UNIVERSITY OF LIÈGE

BIG DATA PROJECT

PROJ0016

Milestone 2: Parameter Estimation

Julien GUSTIN, Romain CHARLES, Joachim HOUYON

October 19, 2021

1 Data

In order to deal with N/A data when merging all data together using the date as reference, we have just replaced them by the mean of the according data.

2 Single-zone Case

2.1 Adaptation of the equation

Our equation was previously the following:

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

However, this formulation is complex since we have to isolate $T_z(t)$ as it appears twice in the right member ($Q(t)$ includes $T_z(t)$, see next subsection). Therefore, we decided that, since the temperature doesn't change a lot between two measures, Our data at time t will be used in order to predict the temperature at time $t+1$, leading us to the following equation:

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

2.2 Heat transferred from the radiator $Q(t)$[1]

$Q(t)$ is not given and needs to be estimated. We do not know the water mass flow rate at time t and we also do not know the temperature of the water leaving the radiator. Our idea is to express the heat transferred from the radiator as a linear function of the difference between the water temperature and the zone temperature:

$$\hat{Q}(t) = cp * m(T_w(t) - T_z(t)) + a \tag{3}$$

Where m and a are unknown parameters that need to be estimated. And cp the heat capacity of water (4810 [kJ/kg/°C]).

2.3 Parameters estimation

We want to find R_z, C_z, m and a such that

$$\hat{T}_z(t+1) = \left(\frac{T_a(t) - T_z(t)}{R_w C_z} + \frac{cp * m(T_w(t) - T_z(t)) + a}{C_z} \right) \times \Delta t + T_z(t) \tag{4}$$

$$\text{and} \quad \delta_z = \sum_t^N (T_z(t+1) - \hat{T}_z(t+1))^2$$

is minimized. $T_z(t+1)$ is the expected output at time $t+1$ and $\hat{T}_z(t+1)$ is the predicted temperature. The noise is assumed to be Gaussian. To minimize δ_z, we apply the *least square method* with the *Trust Region Reflective (TRF)* algorithm in order to find the right values for our parameters.

2.4 Results with the mean temperature of the rooms

Using the root mean square error scoring, the mean difference between the predicted temperature and the real temperature is 1.69 degrees.

[1] https://adgefficiency.com/energy-basics-q-m-cp-dt/

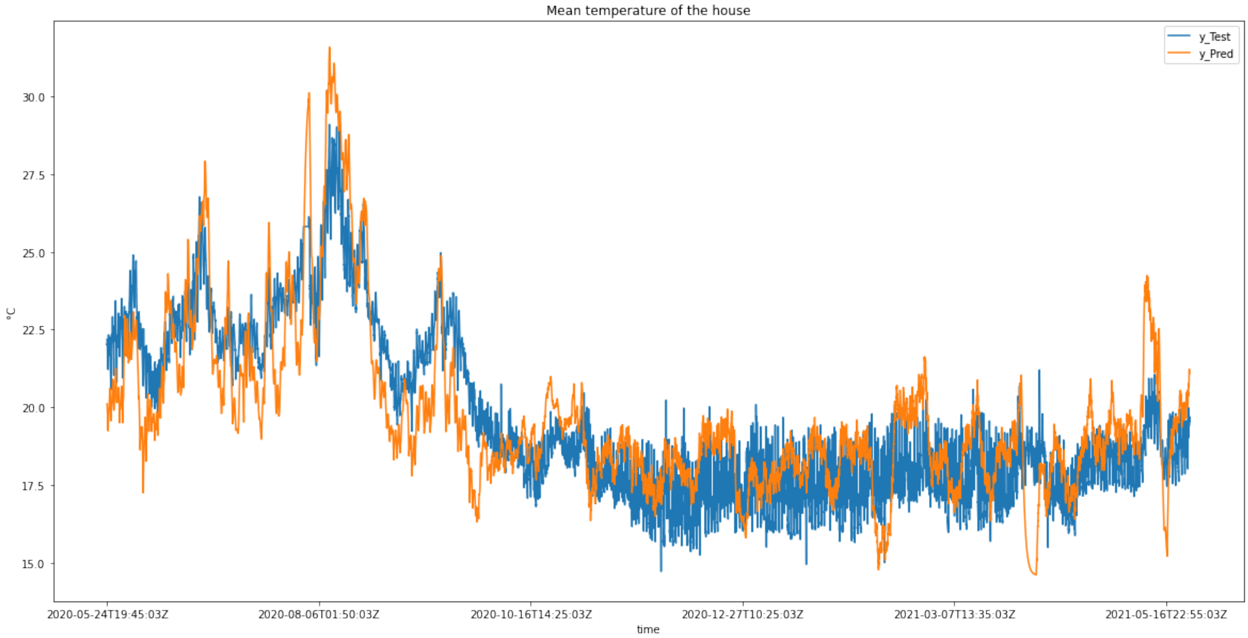

Figure 1: Least square fitting with a learning set of size 75 000

3 Multi-zone Case

3.1 Assumption

To make the problem more tractable we need to make some assumptions:

1. Halls are not considered
2. Rooms without data are collapsed together with the nearest room

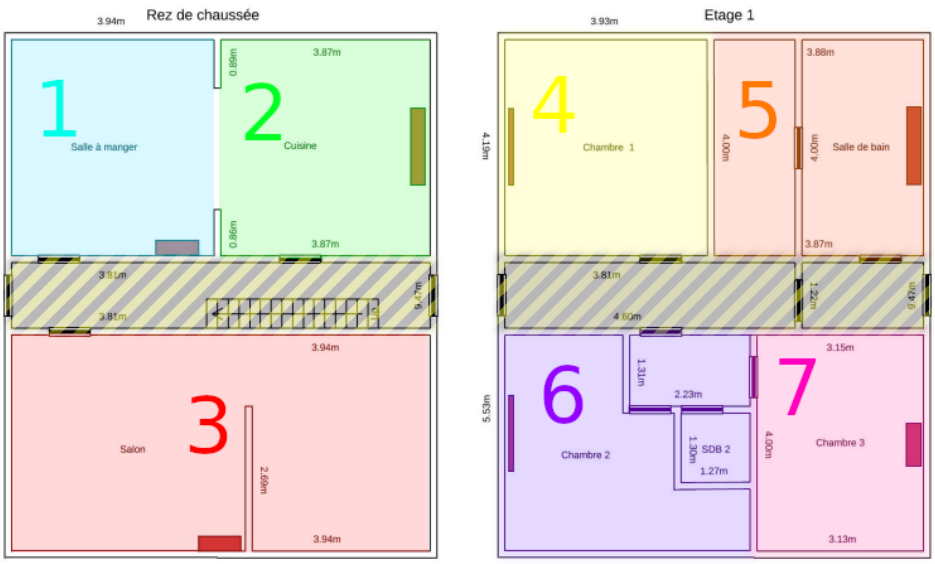

Figure 2: Building schematic, shaded areas are considered non-existent

3.2 Equations

From the previous milestone we know that we can derive equations from circuit that model the heat transfer applied to the house. Here is the equation for the multi-zone case.

$$\frac{T_{zx}(t+1) - T_{zx}(t)}{\Delta t} = \frac{T_a(t) - T_{zx}(t)}{R_{wx}C_{zx}} + \frac{Q_x(t)}{C_{zx}} - \sum_{n \in neighbours} \frac{T_{zx}(t) - T_{zn}(t)}{R_{x,n}C_{zx}} \tag{5}$$

We can now re-use what we have just explained in **2** but adding the thermal influence of adjacent rooms and the shared resistance $R_{x,n}$.

3.3 Parameters estimation

To find the best value for parameters $C_x, R_x, R_{x,y}, m_x, c_x$ we used the **least square method** as explain in **2.3** but using equation **(5)** for computing $\hat{T}_{zx}(t+1)$

$$\delta_{zx} = \sum_{t}^{N} (T_{zx}(t+1) - \hat{T}_{zx}(t+1))^2$$

3.3.1 Shared parameters

But there is still an issue, the value of resistance linking two adjacent rooms (x and y for instance) $R_{x,y}$ is dependent to two optimisation problem, minimizing δ_{zx} and δ_{zy}. Therefore what we have done is to simply merge these two optimization problems in one such that we want to minimize

$$\sum_{n \in G} \delta_{zn} \tag{6}$$

Summing all residual of dependent room into one minimization, and to be more efficient we represent each room as a node for which the resistance $R_{x,y}$ is a link between two nodes, that way we minimize the residual **(6)** for each connected components G separately.

3.4 Results

	Multi-zone-modified		Single-zone	Multi-zone
Room	**RMSE °C**	**Adjacent**	**RMSE °C**	**RMSE °C**
Dining room	2.2144	{Kitchen, Bedroom1}	2.2785	2.2144
Kitchen	2.1381	{Dining room, Bathroom}	2.2285	2.1381
Living room	2.1587	{∅}	2.1587	2.1576
Bedroom 1	1.5749	{Dining room, Bathroom}	1.5707	1.5759
Bathroom	2.0378	{Kitchen, Bedroom1}	2.3109	2.0378
Bedroom 2	1.3139	{Bedroom3}	1.2834	1.4177
Bedroom 3	1.5810	{Bedroom2}	1.6563	1.5832
TOTAL	13.0188		13.4870	13.1237

Table 1: Result with data fit using 75000 samples as training set and RMSE compute on the test set. (Multi-zone consider all adjacent zones)

To best compare results, we tried different configuration of our model, **Single-zone** consider each room as a single-zone model, **Multi-zone** consider all adjacent room as having a thermal influence (1->2, 1->4, 2->5, 3->6, 3->7, 6->7) and finally **Multi-zone modified** is the configuration that give the best result when summing the RMSE of the test set for each room. And results are quite good because as expected when considering thermal influence of adjacent room the model is more accurate.

Figure 3: Results of applying least-square method for parameter estimation, using a training set of size 75000 and plotted using the whole data set. Using *Multi-zone modified* model

When looking at the graph we see that results are quite accurate and do not over-fit much the training set, obviously the model is not perfect due to noise and assumption but however our model seems to fit quite well the curve.

Room (z)	C_z	R_z	R_xy	m	a
Dining room	8.2920	8123.1533	{1331.21, 262427.69}	$1.8727 * 10^{-08}$	-0.0002
Kitchen	48.6936	1541.8719	{1331.21, 1174.47}	$1.0688 * 10^{-07}$	0.0007
Living room	633.1925	267.7173	$\{\emptyset\}$	$3.7903 * 10^{-07}$	0.0057
Bedroom 1	262516.8167	0.7772	{262427.6, 7.36}	0.0001	2.9750
Bathroom	6100.5725	15.9193	{1174.47, 7.36}	$8.6344 * 10^{-06}$	0.0708
Bedroom 2	599.7513	568.9402	{265.28}	$1.1343 * 10^{-07}$	0.0047
Bedroom 3	948.4629	251.6712	{265.28}	$3.3457 * 10^{7}$	0.0127

Table 2: Estimated value of unknown parameters, R_{xy} values are sort in the same order as the one of *Adjacent* column from **Table 1.**