exceeding the threshold when the system has higher order responses (i.e. overshoot or oscillations), checking if the $\dot{Q}_{\text{HVAC}}$ is within the transient region will counter the removal of these data in the case that they are good data.
Due to the fact that in many HVAC systems, the upper and lower $\dot{Q}_{\text{HVAC}}$ saturation limits may not be known a priori, we base the maximum and minimum saturation limits on the $\dot{Q}_{\text{HVAC}}$ measured. In addition, these upper and lower saturation limits can be time varying as illustrated in the figure below:

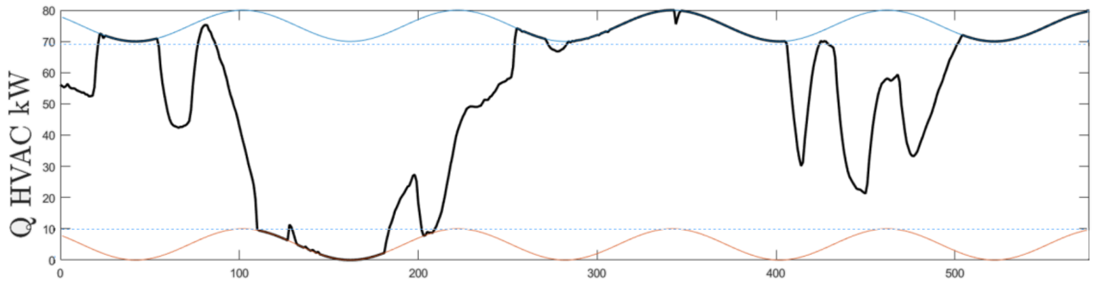

Figure 2. Illustration of time varying $\dot{Q}_{HVAC}$ saturation limits of HVAC systems.

Thus, we take a conservative region of the $\dot{Q}_{\text{HVAC}}$ range to be our transient region of $\dot{Q}_{\text{HVAC}}$. Ideally, we would like to truncate the region to include only the $\dot{Q}_{\text{HVAC}}$ data between the two dashed lines in the figure above. However, since these values are also not known, we will truncate this region by a conservative percentage to avoid including any saturated $\dot{Q}_{\text{HVAC}}$ data. We take the entire set of $\dot{Q}_{\text{HVAC}}$ data for the entire system identification experiment. Then, we find the minimum and maximum values of $\dot{Q}_{\text{HVAC}}$. After that, we truncate the region by a total of 30 %. This truncation is done as follows:

$$\dot{Q}_{truncate} = 0.15(\max(\dot{Q}_{\text{HVAC}}) - \min(\dot{Q}_{\text{HVAC}})) \quad (22)$$

Then we define: $\dot{Q}_{max} = \max(\dot{Q}_{\text{HVAC}}) - \dot{Q}_{truncate}$ and $\dot{Q}_{min} = \min(\dot{Q}_{\text{HVAC}}) + \dot{Q}_{truncate}$

The range of Q_{HVAC} that lies between $\dot{Q}_{min}$ and $\dot{Q}_{max}$ is what we call the transient region of $\dot{Q}_{\text{HVAC}}$. When $AvrgE_i$ exceeds a predefined threshold, we check if more than 50% of the $\dot{Q}_{\text{HVAC}}$ data in that set are in the transient region of $\dot{Q}_{HVAC}$. If they are, we do not remove that set, otherwise, we remove that set and we consider it to have significant saturation. The motivation behind choosing this as a threshold comes from the respond of linear systems under a PI controller to a set point change. Consider a linear system under PI control operating exactly at T_{sp1}. When this system is subject a set-point changes from $T_{sp1}$to T_{sp2}, the system under the PI controller will exponential go towards the desired T_{sp2}. The speed of going to T_{sp2} depends on the linear system time constant as well as the PI controller. These responses will look like the curved lines in the following figure:

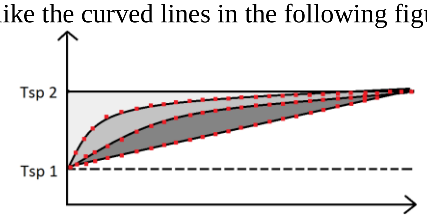

Figure 3. Illustration of linear system response behavior to a step-input (over-damped systems).

However, these lines (in theory and under very ideal circumstances), will never cross the straight line going from $T_{sp1}$to T_{sp2}. Therefore, the average absolute error for the entire period will never be greater than $0.5(T_{sp_{max}} - T_{sp_{min}})$. The reason we check for significant errors in temperatures before checking the $\dot{Q}_{\text{HVAC}}$ is that, in general, the system may be under control even when $\dot{Q}_{\text{HVAC}}$ is near saturation limits.

4.2 Dealing with long set-point periods

The threshold above does not work perfectly for temperature set-points that last for a long time (here long time is a relative term that has to take into account the time constant of the system and the aggressiveness of the PI controller that supplies the $\dot{Q}_{\text{HVAC}}$). For example, if a T_{sp} lasts for a very long time and the system, under ideal condition, reaches the desired set-point very early on, then having the threshold as $0.5(T_{sp_{max}} - T_{sp_{min}})$ is very loose. In system identification, before delivering the input excitation signal, we specify a minimum time duration (denoted as $MinL_{Tsp}$) that a set-point should stay at a specific value before it changes to another value. This minimum time is based on the response of the HVAC equipment and control to step inputs; which could be approximated from experiments prior to system identification. We assume that this minimum time is relatively equal to the time it takes for the system to go from one T_{sp} to the other under ideal condition. Thus, the threshold needs to be scaled accordingly for T_{sp} that last longer than $MinL_{Tsp}$. If we call the time that the system stays at a given set-point i as L_{Tsp_i}. Then, the threshold for each set-point should be scaled by a factor equal to $\frac{MinL_{Tsp}}{L_{Tsp_i}}$. This makes the threshold more sensitive for T_{sp} that last a longer duration than $MinL_{Tsp}$ and accounts for the resulting effects of that on the average temperature error threshold. Therefore, the new threshold becomes:

$$Threshold_i = \frac{MinL_{Tsp}}{L_{Tsp_i}} \frac{\left(T_{sp_{max}} - T_{sp_{min}}\right)}{2} \quad (23)$$

4.3 Dealing with Dead-Band

In the case that the HVAC system has a dead-band in tracking temperature set-points, two changes will occur in the algorithm. The first change is how we compute the average temperature error. Since any temperature within the dead-band is considered tracking, only temperatures outside the dead-band should be considered for the average absolute error computation. Therefore, the computation of the average absolute error becomes:

1) If measured temperatures are within the dead-band , the temperature error ($e_{T_i}(k)$) equals zero
2) If not, compute the error (or absolute error). The error here is computed as the difference between the temperature and the bounds of the dead-band (Not the actual T_{sp}) as follows:

$$e_{T_i}(k) = \min(abs\left(T_{sp_i} + DeadBand - T_{ia}(k)\ ;\ T_{sp_i} - DeadBand - T_{ia}(k)\right)) \quad (28)$$

After that, we compute the average absolute error in the same fashion. The error region that is accounted for in this algorithm can be illustrated by the grey area in the figure below:

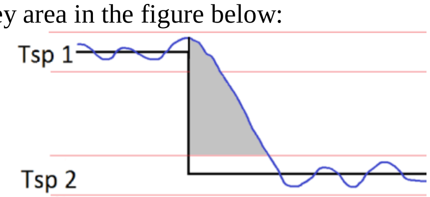

Figure 4. Illustration of temperature error accumulated when HVAC systems have a dead-band.

The second change that has to be accounted for when having a dead-band is the threshold. Since the region of non-zero error has shrunk, the threshold has to be modified in order to account for this change. While it may seem from the first glance that the region should shrink by a factor that is equal to the dead-band, we will shrink the region by twice the dead-band. In order to understand why this is the case, we refer to the figure above. Before doing a set-point change, the temperature could be floating anywhere within the dead-band region of T_{sp}. Therefore, there is a possibility to be at the top or at the bottom of T_{sp1} dead-band before requesting the change to T_{sp2}. Therefore, we take a more conservative threshold that can account for this fact too. The threshold now becomes:

$$Threshold_i = \frac{MinL_{Tsp}}{L_{Tsp_i}} \frac{\left(T_{sp_{max}} - T_{sp_{min}} - 2 * DeadBand\right)}{2} \quad (24)$$

4.4 Adapting to non-PRBS Input Signals

In many cases, the input signal in the system identification experiment has more than two set-point values (non-PRBS signals such as Gaussian signal or PRBS that changes values with time). To make the algorithm account to such changes, the threshold should also adapt to non-PRBS input signals. Therefore, instead of taking $T_{sp_{max}} - T_{sp_{min}}$ as a base for the error, we take $T_{sp_i} - T_{sp_{i-1}}$ at a given i^{th} T_{sp}. This makes the threshold become:

$$Threshold_i = \frac{MinL_{Tsp}}{L_{Tsp_i}} \frac{abs(T_{sp_i} - T_{sp_{i-1}}) - 2 * DeadBand}{2} \quad (25)$$

4.5 Complete Algorithm Summary

The saturation detection and removal algorithm is:

1) If the zone temperature is within the dead-band, the temperature error ($e_{T_i}(k)$) equals zero.
2) If the zone temperature is not within the dead-band, compute the error (or absolute error). The error here is computed as the difference between the temperature and the bounds of the dead-band (Not the actual T_{sp}) as follows:

$$e_{T_i}(k) = \min(abs\left(T_{sp_i} + DeadBand - T_{ia}(k)\ ;\ T_{sp_i} - DeadBand - T_{ia}(k)\right))$$

3) For each set-point change, the average absolute error is:

$$AvrgE_i = \frac{\sum_{k=S_i}^{k=S_i+L_i} e_{T_i}(k)}{L_i}$$

4) We check which set-points have an $AvrgE_i$ that exceeds the following threshold:

$$Threshold_i = \frac{MinL_{Tsp}}{L_{Tsp_i}} \frac{abs(T_{sp_i} - T_{sp_{i-1}}) - 2 * DeadBand}{2}$$

5) For the T_{sp} periods that exceeded this threshold, if more than half of the $\dot{Q}_{\mathrm{HVAC}}$ data during that period are not in the transient region of Q_{HVAC}, then that data set contains significant saturation and it is removed. Otherwise, the data are not kept as useful data for system identification.

6) The transient region of $\dot{Q}_{\mathrm{HVAC}}$ is computed as follows:

$$\dot{Q}_{truncate} = 0.15(\max(\dot{Q}_{\mathrm{HVAC}}) - \min(\dot{Q}_{\mathrm{HVAC}}))$$

$$\dot{Q}_{max} = \max(\dot{Q}_{\mathrm{HVAC}}) - \dot{Q}_{truncate} \text{ and } \dot{Q}_{min} = \min(\dot{Q}_{\mathrm{HVAC}}) + \dot{Q}_{truncate}$$

The range of $\dot{Q}_{\mathrm{HVAC}}$ that lies between $\dot{Q}_{min}$ and $\dot{Q}_{max}$ is what we call the transient region of $\dot{Q}_{\mathrm{HVAC}}$.

7) The remaining multiple data sections are extracted and system identification is performed in order fit one model to all data sections.

5. Efficient Model Generation Algorithm

Due to the non-convexity and complexity of the system identification optimization problem, it is always desired to avoid initial guesses which lead to local optima that result in poor models. It is difficult to know a priori whether a given random initial guess would lead to an accurate or in-accurate model. Therefore, we employ the following algorithm that tries to eliminate bad initial guesses and yield suitable starting guesses that ultimately may lead to acceptable fits of the model to the data.

5.1 Algorithm:

1) Generate N initial guesses of the system parameters where each of these initial guesses construct a stable A matrix and a stable $A - KC$ matrix.
2) Run each system identification problem for a small number of iterations M.
3) Record the cost function for each for the N runs along with the parameters values acquired.
4) Remove initial guesses that led to an unstable A matrix or unstable $A - KC$ matrix after M iterations
5) Remove initial guesses that led to a very high condition number of the A matrix or $A - KC$ matrix
6) Remove initial guesses going towards the same local optimal solution as another initial guess. This can be done by comparing the cost function along with the parameters values after the M iterations. Then, removing initial guesses that have similar cost function value as well as similar parameter values with a small tolerance of variation.
7) Rank the cost function values in an ascending order from lowest to highest.
8) Start from the initial guess that gave the lowest cost function value and run it all the way to local optimality.
9) Evaluate the stability, controllability, observability, and condition numbers of the A matrix and the $A - KC$ matrix. The condition number of these matrices is desired to be smaller than a large number L.
10) If conditions in step 9 are satisfied, terminate the algorithm and the identified model is obtained. If conditions in step are not satisfied, re-run system identification using the remaining initial guesses going from the guesses that gave the lowest cost function to the guesses that gave the highest cost function until conditions in step 9 are satisfied.

Implementing this algorithm avoids running poor initial guess all the way to local optimality. In addition, this algorithm also avoids running different initial guesses that go towards the same local optimal solution. This algorithm turns out to be much more computationally beneficial than simply generating many initial guesses, running each to optimality, checking the quality of the obtained model, then repeating the entire process if the model obtained is not satisfactory.

6. RESULTS

In this section, we analyze the proposed saturation detection and removal algorithm, and evaluate the improvement obtained when conducting system identification. In addition, we also evaluate the computational benefit obtained from implementing the efficient model generation algorithm. A system identification experiment is conducted on an area of a building located in Milwaukee, Wisconsin. Thus, a PRBS temperature set-point T_{sp} signal is generated varying between the values of 73º F and 75º F. The PRBS signal is designed to stay at each set-point value a minimum duration of 1 hour (i.e. $MinL_{Tsp} = 1$). The system identification experiment took place over the course of two days where inputs (T_{sp} and T_{oa}) and outputs ($T_{ia}, \dot{Q}_{\text{HVAC}}$) were collected at a sampling frequency of 10 minutes. The HVAC system considered has a dead-band of 0.5º F. The data collected are processed through the proposed saturation detection and removal algorithm and the resulting data are utilized for parameter estimation. The state-space equations (Eq. 17 and Eq.20) are used to model the building thermal and HVAC dynamics and the parameters of this model are identified first using simulation PEM. Then, the resulting state-space model is augmented with an integrating disturbance model and the Kalman filter gain in the form of equation (Eq. 22) is estimated using the one-step ahead PEM. In order to efficiently obtain initial guesses that lead to accurate model parameters, the proposed efficient model generation algorithm is implemented in both parameter estimation steps where both M and N are equal to 50.

After that, the model was validated over the same two days' worth of input-output data used for system identification by plotting the actual outputs versus the 1-step prediction of the outputs which are demonstrated in the following figures:

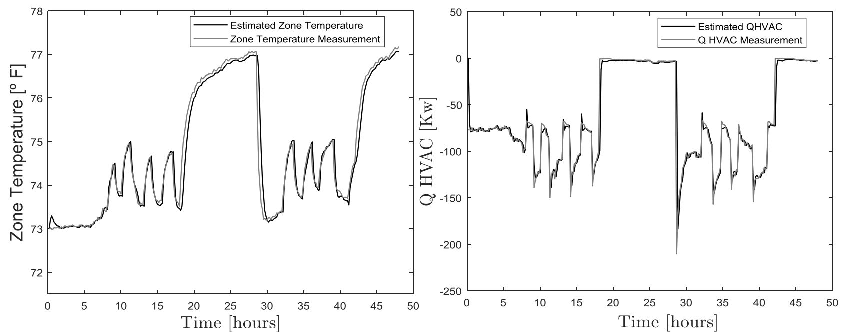

Figure 5. Measured outputs trajectories (in grey) and the 1-step prediction of the outputs (in black).

The identified model demonstrates accurate 1-step prediction. The saturation detection algorithm was successful in detecting and removing the saturation periods which span the periods from hour 18 to 28 and from hour 42 to 48. The 2-norm of the 1-step prediction error in the outputs for the two days is 1.6175 for the zone temperature and 16.6594 for $\dot{Q}_{\text{HVAC}}$.When implementing the efficient model generation algorithm over the excited data sections, the total computational time it took for the system identification algorithm to yield this model is 307 CPU seconds.

For comparison, a second system identification procedure was conducted over the same collected input-output data where the input-output data were not passed through the saturation detection and removal algorithm. In addition, the efficient model generation algorithm was not used and instead, each initial guess was run to local optimality and the identified model's stability, controllability, observability, and condition numbers of the identified A matrix and the $A - KC$ matrix were evaluated. If identified model is not satisfactory, another random guess is generated and the system identification is solved again to local optimality. Eventually, when an acceptable model is identified, it was validated over the same two days' worth data by plotting the actual outputs versus the 1-step prediction of the outputs which are demonstrated in the following figures:

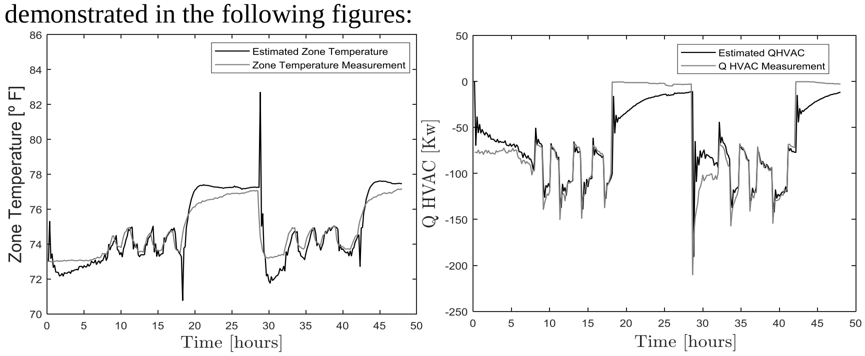

Figure 6. Measured outputs trajectories (in grey) and the 1-step prediction of the outputs (in black).

The 2-norm of the 1-step prediction error in the outputs for the two days is 7.8597 for the zone temperature and 94.4387 for $\dot{Q}_{\mathrm{HVAC}}$.The total computational time for this method to result in an acceptable model is 1923 CPU seconds. As we can see from the figures above, the predictions have a low degree of accuracy.

7. CONCLUSIONS

This paper presented two strategies to overcome some of the difficulties faced in conducting system identification of building thermal and HVAC dynamics. The first strategy is an algorithm that is designed to detect and remove saturation data caused by buildings HVAC systems. The goal of this algorithm is to extract the useful excited data for system identification in order to produce models that accurately capture the crucial buildings HVAC dynamics. The second strategy presents a systematic procedure that eliminates poor initial guesses when solving the non-convex system identification problem to identify building thermal and HVAC dynamics. The procedure presents guidelines to tests the quality of initial guesses and yields guesses that ultimately lead to models that accurately fit the data and captures the system dynamics. The proposed strategies were implemented to identify a grey-box model parameters of building thermal and HVAC model as well as a Kalman gain. The system identification was applied to data collected from real building HVAC systems and have shown to successfully work in practice.

REFERENCES

Mendoza-Serrano, D. I., & Chmielewski, D. J. (2012). HVAC control using infinite-horizon economic MPC. In *Proceedings of the 51st IEEE Conference on Decision and Control* (pp. 6963–6968). Maui, HI.

Afram, A., & Janabi-Sharifi, F. (2014). Afram, A., & Janabi-Sharifi, F. (2014). Theory and applications of HVAC control systems – A review of model predictive control (MPC). *Building and Environment*, 72, 343–355.

Rawlings J. B., Patel N. R., Risbeck M. J., Maravelias C. T., Wenzel M. J. & Turney R. D. (2017, In press). Economic MPC and real-time decision making with application to large-scale HVAC energy systems. *Computers & Chemical Engineering*.

Alanqar, A., Ellis, M. & Christofides, P. D. (2015). Economic Model Predictive Control of Nonlinear Process Systems Using Empirical Models. *AIChE Journal.*, 61, 816-830.

Alanqar, A., Durand, H. & Christofides, P. D. (2015). On Identification of Well-Conditioned Nonlinear Systems: Application to Economic Model Predictive Control of Nonlinear Processes. *AIChE Journal.*, 61, 3353-3373.

Ellis, M. J., Wenzel, M. J., & Turney, R. D. (2016). System Identification for Model Predictive Control of Building Region Temperature. In *Proceedings of the 4th International High Performance Buildings Conference* (p. 3583). West Lafayette, IN.

Lin, Y., Middelkoop, T., & Barooah, P. (2012). Issues in identification of control-oriented thermal models of zones in multi-zone buildings. In *Proceedings of the 51st IEEE Conference on Decision and Control* (pp. 6932–6937). Maui, HI.

Widanage, W. D., Godfrey, K. R., & Tan, A. H. (2004). System identification in the presence of a saturation nonlinearity. *Control, University of Bath, UK, ID-141*, pp. 5.

Wills, A., Schön, T.B., Ljung, L. and Ninness, B., (2013). Identification of Hammerstein–Wiener Models. *Automatica*, *49*, 70-81.

Ljung, L. (1999). *System Identification: Theory for the User, 2nd ed.* (Prentice Hall PTR, Upper Saddle River).

Pannocchia, G. & Rawlings, J. B. (2003) Disturbance Models for Offset-Free Model-Predictive Control. *AIChE Journal*, 49,426-437.