# Data-driven control strategies to enhance energy performance of HVAC system in building

## Case study "E" by Enertech


Armin Hooman

Keyvan Delfarah

# Contents

# Introduction

In order to optimize energy usage, raise comfort levels, and save operational costs, buildings' energy systems must be accurately modeled and controlled. The calibration of an RC (resistance-capacitance) model, which simulates the dynamic behavior of the understudied building, is presented in this article. Additionally, we investigate the Functional Mock-up Interface (FMU)'s ability to connect Simulink to EnergyPlus in order to create and assess two controllers: a basic controller that uses Proportional-Integral-Derivative (PID) control, and an advanced Model Predictive Controller (MPC).
In order to effectively represent the thermal behavior of the building, the RC model needed to have its parameters identified and adjusted. By comparing the simulated results with measured data, we ensured that the model could effectively represent the building's response to external disturbances, occupancy patterns, and heating/cooling system operation.
To enable dynamic simulation and control, we established a connection between Simulink, a powerful control design and simulation environment, and EnergyPlus. The FMU technology facilitated the communication between these software tools, allowing us to exchange information seamlessly and create a closed-loop control system.
We implemented a baseline controller using the PID control strategy as a starting point. This controller aimed to maintain the building's temperature within a desired setpoint range while minimizing energy consumption and maintaining occupant comfort.
We also created an MPC, that leveraged the RC model and predictive optimization algorithms to generate optimal control actions based on future predictions of the building's thermal behavior. This controller aimed to improve energy efficiency further and enhance occupant comfort.
In this report, we will discuss the calibration process of the RC model, the integration of Simulink with EnergyPlus through FMU, the design and implementation of the baseline PID controller, and the development of the advanced MPC. Furthermore, we will present simulation results and evaluate the performance of the controllers in terms of energy consumption, thermal comfort, and control effectiveness.

# Final calibrated reduced order model

The reduced order model was developed through a rigorous calibration process, where the parameters of the model were adjusted to accurately capture the building's thermal dynamics. By comparing the simulated results with measured data we ensured that the model faithfully represented the behavior of the actual building under different operating conditions. The approach used to simulate the reduced model integrated into the MPC controller is discussed in the section that follows.

## Final RC model

In the previous report, a first attempt for the RC circuit was created. Then, the model has been changed in order to better represent the building dynamics; the final circuit is shown in Figure 1.

## Final state space representation

Obtaining the state equations that characterize the system's behavior in the time domain after having an RC model is the next stage. The following states were used for the state space equations:

- Temperature of classrooms $T_1$
- Temperature of gym $T_2$
- Temperature of auditorium $T_3$

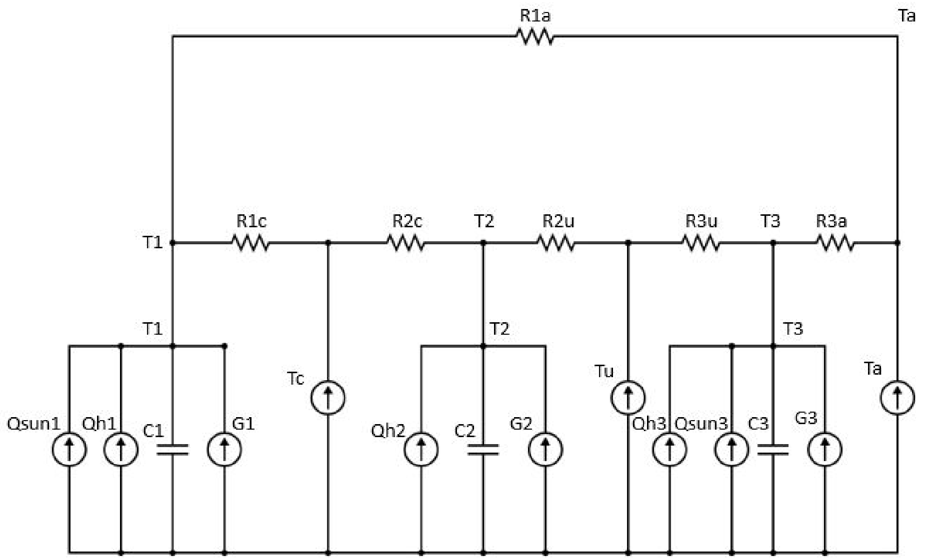


Figure 1: RC circuit model

We may acquire three sets of equations because there is only one state to be determined in each of our three zones (three ODEs of first order):

$$c_1\dot{T}_1 = \dot{Q}_{h_1} + \dot{G}_1 + k_1\dot{Q}_{s_{\text{sun}}} + \frac{T_c - T_1}{R_{1_c}} + \frac{T_a - T_1}{R_{1_a}}$$

$$c_2\dot{T}_2 = \dot{Q}_{h_2} + \dot{G}_2 + \frac{T_u - T_2}{R_{2_u}} + \frac{T_a - T_2}{R_{2_a}}$$

$$c_3\dot{T}_3 = \dot{Q}_{h_3} + \dot{G}_3 + k_3\dot{Q}_{s_{\text{sun}}} + \frac{T_u - T_3}{R_{3_u}} + \frac{T_c - T_3}{R_{3_c}}$$

A more condensed form of this equation system that separates the states from the input values is as follows:

$$\dot{X} = AX + BU$$

$$Y = CX + DU$$

where dynamics matrix A and control matrix B are respectively:

$$A = \begin{bmatrix} -\frac{1}{c_1}\left(\frac{1}{R_{1a}} + \frac{1}{R_{1c}}\right) & 0 & 0 \\ 0 & \frac{-1}{c_2}\left(\frac{1}{R_{2u}} + \frac{1}{R_{z_a}}\right) & 0 \\ 0 & 0 & \frac{-1}{c_3}\left(\frac{1}{R_{3a}} + \frac{1}{R_{3u}}\right) \end{bmatrix}$$

and

$$B = \begin{bmatrix} \frac{1}{c_1} & 0 & 0 & \frac{1}{c_1R_{1a}} & \frac{1}{c_1R_{1c}} & 0 & \frac{1}{c_1} & 0 & 0 & \frac{k_1}{c_1} & 0 \\ 0 & \frac{1}{c_2} & 0 & 0 & \frac{1}{c_2R_{2c}} & \frac{1}{c_2R_{2u}} & 0 & \frac{1}{c_2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{c_3} & \frac{1}{c_3R_{3a}} & 0 & \frac{1}{c_3R_{3u}} & 0 & 0 & \frac{1}{c_3} & 0 & \frac{k_3}{c_3} \end{bmatrix}$$

## Data collection and testing periods

Once the RC model is created it is necessary a calibration phase, in order to better follow the behavior of the building according to its model in EnergyPlus. This phase allows also to find the parameters of the differential equations. In particular, a first attempt is made to calculate the values of the resistances and capacitances involved in the RC model. As explained in the previous report, a synthetic dataset is created in order to calibrate the RC model. The dataset contains the results of the EnergyPlus model from the 8th of January to the 19th of February. In this time-lapse, the heating system schedule is changed every day in order to train the model with different scenarios. To obtain a better calibrated model the inside temperature is kept around the set-point temperature. In this way, there are no frequent and too much high variations in the temperature. The final dataset has the power provided by the heating system, the power of the internal loads, the inside temperature in the different thermal zones and the outside temperature. To improve the results, the temperature in the corridors is set as constant (18 °C). In the last model, the nominal power for the heating system is set as 7 kW for the auditorium, 75 kW for classrooms and 43 kW for the gym. These results arrive from the first model in which the ideal HVAC was considered. During the calibration phase, these values were changed in order to obtain better results.
The heating system's schedule also changes throughout the calibration period. The final schedule for the auditorium is shown in Figure 3. For both classrooms and the gym, the schedules are reported in the appendix.

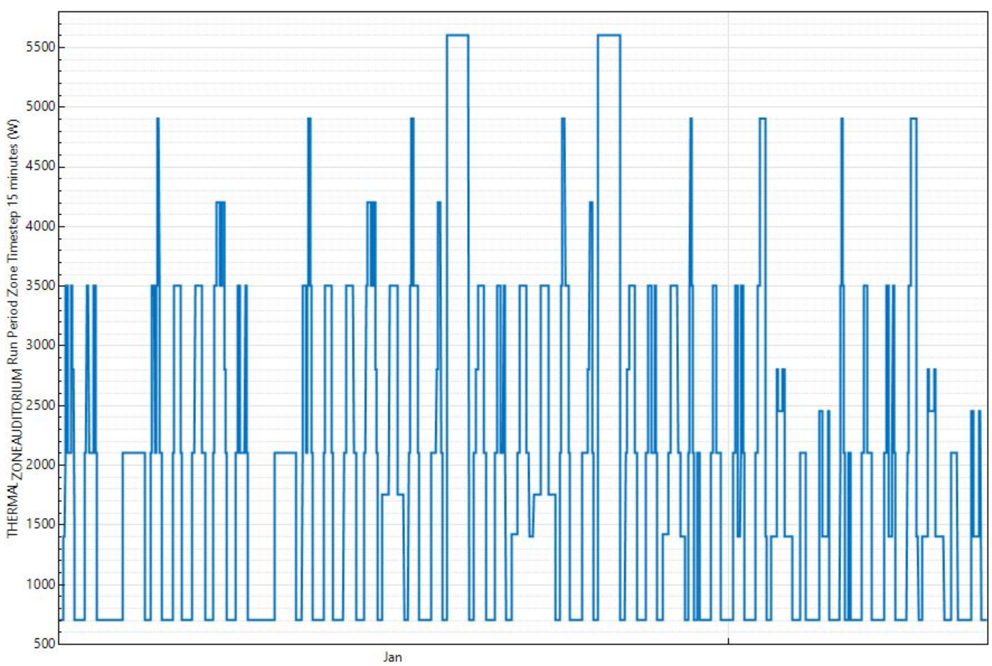


Figure 2: Schedule for the heating system of the auditorium

Furthermore, in the next figure (Figure 3) the trend of auditorium internal temperature is shown. The internal temperature of classrooms and gym is provided in the Appendix. This graph could be useful for a comparison with the results that will be obtain with the simulation in Simulink.

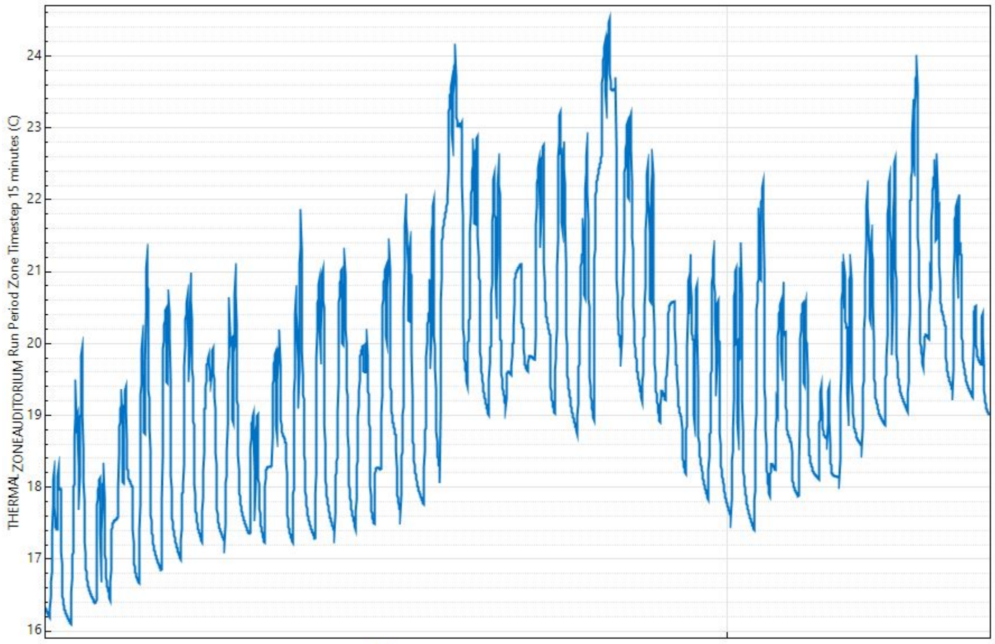


Figure 3: Auditorium internal temperature trend

## Parameters identification

Prior to initiating the parameter identification, a first estimate must be made for the parameters specified in the preceding section including capacitances and resistances in order to reach A and B matrices. This leads to a well-calibrated parameter identification.
The resistances are computed as the mutuals of the transmittance:

- $R_{1a}$ is the resistance between classrooms and outside ambient
- $R_{1c}$ is the resistance between classrooms and corridors
- $R_{2c}$ the resistance between gym and corridors
- $R_{2u}$ is the resistance between gym and unconditioned zones
- $R_{3a}$ is resistance between the auditorium and ambient
- $R_{3u}$ between the auditorium and unconditioned zones

To compute the resistances it is necessary to consider all the walls' layers as resistors in series. Instead, for the capacitances all the elements are considered as capacity in parallel. At first, nine parameters including thermal zones-unconditioned spaces resistances, thermal zones-ambient space resistances, thermal zones-corridor resistances and three thermal zone capacitances have been defined based on stratigraphy of the opaque and transparent components according to the following formulas:

$$R = \frac{1}{h_i} + \sum_j \frac{s_j}{\lambda_j} + \sum_k R_k + \frac{1}{h_e} \quad [m^2 \times K/W]$$

In which resistance of each component is affected by internal surface convective heat transfer coefficient,internal layers thickness and thermal conductivity, and also external surface convective heat transfer coefficient.The above formula provides the resistance of each thermal zones components in $(m^2 \cdot K/W)$. Thus, in order to calculate the total resistance between thermal zones and ambient temperature/unconditioned spaces the indicated formula is utilized to consider also the area:

$$R_{\text{tot}} = \frac{1}{\sum_{i=1}^{n} A_i U_i} \quad [K/W]$$

Regarding the capacitance, the thermal capacitance of each thermal zone is calculated by summation of different component's capacitance. It should be noted that the capacitance of each element is acquired through the following formula:

$$C_{eq} = \sum_{i=1}^{n} \rho_k c_{p,k} s_k \quad [J/m^2 \times K]$$

$$C_{\text{tot}} = \sum_{i=1}^{n} C_i A_i \quad [J/K]$$

In the above formula, capacitance is calculated through multipication of density, specific heat capacity and thickness of different layers at each transparent and opaque component. Finally, total capacitance of each thermal zone is acquired through multiplying each component equivalent capacitance with its area as the capacitances are assumed to have parallel effect.

The parameter identification process for the RC model involves obtaining the necessary data for training and testing the model's parameters. In our study, we adopted a data-driven approach, utilizing measured data from the building under investigation in Energyplus. To ensure robustness and accuracy, we divided the available dataset into two distinct sets: a training set comprising 75% of the data (32 days) and a testing set containing the remaining 25% (11 days). The training set was used to estimate the model's parameters, while the validation set was used to evaluate the model's effectiveness and determine how well it would perform when applied to new data. By employing this partitioning strategy, we aimed to capture the fundamental dynamics of the structure while reducing the risk that the model would become overfit to the particular training dataset.
The outcomes of the system identification process are shown in Table 1, where the parameters' estimated values – those deduced from physical considerations – are listed in the third column, while the identified parameters as a result are shown in the last column. The new A and B matrixes are:

$$A = \begin{bmatrix} -0.758 & 0 & 0 \\ 0 & -0.387 & 0 \\ 0 & 0 & -1.025 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 8.807e^{-5} & 0 & 0 & 0.124 & 0.633 & 0 & 8.807e^{-5} & 0 & 0 & 8.807e^{-5} & 0 \\ 0 & 1.978e^{-4} & 0 & 0 & 0.361 & 0.027 & 0 & 1.978e^{-4} & 0 & 0 & 0 \\ 0 & 0 & 2.675e^{-4} & 0.039 & 0 & 0.986 & 0 & 0 & 2.675e^{-4} & 0 & 2.675e^{-4} \end{bmatrix}$$

| Parameter | Unit | Guessed value | Estimated value |
|---|---|---|---|
| $c_1$ | (kJ/K) | $1.1355 \times 10^4$ | $1.1355 \times 10^4$ |
| $c_2$ | (kJ/K) | $5.0566 \times 10^3$ | $5.0566 \times 10^3$ |
| $c_3$ | (kJ/K) | $3.7382 \times 10^3$ | $3.7382 \times 10^3$ |
| $R_{1a}$ | (K/kW) | $7.0856 \times 10^{-3}$ | $7.0856 \times 10^{-4}$ |
| $R_{1c}$ | (K/kW) | $1.3906 \times 10^{-3}$ | $1.3906 \times 10^{-4}$ |
| $R_{2c}$ | (K/kW) | $2.0252 \times 10^{-3}$ | $5.4833 \times 10^{-4}$ |
| $R_{2u}$ | (K/kW) | $7.4776 \times 10^{-3}$ | $7.4133 \times 10^{-3}$ |
| $R_{3a}$ | (K/kW) | $6.9090 \times 10^{-4}$ | $6.9090 \times 10^{-3}$ |
| $R_{3u}$ | (K/kW) | $2.4752 \times 10^{-3}$ | $2.7126 \times 10^{-4}$ |

Table 1: Parameter estimation results for the RC model

After the acquisition of the parameter estimation for resistances, capacitances and consequently A,B matrix elements, we employed the *idgrey* and *greyest* functions available in MATLAB for the identification of the model parameters. These functions utilize the grey-box modeling approach, which combines the strengths of both black-box and white-box modeling techniques.

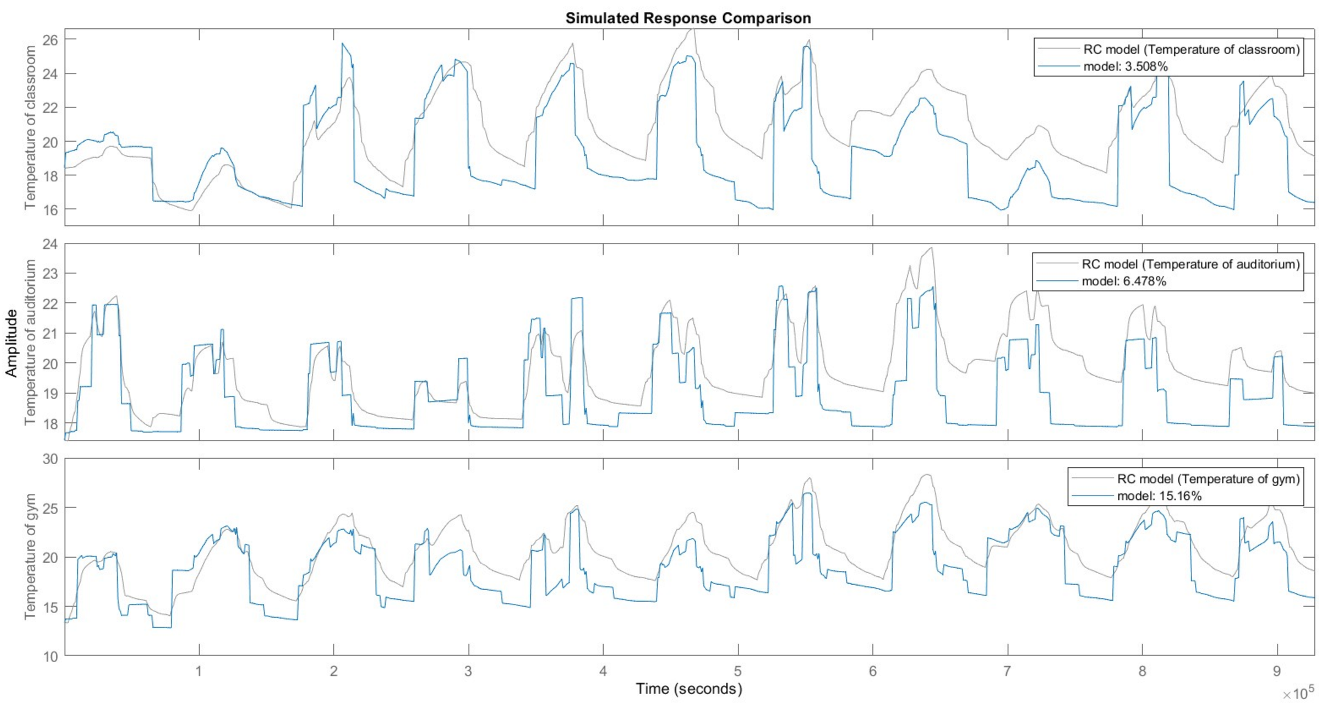


Figure 4: Results obtained with the *comapare* function

In order to evaluate the performance of the identified model, we used the *compare* function in MATLAB, which allows us to compare the simulated response of the identified model with the measured data from the classrooms, auditorium, and gym. The comparison involved assessing the root mean square error (RMSE) performance metric:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

In the formula, $N$ represents the total number of data points, $y_i$ represents the actual (measured) values, and $\hat{y}_i$ represents the predicted (simulated) values. The RMSE is calculated by taking the square root of the mean of the squared differences between the actual and predicted values.
The last 11 days were used as a testing dataset to compare the behavior predicted by the identified model with the actual behavior supplied by the measured dataset modeled in Energyplus. $T_1$, $T_2$, and $T_3$ are the system outputs used in this comparison. Figure 4, shows the actual behavior in grey with respect to the forecast behavior in blue.
The obtained results revealed promising accuracy, with errors of 3.5%, 6.5%, and 15.2% for the classrooms, auditorium, and gym, respectively. These errors indicate the percentage deviation between the simulated responses of the identified model and the actual measurements: lower error percentages signify a closer agreement between the model predictions and the real-world behavior of the building's thermal dynamics.

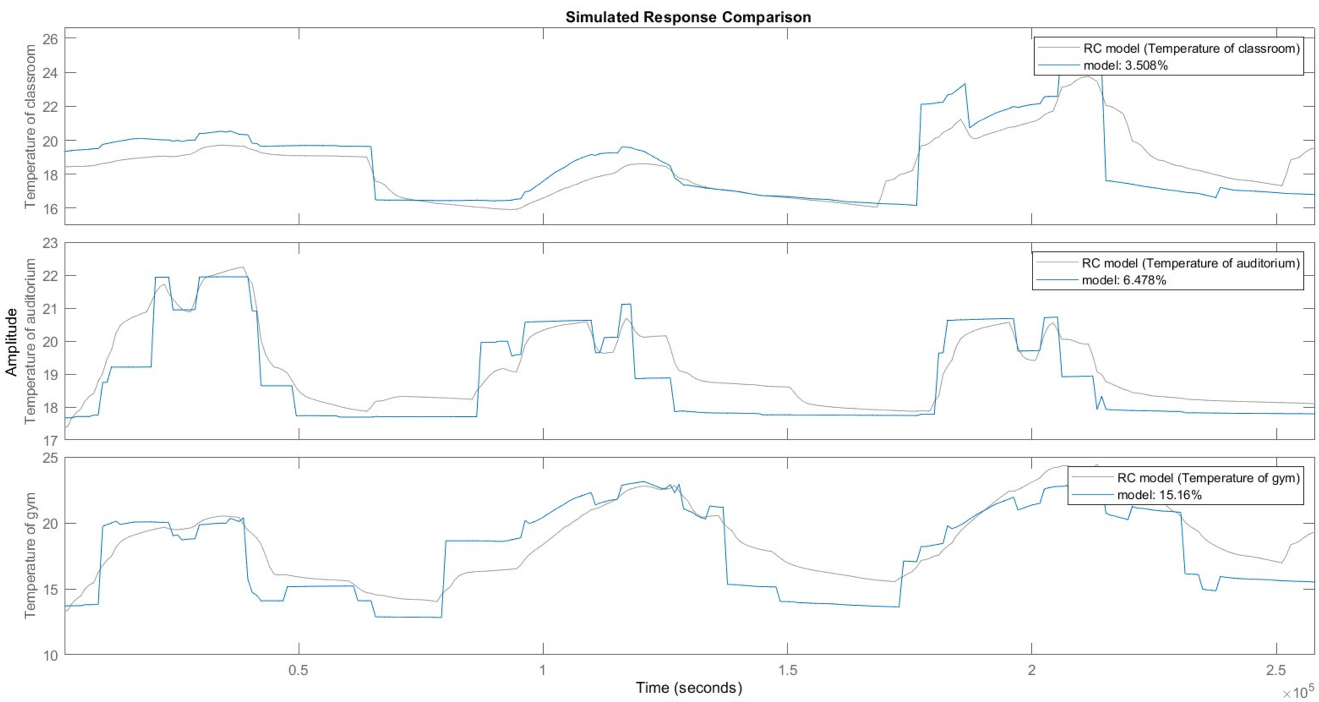


Figure 5: Zoom of the results obtained with the *comapare* function

In Figure 5 a zoom of the previous figure is shown since the most important result was achieving a similar trend between the two models specifically in the first days.

## Connection between Energy+ and Simulink

To establish the connection between Simulink and EnergyPlus, we utilized the EnergyPlusToFMU tool in conjunction with Visual Studio. EnergyPlusToFMU is an open-source software that enables the export of

EnergyPlus models as Functional Mock-up Units (FMUs), which can be seamlessly integrated into Simulink. Once the EnergyPlus model was successfully configured, we exported it as an FMU file, which encapsulated the entire EnergyPlus simulation. This FMU file was then imported into Simulink, allowing us to incorporate the EnergyPlus model as a subsystem within the Simulink environment.

## Baseline control strategy

The baseline control strategy implemented for the building system is the PID (Proportional-Integral-Derivative) control algorithm. PID control is a feedback control mechanism that continuously adjusts the control input based on the error between the desired setpoint and the measured process variable. The PID controller consists of three main components (Figure 6): the proportional, integral, and derivative terms. The proportional term provides a control output proportional to the current error, the integral term accumulates and corrects for past errors, and the derivative term predicts future changes in the error. By combining these three terms, the PID controller can effectively respond to both the steady-state and dynamic behavior of the controlled system. The choice of PID control as the baseline strategy stems from its simplicity, effectiveness, and well-established theoretical foundations.

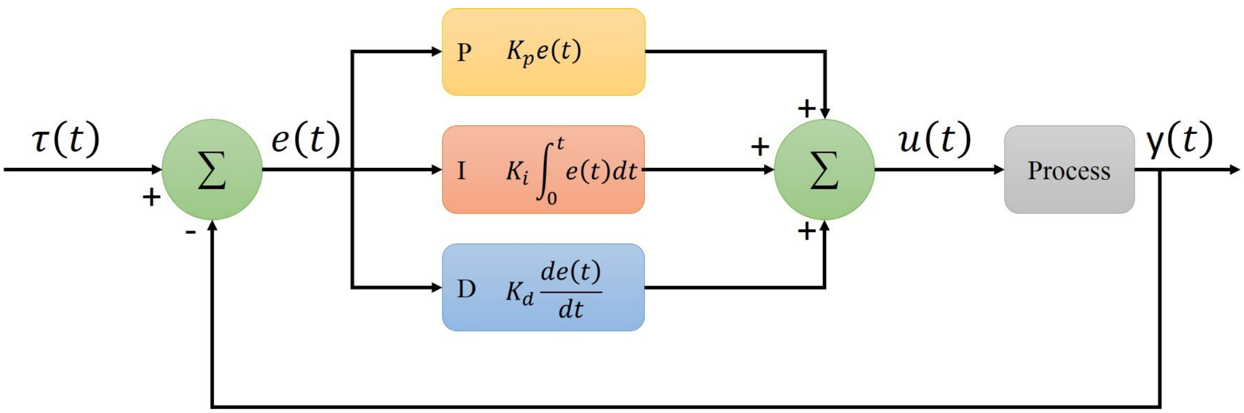


Figure 6: PID controller scheme

### PI control strategy

The initial control strategy implemented for the building system is based on the PI (Proportional-Integral) control algorithm. The decision to begin with a PI controller stems from several considerations, including the system dynamics, control objectives, and practical implementation aspects. In order to obtain the best results different values for the PID controller are used.

The PI control algorithm combines proportional and integral terms to regulate the controlled variables and achieve the desired setpoints. The proportional term provides control action proportional to the current error between the setpoint and the measured variable, while the integral term accumulates and corrects for past errors. By employing both terms, the PI controller can address steady-state errors and improve the system's response to changes and disturbances.

The choice to start with a PI implementation is driven by several factors. Firstly, the building system exhibits dynamics that can be effectively managed by proportional and integral actions. The proportional term ensures a quick response to deviations from the setpoint, while the integral term aids in eliminating steady-state errors and enhancing overall system accuracy.

Furthermore, the PI control strategy strikes a balance between complexity and control performance. Compared to more complex control algorithms such as PID or advanced model-based approaches, the PI implementation offers a simpler structure and ease of tuning.

Additionally, starting with a PI controller provides a solid foundation for future control enhancements and refinements. Several attempts are made in order to obtain the best results. Setting the P constant as 0.56 and the I constant as 0, the output better follows the schedule of the temperature. The final choice is to consider only the proportional contribution because with the integral contribution the internal temperature is too close to the results get from EnergyPlus, instead of following the right schedule.

# First implementation of the advanced controller (MPC)

This section focuses on the application of Model Predictive Control (MPC) in system modeling and design. MPC is an advanced control strategy widely used to optimize the performance of dynamic systems. By utilizing a mathematical model, MPC predicts future system behavior and computes an optimal control sequence, considering constraints and performance objectives. The Reduced Order Model (RC) is employed to enhance the optimization process, addressing the complexity of thermal models. The MATLAB MPC toolbox provides a robust tool for designing optimal control. The real plant, represented by a white box model, receives inputs from the controller, and the reference tracking error is measured. Design steps involve differentiating manipulated variables and disturbances, defining constraints and weights, and modeling the process within the Simulink MPC block. Constraints include upper and lower limits, and weights are assigned to inputs and outputs. Figure 7 represents the system response to a given reference with an MPC controller.

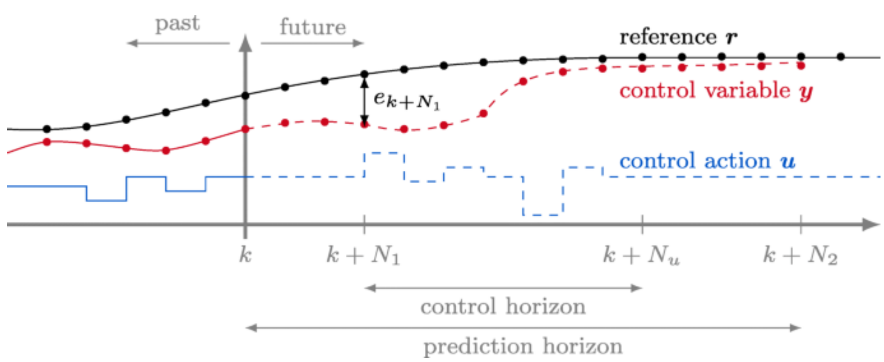


Figure 7: A general system's response to a given reference with an MPC controller

## A conceptual preface to Model Predictive Control strategy

Model Predictive Control (MPC) is an advanced control strategy widely used in engineering and industrial applications to optimize the performance of dynamic systems. Unlike traditional control methods that operate in a feedback loop, MPC takes a predictive approach by using a mathematical model of the system to forecast its future behavior. By formulating an optimization problem, MPC computes an optimal control sequence over a finite time horizon, considering system constraints and performance objectives. This predicted control sequence is then implemented in a receding horizon manner, where only the first control action is applied, and the process is repeated at each time step, continuously updating the predictions and control actions. MPC offers benefits such as handling constraints, handling nonlinear systems, and accommodating varying operating conditions, making it a powerful tool for controlling complex systems in real time.

## System modelling

Despite having a promising ability in dealing with the nonlinearities of the dynamic model, due to the over-complexity of the thermal model, MPC is unable to perform the optimization process on the real plant (white box model), i.e. the task of optimization is cumbersome, and in most cases non-converging. That is the principal motivation behind the RC modeling of the system. Concretely, employing a relatively precise

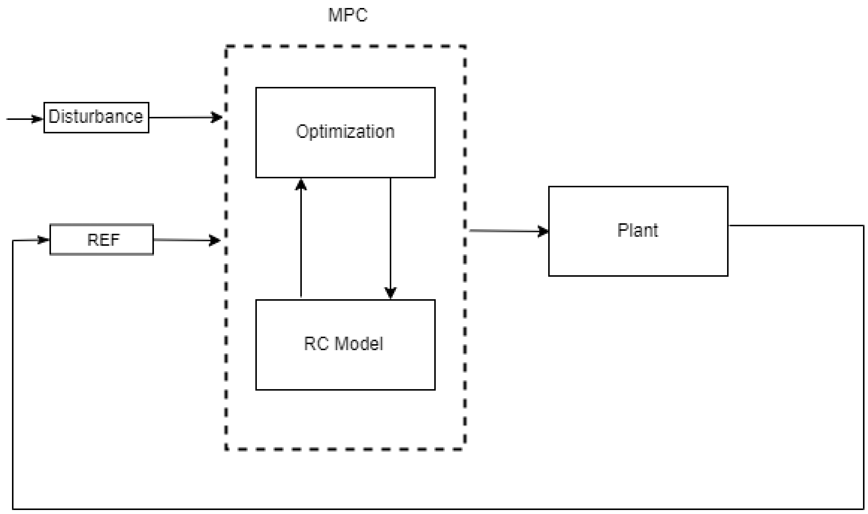


Figure 8: General scheme of the model predictive controller

reduced model in the optimization process is the key to developing such a technique. Figure 8 represents the scheme of the system. Speaking of the mathematics behind MPC is beyond the scope of this challenge. Moreover, the MATLAB MPC toolbox provides a robust tool in order to design the optimal control of the system.
The model which is used in the optimization loop is that of the obtained model in the grey box identification phase.
The real plant which is the white box model from open studio gets fed with the outputs of the controller and in the feedback loop, the reference tracking error is measured between the output of the plant and the temperature schedule given by the company. In the following sections, a more detailed framework of the design is illustrated.

## Design steps

Unlike the identification phase, in which all the 11 inputs are responsible for the system inputs, in the case of control design, one must discriminate between the manipulated variables and the disturbance. This is further coherent with the contractility of the system and having the contractility matrix full ranked. To this regard, the first three inputs of speak, the heat flux introduced in each zone are considered the manipulated variables and the rest that encompassing the internal and solar gain, and the ambient, corridor, and the unconditioned zone temperature.
Once the input and disturbance signals are divided, the next step is to define the prediction and control horizon. It is suggested in the literature that a control horizon of 20 time steps would be enough to embed the system dynamics in its upcoming steps. The control is chosen 7 time steps, since in the sudden reduction of the reference temperature the controller was not able to suitably track the reference.
As spoken earlier, MPC is a powerful tool to deal with the system constraints.In this step the constraints and the weights of system inputs and outputs are determined. Thanks to MPC toolbox, the whole process is being modeled inside the simulink MPC block.
In our case the upper value for the control input is set to 1 and the lower value to 0. This is due to the fact that the FMU model which is implemented in the simulation takes the input as a fraction of the ideal heat load. After a process of trial and error, giving a zero weight to the inputs is the best compromise to keep the robustness of the closed-loop response. However, to keep the smoothness of the input and not having instability issues a value of 0.1 is assigned to the rate of change. This is due to the fact that in the process of

HWT design having a sudden and vigorous drop in the heat load that is cyclic would provide technical issues from the applicability point of view. To give an equal value to each output, the weight matrix is considered an identity matrix. Eventually, a tolerance of 3°C is assigned to the outputs not to exceed the given interval. Figure 9, represents the system model in the Simulink environment.

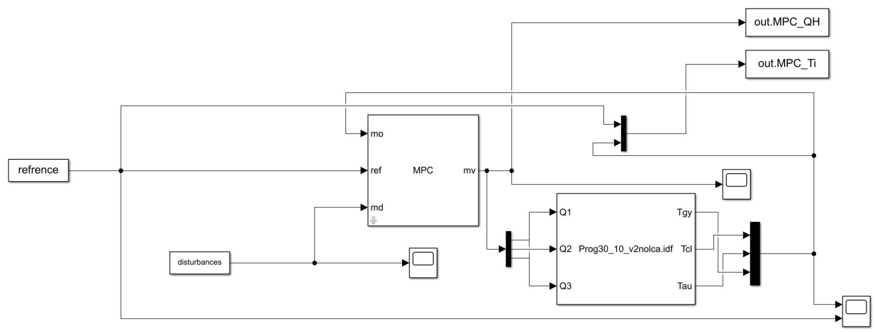


Figure 9: Simulink model of the control system

## Results

After the implementation of different control strategies, a comparison between every result is made, shown in Figure 10 and Figure 11. The simplest case is created using the baseline control strategies. The rules for this kind of control are given by the company. To perform this control a PID system is used. For doing that, it was necessary to create a model in Simulink in order to simulate the behaviour of the building with different PID systems. After the implementation of the PID control, the following step was to perform an MPC. As shown in the next chart, this advanced control strategy allows us to obtain better results with lower energy consumption. In particular, thanks to the MPC it is possible to reduce the so high peak at the beginning of the school day and to turn off the heating system during the night. Instead with PID control in the morning the power peak is extremely high (and energy-consuming) and during the night it is hard to obtain zero as the heating power.

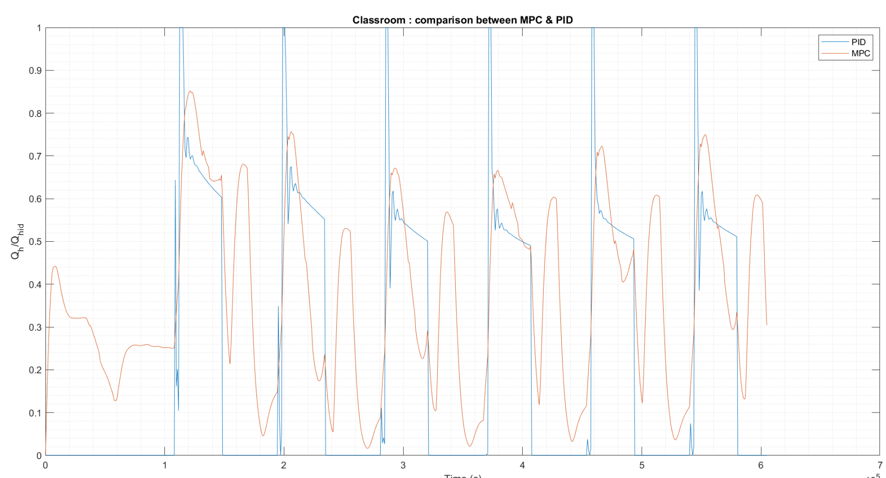


Figure 10: Heating power comparison between PID and MPC control in classrooms

Regarding the temperature trend obtained with different ways of control, it is possible to see that the PID struggles to achieve the setpoint temperature instead of the MPC that overcomes the setpoint.

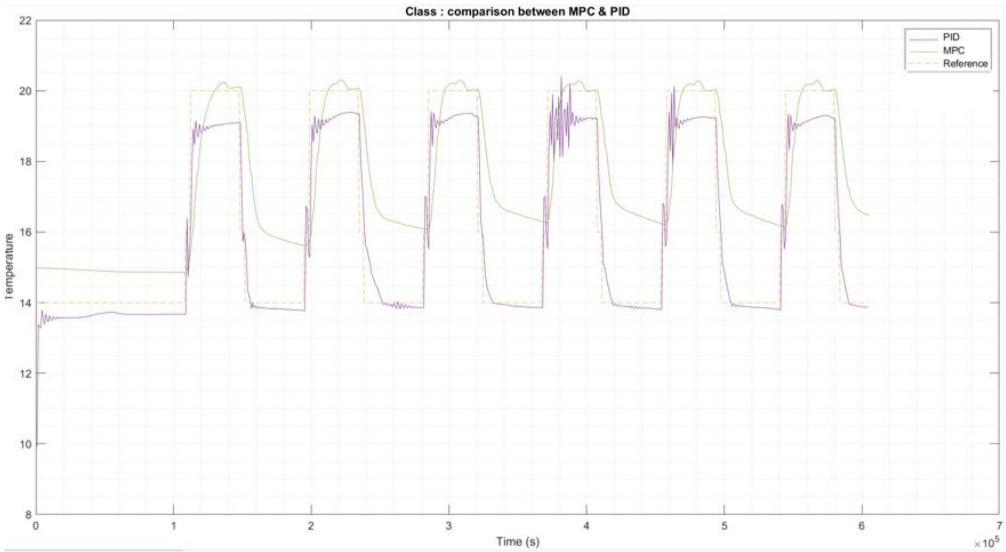


Figure 11: Temperature comparison between PID and MPC control in classrooms

# Conclusions

As discussed earlier, two control techniques are employed in this project, namely, PID and MPC. PID controllers are referred to as the kind of controllers that use a proportional gain, an integrator and a derivative component that compensates the reference tracking error at each time step. The main advantage of these controllers is their convenience to be used in plants since tuning is a relatively simple task. However, several drawbacks of these controllers make them less suitable to be used in most applications. They cannot deal with the nonlinearities of the plants and only their performance is limited to linear time-invariant (LTI) systems. The parameters are fixed with respect to time and only tuned once. They are disabled to deal with the constraints of the systems. These hindrances demand more advanced and optimal control strategies that can be used in more complex systems.

On the other hand, Model Predictive Controller (MPC) is based on an optimization process that at each time step updates the input sequences to minimize the objective cost function that guarantees the minimum tracking error and the optimal energy consumption in the system with subject to system constraints. Furthermore, it holds the generality to most nonlinear systems which in our case the plant is notably nonlinear.

The graphical results from each controller are a reliable litmus paper to came in concluding which strategy performs better. Speaking of system response, in the case of the PID controller, there's approximately 8% static error. In the case of MPC, the deviation is in the form of an overshoot with less than 5% error. It further behaves closely precisely in the lower-bond reference period.

Regarding energy consumption, it can be perceived that in the instances that the PID controller is used, the high-frequency oscillation of the input can introduce several issues to the performance of HWT. Furthermore, over a long period of time, the energy consumption in MPC is less than that of PID.

Referring to the considerations, from accuracy and energy consumption to design, it is confident to say that MPC exhibits a better performance as compared to the classical PID controller.

# Appendix

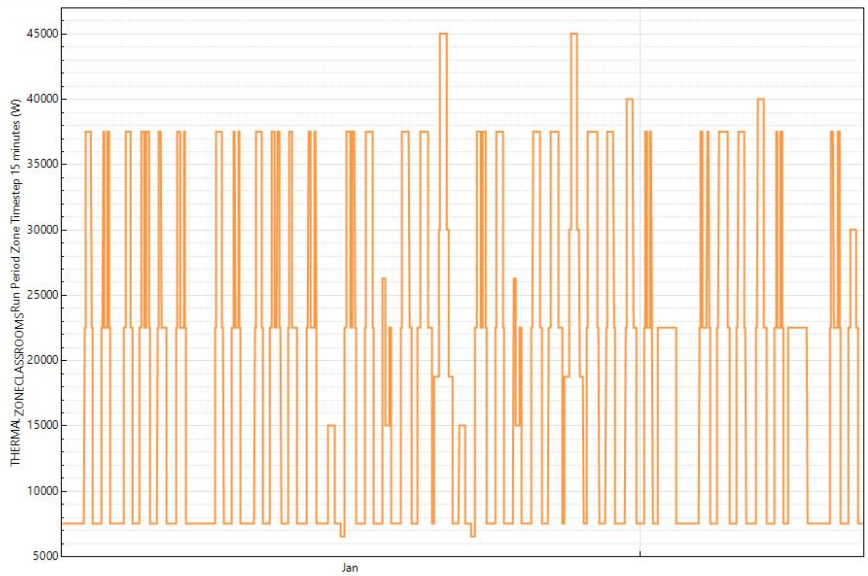


Figure 12: Schedule for heating system of classrooms

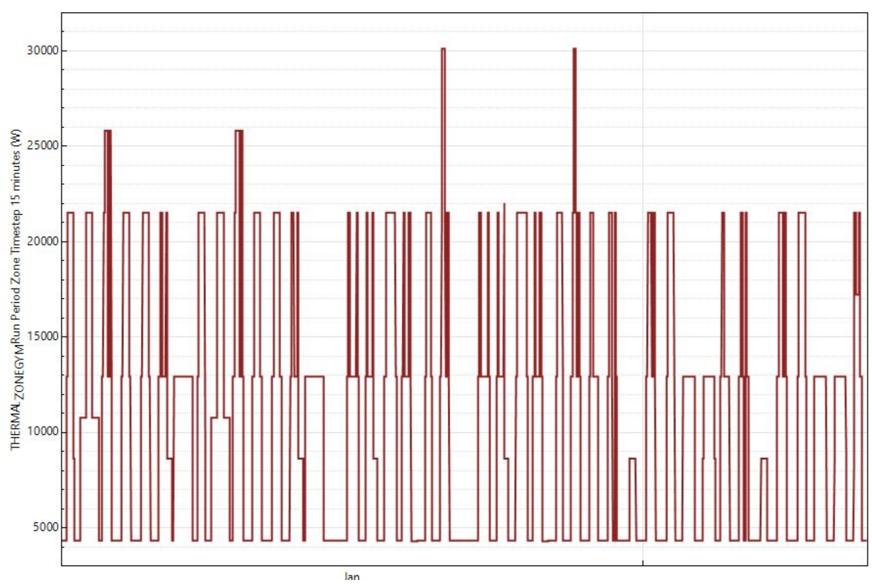


Figure 13: Schedule for heating system of gym

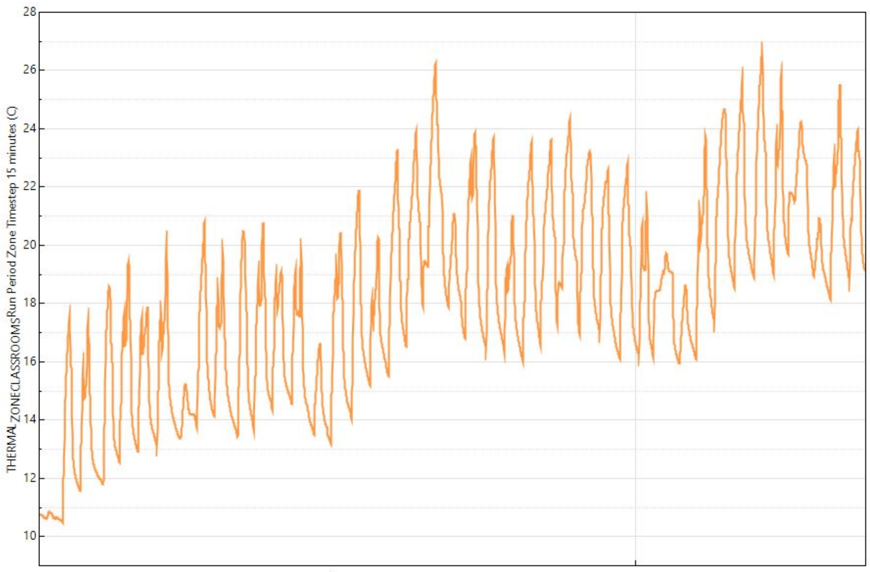


Figure 14: Classrooms internal temperature trend

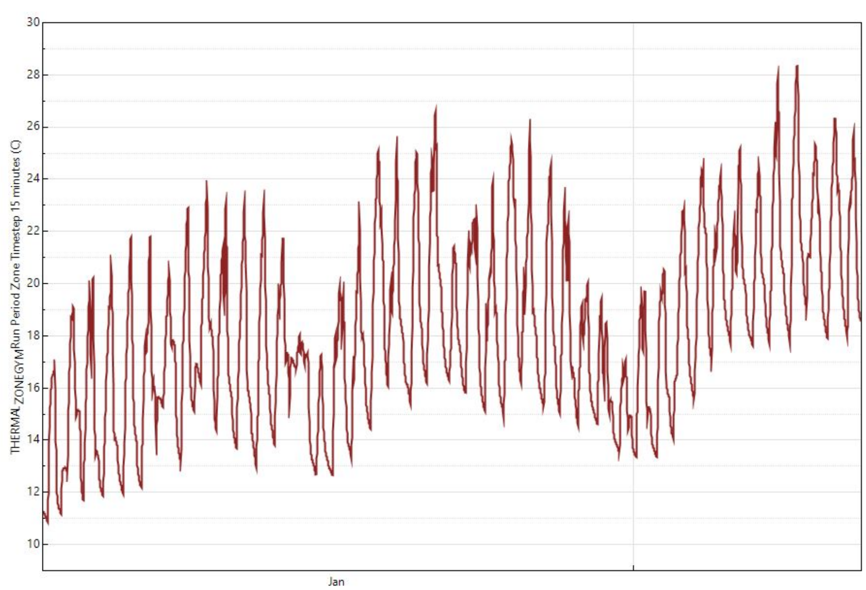


Figure 15: Gym internal temperature trend

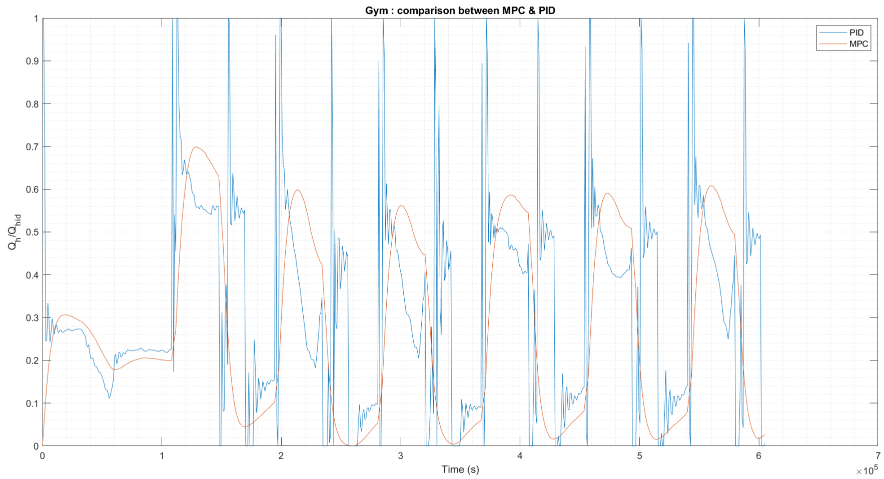


Figure 16: Gym heating power

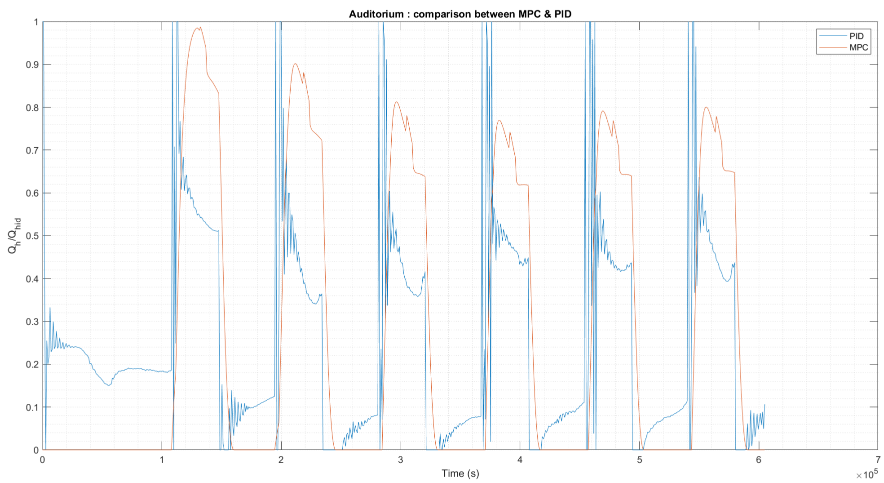


Figure 17: Auditorium heating power

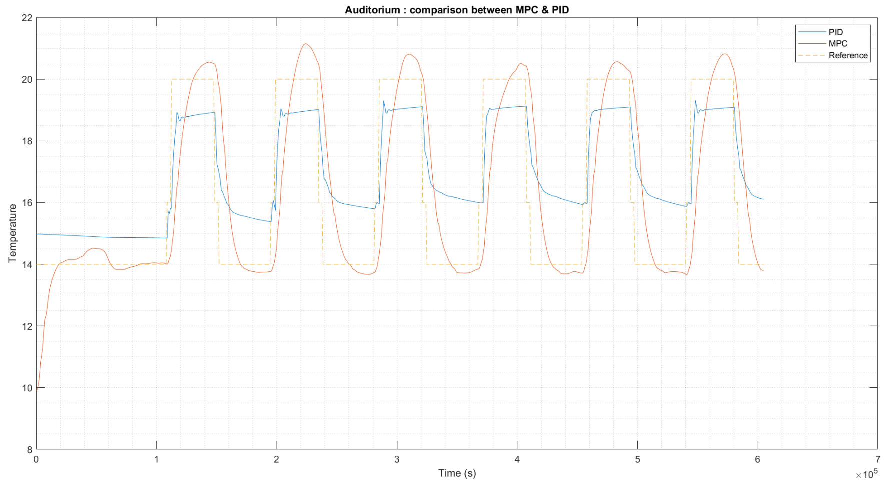


Figure 18: Auditorium temperature

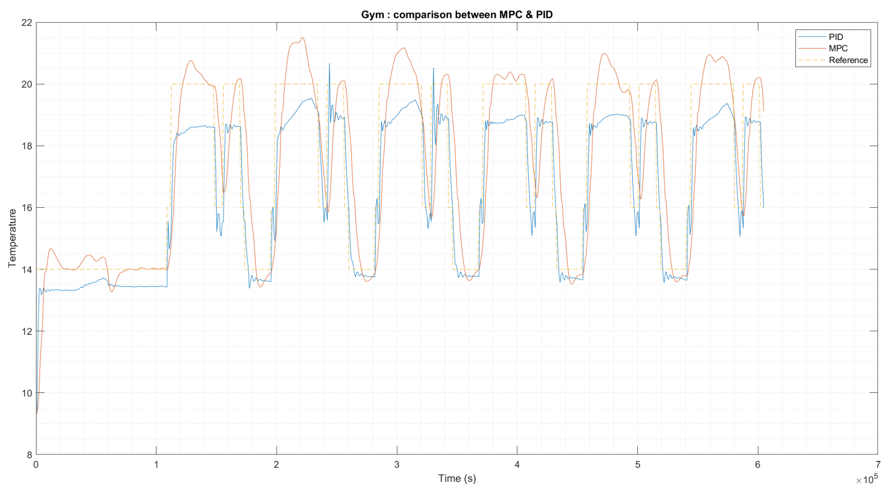


Figure 19: Gym temperature