University of Liège

Big Data project

PROJ0016

Milestone 1: Literature review

Julien Gustin, Romain Charles, Joachim Houyon

September 30, 2021

1 Literature Review

1.1 Scientific publication 1: "Gray-Box Approach for Thermal Modelling Of Buildings for Applications in District Heating and Cooling Networks"

1.1.1 Contextualisation

This article aims to build a Gray-box for thermal modelling of buildings for applications in district heating and cooling networks. Gray-box means that the model will use basic equations taken from the physics based models and we will learn some coefficients using data.

1.1.2 Grey-box building modelling

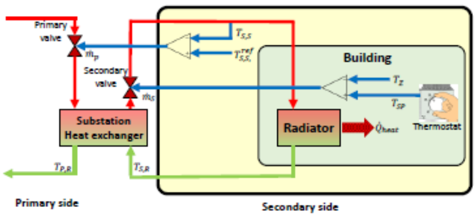

Figure 1: Grey-Box model of the building

Here is the grey-box building modelling. Individual components are categorized as “primary-side” for the ones which connect the building with the thermal grid while all the components contained in the building are classified as the secondary-side components.

1.1.3 Indoor zone temperature

Buildings can be modelled as lumped circuit elements having one resistance R_z and capacitance C_z. Indeed, while the resistance represents the thermal loss and the capacitor represents the total heat capacity of the building. The dynamic evolution of the internal temperature, using a time step of size Δt seconds, can be written as:

$$\frac{T_z(t+1) - T_z(k)}{\Delta t} = \frac{T_\infty(t) - T_z(t)}{R_z C_z} + \frac{Q_{heat}(t)}{C_z} + \frac{Q_{int}(t)}{C_z}$$

where T_∞ is the ambient temperature, T_z is the zone temperature, Q_{heat} is the heat supplied by the HVAC system and Q_{int} is the rate of heat transfer to the building from internal sources.

While T_z, T_∞ and Q_{heat} are known at time instance k, R_z, C_z and Q_{int} can be learned through time-series regression from historical data. Regarding Q_{heat}, we can get its value thanks to the first law of thermodynamics stating that heat lost by the water should be the heat transferred to the zone.

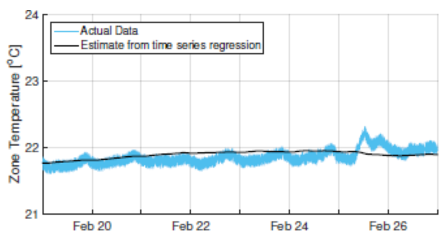

Figure 2: Learning R_z & C_z using time series regression

1.2 Scientific publication 2: "Identification of Multi-Zone Grey-Box Building Models for Use in Model Predictive Control"

1.2.1 Contextualisation

A way is sought to improve the performance of building energy systems using predictive controllers. Therefore, an efficient building model for optimisation is needed: the grey-box model.

For *multi-zone grey-box models*, the challenge is greater as the number of parameters to be estimated increases. However, they are able to output an estimated temperature per zone with the advantage that they can be used by a predictive controller to optimize the heating inputs for each zone individually. To identify multi-zone models, there are 2 main approaches:

- **The decentralized approach** (*decentralized model*): Identifies a single-zone grey-box model for each zone independently of the others. This approach does not consider heat interactions between zones. The accuracy of the models could therefore be affected.
- **The centralized approach** (*centralized model*): Identifies a multi-zone grey-box model directly. Heat interactions between zones are taken into account.

1.2.2 Description of the used method: Centralized approach

Inputs: zone volumes, physical connections and historical data (indoor temperatures and thermal powers).

The complexity of the problem is divided by zones (decentralized approach) keeping the temperatures of adjacent zones as exogenous variables (centralized approach). Then the zones are aggregated and the results of each sub-problem are used as an educated initial guess to estimate the parameters of the fully merged model.

An RC architecture is used in the forward selection without adjacent zone temperature input. This selection compares the results of models of increasing complexity for each individual zone. The model complexities range from 1 to 4 thermal states: the zone air, wall, internal walls and embedded pipes temperatures. The parameters to be estimated are the thermal capacitances, the thermal resistors and the solar transmittance.

In the centralized approach, the physical connection of the zones is introduced using a sparse convex matrix. This information allows each zone to be optimized with the temperatures of the neighbouring zones as input at each stage of the forward selection. A pure resistance branch R1C0 is used to represent the coupling effect between 2 zones. For the thermal resistors of this coupling, the average of the temperature values obtained from each zone is taken.

For each model, 2 validation tests are made to filter out models that do not make sense from a physical point of view: the first test checks whether the heat flow is used dynamically and the second checks whether very large or very small capacities are estimated.

The Root Mean Square Error (RMSE) of the single-zone model attempt is calculated in auto and cross-validation to estimate the model accuracy of each valid z-zone grey-box model attempt. The one with the smallest RMSE in auto-validation is chosen as the representative of the thermal zone.

1.2.3 Comparison with the other methods

The methodology used is compared in the following to a decentralized model and a single-zone model. When comparing, it is noticed that in terms of:

- *Simulation performance*:
  The centralized model slightly outperforms the decentralized model and is similar to the single-zone model.

- *Control performance*:
  The centralized model does not overestimate the temperature in any zone (minimum discomfort) while the decentralized model shows the worst comfort. The single-zone model is outperformed by the centralized model. However, although a perfect hydraulic balance is assumed, the performance of the single-zone model is good.

Therefore, for multi-zone buildings, the thermal interactions between the zones must be modelled. Single-zone models may also be suitable under the assumption mentioned above.

2 Model proposal

2.1 Given data (Input)

T_z: Air temperature of each zone [°C]
T_a: Ambient temperature [°C]
T_s: Temperature set [°C]
T_w: Water temperature [°C]
House measurement

2.2 Way of modeling

As explained in the different articles, they are mainly 3 ways for thermal modelling of buildings.

- **White-box models**: It is pure physics based models

- **Black-box models**: Pure data driven methodology

- **Grey-box models**: Mix between the white and black model (physics and data).

2.3 Our first *simple* model

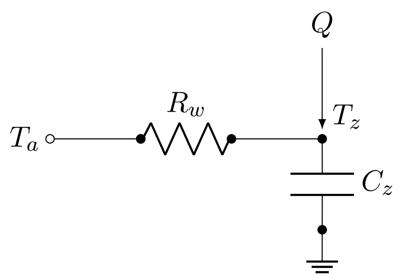

Figure 3: 1R1C Architecture

As a first *simple* model we decided to use a **Grey-box models**. This allows to represent buildings as a lumped circuit as shown in **Fig. 3** . Elements consist of thermal resistance R [K/W] (for thermal loss) and thermal capacities C [J/°C] (Heat capacity of the building), this framework allows to easily deal with thermodynamics applied to the house.

In the literature most of the time[1] $Q = Q_{heat} + Q_{int}$, where Q_{heat} [W] represent the heat transfer from the radiator to the building and Q_{int} [W] the rate of heat transfer from internal sources such as occupants, lighting and appliances. However using only data that we have ***2.1*** it is *impossible* to calculate Q_{heat}. Thus these two parameters would need to be estimated from given data.
From this circuit (**Fig. 3**) and by using the first law of thermodynamics[2] we can derive the following equations

$$\frac{dT_z(t)}{dt} = \frac{T_a(t) - T_z(t)}{R_w C_z} + \frac{Q(t)}{C_z} \tag{1}$$

Our data are given in a discrete time, with a time step of Δt seconds equals to 5s. Thus we can approximate equation (1) by (2)

$$\frac{T_z(t) - T_z(t-1)}{\Delta t} = \frac{T_a(t) - T_z(t)}{R_w C_z} + \frac{Q(t)}{C_z} \tag{2}$$

Using this equation we can now predict the temperature. However we have a bunch of parameters (C_z, R_w, $Q(t)$) that are unknown and need to be estimate. (Milestone 2).

2.4 Dealing with multi-zone

We will use the methodology of **1.2** to deal with multi-zone, saying that we will deal each *zone* as a single zone but keeping the thermal influence of adjacent zones using a resistance $R_{z1,z2}$ to represent the thermal effect of one zone to an other.

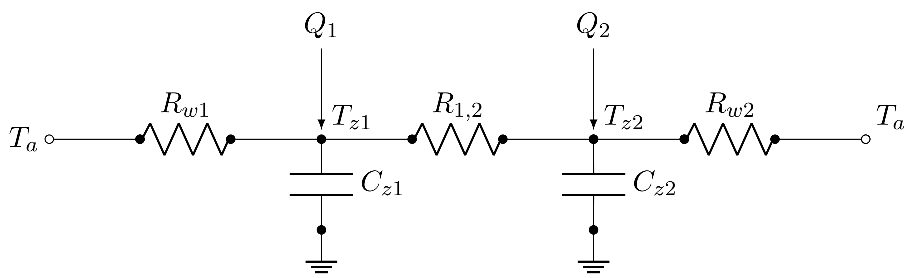

Figure 4: 1R1C Architecture for multi-zone, with zone 1 = "Salle à manger" and zone 2 = "Cuisine"

As an example **Fig. 4** is how we model a multi-zone building. (Here we only take into account the "salle à manger" and "cuisine" from our case). And this can be translate as the following equations.

$$\frac{dT_{z1}(t)}{dt} = \frac{T_a(t) - T_{z1}(t)}{R_{w1} C_{z1}} - \frac{T_{z1}(t) - T_{z2}(t)}{R_{1,2} C_{z1}} + \frac{Q_1(t)}{C_{z1}} \tag{3}$$

$$\frac{dT_{z2}(t)}{dt} = \frac{T_{z1}(t) - T_{z2}(t)}{R_{1,2} C_{z2}} - \frac{T_{z2}(t) - T_a(t)}{R_{w2} C_{z2}} + \frac{Q_2(t)}{C_{z2}} \tag{4}$$

[1] Sometime using also the solar heat gains, but as we have no clue about any window it is not that interesting
[2] Change in zone temperature should be proportional to the net neat gain/loss and the proportionality constant is the thermal capacitance or the heat capacity of the zone. (cf. **1.1**)