A Grey-box Model Based on Unscented Kalman Filter to Estimate Thermal Dynamics in Buildings

A Grey-box Model Based on Unscented Kalman Filter to Estimate Thermal Dynamics in Buildings

Marco Massano*, Enrico Macii*, Edoardo Patti*, Andrea Acquaviva* and Lorenzo Bottaccioli*
*Politecnico di Torino, Turin, Italy. Email: name.surname@polito.it

***Abstract*—Buildings are responsible of about 40% of primary energy consumption. The widespread diffusion of Internet-of-Things devices provide allow collecting large amount of energy related data such as indoor air-temperature and power consumption of heating/cooling systems. Collected information can be used to develop data-driven models to learn building characteristics and to forecast indoor temperature trends. In this paper, we present a *Grey-box model* to estimate thermal dynamics in buildings based on *Unscented Kalman Filter* and thermal network representation. The proposed methodology has been applied to different implementation of building thermal networks to test their accuracy in temperature prediction. Results show the accuracy of the proposed methodology in predicting indoor temperature trends up to next 24-hours with a maximum error of $1.50°C$.**

***Index Terms*—Building simulation, Unscented Kalman Filter, Grey-box model, Parameter estimation, Thermal Dynamics**

I. Introduction

Nowadays, more than half of the overall worlds population lives in urban areas. Studies of the United Nation expects that by 2030 urban areas will host around 68% of people [1]. Furthermore, one-third of the population will live in cities with at least half a million of inhabitants [1]. Urbanization is largely energy-intensive as reported by the United Nations habitat division [2]. Cities consume about 75% of the global primary energy supply and they are responsible for about 50-60% of the worlds total greenhouse gas emissions [2]. In particular, heating systems in buildings are responsible of roughly 40% of the overall energy consumption [3]. In this context, ICT and, in particular, Internet-of-Things (IoT) technologies play a crucial role allowing to monitor and optimize energy consumption [4], hence increasing the efficiency of energy systems. This is confirmed by a widespread diffusion of heterogeneous and pervasive devices in our houses and cities. Such IoT devices allow to collect large amounts of energy data (e.g. indoor air temperature) providing detailed information to derive and model thermal dynamics in buildings. Results of these models can foster novel control strategies and tools for energy management in buildings and cities, for example, by exploiting the flexibility of electro-thermal devices. Thus, heating systems in buildings can be included in Demand/Response [5], [6] and Demand Side Management applications [7], [8]

In this paper, we propose a novel data-driven model based on *Unscented Kalman Filter* to estimate thermal dynamics in buildings. It takes advantages of information sampled by pervasive IoT devices to allow control policies for: i) Optimal Scheduling of system operation, ii) Model Predictive Control, iii) Demand Side Management and iv) Demand/Response.

The rest of the paper is organized as follows. Section II reviews literature solution to simulate thermal dynamics in buildings. Section III introduces the Unscented Kalman Filter and Section IV presents the proposed methodology. Section V debates the experimental results conducted on different representation of the thermal model of a test-case building. Finally, Section VI discusses the concluding remarks and future works.

II. Related works

In the last years, a strong research effort has been given to model thermal energy consumption of buildings. In particular, models to describe the thermal dynamics in buildings, either residential or commercial, have been developed, [9]. Such models can be grouped in three macro-categories: i) *White-box*, ii) *Black-box* and iii) *Grey-box*.

White-box models have a complete knowledge and description of i) physical phenomena, ii) structural and thermal parameters of buildings (e.g. thermal capacity of each element like walls and windows). Among all, *Energy+* [10] is one of the most famous solution belonging to this category that needs thermal capacitance, thermal resistance and thickness of the materials to run simulations [11]. In [12], authors present a methodology to realistically simulate thermal behaviours of building by integrating Building Information Models with data sampled by IoT devices to monitor indoor air temperature and by replacing information about Test Meteorological Years with real meteorological data sampled by weather stations. However, even if most of the fundamental parameters and settings are known, White-box models are not error free. Indeed, some other parameters are not known mainly related to inhabitant behaviours and weather conditions (e.g. window openings and temporary cloud covering, respectively) that can significantly affect the thermal behaviour.

Black-box models are empirical models based on statistical data that neglect information on physical phenomena, structural and thermal parameters of buildings. Black-box models aim at finding relation among input and output variables. In [13], [14], authors present two solutions based on neural networks techniques to estimate and predict indoor air temperature in buildings. Such solutions exploit a data-set of about six years with information sampled by IoT devices deployed in a real-world building. The main limitation of this approach is the need of a consistent data-set to train and validate neural networks including all possible energetic conditions. Indeed, this data-set is not available for every building.

Grey-box models are a combination of both White- and Black-box, where statistical information are blended with physical phenomena, structural and thermal parameters of buildings. Several studies demonstrates that the thermal behaviour of a building can be modeled as an Resistor-Capacitor (RC) circuit [15]–[17]. In this analogy with electric circuits, resistors and capacitors are walls and air mass, electric voltage is the temperature and the electric current is the heat flow.

Bacher et al. [18] proposed a model that couples a set of continuous time stochastic differential equations with a set of discrete time measurement equations. In the model, variation of temperature can be affected by disturbance that can not be measured directly, such as windows openings, presence of people, machines, heating systems and solar gain.

Techniques like Kalman filters and genetic algorithms are used to estimate the effects of unmeasured disturbances on the dynamics of state variables. In [19], authors performed a virtual and mini test-bed experiment to test the capability of Kalman filters to detect process disturbance. In the estimation process, they have summed together solar heat gain and losses from internal sources.

Fux et al. [20] applied the *Extended Kalman Filter* to a 1R1C circuit that models a commercial building. This self-adaptive thermal model provides estimations of the unmeasured heat flows caused by inhabitants, predicting the indoor air temperature with an error ranging from $0.5°C$ to $3°C$ for future 3 and 48 hours, respectively.

Kalman filters has been applied also to design a Parameter-Adaptive Building model by simultaneously tuning the parameters of the model and estimating the states of the system [21].

Finally in [22], authors used the *Unscented Kalman Filter* to define the state parameters of a multi-zone thermal network validating their results with *Energy+* simulation. This model is first trained with a 7-days data-set consisting of known loads to learn the constant parameters; then, it is able to characterize the unknown loads (i.e. sun irradiance). Predictions of temperature trends are quite accurate. Indeed, the resulting Root Mean Square Difference is about $1.48°C$.

With respect to literature solutions, our model estimates the effects of solar heat gain and HVAC (Heating, Ventilation and Air Conditioning) systems on the indoor air temperature in buildings. In particular, solar heat gain is modelled by re-projecting the solar radiation on the building facades. Furthermore, we discuss the performance of our model with i) different configurations of the proposed RC circuit, ii) different parameters representing the building elements and iii) different length of the training set.

III. Unscented Kalman Filter

Kalman filters belongs to *Bayesian filters* family. They describe a general probabilistic approach for estimating an unknown probability density function recursively over the time. They use a mathematical model to describe a phenomena, and real control measurements to validate the model. Kalman filters results particularly efficient when a *dual estimation* is needed, i.e. when they have to estimate both state and parameters [23].

Unscented Kalman Filter (UKF) is based on the assumption that a probability distribution function is easier to be approximated than an arbitrary non linear function [24]. Therefore the distribution is approximated by the gaussian density, and it is represented by a set of deterministic chosen samples. These so called *sigma points* completely get the true mean and covariance of the gaussian distribution. When propagated through nonlinear systems, they get the true mean and covariance accurately to the second order (Taylor series expansion) of any non-linearity. The core of UKF is the *Unscented Transform* that allows, with a nonlinear transformation, to compute statistics of a random variable. The *n*-dimensional random variable x, with mean $\hat{x}(k|k)$ and covariance matrix $P(k|k)$ for k^{th} iteration, is approximated by the following equations through $2n+1$ weighted sigma points and correlated weights:

$$\begin{aligned}
\chi_0(k|k) &= \hat{x}(k|k) && i = 0 \\
W_m^{(0)} &= \frac{\kappa}{(n+\kappa)} \\
\chi_i(k|k) &= \hat{x}(k|k) + \left(\sqrt{(n+\kappa)P(k|k)}\right)_i && i = 1, ..., n \\
W_c^{(0)} &= \frac{1}{2(n+\kappa)} \\
\chi_i(k|k) &= \hat{x}(k|k) - \left(\sqrt{(n+\kappa)P(k|k)}\right)_i && i = n+1, ..., 2n \\
W_m^{(i)} &= W_c^{(i)} = \frac{1}{2(n+\kappa)}
\end{aligned} \tag{1}$$

where n represents the state dimension, κ is a scaling parameter such that $\kappa + n \neq 0$ with $n, \kappa \in \mathbb{R}$. W_i is the weight associated to i^{th} point, and it can be associated to the mean weight (W_m) or to the covariance (W_c). Weights are then normalized, so they satisfy $\sum_{i=0}^{2n} W_i = 1$. These sigma vectors are propagated through the nonlinear function $f(\cdot)$ yielding a transformed set of points.

$$\chi^{\mathbf{i}}_{\mathbf{k+1|k}} = f_k\Big(\chi_i\big(k|k\big)\Big) \tag{2}$$

The **Prediction Step** calculates the a-priori mean and covariance as follows:

$$\begin{aligned}
&\hat{x}(k+1|k) = \sum_{i=0}^{2N} W_m^{(i)} \chi^i_{k+1|k} \\
&P_x(k+1|k) = \\
&\sum_{i=0}^{2N} W_c^{(i)} [\chi^i_{k+1|k} - \hat{x}(k+1|k)][\chi^i_{k+1|k} - \hat{x}(k+1|k)]^T + Q
\end{aligned} \tag{3}$$

where Q represents a noise term.

The **Update Step** is performed in measurement space. Thus the sigma points are transformed into $\mathcal{Y}$ through the measurement function $h(\cdot)$:

$$\mathcal{Y}^i_{k+1|k} = h_k(\chi^i_{k+1|k}) \tag{4}$$

and the a-posteriori mean and covariance are calculated as follow:

$$\hat{y}(k+1|k) = \sum_{i=0}^{2N} W_m^{(i)} \mathcal{Y}_{k+1|k}^i$$
$$P_y(k+1|k) = \tag{5}$$
$$\sum_{i=0}^{2N} W_c^{(i)} [\mathcal{Y}_{k+1|k}^i - \hat{y}(k+1|k)][\mathcal{Y}_{k+1|k}^i - \hat{y}(k+1|k)]^T + R$$

where R represents the measurement noise. Then, the residual of the measurement is computed by the following equation:

$$\mathbf{y} = \mathbf{z} - \hat{y}(k+1|k) \tag{6}$$

Cross-covariance of state and measurement is calculated by:

$$P^{xy}(k+1|k) = \tag{7}$$
$$\sum_{i=0}^{2N} W_c^{(i)} [\chi_{k+1|k}^i - \hat{x}(k+1|k)][\mathcal{Y}_{k+1|k}^i - \hat{y}(k+1|k)]^T$$

Then, the Kalman gain (K) is computed as:

$$K_k = P^{xy}(k+1|k) P^{-1}(k+1|k) \tag{8}$$

Finally, the UKF estimation, and its covariance, is calculated as follows:

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K_k \mathbf{y} \tag{9}$$
$$P(k+1|k+1) = P_x(k+1|k) - K_k P_y(k+1|k) K_k^T$$

Algorithm 1 summaries the steps involved in the Unscented Kalman filter.

Algorithm 1 Unscented Kalman Filter

Initialization
Initialize the belief in the state
Prediction
Generate χ and W_m, W_c
Project χ to $f(\cdot)$ creating $\mathcal{Y} = f(\chi)$
Compute mean and covariance of the prior
$\bar{\mathbf{x}}_- = \sum W^m \mathcal{Y}$
$\bar{\mathbf{P}}_- = \sum W^c (\mathcal{Y} - \bar{\mathbf{x}}_-)(\mathcal{Y} - \bar{\mathbf{x}}_-)^T + \mathbf{Q}$
Update
Project $\mathcal{Y}$ to $h(\cdot)$ creating $\mathcal{Z} = h(\mathcal{Y})$
Compute mean and covariance of the posterior
$\bar{\mathbf{x}}_+ = \sum W^m \mathcal{Z}$
$\bar{\mathbf{P}}_+ = \sum W^c (\mathcal{Z} - \bar{x}_+)(\mathcal{Z} - \bar{\mathbf{x}}_+)^T + \mathbf{R}$
Compute residual $\mathbf{y} = \mathbf{z} - \bar{\mathbf{x}}_+$
Compute $\mathbf{K} = [\sum W^c (\mathcal{Y} - \bar{\mathbf{x}}_+)(\mathcal{Z} - \bar{\mathbf{x}}_+)^T] \mathbf{P}_+^{-1}$
Update mean and covariance of the UKF estimation
$\mathbf{x} = \bar{\mathbf{x}}_- + \mathbf{K}\mathbf{y}$
$\mathbf{P} = \bar{\mathbf{P}}_- + \mathbf{K}\bar{\mathbf{P}}_+ \mathbf{K}$

IV. Methodology

This section describes the proposed methodology used to create a *Grey-box model* that exploits UKF for predicting temperature trends in buildings. As highlighted in [15]–[17], thermal systems in buildings can be modelled and approximated through an electric RC circuit (see Figure 1). The main assumption is that a building is composed by a finite number of parts (n), called *nodes*. Each resistance between two nodes represent a wall, a window or the air mass. Figure 1 (a) shows a single node RC circuit of a building where R_{in} and C_{in} represent the resistance and capacity of all the element in the building itself. Figure 1 (b) shows a two node RC circuit where indoor air mass is between walls resistance and capacity. Furthermore, there may be a direct input heat source applied to each node (e.g. solar radiation, air-heating system and lights). Buildings can be represented by a *single-zone* RC circuit (see Figure 1 (a) and Figure 1 (b)) or with a *multi-zone* RC circuit (see Figure 1 (c)); this depends on the level of detail and aggregation to provide to the model. In a single-zone model the unique heat exchange will be with the external environment. Whilst, a multi-zone model considers also the heat exchange with the adjacent zones.

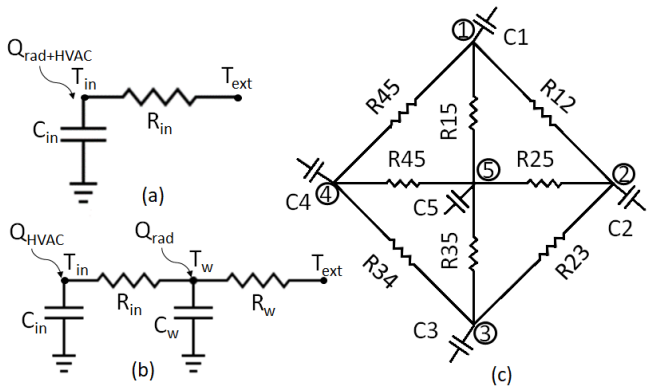

Fig. 1. Thermal networks representation. a) single-zone 1-node; b) single-zone 2-node c) multi-zone 1-node

The variation of temperature at a Node of zone i is determined by heat exchange with adjacent zone(s) j and by external/internal heat supply. This is modelled by the following equation:

$$C_i \frac{dT_i}{dt} = \sum_j \frac{T_j - T_i}{R_{i,j}} + A_i Q_{rad_i} + \dot{Q}_{int_i} \tag{10}$$

where C_i is the thermal capacity ($J/°C$) of zone i, $R_{i,j}$ is the thermal resistance ($°C/W$) between zone i and zone j, A_i represent the fraction of wall irradiated by solar radiation (m^2), Q_{rad_i} is the solar radiation (W/m^2) and Q_{int_i} the interior heat supply of zone i (W). Each zone can be divided in infinite sub-zones. A typical procedure to do so consists in considering also temperature of walls as an heat exchange source.

The UKF task is to train the model with a series of measurements and to estimate unknown parameters and disturbances. Heat transfer equations that describe the physical system yield the following set of continuous time differential equations:

$$\begin{aligned} \dot{x} &= f(x, u, d, p) \\ z &= h(x, u) \end{aligned} \tag{11}$$

where $f(\cdot)$ and $h(\cdot)$ are the same function described in Section III representing respectively state transition and measurement function; $x \in \mathbb{R}^n$ is the state vector representing the temperature of the state and $\dot{x}$ is the state temperature

rate of change; $u \in \mathbb{R}^m$ is the input vector (e.g., outdoor temperature); d is the disturbance vector (e.g. solar load, heater loads, infiltration) and p is the selected uncertain model parameters; $z \in mathbbR^m$ is the output vector of the system and represents the temperature of thermal zones. This model is used to reproduce the true building thermal evolution.

As described in [22], to allow the parameter estimation with UKF, the state x in Equation 11 has been augmented combining together temperatures $\bar{T} = [T_1 \cdots T_n]^T$ with uncertain parameters $\bar{p} = [(RC)_1 \cdots (RC)_k]^T$ and disturbances $\bar{d} = [A_w Q_{rad}/C \qquad Q_{int}/C]^T$. The result is an augmented state vector $\bar{x}$:

$$\bar{x} = \begin{bmatrix} \bar{T} \\ \bar{p} \\ \bar{d} \end{bmatrix}$$

Multiplying $\bar{p}$ and $\bar{d}$ parameters through the state variables $\bar{T}$ and u causes the state update function $f(\bar{x}(k), u(k))$ to becomes non-linear. This non-linearity is needed by UKF to predict the state equations. To implement the filter, Equation 11 is discretized with a 1-min interval steps Euler integration. The full discrete-time stochastic system becomes as follow:

$$\begin{aligned} \bar{x}(k+1) &= f(\bar{x}(k), u(k)) + \bar{\omega}_1(k) \\ \bar{p}(k+1) &= \bar{p}(k) + \bar{\omega}_2(k) \\ \bar{b}(k+1) &= \bar{d}(k) + \bar{\omega}_3(k) \\ \bar{z}(k+1) &= h(\bar{x}(k), u(k)) + \bar{v}(k) \end{aligned} \tag{12}$$

where $\bar{\omega}$ and $\bar{v}$ are process and measurements noise, respectively. They are assumed to be zero mean, multivariate, white gaussian with covariance Q and R, (i.e. $\bar{\omega} \approx N(0, Q)$ and $\bar{v} \approx N(0, R)$), respectively. In particular $\bar{\omega}_1(k)$ represents process noise for temperatures, $\bar{\omega}_2(k)$ represents estimation uncertainty in RC parameters and $\bar{\omega}_3(k)$ represents process noise for disturbances.

The measurement noise is correlated with the accuracy of the measure used to study the system. In our case, this is the accuracy of the temperature sensors. Temperature noise is correlated with the accuracy of the equation that define the state. The more the law is detailed and involves all the possible sources of heat, the smaller has to be the noise value. RC noise parameter must be set to a value greater than zero to allow the filter to vary its estimation over the time. Also disturbance process noise is a parameter, but it differs from the previous because the disturbance noise is considered only when disturbance have to be estimated. The noise level for RC parameters should be much smaller than for disturbances due to RC values that are constant over the time, while disturbance bias may change over the day.

In a first iteration, the filter runs only for night hours to have disturbance (both solar gains and HVAC systems) equal to zero, and to better estimate RC parameters. In a second iteration, the filter runs considering also day hours to estimate both RC parameters and disturbance. After these two iterations, all the RC parameters and disturbance patterns are used together to predict the indoor air-temperature trends for the next 24-hours. During first iteration, the disturbance process covariance Q is set equal to zero to not estimate disturbance parameters. Whilst, during the second iteration, Q is inflated to allow the disturbance parameters to change over the time and follow the real behavior of the phenomena.

Another important parameter to be initialized is the covariance matrix P. It represents the uncertainty on the $\hat{\mathbf{x}}_\mathbf{0}$ initial state estimation, and, in contrast to measurement noise covariance, it cannot be estimated experimentally. In this work, we apply the approach proposed in [25], [26] to initialize $\mathbf{P_0}$ as follows:

$$\mathbf{P_0} = diag((\hat{\mathbf{x}}_\mathbf{0} - \mathbf{x_0})^T (\hat{\mathbf{x}}_\mathbf{0} - \mathbf{x_0})) \tag{13}$$

The exact value of state $\mathbf{x_0}$ is rarely known in practice, but more often are known upper $(\mathbf{x_u})$ and lower $(\mathbf{x_l})$ bounds that are used to approximate $\hat{\mathbf{x}}_\mathbf{0}$. Using this approximation, it is possible to set $\hat{\mathbf{x}}_\mathbf{0} = 0.5(\mathbf{x_u} + \mathbf{x_l})$ and $\hat{\mathbf{x}}_\mathbf{0} - \mathbf{x_0} = 0.5(\mathbf{x_u} - \mathbf{x_l})$.

In this brief, both solar radiation and interior heat gain are treated as external disturbances. As a matter of fact, it is a frequent scenario to not have exhaustive information about their shape. On the contrary, it is common to have information about outdoor temperature profiles, so they have been treated as a known input.

V. Experimental Results

This section presents the experimental results of the proposed methodology i) in estimating thermal parameters of a building and ii) in predicting indoor air temperature trends up to future 24 hours. Tests are performed considering a mathematical solution of the investigated thermal networks and a building simulated with *Energy+*. For all the tests, weather conditions (i.e. outdoor temperature and solar radiation) are taken from the *Energy+* weather file for Turin.

The first test is conducted on a simple *1-node* thermal network (Figure 1 (a)) where temperature profile are created *ad-hoc* with Eulerian 1-minute discretization of Equation 10. This test-case is used to check the filter's performance on the estimation of the RC parameters. Figure 2 shows results on estimating the RC parameter. After two or three simulated days, RC parameters converges to the real value. After ten simulated days, it reaches a complete convergence.

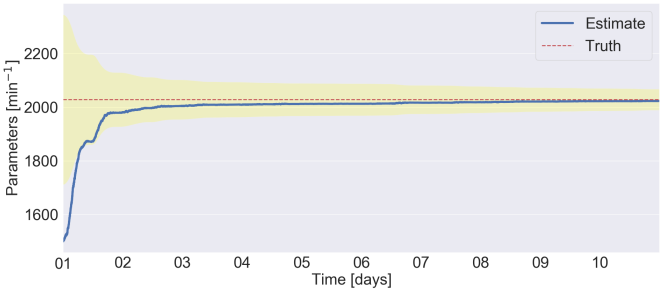

Fig. 2. Parameter Convergence

Figure 2 shows how the covariance associated to the *RC*-parameters converges. This is confirmed by the decreasing *P* trend that is the area highlighted in yellow in the plot. This is also confirmed by the covariances evolution of both temperature and parameters in Figure 3 that drop to zero in few simulation-steps.

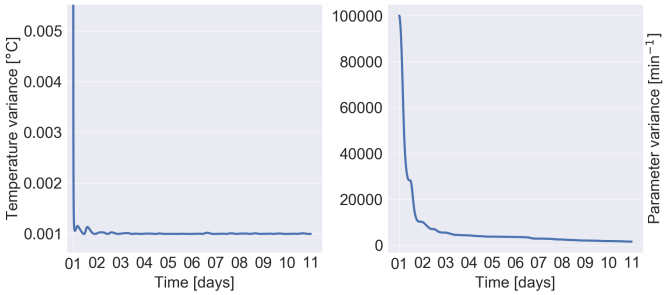

Fig. 3. Covariances Convergence

Furthermore, we tested the ability of our solution in predicting the indoor air temperature trends for the future 24 hours. We performed our test considering a "virtual and realistic building", which is an *industrial-like* building spreading over a base surface of $10\,000m^2$. The RC circuit is designed to dived the whole area into five thermal zones (see Figure 1 (c)). The first four are oriented towards the cardinal directions, while the fifth is located in the center of the building. To compare our results with realistic indoor temperature trends, we simulated the same virtual building with *Energy+*. The outdoor temperature trend is considered as the input of the thermal network (see Section IV). To take into account the effect of solar radiation on outdoor air temperature, we used the the $T_{sol-air}$ calculated as:

$$T_{sol-air} = T_{ext} + k \cdot GTI \tag{14}$$

where GTI is the incident global solar radiation on the tilted surface and k is the radiative loss factor of the surface. As presented by Bottaccioli et al. [27], $T_{sol-air}$ represents the "projection" of the external temperature over a surface considering the effect of solar radiation.

We have performed eight tests where, for each test, we gradually increased the complexity of the model as follows: *i)* considering solar radiation as disturbance; ii) considering solar radiation and HVAC as disturbance; iii) estimating together R and C; iv) estimating separately R and C. For each level of complexity, we performed the *1-node* circuit (considering only zone temperature as a node), and the *2-node* circuit (considering also wall temperature). To evaluate the accuracy of our results compared with results of *Energy+* simulations (our benchmark), we adopted the following performance indicators. i) *Mean Absolute Error (MAE)* is defined as the average of the absolute difference between the state estimation and the reference data. ii) *Root Mean Square Error (RMSE)* is the standard deviation of differences between state estimation and observed values. iii) *Correlation Coefficient* is a 0 to 1 number representing the strength of the correlation between the state estimation and observed values.

Figure 4 shows the prediction for the future 24 hours of indoor air temperature for *1-node* circuit. In this scenario, only solar radiation is considered as a disturbance and RC is estimated as coupled parameters. This plot highlights that the predicted air temperature follows with a good accuracy the benchmark trend.

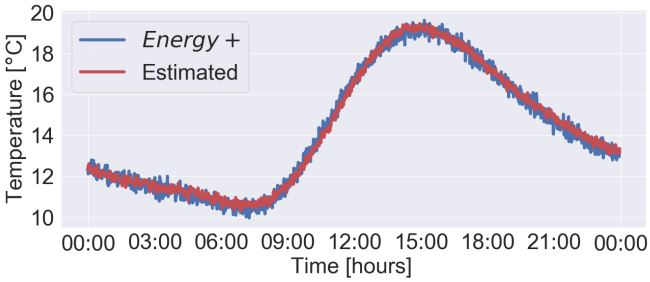

Fig. 4. 24-hours prediction with solar gain as disturbance for *1-node* circuit

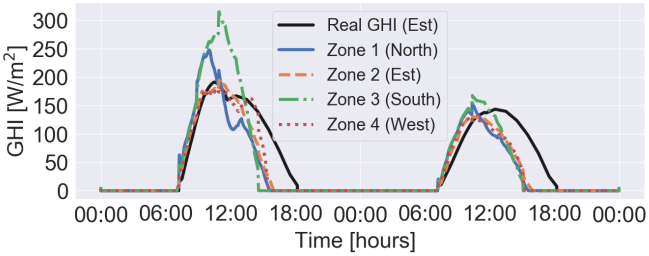

Fig. 5. Solar gain disturbance pattern of thermal network of the five zones and GHI of the same days

Figure 5 shows the estimated disturbances pattern due to solar gain in the thermal network of the zones. With the black line is reported, as a comparison, the Global Horizontal Irradiance (GHI). As shown in the plot, trends are very similar to the real solar radiation profile. The few differences are due to the different orientations.

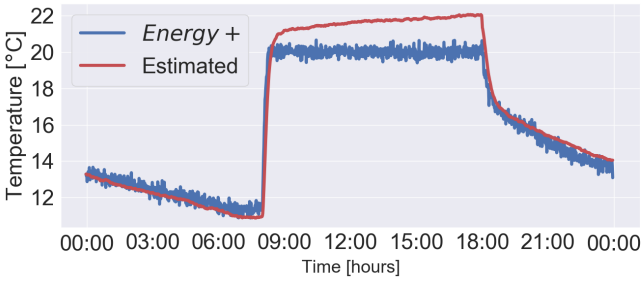

Fig. 6. 1-zone 24-hours prediction with solar gain and HVAC as disturbance

Figure 4 shows the prediction for the future 24 hours of indoor air temperature for a second *1-node* circuit. In this second scenario, solar radiation and HVAC are considered as a disturbance and RC is estimated as coupled parameters. In this

case, our solution estimates two different disturbance profiles, causing some instability in the recognition of disturbance pattern and decreasing the accuracy of the prediction. However also in this case, our forecasts follows with a good accuracy the benchmark trend.

Table I reports the performance indicator for all the eight tests we performed. In the table, GTI stands for solar radiation disturbance. CP and SP refers to RC estimated as coupled or separate parameters, respectively. The best performance is given by the combination of *1-node* CP circuits that considers only solar radiation as disturbance. However in case of *1-node* SP circuits, errors slightly get worst. Even if we introduce the HVAC system in our computations, the error rate increases but performance is still accurate. For *2-node* circuits, performance is slightly worse than *1-node* circuits. This is due to the temperature of walls that is an additional parameter to be estimated and not an input of the model as for *1-node* circuits.

TABLE I
PERFORMANCE INDICATOR FOR THE EIGHT TESTS

Test	RMSE [°C]	MAE [°C]	CORR
1-node, GTI, SP	0.464	0.384	99.2
1-node, GTI, CP	0.370	0.295	99.1
1-node, GTI+HVAC, SP	1.50	1.14	98.3
1-node, GTI+HVAC, CP	1.24	0.90	98.3
2-node, GTI, SP	1.16	1.01	98.4
2-node, GTI, CP	0.73	0.63	98.9
2-node, GTI+HVAC, SP	1.75	1.40	96.9
2-node, GTI+HVAC, CP	1.09	1.32	91,2

VI. CONCLUSION

In this paper, we presented a *Grey-box model* to predict indoor air temperature in buildings by learning their thermal characteristic and dynamics. The grey-box model is trained using the UKF to learn parameter of the thermal network and the disturbances due to solar gain and HVAC systems. First, we discuss the motivation that drive the development of such models and presented the literature solution for such problem. We than present the formulation of the UKF filter and the methodology used to test several thermal networks. The experimental results show the accuracy of the approach in forecasting indoor-air temperature trends. As future work, we plan to test this method with on filed measurements in complex buildings with more detailed thermal networks and compare the accuracy of different learning algorithms.

REFERENCES

[1] United Nations, "World Urbanization Prospects, Population Division." [Online]. Available: https://population.un.org/wup/

[2] United Nations, "Energy, UN-Habitat." [Online]. Available: https://unhabitat.org/urban-themes/energy/

[3] European Parliament, "Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings," 2010.

[4] L. Wigle, "How the internet of things will enable vast new levels of efficiency alan rose, intel corporation dr. subramanian vadari, modern grid solutions," 2014.

[5] P. Siano, "Demand response and smart gridsa survey," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 461–478, 2014.

[6] J. Lizana, D. Friedrich, R. Renaldi, and R. Chacartegui, "Energy flexible building through smart demand-side management and latent heat storage," *Applied energy*, vol. 230, pp. 471–485, 2018.

[7] J. L. Cremer, M. Pau, F. Ponci, and A. Monti, "Optimal scheduling of heat pumps for power peak shaving and customers thermal comfort," in *SMARTGREENS*, 2017, pp. 23–34.

[8] J. Kensby, A. Trüschel, and J.-O. Dalenbäck, "Potential of residential buildings as thermal energy storage in district heating systems–results from a pilot test," *Applied Energy*, vol. 137, pp. 773–781, 2015.

[9] F. Amara, K. Agbossou, A. Cardenas, Y. Dubé, and S. Kelouwani, "Comparison and simulation of building thermal models for effective energy management," *Smart Grid and renewable energy*, vol. 6, no. 04, p. 95, 2015.

[10] D. B. Crawley, L. K. Lawrie, C. O. Pedersen, and F. C. Winkelmann, "Energy plus: energy simulation program," *ASHRAE journal*, vol. 42, no. 4, pp. 49–56, 2000.

[11] C. Ellis, M. Hazas, and J. Scott, "Matchstick: A room-to-room thermal model for predicting indoor temperature from wireless sensor data," in *Proceedings of the 12th international conference on Information processing in sensor networks*. ACM, 2013, pp. 31–42.

[12] L. Bottaccioli, A. Aliberti, F. Ugliotti, E. Patti, A. Osello, E. Macii, and A. Acquaviva, "Building energy modelling and monitoring by integration of iot devices and building information models," in *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, vol. 1. IEEE, 2017, pp. 914–922.

[13] A. Aliberti, F. M. Ugliotti, L. Bottaccioli, G. Cirrincione, A. Osello, E. Macii, E. Patti, and A. Acquaviva, "Indoor air-temperature forecast for energy-efficient management in smart buildings," in *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2018, pp. 1–6.

[14] S. Royer, S. Thil, T. Talbert, and M. Polit, "Black-box modeling of buildings thermal behavior using system identification," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 10 850–10 855, 2014.

[15] R. Kramer, J. Van Schijndel, and H. Schellen, "Simplified thermal and hygric building models: A literature review," *Frontiers of architectural research*, vol. 1, no. 4, pp. 318–325, 2012.

[16] A. Bagheri, V. Feldheim, D. Thomas, and C. S. Ioakimidis, "Coupling building thermal network and control system, the first step to smart buildings," in *2016 IEEE International Smart Cities Conference (ISC2)*. IEEE, 2016, pp. 1–6.

[17] A. Bagheri, V. Feldheim, and C. Ioakimidis, "On the evolution and application of the thermal network method for energy assessments in buildings," *Energies*, vol. 11, no. 4, p. 890, 2018.

[18] P. Bacher and H. Madsen, "Identifying suitable models for the heat dynamics of buildings," *Energy and Buildings*, vol. 43, no. 7, pp. 1511–1522, 2011.

[19] D.-W. Kim and C.-S. Park, "Application of kalman filter for estimating a process disturbance in a building space," *Sustainability*, vol. 9, no. 10, p. 1868, 2017.

[20] S. F. Fux, A. Ashouri, M. J. Benz, and L. Guzzella, "Ekf based self-adaptive thermal model for a passive house," *Energy and Buildings*, vol. 68, pp. 811–817, 2014.

[21] M. Maasoumy, B. Moridian, M. Razmara, M. Shahbakhti, and A. Sangiovanni-Vincentelli, "Online simultaneous state estimation and parameter adaptation for building predictive control," 2014.

[22] P. Radecki and B. Hencey, "Online model estimation for predictive thermal control of buildings," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 4, pp. 1414–1422, 2017.

[23] E. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," vol. 153-158, 02 2000, pp. 153 – 158.

[24] S. Banani and M. Masnadi-Shirazi, "A new version of unscented kalman filter," *Proceedings of World Academy of Science, Engineering and Technology*, vol. 20, pp. 192–197, 01 2007.

[25] J. Valappil and C. Georgakis, "Systematic estimation of state noise statistics for extended kalman filters," *AIChE Journal*, vol. 46, no. 2, pp. 292–308, 2000.

[26] A. Jazwinski, "Stochastic process and filtering theory, academic press," *A subsidiary of Harcourt Brace Jovanovich Publishers*, 1970.

[27] L. Bottaccioli, A. Estebsari, E. Pons, E. Bompard, E. Macii, E. Patti, and A. Acquaviva, "A flexible distributed infrastructure for real-time cosimulations in smart grids," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3265–3274, 2017.