Testing building controls with the BLDG toolbox

Invited Paper, 2016 American Control Conference, Boston, MA

Kevin J. Kircher and K. Max Zhang

Abstract—We introduce the open-source MATLAB® toolbox bldg. The main function, bsim, simulates a simple building under forcing from the outdoor air temperature, sun, and internal heat sources. bsim retains the mathematical structure of state-of-the-art building simulators, including time-varying solar forcing, PDE-governed wall temperatures, and nonlinearities arising from thermal radiation and temperature-dependent convection coefficients. Supporting functions are provided to instantiate a default building object, import weather data, and generate internal heat source signals. The bldg toolbox is intended to eliminate the need to model a building or leave the MATLAB® environment before testing an estimation or control algorithm. We demonstrate the software through the examples of system identification, online state and parameter estimation, and model predictive control. Open research questions that the bldg toolbox could help explore are noted.

I. INTRODUCTION

The world’s buildings have many different sizes, shapes, locations and weather conditions, material properties, mechanical systems, and patterns of use. This variety makes it difficult to draw general conclusions about the performance of a building estimation or control algorithm. Typical practice, therefore, is to test algorithms through case studies and hope that observed trends generalize.

Many building control algorithms have been proposed in recent years. Since 2010, researchers have tested certainty-equivalent model predictive control (MPC) [1–3], learning-based MPC [4], economic MPC [5, 6], stochastic MPC [7–9], robust MPC [3], as well as subspace system identification [10], MPC-relevant system identification [10], and the linear, extended, and unscented Kalman filters [2, 11–15].

Some research, such as investigating an algorithm’s scalability to large multizone buildings, requires a complex building geometry. For many purposes, however, a simple geometry suffices. Indeed, the model in the majority of the case studies mentioned above consisted of either one [3, 4, 7, 8, 14, 15] or two [2, 10] rooms. These simple test cases have the advantages of easy modeling, fast simulation, and transparency.

Model fidelity also matters. For a case study to be meaningful, the system from which control-oriented models are learned, and in which controls are implemented, should behave like a real (i.e., nonlinear and time-varying) building. Empirical case studies such as [1, 3, 4, 9, 10, 13, 14] certainly pass this test. Simulation-based studies, on the other hand, commonly use semiphysical linear systems as truth models [2, 6–8, 11, 15]. Only a few [5, 10, 12] have learned control-oriented models from, or implemented controls in, high-fidelity simulators such as EnergyPlus, TRNSYS, or IDAICE. This is likely due to the nontrivial effort required to learn a simulation program, develop a model, and integrate the program with MATLAB®, the standard environment for control design.

We believe that building control researchers could benefit from a simple, control-oriented, high-fidelity MATLAB® building simulator. As a step toward that end, in this paper we introduce the bldg toolbox. We review related work in §II. In §III, we give an overview of the main function bsim, which mimics the syntax and behavior of the MATLAB® linear system simulator lsim. Examples are presented in §IV, including system identification, online state and parameter estimation, and MPC. We conclude in §V.

II. RELATED WORK

Two recent projects overlap with our work on control-oriented building simulation in MATLAB®. These are the Building Resistance-Capacitance Modeling (BRCM) [16] and OpenBuild [17] toolboxes. Both are built on the Building Controls Virtual Test-Bed [18] and MLE+ [19], which enable communication between MATLAB® and EnergyPlus.

The main factor that distinguishes bldg from BRCM is fidelity. BRCM maps an EnergyPlus model to an RC network and generates cost and constraint matrices for MPC. The resulting RC network neglects solar forcing and the nonlinearities arising from longwave radiation and temperature-dependent convection coefficients. These departures from the best practices of white-box simulators make BRCM a poor truth model for algorithm testing. This is to be expected, since BRCM was not designed as a test-bed, but rather to enable MPC in real, complex buildings. The effectiveness of BRCM for MPC is demonstrated in [20].

Like BRCM, OpenBuild integrates MATLAB® and EnergyPlus. However, OpenBuild uses EnergyPlus as its physics engine, and can therefore be regarded as a valid truth model. The key difference between bldg and OpenBuild is simplicity. While OpenBuild’s link to EnergyPlus allows researchers to study a wide range of building geometries and HVAC systems, it also introduces all of the physical and numerical complexity described in the 1,444-page EnergyPlus engineering reference [21] and the 2,234-page EnergyPlus input-output reference [22]. By contrast, we propose a simple building with few parameters and explain all of our assumptions and numerical methods in the 14-page
A room whose length and height are much greater than the glass thickness $T$ distribution, and the subscripts $-\text{ and } +$ indicate the left and right wall surfaces, respectively. The fluxes $q_-$ and $q_+$ (W/m$^2$), which are nonlinear and time-varying, include convection with temperature-dependent coefficients, longwave radiation exchange with the window and surrounding environment, and shortwave radiation from the sun and artificial lighting.

As demonstrated in [24], the window temperature is well-modeled by

$$C_g \dot{T}_g = A(q_{g-} - q_{g+}), \quad T_g(0) = T^0_g, \quad (2)$$

where $C_g$ (J/K) is the thermal capacitance of the window and $A$ (m$^2$) is its surface area. As with the wall, the window surface fluxes $q_{g-}$ and $q_{g+}$ include nonlinear convection and radiation. The model also accounts for the reflection, absorption, and transmission of visible light. The internal source $Q_a(t)$ includes heat flows from people, control systems, lights, and other equipment (appliances, electronics, etc.).

III. SIMULATOR OVERVIEW

The bldg toolbox simulates the building illustrated in Figure 1. It is a box with two dimensions sufficiently greater than the third that all heat transfer can be considered one-dimensional. For different choices of the geometric and physical parameters, this building can be thermally light or heavy, with weak or strong solar forcing and a drafty or tight envelope. In this section, we briefly review the governing equations and numerical solution scheme, input signals, and simulator syntax. Further details, including empirical validation and sensitivity analysis, can be found in [23].

A. Governing equations

The wall temperature distribution $T(x,t)$ satisfies the heat equation,

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2}, \quad T(x,0) = T^0(x)$$

$$-k \frac{\partial T}{\partial x}_- = q_-, \quad -k \frac{\partial T}{\partial x}_+ = q_+.$$ \quad (1)

Here $\alpha$ (m$^2$/s) and $k$ (W/m·K) are the wall’s thermal diffusivity and conductivity, $T^0(x)$ is the initial temperature distribution, and the subscripts $-\text{ and } +$ indicate the left and right wall surfaces, respectively. The fluxes $q_-$ and $q_+$ (W/m$^2$), which are nonlinear and time-varying, include convection with temperature-dependent coefficients, longwave radiation exchange with the window and surrounding environment, and shortwave radiation from the sun and artificial lighting.

As demonstrated in [24], the window temperature is well-modeled by

$$C_g \dot{T}_g = A(q_{g-} - q_{g+}), \quad T_g(0) = T^0_g,$$ \quad (2)

where $C_g$ (J/K) is the thermal capacitance of the window and $A$ (m$^2$) is its surface area. As with the wall, the window surface fluxes $q_{g-}$ and $q_{g+}$ include nonlinear convection and radiation. The model also accounts for the reflection, absorption, and transmission of visible light by the window.

The indoor air is assumed to be well-mixed, with temperature $T_a$ governed by

$$C_a \dot{T}_a = A(q_{a-} - q_{a+}) + Q_a, \quad T_a(0) = T^0_a,$$ \quad (3)

where $C_a$ (J/K) is the thermal capacitance of the air and any material that is isothermal with it (e.g., ducts, dampers, or furniture). The fluxes $q_{a-}$ and $q_{a+}$ couple the air to the window and wall. The internal heat flow $Q_a$ (W) includes infiltration of outdoor air, as well as heat flows from control systems, people, lights, and other equipment.

By carefully discretizing spatial derivatives, the PDE (1) can be written with $O((\Delta x)^2)$ truncation error as a system of $N$ coupled ODEs, where $N$ is the number of finite difference nodes in the wall and $\Delta x = l/(N - 1)$. These ODEs are coupled to Equations (2) and (3) through $q_-$, $q_{g+}$, and $q_{a+}$, giving a nonlinear, time-varying dynamical system that can be simulated by a MATLAB® stiff ODE solver. The system state is $(T_1, \ldots, T_N, T_g, T_a)$, where $T_i$ is the temperature of the $i^{th}$ wall node.

B. Input signals

We divide the input signals into disturbances and controls, though this distinction is somewhat arbitrary. The disturbance is $(T_{\infty}, I_b, I^h_b, Q_{p}, Q_{l}, Q_{e})$, where $I_b$ (W/m$^2$) is the total solar irradiance on a horizontal surface, $I^h_b$ (W/m$^2$) is the beam irradiance on a surface normal to the sun, and $Q_{p}$, $Q_{l}$, and $Q_{e}$ (W) are the internal gains from people, lights, and other equipment.

1For example, with flexible plug loads $Q_e$ could be controllable.
Fig. 2. Input signals from weather (top) and internal heat sources (bottom). The outdoor air temperature ($T_{\infty}$), total solar irradiance on a horizontal surface ($I_{h}$), and beam solar irradiance on a surface normal to the sun ($I_{b}$) show that the simulation day is cold and partly cloudy. The internal heat sources model a home on a weekday, with decreased heat flows from people ($Q_{p}$), lights ($Q_{l}$), and equipment ($Q_{e}$) overnight and during work hours.

The control can be either $Q_{e}$ (W), the heat supplied by control systems, or ($\dot{m}_{c}, T_{c}$), where $\dot{m}_{c}$ (kg/s) is the mass flow rate of supply air and $T_{c}$ is the supply air temperature. In the latter case, the heat supplied by control systems is state-dependent and nonlinear: $Q_{c} = \dot{m}_{c} c_{u}(T_{c} - T_{a})$, where $c_{u}$ (J/kg·K) is the specific heat of air at constant pressure. If no control is specified, bsim returns the heat flows required to perfectly regulate the air temperature at its initial value. This is useful for quick simulations or to size heating and cooling systems.

C. Syntax

Under the default parameter values, bsim simulates a passive solar building that stores the sun’s energy in the thermally massive wall and releases it overnight. The user can modify this behavior by adjusting parameters in the bldg object. For example, the thermal mass can be decreased by increasing the wall diffusivity $\alpha$ (bldq.a), and the solar forcing can be weakened by decreasing the wall shortwave absorptivity $\alpha_{s}$ (bldq.as). The sensitivity analysis in [23] contains more examples.

The following code instantiates a default building object, imports and interpolates a year of hourly weather data from a TMY3 file for New York City, generates plausible internal gains for a home, and defines a 24-hour simulation time span with step $\Delta t = 15$ minutes, starting at midnight ($t = 0$) on February 7 (day number $n_{d} = 38$). The last line packs the disturbances into the matrix form accepted by bsim.

```
b = bldg; dt = 15*60;
gains = generateGains(b,weather.tw);
nd = 38; t0 = 0; tf = 24*3600;
t = getTiming(weather.tw,nd,t0,tf);
W = getDisturbances(weather,gains,t);
```

Figure 2 shows the input signals generated by this code. Simulating the building requires an initial state, which can be difficult to produce a priori. One is easily generated, however, using

```
x0 = precondition(...
b,t,weather,gains,N,Ts);
```

The precondition function starts the building at a naive initial state two weeks before $t_{0}$ and integrates it forward with $T_{a}$ perfectly regulated at setpoint $T_{s}$. The building can then be simulated by

```
[X,Qc] = bsim(b,t,W,x0);
```

With $N = 50$ wall nodes, this 24-hour simulation takes about a tenth of a second to run on a 2 GHz Intel Core 2 Duo processor. Figure 3 shows the output.

IV. Examples

A. System identification

In this example, we consider the problem of learning a low-order, linear model of the building dynamics from measurements of the indoor air temperature and weather signals. We specify the first-order ARX structure

$$T_{a}^{k+1} = \beta_{1} T_{a}^{k} + \beta_{2} Q_{c}^{k} + \beta_{3} I_{\infty}^{k} + \beta_{4} I_{h}^{k} + w_{a}^{k},$$

(4)

where the $w_{a}^{k}$ are independently, identically $N(0, \sigma_{w}^{2})$ distributed. This model is naive, since it neglects the internal gains and the dynamics of the building envelope, and since the true solar forcing is time-varying and cannot be determined from $I_{h}$ alone. Nevertheless, the model is sufficiently accurate to give fair controller performance (see §IV-C).

To fit the model, we simulate the (nonlinear, time-varying, high-dimensional) dynamics in bsim for the last three weeks of January under a sequence of pseudorandom binary control inputs. We assume perfect knowledge of $Q_{e}$, $T_{\infty}$, and $I_{h}$,
but corrupt $T_a$ with zero mean, white, Gaussian noise with standard deviation $\sigma_c = 1/6$ C. The model parameters $\beta$ and $\sigma_u$ are estimated using linear regression.

Figure 4 shows the one-step prediction errors during the three training weeks, and in two subsequent test weeks under thermostatic control with deadband [18 C, 22 C]. The maximum absolute error of about 2 C is large. A higher-order ARX, ARMAX, or RC network model could achieve better predictions at the cost of introducing more states and parameters. This accuracy/complexity trade-off was explored for a family of RC networks in [15]. The authors found that a second-order RC network with five parameters is able to capture the essential building behavior (a result that dates back at least to 1980; see [25, 26]). Less is known about accuracy/complexity trade-offs for buildings with thermally massive construction or strong solar forcing. These questions could be explored in a simulation by varying physical parameters such as the wall diffusivity $\alpha$ (bldg.a) and shortwave absorptivity $\alpha_s$ (bldg.as) and studying the model fit.

Another interesting topic is the relationship between gray- and black-box models. The parameters in model (4), for example, can be interpreted in terms of the first-order RC network

$$C_{\text{eff}}T_a = \frac{T_{\infty} - T_a}{R_{\text{eff}}} + Q_c + A_{\text{eff}}I_h,$$

where $C_{\text{eff}}$, $R_{\text{eff}}$, and $A_{\text{eff}}$ are the effective capacitance, resistance, and solar absorption area. Forward Euler discretization gives

$$T_{a,k+1} = \left(1 - \frac{\Delta t}{R_{\text{eff}}C_{\text{eff}}}\right)T_{a,k} + \frac{\Delta t}{C_{\text{eff}}}Q_{c,k} + \frac{\Delta t}{R_{\text{eff}}C_{\text{eff}}}T_{\infty}^k + \frac{\Delta tA_{\text{eff}}}{C_{\text{eff}}}I_{h,k},$$

2 which is consistent with model (4) only if $\beta_1 + \beta_3 = 1$. In this case, the ‘cost of grayness’ is the accuracy loss by imposing the constraint $\beta_1 + \beta_3 = 1$ on the least-squares fitting problem, along with the resulting controller performance reduction (if any). While several authors have argued that the RC network structure has benefits – e.g., it can provide initial guesses to estimation algorithms and sanity checks on their output – to our knowledge the cost of grayness has not been well-studied.

### B. Online estimation

Building dynamics are naturally time-varying: Solar forcing depends on the season and time of day, infiltration rates change as occupants open and close windows and doors, and control setpoints and delivery mechanisms differ in heating and cooling modes. We therefore expect model parameters to change over several time scales. The framework of online estimation allows parameters to be continuously calibrated to measurements, enabling adaptive control. Online estimation of building model parameters has been studied in [2, 11–14].

In this example, we consider the problem of adapting a subset of the parameters of the ARX model identified in §IV-A, in order to reduce the prediction bias apparent in Figure 4. We follow the typical approach of simultaneously estimating the states and parameters with an unscented Kalman filter. [27] We allow the parameters $\beta_1$ and $\beta_3$, initialized with the fit from §IV-A, to vary under the random walk model $\beta^{k+1} = \beta^k + w_{\beta}^k$. This gives the augmented system model

$$x^{k+1} = f(x^k, u^k) + w^k,$$

$$y^k = x^k + v^k,$$

5

where $x = (T_a, \beta_1, \beta_3)$, $u = (Q_c, T_{\infty}, I_h)$, $w = (w_a, w_{\beta_1}, w_{\beta_3})$, $f_1(x, u) = x_2x_1 + \beta_2u_1 + x_3u_2 + \beta_4u_3$, and $f_i(x, u) = x_i$ for $i = 2, 3$. We model the disturbance $w$ as zero-mean, white, and Gaussian. The RC network analogy developed in §IV-A suggests that the parameters $\beta_1$ and $\beta_3$ should (roughly) sum to one, so we specify a strong negative correlation between $w_{\beta_1}$ and $w_{\beta_3}$.

Figure 5 shows histograms of the one-step prediction errors in the test data with and without parameter adaptation. Adaptation nearly eliminates the prediction bias. The unscented Kalman filter accomplishes this by decreasing $\beta_1$ and increasing $\beta_3$ over the course of about twelve hours, as shown in Figure 6. The parameters stabilize after the initial adjustment. This is expected, since the underlying physical model remains nearly constant over the estimation period. It is less clear how the filter would respond to a large,
sudden change in the underlying system, e.g., a window being opened or a mechanical component failing. These questions, which lie in the domain of fault detection, could be explored in bsim by perturbing the mass flow rate of infiltration air, $\dot{m}_\infty$ (bldg.mdot), or the fraction of the control heat that convects to the indoor air, $\zeta_c$ (bldg.zc).

### C. Model predictive control

This example involves efficiently heating a building. We consider the stochastic optimal control problem

$$\begin{align*}
\text{minimize} & \quad E \left[ \Delta t \sum_{k=0}^{M} Q_c^k \right] \\
\text{subject to} & \quad T_{a}^k = T_{a}^{k-1} + a^k \\
& \quad T_{a}^{k+1} \geq T_{\text{min}} \quad \text{almost surely} \\
& \quad Q_c^k = \mu^k (y_0^k, \ldots, y_M^k) \in [0, Q_{c_{\text{max}}}],
\end{align*}$$

where the constraints hold for $k = 0, \ldots, M$. The optimization variable is the control policy $(\mu^0, \ldots, \mu^M)$, where $\mu^k : \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ maps observations into controls. The expectation is taken with respect to the joint distribution of the disturbance and noise sequences and the initial state.

Problem (6) is analytically intractable due to the nonlinear dynamics, imperfect state information, and optimization over infinite-dimensional objects (the functions $\mu^k$). While it is not difficult to generate good approximate solutions—a well-tuned thermostat works—we apply two variants of MPC to Problem 6 in order to illustrate the research value of the bldg toolbox. The first MPC variant uses the ARX model from §IV-A, with no parameter adaptation. At each time step, we estimate the temperature using the linear Kalman filter, then solve a truncated, certainty-equivalent version of problem (6) with horizon $H = 6$ hours. Each MPC subproblem is a deterministic linear program that generates a planned control trajectory, of which the first control is implemented. We then allow the system to evolve according to the nonlinear, time-varying dynamics in bsim and repeat the process.

The second MPC variant is identical, except that state and parameter estimates are simultaneously updated at each time step using the unscented Kalman filter from §IV-B. Simulating the building under MPC for one day with fifteen-minute time steps takes about 42 seconds, with linear programs solved by Gurobi on a 2 GHz Intel Core 2 Duo processor. About 93% of that time is spent in optimization, 6% in bsim, and 1% in the unscented Kalman filter.

Figure 7 shows the indoor air temperatures and control heat flows under both MPC variants. The simulation takes place on February 7, under the exogenous input signals shown in Figure 2. Both MPC variants attempt to regulate $T_a$ at the minimum feasible temperature of $20^\circ C$. This generates control trajectories that resemble noisy versions of the ‘perfect regulation’ case discussed in §III-B. Due to model error, both MPC variants frequently underheat due to model error. Parameter adaptation mitigates, but does not eliminate, constraint violation.

The constraint violations are an artifact of model mismatch between the controller and the underlying bsim dynamics. They could be reduced by using a more accurate, higher-order model, by modifying the MPC objective function (e.g., by replacing $T_{\text{min}}$ with $T_{\text{min}} + \delta$ for some $\delta > 0$), or by moving the MPC optimization to a stochastic or robust framework. While several studies have applied stochastic and/or robust MPC to buildings [3, 7–9], it is not clear
how these approaches, which add significant conceptual and computational complexity, compare to increasing the model order or modifying the objective function.

V. CONCLUSION

In this paper, we introduced the MATLAB® building simulation toolbox bldg. We reviewed the building geometry and governing equations, then demonstrated the use of the main function bsim as a nonlinear, time-varying truth model for system identification, online estimation, and MPC. In MPC simulations, the computation time added by bsim was a modest 6% of the time spent solving linear programs.

Many research questions, such as those involving scalability to large multizone buildings, are beyond the scope of bldg. We also emphasize that simulation is no substitute for empirical validation, regardless of the simulator quality. Nevertheless, our experience suggests that bldg’s speed, fidelity, and easily varied parameters could make it a useful tool for exploring a variety of open research questions, e.g.:

1) How does the trade-off between a learned model’s accuracy and complexity depend on the building’s thermal mass and the strength of solar forcing?
2) What are the costs, in terms of model accuracy or controller performance, of imposing gray-box structure on learned models?
3) Of the various model structures and filtering algorithms in the literature, which are best for fault detection?
4) How do the costs and benefits of using robust or stochastic MPC compare to those of using a more accurate model or of modifying the certainty-equivalent MPC optimization problem?

The bldg toolbox is free and open source. Readers who are interested in using or modifying it can find bldg online at [28].

ACKNOWLEDGMENTS

The authors are grateful to the Hydro Research Foundation and the U.S. Department of Energy’s Office of Energy Efficiency and Renewable Energy for support.

REFERENCES