Modeling the Thermal Dynamics of Buildings: A Latent-Force-Model-Based Approach

SIDDHARTHA GHOSH, University of Southampton
STEVE REECE, University of Oxford
ALEX ROGERS, University of Southampton
STEPHEN ROBERTS, University of Oxford
AREEJ MALIBARI, King Abdulaziz University
NICHOLAS R. JENNINGS, University of Southampton and King Abdulaziz University

Minimizing the energy consumed by heating, ventilation, and air conditioning (HVAC) systems of residential buildings without impacting occupants' comfort has been highlighted as an important artificial intelligence (AI) challenge. Typically, approaches that seek to address this challenge use a model that captures the thermal dynamics within a building, also referred to as a thermal model. Among thermal models, gray-box models are a popular choice for modeling the thermal dynamics of buildings. They combine knowledge of the physical structure of a building with various data-driven inputs and are accurate estimators of the state (internal temperature). However, existing gray-box models require a detailed specification of all the physical elements that can affect the thermal dynamics of a building a priori. This limits their applicability, particularly in residential buildings, where additional dynamics can be induced by human activities such as cooking, which contributes additional heat, or opening of windows, which leads to additional leakage of heat. Since the incidence of these additional dynamics is rarely known, their combined effects cannot readily be accommodated within existing models.

To overcome this limitation and improve the general applicability of gray-box models, we introduce a novel model, which we refer to as a latent force thermal model of the thermal dynamics of a building, or LFM-TM. Our model is derived from an existing gray-box thermal model, which is augmented with an extra term referred to as the learned residual. This term is capable of modeling the effect of any a priori unknown additional dynamic, which, if not captured, appears as a structure in a thermal model's residual (the error induced by the model). More importantly, the learned residual can also capture the effects of physical elements such as a building's envelope or the lags in a heating system, leading to a significant reduction in complexity compared to existing models.

To evaluate the performance of LFM-TM, we apply it to two independent data sources. The first is an established dataset, referred to as the FlexHouse data, which was previously used for evaluating the efficacy of existing gray-box models [Bacher and Madsen 2011]. The second dataset consists of heating data logged within homes located on the University of Southampton campus, which were specifically instrumented to collect data for our thermal modeling experiments. On both datasets, we show that LFM-TM outperforms existing models in its ability to accurately fit the observed data, generate accurate day-ahead internal temperature predictions, and explain a large amount of the variability in the future observations. This,

This work was funded in the United Kingdom by the EPSRC Intelligent Agents for Home Energy Management project (EP/I000143/1) and ORCHID programme grant (EP/I011587/1), and in the Kingdom of Saudi Arabia by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah (9-15-1432-HiCi).
Authors' addresses: S. Ghosh, A. Rogers, and N. R. Jennings, School of Electronics and Computer Science, University of Southampton; email: sidsmailin@gmail.com; S. Reece and S. Roberts, Machine Learning Research Group, University of Oxford; A. Malibari and N. R. Jennings, Department of Computing and Information Technology, King Abdulaziz University, Saudi Arabia.

DOI: http://dx.doi.org/10.1145/2629674

along with the fact that we also use a corresponding efficient sequential inference scheme for LFM-TM, makes it an ideal candidate for model-based predictive control, where having accurate online predictions of internal temperatures is essential for high-quality solutions.

Categories and Subject Descriptors: G.3 [**Mathematics of Computing**]: Probability and Statistics and Stochastic Processes; G.1.10 [**Mathematics of Computing**]: Applications; J.2.8 [**Physical Sciences and Engineering**]: Mathematics and Statistics; J.7.5 [**Computers in Other Systems**]: Process Control

General Terms: Experimentation, Measurement, Performance, Theory, Validation

Additional Key Words and Phrases: Artificial intelligence, sustainability, HVAC, smart homes, residential buildings, energy savings, thermal models, latent force models

ACM Reference Format:
Siddhartha Ghosh, Steven Reece, Alex Rogers, Stephen Roberts, Areej Malibari, and Nicholas R. Jennings. 2015. Modeling the thermal dynamics of buildings: A latent-force-model-based approach. ACM Trans. Intell. Syst. Technol. 6, 1, Article 7 (March 2015), 27 pages.
DOI: http://dx.doi.org/10.1145/2629674

1. INTRODUCTION

On the path to our low-carbon future, increased attention is being paid to approaches that optimize and reduce the energy consumption of the HVAC systems of buildings, which can account for up to 40% of the total energy consumed and up to 33% of the total carbon emissions in some countries [Aswani et al. 2012; Privara et al. 2013]. Heating or cooling is necessary to maintain a comfortable temperature within a home. However, this comfort comes at a price, both in terms of the monetary cost to the householder and in terms of the carbon emissions resultant from the process. Consequently, there is a need to optimize the operation of the HVAC system to minimize the energy consumed by heating or cooling without impacting the householder's comfort. This has been highlighted as an important artificial intelligence (AI) challenge [Ramchurn et al. 2012; Evans 1991].

Toward this end, an important strand of work consists of approaches that treat HVAC optimization as an optimal control problem. A model-based predictive control (MPC) algorithm is then used to obtain a solution in the form of a sequence of accurate and efficient control actions [Mady et al. 2011; Aswani et al. 2012; Privara et al. 2013; Oldewurtel et al. 2012; Rawlings 2000]. Researchers have also modeled heating optimization as a Markov decision process (MDP), whose solution yields a sequence of optimal heating or cooling actions [Urieli and Stone 2013; Dounis and Caraiscos 2009]. Researchers have also explored the idea of exploiting occupancy information to more efficiently control HVAC systems [Lu et al. 2010; Scott et al. 2011].

A key component of these approaches is a model that captures the thermal dynamics of a building, commonly referred to as a thermal model. This is typically a heat balance equation that describes the evolution of the internal temperature within a building in response to various physical processes [Bacher and Madsen 2011]. Having a thermal model enables an HVAC controller to make more informed decisions on when to turn on the HVAC system; it can determine how a building would respond to heating or cooling and how long it would take for a building to reach the set-point temperature. Within the AI literature, little research exists on the development of accurate, robust, and adaptive thermal models that can accurately model the thermal dynamics of homes. This is important because a poor thermal model, if used within an MPC algorithm or within an MDP for HVAC optimization, will generate suboptimal heating or cooling plans. In other words, these algorithms will fail to optimally control the heating or cooling actions. Consequently, a householder may experience high discomfort. In both cases, this would lead to a loss of confidence in the operation of the HVAC system.

Substantive research on the development of thermal models does exist in the building science literature, where one of the most common approaches for statistically modeling the thermal dynamics of buildings is to explicitly include all the physical elements that influence the evolution of the internal temperature. These models are referred to as gray-box models and have been shown to be accurate at modeling the thermal dynamics in various settings [Mady et al. 2011; Jimenez and Madsen 2008; Aswani et al. 2012; Andersen et al. 2000; Privara et al. 2013].[1] Although other statistical modeling approaches, such as subspace identification and MPC-relevant identification, have been proposed [Privara et al. 2013], gray-box models have been shown to be competitive with these approaches at generating predictions of internal temperature [Privara et al. 2013]. In addition, all elements within a gray-box model correspond to a specific physical element within a building, which helps in the interpretability of the models [Bacher and Madsen 2011]. Furthermore, efficient techniques exist for model identification, inference, and parameter estimation in gray-box models, which facilitates their application to data from a wide variety of buildings [Madsen and Holst 1995; Bacher and Madsen 2011; Andersen et al. 2000].

The simplest gray-box model, subsequently referred to as the *Ti* model, captures two key dynamics: the heat from the heating system and the leakage of heat to the ambient environment outside the building. However, more complex models have also been proposed, which incrementally incorporate the effect of various physical factors such as the lags in the building's heating system (denoted as the *TiTh* model), its envelope (denoted as the *TiTeTh* model), and sensor lags (denoted as the *TiTeThTs* model) [Bacher and Madsen 2011] (we refer the reader to Appendix A for a detailed description of these models). Such models can have several parameters, and as they become increasingly complex, the number of parameters in the models proportionately increases. Now, these parameters are typically learned from the data collected from a building using least-squares fitting or maximum likelihood estimation [Privara et al. 2012]. However, such learning becomes more difficult as more parameters are incorporated in a model [Fernández Slezak et al. 2010]. Moreover, models with poor estimated parameters generate poor predictions of the internal temperature.

Another key limitation of existing gray-box models is that they require a complete a priori specification of all the physical elements that influence the thermal dynamics of a building [Privara et al. 2013]. This can be problematic, especially when modeling the thermal dynamics of residential buildings, where additional dynamics can be induced by human activities. To explain this further, domestic activities such as cooking or the deployment of an additional heater during a cold spell can contribute a significant amount of extra heat, while an open window can cause additional heat loss. Since the incidence of these additional dynamics cannot be fully known a priori, they cannot be explicitly represented within a model. Typically, any additional factor that influences the thermal dynamics but that has not been explicitly included in the model gets accommodated as the variance of a Wiener noise process, or the stochastic noise term [Yu et al. 2012; Bacher and Madsen 2011]. However, this is only legitimate if the additional factor is uncorrelated across time lags and has white noise properties. Unfortunately, in practice, this assumption is often violated. For instance, a thermal lag between the heater coming on and the temperature rising, if not explicitly included in

[1]Previously, gray-box models have been classified into two categories [Privara et al. 2013]. Probabilistic semiphysical models have a stochastic noise term [Bacher and Madsen 2011]. In contrast, models that do not include a stochastic noise term are termed deterministic gray-box models. In this article, we only consider the former. Explicitly including a noise term in the model allows us to more effectively cope with any process or measurement noise that can impair the ability of a model to generate accurate internal temperature predictions.

the process model, will induce structure in the model's residual [Durbin and Koopman 2001]. Such a residual will have statistically significant correlation across several time lags. Similarly, during cold spells, a householder may employ an additional heater that will contribute additional heat, which, if not explicitly modeled, will also result in the residual having structure and violating the white noise property. In contrast, we believe that an adaptive model should be capable of modeling any additional factors that have not been specified initially from the data. Existing gray-box models lack such flexibility.

Finally, the most mature work on gray-box modeling to date has empirically demonstrated the accuracy of the models only at fitting observed data [Bacher and Madsen 2011; Andersen et al. 2000]. However, it is possible for a model to fit the observed data very well but generate poor predictions on unseen inputs [Bishop 2006; Cawley and Talbot 2010]. This process is known as overfitting. Existing gray-box models have not been thoroughly evaluated on their ability to generate accurate predictions of internal temperature. Consequently, they may not generalize well.

To address the aforementioned limitations, we introduce a novel model for the thermal dynamics of buildings. In particular, we employ latent force models (LFMs) to model the thermal dynamics of a building [Álvarez et al. 2009]. LFMs have previously demonstrated their utility in diverse application areas of AI such as computational biology, understanding motion patterns, and target tracking [Álvarez et al. 2009; Álvarez et al. 2010; Hartikainen and Särkkä 2011]. In LFMs, the physical knowledge of a dynamic system is represented as one or more coupled differential equations. These are combined with a data-driven framework in which physical factors that affect the dynamics of the system are modeled using Gaussian processes (GPs) that are drawn from distributions over the space of suitable functions [Álvarez et al. 2009; Álvarez et al. 2010; Hartikainen and Särkkä 2011]. This is particularly useful in situations where a detailed specification of all physical factors that can affect the dynamics is not always possible due to data being unavailable [Hartikainen and Särkkä 2011]. This is often the situation in buildings, where the effect of physical elements such as the envelope or the lags in the heating system are difficult to measure and recording data on human activities that can affect the thermal dynamics is impractical.

In this article, we extend the standard LFM framework to develop latent force thermal models of buildings, which we refer to as LFM-TM. This is based on augmenting the simple gray-box thermal model described previously with a time-varying residual, $\epsilon(\mathbf{t})$, which attempts to model the latent forces in a building that can influence the evolution of the internal temperature and cause alterations in the thermal dynamics. These latent forces include (1) all a priori unknown residual dynamics induced by human activity and (2) the effect of the physical elements such as a building's envelope or the lags in its heating system. To explain this further, in the LFM-TM, some knowledge of the physics of thermal dynamics in a building is explicitly incorporated, with $\epsilon(\mathbf{t})$ modeling the effect of the unobserved physical elements such as envelopes and lags, as well as human-induced additional thermal dynamics. To achieve this, we assume that $\epsilon(\mathbf{t})$ is a function drawn from a GP prior, which is subsequently learned when the LFM-TM is applied to real data. At the end of the learning stage, $\epsilon(\mathbf{t})$ is able to capture a model of the residual dynamics. Overall, the LFM-TM has fewer parameters than some of the more complex gray-box models discussed earlier. As shown later in this article (in Sections 5 and 6, respectively), its reduced representation does not have a detrimental effect on its ability to generate accurate predictions of internal temperature in comparison with the more complex gray-box models.

We use a sequential inference scheme for the LFM-TM that was originally developed for the LFM [Hartikainen and Särkkä 2010]. Standard inference in LFM is computationally expensive and scales as $\mathcal{O}(D^3N^3)$, where D is the dimensionality of the state

in the LFM and N is the total time over which observations are recorded [Hartikainen and Särkkä 2011]. In contrast, sequential inference for the LFM is more efficient as it reduces the complexity to $\mathcal{O}(D^3N)$. The inference scheme we adopt is based on a state-space representation of the LFM, which can be solved using standard tools such as the Kalman filter [Hartikainen and Särkkä 2010].

To demonstrate the utility of our LFM-TM, we apply it to model the thermal dynamics of residential buildings. In contrast with office or commercial buildings, modeling the thermal dynamics of residential buildings has received relatively little attention. In particular, most existing work has only been tested on simulated data [Yu et al. 2012; Rogers et al. 2011]. Comparatively less research exists on thermal modeling of residential buildings using real data [Mozer et al. 1996].

Noting this, we first apply LFM-TM to an existing dataset (referred to as the data from FlexHouse), which has previously been used in the development of a range of gray-box thermal models [Bacher and Madsen 2011]. However, FlexHouse is an experimental residential building and does not have any occupants. Consequently, it is not possible to test how human-induced activity in a home can affect the performance of our LFM-TM. To address this issue, we also address the development of thermal models of residential buildings inhabited by householders. Specifically, we instrumented homes of university staff to collect home heating data. In both cases, we show that LFM-TM is able to explain the observed data much better than previously specified gray-box models. This results in LFM-TM yielding significantly higher log-likelihoods over the forecast sample than existing models. Furthermore, not only does the LFM-TM have a simpler structure, with fewer parameters and no loss of efficacy in comparison with existing gray-box models, but also we show that it outperforms them by generating much better internal temperature predictions.

In summary, we advance the state of the art in the following ways:

—We extend the LFM paradigm to develop LFM-TMs, an adaptive gray-box model of the thermal dynamics of a building. It can be used to model the thermal dynamics of a building without requiring a priori knowledge of all the physical elements that can have an effect on the thermal dynamics. It generates accurate internal temperature predictions in the face of human activities, which can interplay with the physical elements in altering the thermal dynamics. This makes it particularly suited for modeling the thermal dynamics of residential buildings.
—On an established dataset [Bacher and Madsen 2011], we demonstrate that LFM-TM yields log-likelihoods that are almost 17% greater than the best-performing gray-box model. Furthermore, it generates one-step predictions that are at least 25% more accurate.
—On data collected from two different homes with occupants, we show that LFM-TM yields log-likelihoods that are at least 38% greater than the best-performing gray-box model. It also generates day-ahead predictions, which are at least 14% more accurate in comparison with existing gray-box models. Finally, we observe that LFM-TM is able to explain a greater proportion of the variance in the future observations.

The remainder of this article is structured as follows. In Section 2, we introduce standard LFMs. In Section 3, we described the development of LFM-TM. Next, in Section 4, we describe how we acquired the data for all the experiments presented in the article. In Section 5, we apply LFM-TM to a standard previously published dataset and demonstrate its superiority over competing gray-box models. We then apply the LFM-TM to modeling the thermal dynamics of real homes, which is described in Section 6. Section 7 concludes.

2. BACKGROUND ON LATENT FORCE MODELS

Latent force models have received considerable interest as they combine underlying physical knowledge of a system with data-driven models expressed as Bayesian nonparametric Gaussian process priors [Álvarez et al. 2009; Hartikainen and Särkkä 2010]. We describe the LFMs in detail in Section 2.1. Subsequently, in Section 2.2, we explain an efficient sequential inference scheme for inference in an LFM. Finally, in Section 2.3, we describe how inference and predictions are undertaken as per this scheme.

2.1. Latent Force Models

Consider a physical system consisting of a single output process $x(t)$ that is modeled as a linear first-order differential equation:

$$A\frac{\mathrm{d}x(t)}{\mathrm{d}t} + \kappa x(t) = \sum_{r=1}^{R} S_{1,r} u_r(t), \tag{1}$$

where $u_r(t)$ are driving processes that have an effect on the behavior of $x(t)$. These processes are not observed and hence they are also referred to as latent forces. In order to correctly accommodate these unknown forces, they are given independent GP priors $u_r(t) \sim \mathcal{GP}(m(t), k_{u_r}(t,t'))$, $r = 1, \ldots, R$, where $m(t)$ is an appropriate mean function (taken usually to be zero without loss of generality) and $k_{u_r}(t,t')$ a suitably chosen covariance function.

As described in Álvarez et al. [2009] inference in this approach is based on closed-form computation of the covariance functions of $x(t)$ and $dx(t)/dt$ and all the required cross-covariances by solving the differential equation. It is possible to represent differential equations of the type in Equation (1) as state-space models, which in case of Equation (1) can be done as follows:

(1) Define state and input vectors as $\mathbf{x}(t) = (x(t), dx(t)/dt)$ and $\mathbf{u}(t) = (u_1(t) \ldots u_R(t))^T$.
(2) Define matrices

$$\mathbf{F} = \begin{pmatrix} 0 & 1 \\ -\frac{\kappa}{A} & 0 \end{pmatrix} \tag{2}$$

and

$$\mathbf{L} = \begin{pmatrix} 0 & \cdots & 0 \\ \frac{S_{1,1}}{A} & \cdots & \frac{S_{1,R}}{A} \end{pmatrix}. \tag{3}$$

This model can be written in the form

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{F}\mathbf{x}(t) + \mathbf{L}\mathbf{u}(t). \tag{4}$$

The differential equation then has the following solution:

$$\mathbf{x}(t) = \Phi(t)\mathbf{x}(t_0) + \int_{t_0}^{t} \Phi(t-s)\mathbf{L}\mathbf{u}(s)ds, \tag{5}$$

where $\Phi(\tau)$ denotes the matrix exponential $\Phi(\tau) = \exp(\mathbf{F}\tau)$. The required covariance terms can now be evaluated as follows:

$$\begin{aligned} E[\mathbf{x}(t)\mathbf{x}(t')] &= \Phi(t-t_0)\mathbf{P}_x^0\Phi(t'-t_0)^T \\ &\quad + \int_{t_0}^{t'}\int_{t_0}^{t} \Phi(t-s)\mathbf{L}\mathbf{K}_{uu}(s,s')\mathbf{L}^T\Phi(t'-s')^T dsds', \end{aligned} \tag{6}$$

where $\mathbf{P}_x^0$ is the prior covariance of $\mathbf{x}(t)$ and $\mathbf{K}_{uu}(s, s')$ is the joint covariance of all the latent forces between time instants s and s'. Since we assume independence across forces, $\mathbf{K}_{uu}(s, s')$ is block diagonal.

If the covariance functions of the latent forces are modeled by squared exponential covariance functions,

$$k_{u_r}(\tau) = \exp\left(-\frac{\tau^2}{l_r^2}\right), \quad \tau = t - t',\ r = 1 \ldots R, \tag{7}$$

then the covariance functions $k_{y_i,x_j}(t, t')$, $k_{x_i,x_j}(t, t')$, $k_{x_i,u_r}(t, t')$, and $k_{y_i,u_r}(t, t')$ can be solved analytically for certain output models, such as Equation (1). This enables the usage of standard GP regression techniques for predicting the values of the state $\mathbf{x}(t)$ as well as estimating the necessary covariance functions [Rasmussen and Williams 2005].

However, for most other covariance functions, difficulties reside in the evaluation of the double integral over the driving force covariances in Equation (6) that are required to build the GP prior over the target variable, $\mathbf{x}(t)$. One has to always solve for the necessary covariance functions when constructing new output models, and currently such solutions exist only for squared exponential covariance functions. In addition, integrating over matrix exponentials is also computationally expensive. A further drawback of the direct GP regression solution is that the computational complexity scales as $\mathcal{O}(D^3N^3)$, where N is the number of time instances in the observations and D is the number of data dimensions. All these reasons can impose serious restrictions on the generality of this modeling framework.

To address the issues of tractability, as well as the high computational cost associated with inference in LFMs, a sequential inference scheme based on computationally efficient techniques is presented in Hartikainen and Särkkä [2011]. In the next section, we briefly introduce this core concept of sequential inference in LFMs in general.

2.2. Sequential Inference for Latent Force Models

To remedy the problems described in the previous section, in Hartikainen and Särkkä [2010], a technique is proposed for formulating GP priors on the components $r = 1, \ldots, R$ of $\mathbf{u}(t)$ as a multivariate linear time-invariant (LTI) stochastic differential equation (SDE) model of the form

$$\frac{\mathrm{d}\mathbf{z}_r(t)}{\mathrm{d}t} = \mathbf{F}_{z,r}\,\mathbf{z}_r(t) + \mathbf{L}_{z,r}\,w_{z,r}(t), \tag{8}$$

where $\mathbf{z}_r(t) = (u_r(t)\ \frac{du_r(t)}{dt}\ \cdots\ \frac{d^{d_r-1}u_r(t)}{dt^{d_r-1}})^T$ and

$$\mathbf{F}_{z,r} = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -a_r^0 & \cdots & -a_r^{d_r-2} & -a_r^{d_r-1} \end{pmatrix}, \mathbf{L}_{z,r} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

By choosing the coefficients $a_r^0, \ldots, a_r^{d_r-1}$, the spectral density q_r of white noise process $w_{z,r}(t)$, and the dimensionality d_r of $\mathbf{z}_r(t)$ appropriately, the dynamic model on $u_r(t)$ can be chosen to correspond to a GP prior with a certain stationary covariance function.

As described in Hartikainen and Särkkä [2010], the coefficients $a_r^0, \ldots, a_r^{d_r-1}$ are found by initially taking the Fourier transform of both sides of Equation (8). The coefficients can then be expressed in terms of the spectral density of the latent force kernel, k, provided that its spectral density, $S(\omega)$, can be written as a rational function

of ω^2:

$$S(\omega) = \frac{(\text{constant})}{(\text{polynomial in } \omega^2)} \tag{9}$$

The inverse power spectrum is then approximated by a polynomial series from which the transfer function of an equivalent stable Markov process for the kernel can be inferred along with the corresponding spectral density of the white noise process. The stochastic differential equation coefficients are then calculated from the transfer function.

Here, we consider covariance functions in the Whittle-Matérn family [Hartikainen and Särkkä 2010; Rasmussen and Williams 2005]:

$$k_\nu(\tau) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}}{l} \tau \right)^\nu K_\nu \left(\frac{\sqrt{2\nu}}{l} \tau \right),$$

where l and σ^2 are the length scale and magnitude hyperparameters controlling the overall correlation scale and variability of the process, K_ν is a modified Bessel function of the second kind, and ν is a parameter controlling the smoothness of the process. In line with previous applications, we limit our view to cases in which $\nu = d_r + 1/2$ and Γ is the gamma function. This class of covariance functions is particularly useful since it contains the exponential and the squared exponential covariance functions as special cases ($\nu = 1/2$ and $\nu \to \infty$). The key property of this model class is that it has an analytic state-space representation since its spectral density $S(\omega)$ can be written as a rational function of ω^2 [Hartikainen and Särkkä 2010]. The mathematical forms for $S(\omega)$ and $k_{u_r}(\tau)$ that result from choosing small values of d_r for the Matérn class of covariance functions can be found in Hartikainen and Särkkä [2010].

The GP prior models of the form in Equation (8) can be straightforwardly combined with the output model of Equation (4) to form a joint model:

$$\frac{\mathrm{d}\mathbf{x}_a(t)}{\mathrm{d}t} = \mathbf{F}_a \mathbf{x}_a(t) + \mathbf{L}_a \, \mathbf{w}_a(t), \tag{10}$$

where we have defined an augmented state vector $\mathbf{x}_a(t) = (\mathbf{x}(t)^T \ \mathbf{z}_1(t)^T \ \cdots \ \mathbf{z}_R(t)^T)^T$, and the matrices $\mathbf{F}_a$ and $\mathbf{L}_a$ are constructed such that they operate on the augmented state appropriately.

2.3. Posterior Inference and Predictions

The LTI SDE model in Equation (10) has the desirable property that it can be analytically converted to a discrete-time dynamic model as

$$\mathbf{x}_k = \mathbf{A}(\Delta t_k)\mathbf{x}_{k-1} + \mathbf{q}_{k-1}, \ \mathbf{q}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(\Delta t_k)), \tag{11}$$

where the transition and process noise matrices can be solved on the time instances $\mathcal{T} = \{t_k\}_{k=1}^N$ as

$$\begin{aligned} &\mathbf{A}(\Delta t_k) = \Phi_a(\Delta t_k), \ \Delta t_k = t_k - t_{k-1}, \ \Phi_a(\tau) = \exp(\mathbf{F}_a \, \tau), \\ &\mathbf{Q}(\Delta t_k) = \int_0^{\Delta t_k} \Phi_a(\Delta t_k - \tau) \, \mathbf{L}_a \, \mathbf{Q}_c \, \mathbf{L}_a^T \, \Phi_a(\Delta t_k - \tau)^T \, \mathrm{d}\tau, \end{aligned} \tag{12}$$

where $\mathbf{Q}_c$ is the spectral density of white noise process $\mathbf{w}_a(t)$ in Equation (10).

So far we have not discussed how the output process is observed. The standard approach is to use the linear Gaussian model:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{r}_k, \ \mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k), \tag{13}$$

where the matrix $\mathbf{H}_k$ collects the observed components from the state vector. Now the filtered posterior distribution of the state $p(\mathbf{x}_k|\mathbf{y}_{1:k}, \theta) = \mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$ on the selected time points can be solved exactly with the classical Kalman filter (KF) and the smoothing distribution $p(\mathbf{x}_k|\mathbf{y}_{1:N}, \theta) = \mathcal{N}(\tilde{\mathbf{m}}_k, \tilde{\mathbf{P}}_k)$ with the Rauch-Tung-Striebel smoother (RTSS) (see Simon [2006]). Both the KF and RTSS scale as $\mathcal{O}(D^3N)$, where D is the dimensionality of $\mathbf{x}$ and N the number of time points. The estimation should be started from the Gaussian prior $p(\mathbf{x}_0|\theta) = \mathcal{N}(\mathbf{m}^0, \mathbf{P}^0)$, where it is reasonable to set the covariance matrix to be block diagonal of the form $\mathbf{P}^0 = \text{blkdiag}(\mathbf{P}^0_x, \mathbf{P}^0_{u_1}, \ldots, \mathbf{P}^0_{u_R})$, where $\mathbf{P}^0_x$ is the joint prior covariance for the nonaugmented output process $\mathbf{x}(t)$ chosen according to a priori knowledge. The blocks $\mathbf{P}^0_{u_r}$ for the R latent forces can be set to stationary covariances by numerically solving the algebraic Riccati equations as [Grewal and Andrews 2001]

$$\frac{\mathrm{d}\mathbf{P}_{u_r}}{\mathrm{d}t} = \mathbf{F}_{z,r}\,\mathbf{P}_{u_r} + \mathbf{P}_{u_r}\,\mathbf{F}^T_{z,r} + \mathbf{L}_{z,r}\,q_r\,\mathbf{L}^T_{z,r} = \mathbf{0}. \tag{14}$$

In this context, the hyperparameters of the covariance function are learned by optimizing the marginal likelihood of the observations as

$$p(\mathbf{y}|\theta) = \prod_{k=1}^{N} p(\mathbf{y}_k|\mathbf{y}_{k-1}, \theta). \tag{15}$$

3. LATENT FORCE THERMAL MODELS

Fundamentally, LFM-TM is derived from the *Ti* model described in Bacher and Madsen [2011], to which an extra term, which we refer to as the learned residual, is added. This term can model any a priori unknown additional dynamics. This makes it possible to use the LFM-TM in situations where human-induced activity can suddenly change the thermal dynamics of a building, as is often the case in domestic buildings or residential homes. LFM-TM, in essence, is a gray-box model, which is augmented with a residual that can accommodate the effect of the latent forces seen in real data.

To illustrate the LFM-TM, we start by specifying the following thermal model of a home represented as a linear first-order DE, which is adapted from *Ti* (refer to Appendix A) in Bacher and Madsen [2011] as

$$\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} = \frac{1}{C_i}\eta_h(t) + \frac{1}{C_i R_{ia}}\big(T_{ext}(t) - T_{int}(t)\big) + \sigma_i d\omega_i + \epsilon(t). \tag{16}$$

Equation (16), in essence, is a stochastic linear first-order differential equation, where we have introduced an additional term $\epsilon(\mathbf{t})$. This is an unknown time-varying residual function that is introduced to capture all factors in the built environment, which influence the thermal dynamics but are not accommodated within *Ti*. Such phenomena include lags in a heating system or the effect of an additional heater. In particular, Equation (16) can be viewed as an LFM with an unknown time-varying residual latent force $\epsilon(\mathbf{t})$, and not random noise.

LFM-TM is also an extension of the *Ti* gray-box model that avoids the need to introduce additional states such as envelopes and sensors. In existing gray-box models, these elements are introduced mainly to derive a better fit to the data obtained from buildings, but they yield models of increased complexity [Andersen et al. 2000]. We will demonstrate that by eliminating additional states and replacing them by one single time-varying driving force, LFM-TM performs better than previous models at fitting the data observed from residential buildings.

Now as per the LFM machinery described in Álvarez et al. [2009], we assume that the residual is a GP having a stationary covariance function as per $\epsilon(\mathbf{t}) \sim \mathcal{GP}(0, k(\mathbf{t}, \mathbf{t}'))$, where $k(\mathbf{t}, \mathbf{t}')$ can be any stationary covariance function. For the GP prior on $\epsilon(\mathbf{t})$, we

use a Matérn covariance function described in Rasmussen and Williams [2005]. We make this choice because this covariance function does not impose any unrealistic smoothness constraints and has a more flexible parameterization compared to other covariance functions such as the squared-exponential covariance function [Rasmussen and Williams 2005]. As demonstrated later, this confers a definite advantage in our application, where $\epsilon(\mathbf{t})$ can show precipitous changes in structure, which a GP with a squared-exponential covariance function will not accurately model. Furthermore, by altering the parameterization of the Matérn covariance function (by varying the ν parameter of a Matérn covariance function), it is possible to model residuals that range from having white noise properties ($\nu = 1/2$) to those that are correlated across long time scales ($\nu = 5/2$). Finally, in Equation (16), $\mathbf{T}_{ext}$ is a vector of known inputs representing the external temperature, which needs to be accommodated with the process model.

Once LFM-TM has been specified, as per the LFM machinery, we derive an augmented model for the LFM-TM given as

$$\frac{\mathrm{d}\mathbf{x}_a(t)}{\mathrm{d}t} = \mathbf{F}_a\mathbf{x}_a(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}_a\,\mathbf{w}_a(t), \tag{17}$$

where the augmented state vector is $\mathbf{x}_a(t) = [T_{int}(t), \mathbf{z}(t)^T)]^T$. Here, $\mathbf{z}(t)$ incorporates the residual within the process model, $\mathbf{u}(t)$ represents the input vector at t, and $\mathbf{G}$ is the input coupling matrix.

In Equation (16), $\mathbf{u}(t) = [T_{ext}(t)]$ and $\mathbf{G} = [\frac{1}{C_i R_{ia}}]$. The matrices $\mathbf{F}_a$ and $\mathbf{L}_a$ for the LFM-TM are

$$\mathbf{F}_a = \begin{pmatrix} -\frac{1}{C_i} & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -a^0 & -a^1 & -a^2 \end{pmatrix}, \mathbf{L}_a = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Here, $d_r = 3$; that is, the dimensionality of $\mathbf{z}_r(t) = 3$. Equation (16) can be extended to accommodate any known additional inputs. For instance, one might incorporate solar radiation in Equation (16). In this case, Equation (16) becomes

$$\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} = \frac{1}{C_i}\eta_h(t) + \frac{1}{C_i R_{ia}}\left(T_{ext}(t) - T_{int}(t)\right) + \frac{A_w}{C_i}\eta_s(t) + \epsilon(t) + \sigma_i d\omega_i, \tag{18}$$

where $\mathbf{u}(t) = [T_{ext}(t), \eta_s]^T$ and $\mathbf{G} = [\frac{1}{C_i R_{ia}}, \frac{A_w}{C_i}]^T$.

Next, we describe how inference is done for Equation (17) as per the sequential inference scheme explained previously. In doing so, Equation (17) is analytically converted to a discrete-time dynamic model as per Simon [2006]:

$$\mathbf{x}_a(k) = \mathbf{A}(\Delta t_k)\mathbf{x}_a(k-1) + \mathbf{B}(\Delta t_k)\mathbf{u}(k-1) + \mathbf{q}_{k-1},\ \mathbf{q}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(\Delta t_k)), \tag{19}$$

where the corresponding transition matrix $\mathbf{A}$ and the input coupling matrix $\mathbf{B}$ can be solved on the discrete-time instances $\mathcal{T} = \{t_k\}_{k=1}^N$ as

$$\mathbf{A}(\Delta t_k) = \exp(\mathbf{F}_a\,\tau), \mathbf{B}(\Delta t_k) = \mathbf{A}[\mathbf{I} - \exp(-\mathbf{F}_a\Delta t_k)]\mathbf{F}_a^{-1}\mathbf{G}.$$

Furthermore, the process noise matrices $\mathbf{Q}$ is

$$\mathbf{Q}(\Delta t_k) = \int_0^{\Delta t_k} \Phi_a(\Delta t_k - \tau)\,\mathbf{L}_a\,\mathbf{Q}_c\,\mathbf{L}_a^T\,\Phi_a(\Delta t_k - \tau)^T\,\mathrm{d}\tau, \tag{20}$$

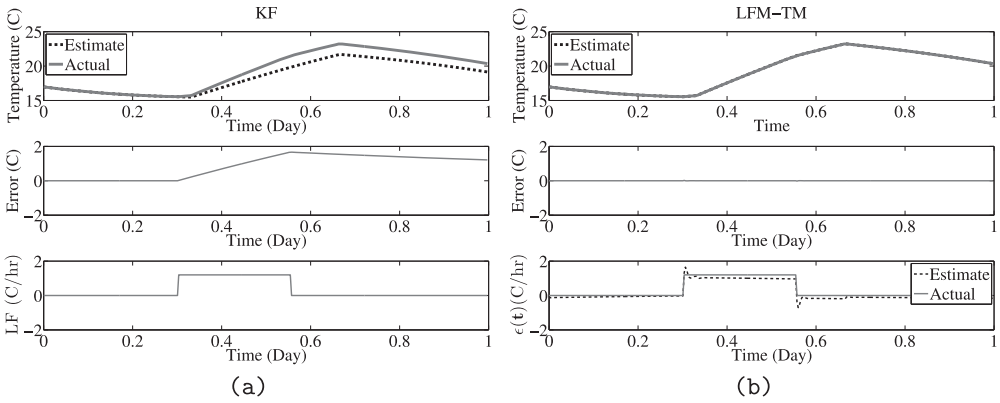

Fig. 1. Simulated example showing the performance of a KF (left) and LFM-TM (right) at modeling an extra heat source.

where $\mathbf{Q}_c$ is the spectral density of the white noise process $\mathbf{w}_a(t)$ in Equation (17). The observation model is given by

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{r}_k,\ \mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k), \tag{21}$$

where the matrix $\mathbf{H}_k$ collects the observed components from the state vector.

In all applications of the LFM-TM described subsequently in this article, the time evolution of the internal temperature during parameter learning and for the purposes of generating predictions at selected time points is solved analytically with the KF. First of all, maximum likelihood is used to learn the thermal parameters θ, as well as the GP's hyperparameters, $\{\sigma,\ l\}$. The internal temperature is estimated as follows. First, the KF is initialized with a best-guess internal temperature value. The initial covariance as explained earlier is block diagonal with a diagonal matrix over the temperature components including the solution to the appropriate Riccati equation for the Matérn model of the residual as presented in Hartikainen and Särkkä [2010]. The KF is then used to predict the next state $\mathbf{x}_a$ at the end of the current time interval t as per Equation (19). The prediction step is conditioned on the value of the predicted heating action at the previous time step. The KF predict and update steps are then repeated, $\forall t \in N$. The culmination of these steps also results in the residual being learned from the data. In later sections, we will see that the residual in effect models the innovation or the error at the end of each KF predict step. We illustrate the core idea behind LFM-TM with a simple example. Figure 1(a) shows a simulated example of running a simple gray-box model, where a latent force (LF) in the form of a piecewise constant heat source is introduced to represent an additional heater in a home between Time = 0.3 and Time = 0.6. This latent force is shown in the lowest subplot of Figure 1(a). The estimated internal temperature from this model (based on inference using a KF) is shown in the top subplot, along with the error, labeled as Error (C), in the middle subplot of Figure 1(a). This error is in effect the KF's innovation and can also be construed as the model's residual. We see that the extra heat source induces a large error in the estimation of the internal temperature as shown in the top subplot. The KF is unable to fully accommodate this additional dynamic in its process model and induces a pronounced error.

Next, an LFM-TM is applied to the same data and it is able to accurately model this constant heat source as shown in Figure 1(b). The learned residual, $\epsilon(\mathbf{t})$, in the lowest subplot of Figure 1(b) accurately models the constant heat source. As a result, the error as shown in the middle subplot of Figure 1(b) is close to zero. $\epsilon(\mathbf{t})$ is able to capture the

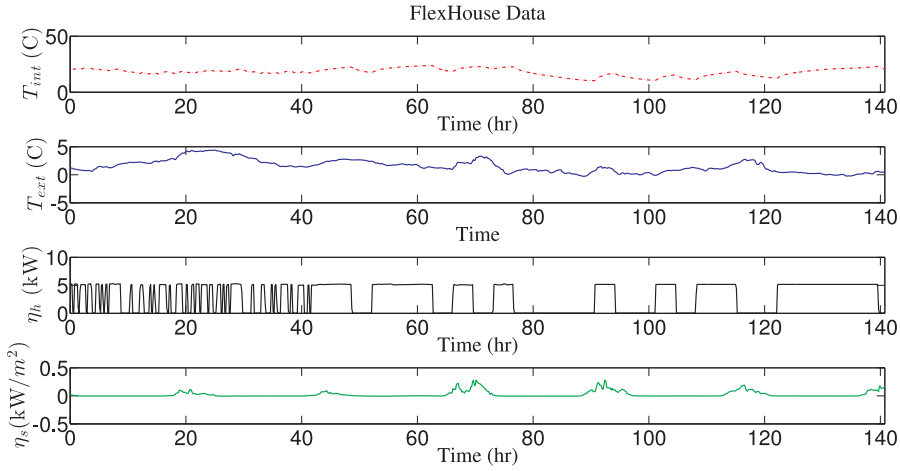

Fig. 2. FlexHouse data.

inherent (piecewise constant) structure of the latent force. It is also correlated across all the time lags over which the additional heat is introduced. To generate Figure 1(b), we use the sequential inference scheme discussed earlier.

4. DATA COLLECTION

We demonstrate the efficacy of the LFM-TM approach at accurately modeling the thermal dynamics of buildings. We evaluate our approach on two different datasets:

(1) Our first dataset is composed of data collected from an experimental energy system, Syslab, that consists of a specially designed building referred to as FlexHouse. Data from FlexHouse has been previously used for thermal modeling purposes (see Bacher and Madsen [2011] for more details). Although this data was not publicly made available, we digitized the data plots and extracted the time series from the digitized plots. Figure 2 plots this data, which is composed of the following time series:
   —A vector of observations of the internal temperature, T_{int} (in $^\circ$C)
   —Vector of logged external air temperature readings, T_{ext} (in $^\circ$C)
   —Total heat input from the electrical heaters in the building, η_h (in kW)
   —Total solar irradiance measured on-site, η_s (in kW per m^2)

(2) Our second dataset consists of data collected from a bespoke installation of heating equipment within multiple residential buildings undertaken in Southampton, United Kingdom. These homes represent standard UK 1930s building stock, which are heated by gas boilers and water radiators. We also collected historic as well as forecasts of external temperature at the location of these homes from an external website.[2] A GP regression model is used to interpolate over missing data points in the external temperature time series. The following time series are created and subsequently used in the development of thermal models of the homes as described in Section 6.
   —Observations of the internal temperature (in $^\circ$C) are used in the measurement model of *Ti*, *TiTh*, and *TiTeTh* and the LFM-TM models.
   —Logged external temperature readings are used as an input (T_{ext}) to every thermal model.

[2]http://www.wunderground.com.

—Logged data on boiler activity in each home yields $\eta_{on}(t) \in [0, 1], \forall t \in N$—the proportion of each 10-minute thermostat cycle during which the boiler on.

All data is down-sampled to a 10-minute resolution to be in sync with a boiler cycle.

5. EXPERIMENTS ON THE FLEXHOUSE DATA

To evaluate the efficacy and the accuracy of LFM-TM, we first apply it to model the FlexHouse data. For this dataset, we use the LFM-TM described in Equation (18), as solar irradiance is logged and subsequently used as an input. We also reimplemented their previously published *Ti*, *TiTh*, *TiTeTh,* and *TiTeThTs* gray-box models for comparison [Bacher and Madsen 2011]. Our aim is to remain consistent with the experiments presented in this article, and hence, we present a like-for-like comparison between their work and ours.

Three different metrics are used to compare the LFM-TM with the aforementioned gray-box models in this context:

(1) Log-likelihood (logL): The log-likelihood assesses how well a model that is parameterized by θ fits the observed data. Since a Kalman filter is used for the inference of all models, we assume a Gaussian likelihood in all cases. Then the likelihood is calculated using the prediction $\hat{y}(t|t-1)$ as per

$$\text{logL} = \log \prod_{t=2}^{N} \frac{1}{(2\pi)^{1/2}(\sigma_{y(t|t-1)} + \sigma_r)} e^{-\frac{(y(t)-\hat{y}(t|t-1))^2}{2(\sigma_{y(t|t-1)}+\sigma_r)^2}}, \tag{22}$$

where $\sigma_{y(t|t-1)}$ is the variance of $\hat{y}(t|t-1)$ and σ_r is the observation noise variance.

(2) One step-ahead prediction error (pred-rmse): We measure the one-step-ahead prediction error between the predicted and observed internal temperature given as

$$\text{pred-rmse} = \sqrt{\frac{\sum_{t=2}^{N}(y(t) - \hat{y}(t|t-1))^2}{N-2}}. \tag{23}$$

pred-rmse estimates the distance between the mean predictions and the observations. The closer the pred-rmse is to zero, the better is the model.

(3) Normalized prediction error (npe): We use a third metric to evaluate how much of the variance in the observations is explained by each thermal model. We refer to this metric as the normalized prediction error or npe, which is given as per

$$\text{npe} = \frac{\text{pred-rmse}^2}{\frac{\sum_{t=2}^{N}(y(t)-\bar{y}(t))^2}{N-2}}, \tag{24}$$

where pred-rmse is calculated as per Equation (23). In effect, npe explains how much of the variability in the observations our model captures. By using this metric, it is possible assess how uncertain a model's predictions are. The lower the value of npe is, the higher the confidence one can have in a model's predictions; models for which npe > 1 have poor predictive ability, whereas models with npe $\ll 1$ are excellent.

We now turn to the results in Table I. As can be seen, LFM-TM yields the best results on all metrics. It yields the highest log-likelihood and the least one-step prediction error and explains a far greater percentage of the variability in the future observations in comparison with all the other gray-box models. As the complexity of the gray-box models is increased, the model fit improves as well. The best-performing gray-box model is *TiThTeTs*. LFM-TM improves upon this model on all metrics. The logL is

Table I. Comparison of LFM-TM with Existing Gray-Box Models Applied to FlexHouse Data

Model	logL	pred-rmse	npe
Ti	544	0.1410	0.0017
TiTh	1301	0.1335	0.0016
TiThTe	1480	0.1195	0.0012
TiThTeTs	1541	0.1143	0.00099
LFM-TM	**1802**	**0.0852**	**0.00062**

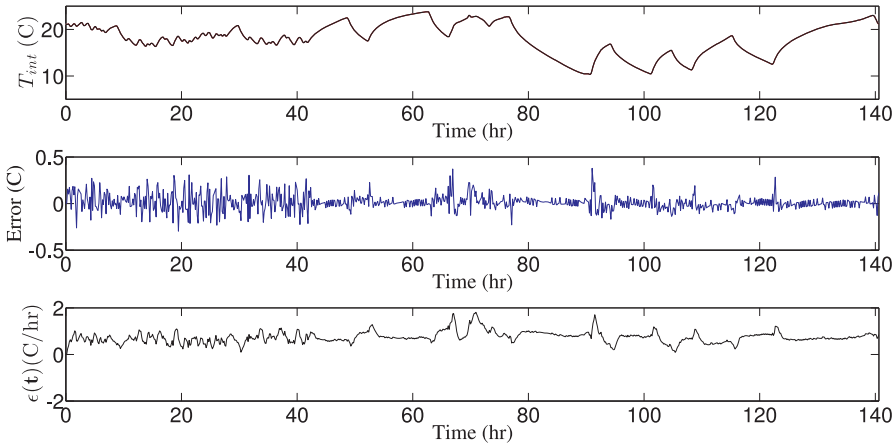

Fig. 3. Estimate of the internal temperature, the estimation error, and the learned residual for LFM-TM.

14.5% greater; the prediction error, pred-rmse, is 25% lower; and npe is 37% lower. Figure 3 plots the estimate of the internal temperature (T_i) as per LFM-TM, along with the one-step-ahead prediction error at each time step. Furthermore, the plot also shows the time-varying residual $\epsilon(\mathbf{t})$ that is inferred by LFM-TM. Effectively, the addition of this GP process model to *Ti* results in the LFM-TM being able to model the residual process explicitly, unlike the *Ti* model without the GP process, which attempts to accommodate the residual process as variance in the process noise. To explain this further, note that there is a small spike every time the heater is switched on because of a lag between the heater coming on and the temperature rising. LFM-TM accurately models this lag. This is seen as corresponding spikes in the learned residual. Visually the learned residual appears to be smooth, which leads us to conclude that the LFM-TM is able to distinguish between modeling essential structure from modeling noise.

In comparison, Figure 4 plots the outputs of the *Ti* and *TiTh* models without the GP process model. Also, in Figure 5, the outputs of the *TiTeTh* and *TiTeThTs* models are shown. Here, additional states representing the building envelope and the sensor lead to further improvements in the model fit, with corresponding reductions in pred-rmse. Specifically, it can be noted that it is the introduction of the building envelope that results in the most significant reduction in pred-rmse. Thus, a building's envelope significantly affects the thermal dynamics of the internal temperature through conduction and convection effects. Modeling the temperature sensor leads to a small further improvement.

However, the LFM-TM not only fits the data more accurately in comparison with the competing gray-box models but also requires fewer parameters and little knowledge of the physical elements of a building, which have an effect the thermal dynamics. The residual $\epsilon(\mathbf{t})$ captures the effects of the lags in the heating system, the envelope, and the sensor that are explicitly represented in the comparable gray-box models. Table II

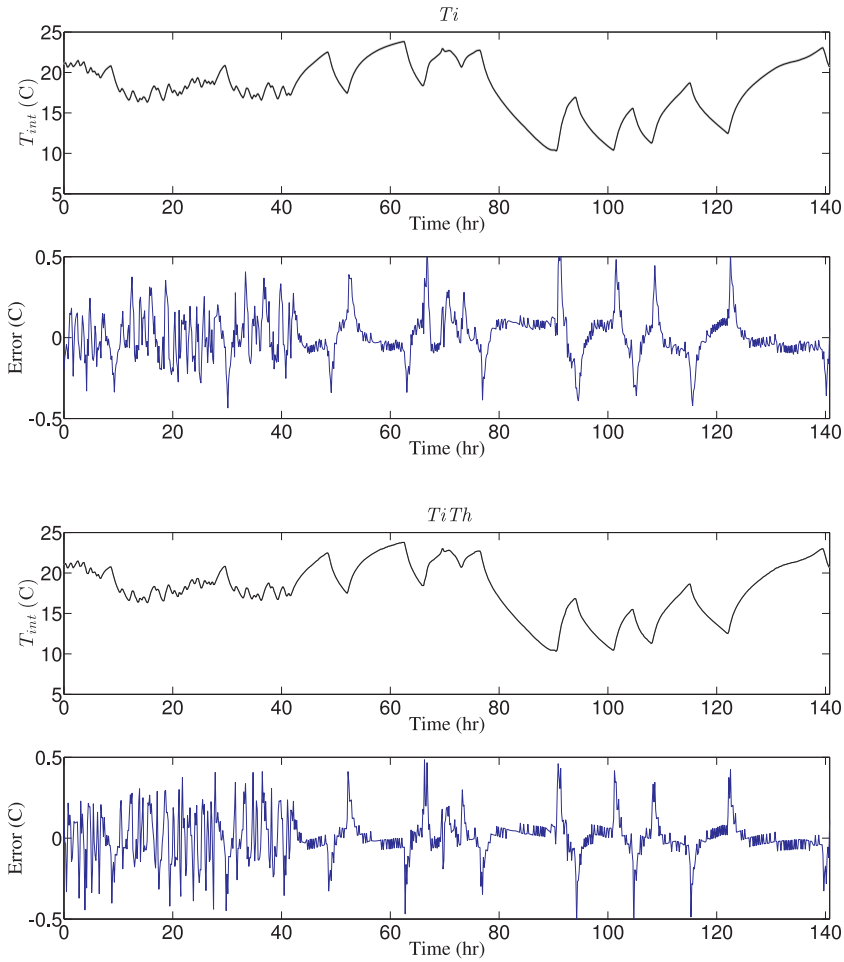

Fig. 4. Estimate of the internal temperature and the resultant estimation error for *Ti* (top) and *TiTh* (bottom) models.

presents the total number of parameters in the various models. From this, it is possible to conclude that the LFM-TM, in spite of having a relatively simple structure and fewer parameters, is able to accurately model the data.

6. EXPERIMENTS ON RESIDENTIAL DATA

In this section, we present the results of applying the LFM-TM approach to thermal modeling of real homes. We first describe how the LFM-TM is applied to the data collected from real homes described in Section 4. Next, we compare the LFM-TM with gray-box models. Specifically, for each home, we develop bespoke of *Ti*, *TiTh,* and *TiTeTh* gray-box models and then demonstrate that the LFM-TM has improved efficacy as compared to these models.

In Section 4, we explained how the logged data on the boiler activity in each home yields $\eta_{on}(t) \in [0, 1]$—the proportion of each 10-minute thermostat cycle during which the boiler is on. This is multiplied by a parameter, r_h (the total heater power output in kW), to yield $\eta_h(t), t \in N$ for each thermal model. r_h for a home is unknown a priori but is learned from the data.

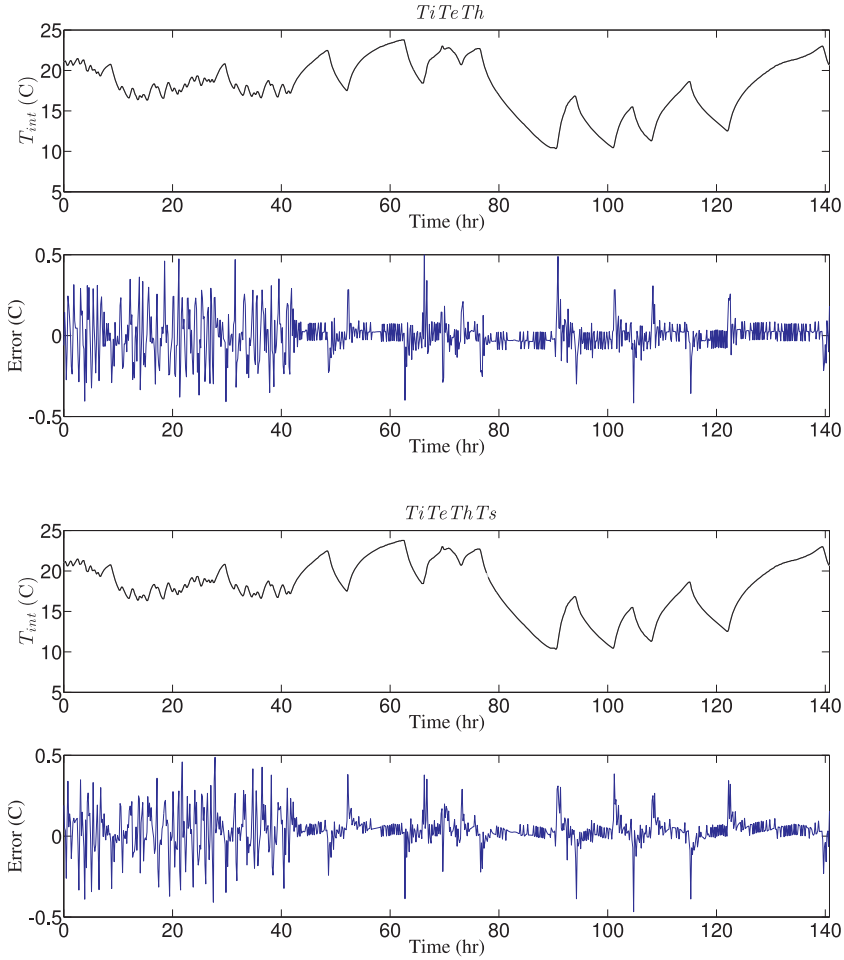

Fig. 5. Estimate of the internal temperature and the resultant estimation error for *TiTeTh* (top) and *TiTeThTs* (bottom) models.

Table II. Number of Parameters in the Thermal Models

Model	No. Parameters
Ti	5
TiTh	9
TiThTe	14
TiThTeTs	18
LFM-TM	7

We did not log solar irradiance at the site where the homes are located since most of the data collection was done over winter months. Hence, in all the thermal models considered henceforth, solar irradiance is not used as an input. More specifically, we use *Ti*, *TiTh,* and *TiTeTh*. We do not explicitly model the effect of the temperature sensor in this setting either, since all internal temperature observations are directly logged by the thermostat. In the absence of solar irradiance, the LFM-TM in Equation (16) is chosen for all experiments. In Equation (16), $\epsilon(\mathbf{t})$ is drawn from a GP

Table III. Comparison of LFM-TM with Existing Gray-Box Models

	Model	Mean logL	Mean pred-rmse	Mean npe
Home 1	Ti	829	0.8534	0.9111
	TiTh	1049	1.2215	1.0428
	TiThTe	1149	1.0938	0.5424
	LFM-TM	**1584**	**0.7292**	**0.5034**
Home 2	Ti	703	1.6354	0.76
	TiTh	831	1.7442	0.8255
	TiThTe	891	1.7985	0.7315
	LFM-TM	**1340**	**1.4787**	**0.6361**

prior, that is, $\epsilon(t) \sim \mathcal{GP}(0, k(\mathbf{t}, \mathbf{t}'))$, where $k(\mathbf{t}, \mathbf{t}')$. Since $\epsilon(\mathbf{t})$ has a zero-mean covariance function, any additional dynamics in the data are fully accounted for by $k(\mathbf{t}, \mathbf{t}')$.

Three different metrics are used to compare our LFM-TM with the existing gray-box models (*Ti*, *TiTh,* and *TiTeTh*). In addition to the logL described in the previous section, we first consider the pred-rmse, which is slightly reformed in this setting to deal with day-ahead predictions and is given as

$$\text{pred-rmse} = \sqrt{\frac{\sum_{t=2}^{N}(y(t) - \hat{y}(t))^2}{N-2}}. \tag{25}$$

Equation (25) calculates the error between the predicted and actual internal temperatures for the entire day ahead. This is estimated by running the thermal model forward over the day-ahead period. Unlike the experiments on the FlexHouse data, we choose to perform day-ahead predictions in this setting, in order to thoroughly evaluate how well each model generalizes to unseen inputs over a future time horizon. More specifically, day ahead is chosen as the horizon in accordance with previous work, where the predictive ability of thermal models has been assessed on the basis of their ability to generate accurate predictions of internal temperatures for the entire day ahead [Yu et al. 2012; Rogers et al. 2011]. Finally, we use the normalized prediction error or npe, which is now calculated as

$$\text{npe} = \frac{\text{pred-rmse}^2}{\frac{\sum_{t=2}^{N}(y(t) - \bar{y}(t))^2}{N-2}}, \tag{26}$$

where pred-rmse is calculated as per Equation (25). For all experiments on the data from homes, we follow a standard sliding-window approach. Each model is trained on 4 days of training data. This is followed by day-ahead predictions generated for the fifth day. The window is then moved forward by a day, and day-ahead predictions for the sixth day are estimated. At the start, all parameters are learned using maximum likelihood estimation. In Section 3, we described how the estimation step is undertaken using a KF. The culmination of the training period results in a model of the residual being learned. The log-likelihood of each model is subsequently recorded.

Next, the thermal model is run forward 1 day to generate day-ahead predictions of the internal temperature. We generate temperature predictions over each 10-minute interval corresponding to a boiler cycle resulting in 144 time instants in total for each day. Day-ahead predictions are performed using the KF. The KF predicts the next state $\mathbf{x}_a(t)$ conditioned on $\mathbf{x}_a(t-1)$ and the inputs $T_{ext}(t-1)$ and $\eta_h(t-1)$. This process is iterated $t = 2$ to 144 to generate predictions for the entire day ahead. At the end of the prediction step, we estimate pred-rmse and npe to assess the accuracy of the day-ahead predictions and therefore quantify the predictive ability of each model. We generate day-ahead predictions for 14 days in total for each home. Table III presents the mean

Table IV. The Total Number of Parameters in Each Model

Model	No. Parameters
Ti	5
TiTh	9
TiThTe	13
LFM-TM	6

Table V. Mean CPU Time in Seconds Required to Learn Parameters in Each Model for Varying Number of Days of Training Data

	No. Days Training Data			
Model	1	2	3	4
Ti	3.9	8.1	11.7	19.2
TiTh	23.7	55.9	89.6	82.9
TiThTe	67.3	303.4	370.7	463.1
LFM-TM	10.9	22.9	30	35.4

log-likelihood, logL; the mean prediction error, pred-rmse; and the mean normalized prediction error, npe, obtained from two different homes. These results are averages over all 14 days.

We observe that as the complexity of the gray-box models is increased by adding extra states, the model fit shows corresponding improvements. In Home 1, the best-performing gray-box model is *TiThTe*. Our LFM-TM outperforms *TiThTe* on all three metrics. It yields higher log-likelihoods and lower prediction errors and explains more of the variability in the future (day-ahead) internal temperature observations in comparison with the gray-box models. In the case of Home 1, the improvements in logL, pred-rmse, and npe are 38%, 14.5%, and 7%, respectively, over *TiThTe*.

In the case of Home 2, among all competing gray-box models, *Ti* yields the best mean pred-rmse estimate, while *TiThTe* yields the best mean logL and npe estimates. LFM-TM outperforms both models: logL is 33.5% greater, pred-rmse is 9.6% lower, and npe is 13% lower in comparison with the corresponding gray-box model that yields the best estimate on each metric. These improvements are seen in spite of LFM-TM having significantly fewer parameters than the more complex *TiTh* and *TiThTe* models as shown in Table IV. The time taken to learn the parameters of the LFM-TM is significantly less than the more complex grey-box models as shown in Table V. Figure 6 is a plot from applying the LFM-TM to data from one of the homes. The top subplot shows the estimate of the internal temperature during the training period over the first 4 days, followed by the day-ahead predictions of the internal temperature on the fifth day. The vertical line divides the training and prediction periods. In the subplot in the middle, the prediction error over both training and prediction periods is plotted. The lowest subplot displays the time-varying residual $\epsilon(\mathbf{t})$ that is learned by the LFM-TM during the training period. Also shown is $\epsilon(\mathbf{t})$ over the prediction period. Along similar lines to what was observed in the experiments in Section 5, $\epsilon(\mathbf{t})$ accurately models the error during the training periods. In effect, the residual models the error or the innovation induced by the KF at each time step. This is equivalent to modeling the errors that are induced by the inadequacies in the *Ti* when residual dynamics are present.

Figure 6 also shows that the predicted residual in the lowest subplot is smoother than the residual over the training phase, with the mean value reaching zero in Figure 6 within a few time steps. This is because during the prediction phase, no observations of internal temperature are available to the KF and the update step at each time iteration is effectively skipped. This results in just the initial residual values getting propagated through the KF process model over the first few steps in the prediction

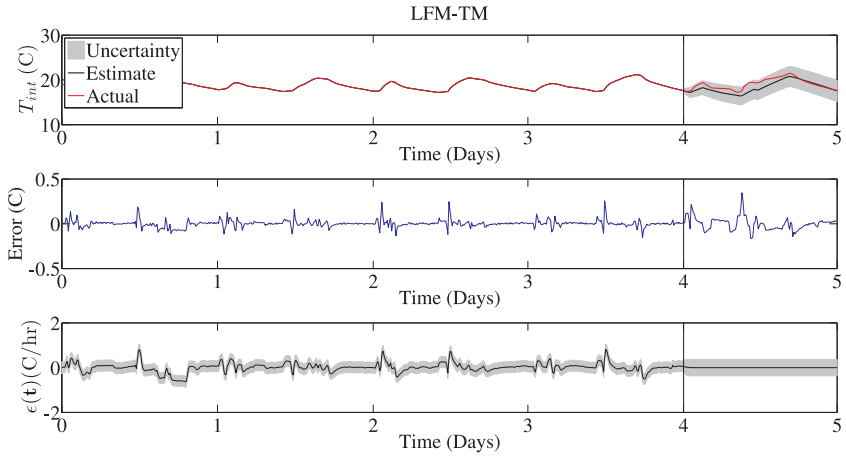

Fig. 6. Estimate of the internal temperature, the estimation error, and the learned residual for LFM-TM.

period. The reason that the residual reverts to zero in Figure 6 is because the length scale of the GP's (Matérn) covariance function learned during the training period is small. Consequently, when no observations are available, the correlations between successive states of the residual during the prediction phase quickly decay to zero.

From these empirical tests, we observed that overall, introducing a residual has a significant impact on the accuracy of the day-ahead predictions. In Table III, the *Ti* model, which is essentially the LFM-TM without a residual, generates significantly poorer predictions in comparison to the LFM-TM. The introduction of the residual in LFM-TM leads to better process models being learned over the training phase. Subsequently, when the same process model is used to generate day-ahead predictions, it yields more accurate predictions. In comparison, Figures 7 and 8 plot the output of all gray-box models. As shown in Figure 7, the addition of the heater state in *TiTh* leads to an improvement in the model fit and reduction in the errors over *Ti*. Furthermore, as shown in Figure 8, the addition of a further state representing the building envelope yields even greater improvements in model fit. This is similar to the trend observed in the previous section. However, despite these improvements in the physical thermal model, they are unable to match the predictive accuracy of LFM-TM. As evident in Figures 7 and 8, there are large deviations between the predictions of the gray-box models and the observed internal temperature. Furthermore, the uncertainty in the predictions grows over time, with both *TiTh* and *TiTeTh* inducing large uncertainty in predictions.

Finally, our experiments in the home setting also show that the performance of all models at generating one-step (10-minute)-ahead predictions are comparable and the differences are not statistically significant. Consequently, no definitive conclusion can be drawn as to which model is better, in this setting, based on simply evaluating their predictive performance over such a short time window. In contrast, by comparing the models over a longer day-ahead window, we are clearly able to highlight the significant improvements yielded by LFM-TM.

Table IV presents the total number of parameters in all the models that are tested. As can be seen, LFM-TM needs fewer parameters than *TiTh* and *TiTeTh*. The extra parameter in LFM-TM, in comparison with the *Ti* model, is due to the introduction of the hyperparameters in the Matérn covariance function, as described in Section 2.1. Table IV presents the results of assessing the impact of the number of parameters in each model on the time it takes to train each model. The results presented here are the mean cpu time in seconds that is required to train each model for varying number

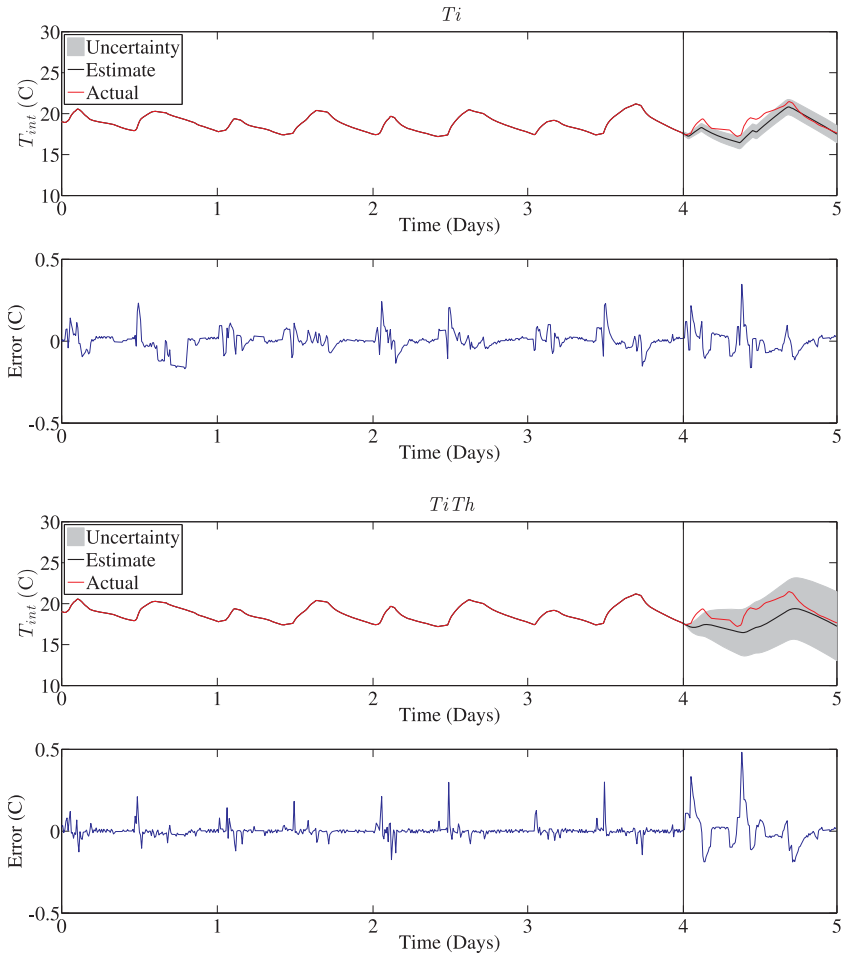

Fig. 7. Estimate of the internal temperature and the estimation error for *Ti* and *TiTh* models.

of days of training data. The results are averages over 14 days. For estimating the cpu times, we implemented each model in MATLAB. We used maximum likelihood estimation using conjugate gradients for learning the parameters. The configuration of the learning algorithm was the same for all runs.

It is clearly evident from the results presented in Table IV and Table IV that as the complexity of the gray-box models increases, the number of parameters correspondingly increases and more time is required to learn the parameters. In a model such as TiThTe, learning parameters takes longer than 7 minutes. This is because adding complexity by introducing additional parameters results in more degrees of freedom in a model, which results in parameter estimation taking longer to converge. It is important to emphasize that although 7 minutes may not be prohibitively long, in deployment, this computation is likely to be done in the cloud for multiple homes.[3] Hence, the LFM-TM will yield a 10-fold reduction in the cloud computation used in comparison with TiTh and TiThTe, which across multiple homes is significant. Finally, the results also show that LFM-TM's performance is a tradeoff between the Ti model, whose parameters can

[3]For instance, all computation in the Nest thermostat is done in the cloud.

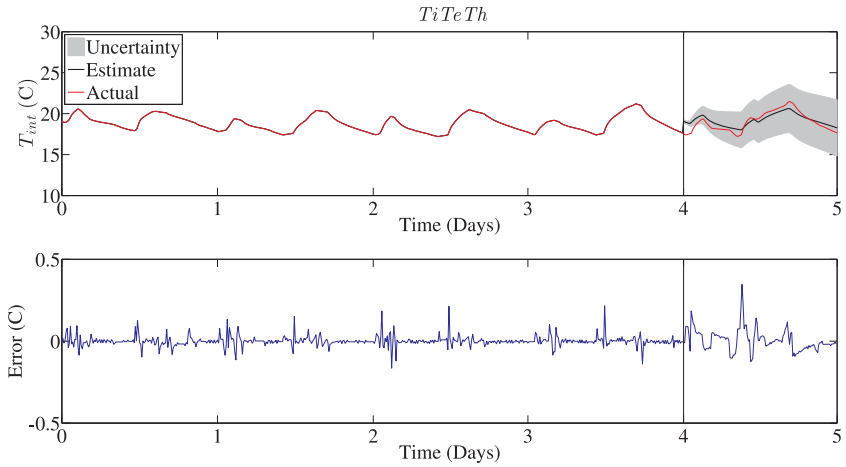

Fig. 8. Estimate of the internal temperature and the estimation error for *TiTeTh* model.

be learned quickly, and the more complex TiTh and TiThTe models, whose parameters take longer to run.

From these experiments, it is possible to conclude that although the residual $\epsilon(\mathbf{t})$ is not defined on the basis of the actual physical elements of a home, it is in effect capturing the effect of these elements, such as the structural envelope and lags in the heating system, which need to be explicitly represented in the process models of existing gray-box models. Consequently, the LFM-TM, in spite of having a relatively simple structure and fewer parameters than the more complex gray-box models, is accurate and improves on their predictive performance. This makes it an attractive model to implement in practice. Further speedups can be achieved by implementing it in compiled code.

To conclude, our experiments have enabled us to systematically evaluate the performance of existing gray-box models at modeling data collected from real homes. In previous work, such evaluation either was limited to data from a single residential building or used simulated data only [Mozer et al. 1996; Yu et al. 2012; Rogers et al. 2011].

7. CONCLUSIONS

We propose a novel model for the thermal dynamics of buildings, which we refer to as the latent force thermal model, or LFM-TM. We showed how the LFM-TM is an adaptation of an existing *Ti* gray-box model that is augmented with a time-varying residual, $\epsilon(\mathbf{t})$, which is modeled as a Gaussian process. We thoroughly evaluated the ability of the LFM-TM to produce accurate thermal models based on data collected from two different settings: (1) data from FlexHouse and (2) data collected from homes located on our university campus, which we specifically instrumented to collect heating data. The first setting represents an experimental house without any occupants. Thus, only the physics of the building is of salience in this setting. In contrast, the homes in the second setting are inhabited by occupants. This enables us to evaluate how well LFM-TM performs at modeling the thermal dynamics when human-induced dynamics interplay with the physics of a residential building.

In both these settings, we showed that our LFM-TM outperforms all existing gray-box models in terms of fitting observed data, generating accurate day-ahead predictions and explaining more of the variability in the future observations. Furthermore, we showed

how LFM-TM yields improvements in spite of having a simpler structure compared to the more complex gray-box models, as the residual is able to accommodate the effect of various physical elements that are explicitly modeled in gray-box models.

From the evidence presented in this article, it is possible to conclude that the key advantage of the LFM-TM is that it can be used of -the shelf to model the thermal dynamics of a building without requiring any detailed understanding of the structure of a building. A possible limitation of LFM-TM is that unlike existing gray-box models, where the physical elements within a building are explicitly represented, LFM-TM's residual soaks up the effect of these physical elements, leading to a slight loss of interpretability. However, in the development of intelligent controllers for HVAC systems of buildings, having accurate internal temperature predictions from a thermal model is more important than having a complete understanding of all the physical elements within a building or an accurate specification of the physical parameters [Privara et al. 2013]. This enables an intelligent controller to formulate a more efficient heating plan that can potentially save money, as well as reduce carbon emissions, without impacting the comfort of the occupants of a building.

As a next step, we intend to embed LFM-TM within an intelligent-model-based predictive controller for home heating that will be deployed in each home, which will attempt to minimize consumption without impacting householders' comfort [Mozer et al. 1996]. We believe that a controller equipped with LFM-TM will yield significantly better heating plans, which achieve a better tradeoff between cost and comfort in comparison with existing work. Moreover, an accurate thermal model is also important for cooling. We are currently investigating how LFM-TM can be embedded within an HVAC controller for cooling a home in a hot country. More specifically, we are calibrating LFM-TM to data collected from a set of homes in Saudi Arabia. However, discussion of this specific application is beyond the scope of this article and will be covered in detail in a separate publication. Finally, validation of the learned residual by communicating with the householders to understand the correspondence between their activities and the residual induced in the data will also be addressed as future work.

APPENDIX

A. GRAY-BOX MODELS

We describe gray-box models of the thermal dynamics of buildings. This approach combines physical knowledge of a building (thermodynamics within a building) with actual data collected from the same setting in the development of a thermal model [Jimenez and Madsen 2008]. This combination facilitates insight into otherwise hidden information about the physical properties of a building [Bacher and Madsen 2011]. Although it is possible to represent a gray-box model graphically as a resistor capacitor network (RC network) [Bacher and Madsen 2011], for consistency and ease of interpretability, we present their differential forms in this article.

We start with a description of the overall heat dynamics within any built environment that can be expressed as a heat balance equation as per Madsen and Holst [1995]:

$$C_a \frac{\mathrm{d}T_{int}}{\mathrm{d}t} = \sum \emptyset_{in} + \sum \emptyset_{out}, \tag{27}$$

where a day has been divided into a set of discrete time slots, $t \in N$. Here, C_a (J/°C) represents the heat capacity, and the internal temperature of a residential building at time t is $T_{int}(t) \in \mathbb{R}^+$. $\emptyset$ (W) denotes the heat transfer that influences the overall heat dynamics. To capture the heat dynamics more accurately, it is necessary to deconstruct $\emptyset$ into the appropriate convective, conductive, and radiative heat transfer components [Madsen and Holst 1995]. We now provide a brief description of a number of gray-box

models that have been proposed previously [Bacher and Madsen 2011], which will be used for the experiments described later in this article.

A.1. The Ti Model

One can derive a simple thermal model that only considers the heat input by the heating system and the convective heat transfer (leakage) to the ambient air outside as per Equation (28):

$$\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} = \frac{1}{C_i}\eta_h(t) + \frac{1}{C_i R_{ia}}\left(T_{ext}(t) - T_{int}(t)\right) + \sigma_i \mathrm{d}\omega_i. \tag{28}$$

In this article, we refer to this model as the Ti model and the coefficients in this thermal model are

—C_i: the heat capacity for the interior of a built environment (kWh/°C)
—R_{ia}: thermal resistance between the interior and the ambient (°C/kW)

We also denote the external temperature (in °C) as $T_{ext}(t) \in \mathbb{R}^+$ and define the variable $\eta_h(t)$ to be the total heater output (in kW) at t. Equation (28) is a linear first-order stochastic differential equation. ω_i is a standard Wiener noise process having incremental variance σ_i^2 that represents the noise in the physical process.

In some situations, it might be possible to record observations of the global solar irradiance incident on a building. In such cases, Ti can be augmented as per

$$\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} = \frac{1}{C_i}\eta_h(t) + \frac{1}{C_i R_{ia}}\left(T_{ext}(t) - T_{int}(t)\right) + \frac{A_w}{C_i}\eta_s(t) + \sigma_i \mathrm{d}\omega_i, \tag{29}$$

where $\eta_s(t)$ is the measured solar irradiance at a building (kW/m^2) and A_w is the effective window area of the building (m^2). One can continue to add more complexity by adding additional states, resulting in thermal models of increasing complexity, all represented as coupled stochastic first-order differential equations.

A.2. The TiTh Model

This model is derived by augmenting Ti with an additional state that captures the temperature of the heaters as per

$$\begin{aligned}\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} &= \frac{1}{C_i R_{ia}}\left(T_{ext}(t) - T_{int}(t)\right) + \frac{1}{C_i R_{ih}}\left(T_h(t) - T_{int}(t)\right) + \sigma_i \mathrm{d}\omega_i \\ \frac{\mathrm{d}T_h(t)}{\mathrm{d}t} &= \frac{1}{C_h R_{ih}}\left(T_{int}(t) - T_h(t)\right) + \frac{1}{C_h}\eta_h(t) + \sigma_h \mathrm{d}\omega_h.\end{aligned} \tag{30}$$

We refer to this model as the $TiTh$ model. $T_h(t)$ represents the temperature of the internal heaters (in °C) at t. It is a latent variable that is not directly observed, but inferred from the data. The additional parameters in this model are

—C_h: the heat capacity of the internal heaters in a building (kWh/°C)
—R_{ih}: thermal resistance between the interior and the heaters (°C/kW)

Again, if observations of the global solar irradiance incident on a building are available, Equation (30) can be augmented to derive

$$\begin{aligned}\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} &= \frac{1}{C_i R_{ia}}\left(T_{ext}(t) - T_{int}(t)\right) + \frac{1}{C_i R_{ih}}\left(T_h(t) - T_{int}(t)\right) + \frac{A_w}{C_i}\eta_s(t) + \sigma_i \mathrm{d}\omega_i \\ \frac{\mathrm{d}T_h(t)}{\mathrm{d}t} &= \frac{1}{C_h R_{ih}}\left(T_{int}(t) - T_h(t)\right) + \frac{1}{C_h}\eta_h(t) + \sigma_h \mathrm{d}\omega_h.\end{aligned} \tag{31}$$

Again, in Equations (30) and (31), ω_i and ω_h are the corresponding Wiener noise processes.

A.3. The TiTeTh Model

This model is derived by augmenting *TiTh* with a further state that captures the state of the building envelope as per

$$\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} = \frac{1}{C_i R_{ih}}\left(T_h(t) - T_{int}(t)\right) + \frac{1}{C_i R_{ie}}\left(T_{env}(t) - T_{int}(t)\right) + \sigma_i \mathrm{d}\omega_i$$

$$\frac{\mathrm{d}T_{env}(t)}{\mathrm{d}t} = \frac{1}{C_e R_{ie}}\left(T_{int}(t) - T_{env}(t)\right) + \frac{1}{C_e R_{ea}}\left(T_{ext}(t) - T_{env}(t)\right) + \sigma_e \mathrm{d}\omega_e$$

$$\frac{\mathrm{d}T_h(t)}{\mathrm{d}t} = \frac{1}{C_h R_{ih}}\left(T_{int}(t) - T_h(t)\right) + \frac{1}{C_h}\eta_h(t) + \sigma_h \mathrm{d}\omega_h. \tag{32}$$

Here, $T_{env}(t)$ represents the temperature of the building envelope (in $^\circ$C) at t. It too is not directly observed, but inferred from the data. The additional parameters in this model are

—C_e: the heat capacity of the building envelope (kWh/$^\circ$C)
—R_{ie}: thermal resistance between the interior and the envelope ($^\circ$C/kW)
—R_{ea}: thermal resistance between the envelope and the ambient air ($^\circ$C/kW)

Again, if observations of the global solar irradiance incident on a building are available, Equation (30) can be augmented as per

$$\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} = \frac{1}{C_i R_{ih}}\left(T_h(t) - T_{int}(t)\right) + \frac{1}{C_i R_{ie}}\left(T_{env}(t) - T_{int}(t)\right) + \sigma_i \mathrm{d}\omega_i$$

$$\frac{\mathrm{d}T_{env}(t)}{\mathrm{d}t} = \frac{1}{C_e R_{ie}}\left(T_{int}(t) - T_{env}(t)\right) + \frac{1}{C_e R_{ea}}\left(T_{ext}(t) - T_{env}(t)\right) + \frac{A_e}{C_e}\eta_s(t) + \sigma_e \mathrm{d}\omega_e$$

$$\frac{\mathrm{d}T_h(t)}{\mathrm{d}t} = \frac{1}{C_h R_{ih}}\left(T_{int}(t) - T_h(t)\right) + \frac{1}{C_h}\eta_h(t) + \sigma_h \mathrm{d}\omega_h. \tag{33}$$

Here, A_e is the effective area through which solar radiation enters the building envelope (in m^2). In Equations (32) and (33), ω_i, ω_e, and ω_h are the corresponding Wiener noise processes.

A.4. The TiTeThTs Model

This model is derived by augmenting *TiTeThTs* with the state of the sensor that measures the ambient internal temperature as per

$$\begin{aligned}
\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} &= \frac{1}{C_i R_{ih}}\left(T_h(t) - T_{int}(t)\right) + \frac{1}{C_i R_{ie}}\left(T_{env}(t) - T_{int}(t)\right) \\
&\quad + \frac{1}{C_i R_{is}}\left(T_s(t) - T_{int}(t)\right) + \sigma_i \mathrm{d}\omega_i \\
\frac{\mathrm{d}T_{env}(t)}{\mathrm{d}t} &= \frac{1}{C_e R_{ie}}\left(T_{int}(t) - T_{env}(t)\right) + \frac{1}{C_e R_{ea}}\left(T_{ext}(t) - T_{env}(t)\right) + \sigma_e \mathrm{d}\omega_e \\
\frac{\mathrm{d}T_h(t)}{\mathrm{d}t} &= \frac{1}{C_h R_{ih}}\left(T_{int}(t) - T_h(t)\right) + \frac{1}{C_h}\eta_h(t) + \sigma_h \mathrm{d}\omega_h \\
\frac{\mathrm{d}T_s(t)}{\mathrm{d}t} &= \frac{1}{C_s R_{is}}\left(T_{int}(t) - T_s(t)\right) + \sigma_s \mathrm{d}\omega_s.
\end{aligned} \tag{34}$$

Here, $T_s(t)$ represents the temperature of the temperature sensor (in $^\circ$C) at t. The additional parameters in this model are

—C_s: the heat capacity of the sensor (kWh/$^\circ$C)
—R_{is}: thermal resistance between the interior and the sensor ($^\circ$C/kW)

Again, if observations of the global solar irradiance incident on a building are available, Equation (30) can be augmented to derive

$$\begin{aligned}
\frac{\mathrm{d}T_{int}(t)}{\mathrm{d}t} &= \frac{1}{C_i R_{ih}}\left(T_h(t) - T_{int}(t)\right) + \frac{1}{C_i R_{ie}}\left(T_{env}(t) - T_{int}(t)\right) \\
&\quad + \frac{1}{C_i R_{is}}\left(T_s(t) - T_{int}(t)\right) + \sigma_i \mathrm{d}\omega_i \\
\frac{\mathrm{d}T_{env}(t)}{\mathrm{d}t} &= \frac{1}{C_e R_{ie}}\left(T_{int}(t) - T_{env}(t)\right) + \frac{1}{C_e R_{ea}}\left(T_{ext}(t) - T_{env}(t)\right) \\
&\quad + \frac{A_e}{C_e}\eta_s(t) + \sigma_e \mathrm{d}\omega_e \\
\frac{\mathrm{d}T_h(t)}{\mathrm{d}t} &= \frac{1}{C_h R_{ih}}\left(T_{int}(t) - T_h(t)\right) + \frac{1}{C_h}\eta_h(t) + \sigma_h \mathrm{d}\omega_h \\
\frac{\mathrm{d}T_s(t)}{\mathrm{d}t} &= \frac{1}{C_s R_{is}}\left(T_{int}(t) - T_s(t)\right) + \sigma_s \mathrm{d}\omega_s.
\end{aligned} \tag{35}$$

In Equation (34) and Equation (35), ω_s is an additional Wiener noise process. For all the gray-box models described previously, barring the *TiTeThTs* model, the corresponding observation model is given by

$$y(t) = T_{int}(t) + e(t), \tag{36}$$

where $e(t)$ is the measurement noise and is assumed to be i.i.d Gaussian.

For all the *TiTeThTs* models, the corresponding observation model is given by

$$y(t) = T_s(t) + e(t). \tag{37}$$

A.5. Estimation and Parameter Learning in Existing Gray-Box Models

The stochastic differential equation representations of the aforementioned gray-box models can also be represented as continuous time state-space models as per

$$\mathrm{d}\mathbf{X} = A\mathbf{X}dt + B\mathbf{U}dt + \mathrm{d}\mathbf{W} \tag{38}$$

$$\mathbf{Y} = C\mathbf{X} + D\mathbf{U} + \mathbf{e}. \tag{39}$$

Here, Equations (38) and (39) represent the process and observation models, respectively; $\mathbf{X}$ is the state vector that contains all the state variables; and $\mathbf{U}$ contains all the exogenous inputs (external temperature, solar radiation, etc.). The matrices A, B, C, and D are all parameters (θ) of the state-space model [Simon 2006], while $\mathbf{W}$ and $\mathbf{e}$ represent the process and observation (Wiener) noise processes. In Equation (39), the observations over time t, $\mathbf{Y} = [Y(1), \ldots, Y(t)]$, are assumed to be noise-corrupted observations of the state $\mathbf{X} = [X(1), \ldots, X(t)]$. Since Equations (38) and (39) are linear continuous-time models, they can be appropriately discretized, with states estimated using a Kalman filter [Jimenez and Madsen 2008]. The parameters θ can be learned from data using a range of different techniques such as least squares, maximum likelihood, or subspace identification [Privara et al. 2012]. Further details on parameter estimation in state-space models can be found in Madsen and Holst [1995], and Bacher and Madsen [2011].

All the aforementioned gray-box models described in this section are developed based on an understanding of the various physical elements that make up the physical layout of a building. Each model introduces additional dynamics through extra states and parameters. In effect, the physical factors that influence thermal dynamics are gradually accommodated in models of increasing complexity. However, for these models to be deployed in practice, all the factors that influence the evolution of the internal temperature have to be specified a priori. In the event that there are any unspecified residual dynamics, existing gray-box models will erroneously attempt to accommodate them as the variance in the process noise having white noise properties. However, as explained in Section 1, this assumption is invalid in many cases, causing structure to appear in the model's residual. To address this limitation, we introduced LFM-TM in Section 3, which is essentially the *Ti* model augmented with a time-varying term that is specifically introduced to model a residual having structure.

REFERENCES

Mauricio A. Álvarez, David Luengo, and Neil D. Lawrence. 2009. Latent force models. *Journal of Machine Learning Research - Proceedings Track* 5 (2009), 9–16.

Mauricio A. Álvarez, Jan Peters, Bernhard Schölkopf, and Neil D. Lawrence. 2010. Switched latent force models for movement segmentation. In *Proceedings of Neural Information Processing Systems (NIPS'10)*. 55–63.

Klaus Kaae Andersen, Henrik Madsen, and Lars H. Hansen. 2000. Modelling the heat dynamics of a building using stochastic differential equations. *Energy and Buildings* 31, 1 (2000), 13–24.

Anil Aswani, Neal Master, Jay Taneja, David E. Culler, and Claire Tomlin. 2012. Reducing transient and steady state electricity consumption in HVAC using learning-based model-predictive control. *Proceedings of the IEEE* (2012), 240–253.

Peder Bacher and Henrik Madsen. 2011. Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings* 43, 7 (2011), 1511–1522.

Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning*. Springer.

Gavin C. Cawley and Nicola L. C. Talbot. 2010. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research* 11 (2010), 2079–2107.

Anastasios I. Dounis and Christos Caraiscos. 2009. Advanced control systems engineering for energy and comfort management in a building environment—A review. *Renewable and Sustainable Energy Reviews* 13, 6 (2009), 1246–1261.

James Durbin and Siem Jan Koopman. 2001. *Time Series Analysis by State Space Methods*. Number 24. Oxford University Press.

Grayson Evans. 1991. Solving home automation problems using artificial intelligence techniques. *IEEE Transactions on Consumer Electronics* 37, 3 (1991), 395–400.

Diego Fernández Slezak, Cecilia Suárez, Guillermo A. Cecchi, Guillermo Marshall, and Gustavo Stolovitzky. 2010. When the optimal is not the best: Parameter estimation in complex biological models. *PLoS ONE* 5, 10 (2010).

Mohinder S. Grewal and Angus P. Andrews. 2001. *Kalman Filtering: Theory and Practice Using MATLAB*. Wiley-Interscience.

Jouni Hartikainen and Simo Särkkä. 2010. Kalman filtering and smoothing solutions to temporal gaussian process regression models. In *Proceedings of IEEE Workshop on Machine Learning for Signal Processing*. 379–384.

Jouni Hartikainen and Simo Särkkä. 2011. Sequential inference for latent force models. In *Proceedings of Uncertainty in Artificial Intelligence (UAI'11)*. 311–318.

Miguel J. Jimenez and Henrik Madsen. 2008. Models for describing the thermal characteristics of building components. *Building and Environment* 43, 2 (2 2008), 152–162.

Jiakang Lu, Tamim I. Sookoor, Vijay Srinivasan, Ge Gao, Brian Holben, John A. Stankovic, Eric Field, and Kamin Whitehouse. 2010. The smart thermostat: Using occupancy sensors to save energy in homes. In *SenSys*. 211–224.

Henrik Madsen and Jan Holst. 1995. Estimation of continuous-time models for the heat dynamics of a building. *Energy and Buildings* 22, 1 (1995), 67–79.

Alie El-Din Mady, Gregory Provan, Conor Ryan, and Kenneth N. Brown. 2011. Stochastic model predictive controller for the integration of building use and temperature regulation. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI'11)*.

Michael Mozer, Lucky Vidmar, and Robert H. Dodier. 1996. The neurothermostat: Predictive optimal control of residential heating systems. In *Proceedings of Neural Information Processing Systems (NIPS'96)*. 953–959.

Frauke Oldewurtel, Alessandra Parisio, Colin N. Jones, Dimitrios Gyalistras, Markus Gwerder, Vanessa Stauch, Beat Lehmann, and Manfred Morari. 2012. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings* 45 (2012), 15–27.

Samuel Privara, Jiri Cigler, Zdenek Vana, Frauke Oldewurtelb, Carina Sagerschnigc, and Eva Zacekova. 2013. Building modeling as a crucial part for building predictive control. In *Energy and Buildings*, Vol. 56. 8–22.

Samuel Privara, Zdenek Vana, Eva Zacekova, and Jiri Cigler. 2012. Building modeling: Selection of the most appropriate model for predictive control. In *Energy and Buildings*, Vol. 55. 341–350.

Sarvapali D. Ramchurn, Perukrishnen Vytelingum, Alex Rogers, and Nicholas R. Jennings. 2012. Putting the 'smarts' into the smart grid: A grand challenge for artificial intelligence. *Communications of the ACM* 55, 4 (2012), 86–97.

Carl Rasmussen and Christopher Williams. 2005. *Gaussian Processes for Machine Learning*. MIT Press.

James Rawlings. 2000. Tutorial overview of model predictive control. *IEEE Control Systems,* 20, 3 (2000), 38–52.

Alex Rogers, Sasan Maleki, Siddhartha Ghosh, and Nick Jennings. 2011. Adaptive home heating control through Gaussian process prediction and mathematical programming. In *Proceedings of the 2nd Agent Technologies for Energy Systems Workshop (AAMAS'11)*.

James Scott, A. J. Bernheim Brush, John Krumm, Brian Meyers, Mike Hazas, Stephen Hodges, and Nicolas Villar. 2011. PreHeat: Controlling home heating using occupancy prediction. In *Ubicomp*. 281–290.

Dan Simon. 2006. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley.

Daniel Urieli and Peter Stone. 2013. A learning agent for heat-pump thermostat control. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'13)*. 1093–1100.

Zhe Yu, Linda McLaughlin, Liyan Jia, Mary C.Murphy-Hoye, Annabelle Pratt, and Lang Tong. 2012. Modeling and stochastic control for home energy management. In *Proceedings of the 2012 IEEE Power and Energy Society General Meeting*. 1–9.

Received October 2013; revised February 2014; accepted March 2014