POLITECNICO DI TORINO

Corso di Laurea in Ingegneria Energetica e Nucleare

Indirizzo Termotecnico ed Uso Razionale dell'Energia

# STUDY OF THE ADAPTABILITY AND EFFICIENCY OF REINFORCEMENT LEARNING BASED CONTROL FOR HVAC SYSTEMS THROUGH ENERGYPLUS DYNAMIC SIMULATIONS.


**Relatore:**
Prof. Enrico Fabrizio

**Candidato:**
Fabio Castiglione

## Acknowledgements.

For this Master Thesis Degree my gratitude goes especially to Professor Enrico Fabrizio who helps me in all the steps of this project from the beginning and all Politecnico of Turin that gave me the possibility to join Addfor.

I would also like to thank my tutor Enrico Busto and all the team from Addfor who followed me in this project especially Stefano, Antonio and Matteo but also all the big family of Addfor for receive me in the company as one of them and help me feel comfortable.

My heartfelt thanks to my family for the sustain in all my life and academic experience here in Turin, in Palermo, in Vigo but also at home and in the future with my brother and my sister. I would like to dedicate all this work to my father who transmit me the passion for technology and engineering and still is source of inspiration and knowledge despite my degree.

For my beautiful experience in Turin my gratitude goes to my best friends Daniele and Giuseppe for enjoying the everyday stay. They will be anything without all the others friends and university colleagues for sharing with me two years of their life and hopefully also many years that will come in the future.

Finally thanks also to all other my friends in Caltanissetta and San Cataldo that I left many years ago and the other friends who came from there to celebrate this result with me, the friends I left in Palermo, in Vigo and all around the world that I still keep in my mind and still are a landmark for me.

## Summary of Notation, Reinforcement Learning.

| | |
|---|---|
| $a$ | General element of $A$ |
| $a_t$ | Action of the agent at timestep $t$ |
| $c_p$ | Specific heat |
| $e$ | Offset or error signal |
| $g$ | Solar Factor |
| $h$ | Enthalpy |
| $h_c$ | Convection Coefficient |
| $m_a$ | Air Mass |
| $r$ | General element of $R$ |
| $r_t$ | Reward at timestep $t$ |
| $s$ | General element of $S$ |
| $s_t$ | State of the environment at timestep $t$ |
| $t$ | Thickness |
| $\underline{x}$ | Vector with $x_j$ components |
| $x_j$ | General input data of Machine Learning algorithm |
| $\dot{x}$ | Time derivative of variable $x$ |
| $\underline{y}$ | Vector with $y_j$ components |
| $y_j$ | General output of Machine Learning algorithm |
| $w$ | Wind Velocity |
| $A$ | Space of the possible actions |
| $B$ | Set of temperature in between heating and cooling set-point |
| $E[]$ | Expected Value |
| $K_d$ | Derivative Gain |
| $K_i$ | Integral Gain |
| $K_p$ | Proportional Gain |
| $P[x_t = x]$ | Probability of $x_t$ being equal to $x$ |
| $P[x\|y]$ | Conditional Probability |
| $P[]$ | Probability function |
| $Q$ | Net Heat Transfer |
| $Q^\pi$ | Action-value function for the policy $\pi$ |
| $R$ | Space of the possible rewards |
| $R_t$ | Thermal Resistance |
| $S$ | Space of the possible states |
| $T$ | Temperature |
| $T_0$ | Adjustment Constant Temperature |
| $T_a$ | Ambient Temperature |

| | |
|---|---|
| $T_i$ | Indoor Air Temperature |
| $T_o$ | Outdoor Air Temperature |
| $T_p$ | Controlled Temperature |
| $T_s$ | Supply Temperature |
| $T_{s,j}$ | Surface j Temperature |
| $T_z$ | Air Temperature of the $z$ thermal zone |
| $V^\pi$ | Value Function for the policy $\pi$ |
| $U$ | U-value |
| $W$ | Mechanical Work |
| $\alpha$ | Learning Rate |
| $\gamma$ | Discount Factor |
| $\theta_j$ | Parameter of Machine Learning algorithm |
| $\lambda$ | Thermal conductivity |
| $\pi_t$ | Policy function at timestep $t$ |
| $\rho$ | Density |
| $\psi_g$ | Linear Thermal Bridge coefficient |
| $\Delta U$ | Internal Energy variation |
| $\Phi$ | Heat Power |
| $\mathcal{P}$ | Transition Function |
| $\mathcal{R}$ | Reward Function |

# Contents

# 1 Introduction.

In last years in which pollution and energy consumption is one of the firsts topic in research and industrial development energy savings have been applied to all fields like space heating and air conditioning. This achievement can be reached by modifying the thermal envelope of the building (for example increase thermal insulation or advanced glazing system), by improving the HVAC system (Heating, Ventilation and Air Conditioning that is the system designed to satisfy the thermal needs of the building) that allows to have the same output with a reduction of the energy input in the system, by modifying the control strategy of the HVAC system in order to more efficiently use the free internal, external gains of the building joint with future weather prediction and thermal requirements.

The aim of this Master Thesis Degree in Energy Engineering is to test a control algorithm based on Reinforcement Learning to reduce the energy consumption for heating and air conditioning guaranteeing the thermal comfort requirements. The test has been done through the development of some different dynamic models with the software EnergyPlus that allows to create dynamic models for buildings.

The second chapter is a brief introduction to the most common control system applied to HVAC system with the description of some logic and controls variable used in different cases. Then, in the third chapter, is described the general rationale behind Machine and Reinforcement Learning with the explanation of the algorithm used in the simulations.
In the fourth chapter are described the equation used by EnergyPlus to solve heat balance equation in buildings and are described the models used for the dynamic simulations. Chapter five shows results obtained with the comparison of the Reinforcement Learning based control in comparison with the most common control system explained in chapter one to quantify the energy savings given by the algorithm.

## 2 Control of HVAC System: State of the art.

An **HVAC** (Heating, Ventilation and Air Conditioning) is a system that is designed to accomplish some comfort and healthy needs of the users of all the buildings. The main objectives HVAC systems are to guarantee a certain air temperature and air exchange with mechanical ventilation. The objective quantity that the HVAC want to keep in an environment is called **set-point** while the controlled variable is the quantity that is modified by the controller to reach the set-point. The logic from which this variable is modified to reach the desired set-point is called control action.

A **closed loop** or **feedback** control (Figure 1) measures what are the changes in the controlled variable and modifies the output until the set-point is reached.

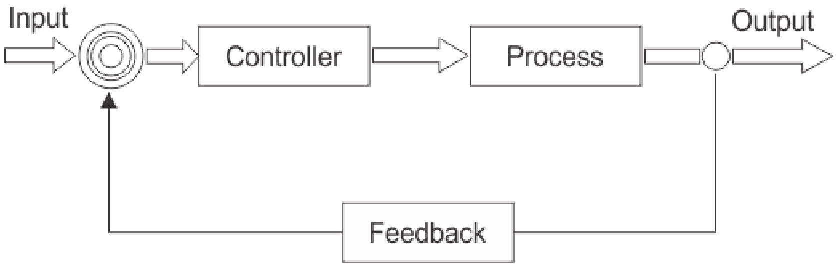



Figure 1: Closed Loop Scheme - Electrical4u

Contrarily the **Open loop** control does not have a direct link between the controller and the controlled variable but the adjustment is done based on an external variable in order not to have a too big offset between the set-point and the controlled variable. Typical example of this strategy is the thermostat control with the external temperature while the controlled variable is the internal temperature that is not direct compared with the set-point.

Every closed loop must include some devices called **sensor** which measures the controlled variable, compares it with the input of the chain and modifies it, by some equation, through an **actuator**. In most cases actuators, for HVAC Systems are valves, dumpers, variable-speed pumps or fans.

### 2.1 Types of Control Actions.

A first classification can be done according to the adjustability of the controlled variable, that can have just two operational phases (for example a valve fully opened or fully closed, called **Two-Position Control**) or more than two and the control action type is called modulating)[1]. Two-Position Control is used for example in hot water storage in which water is heated by a boiler until a certain temperature is reached, then the boiler is turned off until water temperature decrease below a certain threshold and boiler is turned on again as shown in Figure 2.

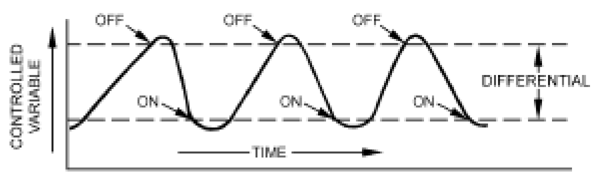



Figure 2: Two Position Control - 2017 ASHRAE Handbook-Fundamentals (SI)

The simplest modulating control is the **Proportional Control:** the output is modified proportionally with the **offset** (or error signal which is the difference between the set-point and the actual value of the controlled variable). It can be mathematically described as follows in the example of temperature control.

$$T_p = K_p e + T_0 \tag{1}$$

- $T_p$: is the temperature of the controlled variable to be set (for example the inlet water temperature of a fan coil);
- $K_p$: proportional gain;
- $e$: is the offset;
- $T_0$: is the adjustment constant for the output $T_p$.

For all modulating control strategies a **Proportional Band (PB)** can also be defined as the range in which the controlled variable can vary and is inversely proportional to the proportional gain $K_p$. A limit of this control strategy is that if a constant offset in time is reached, the output $T_p$ will be no longer modified so that the set-point will not be reached and a steady-state error can be observed as shown in Figure 3.

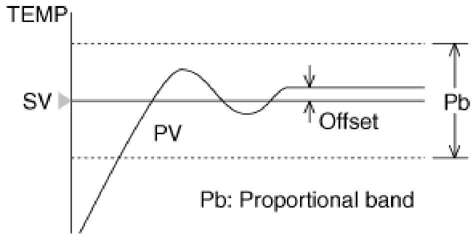



Figure 3: Proportional Control - RKC Instrument INC.

Improvement to the Proportional Control is the **Proportional Integral (PI) Control** that add an integral term to the simple proportional adjustment:

$$T_p = K_p e + K_i \int e dt + T_0 \tag{2}$$

- $K_i$: integral gain;
- $t$: time variable.

This type of control implies that in case of constant offset the integral error will increase and result in a modification in the output $T_p$ because if $e$ is different from 0 the output temperature continue in changing until the error get to 0 (Figure 4).

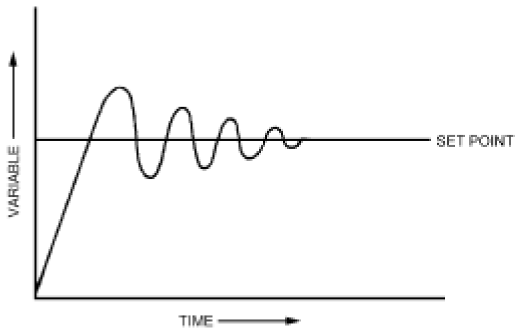



Figure 4: Proportional-Integral Control - 2017 ASHRAE Handbook-Fundamentals (SI)

Further improvement is the **Proportional-Integral-Derivative (PID) Control** that includes also a derivative term added to compute the output:

$$T_p = K_p e + K_i \int edt + K_d \frac{de}{dt} + T_0 \tag{3}$$

- $K_d$: derivative gain.

Adding this last derivative term gives an anticipatory control so that if the error is increasing rapidly in time will be strongly modified while if it does not change in time the derivative term is null and the steady-state error is avoided. Disadvantages of the PID control logic is that is hard to tune because is very sensitive to perturbations so in most of cases, for HVAC System control, PI logic is sufficient.[1]

Proportional, PI and PID Controllers introduce some coefficients ($K_p$, $K_i$, and $K_d$) that need to be tuned in order to minimize steady-state error of the set-point, responds quickly to disturbances and increase stability. Tuning of these controllers includes closed and open loop methods and trial-error methods. Two of the most widely used approaches are ultimate oscillation and first-order-plus-dead-time methods optimized with **Ziegler-Nichols** method described in AHSRAE Handbook [1].

**Ultimate Oscillation Method** consist in increasing $K_p$ value until the controlled variable continuously cycle around the set-point (Figure 5). Then proportional and integral terms are computed from the period of oscillation ($T_u$) and the $K_p$ value that got rise to periodic behaviour that neither grow nor diminish in amplitude called $K_{p,u}$ (for example equal to 40.0 in Figure 5).

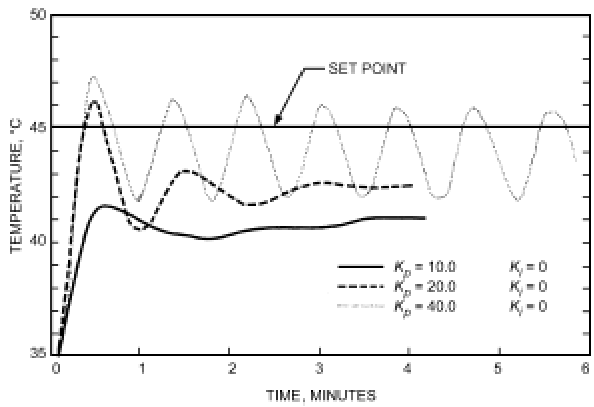



Figure 5: Response of Discharge Air Temperature to Step Change in Set Points at Various Proportional Constants with No Integral Action - 2017 ASHRAE Handbook-Fundamentals(SI)

Proportional only:

$$K_p = \frac{K_{p,u}}{1.8} \tag{4}$$

Proportional plus Integral:

$$K_p = \frac{K_{p,u}}{2.22} \tag{5}$$

$$K_i = \frac{K_{p,u}}{0.83T_u} \tag{6}$$

Proportional plus integral plus derivative:

$$K_p = \frac{K_{p,u}}{1.67} \tag{7}$$

$$K_i = \frac{K_{p,u}}{0.50T_u} \tag{8}$$

$$K_d = \frac{K_{p,u}}{0.125T_u} \tag{9}$$

**First-Order-plus-Dead-Time Method** introduce a step discontinuity in the set-point of the controlled variable and then graphically are obtained dead time ($TD$), that is the time passed from the giving of the new set-point and the time in which the controlled variable begin to change , and time constant ($TC$), that is the time that pass from $TD$ to the time in which the controlled variable reach 95% of the set-point (Figure 6). Then the coefficients are obtained as follows:

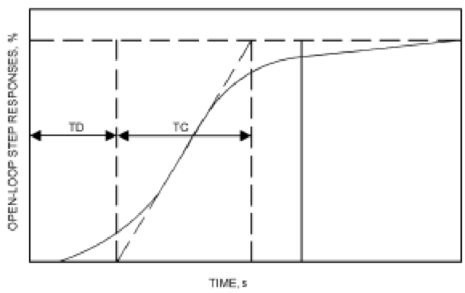



Figure 6: Open-Loop Step Response Versus Time - 2017 ASHRAE Handbook-Fundamentals(SI)

$$Gain = \frac{change\,in\,controlled\,variable}{change\,in\,control\,signal} \tag{10}$$

Proportional only:

$$K_p = \frac{TC}{TD * Gain} \tag{11}$$

Proportional plus integral:

$$K_p = \frac{TC}{TD * Gain * 0.9} \tag{12}$$

$$K_i = \frac{K_p}{3.33(TD)} \tag{13}$$

Proportional plus integral plus derivative:

$$K_p = \frac{TC}{TD * Gain * 1.2} \tag{14}$$

$$K_i = \frac{K_p}{2(TD)} \tag{15}$$

$$K_d = \frac{K_p}{0.5(TD)} \tag{16}$$

Finally **trial and error** method adjust the gain of proportional controller only until the desired response to the set-point is obtained. The integral term is the adjusted to met the objective of minimizing the steady-state error and increase stability. Figure 7 shows the effect of, after having increased $K_p$, $K_i$ is increased to meet the desired behaviour.

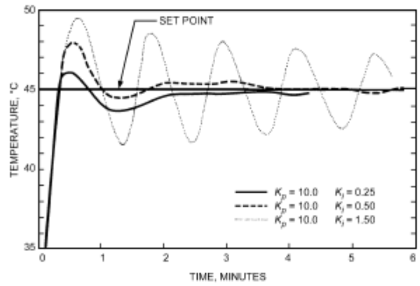



Figure 7: Response of Discharge Air Temperature to Step Change in Set Points at Various Integral Constants with Fixed Proportional Constant - 2017 ASHRAE Handbook-Fundamentals(SI)

### 2.1.1 Advanced Types of Control Action.

Some advanced control strategies are available for both two-position and for modulating control.

An example of advanced control algorithm is **Fuzzy Logic** that uses a series of *if-then* logics that emulate what a human user will do when controlling the output variable. In Fuzzy Control the designer must at first design the value of some elements such as *very low temperature, high temperature, ok temperature* and so on and *if* this values are reached *then* apply the corresponding increase or decrease in the output variable. Each of this couple of state of the controlled variable and modification in the output variable is called **rule** and a fuzzy algorithm can contain several rules defined by the designer. The result of this algorithm is a table in which all the rules are defined by *if* statement and *then* the actions occurs, if more than one statement have to be satisfied also some *and* logic could be included as shown in Table 1.

| Rule Nmb. | IF | AND | THEN |
|---|---|---|---|
| Rule 1 | T is very low | T is decreasing | Increase a lot heating |
| Rule 2 | T is OK | T is increasing | Decrease a little heating |
| ... | ... | ... | ... |

Table 1: Example of Fuzzy Logic.

Other example of advanced control strategies is the **Model Predictive Control (MPC)** in which a the current output variable is modified taking into account load forecast in the building, weather forecast and all those parameters that could affect the chosen controlled variable [3]. This control is able in principle to give at each moment the right amount of heat needed by the space but needs a specific model to be built for each conditioned space.

## 2.2 Controlled Variables for HVAC Systems.

In a typical HVAC system two control levels are present. The upper level control named **Supervisory Control**, is the level in which the set-points and the schedules of operation are set; the lower level control is called **Local-loop Control**, it is the control that provide from each single set-point to adjust the actuator. Both levels can be significantly improved in order to achieve a reduction in energetic consumption in a building for example by correct tuning of local-loop control or right choice in supervisory control to avoid energy waste.[2]

Figure 8 shows a scheme of a typical **Cooling System** in which an electric chiller produces chilled water that feeds a coil system where heat is exchanged whit supply air that has to be sent to the zone. Figure 9 shows, instead, a scheme of a typical centralized **Heating System** in which a boiler fuel driven produce hot water that in an heating coil exchange heat with the supply air to feed the thermal zones. Both these systems together guarantee the object of an

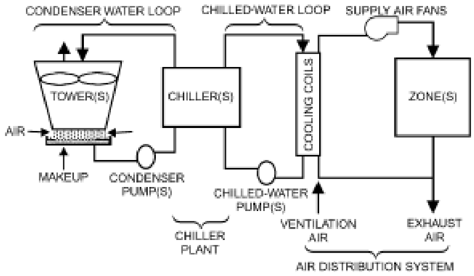



Figure 8: Chilled-Water Cooling System Scheme - 2007 ASHRAE Handbook-HVAC Applications (SI)

HVAC System, so provide heating, cooling and ventilation as the supply air is taken outdoors as fresh air.[2]

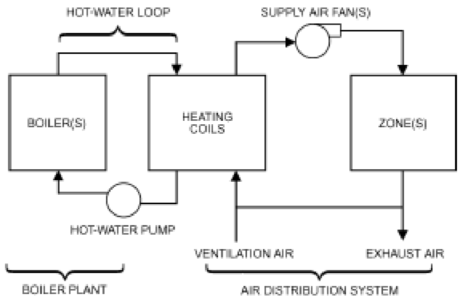



Figure 9: Hot-Water Heating System Scheme - 2007 ASHRAE Handbook-HVAC Applications (SI)

According to the type of HVAC system many types of subsystems may occur. One of most common system is **Air Distribution System** that includes terminals, like diffusers, one or more **Air Handling Unit** (AHU) and ducts to drive the air to the conditioned zones. In the AHU fresh air is mixed with external air forced by a fan and cross at least two heat exchanger fed by hot or chilled water. The hot water exchanger is usually called **hot deck** while the chilled water exchanger is called **cold deck**. Figure 10 represents a simplified scheme of an AHU with

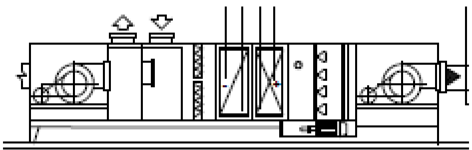


Figure 10: AHU Scheme.

two fans, one hot and one cold deck and a humidifier. The hot/chilled water flow is adjusted by a local-loop controller that actuates a three-way valve to control the **supply air temperature**.

If the fans have fixed speed the airflow rate cannot be adjusted, for this system the control regulates the amount of local reheat in each zone, this system is called **Constant Air Volume** or CAV system. If instead the fan speed can be adjusted the airflow rate can be controlled in order to maintain the set-point in each zone, this system is called **Variable Air Volume** or VAV system.[2]

For the water loop several control strategies may be applied. With a **primary pumping** system the a single water pipe is used from the boiler or chiller to the coils, often fixed-speed pumps are used that are cycled alternatively on and off, if the boiler or the chiller that a pump serves is cycled contemporary with the pump the control is called **dedicated**. **Water by-pass valve** are added to maintain almost constant water flow and avoid low loads for pump system.

In some cases can be find both primary and secondary water loop, these systems are designed for variable-speed pumps. While the primary loop provide almost constant water flow, as described before, with a fixed-speed pump; the second loop has one or more variable-speed pumps that provide different flow rates.

Finally for boilers typical control variable is **supply water temperature** ($T_s$) often adjusted with an Open-Loop control driven by ambient temperature. The two variables $T_s$ and $T_a$ (ambient temperature) are correlated with a curve, called **thermoregulation curve** obtained by interpolation between two chosen limit point. The kind of curve used for interpolation are function of the water equipment and the climatic zone of the building to control.[4]

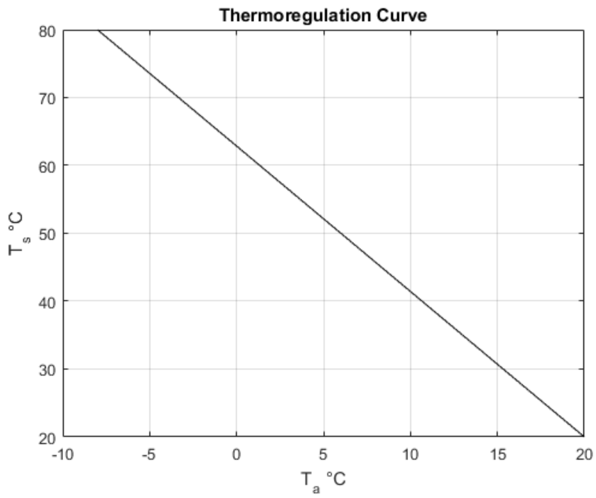



Figure 11: Example of Thermoregulation Curve with linear interpolation

# 3 Machine Learning and Reinforcement Learning.

The objective of Machine Learning Algorithms is finding patterns in a large set of data $(x_1, x_2, ..., x_N)$ called **training set** used to tune a set of parameters $(\theta_1, \theta_2, ..., \theta_M)$. The characteristics of each data in the training set is known in advance, by single inspection for example, and are labelled. The result of the machine learning algorithm is a function $y(x)$ which take as an input the set of data $\underline{x}$ and give as an output a vector $\underline{y}$. The form of the function $y(x)$ is defined during the training (or learning) phase of the algorithm and once the algorithm has been trained it will be able to characterize correctly a new set of data, called **test set**, without being manually labelled.

Before the training of the algorithm the input data are generally preprocessed to reduce the variability with scaling and translation to let the algorithm learn easier.

Machine Learning Algorithms can be divided in three main categories: **Supervised Learning** (**SL**), **Unsupervised Learning** (**UL**) and **Reinforcement Learning** (**RL**) (Figure 12).

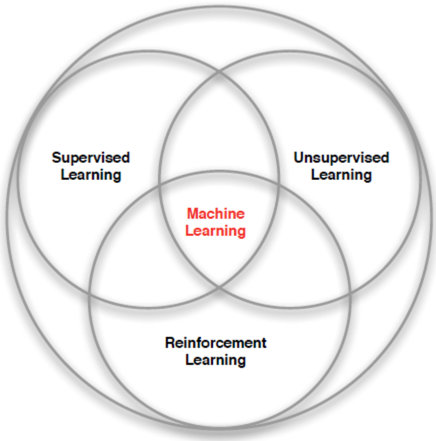



Figure 12: Machine Learning problems - David Silver

Applications in which the training set contains couples of input data and the corresponding output are called **Supervised** problems, they belong to the branch of Machine Learning called **Supervised Learning**. Inside the Supervised Learning, cases in which the problem is to assign to each single input vector a class over a finite number of classes are called **classification** problems; while if the problem is to assign as output one or more continuous variables are called **regression** problems.[5]

In another class of Machine Learning algorithm the input data contain just the $\underline{x}$ without any $\underline{y}$ associated. These kind of problems belong to **Unsupervised Learning** and their goal is to find groups of similar input data, this goal is called **clustering**.

The third class of Machine Learning algorithm that will be discussed here is called **Reinforcement Learning**. In this case the issue is to find the best actions to take in a given environment in order to maximize the reward. Differently from Supervised Learning, here the best actions are not given as training set but must be discovered by trial and error [5].

## 3.1 Reinforcement Learning.

In the previous brief introduction to Reinforcement Learning it emerges that the key of this class of algorithm is learning by experience that already is the common practice in human learning. The human experience provide information about the effects of our actions in a certain environment.

The computational corresponding in learning how good are the actions in relation with the environment is the maximization of a numerical reward related to the purpose of the algorithm (for example in the case of HVAC system the reward can be related to the energy consumption). The reward came directly from the environment and the best choices are discovered by the agent itself by trial and error. In addition, in the most general case the action of the algorithm may affect the environment also for many timesteps after the action has been taken. This is the problem of **immediate reward** that is given in the timestep just after the action, and the **delayed reward** described above.[6]

A Reinforcement Learning algorithm has to receive the state of the environment and must be able to take actions that directly affect it, finally it has a goal that is strictly related with the state of the environment. In this process the decision maker is called **agent** and the thing it interacts with is the **environment** (Figure 13).

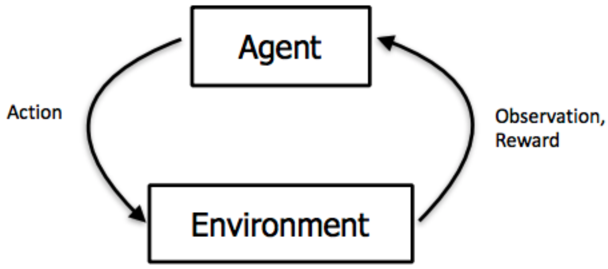



Figure 13: Reinforcement Learning scheme - Reinforcement Learning with Q tables, Mohit Mayak

This learning approach get rise to some problems like **exploitation-vs-exploration** trade-off. To obtain a big reward, the agent may chose some actions already tried in the past that are known to be highly rewarded but in order to find these actions it has to try new choice never tried before. So the agent has to both **explore** in order to find new highly rewarded and take better actions that the previous ones but also **exploit** what has been already experienced before. The agent may experience a lot of actions and, on them, chose the ones that has greater rewards.

A Reinforcement Learning system has, in most cases, four elements:

- A **policy**, which defines the way the agent behave at a single time.
- The **reward**, which is a single scalar value given to the agent by the environment at a single timestep that define the goal of the Reinforcement Learning system, so it identifies good and bad behaviour.
- While the reward is an indicator of the goodness of the action at a single timestep, the **value function** is a forecast of the total amount of rewards that a system can obtain in the future timesteps.
- In some systems there is the possibility for the algorithm to predict the behaviour of the environment by a **model**. This distinguish the **model-based** algorithms by the much simpler category of **model-free** algorithm that rely only on trial-and-error.

In mathematical terms at each timestep t, the agent receives observation about the state $s_t$ and the reward $r_t$ by the environment and chose an action $a_t$. The environment receives the action $a_t$ and send to the agent next state $s_{t+1}$ and scalar reward $r_{t+1}$ (Figure 14).

The agent policy $\pi_t(s, a)$ is a function from states to action probabilities. It is the conditional probability of choosing a certain action $a$ given the state $s$:

$$\pi_t(s, a) = P[a_t = a | s_t = s] \tag{17}$$

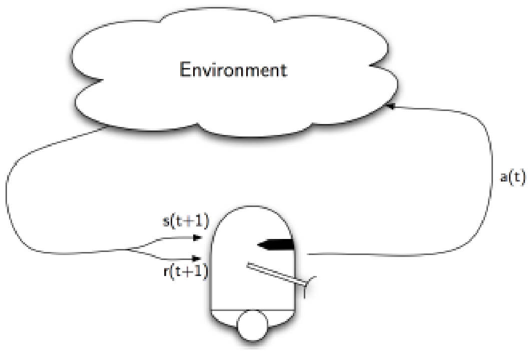



Figure 14: Agent-Environment Interaction - Fundamentals of Reinforcement Learning Yann-Michael De Hauwere

Reinforcement Learning system specifies how the policy changes in time according to trial-and-error procedure.

A technique give higher importance to immediate and near reward against long-term ones is the **discounting**. It introduce the **discount factor** $\gamma \in [0, 1]$ that is useful to compute the **discounted return** $R_t$:

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + ... = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \tag{18}$$

Since $\gamma \leq 1$ a reward given at timestep $k$ will worth less with respect of an equal reward given in a previous timestep. If $\gamma = 0$ the agent will focus on maximise only the immediate reward while as $\gamma$ approaches 1 the agent will be more and more farsighted taking strongly in account also future rewards.

As usually the amount of state, actions and transitions is finite and the current state contains all important information from previous states (**Markov property**); the Reinforcement Learning task is a **Markovian Decision Process (MDP)**.

$$P(s_{t+1}, r_{t+1} | s_t, a_t) = P(s_{t+1}, r_{t+1} | s_t, a_t, r_t, s_{t-1}, a_{t-1}, ..., s_0) \tag{19}$$

In a finite MDP the sets of states $S$, actions $A$ and rewards $R$ has all finite number of element and their probability is also described in a discrete way.

The transition between two subsequently states of the MDP is described by the **transition function**:

$$\mathcal{P}^a_{ss'} = P[s_{t+1} | s_t = s, a_t = a] \tag{20}$$

And a **reward function** defined as expected value of future reward for a certain number of timestep considered:

$$\mathcal{R}^a_{ss'} = E[r_{t+1} | s_t = s, a_t = a, s_{t+1} = s'] \tag{21}$$

The goal is to learn the correct policy $\pi : S \rightarrow A$ that maximises the reward. The policy can be deterministic, so the action chosen is function of $s$ and every time agent receives from the environment the state $s$, the action will be the same:

$$a = \pi(s) \tag{22}$$

Stochastic policy is instead defined as probability of having action $a$ at time t given state $s$:

$$\pi(a|s) = P[a_t = a | s_t = s] \tag{23}$$

Fixed the policy $\pi$ is possible to define the value function $V^{\pi}(s)$ as the expected value of the discounted return for an input state $s$, that identify the value for the state $s$:

$$V^{\pi}(s) = E_{\pi}[R_t|s_t = s] = E_{\pi}[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1}|s_t = s] \tag{24}$$

Similarly is possible to define the value for the given action $a$ in state $s$ called **action-value function**:

$$Q^{\pi}(s, a) = E_{\pi}[R_t|s_t = s, a_t = a] \tag{25}$$

The goal of the agent is to find the policy that maximise the value in the long-term behaviour:

$$\pi_{opt} = argmax_{\pi}[V^{\pi}(s)] \tag{26}$$

Given a state $s$, for each $a \in A$ a value $Q^{\pi}(s, a)$ may be defined, according to this value the policy assign a certain probability of being chosen to each action in order to maximise it and finally the value $V^{\pi}(s)$ can be computed as follows:

$$V^{\pi}(s) = \sum \pi(a|s) q_{\pi}(s, a) \tag{27}$$

Between all the possible action the agent will higly prefer the ones that lead to high value, so maximise the reward in long-term.[7]

## 3.2 Decision Tree and Random Forest.

For the object of this thesis the number of possible states of the environment and action given by the agent is extremely large. In such case it is impossible to find an optimal policy or an exact value function. In order to make correct decision in this case it is needed to have generalization tools that allows to reconstruct the approximated functions needed for the Reinforcement Learning system to work. This theme has already been studied by Supervised Learning algorithms so they will be integrated in the general system.

### 3.2.1 Decision Tree.

**Decision trees** are Supervised Learning algorithm used both for classification and regression problems. These involve stratifying or segmenting the **feature space**, that is the space of possible inputs $\underline{x}$ into $J$ distinct and non-overlapping regions $R_1, R_2, ...R_J$ (example in Figure 15) and for each combination of input data that fall in the region $R_j$ the same prediction is given that is the average value of that region.[8]

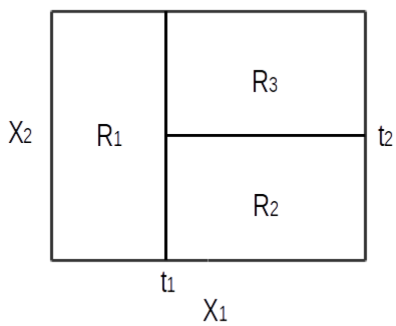



Figure 15: Segmentation of feature space in the example of three regions with two input - Pattern Recognition and Machine Learning

The goal is to find rectangular regions that minimize the **residual sum of square (RSS)** defined as follows:

$$\sum_{j=1}^{J}\sum_{i\in\mathcal{R}}(y_i - y_{R_i})^2 \tag{28}$$

Where $y_{R_i}$ is the mean response of region $R_i$.

Of course it is impossible to take into account and compare each possible segmentation of the feature space so is used a procedure called **recursive binary splitting**. This approach is defined **top-down** because start from the top of an imaginary tree and successively split the feature space into two region that identify the branches of the tree and repeat for each of these branches to continue down in the tree. It is also called **greedy** because at each step take into account only the best possibility at that particular step without considering next subdivision.

To perform this algorithm first a feature $x_j$ is chosen with a cutpoint $t$ that splits the feature space into two regions $[x|x_j < t]$ and $[x|x_j \geq t]$ that leads to the best possible reduction in RSS as shown in Figure 16. Then again for each branch look for the best feature $x_j$ and cutpoint $t$ that allow to minimize the RSS. Then the algorithm continues recursive until a certain stopping criteria has been reached.

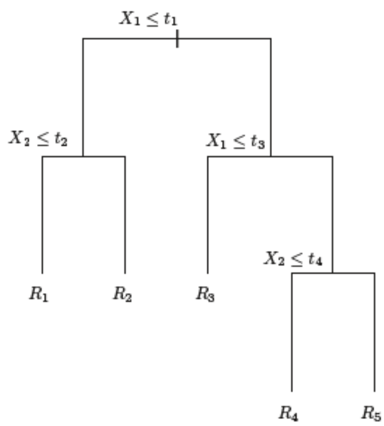



Figure 16: Example of recursive binary splitting procedure - An Introduction to Statistical Learning

### 3.2.2 Random Forest.

Decision tree often suffer for **high variance** it means that for different training set the result could be quite different. **Bootstrap aggregation Bagging** or **Averaging** is a procedure that allows to overcome this problem that may occur. It consist in training different models in various random sub-sets of the total training set and then the final prediction will be obtained averaging the predictions obtained by the models.[8]

In **Random Forest** algorithm many Decision Tree are built according to bagging procedure. But when building the trees, every time a split in the tree is conducted, the choice is conducted in just a subset of $m$ features instead of the total amount $J$ of features where in general $m \simeq \sqrt{J}$, so the majority of features is not even taken into account in making binary splitting. This procedure allows to **decorrelate** the predictions in the different trees that will create a more reliable Random Forest.[8]

A step forward are **Extremely Randomized Trees** in which not only, as in Random Forest, at each step is considered a random subset of the total features but instead of looking for the best cutpoint $s$ as before, cutpoints are randomly chosen for each feature. Then among all the couples (feature, cutpoint) that has been created the algorithm choose the best option in reducing the RSS.

### 3.3 Neural Network.

The **Neural Network** concept is inspired by the behaviour of human neural system. The fundamental unit of the neural system is the neuron that is formed by a cell body, an axon that give the output response and a dendritic tree that connect it with other nearby neurons (Figure 17)[5].

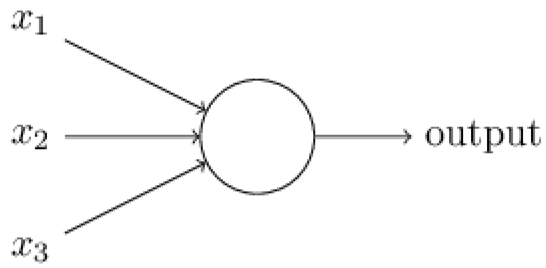



Figure 17: Fundamental Unit of Neural Network - Neural Networks and Deep Learning

In the same way the fundamental unit of neural network takes as input the features $\underline{x}$ and provides an output $y$. A Neural Network is a complex architecture composed by several fundamental units, the first vector of features is called **input layer**, the output is called **output layer** and between them can be placed one or more **hidden layer** that compose the neural network. Figure 18 shows an example of a simple architecture composed by two hidden layer, the first hidden layer is composed by two neurons while the second by three neurons.

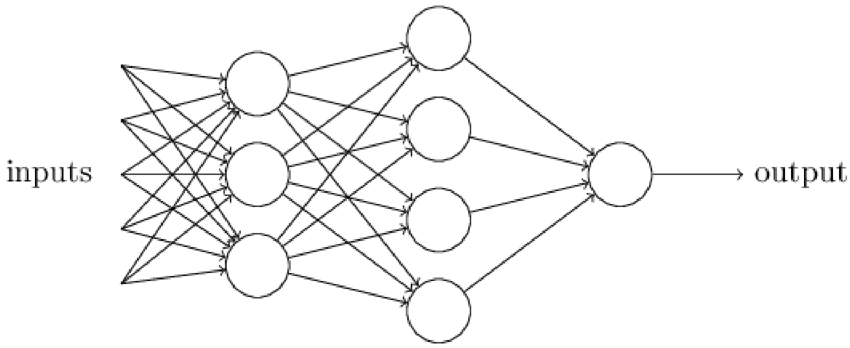



Figure 18: Example of Neural Network architecture with two hidden later - Neural Networks and Deep Learning

For the generic fundamental unit is $a_i^{(j)}$ is the activation of unit $i$ in layer $j$. Then is defined $\Theta^{(j)}$ called **matrix of weights**. Each activation unit is obtained as function of the previous layer's unit times the coefficient of the matrix of weights up to the final output $y$.
So for example for the first unit of the second layer:

$$a_1^{(2)} = f(\Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + ...) \tag{29}$$

If network has $s_j$ units in layer $j$ and $s_{j+1}$ in layer $j+1$ then $\Theta^{(j)} \in \mathcal{R}^{s_{j+1} * s_j}$.

For simplicity of notation we can define as follows, through an additional variable $z_i^{(j)}$:

$$a_i^{(j)} = f(z_i^j) \tag{30}$$

That indicates the summation of the input from the layer before times the coefficient of the weight matrix.

Then proceeding layer by layer the final output is obtained, this procedure is called **forward propagation** and is used to approximate complex and non-linear functions useful for Supervised Learning algorithms.

The problem is know the choice of $\Theta^{(j)}$ obtained by best-fit. The objective is to minimize the **cost function** that express how far our predictions are from the real outputs in the training set.

An example of cost function is the quadratic cost function:

$$J(\Theta) = \frac{1}{2n} \sum_{i=1}^{n} ||y - f(x)||^2 \tag{31}$$

And the goal is reach its global minimum, the algorithm that does this is **gradient descent**. In this algorithm all parameters are updated simultaneously by:

$$\Theta_{ij}^{(l)} := \Theta_{ij}^{(l)} - \alpha \frac{\partial}{\partial \Theta_{ij}^{(l)}} J(\Theta) \tag{32}$$

The coefficient $\alpha$ is called **learning rate** and has to be tuned in order to have a fast convergence of the iterative algorithm. In addition the derivative term must be computed and using discretization numeric tools like Finite Difference may be computationally too expensive. [5]

A useful tool in order to compute the derivative term is the **Backpropagation Algorithm**: for each node a new term $\delta_j^{(l)}$ is computed. This term is the error of node $j$ in layer $l$.[9]

$$\delta_j^{(l)} = a_j^{(l)} - y \tag{33}$$

And then from the last layer the algorithm goes backwards computing the errors on the previous layer up to the input layer.

$$\underline{\delta}^{(l-1)} = (\Theta^{(l-1)})^T \underline{\delta}^{(l)} f'(\underline{z}^{(l-1)}) \tag{34}$$

Finally is possible to prove [5] that:

$$\frac{\partial}{\partial \Theta_{ij}^{(l)}} J(\Theta) = a_j^{(l)} \delta_i^{(l+1)} \tag{35}$$

### 3.4 Q-learning.

**Q-learning** (Watkins, 1989) is a group of control algorithm defined by:

$$Q(s_t, a_t) := Q(s_t, a_t) + \alpha[R_{t+1} + \gamma max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \tag{36}$$

The action-value function $Q$ is an approximation of the optimal value function q, and this approximation can be done using Random Forest or Neural Network. This simplifies the analysis of the algorithm and guarantee faster convergence [6]. The action-value function is updated at each timestep of the episode (the time between the initialization and the terminal state of the algorithm).

Figure 19 shows the way the algorithm may be implemented, where action-value function $Q$ is iteratively updated.

**Q-learning (off-policy TD control) for estimating $\pi \approx \pi_*$**

Initialize $Q(s,a)$, for all $s \in \mathcal{S}, a \in \mathcal{A}(s)$, arbitrarily, and $Q(\textit{terminal-state}, \cdot) = 0$
Repeat (for each episode):
  Initialize $S$
  Repeat (for each step of episode):
    Choose $A$ from $S$ using policy derived from $Q$ (e.g., $\epsilon$-greedy)
    Take action $A$, observe $R$, $S'$
    $Q(S,A) \leftarrow Q(S,A) + \alpha\big[R + \gamma \max_a Q(S',a) - Q(S,A)\big]$
    $S \leftarrow S'$
  until $S$ is terminal

Figure 19: Q-learning Algorithm - Reinforcement Learning: An Introduction

### 3.5 HVAC application of Reinforcement Learning.

Thermostats for Supervisory Control of HVAC can be generally divided into manual and programmable [10]. In manual thermostats, the set-point is chosen by an operator and can be changed only by human intervention; in programmed ones the set-point can be set in order to achieve the best schedule for the requirements of the building.

The aim of this Master Thesis Degree is to test an autonomous HVAC system based on a Reinforcement Learning architecture with Q-learning algorithm, the $Q$ function will be approximated both with Random Forest and Neural Network.

As Reinforcement Learning approach the algorithm will need an objective that define the reward. The objective of this autonomous HVAC system is minimizing the energy consumption still guaranteeing thermal comfort. The chosen reward is always negative and the more is the energy consumption, the lower will be the reward. Furthermore will be considered another negative term for the thermal comfort chosen based only on air temperature according the design temperature values defined by manuals [14].

The state of the environment will be described by some parameters that could be obtained by common sensors already present in HVAC systems as temperature and humidity sensors, flow meter, multimiter and so on.

In order to test the Reinforcement Learning algorithm is necessary to build a model of the building that can give in output all the values that the standard sensoristic can provide.

When dealing with reinforcement learning algorithm is essential to let the algorithm explore the state and the space of possible actions in order to achieve the best energy savings. During the application on a real environment free exploration is not possible as the actions chosen by the algorithm can lead to dangerous or uncomfortable condition to occupants. The problem can be overcome by introduce a supervisor component in the environment that limit the set of actions made by the agent or impose default safe actions when the agent is likely to make big mistakes while explorating.

The problem of thermostat control belongs to the set of constrained optimization problem in which the objective of minimise the energy consumption can be achieved only fulfilling also air temperature requirements of the occupant. To define the constraint air temperature has been considered [10] while energy cost can be easily obtained by simulation results.

The issue of introduce these two different aspect of the constrained optimization problem in a single reward has already been faced up [11], the reward considered needs to strongly penalise the agent when its actions lead to do not respect air temperature constraint but also encourage energy savings. The reward obtained is expressed as follows:

$$r_t(s_t, a_t) = -cost(a_{t-1}) - PI_{t\in[t_s,t_e]}(I_{T_i \notin B}) \tag{37}$$

Where $I_{t\in[t_s,t_e]}$ is the indicator of the opening hour of the building $[t_s, t_e]$. $cost$ is the function that calculate the cost of energy consumption depending on the type of HVAC system that serve the occupied space. $P$ is a penalization factor that is given to the algorithm if the internal air temperature $T_i$ gets out of the boundary constraint $B$ that corresponds to the set of temperature between the heating and cooling set-points in dual set-point with dead band thermostat control.

In multiple zone thermostat control problem [12], the reward takes into account the presence of several zones by summation over the constraint violation of each zone. Reward formulation is slightly modified as follows:

$$r_t(s_{t-1}, a_{t-1}) = -cost(s_{t-1}, a_{t-1}) - PI_{t\in[t_s,t_e]}[(T_i - \overline{B})_+ + (\underline{B} - T_i)_+] \tag{38}$$

Where the operator $(x)_+ = max(x, 0)$. $\underline{B}$ is the lower bound of $B$ and $\overline{B}$ is its upper bound.

In this work the following reward function is assumed:

$$r_t(s_{t-1}, a_{t-1}) = -\alpha cost(s_{t-1}, a_{t-1}) - \beta I_{t\in[t_s,t_e]} e^{[(T_i - \overline{B})_+ + (\underline{B} - T_i)_+]} \tag{39}$$

The coefficients $\alpha$ and $\beta$ are chosen in order to adjust the energy savings components of the reward with respect to the air temperature constraint. The presence of the negative exponential function represents better the discrimination between the case in which air temperature $T_i$ is inside the boundary and the case in which it gets out from the set of $B$ as its value is at least $-\beta$.

The *cost* function is computed as the sum of the electric energy cost plus natural gas. Natural gas price is considered $cost_{gas} = 0.255994\, e/Sm^3$, that is the price of enhanced protection service. Electric energy cost is divided into three bands as represented in Figure 20 and each band has the cost specified in Figure 21.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lun - Ven | F3 | | | | | | | F2 | F1 | | | | | | | | | | | F2 | | | | F3 |
| Sab | F3 | | | | | | | F2 | | | | | | | | | | | | | | | | F3 |
| Dom / fest. | F3 | | | | | | | | | | | | | | | | | | | | | | | |

Figure 20: Division in bands of electric energy cost function of hour and day of the week.

| Voce / Fascia | | F1 | F2 | F3 |
|---|---|---|---|---|
| TRAS (c€/kWh) | | 0.95 | 0.50 | 0.15 |
| Distribuzione | Potenza max (€/kW/anno) | | 28 | |
| | Energia (c€/kWh) | | 0.15 | |
| MIS (€/punto/anno) | | | 650.00 | |
| Prezzo energia (corrisp. A CCA) (c€/kWh) | | 13.40 | 10.35 | 8.00 |
| COV (€/punto/anno) | | | 250.00 | |
| Σ componenti A, UC, ecc… | €/anno | | 150.00 | |
| | (c€/kWh) | | 1.00 | |
| Σ imposte (c€/kWh) | | | 1.00 | |

Figure 21: Electric energy cost divided in three bands.

Finally the general *cost* function can be expressed as sum of $cost_{gas}$ and $cost_{el}$ (sum of all the contribution for each band) starting from thermal and electric power consumed by the HVAC system $Q_{th}$ and $W_{el}$:

$$cost = cost_{gas}\frac{Q_{th}}{H_i} + cost_{el}W_{el} \tag{40}$$

Where $H_i$ is the lower heating value of natural gas equal to $9.61kWh/Sm^3$

# 4 Materials and Methods.

**EnergyPlus V. 8.5** is the software used to develop suitable building models to test the Reinforcement Learning control system. This older release has been used because is compatible with latest version of **Building Controls Virtual Test Bed 1.5.0** that allows external comunication of EnergyPlus to simulate complex control for HVAC systems [13]. The intent of EnergyPlus is to build model for HVAC designers and develop retrofit analysis for energy optimization in building [15].

## 4.1 Energy Balance Model.

In a complex environment as the dynamic of building in transient calculation, several equation and effects have to be taken into account [16].

The energy balance for **indoor air** is based on First Principle of Thermodynamics specified for each zone of the building simulated:

$$\begin{aligned} C_z \frac{dT_z}{dt} = & \sum_{i=1}^{N_{sl}} \dot{Q}_i + \sum_{i=1}^{N_{surfaces}} h_i A_i (T_{si} - T_z) \\ & + \sum_{i=1}^{N_{zones}} \dot{m}_i c_p (T_{zi} - T_z) + \dot{m}_{inf} c_p (T_\infty - T_z) + \dot{Q}_{sys} \end{aligned} \tag{41}$$

Where:

- $\sum_{i=1}^{N_{sl}} \dot{Q}_i$, is the sum of the convective heat loads;
- $\sum_{i=1}^{N_{surfaces}} h_i A_i (T_{si} - T_z)$, is the sum of convective heat transfer from the zone surfaces;
- $\sum_{i=1}^{N_{zones}} \dot{m}_i c_p (T_{zi} - T_z)$, is the heat transfer due to interzone air mixing;
- $dotm_{inf} c_p (T_\infty - T_z)$, is the heat transfer due to external air infiltrations;
- $\dot{Q}_{sys}$, is the air system output;
- $C_z \frac{dT_z}{dt}$, is the energy stored in the thermal zone. [16]

The energy balance of **indoor surface** (Figure 22) of building envelope that can be split in two parts the external energy is transferred across the envelope due to **conduction heat transfer** while reaching the internal side of the envelope splits into **convection heat transfer** and **radiation heat transfer**. While the heat fraction of convection is directly transferred to the indoor air, the radiation part is transmitted to the air and to the other walls or internal mass of the building and then reach the indoor air.

About the conduction heat in transient regime can be solved by correlation techniques, as **Conduction Transfer Function**, or by space discretization, **Finite Difference** or **Finite Elements**. EnergyPlus is able to use both of them but for this Thesis the space discretization technique has been used as provide higher stability for small timestep simulation.

Indoor air exchange heat by convection with the envelope's surfaces, with internal equipments and with the external air provided by ventilation or infiltration through windows. In order to guarantee thermal comfort an extra heat term must by provided or subtracted by the HVAC System.

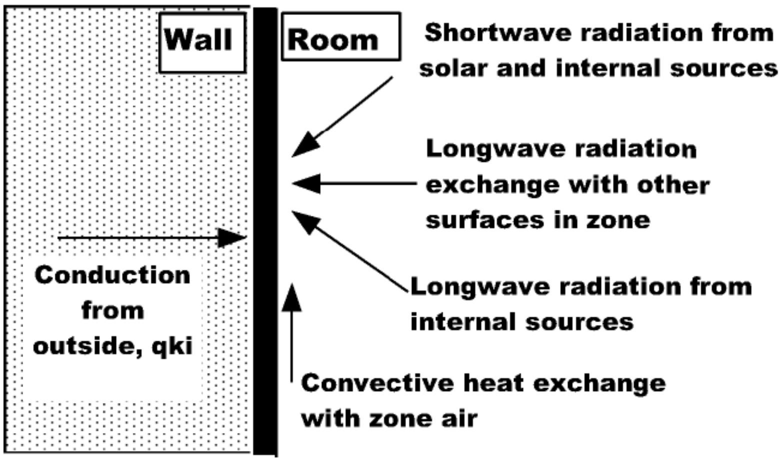



Figure 22: Interior Wall Heat Balance - EnergyPlus Version 8.5 Documentation - Engineering Reference

The heat power exchanged by ventilation can be computed knowing the air mass flow rate imposed by ventilation system or calculated if infiltration occurs.

$$\Phi_v = \dot{m}_a c_{p,a}(T_o - T_i) \tag{42}$$

Where:

- $\Phi_v$ is the heat power exchanged by ventilation or infiltration;
- $\dot{m}_a$ is the air mass flow rate;
- $c_{p,a}$ is the air specific heat;
- $T_o$ and $T_i$ are the outdoor and indoor air temperature.

Indoor air exchange also with internal surfaces, like envelope surface or other equipment. This convective heat exchange calculation is simplified by the convective coefficient $h_c$. The following equation represent heat exchange between indoor air and $j$ surface.

$$\Phi_c = h_{c,j} A_j (T_{s,j} - T_i) \tag{43}$$

Where:

- $\Phi_c$ is the heat power exchanged by convection;
- $h_{c,j}$ is the convective coefficient with the $j$ surface;
- $A_j$ is the area;
- $T_{s,j}$ is the internal surface temperature of $j$.

The internal surface temperature is influenced by the heat balance on **external surfaces** (Figure 23) considering convection with external air, conduction inside the wall, radiation exchange with other nearby buildings and **solar radiation**. Solar radiation also enter the indoor environment by glazing system and exchange directly with internal surfaces.

Inside the building environment **internal heat gains** such as occupants latent and sensible heat, lighting, gas and electrical equipment, refrigeration case and all the heat

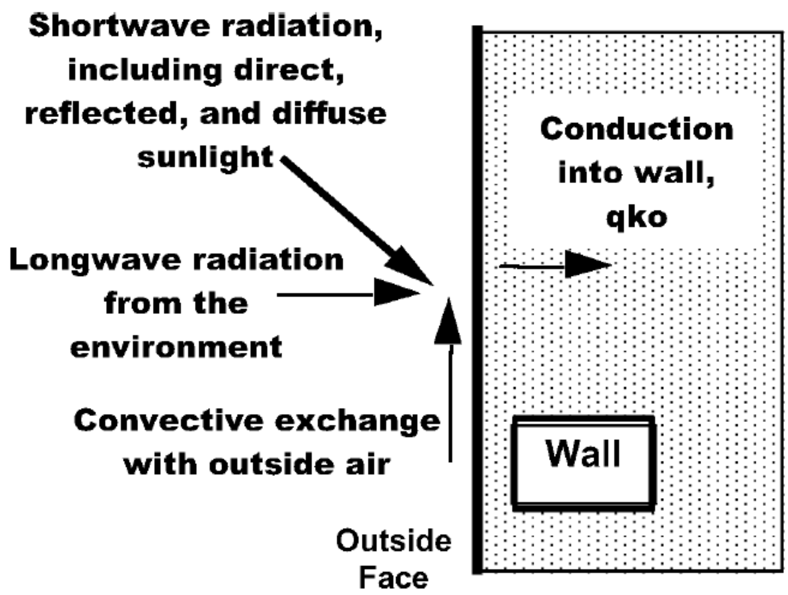



Figure 23: Exterior Wall Heat Balance - EnergyPlus Version 8.5 Documentation - Engineering Reference

power generated by human activity are taken into account and exchange by convection and radiation with indoor air.

The overall indoor air balance is solved integrated with system and HVAC loop equations in a simplified **Thermal Zone** approach where each zone is seen as a single point as well as all the surfaces constituting the building envelope.

## 4.2 Commercial Retail.

The first model that has been developed is a real existing **Commercial Retail** located near Bergamo:

- Latitude 45.67°
- Longitude 9.7°
- Elevation 238 $m$
- Climate Zone **E** [26]
- Degree Days 2533°$C$

Climate data for Bergamo has been acquired by the "Agenzia Regionale per la Protezione dell'Ambiente" ARPA for the region Lombardia (www.arpalombardia.it). The data acquired cover last 3 years of Temperature, Relative Humidity and Solar Radiation; the other input data necessary for the EnergyPlus simulation are taken by the .epw file built by DOE that represent medium value in last 10 years for the city of Bergamo.

### 4.2.1 Geometric Model.

EnergyPlus allows to use external software to easily develop the geometric, structural and utilization model. The geometric model has been developed with **SketchUp 2016** (`www.sketchup.com`) and is represented in Figure 24.

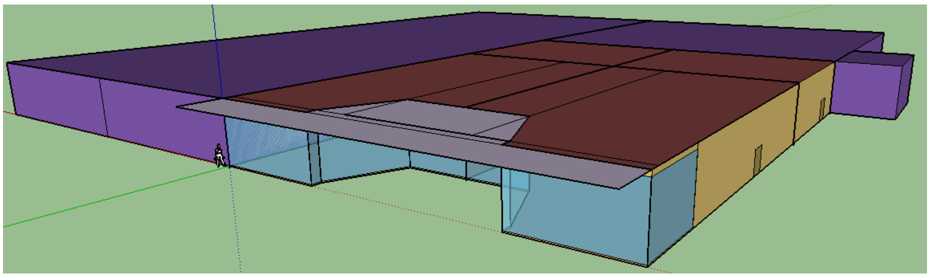


Figure 24: Commercial Retail, Geometric Model.

In Figure 24 the green line represent North direction and is already possible to distinguish opaque walls, windows, frames and ceiling. The purple buildings are the buildings nearby that could shadow from solar radiation in West and South direction.

The Commercial Retail model has an area of 1497 $m^2$ in one floor with a height of around 4 $m$ in a unique large room with no internal partitions.

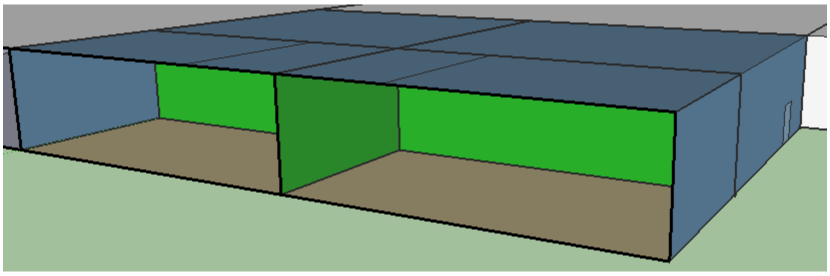


Figure 25: Commercial Retail, Boundary Condition.

Figure 25 represent the boundary condition for each surface necessary for energy balance calculation. Blue surfaces are external surfaces, green walls are internal partitions, brown surfaces are in contact with the ground. Boundary conditions for external walls that are in contact with nearby buildings have been then modified with adiabatic boundary condition supposing that the other buildings are subjected also to air conditioning so that their air temperature is different from outdoors.

Final step is to identify Thermal Zones. The commercial retail model has been divided into four thermal zones (as shown in Figure 26) centered in the position of thermometer in order to compare the model with the real building for validation.

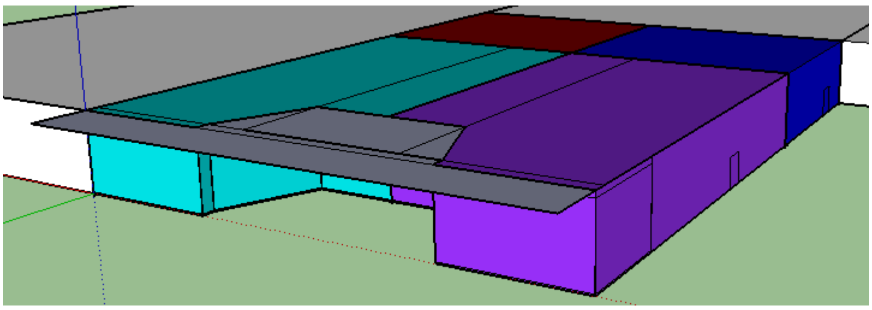


Figure 26: Commercial Retail, Thermal Zones.

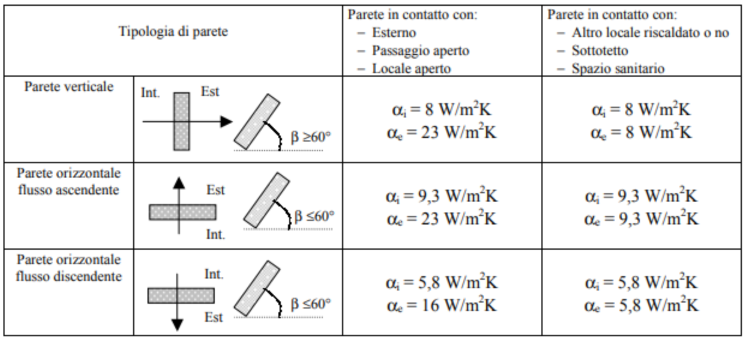
### 4.2.2 Structural Model.

Thermal properties of the envelope are resumed by the following tables in terms of conductivity ($\lambda$), thickness ($t$), specific heat ($c_p$) and density ($\rho$) for exterior walls (Table 3), floor (Table 5) and ceiling (Table 4).

Thermal resistance and U-value are computed according to the follows:

$$R_t = \frac{t}{\lambda} \tag{44}$$

Considering also the additional thermal resistances that could also be described by a coefficient $\alpha$ that assumes different values for the inclination and if regards internal or external layer [20] with the values represented in Table 2.



| Tipologia di parete | | Parete in contatto con: – Esterno – Passaggio aperto – Locale aperto | Parete in contatto con: – Altro locale riscaldato o no – Sottotetto – Spazio sanitario |
|---|---|---|---|
| Parete verticale | Int. Est $\beta \geq 60°$ | $\alpha_i$ = 8 W/m²K<br>$\alpha_e$ = 23 W/m²K | $\alpha_i$ = 8 W/m²K<br>$\alpha_e$ = 8 W/m²K |
| Parete orizzontale flusso ascendente | Est Int. $\beta \leq 60°$ | $\alpha_i$ = 9,3 W/m²K<br>$\alpha_e$ = 23 W/m²K | $\alpha_i$ = 9,3 W/m²K<br>$\alpha_e$ = 9,3 W/m²K |
| Parete orizzontale flusso discendente | Int. Est $\beta \leq 60°$ | $\alpha_i$ = 5,8 W/m²K<br>$\alpha_e$ = 16 W/m²K | $\alpha_i$ = 5,8 W/m²K<br>$\alpha_e$ = 5,8 W/m²K |

Table 2: Adduction Coefficient - Elementi di Termofisica Generale e Applicata.

$$R_t = \frac{1}{\alpha} \tag{45}$$

Finally the U-value that is representative for the global heat transfer by conduction in stationary regime can be computed as follows where $n$ is the total number of layers:

$$U = \frac{1}{\sum_{i=1}^{n} R_t} \tag{46}$$

Values obtained for external walls, ceiling and floor are described respectively in Tables 3, 4 and 5.

The glazing system is composed by a two-layer window of 4 mm plus an internal air gap of 16 mm. According to the data sheet (www.agc-yourglass.com) of the glazing system it has a U-value of $U_g = 0.8\frac{W}{m^2K}$ but also **thermal bridges** have been considered [21]. For this application has been considered the linear thermal bridge that get rise in the

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.130 |
| Interior Concrete | 0.060 | 1.910 | 1000 | 2450 | 0.031 |
| Lightened Layer | 0.100 | 0.044 | 845 | 30 | 5.000 |
| Air Gap | - | - | - | - | 0.160 |
| External Concrete | 0.060 | 2.080 | 1000 | 2450 | 0.029 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.376 |

Table 3: Vertical Wall Stratigraphy for Commercial Retail model.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.100 |
| Mortar | 0.010 | 1.400 | 840 | 2000 | 0.007 |
| Insulation Layer | 0.080 | 0.034 | 850 | 30 | 5.882 |
| Ceiling Bricks | 0.180 | 0.599 | 920 | 950 | 0.301 |
| Expanded Clay | 0.020 | 0.756 | 880 | 1700 | 0.026 |
| Shingles | 0.020 | 0.260 | 880 | 1300 | 0.077 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.344 |

Table 4: Ceiling Stratigraphy for Commercial Retail model.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.170 |
| Floor Tiles | 0.010 | 1.000 | 840 | 2300 | 0.010 |
| Concrete | 0.050 | 1.160 | 1000 | 2000 | 0.043 |
| Insulation Layer | 0.050 | 0.041 | 1030 | 97 | 1.220 |
| Backing | 0.060 | 0.093 | 880 | 2200 | 0.645 |
| Shingles | 0.220 | 0.667 | 920 | 1214 | 0.330 |
| Mortar | 0.015 | 0.900 | 910 | 1800 | 0.017 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.404 |

Table 5: Floor Stratigraphy for Commercial Retail model.

connection between the window and the frame with a coefficient $\psi_g = 0.03$. The frame is made of wood with a U-value of $U_f = 3.7\frac{W}{m^2K}$. Then the total U-value of the glazing system has been computed as average surface value between the glazing system and the frame [21].

$$U_w = \frac{U_g A_g + U_f A_f + \psi_g l_w}{A_g + A_f} = 1.990\frac{W}{m^2K} \tag{47}$$

Where $l_w$ is the perimeter of the glaze and $U_w$ is the U-value of the total glazing system. Finally from data sheet the **solar factor** has been set $g = 0.75$.

According to UNI EN 1634-1:2018 fire doors stratigraphy must respect REI 60 condition so resist to fire, heat and gas release for at least 60 minutes. The stratigraphy as been chosen as expressed in Table 6 [22].

For each material, other data required by EnergyPlus are:

- Thermal absorbance, set to 0.9;
- Roughness, in qualitative terms from *very smooth* to *very rough*;

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.130 |
| Galvanized Steel | 0.010 | 50.00 | 502 | 7300 | 0.0002 |
| Insulation Layer | 0.200 | 0.034 | 850 | 30 | 5.882 |
| Galvanized Steel | 0.010 | 50.00 | 502 | 7300 | 0.0002 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.169 |

Table 6: Door Stratigraphy for Commercial Retail model.

- Solar and visible absorbance, set to 0.2 for metallic material and 0.8 for all the others [17].

#### 4.2.3 Utilization Model.

As already seen in the analysis of energy balance of building are considered five different type of loads: occupancy, lighting, internal equipment, infiltration and ventilation.

**Occupancy** is an internal gain so is helpful in winter while increase energy consumption in summer season. The sensible and latent heat exchanged with indoor air depends on the **metabolic rate** or **activity level** that depends on the activity people are performing, it normally range from 100 to 150 W per person during office activities but can reach up to 900 W per person during hard physical activities [17]. Table 7 shows most common metabolic rates for various activities, the value is also represented in $\frac{W}{m^2}$ of surface body computed by the previous dividing by 1.8 $m^2$ that is the average surface for an average person. Finally another unit of measurement is introduced: 1 MET = 58.1 $\frac{W}{m^2}$ that correspond to a seated person's activity level [23].

| Activity | Activity Level w/ Person EnergyPlus Schedule Value | Activity Level W/m² | met* |
|---|---|---|---|
| *Resting* | | | |
| Sleeping | 72 | 40 | 0.7 |
| Reclining | 81 | 45 | 0.8 |
| Seated, quiet | 108 | 60 | 1 |
| Standing, relaxed | 126 | 70 | 1.2 |
| *Walking (on level surface)* | | | |
| 3.2 km/h (0.9 m/s) | 207 | 115 | 2 |
| 4.3 km/h (1.2 m/s) | 270 | 150 | 2.6 |
| 6.4 km/h (1.8 m/s) | 396 | 220 | 3.8 |
| *Office Activities* | | | |
| Reading, seated | 99 | 55 | 1 |
| Writing | 108 | 60 | 1 |
| Typing | 117 | 65 | 1.1 |
| Filing, seated | 126 | 70 | 1.2 |
| Filing, standing | 144 | 80 | 1.4 |
| Walking about | 180 | 100 | 1.7 |
| Lifting/packing | 216 | 120 | 2.1 |
| *Miscellaneous Occupational Activities* | | | |

Table 7: Metabolic Rates for Various Activities - EnergyPlus Version 8.5 Documentation - Input Output Reference.

Activity Level for typical commercial retail building so is 180 W per person that correspond to walking about activity like shopping.

Another important parameter related to people occupation in building to asses indoor air quality is **Carbon Dioxide generation rate**. The default value suggested by UNI norms is 3.82E-8 $\frac{m^3}{sW}$ related to activity people but the real value could actually be ten times bigger [17], anyway the default value is been considered in this simulation.

Last parameter considered is the thermal insulation of people due to **clothing**, the unit of measurement is called CLO where 1 CLO = 0.155 $\frac{m^2K}{W}$ corresponding to a normal winter suite [23]. The ASHRAE proposal for the clothing calculation takes into account air outdoor temperature at 6:00 AM and the corresponding CLO value is obtained by Figure 27 (**Dynamic Clothing Simulation Method**).

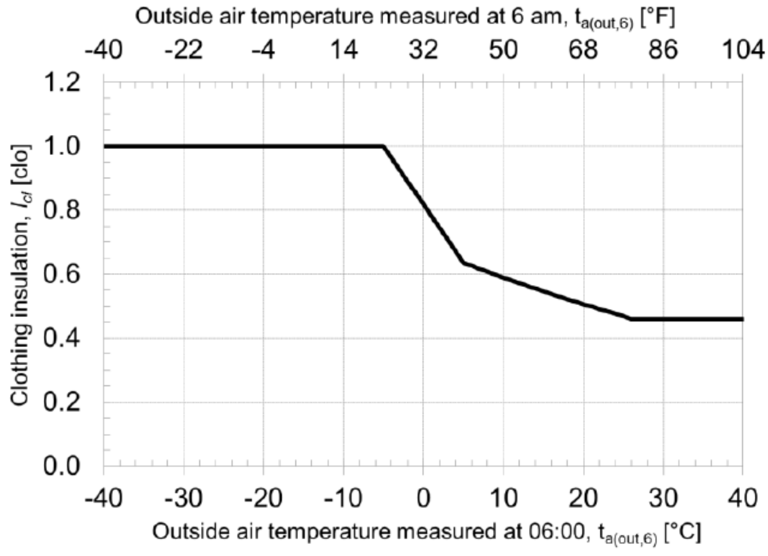



Figure 27: Dynamic Clothing Simulation Method - EnergyPlus Version 8.5 Documentation - Input Output Reference.

For commercial retail application the occupants of the building have been divided into two groups: customers and employees. A total amount of 80 customers and 10 employees can occupy the building simultaneously. Imitating the real behaviour of the schedule of occupancy of a real commercial retail half people are considered presents during mornings in weekdays. The retail is open from 8:00 to 21:00 during weekdays and from 8:00 to 20:00 on Sunday.

**Lighting load** has been set to 45.3 $\frac{W}{m^2}$ of floor area by design while no **electric equipments** are considered as in a typical commercial retail electric equipments are not strongly present. Lighting schedule respect opening schedule of the commercial retail located near Bergamo as already seen for occupancy load.

**Infiltration** is a mass flow that enter from outside the building to the inside, passing trough cracks in windows or in frames junction. As it is at exterior air temperature always cause an increase in energy consumption both in winter and in summer season. It is only present when the commercial retail is closed as the HVAC system is usually designed in order to have a small over-pressure inside the building to avoid infiltration losses. So will be present from 21:00 to 8:00 during weekdays and from 20:00 on Sunday. Infiltration mass flow rate depends both on the air temperature difference between indoor and outdoor and wind velocity ($w$) according the following equation [17]:

$$\dot{m}_{inf} = I_{design} F_{schedule} [A + B(T_i - T_o) + Cw + Dw^2] \tag{48}$$

The result of this equation must be non negative as the infiltration always goes from the outside of the building to the inside while exfiltration should be modelled differently. Suggested values for coefficients A, B, C and D are [17]:

- $A = 0.606$;
- $B = 0.03636$;
- $C = 0.1177$;
- $D = 0$.

The value $F_{schedule}$ identifies the schedule in which infiltrations are present and is 1.0 when HVAC system is in off mode, 0.0 otherwise. $I_{design}$ is the maximum infiltration flow rate allowed set by default 0.00030226 in terms of flow rate per exterior wall area [17].

Also **ventilation** is an external air flow rate but is voluntary in order to guarantee indoor air quality. According to UNI 10339:1995 this quantity has been set to 40 $\frac{m^3}{h}$ [24], fraction of outdoor air with respect to total flow is set to 0.3 so 70% of total mass flow rate is taken by the inside of the building.

#### 4.2.4 HVAC system.

HVAC system for the commercial retail model is an **all-air** system, in which both ventilation and air conditioning are provided by supply air. HVAC system has two separated air loops that handle half of the total amount of ventilation supply air. Each loop is driven by a **Fan** with constant velocity and uncontrolled **Air Diffusers** as terminals to the zones. The fan, that is able to give a pressure rise of 250 Pa has a total efficiency equal to 0.7 that takes into account mechanic, volumetric and hydraulic losses and is driven by an electric motor with efficiency equal to 0.9; all the thermal losses of the electric motor are transferred to the air flow.

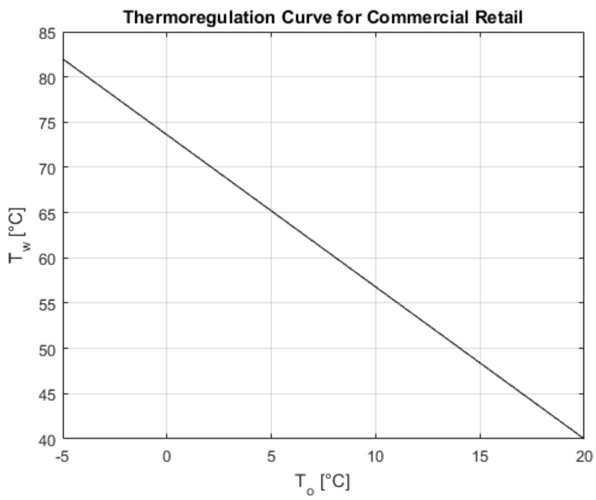



Figure 28: Thermoregulation Curve for Commercial Retail.

All data have been obtained by plant data sheets available while all the other data needed for the simulation have been obtained by **sizing** of the system considering Bergamo design day with minimum design temperature of -5 $°C$ in winter and maximum temperature of 32.2 $°C$ in summer with daily variation $\Delta T = 10.1°C$ [**?**].

Then the plant model is completed by two secondary water loop for hot water and chilled water that send water to decks in AHU as explained before with by-pass valve.

Hot water **boiler** considered is a standard boiler of 60 kW with a design efficiency of 0.8, the design water outlet temperature is 82 $°C$ but is modulated according thermoregulation curve represented in Figure 28, it range from 82 $°C$ in design condition to 40 $°C$ at minimum load. All the relevant parameter are reported in Table 8.

The electric chiller is water cooled with a design cooling capacity of 50 kW and **COP** = 3, the COP (Coefficient of Performance) is the ratio between thermal energy subtracted to the air and elctric energy in input to the chiller. All the relevant parameter are reported in Table 9 [17].

Behaviours at **partial load** is obtained by three curves. The first is used to obtain

| | |
|---|---|
| Fuel Type | Natural Gas |
| Nominal Capacity | 150 kW |
| Nominal Efficiency | 0.8 |
| Design Outlet Water Temperature | 82 $°C$ |
| Design Water Flow Rate | 0.005 $\frac{m^3}{s}$ |

Table 8: Design values Boiler for Commercial Retail Model.

| | |
|---|---|
| Nominal Capacity | 120 kW |
| Nominal COP | 3 |
| Design Outlet Chilled Water Temperature | 6.67 $°C$ |
| Design Inlet Condenser Water Temperature | 29.4 $°C$ |
| Design Chilled Water Flow Rate | 0.002 $\frac{m^3}{s}$ |
| Design Condenser Water Flow Rate | 0.041 $\frac{m^3}{s}$ |

Table 9: Design values Chiller for Commercial Retail Model.

the actual cooling capacity ($\dot{Q}_{cool}$) function of the outlet chilled water temperature ($T_{out}$) and inlet condenser water temperature ($T_{in}$) [17]:

$$\dot{Q}_{cool} = \dot{Q}_{design}[A + B_1 T_{out} - B_2 T_{out}^2 + C_1 T_{in} - C_2 T_{in}^2 - D T_{out} T_{in}] \tag{49}$$

Where the coefficients are the following [17]:

- $A = 0.258$;
- $B_1 = 0.0389$;
- $B_2 = -0.000217$;
- $C_1 = 0.0469$;
- $C_2 = -0.000943$;
- $D = -0.000343$.

The same equation is also used to obtain the real electric output with other coefficients:

- $A = 0.934$;
- $B_1 = -0.0582$;
- $B_2 = 0.0045$;
- $C_1 = 0.00243$;
- $C_2 = 0.000486$;
- $D = -0.00122$.

Last adjustment for part load is made on EIR (**Energy Input Ratio**) that is the inverse of the COP, function of the part load ratio (PLR) that is the ratio between the actual cooling capacity and the design cooling capacity [17]:

$$EIR = EIR_{design}[0.222903 + 0.313387PLR + 0.46371PLR^2] \tag{50}$$

Piping system has been modelled in order to consider also heat exchange between water and internal air. Pipe are in cross-linked polyethylene (PEX) and have been sized in terms of thickness according to [25]. Thermal properties are listed in Table 10.

| | |
|---|---|
| $s\,[m]$ | 0.07 |
| $\lambda\,[\frac{W}{m^2K}]$ | 0.34 |
| $\rho\,[\frac{kg}{m^3}]$ | 940 |
| $c_p\,[\frac{J}{kgK}]$ | 1500 |

Table 10: PEX thermal properties

| | Heating Mode | Cooling Mode |
|---|---|---|
| Lower SP | 16 | 23 |
| Upper SP | 19 | 25 |

Table 11: Design Setpoint for Commercial Retail.

#### 4.2.5 Control System.

The HVAC System is controlled by a PI control (see Equation 2) that has as controlled variable the supply air temperature to reach the set-point obtained by [14] in order to reach thermal comfort requirement for different building application.

Thermostat provide a **Single Set-point**; two temperatures values are set and no action occurs while internal air temperature is in between these two values. Heating is available from 15 October to 15 April for 14 hours per day maximum while the rest of the year the HVAC system is in cooling mode [26].

For commercial retail applications the values are represented in Table 11.

The error signal is computed between the set-point and the actual internal air temperature.

The control system in heating mode has been tuned with Ultimate Oscillation method while for cooling mode has been tuned with trial and error. The results are listed in Table 12.

| | Heating Mode | Cooling Mode |
|---|---|---|
| $K_p$ | 7.21 | 6.5 |
| $K_i$ | 0.434 | 0.04 |

Table 12: Commercial Retail model with boiler and water chiller tuning.

### 4.3 Reference Building: Supermarket.

The **U.S. Department of Energy** (DOE) developed **commercial reference buildings** or commercial building benchmark [18]. These buildings model, defined by the National Renewable Energy Laboratory (NREL) [19], are designed to cover among 70% of commercial buildings types in United States with 16 different models (listed in Table 13).

The aim of these models is to represent both new and existing buildings to simulate the result provided by new technologies in energy efficiency in buildings, analyse advanced control system, daylighting studies and indoor air quality studies [19].

Location were selected to cover significantly high portion of U.S. buildings along all the climate zones. For each climate zone the most populous city was chosen [19] as representative to locate the reference building. Climate zones were divided using a number that identify the degree days plus a letter that subdivide in moist, dry and marine clime. Final division is represented in Figure 29 and Table 14.

| BUILDING TYPE NAME | FLOOR AREA (FT$^2$) | NUMBER OF FLOORS |
|---|---|---|
| Large Office | 498,588 | 12 |
| Medium Office | 53,628 | 3 |
| Small Office | 5,500 | 1 |
| Warehouse | 52,045 | 1 |
| Stand-alone Retail | 24,962 | 1 |
| Strip Mall | 22,500 | 1 |
| Primary School | 73,960 | 1 |
| Secondary School | 210,887 | 2 |
| Supermarket | 45,000 | 1 |
| Quick Service Restaurant | 2,500 | 1 |
| Full Service Restaurant | 5,500 | 1 |
| Hospital | 241,351 | 5 |
| Outpatient Health Care | 40,946 | 3 |
| Small Hotel | 43,200 | 4 |
| Large Hotel | 122,120 | 6 |
| Midrise Apartment | 33,740 | 4 |

Table 13: Reference Buildings List - Commercial Reference Building

| CLIMATE ZONE | REPRESENTATIVE CITY |
|---|---|
| 1A | Miami, Florida |
| 2A | Houston, Texas |
| 2B | Phoenix, Arizona |
| 3A | Atlanta, Georgia |
| 3B-Coast | Los Angeles, California |
| 3B | Las Vegas, Nevada |
| 3C | San Francisco, California |
| 4A | Baltimore, Maryland |
| 4B | Albuquerque, New Mexico |
| 4C | Seattle, Washington |
| 5A | Chicago, Illinois |
| 5B | Boulder, Colorado |
| 6A | Minneapolis, Minnesota |
| 6B | Helena, Montana |
| 7 | Duluth, Minnesota |
| 8 | Fairbanks, Alaska |

Table 14: Climate Zone U.S. - Commercial Reference Building

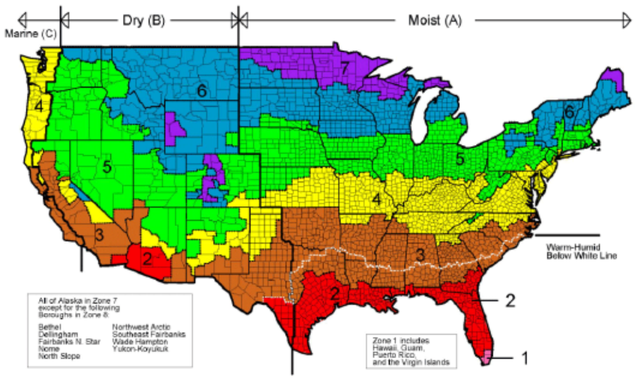



Figure 29: Climate Zone U.S. - Commercial Reference Building.

Among all the models listed in Table 13 **Supermarket** has been chosen because it merge different zones with different requirements (bakery, office, storage, sales area) so the Reinforcement Learning algorithm will take different decision and strategies according to different conditions and schedules that could be taken into account.

In order to test the robustness of the algorithm to different climate condition the model has been located first in Bergamo with the same data taken by ARPA used for the Commercial Retail model. Then the model has been placed in a warmer zone in south Italy:

- Latitude 38.19°
- Longitude 15.55°
- Elevation 29 $m$
- Climate Zone **B** [26]
- Degree Days 707°$C$

As no climate data for recent years were available for this location for the simulation where considered three different **EnergyPlus Weather Format** files available in Sicily to model warmer climate zones. EnergyPlus Weather Format files (.epw extension) are available in `www.energyplus.net` and includes climate data in terms of air temperature, relative humidity, wind velocity, solar radiation and all the other parameters necessary for the EnergyPlus model. In this project .epw files regarding *Messina*, *Catania* and *Gela* cities were used.

### 4.3.1 Geometric Model.

The Supermarket his developed in one floor with total area of 4181 $m^2$ and height of 6.10 $m$.

As Figure 30 shows the **Supermarket** is composed by 5 different environments. The biggest room is the central **Sales** area with a wide window South oriented. On the left of the picture there is the **Produce** area, North exposed there are the **Dry Storage** and the **Office** on the North-West corner, then on the right **Bakery** area on the South-West corner and **Deli**.

Figure 31 shows the rendering of the model by thermal zones:

- Produce, in orange;

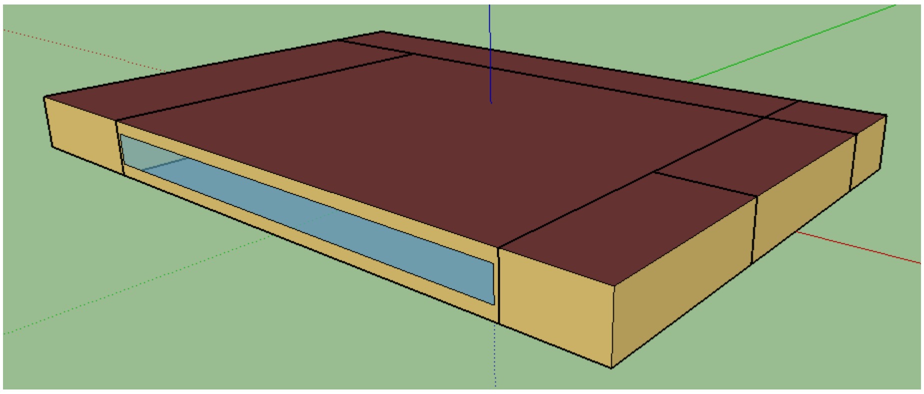


Figure 30: Supermarket, Geometric Model.

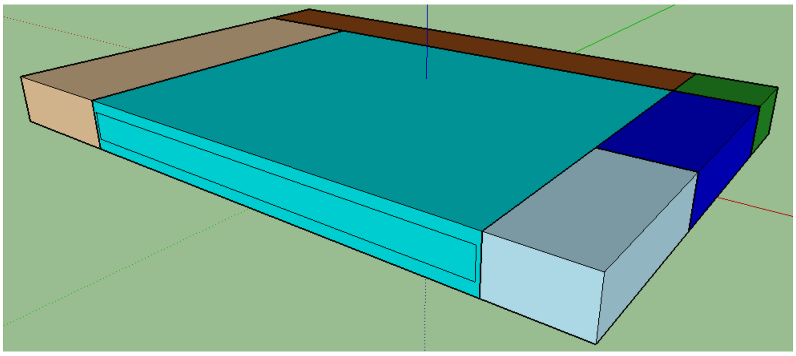


Figure 31: Supermarket, Thermal Zones.

- Dry Storage, in brown;
- Sales, in cyan;
- Bakery, in light blue;
- Deli, in Blue;
- Office, in Green.

#### 4.3.2 Structural Model.

Original stratigraphy of the model has been modified in order to accomplish law requirements [27] that are different for the climate zones taken into accounts. Table 15 shows

| Zona climatica | strutture opache verticali | strutture opache orizzontali o inclinate | | finestre comprensive di infissi |
|---|---|---|---|---|
| | | Coperture | Pavimenti (*) | |
| A | 0,56 | 0,34 | 0,59 | 3,9 |
| B | 0,43 | 0,34 | 0,44 | 2,6 |
| C | 0.36 | 0.34 | 0,38 | 2,1 |
| D | 0,30 | 0,28 | 0,30 | 2,0 |
| E | 0,28 | 0,24 | 0,27 | 1,6 |
| F | 0,27 | 0,23 | 0,26 | 1,4 |

Table 15: Law requirments for U-value of building built after $1^{st}$ January 2010 - Decreto ministeriale 11 Marzo 2008.

the maximum U-value requirement for new buildings (built after $1^{st}$ January 2010).

Two different envelopes have been modelled in order to accomplish the two different requirements for climate zone **E** (Bergamo) and **B** (Messina). Material's properties and layers are constant but change in thickness to reach different requirements in term of U-values given by Table 15.

Tables 16, 17, 18, shows the envelope properties for the model located in climate zone **E**, furthermore glazing system has a U-value equal to 1.6 $[\frac{W}{m^2K}]$ with a solar factor of 0.39.

Tables 19, 20, 21, shows the envelope properties for the model located in climate zone **B**, furthermore glazing system has a U-value equal to 2.6 $[\frac{W}{m^2K}]$ with a solar factor of 0.39.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.130 |
| Interior Concrete | 0.080 | 1.160 | 880 | 2000 | 0.069 |
| Polystyrene Vapour Barrier | 0.001 | 0.060 | 2100 | 50 | 0.017 |
| Expanded Polystyrene | 0.12 | 0.038 | 1250 | 30 | 3.158 |
| Air Gap | 0.08 | - | - | - | 0.150 |
| External Concrete | 0.040 | 1.160 | 880 | 2000 | 0.034 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.278 |
| | | | | Limit Value [27] | 0.280 |

Table 16: Vertical Wall Stratigraphy for Supermarket model in climate zone **E**.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.100 |
| Stoneware Floor | 0.002 | 1.000 | 840 | 2300 | 0.020 |
| Concrete Casting Underlay | 0.050 | 0.250 | 920 | 120 | 0.200 |
| Polystyrene Vapour Barrier | 0.001 | 0.060 | 2100 | 50 | 0.017 |
| Expanded Polystyrene | 0.12 | 0.038 | 1250 | 30 | 3.158 |
| Concrete Layer | 0.050 | 1.91 | 880 | 2000 | 0.026 |
| Alveolar Slab | 0.080 | 0.667 | 880 | 2000 | 0.120 |
| Waterproof Layer | 0.001 | 0.230 | 1000 | 1100 | 0.004 |
| Lean Concrete | 0.050 | 0.250 | 920 | 1200 | 0.200 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.257 |
| | | | | Limit Value [27] | 0.270 |

Table 17: Ceiling Stratigraphy for Supermarket model in Climate Zone **E**.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.170 |
| Alveolar Slab | 0.150 | 0.667 | 880 | 2000 | 0.120 |
| Concrete Layer | 0.050 | 1.91 | 880 | 2000 | 0.026 |
| Polystyrene Vapour Barrier | 0.001 | 0.060 | 2100 | 50 | 0.017 |
| Expanded Polystyrene | 0.12 | 0.038 | 1250 | 30 | 3.158 |
| Concrete Casting Underlay | 0.050 | 0.250 | 920 | 120 | 0.200 |
| Bitumen | 0.080 | 0.170 | 920 | 1200 | 0.471 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.238 |
| | | | | Limit Value [27] | 0.240 |

Table 18: Floor Stratigraphy for Supermarket model in Climate Zone **E**.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.130 |
| Interior Concrete | 0.100 | 1.160 | 880 | 2000 | 0.086 |
| Polystyrene Vapour Barrier | 0.001 | 0.060 | 2100 | 50 | 0.017 |
| Expanded Polystyrene | 0.070 | 0.038 | 1250 | 30 | 1.842 |
| Air Gap | - | - | - | - | 0.150 |
| External Concrete | 0.080 | 1.160 | 880 | 2000 | 0.069 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.429 |
| | | | | Limit Value [27] | 0.430 |

Table 19: Vertical Wall Stratigraphy for Supermarket model in climate zone **B**.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.100 |
| Stoneware Floor | 0.001 | 1.000 | 840 | 2300 | 0.010 |
| Concrete Casting Underlay | 0.050 | 0.250 | 920 | 120 | 0.200 |
| Polystyrene Vapour Barrier | 0.001 | 0.060 | 2100 | 50 | 0.017 |
| Expanded Polystyrene | 0.070 | 0.038 | 1250 | 30 | 1.842 |
| Concrete Layer | 0.050 | 1.91 | 880 | 2000 | 0.026 |
| Alveolar Slab | 0.010 | 0.667 | 880 | 2000 | 0.015 |
| Waterproof Layer | 0.001 | 0.230 | 1000 | 1100 | 0.004 |
| Lean Concrete | 0.020 | 0.250 | 920 | 1200 | 0.080 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.428 |
| | | | | Limit Value [27] | 0.440 |

Table 20: Ceiling Stratigraphy for Supermarket model in Climate Zone **B**.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.170 |
| Alveolar Slab | 0.300 | 0.667 | 880 | 2000 | 0.240 |
| Concrete Layer | 0.050 | 1.91 | 880 | 2000 | 0.026 |
| Polystyrene Vapour Barrier | 0.001 | 0.060 | 2100 | 50 | 0.017 |
| Expanded Polystyrene | 0.070 | 0.038 | 1250 | 30 | 1.842 |
| Concrete Casting Underlay | 0.050 | 0.250 | 920 | 120 | 0.200 |
| Bitumen | 0.080 | 0.170 | 920 | 1200 | 0.471 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.333 |
| | | | | Limit Value [27] | 0.340 |

Table 21: Floor Stratigraphy for Supermarket model in Climate Zone **B**.

### 4.3.3 Utilization Model.

The occupancy profile has been maintained from the original DOE model with closing day on Sunday and opening from 8:00 to 23:00 but number of people (or number of people per $m^2$) has been chosen differently according to Italian standard UNI EN 10339 [24]. Occupancy activity level has been chosen according to Table 7 looking at the different activities conducted in the building, all quantities are resumed in Table 22.

| Thermal Zone | Occupancy | Activity | Activity Level [MET] |
|---|---|---|---|
| Bakery | 3 [people] | Cooking | 1.8 |
| Deli | 3 [people] | Packing | 2 |
| Dry Storage | 10 $[\frac{m^2}{people}]$ | Packing | 2 |
| Office | 16.667 $[\frac{m^2}{people}]$ | Reading, seated | 1 |
| Produce | 3 [people] | Packing | 2 |
| Sales | 5 $[\frac{m^2}{people}]$ | Walking about | 1.7 |

Table 22: Occupancy and Activity Level for Supermarket Model.

Dynamic Clothing Model has been adopted [17] for all the zones with two exceptions. In Bakery and Deli that have great internal gains is supposed to maintain for workers a constant clothing along the year equal to 0.5 CLO.

Lighting and electric equipment load has been modified according to Italian standard UNI EN 15232 [28], value used for each zone are resumed in Table 23. Deli and Bakery have also additional internal gains due to specific activities respectively of 12105 W and 11244 W.

| Thermal Zone | Lighting $[\frac{W}{m^2}]$ | Electric Equipments $[\frac{W}{m^2}]$ |
|---|---|---|
| Bakery | 15 | 15 |
| Deli | 15 | 15 |
| Dry Storage | 10 | 3.5 |
| Office | 15 | 10 |
| Produce | 15 | 15 |
| Sales | 15 | 3.5 |

Table 23: Lighting and Electric Equipment Loads for Supermarket Model.

Other equipment modelled are **Refrigeration Cases** (Table 24).

Ventilation requirement in terms of outdoor air have been specified according to UNI EN 10339 [24], these values are specified in Table 25 with recirculation fraction of 0.3.

| Thermal Zone | Unit of Measurement | Outdoor Air Flow |
|---|---|---|
| Bakery | Flow/Area | 0.0165 |
| Deli | Flow/Area | 0.0065 |
| Dry Storage | Flow/Area | 0.0065 |
| Office | Flow/Person | 0.011 |
| Produce | Flow/Area | 0.00065 |
| Sales | Flow | 0.0065 |

Table 25: Ventilation Requirement for Supermarket model.

| Thermal Zone | Length [m] | Cooling Capacity [$\frac{W}{m}$] | Operating Temperature [$^\circ C$] |
|---|---|---|---|
| Bakery | 2.40 | 461.52 | 2.2 |
| Deli | 3.00 | 1442.25 | 2.2 |
| Deli | 4.80 | 461.52 | 2.2 |
| Produce | 30.00 | 1442.25 | 2.2 |
| Sales | 33.58 | 1442.25 | 2.2 |
| Sales | 49.25 | 1442.25 | 2.2 |
| Sales | 81.69 | 538.44 | -15.0 |
| Sales | 39.01 | 528.83 | -12.0 |
| Sales | 10.44 | 461.52 | 2.2 |
| Sales | 96.66 | 461.52 | 2.2 |
| Sales | 38.10 | 615.36 | -23.3 |

Table 24: Refrigeration Cases for Supermarket model.

#### 4.3.4 HVAC system.

HVAC system for Supermarket model is an all-air system, in which both ventilation and air conditioning are provided by supply air. HVAC system has six different air loops (one for each thermal zone) that handle the outdoor air flow specified in Table 25. Each loop is driven by a fan with constant velocity and air diffusers as terminals to the zone, fans characteristics are specified in Table 26. All the thermal power lost by each fan is transferred to the air flow.

| Thermal Zone | Total Efficiency | Pressure Rise [Pa] | Motor Efficiency |
|---|---|---|---|
| Bakery | 0.56875 | 622.0 | 0.875 |
| Deli | 0.56875 | 622.0 | 0.875 |
| Dry Storage | 0.56875 | 622.0 | 0.875 |
| Office | 0.53625 | 622.0 | 0.825 |
| Produce | 0.56875 | 622.0 | 0.875 |
| Sales | 0.6006 | 1017.592 | 0.924 |

Table 26: Fans characteristic for Supermarket model [19].

Heating load is supplied by **Furnaces** and cooling load is supplied by **Direct Expansion Chillers (DX Chiller)**, one for each AHU. Furnace provide heating by natural gas combustion and the combustion product are driven into a heat exchanger where heat is transferred to the air stream. In the DX chiller refrigerant fluid expansion step take place directly in the cold deck where exchange heat with the air stream, these chillers can operate with two-speed (higher and lower speed) without continuous modulation, at low speed the machine operates at around 70% of the design capacity. Values are obtained by sizing for each climate zone modelled, for Bergamo the design days considered are the same of the ones considered in Commercial Retail model [24]. Table 27 shows design values of the furnaces while Table 28 shows design values of the DX chillers for the Supermarket model located in climate zone **E**.

| | High Speed | | Low Speed | |
|---|---|---|---|---|
| Thermal Zone | Cooling Capacity [kW] | COP | Cooling Capacity [kW] | COP |
| Bakery | 40 | 3.500 | 28 | 3.552 |
| Deli | 40 | 3.300 | 28 | 3.349 |
| Dry Storage | 32 | 3.234 | 22.4 | 3.282 |
| Office | 5.6 | 3.667 | - | - |
| Produce | 28 | 3.234 | 20 | 3.282 |
| Sales | 80 | 3.134 | 56 | 3.180 |

Table 28: DX chillers design values for Supermarket model.

| Thermal Zone | Nominal Capacity [kW] | Burner Efficiency |
|---|---|---|
| Bakery | 30 | 0.80 |
| Deli | 30 | 0.78 |
| Dry Storage | 65 | 0.78 |
| Office | 10 | 0.80 |
| Produce | 90 | 0.78 |
| Sales | 300 | 0.78 |

Table 29: Furnaces design values for Supermarket model in climate zone **B**.

| Thermal Zone | Nominal Capacity [kW] | Burner Efficiency |
|---|---|---|
| Bakery | 32 | 0.80 |
| Deli | 40 | 0.78 |
| Dry Storage | 64 | 0.78 |
| Office | 12 | 0.80 |
| Produce | 96 | 0.78 |
| Sales | 400 | 0.78 |

Table 27: Furnaces design values for Supermarket model.

Winter design day for climate zone **B** has a minimum outdoor temperature of $5°C$ [29] and in summer a maximum temperature of $32.8°C$ with daily variation $\Delta T = 5.4°C$. Table 29 shows design values of the furnaces while Table 30 shows design values of the DX chillers for the Supermarket model located in climate zone **B**.

| | High Speed | | Low Speed | |
|---|---|---|---|---|
| Thermal Zone | Cooling Capacity [kW] | COP | Cooling Capacity [kW] | COP |
| Bakery | 45 | 3.500 | 31.5 | 3.552 |
| Deli | 90 | 3.300 | 63 | 3.349 |
| Dry Storage | 65 | 3.234 | 45 | 3.282 |
| Office | 9 | 3.667 | - | - |
| Produce | 63 | 3.234 | 45 | 3.282 |
| Sales | 225 | 3.134 | 157.5 | 3.180 |

Table 30: DX chillers design values for Supermarket model in climate zone **B**.

#### 4.3.5 Control System.

The HVAC System is controlled by a PI control (see Equation 2) that has as controlled variable the supply air temperature to reach the set-point obtained by [14] in order to reach thermal comfort requirement for different building application.

Thermostat provide a **Dual Set-point plus Dead-band**; two temperatures values are set and no action occurs while internal air temperature is in between these two values. Heating is available from 15 October to 15 April for 14 hours per day maximum while the rest of the year the HVAC system is in cooling mode [26] in climate zone **E** and from 1 December to 31 March for 8 hours per day in climate zone **B**.

For the thermal zones considered the values are represented in Table 31 and are valid both for climate zone **B** and **E**.

| | | Heating Mode | Cooling Mode |
|---|---|---|---|
| Bakery | Lower SP | 16 | 25 |
| | Upper SP | 19 | 27 |
| Deli | Lower SP | 16 | 25 |
| | Upper SP | 19 | 27 |
| Dry Storage | Lower SP | 14 | - |
| | Upper SP | 16 | - |
| Office | Lower SP | 20 | 25 |
| | Upper SP | 22 | 27 |
| Produce | Lower SP | 16 | 25 |
| | Upper SP | 19 | 27 |
| Sales | Lower SP | 16 | 23 |
| | Upper SP | 19 | 25 |

Table 31: Design Setpoint for Supermarket model.

The error signal is computed between the set-point and the actual internal air temperature.

The control system both for heating and cooling have been tuned with Ultimate Oscillation method. The results are listed in Table 32 for climate zone **E** and Table 33 for climate zone **B**.

| | | Heating Mode | Cooling Mode |
|---|---|---|---|
| Bakery | $K_p$ | 1.802 | 2.8 |
| | $K_i$ | 0.217 | 0.3 |
| Deli | $K_p$ | 1.802 | 2.8 |
| | $K_i$ | 0.217 | 0.3 |
| Dry Storage | $K_p$ | 1.802 | 1.802 |
| | $K_i$ | 0.217 | 0.217 |
| Office | $K_p$ | 1.802 | 1.802 |
| | $K_i$ | 0.217 | 0.217 |
| Produce | $K_p$ | 1.802 | 1.6 |
| | $K_i$ | 0.217 | 0.3 |
| Sales | $K_p$ | 1.802 | 1.6 |
| | $K_i$ | 0.217 | 0.3 |

Table 32: Supermarket model HVAC Tuning, climate zone **E**.

| | | Heating Mode | Cooling Mode |
|---|---|---|---|
| Bakery | $K_p$ | 1.802 | 2.8 |
| | $K_i$ | 0.217 | 0.3 |
| Deli | $K_p$ | 1.802 | 1.5 |
| | $K_i$ | 0.217 | 0.3 |
| Dry Storage | $K_p$ | 1.802 | 1.802 |
| | $K_i$ | 0.217 | 0.217 |
| Office | $K_p$ | 1.802 | 1.802 |
| | $K_i$ | 0.217 | 0.217 |
| Produce | $K_p$ | 1.5 | 1.6 |
| | $K_i$ | 0.3 | 0.3 |
| Sales | $K_p$ | 1.802 | 2.8 |
| | $K_i$ | 0.217 | 0.3 |

Table 33: Supermarket model HVAC Tuning, climate zone **B**.

# 5 Results and Comparisons.

Results of simulation were compared with constant set-points imposed by design conditions. To quantify energy savings were compared monthly energy consumption for heating and cooling by an histogram and then the summation over the two years in terms of kWh during heating and cooling season. To verify the ability of the Reinforcement Learning algorithm to maintain internal air temperature between the design bounds, has been plotted the mean hourly air temperature in heating and cooling season.

Constant set-point has been compared with three different Reinforcement Learning algorithm that have some difference in the calculation of the reward term. The three algorithms use three different slack parameters as described in Chapter 3.5 that is the ratio between the coefficient of the energy consumption term and the term related to air temperature in the reward to change the importance given by the algorithm to energy savings in contrast with the thermal needs of occupants.

Three different algorithm were used with slack equal to 0.1, 1 and 10 where energy savings respectively increase but the percentage of time in which the air temperature remains in the required bounds should decrease. The performances of these three controls were compared to a constant set-point with HVAC shut down one hour before the closing time as previously real common practice in the Commercial Retail located in climate zone **E**.

## 5.1 Commercial Retail.

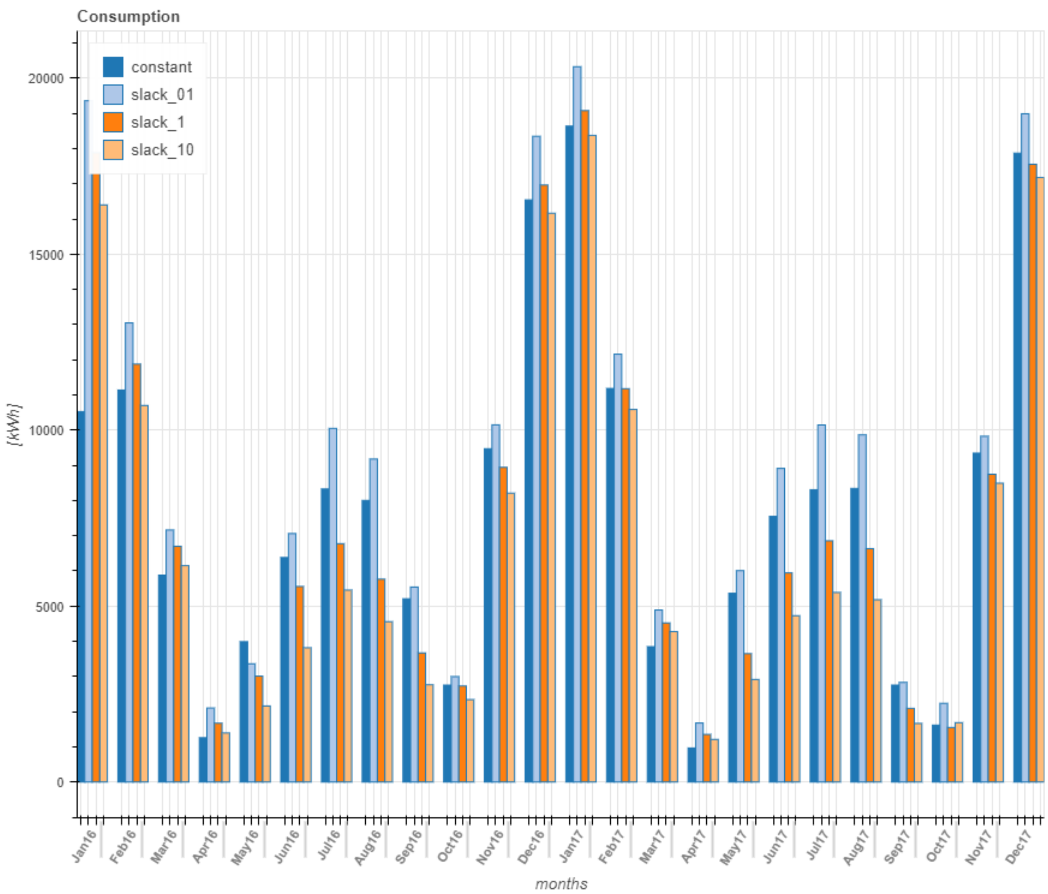



Figure 32: Primary Energy Consumption for HVAC system, Commercial Retail.

Figure 32 shows monthly primary energy consumption for the four simulation conducted with EnergyPlus. As expected the algorithm with highest slack is able to achieve more energy savings especially during middle season and summer season where an effective control strategy can much more improve thermal comfort and reduce energy consumption.

Figure 33 shows the absolute difference between each control algorithm tested in EnergyPlus environment and traditional control in kWh of primary energy for every month of simulation starting from January 2016 to December 2017.

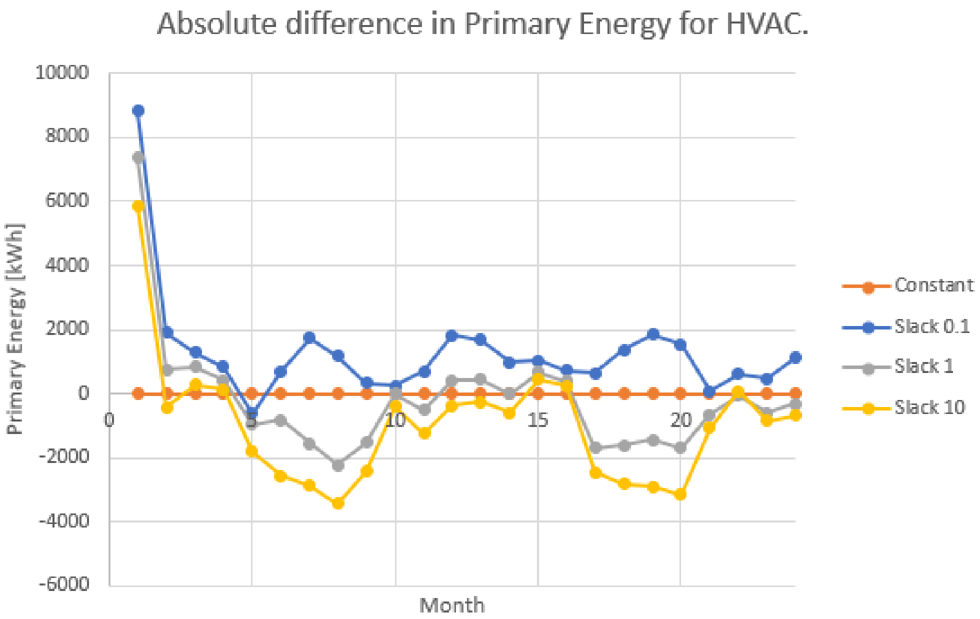



Figure 33: Absolute difference in Primary Energy Consumption for HVAC system, Commercial Retail.

Figures 34 and 35 show for each simulation the hourly mean internal air temperature respectively during heating and cooling season. The horizontal line inside the coloured bars represents the mean value of the internal air temperature, the coloured bar itself contains the first variance of values while the black line represent maximum and minimum temperature obtained in the two years long simulation. The grater variance of Reinforcement Learning control algorithm is justified by the *exploration* phase of the agent conducted every month.

The previous strategy of turning off HVAC system one hour before the closure of the Commercial Retail although allows reduction of energy consumption do not satisfy thermal requirement, especially in cooling season where mean internal air temperature of the last hour of opening gets out the requested temperature bound.

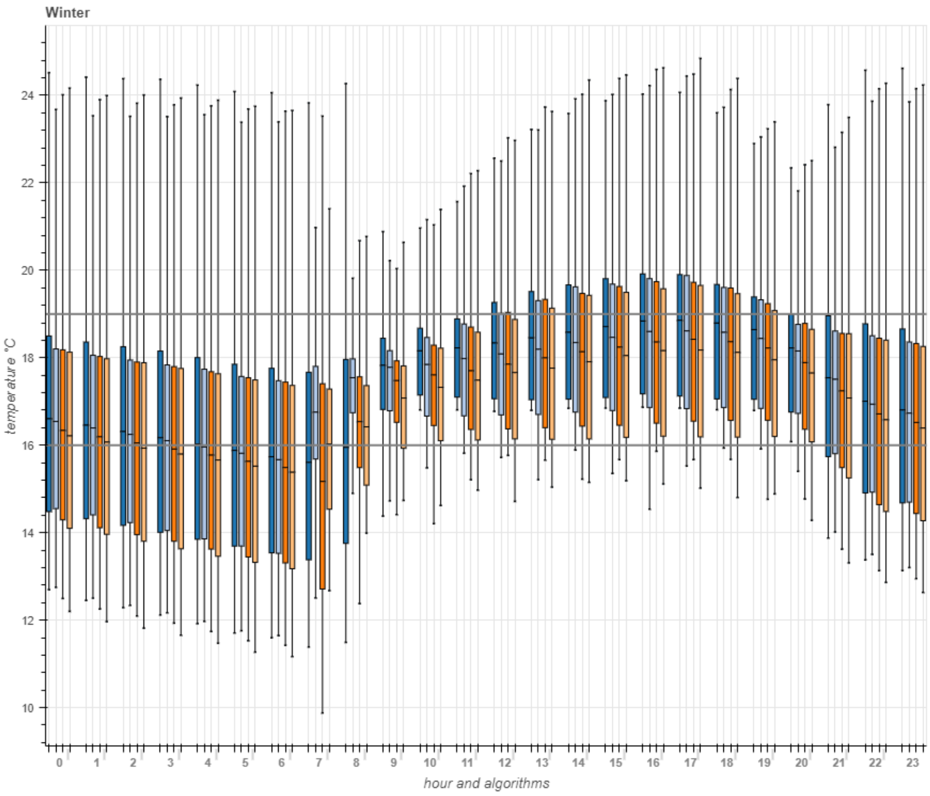



Figure 34: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Commercial Retail.

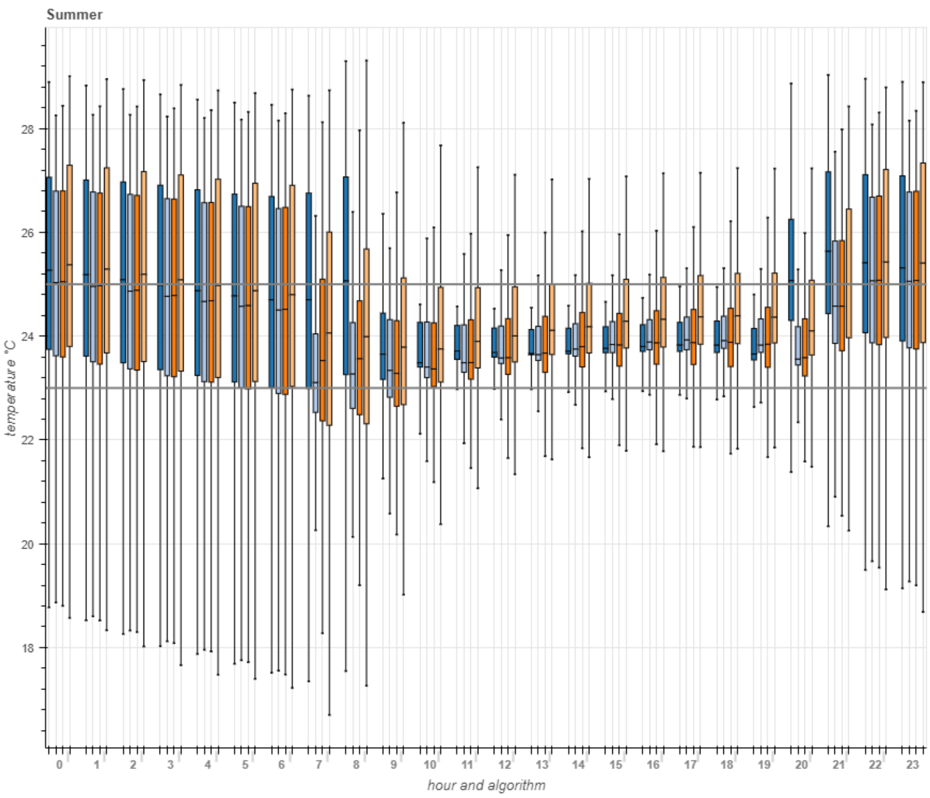



Figure 35: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Commercial Retail.

Same information is given in terms of hourly distribution for each algorithm used by following figures:

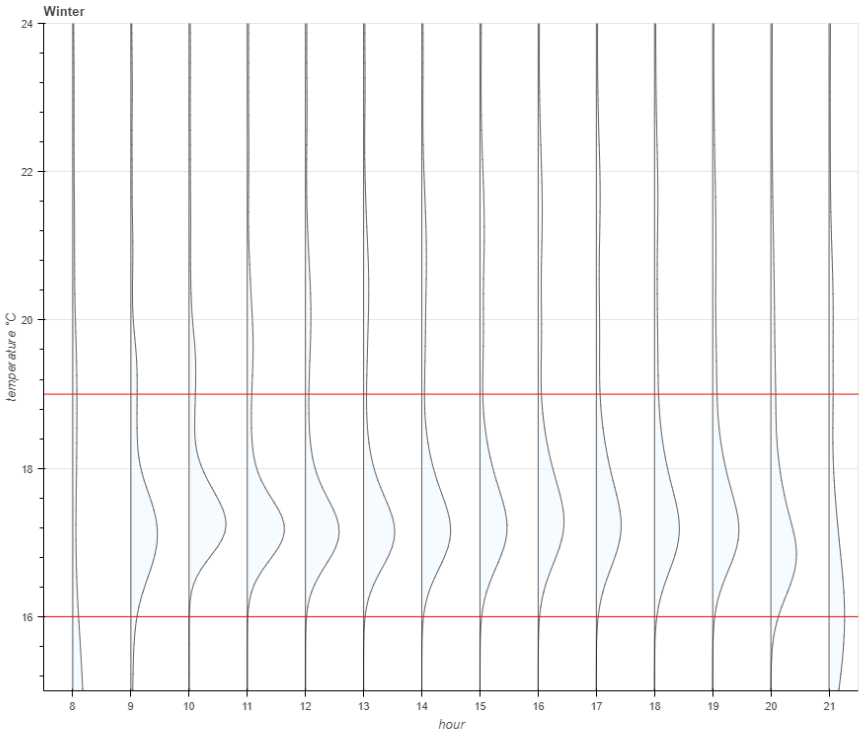



Figure 36: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail.

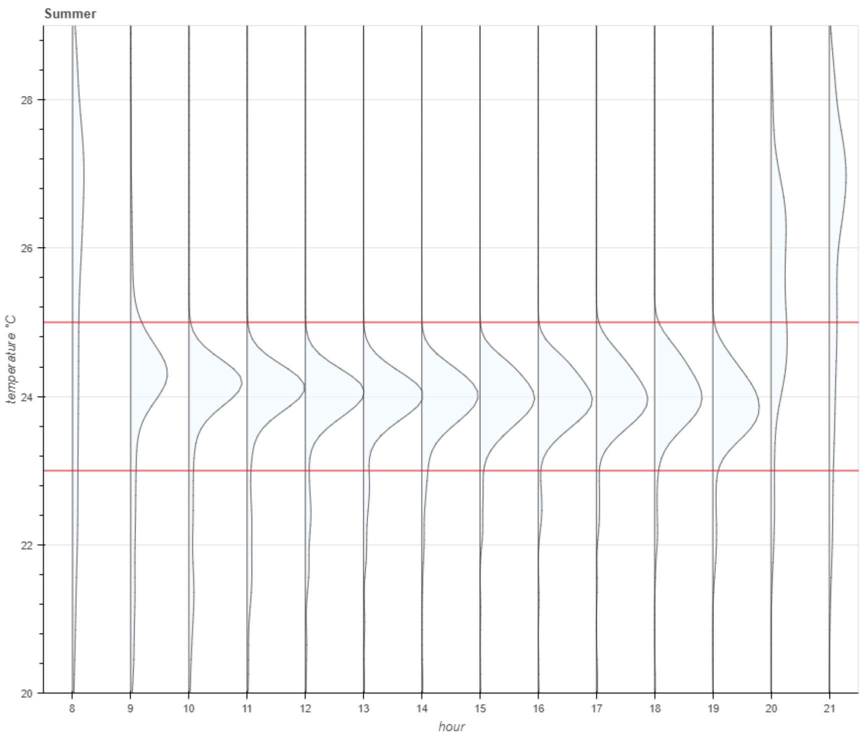



Figure 37: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail.

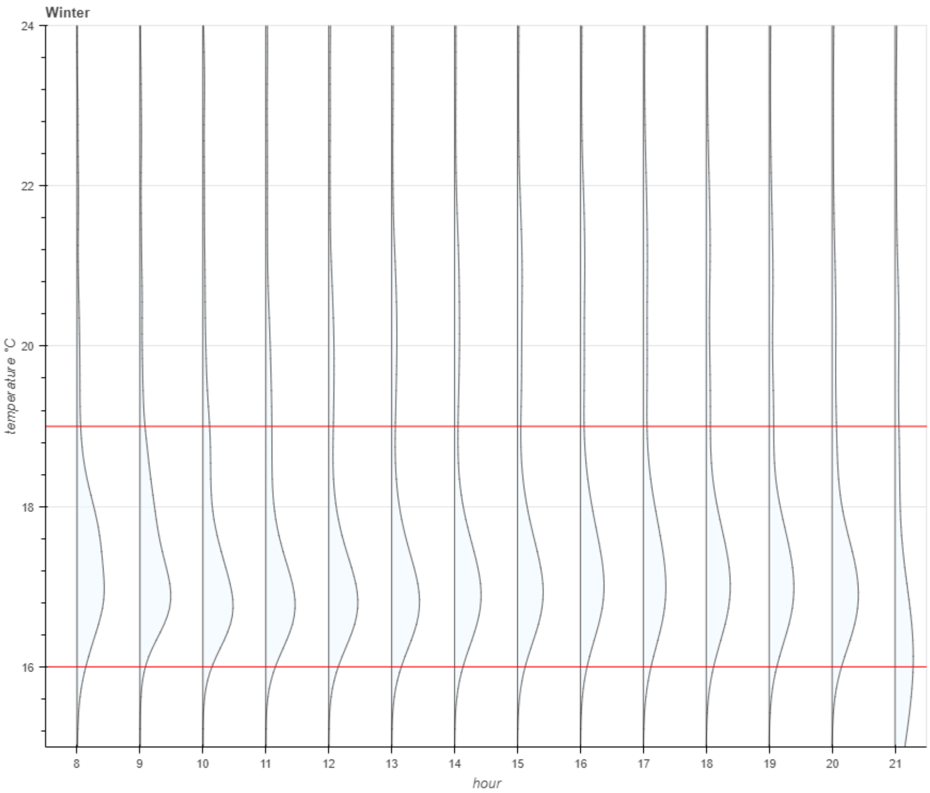



Figure 38: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail.

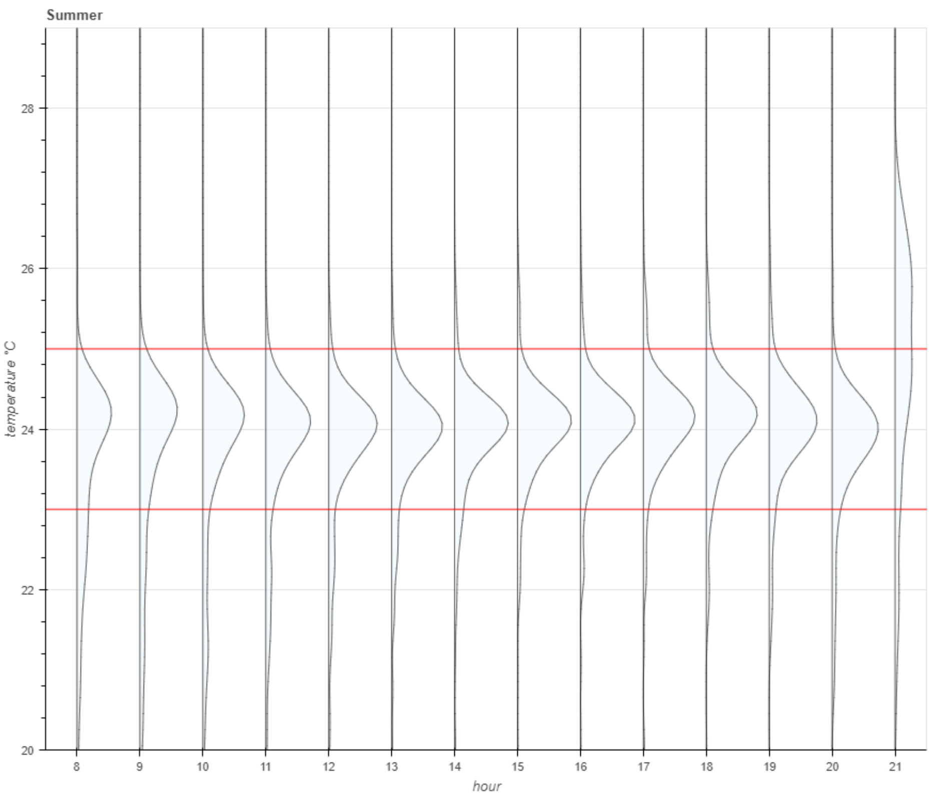



Figure 39: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail.

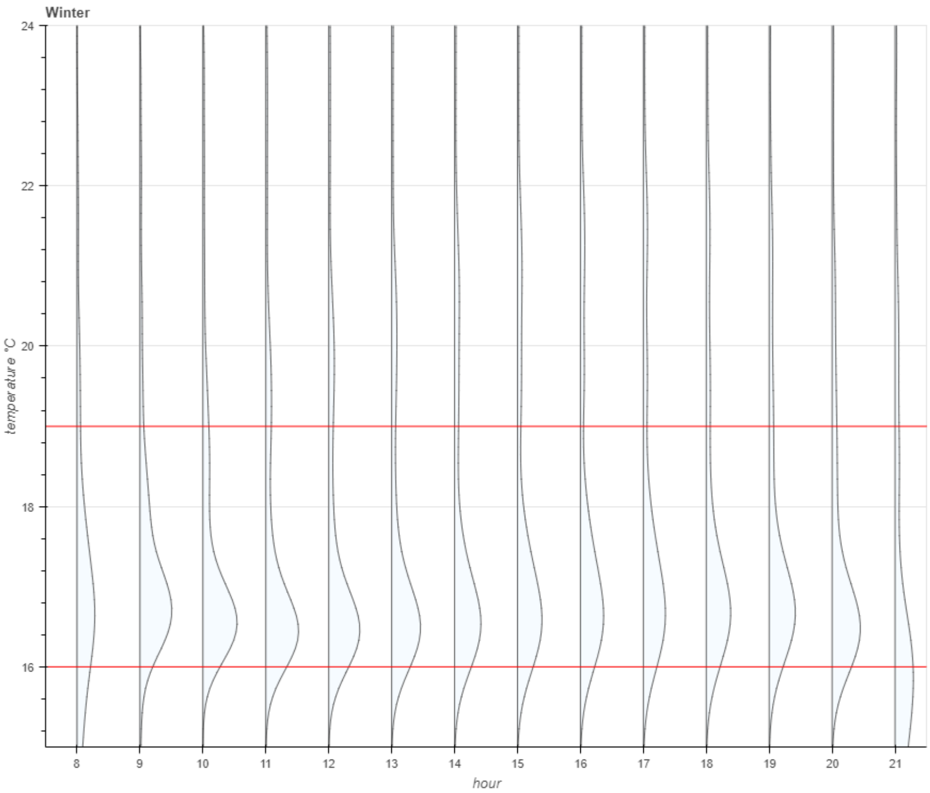



Figure 40: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail.

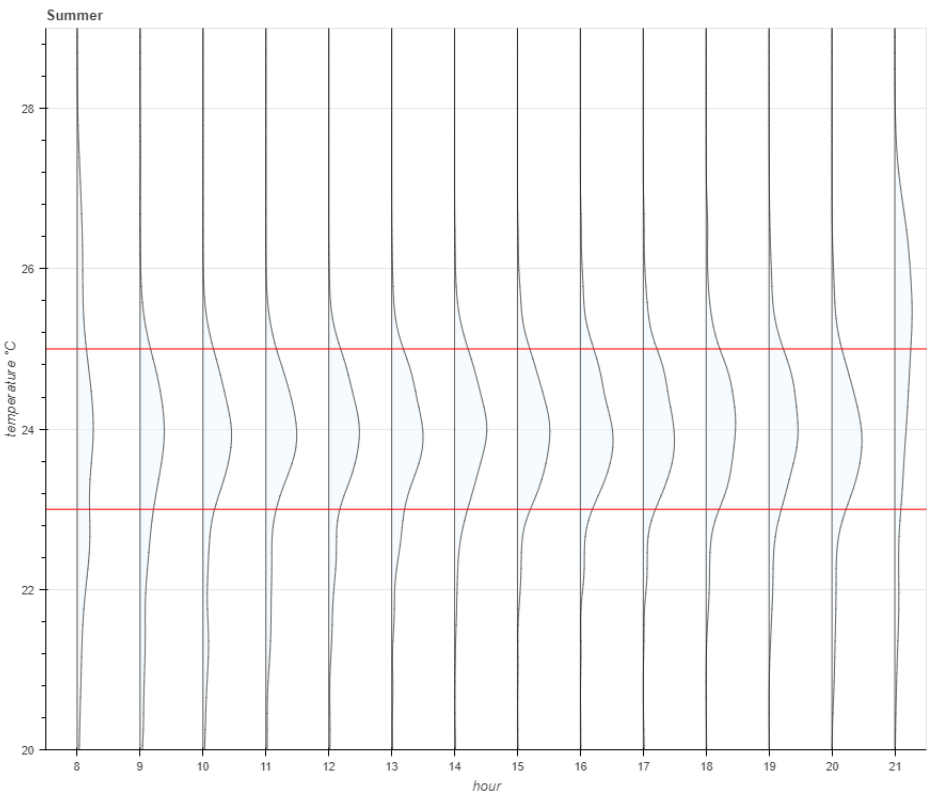



Figure 41: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail.

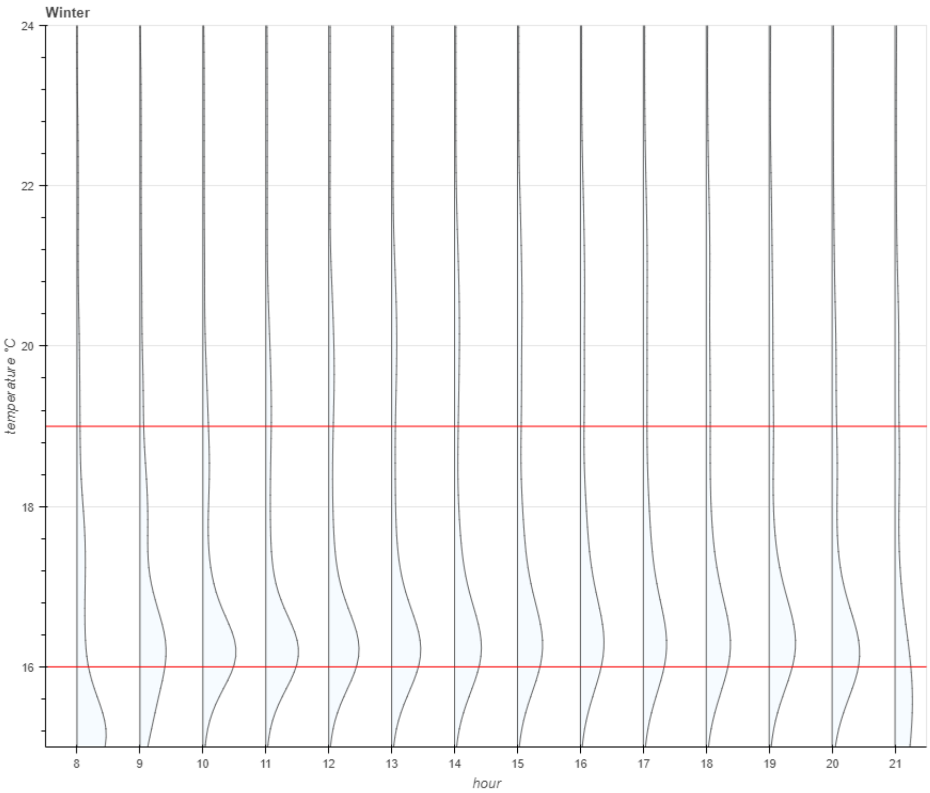



Figure 42: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail.

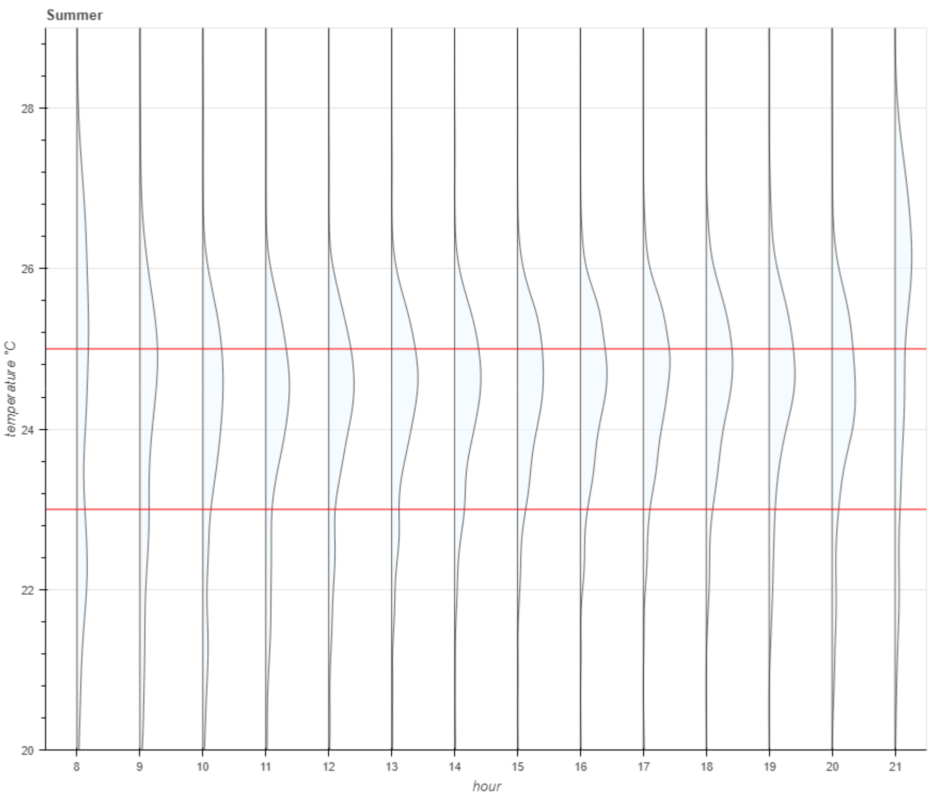



Figure 43: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail.

Percentage of exceeding of the air temperature bound given by design manual is resumed by Table .

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant | 3.30% | 6.95% | 28.62% | 25.87% |
| slack_01 | 1.78% | 6.73% | 28.29% | 25.09% |
| slack_1 | 4.40% | 13.42% | 27.18% | 24.35% |
| slack_10 | 13.01% | 29.20% | 27.13% | 23.52% |

Table 34: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Commercial Retail.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant | 19.34% | 23.70% | 51.25% | 38.57% |
| slack_01 | 21.89% | 27.77% | 2.39% | 2.06% |
| slack_1 | 23.20% | 32.21% | 5.14% | 3.11% |
| slack_10 | 22.96% | 30.01% | 19.24% | 10.08% |

Table 35: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Commercial Retail.

| | saving | |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant | 0.00% | 0.00% |
| slack_01 | -12.84% | -19.04% |
| slack_1 | 22.06% | -10.44% |
| slack_10 | 39.28% | -5.25% |

Table 36: Relative savings with respect of constant set-point, Commercial Retail.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant | 185256.032 | 123.752 |
| Slack 0.1 | 216212.418 | 144.430 |
| Slack 1 | 180678.862 | 120.694 |
| Slack 10 | 161788.830 | 108.075 |

Table 37: Total primary energy consumption for HVAC, Commercial Retail.

Figure 36 and Table 37 shows relative savings in kWh of primary energy with respect constant imposing of set-point. As the slack parameter increases the energy savings also increases but with an increasing in exceeding the thermal bounds of internal air temperature. Energy consumption reduction is achieved during cooling season whit almost 40% of savings.

After the simulations conducted with the model described before, to assess the robustness of the algorithm under different operational configuration some modification have been conducted keeping constant all the other parameters. In particular modifications regards:

- Infiltration;
- Recirculation fraction;
- Climate zone;
- Occupation profile;
- Building envelope;
- Outdoor temperature control.

Infiltration load has been modified supposing that HVAC system is not able to guarantee overpressure inside the building and infiltration occurs also during opening hours according to the equation exposed in Chapter 4.2.3. Results are expressed using the same graphics used before:

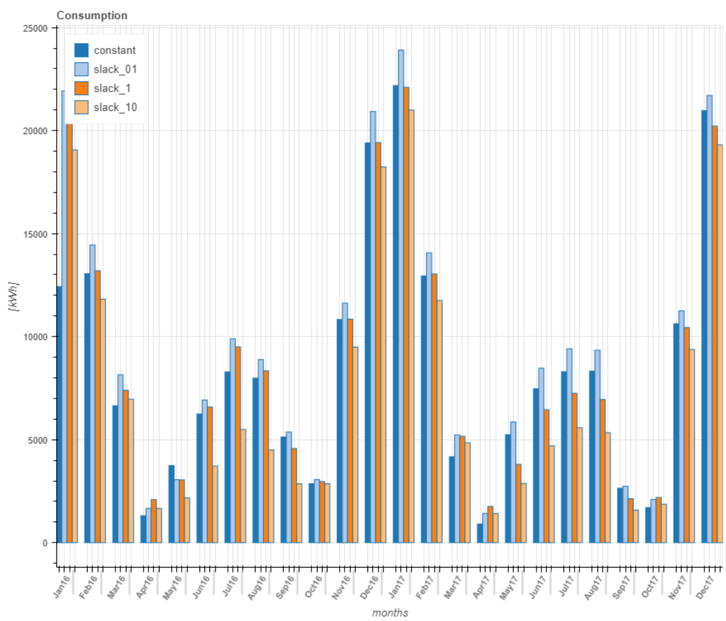



Figure 44: Primary Energy Consumption for HVAC system, Commercial Retail with infiltration during opening hours.

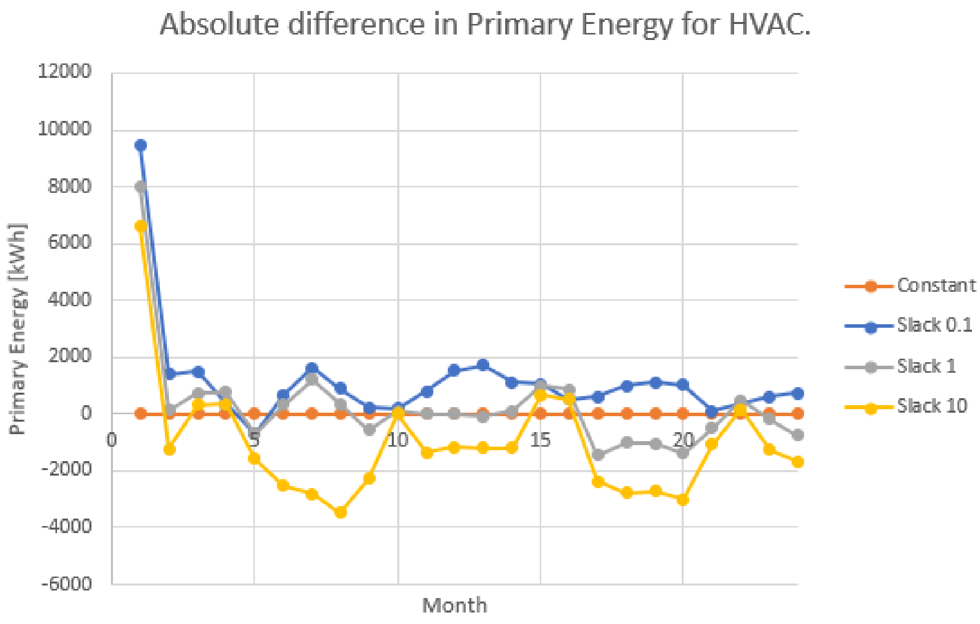



Figure 45: Absolute difference in Primary Energy Consumption for HVAC system, Commercial Retail with infiltration during opening hours.

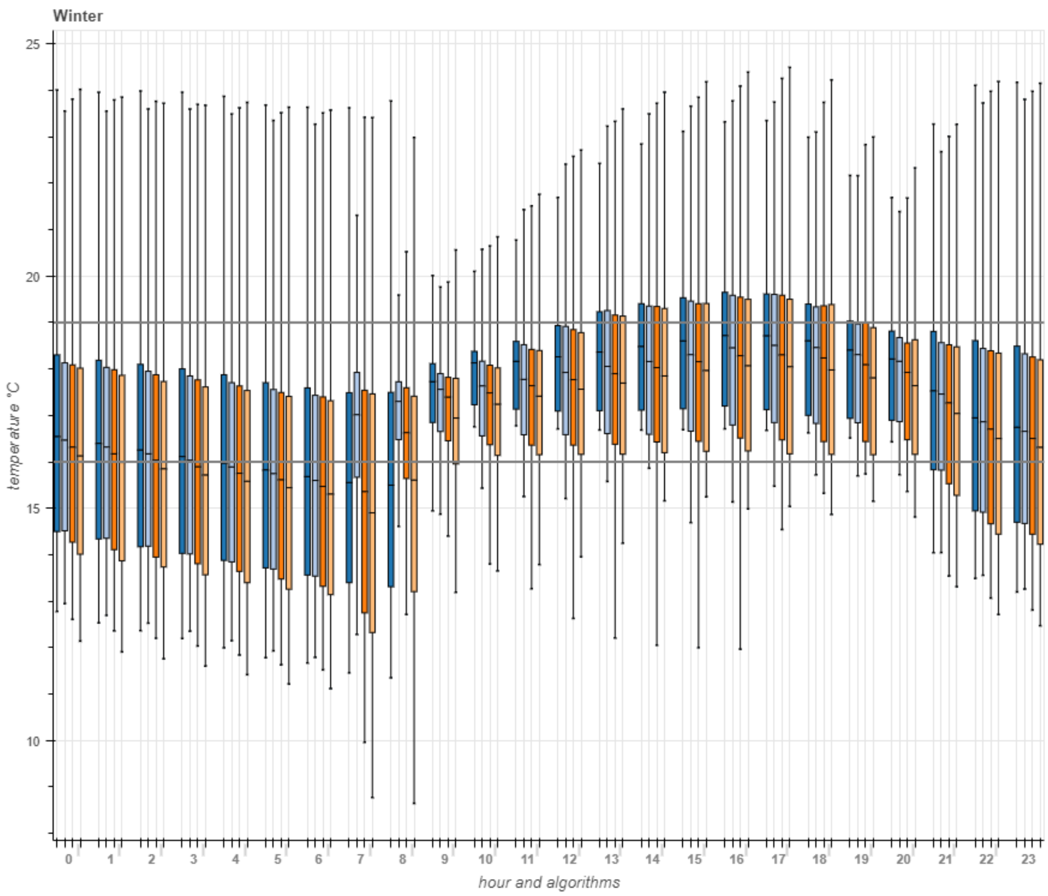



Figure 46: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Commercial Retail with infiltration during opening hours.

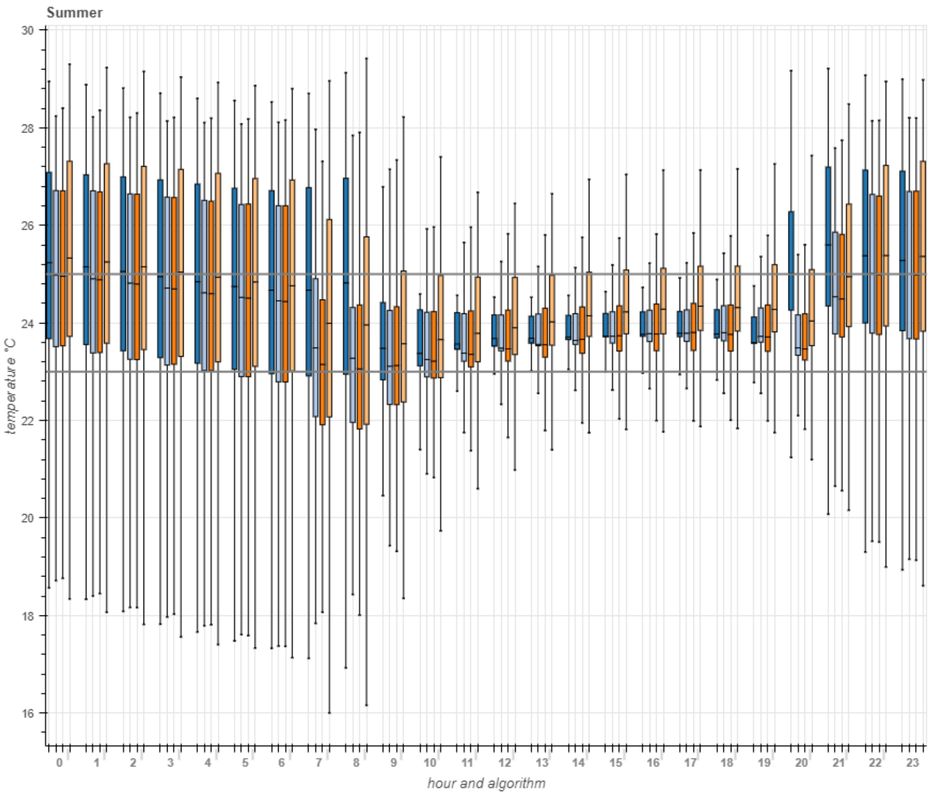



Figure 47: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Commercial Retail with infiltration during opening hours.

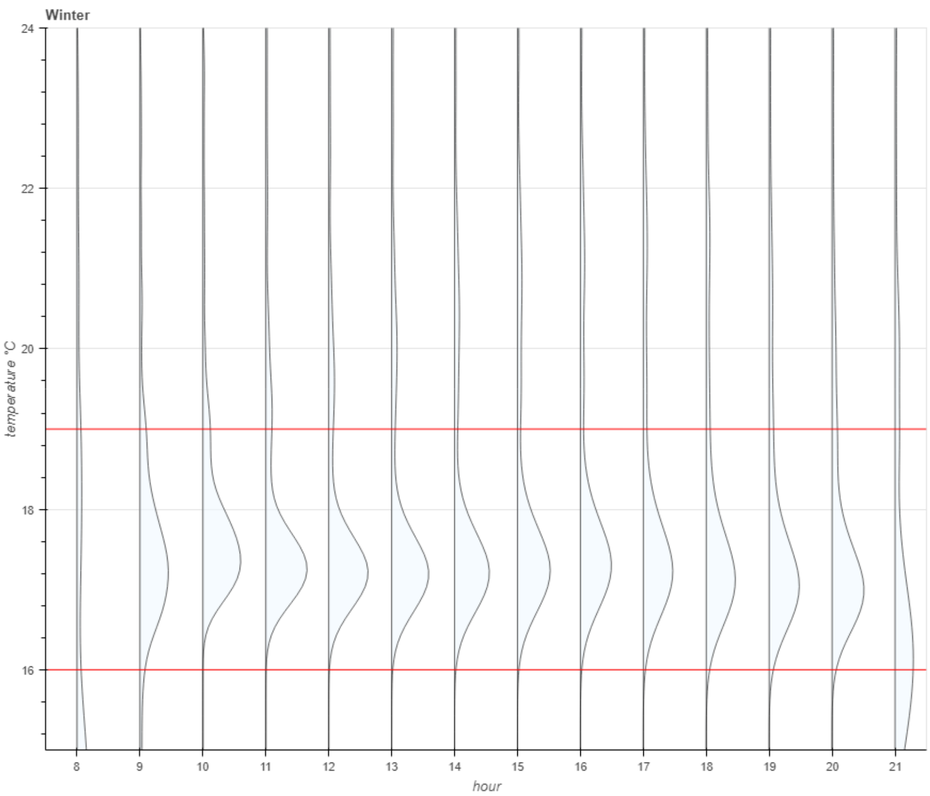



Figure 48: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with infiltration during opening hours.

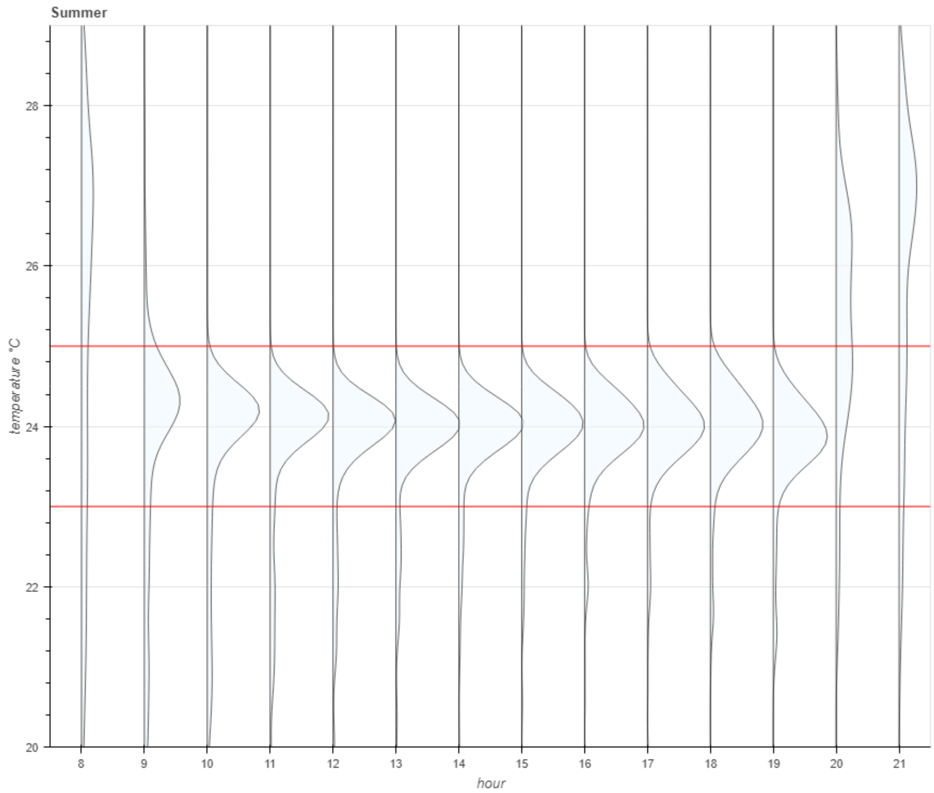



Figure 49: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with infiltration during opening hours.

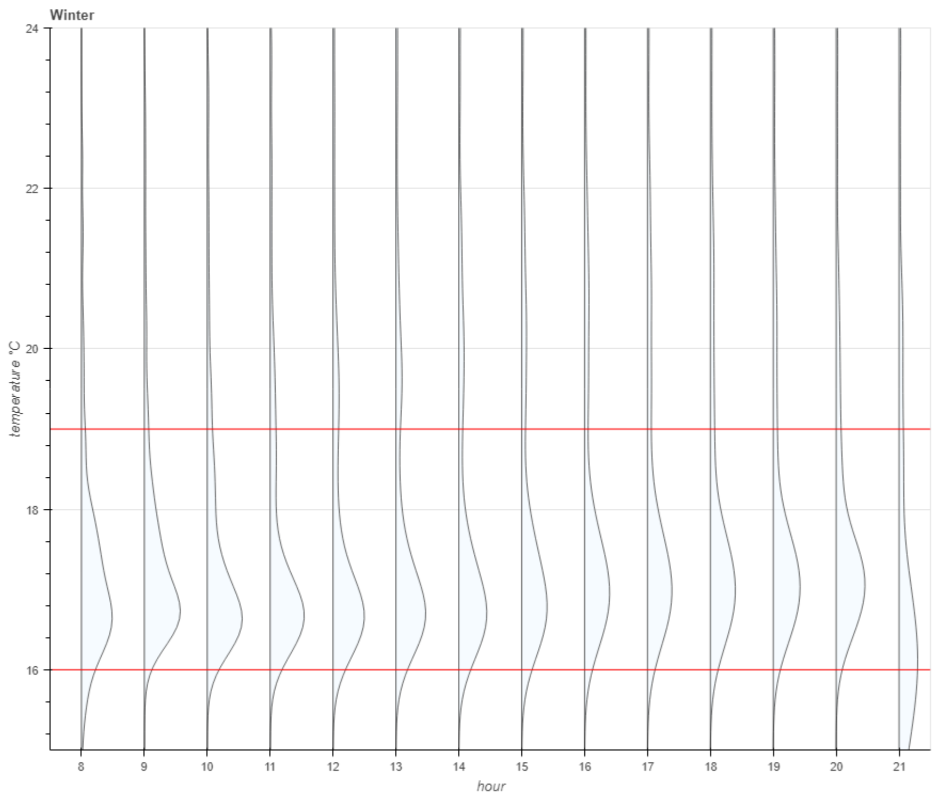



Figure 50: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with infiltration during opening hours.

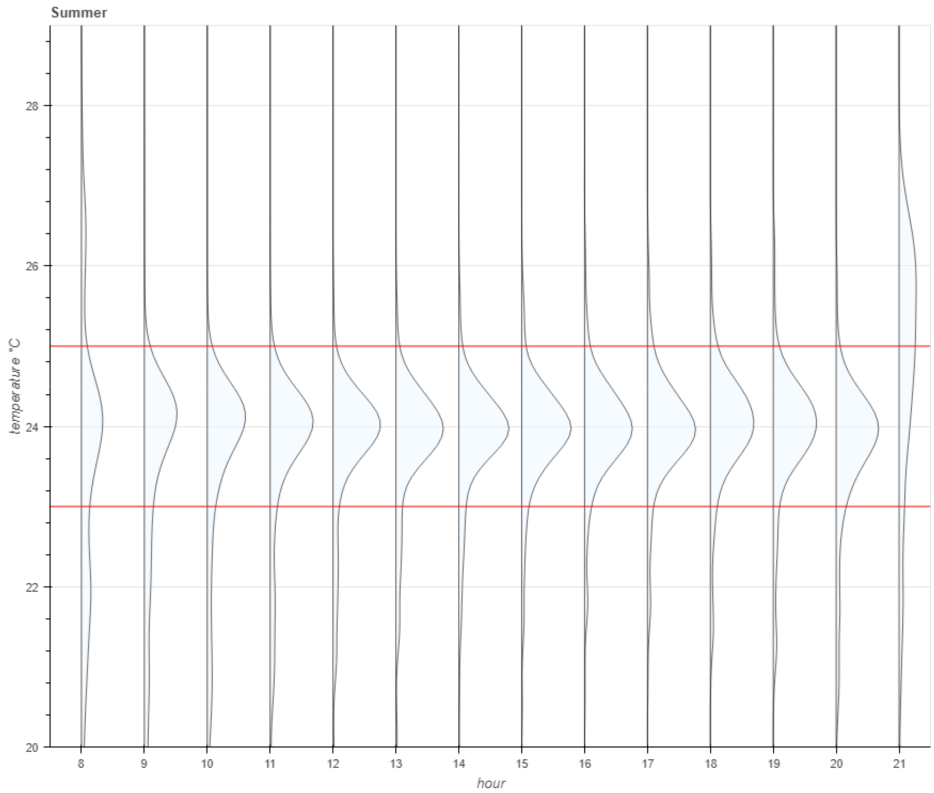



Figure 51: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with infiltration during opening hours.

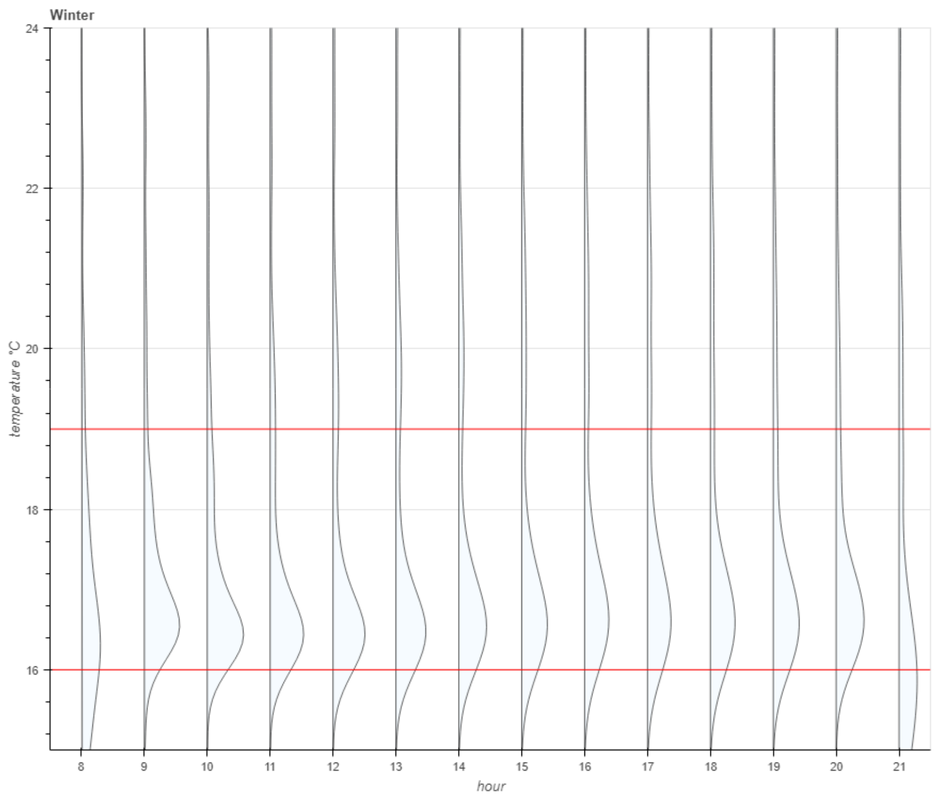



Figure 52: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with infiltration during opening hours.

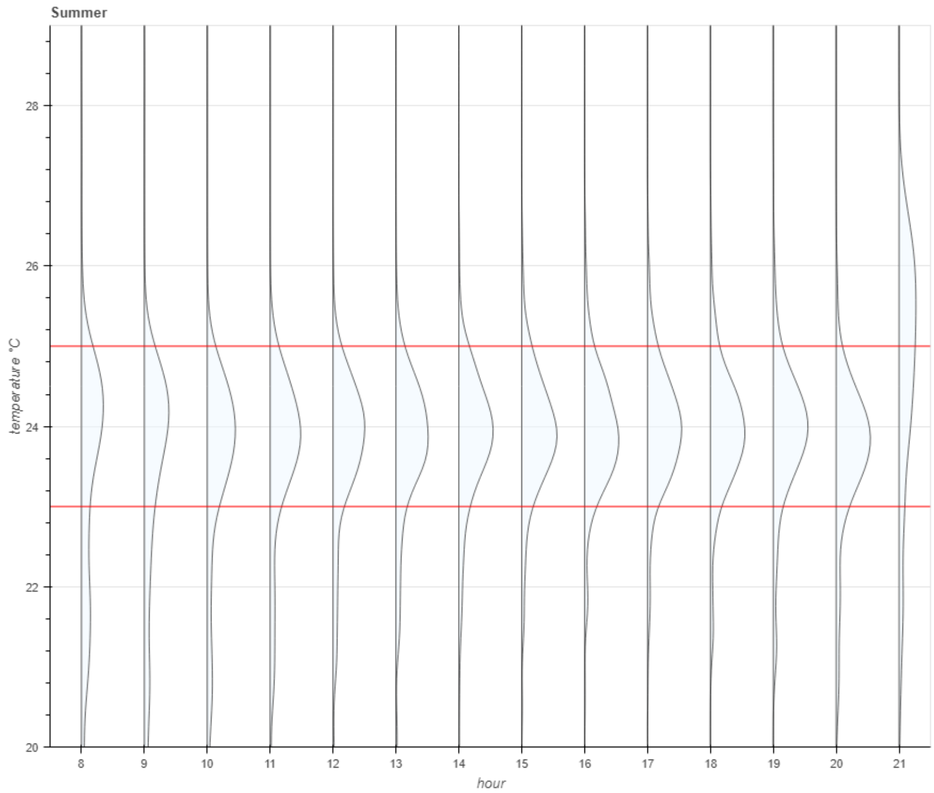



Figure 53: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with infiltration during opening hours.

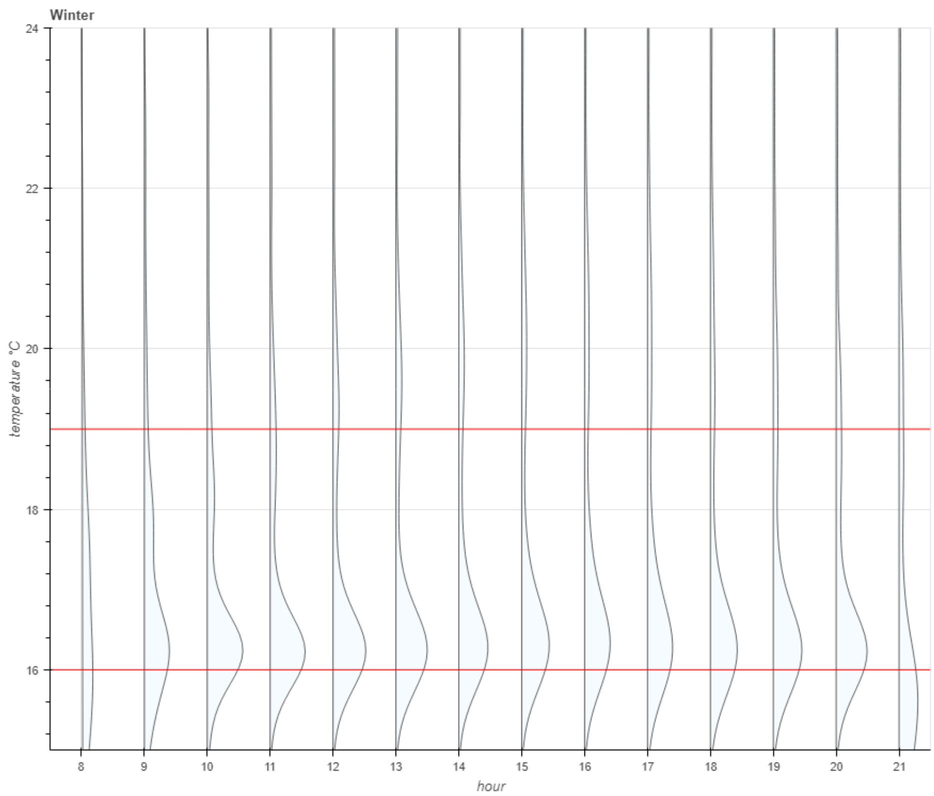



Figure 54: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with infiltration during opening hours.

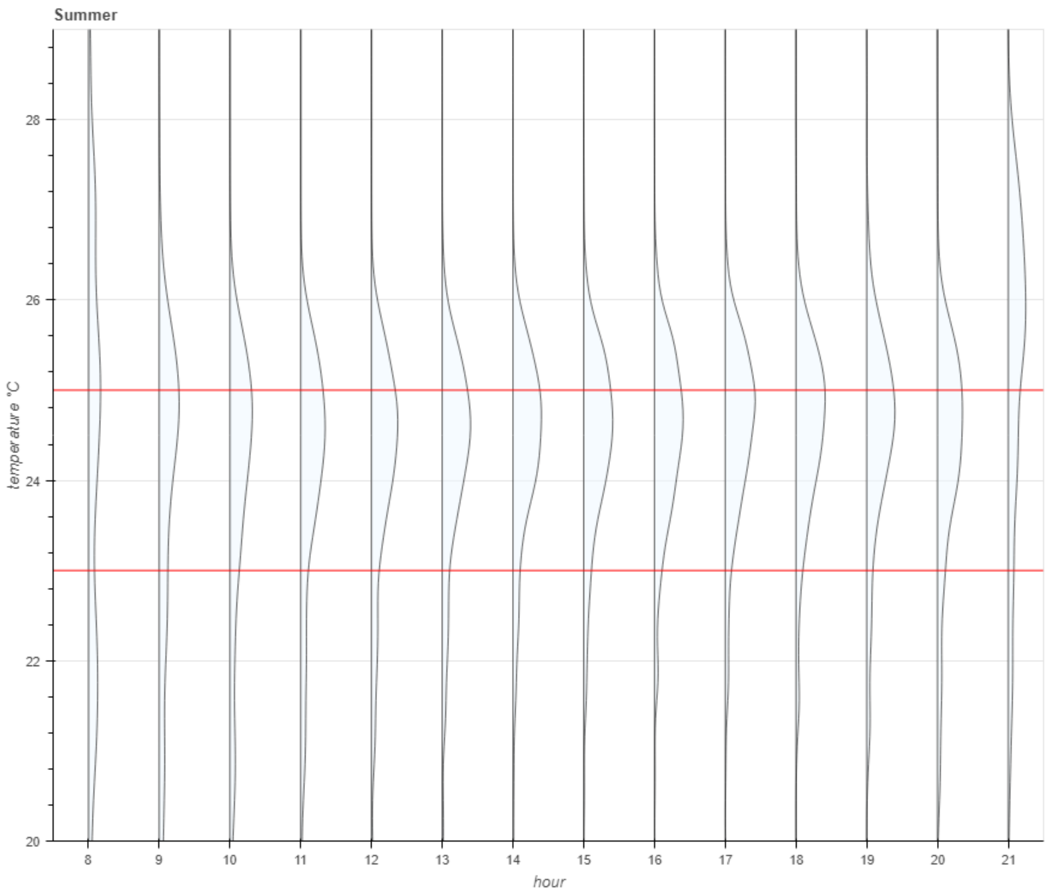



Figure 55: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with infiltration during opening hours.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant | 4.10% | 7.37% | 27.58% | 24.81% |
| slack_01 | 1.71% | 6.71% | 27.11% | 23.46% |
| slack_1 | 4.96% | 14.97% | 26.29% | 22.95% |
| slack_10 | 19.92% | 32.04% | 25.18% | 21.74% |

Table 38: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Commercial Retail with infiltration during opening hours.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant | 22.67% | 27.20% | 50.88% | 39.42% |
| slack_01 | 27.06% | 33.40% | 3.17% | 1.36% |
| slack_1 | 27.44% | 35.50% | 3.28% | 1.86% |
| slack_10 | 26.94% | 32.95% | 18.30% | 7.69% |

Table 39: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Commercial Retail with infiltration during opening hours.

| | saving | |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant | 0.00% | 0.00% |
| slack_01 | -9.71% | -16.31% |
| slack_1 | 7.56% | -10.04% |
| slack_10 | 38.14% | -3.01% |

Figure 56: Relative savings with respect of constant set-point, Commercial Retail with infiltration during opening hours.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant | 203454.700 | 135.908 |
| Slack 0.1 | 231449.262 | 154.609 |
| Slack 1 | 209934.670 | 140.237 |
| Slack 10 | 178491.616 | 119.233 |

Table 40: Total primary energy consumption for HVAC, Commercial Retail with infiltration during opening hours.

Recirculation fraction has been modified from 0.3 to 0.2, results are following reported.

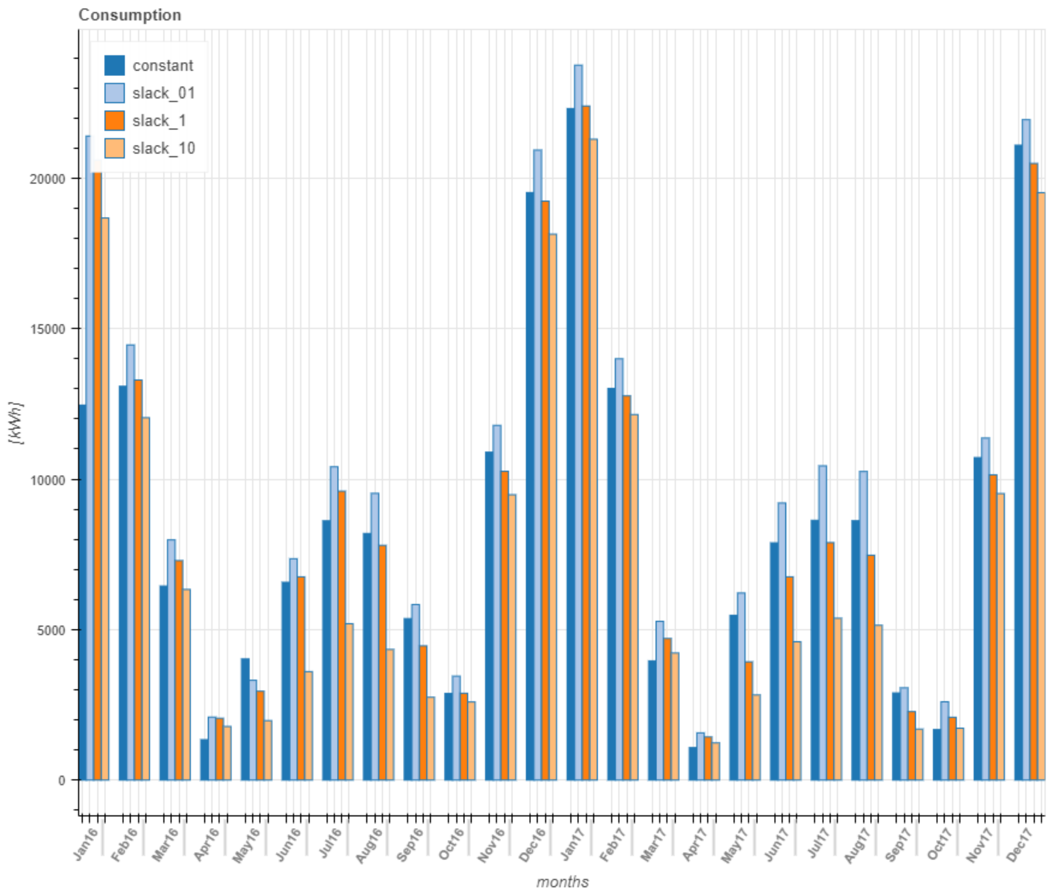



Figure 57: Primary Energy Consumption for HVAC system, Commercial Retail with recirculation fraction equal to 0.2.

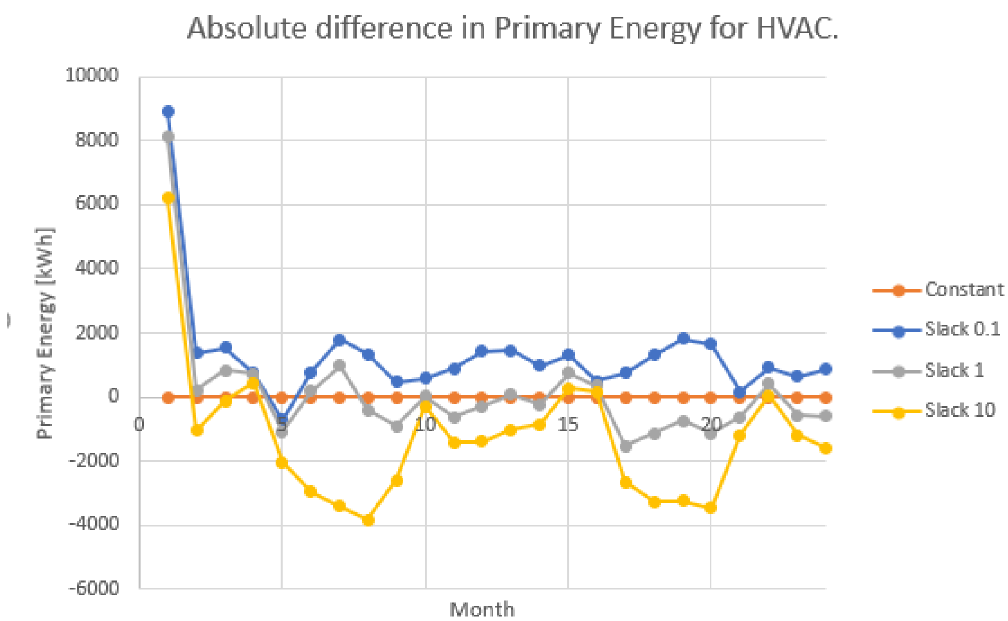



Figure 58: Absolute difference in Primary Energy Consumption for HVAC system, Commercial Retail with recirculation fraction equal to 0.2.

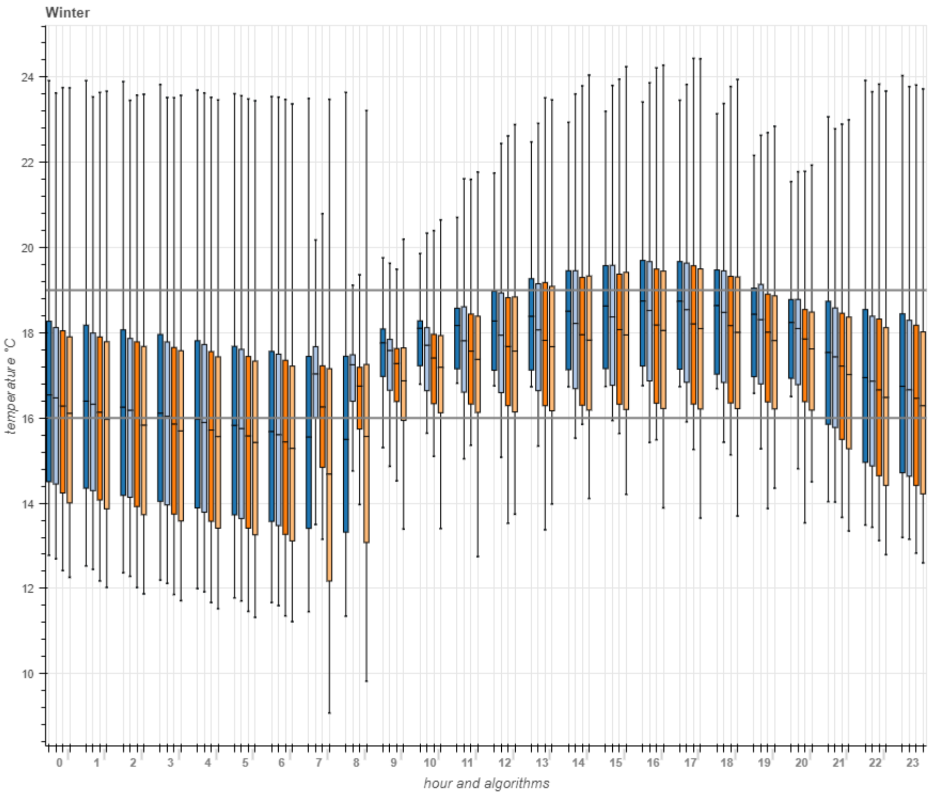



Figure 59: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Commercial Retail with recirculation fraction equal to 0.2.

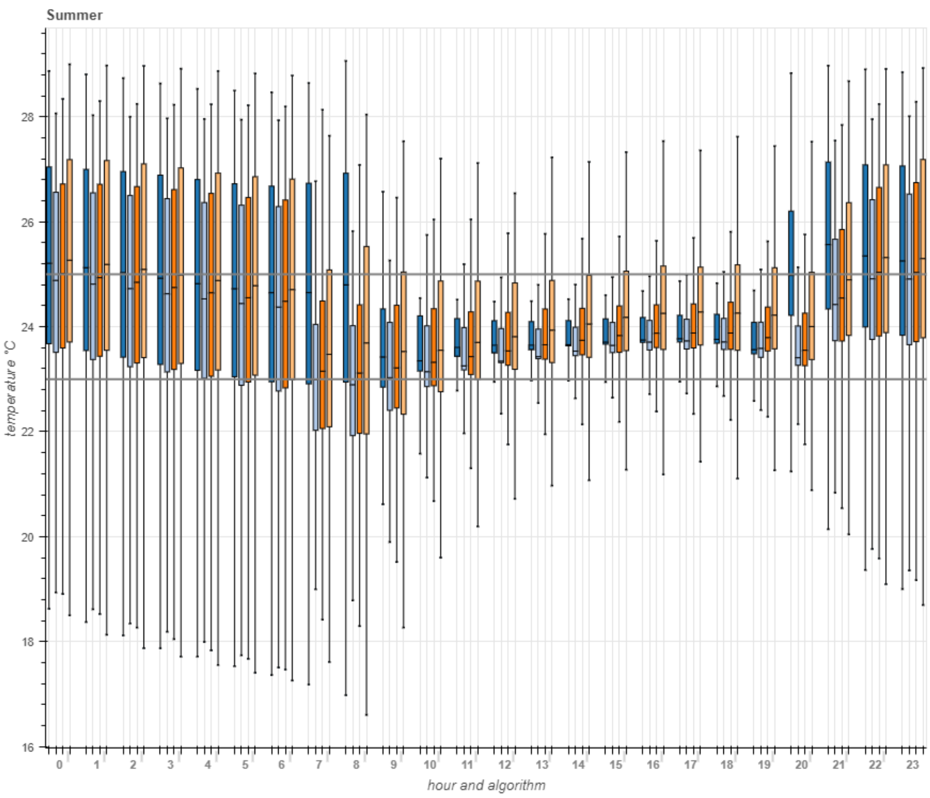



Figure 60: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Commercial Retail with recirculation fraction equal to 0.2.

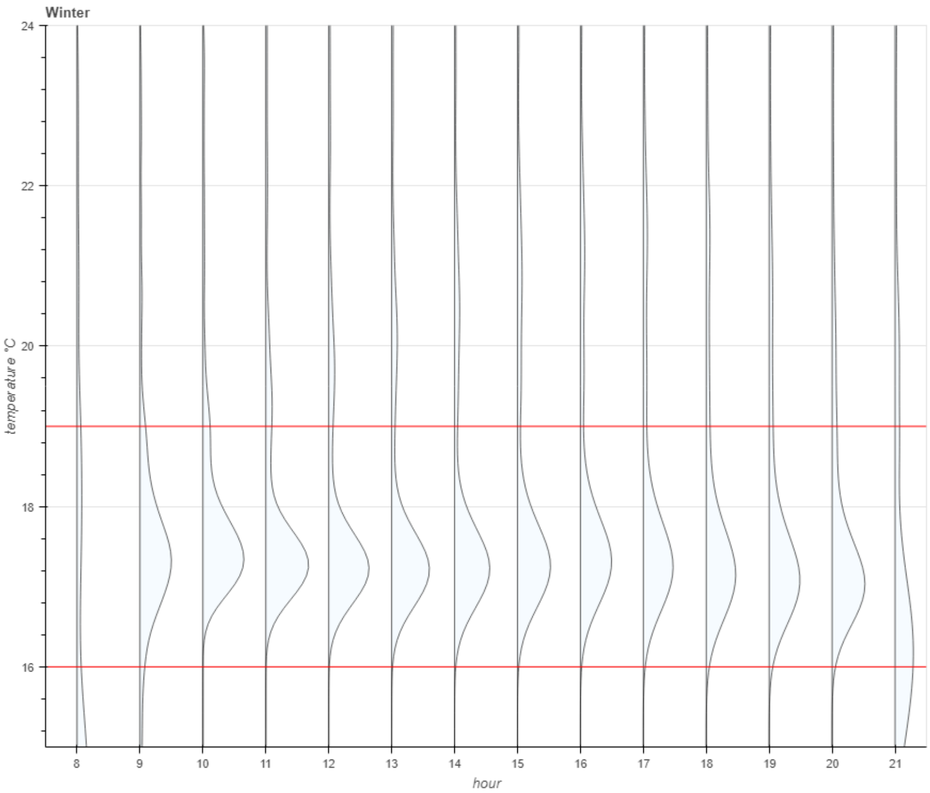



Figure 61: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with recirculation fraction equal to 0.2.

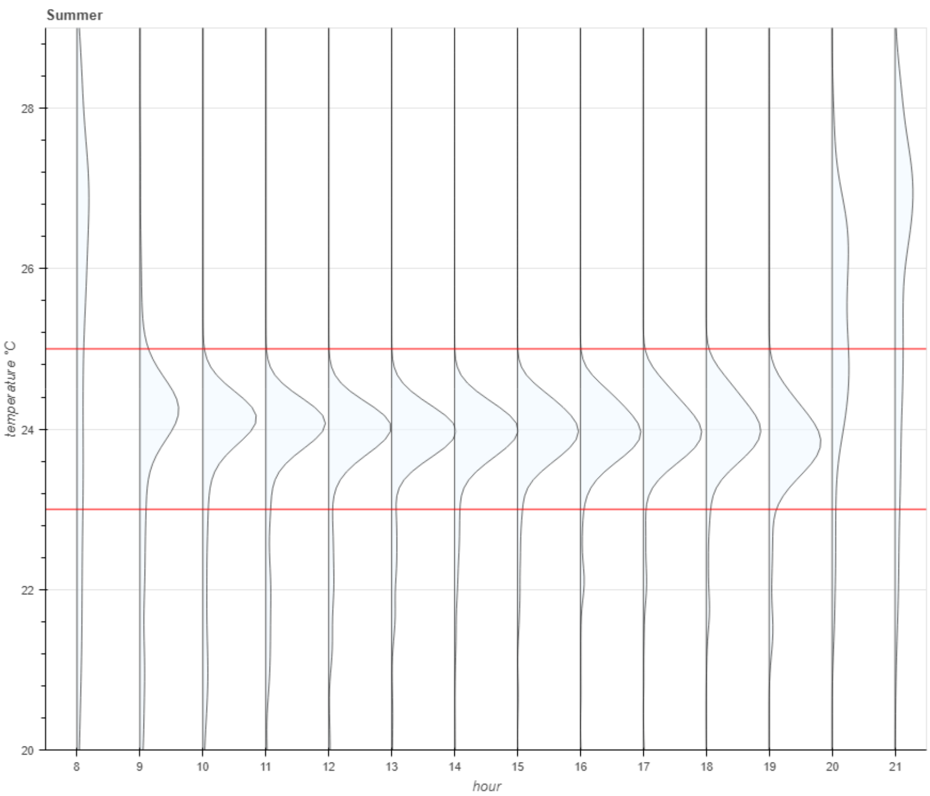



Figure 62: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with recirculation fraction equal to 0.2.

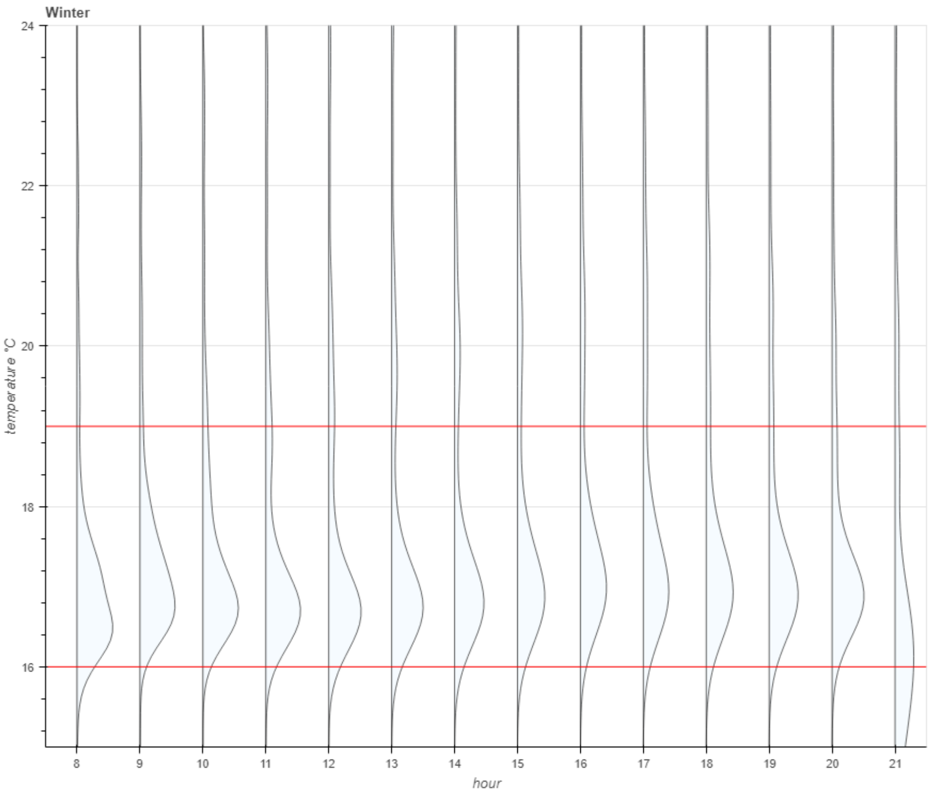



Figure 63: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with recirculation fraction equal to 0.2.

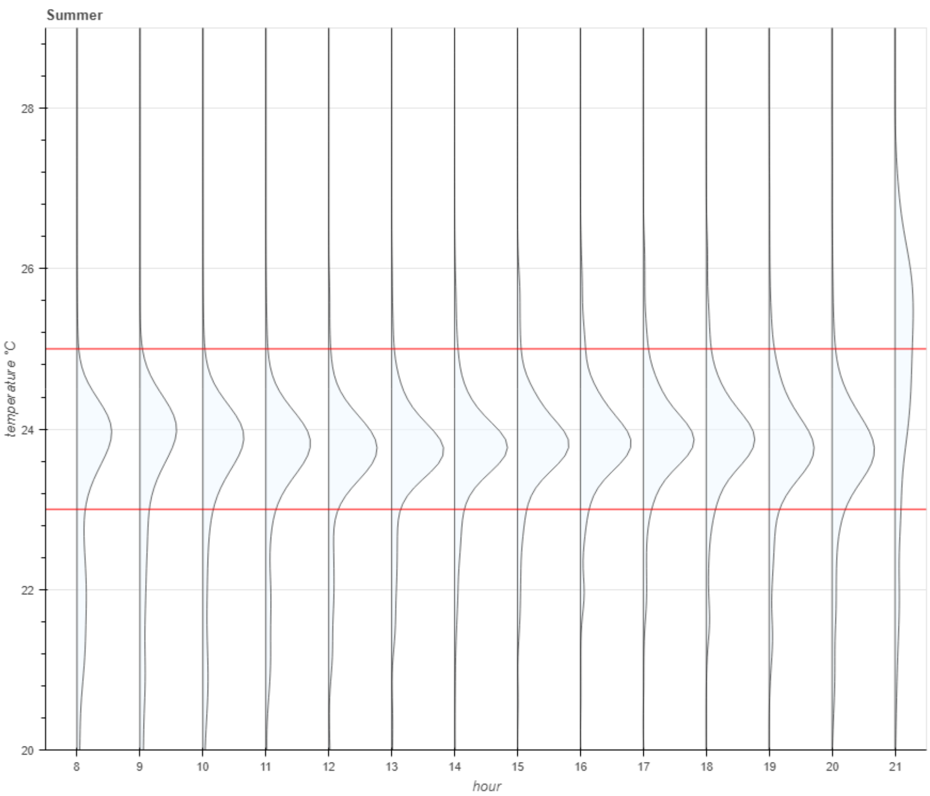



Figure 64: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with recirculation fraction equal to 0.2.

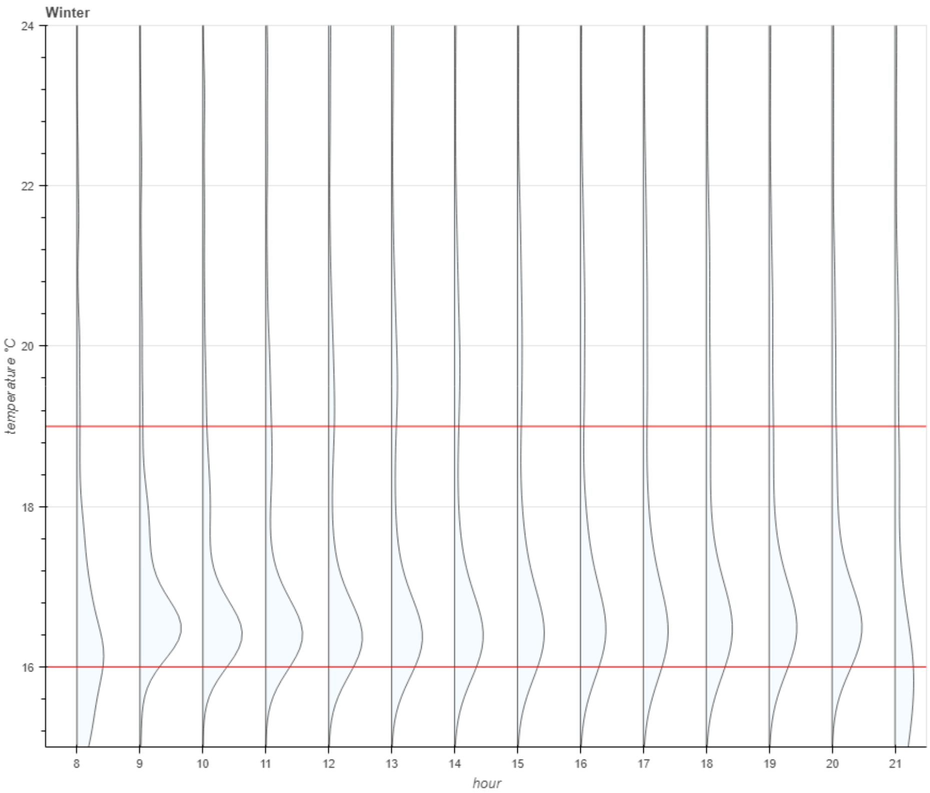



Figure 65: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with recirculation fraction equal to 0.2.

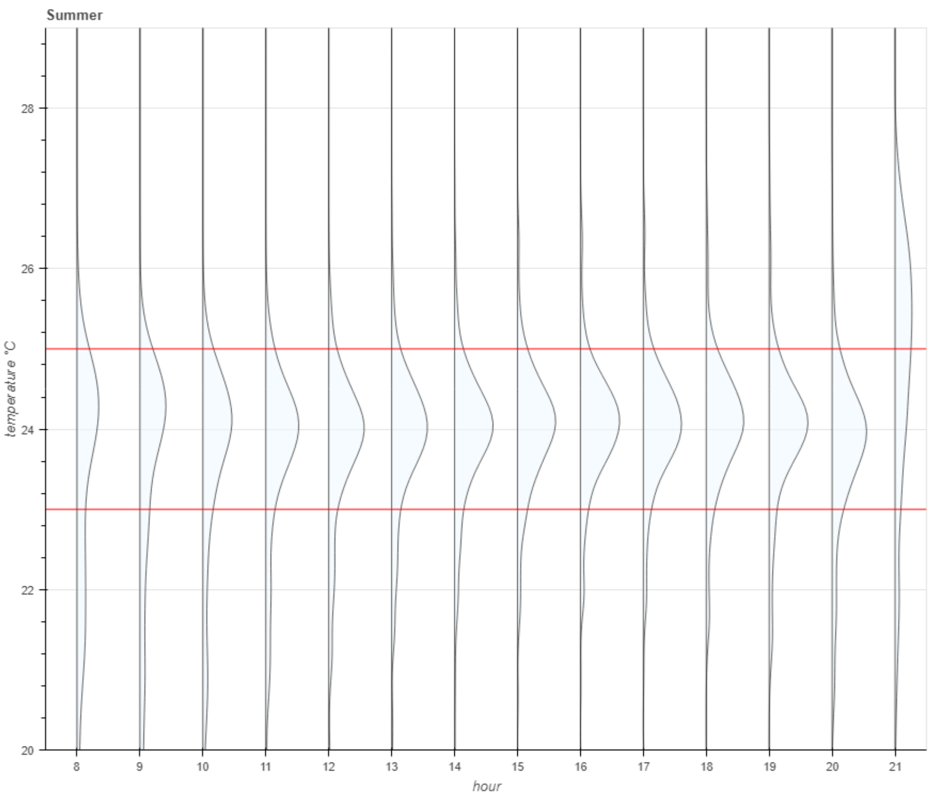



Figure 66: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with recirculation fraction equal to 0.2.

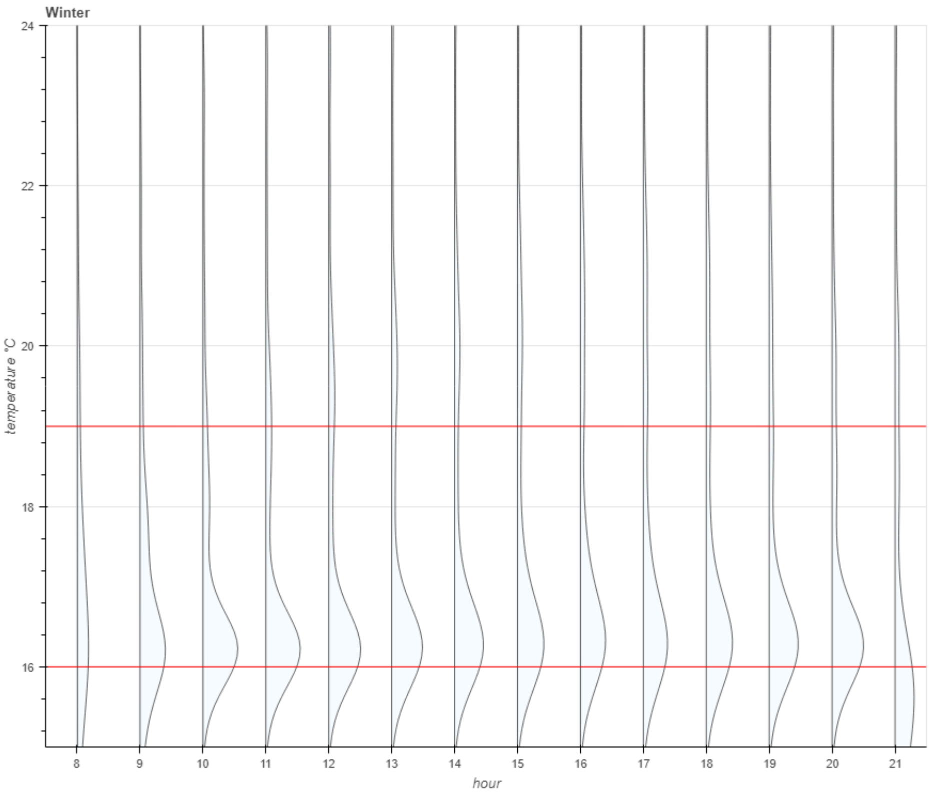



Figure 67: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with recirculation fraction equal to 0.2.

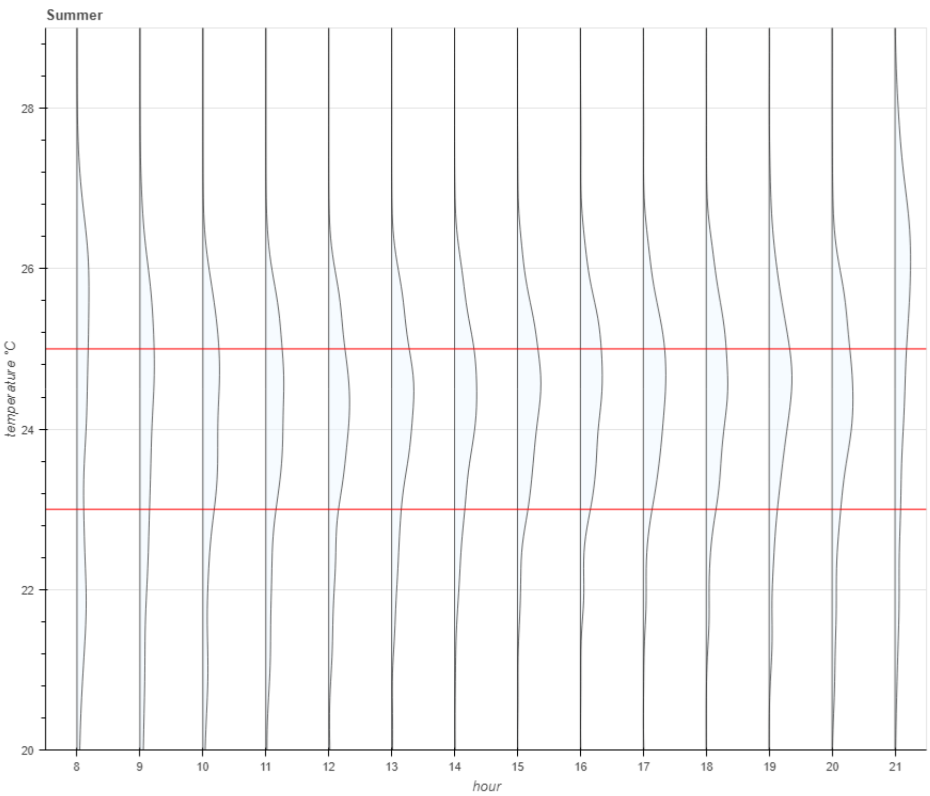



Figure 68: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with recirculation fraction equal to 0.2.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant | 3.63% | 6.69% | 27.52% | 24.85% |
| slack_01 | 2.19% | 7.54% | 27.06% | 23.78% |
| slack_1 | 6.69% | 17.07% | 26.47% | 22.89% |
| slack_10 | 17.41% | 30.30% | 25.33% | 22.03% |

Table 41: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Commercial Retail with recirculation fraction equal to 0.2.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant | 22.38% | 26.90% | 49.59% | 37.96% |
| slack_01 | 26.66% | 33.14% | 2.45% | 1.28% |
| slack_1 | 26.28% | 33.27% | 7.07% | 3.11% |
| slack_10 | 27.79% | 35.00% | 18.50% | 9.11% |

Table 42: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Commercial Retail with recirculation fraction equal to 0.2.

| | | saving |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant | 0.00% | 0.00% |
| slack_01 | -13.36% | -16.80% |
| slack_1 | 9.73% | -9.15% |
| slack_10 | 42.51% | -2.71% |

Figure 69: Relative savings with respect of constant set-point, Commercial Retail with recirculation fraction equal to 0.2.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant | 206744.443 | 138.106 |
| Slack 0.1 | 238210.633 | 159.125 |
| Slack 1 | 209505.785 | 139.950 |
| Slack 10 | 176199.760 | 117.702 |

Table 43: Total primary energy consumption for HVAC, Commercial Retail with recirculation fraction equal to 0.2.

Climate zone has been modified from **E** to **D**, for this purpose has been used the data set of Castellari weather station located in Savona (Liguria) provided by `https://www.arpal.gov.it`. Heating season goes from 1$^{st}$ November to 15$^{th}$ April [26].

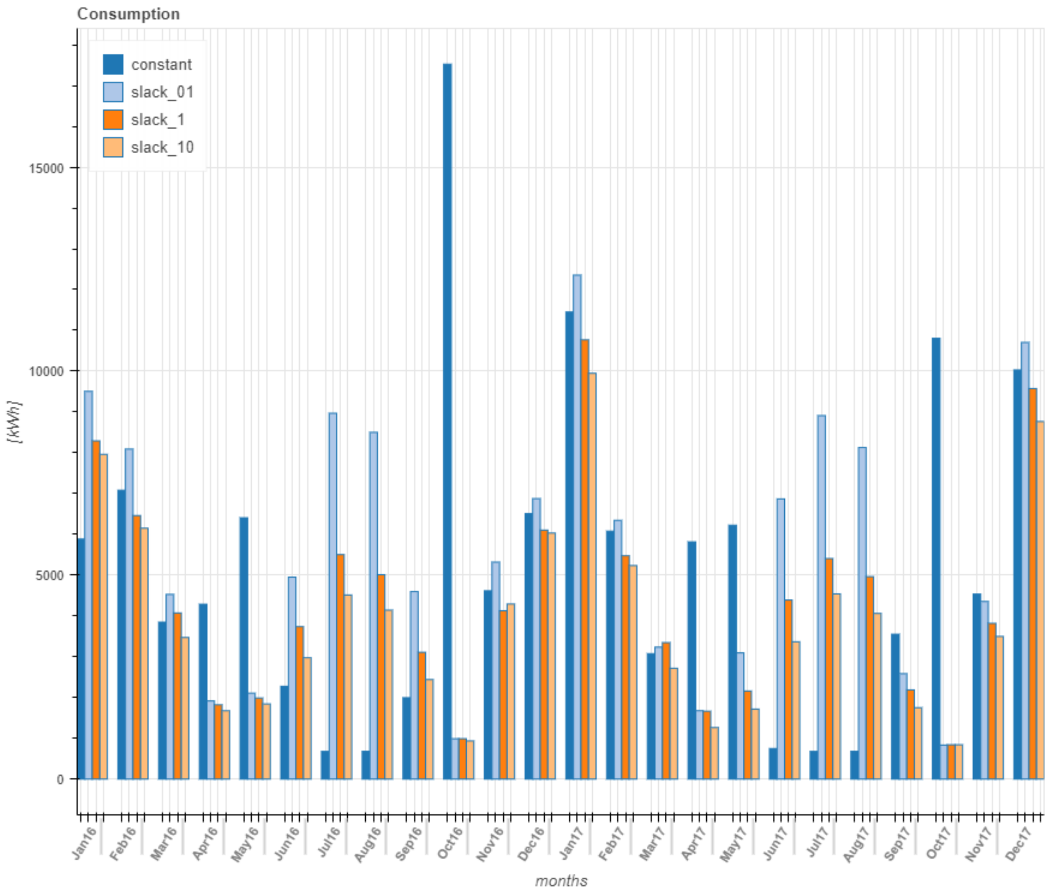



Figure 70: Primary Energy Consumption for HVAC system, Commercial Retail in climate zone **D**.

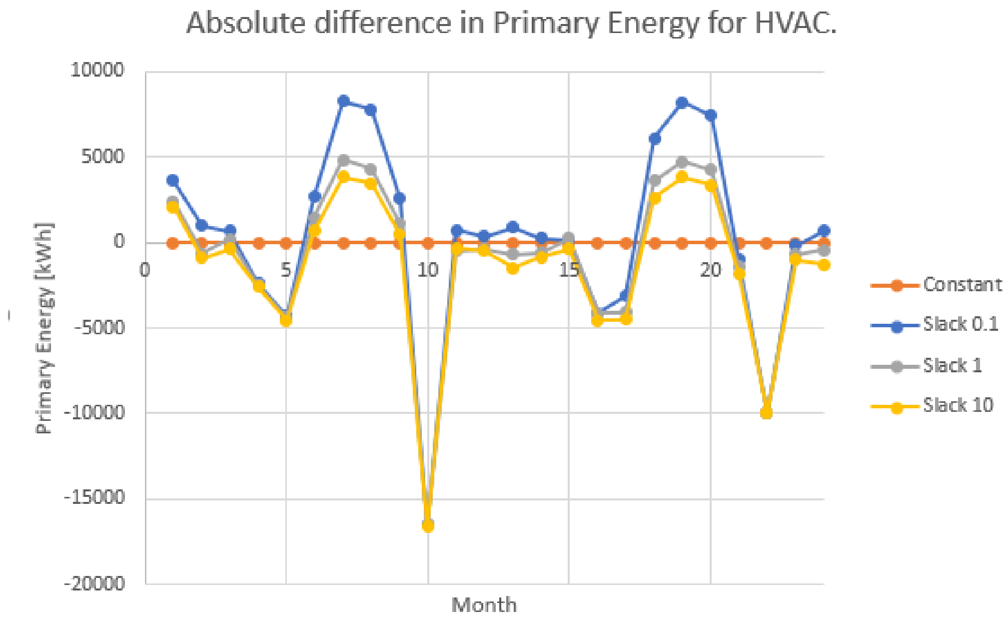



Figure 71: Absolute difference in Primary Energy Consumption for HVAC system, Commercial Retail in climate zone **D**.

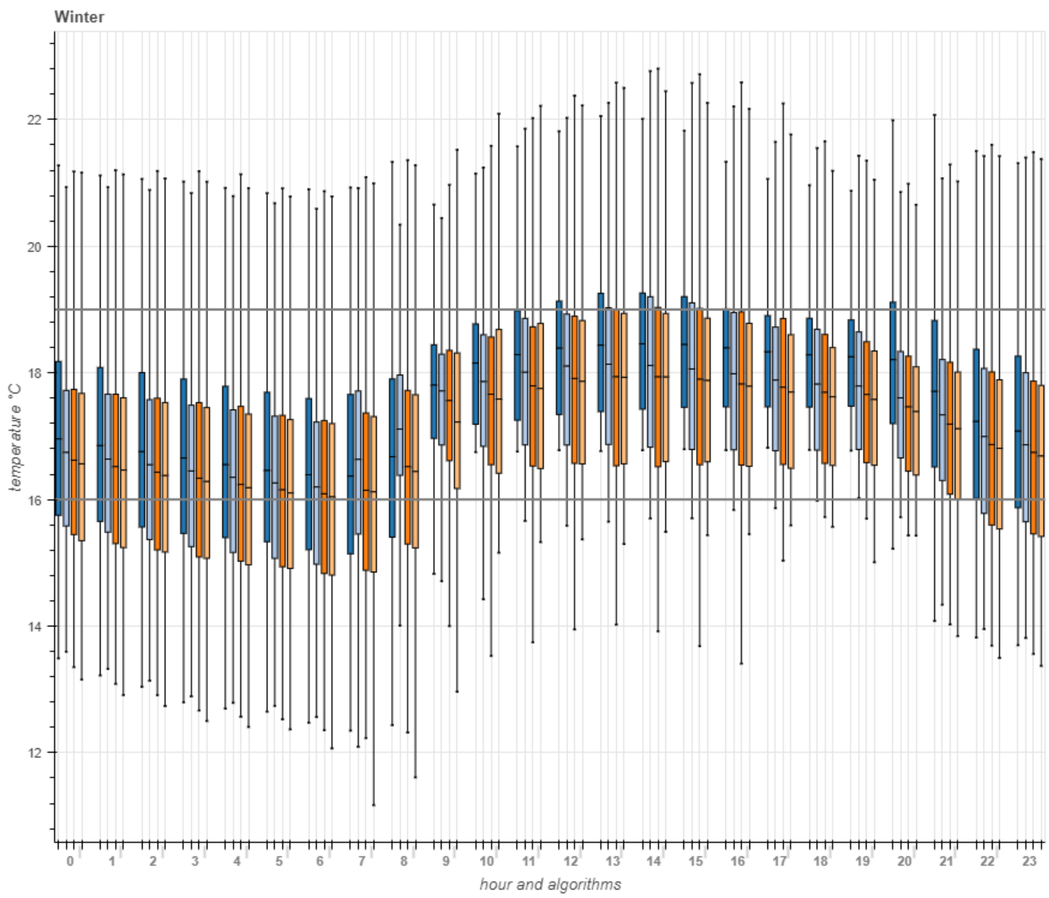



Figure 72: Heating season (1 Nov. to 15 Apr.) hourly mean internal air temperature, Commercial Retail in climate zone **D**.

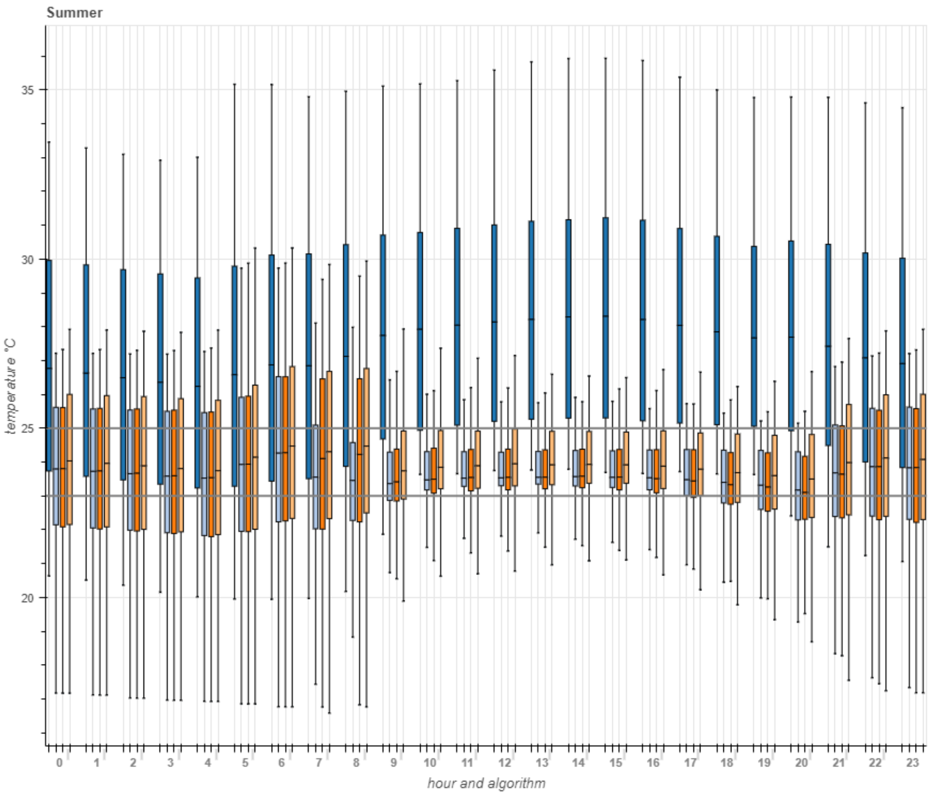



Figure 73: Cooling season (15 Apr. to 1 Nov.) hourly mean internal air temperature, Commercial Retail in climate zone **D**.

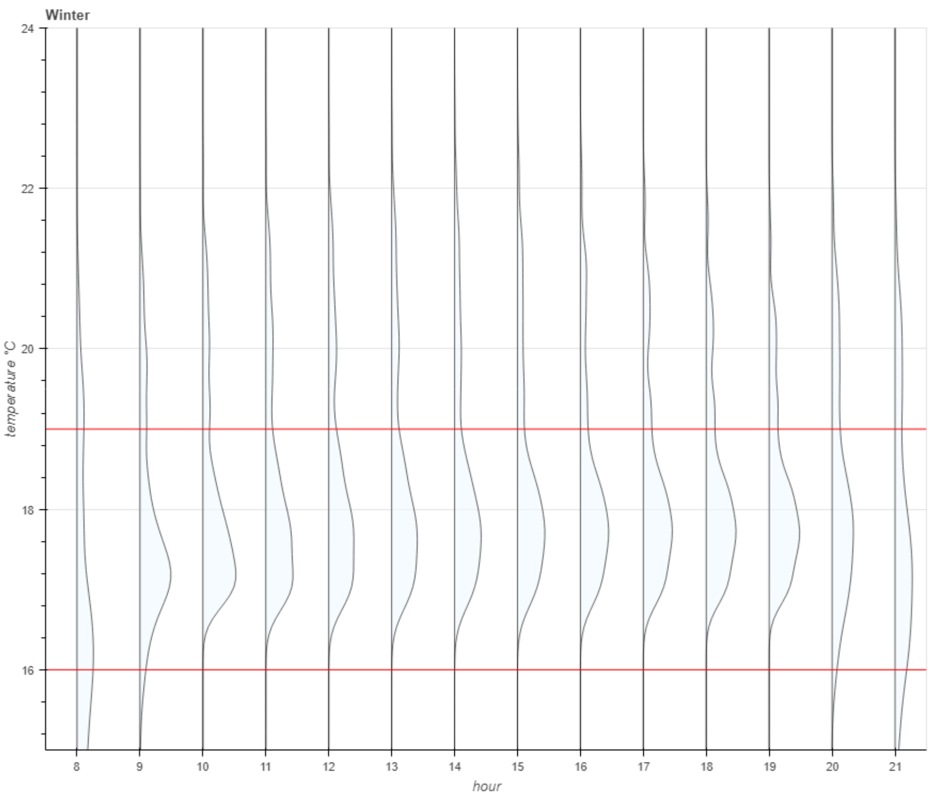



Figure 74: Heating season (1 Nov. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail in climate zone **D**.

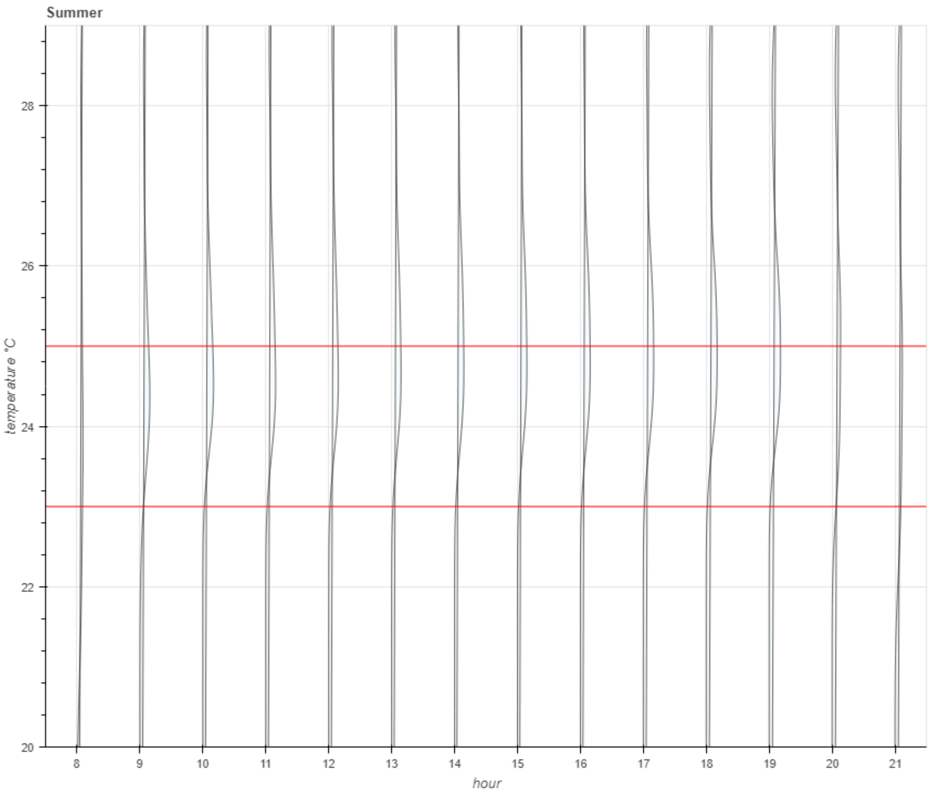



Figure 75: Cooling season (15 Apr. to 1 Nov.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail in climate zone **D**.

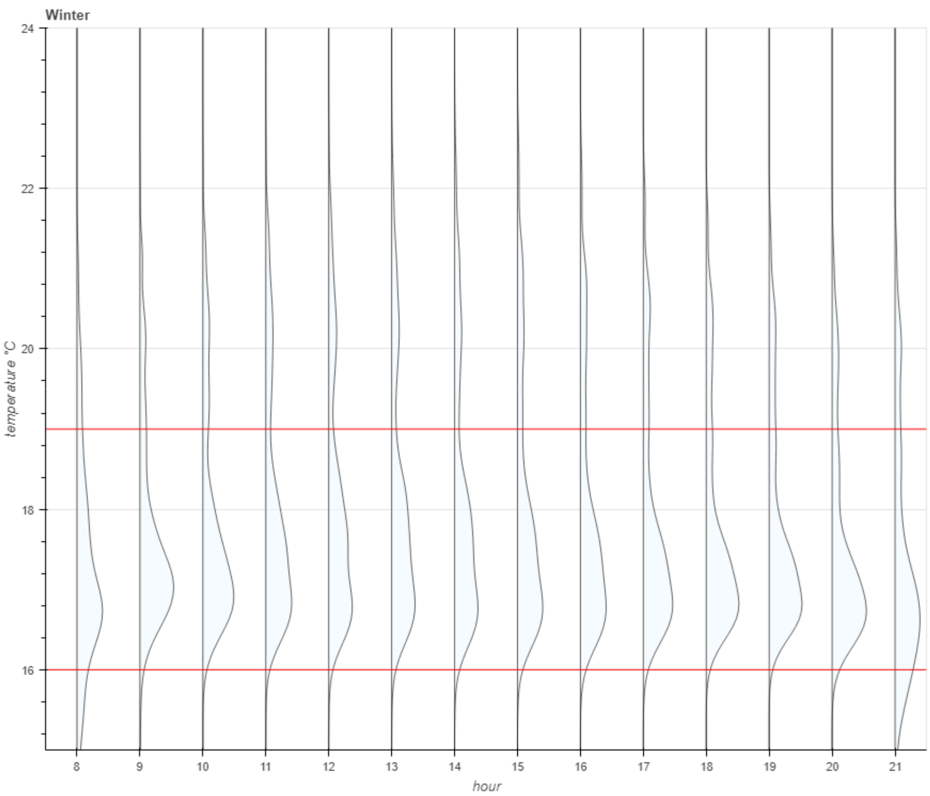



Figure 76: Heating season (1 Nov. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail in climate zone **D**.

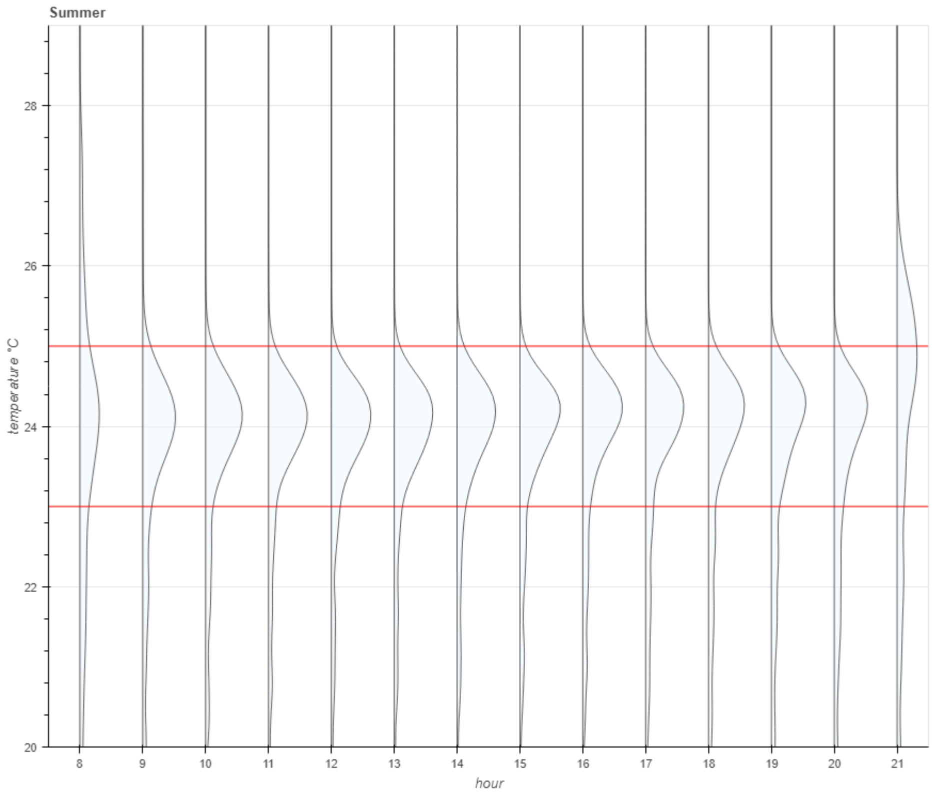



Figure 77: Cooling season (15 Apr. to 1 Nov.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail in climate zone **D**.

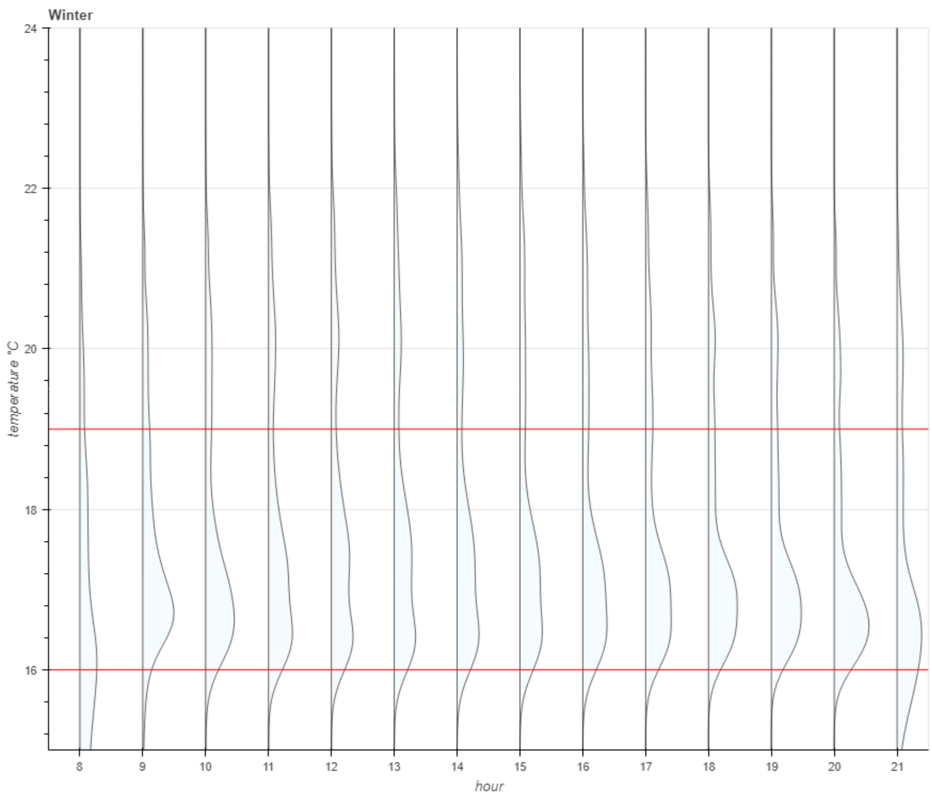



Figure 78: Heating season (1 Nov. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail in climate zone **D**.

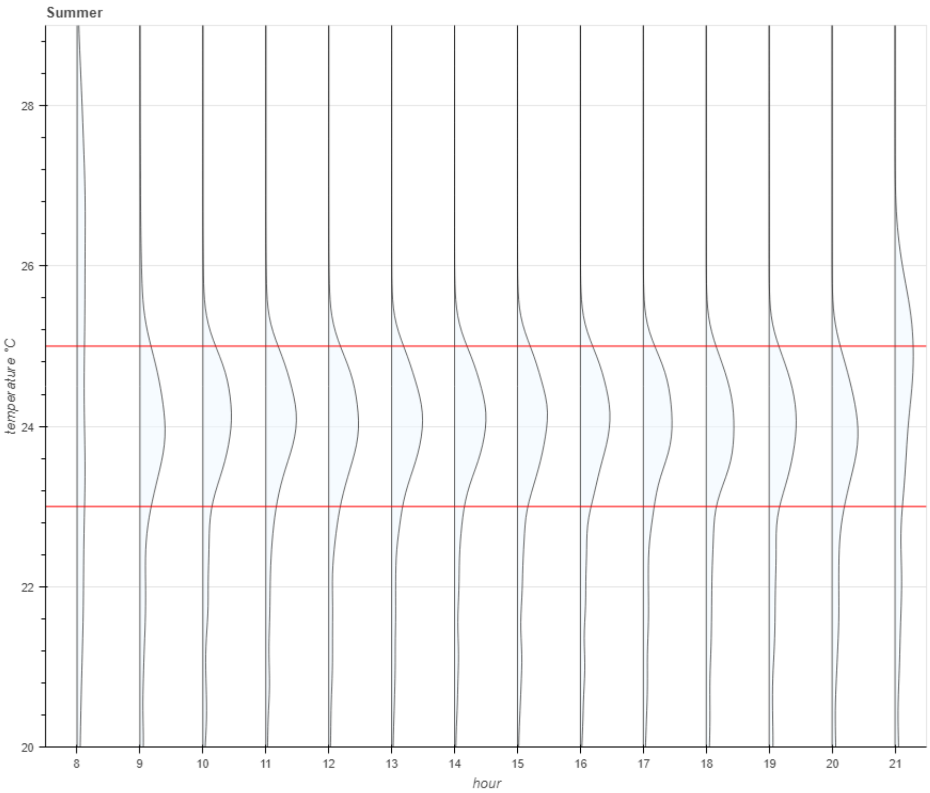



Figure 79: Cooling season (15 Apr. to 1 Nov.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail in climate zone **D**.

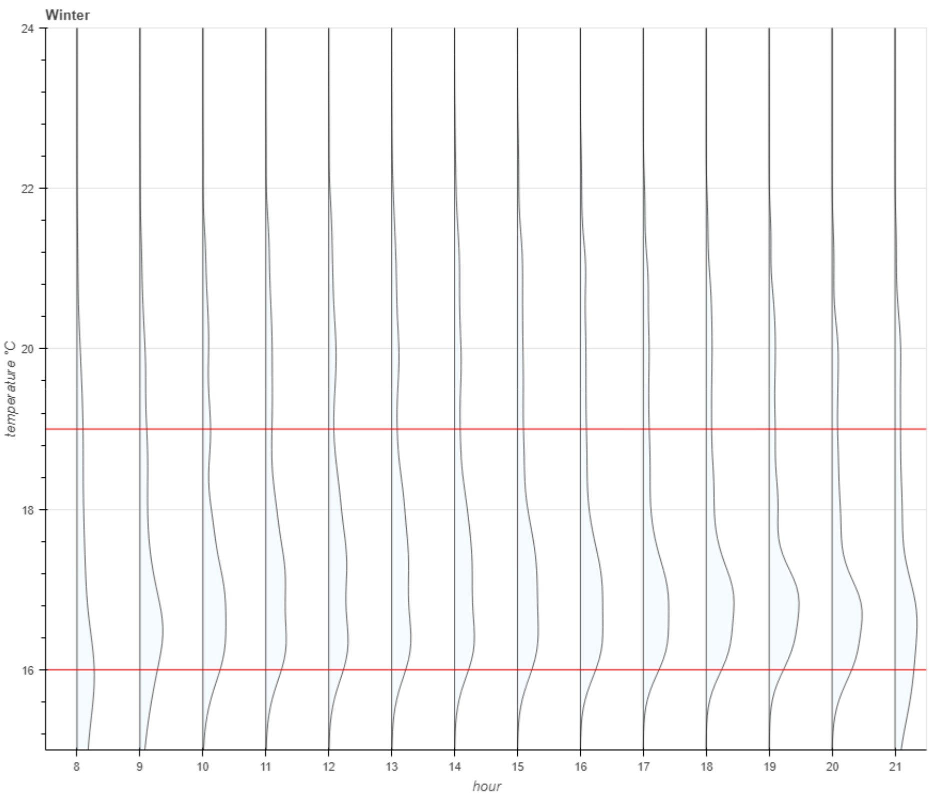



Figure 80: Heating season (1 Nov. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail in climate zone **D**.

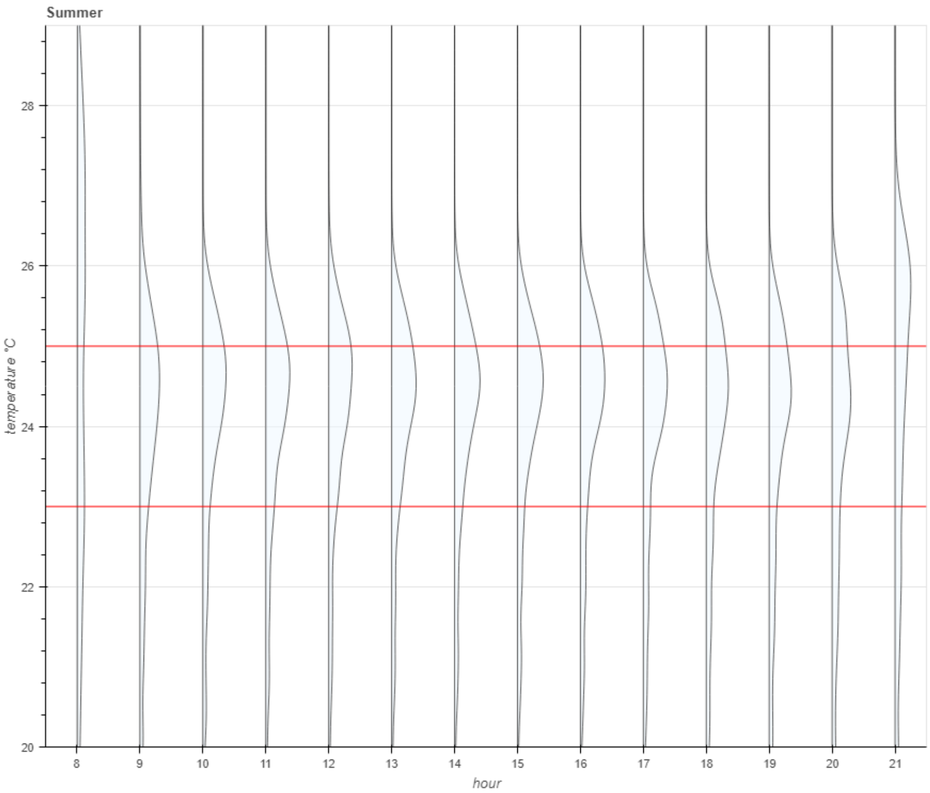



Figure 81: Cooling season (15 Apr. to 1 Nov.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail in climate zone **D**.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant | 1.04% | 4.99% | 22.67% | 17.27% |
| slack_01 | 1.00% | 2.79% | 22.17% | 17.44% |
| slack_1 | 2.48% | 7.87% | 21.33% | 16.35% |
| slack_10 | 10.24% | 23.13% | 20.65% | 15.37% |

Table 44: Percentage of exceeding air temperature in Heating Season (1 Nov. to 15 Apr.), Commercial Retail in climate zone **D**.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant | 2.87% | 6.41% | 56.54% | 46.15% |
| slack_01 | 27.06% | 34.37% | 1.09% | 0.50% |
| slack_1 | 27.35% | 37.75% | 3.85% | 2.07% |
| slack_10 | 26.63% | 32.79% | 10.94% | 4.57% |

Table 45: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 1 Nov.), Commercial Retail in climate zone **D**.

| | saving | |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant | 0.00% | 0.00% |
| slack_01 | -174.43% | -16.81% |
| slack_1 | -83.97% | -4.77% |
| slack_10 | -51.93% | 0.84% |

Figure 82: Relative savings with respect of constant set-point, Commercial Retail in climate zone **D**.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant | 125446.098 | 83.798 |
| Slack 0.1 | 135427.215 | 90.466 |
| Slack 1 | 105800.649 | 70.675 |
| Slack 10 | 94129.494 | 62.879 |

Table 46: Total primary energy consumption for HVAC, Commercial Retail in climate zone **D**.

**Occupation profile** has been modified introducing the stochastic model proposed by J. Page, D. Robinson, N. Morel, J.-L. Scartezzini [30] where occupancy behaviour in buildings is simulated as a random walk in which the only one free parameter is the **parameter of mobility** $\mu$, defined as the ratio between probability of changing state of presence over probability of no change, set equal to 0.5 for our purpose corresponding to "medium" mobility. The process is simulated as a discrete random walk with two position: 0, corresponding to "person outside the building", and 1, corresponding to "person inside the building". According these two positions of the random walk, are defined the **transition probabilities** as the probability of being in a certain state and pass to another state in the next timestep: $T_{00}, T_{01}, T_{10}, T_{11}$. From the total amount of transition probabilities only two needs to be assigned as can be deduced that:

$$T_{00} + T_{01} = 1 \tag{51}$$

$$T_{10} + T_{11} = 1 \quad (52)$$

Starting from $P(t)$, probability for the occupant to be present at timestep t, is possible to compute, with transition probabilities, $P(t+1)$, probability for the occupant to be present at timestep t+1 as follow:

$$P(t+1) = P(t)T_{11}(t) + (1 - P(t))T_{01}(t) \quad (53)$$

From which we can deduce that:

$$T_{11}(t) = \frac{P(t) - 1}{P(t)}T_{01}(t) + \frac{P(t+1)}{P(t)} \quad (54)$$

Having already defined parameter of mobility as the ratio between probability of changing state of presence and probability of no change:

$$\mu = \frac{T_{01}(t) + T_{10}(t)}{T_{11}(t) + T_{00}(t)} \quad (55)$$

Finally is possible to obtain $T_{01}(t)$ and $T_{11}(t)$ from which is possible to reconstruct all the probability density function and the random phenomenon itself:

$$T_{01}(t) = \frac{\mu - 1}{\mu + 1}P(t) + P(t+1) \quad (56)$$

$$T_{11}(t) = \frac{P(t) - 1}{P(t)}[\frac{\mu - 1}{\mu + 1}P(t) + P(t+1)] + \frac{P(t+1)}{P(t)} \quad (57)$$

The occupation profile used has been taken from Google Maps popular times based on visits to the Commercial Retail located near Bergamo. Occupation fraction profile is reported in Table 47.

| Hour | Mon. | Tu. | Wed. | Th. | Fri. | Sat. | Sun. |
|---|---|---|---|---|---|---|---|
| 9 | 0.303 | 0.250 | 0.276 | 0.408 | 0.303 | 0.421 | 0.408 |
| 10 | 0.382 | 0.263 | 0.368 | 0.263 | 0.329 | 0.487 | 0.579 |
| 11 | 0.184 | 0.237 | 0.368 | 0.158 | 0.276 | 0.355 | 0.447 |
| 12 | 0.092 | 0.184 | 0.342 | 0.224 | 0.211 | 0.355 | 0.276 |
| 13 | 0.250 | 0.197 | 0.382 | 0.276 | 0.276 | 0.461 | 0.289 |
| 14 | 0.421 | 0.316 | 0.382 | 0.355 | 0.421 | 0.579 | 0.605 |
| 15 | 0.447 | 0.408 | 0.316 | 0.382 | 0.500 | 0.803 | 0.842 |
| 16 | 0.329 | 0.513 | 0.316 | 0.289 | 0.447 | 0.947 | 0.895 |
| 17 | 0.368 | 0.434 | 0.368 | 0.237 | 0.447 | 0.882 | 0.763 |
| 18 | 0.316 | 0.303 | 0.263 | 0.184 | 0.553 | 0.632 | 0.408 |
| 19 | 0.132 | 0.171 | 0.145 | 0.105 | 0.342 | 0.329 | 0.092 |
| 20 | 0.000 | 0.118 | 0.000 | 0.000 | 0.105 | 0.132 | 0.000 |

Table 47: Occupancy profile for Random Occupation Commercial Retail Model.

Using this equation is possible to build a random occupation profile over a fixed profile that can be used in building simulation with EnergyPlus with the following results:

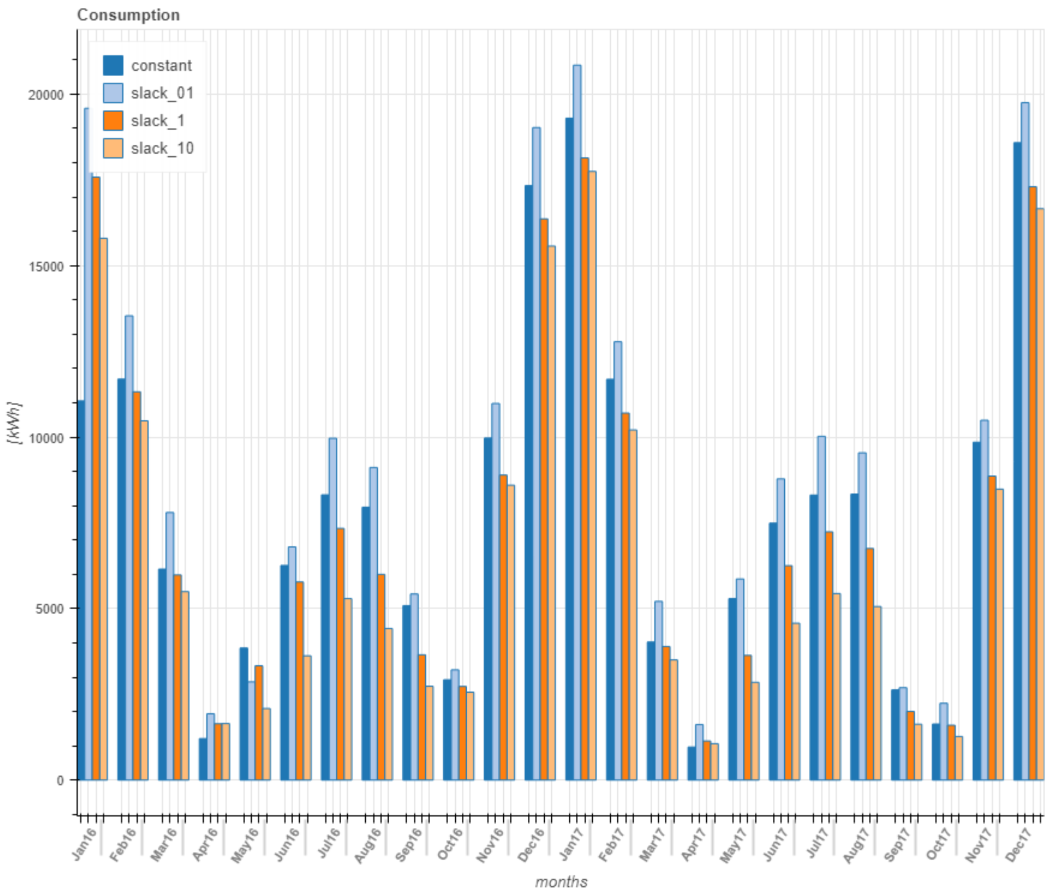



Figure 83: Primary Energy Consumption for HVAC system, Commercial Retail with stochastic model of occupancy.

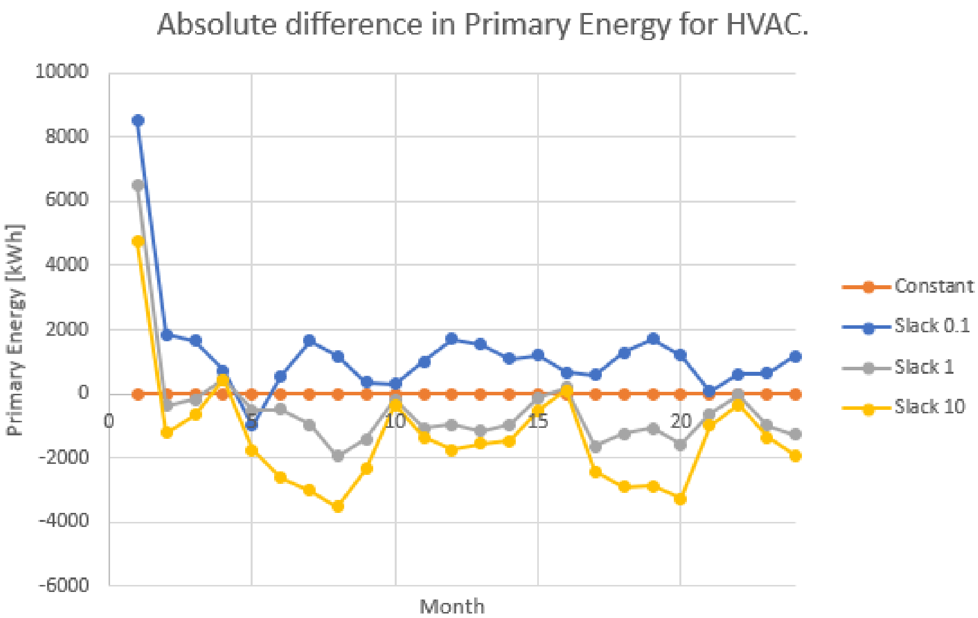



Figure 84: Absolute difference in Primary Energy Consumption for HVAC system, Commercial Retail with stochastic model of occupancy.

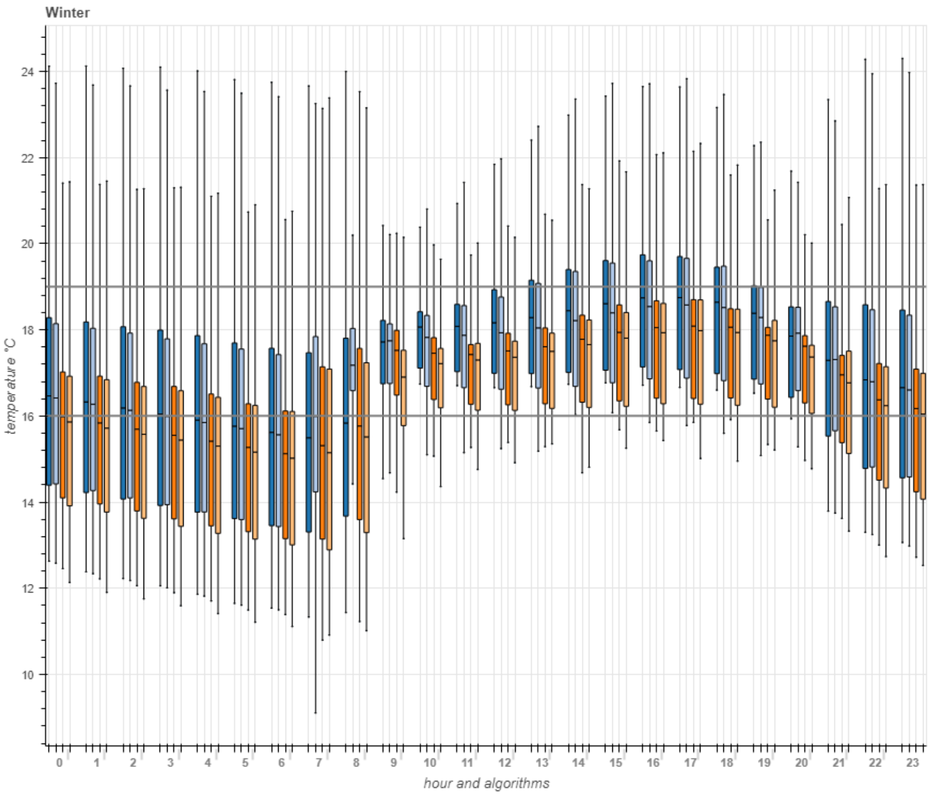



Figure 85: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Commercial Retail with stochastic model of occupancy.

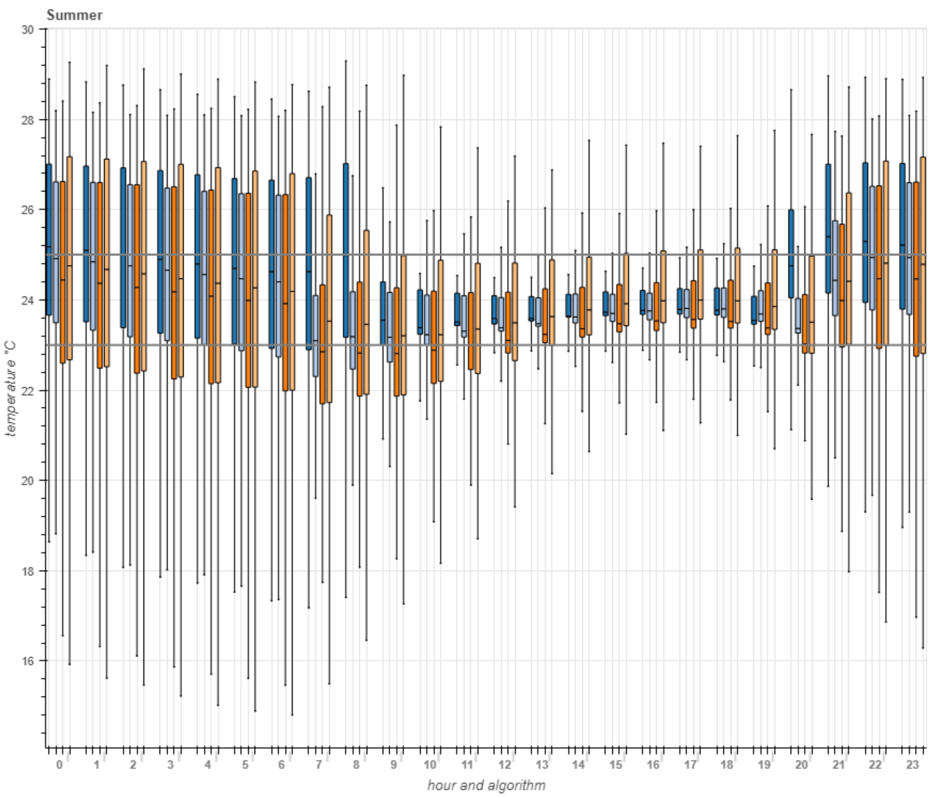



Figure 86: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Commercial Retail with stochastic model of occupancy.

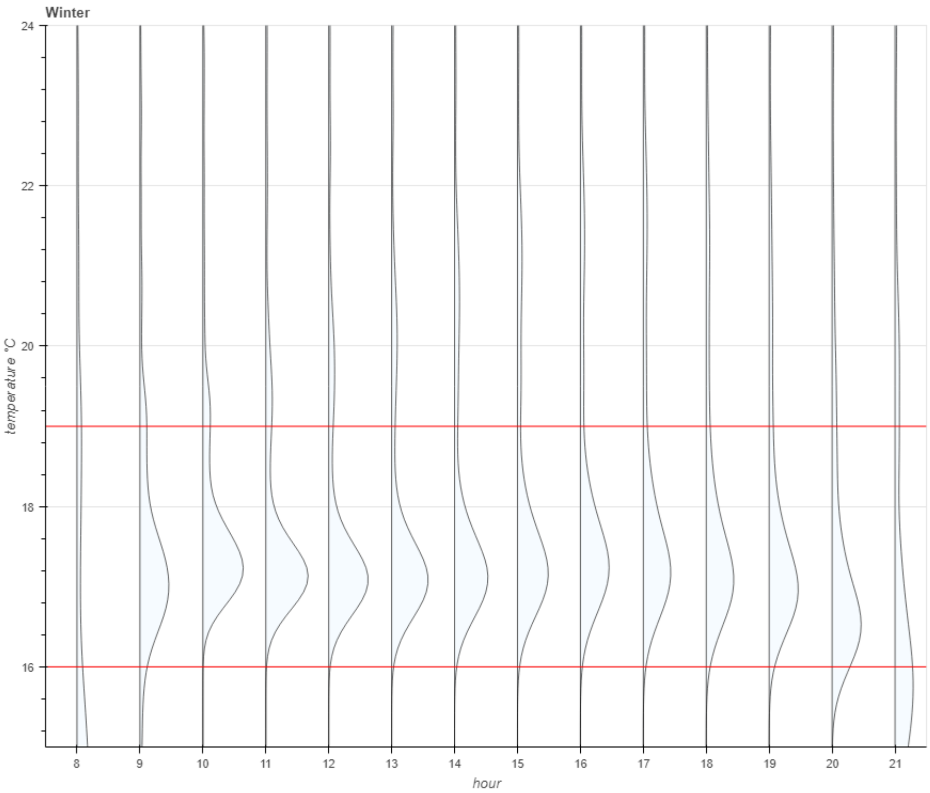



Figure 87: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with stochastic model of occupancy.

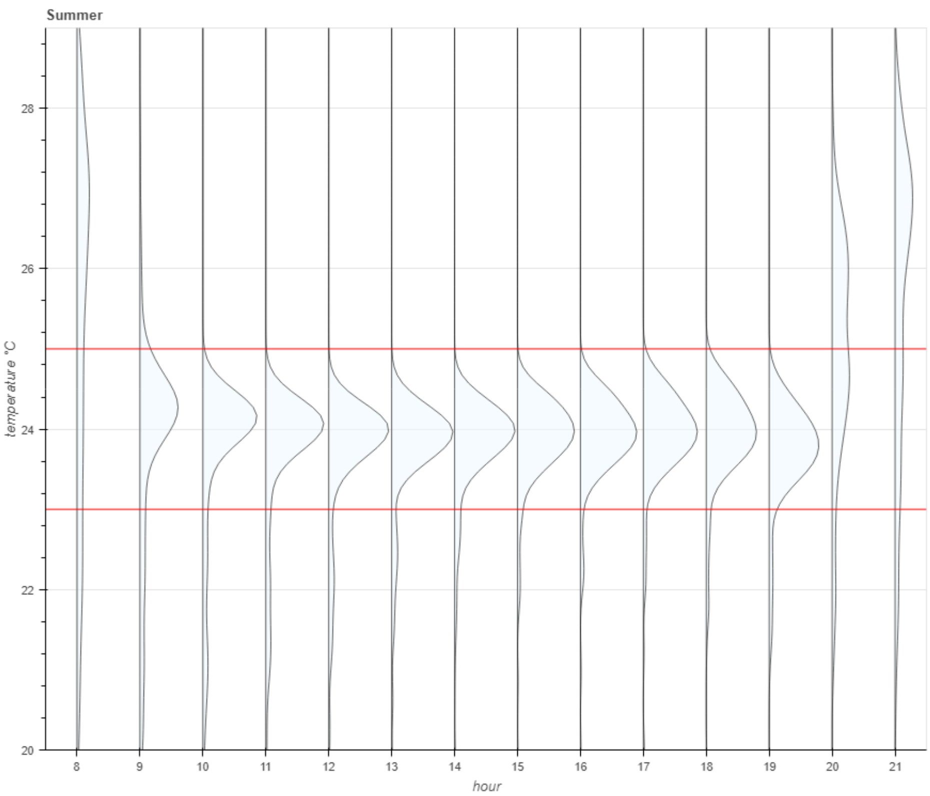



Figure 88: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with stochastic model of occupancy.

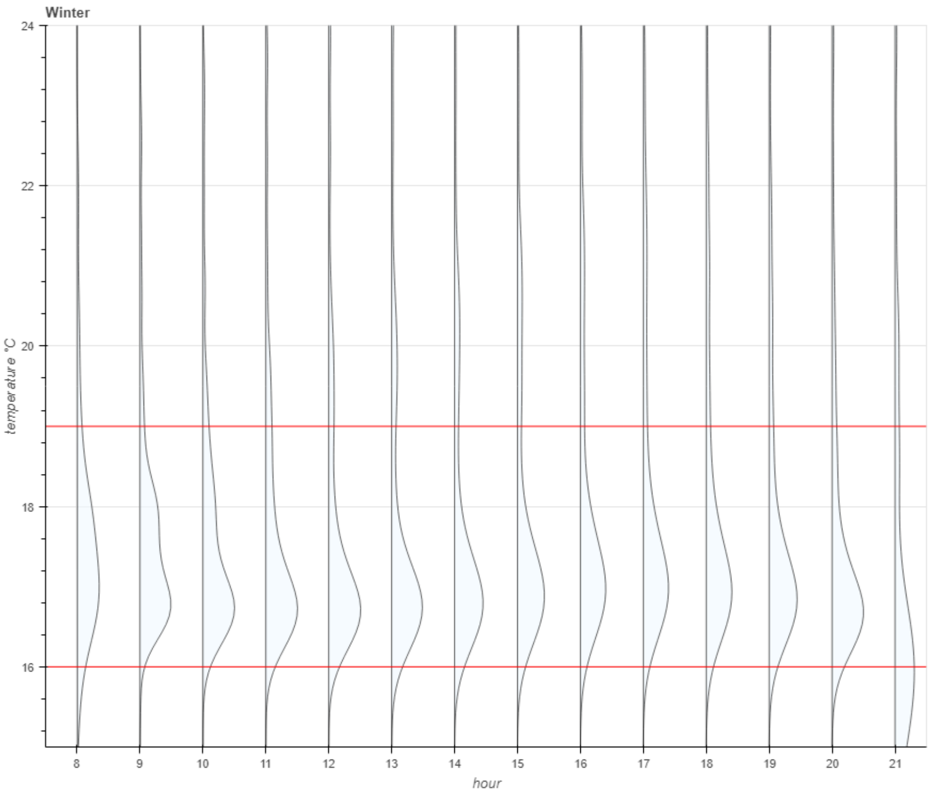



Figure 89: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with stochastic model of occupancy.

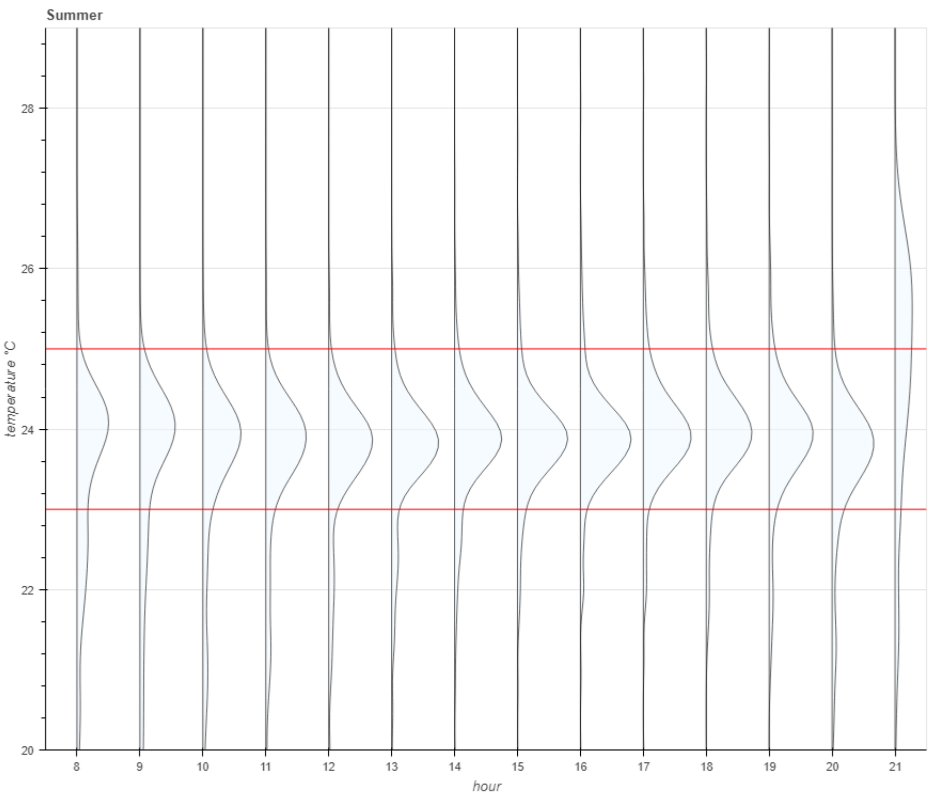



Figure 90: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with stochastic model of occupancy.

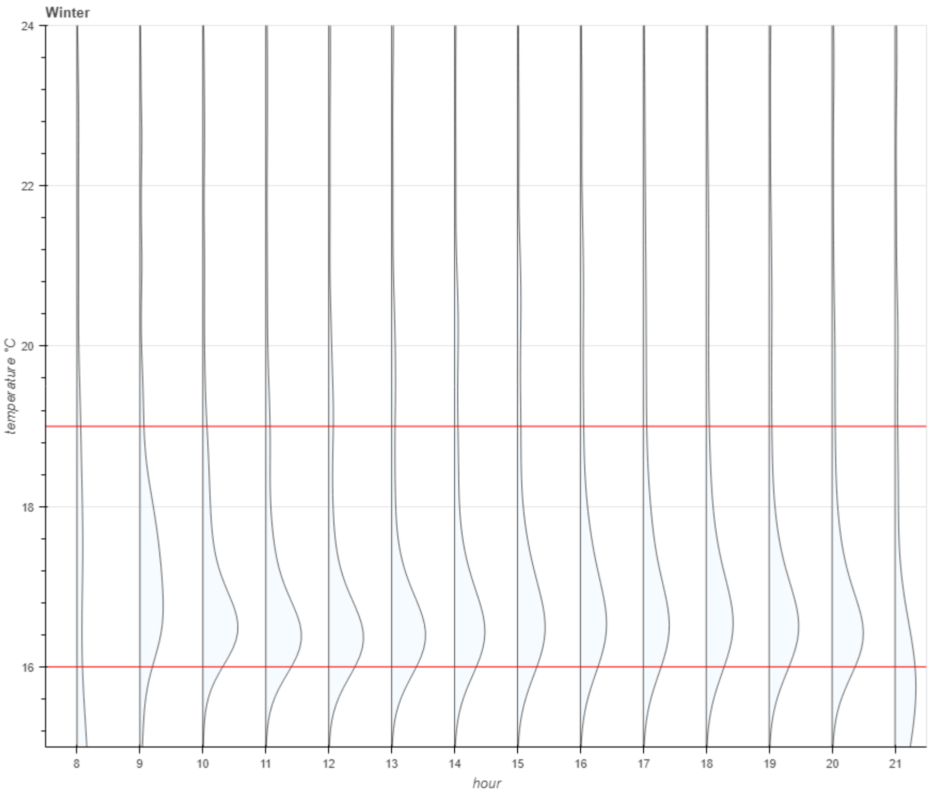



Figure 91: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with stochastic model of occupancy.

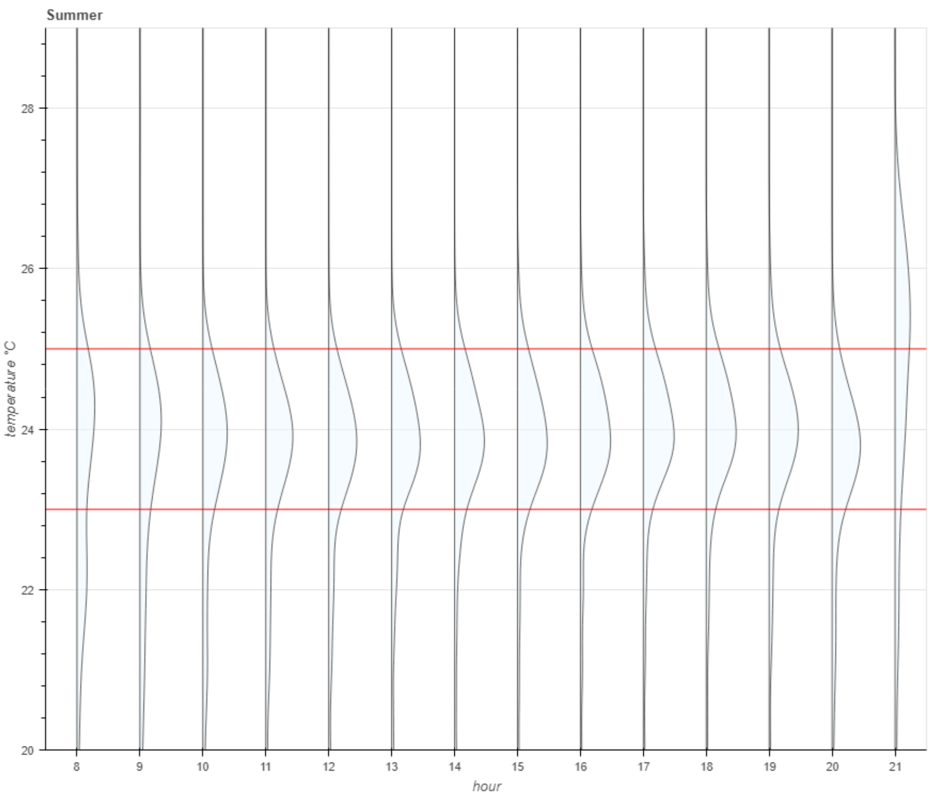



Figure 92: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with stochastic model of occupancy.

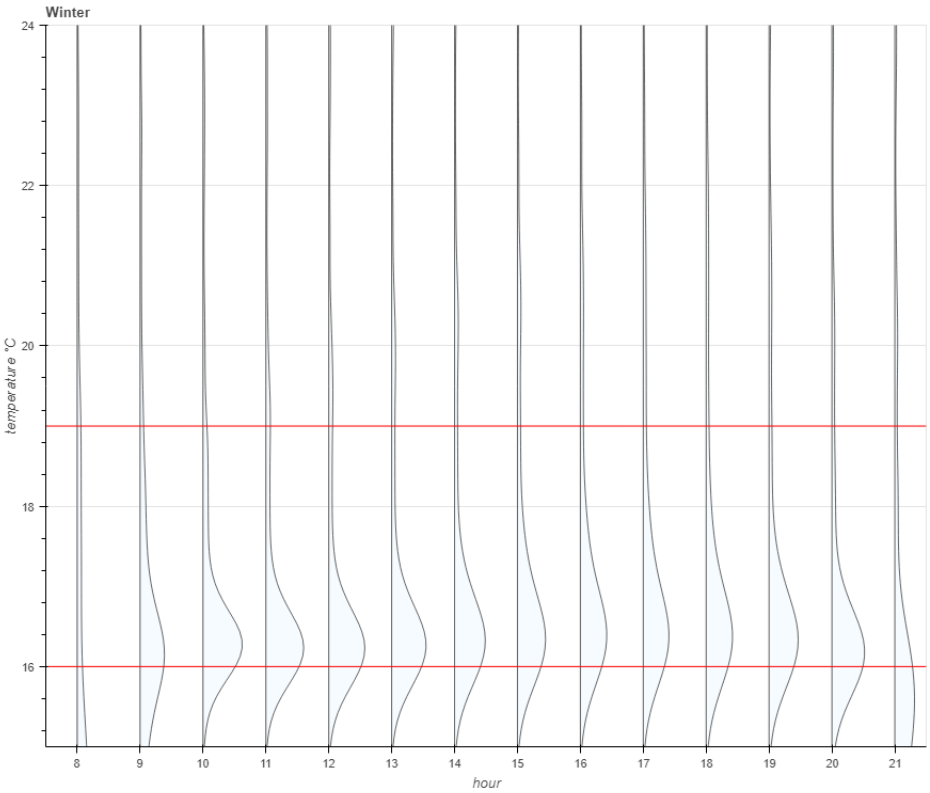



Figure 93: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with stochastic model of occupancy.

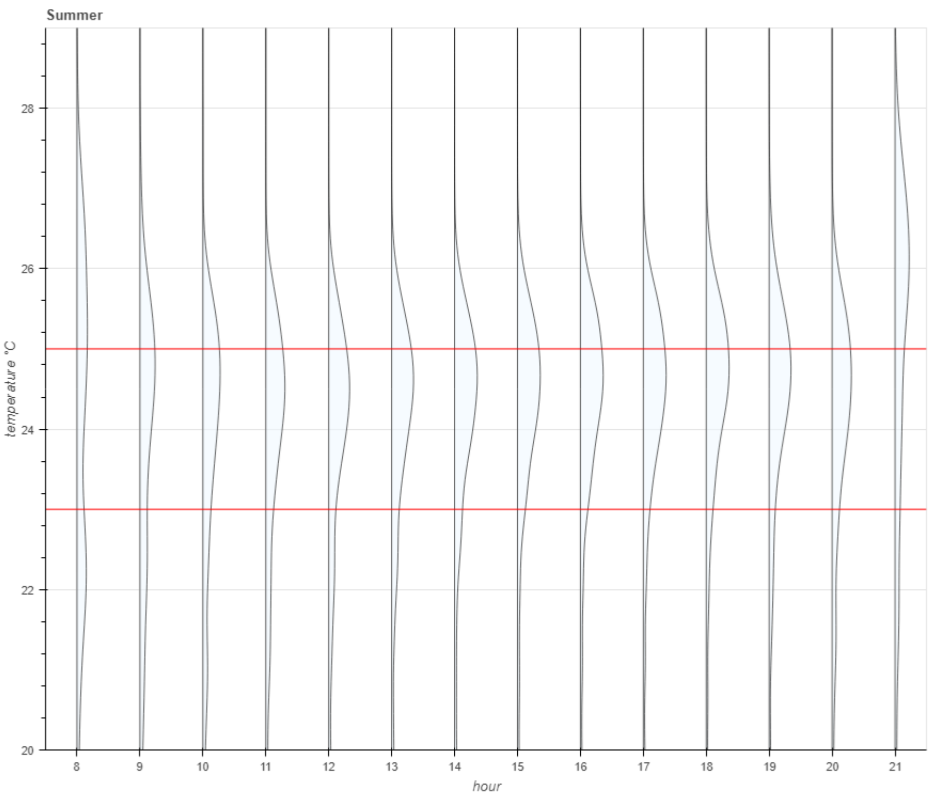



Figure 94: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with stochastic model of occupancy.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant | 3.54% | 8.69% | 29.98% | 27.13% |
| slack_01 | 2.35% | 8.00% | 29.34% | 26.43% |
| slack_1 | 5.56% | 17.30% | 22.18% | 19.67% |
| slack_10 | 21.53% | 36.94% | 21.36% | 19.29% |

Table 48: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Commercial Retail with stochastic model of occupancy.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant | 20.91% | 25.28% | 46.47% | 34.07% |
| slack_01 | 23.90% | 30.76% | 3.66% | 2.66% |
| slack_1 | 31.86% | 39.92% | 3.03% | 0.85% |
| slack_10 | 32.62% | 38.48% | 19.35% | 9.15% |

Table 49: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Commercial Retail with stochastic model of occupancy.

| | saving | |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant | 0.00% | 0.00% |
| slack_01 | -11.16% | -18.62% |
| slack_1 | 15.57% | 3.09% |
| slack_10 | 37.79% | 7.51% |

Figure 95: Relative savings with respect of constant set-point, Commercial Retail with stochastic model of occupancy.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.130 |
| Interior Concrete | 0.060 | 1.910 | 1000 | 2450 | 0.031 |
| Lightened Layer | 0.150 | 0.044 | 845 | 30 | 5.000 |
| Air Gap | - | - | - | - | 0.160 |
| External Concrete | 0.060 | 2.080 | 1000 | 2450 | 0.029 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.263 |
| | | | | Limit Value [27] | 0.280 |

Table 51: Vertical Wall Stratigraphy for Commercial Retail model insulated according DECRETO 11 marzo 2008.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.100 |
| Mortar | 0.010 | 1.400 | 840 | 2000 | 0.007 |
| Insulation Layer | 0.130 | 0.034 | 850 | 30 | 5.882 |
| Ceiling Bricks | 0.180 | 0.599 | 920 | 950 | 0.301 |
| Expanded Clay | 0.020 | 0.756 | 880 | 1700 | 0.026 |
| Shingles | 0.020 | 0.260 | 880 | 1300 | 0.077 |
| External Surface $R_t$ | - | - | - | - | 0.040 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.229 |
| | | | | Limit Value [27] | 0.230 |

Table 52: Ceiling Stratigraphy for Commercial Retail model insulated according DECRETO 11 marzo 2008.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant | 189887.423 | 126.845 |
| Slack 0.1 | 220056.052 | 146.998 |
| Slack 1 | 178107.804 | 118.976 |
| Slack 10 | 156770.084 | 104.723 |

Table 50: Total primary energy consumption for HVAC, Commercial Retail with stochastic model of occupancy.

Results show how increasing stochasticity in the model, energy savings for Reinforcement Learning algorithm increase while traditional control system lacks in performance.

Building envelope has been modified to accomplish Italian requirements for climate zone **E** [27]. New building envelope is resumed in the followings tables and U- factor of glazing system has been set to $1.6\frac{W}{m^2K}$.

| | $t$ [m] | $\lambda[\frac{W}{mK}]$ | $c_p[\frac{J}{kgK}]$ | $\rho\frac{kg}{m^3}$ | $R_t[\frac{m^2K}{W}]$ |
|---|---|---|---|---|---|
| Internal Surface $R_t$ | - | - | - | - | 0.170 |
| Floor Tiles | 0.010 | 1.000 | 840 | 2300 | 0.010 |
| Concrete | 0.060 | 1.160 | 1000 | 2000 | 0.043 |
| Insulation Layer | 0.100 | 0.041 | 1030 | 97 | 1.220 |
| Backing | 0.060 | 0.093 | 880 | 2200 | 0.645 |
| Shingles | 0.220 | 0.667 | 920 | 1214 | 0.330 |
| Mortar | 0.015 | 0.900 | 910 | 1800 | 0.017 |
| | | | | $U[\frac{W}{m^2K}]$ | 0.270 |
| | | | | Limit Value [27] | 0.270 |

Table 53: Floor Stratigraphy for Commercial Retail model insulated according DECRETO 11 marzo 2008.

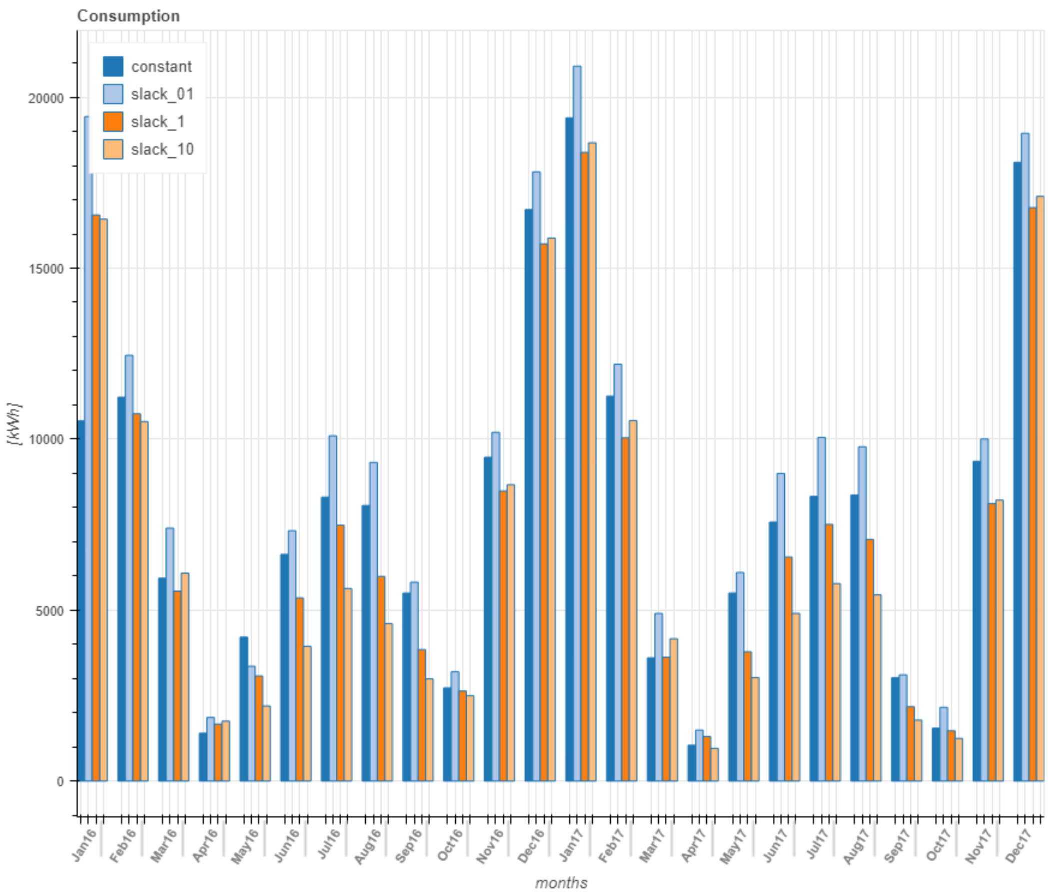



Figure 96: Primary Energy Consumption for HVAC system, Commercial Retail with insulation according DECRETO 11 marzo 2008.

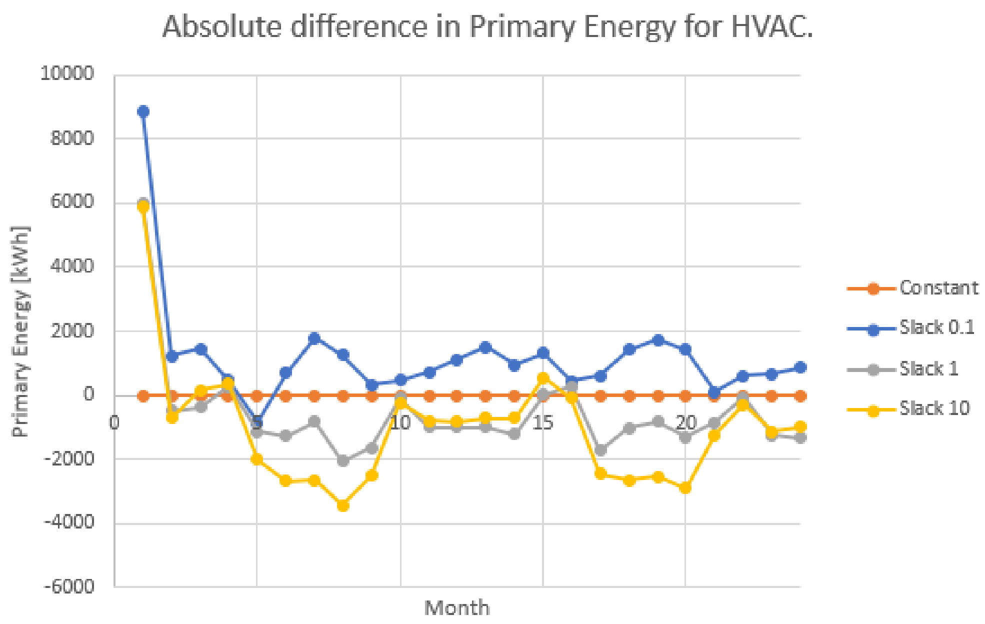



Figure 97: Absolute difference in Primary Energy Consumption for HVAC system, Commercial Retail with insulation according DECRETO 11 marzo 2008.

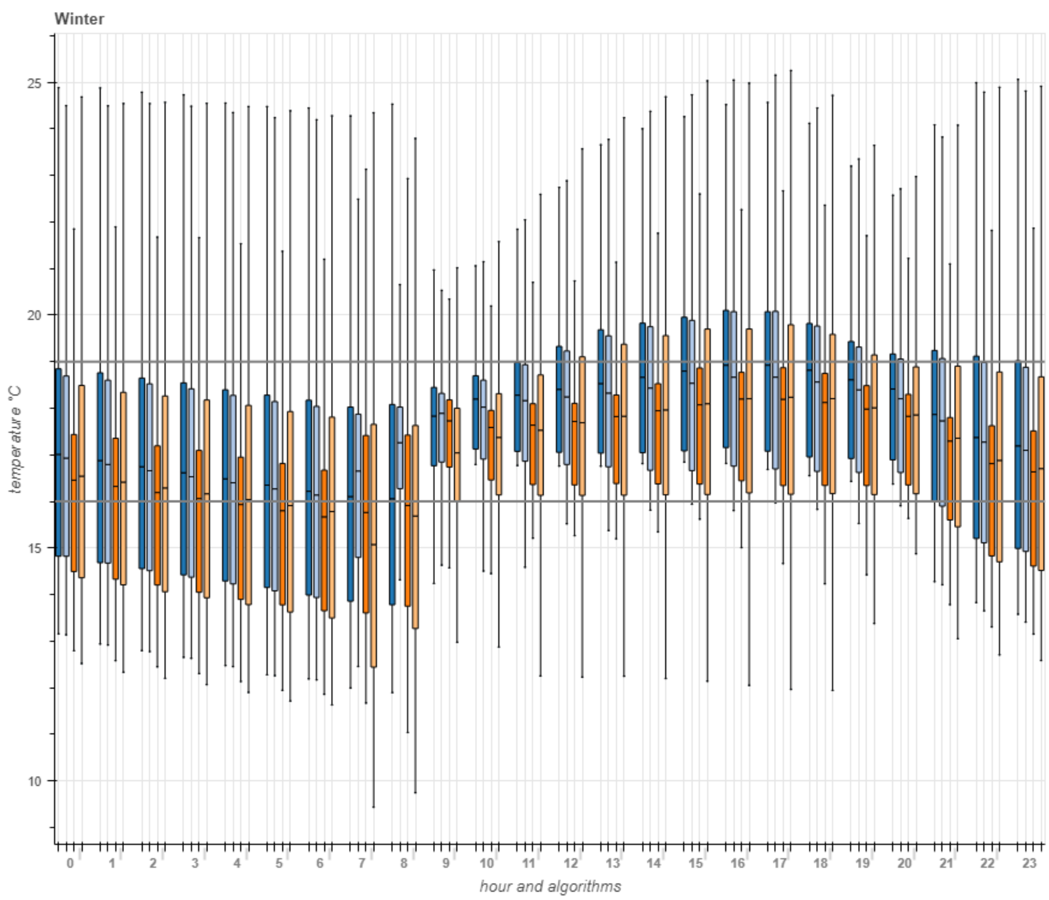



Figure 98: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Commercial Retail with insulation according DECRETO 11 marzo 2008.

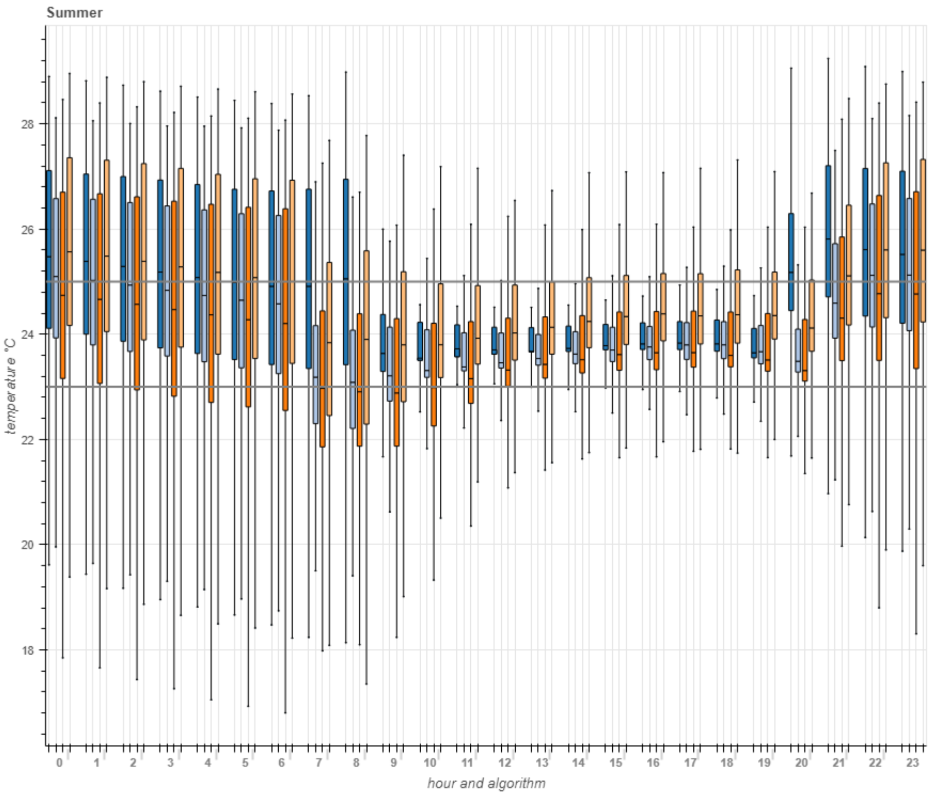



Figure 99: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Commercial Retail with insulation according DECRETO 11 marzo 2008.

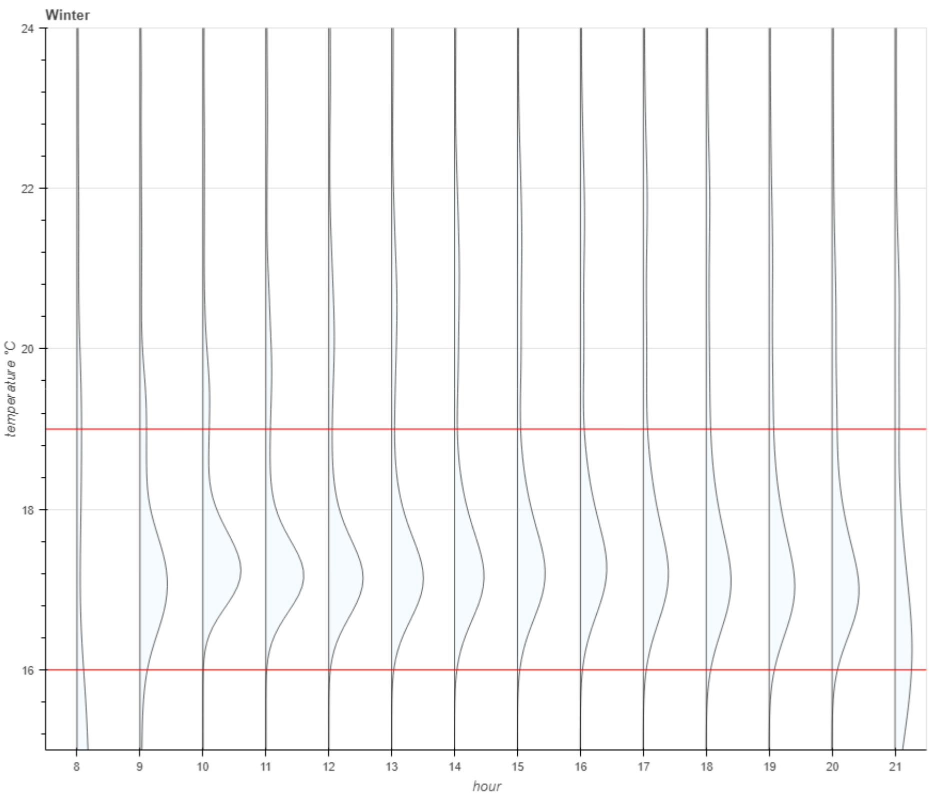



Figure 100: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with insulation according DECRETO 11 marzo 2008.

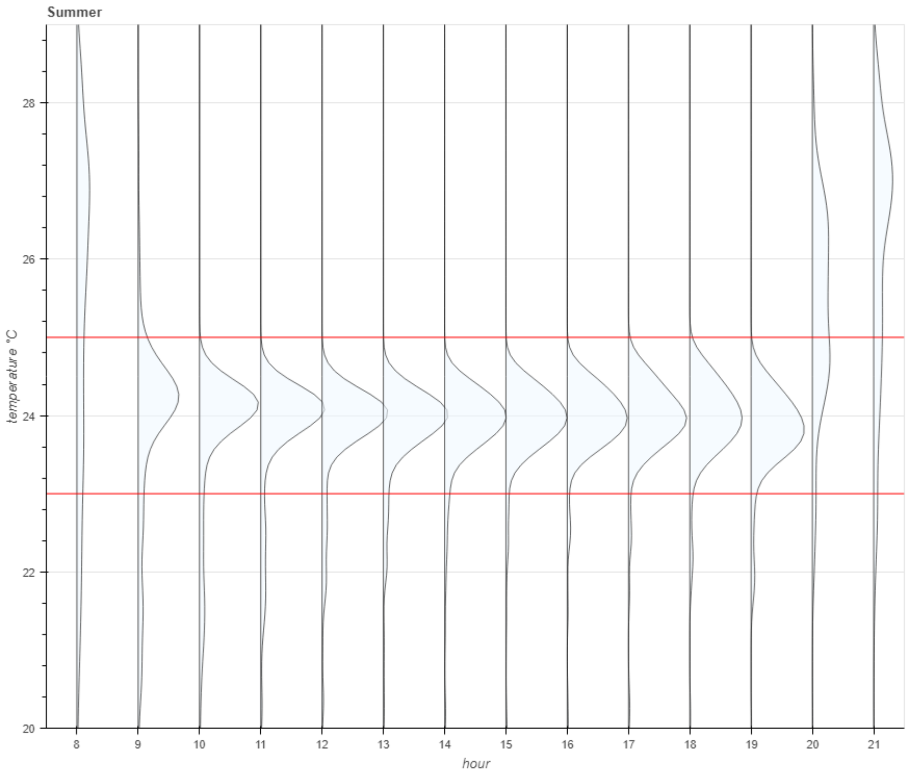



Figure 101: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with insulation according DECRETO 11 marzo 2008.

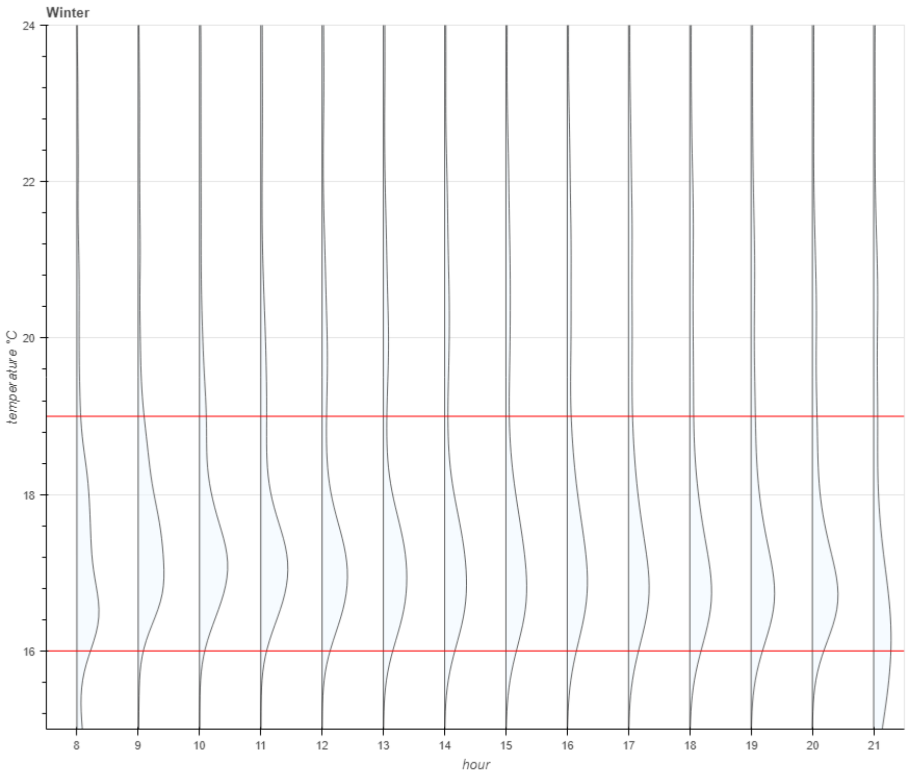



Figure 102: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with insulation according DECRETO 11 marzo 2008.

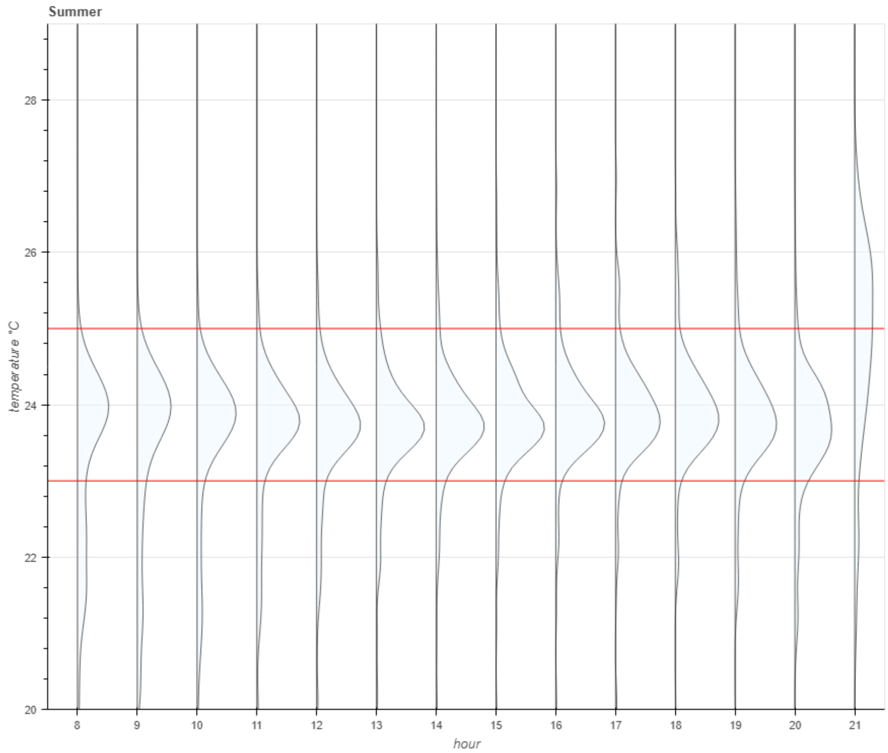



Figure 103: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with insulation according DECRETO 11 marzo 2008.

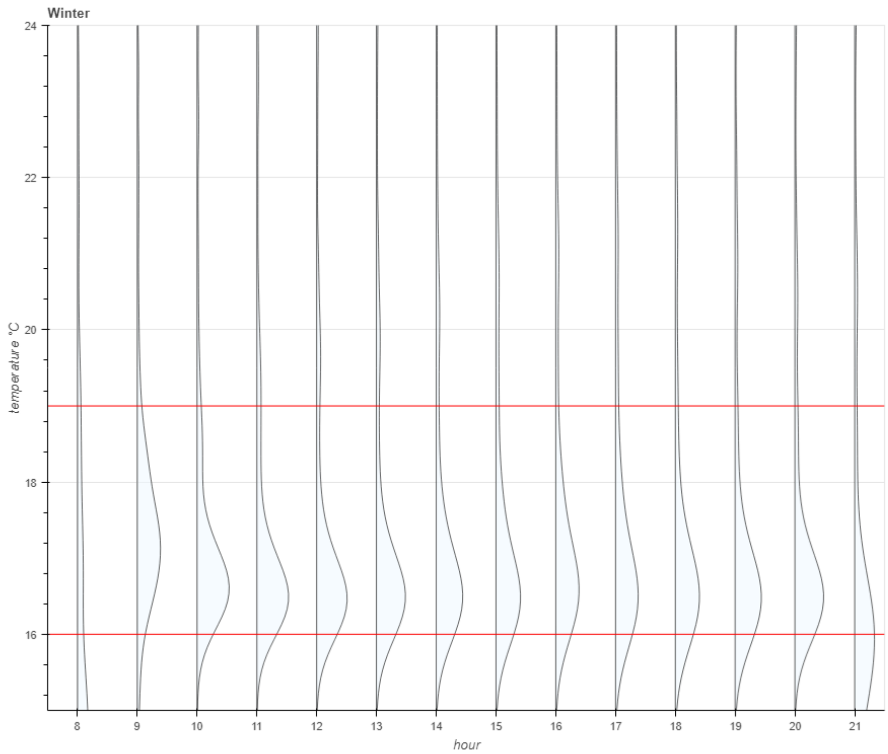



Figure 104: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with insulation according DECRETO 11 marzo 2008.

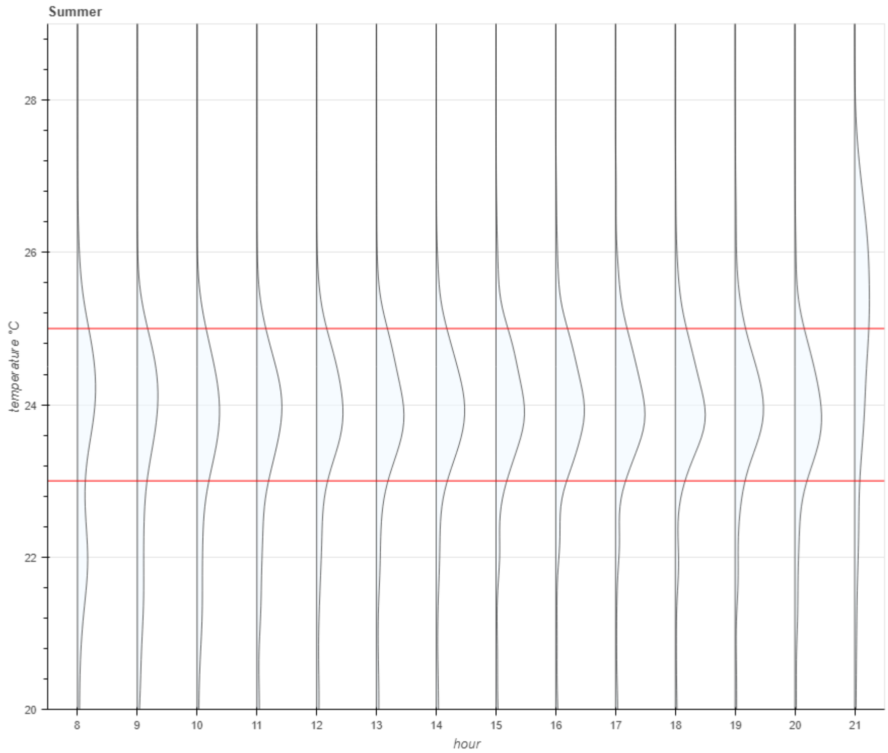



Figure 105: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with insulation according DECRETO 11 marzo 2008.

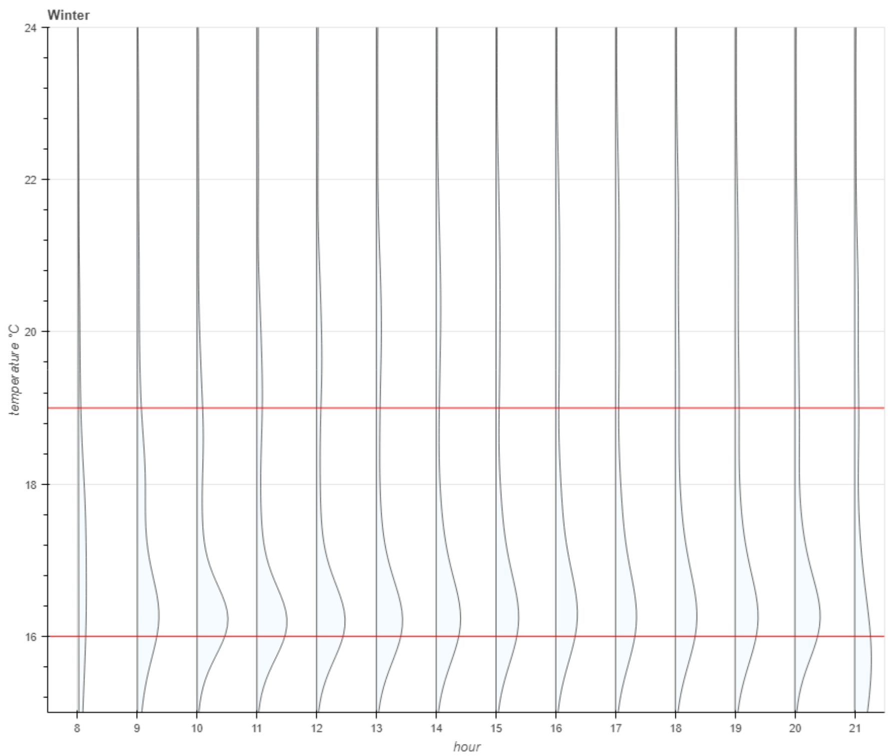



Figure 106: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with insulation according DECRETO 11 marzo 2008.

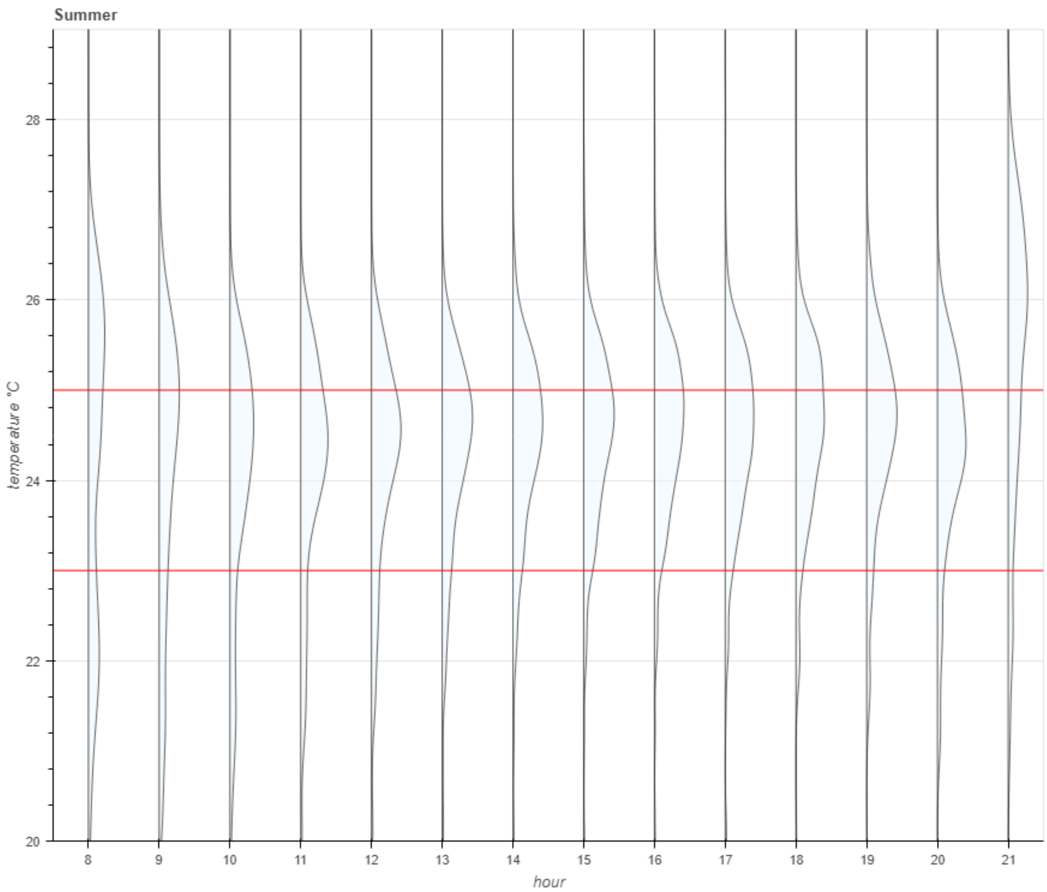



Figure 107: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with insulation according DECRETO 11 marzo 2008.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant | 3.78% | 7.81% | 29.38% | 26.85% |
| slack_01 | 3.29% | 9.99% | 28.78% | 26.19% |
| slack_1 | 7.68% | 18.55% | 22.37% | 20.10% |
| slack_10 | 16.33% | 28.27% | 27.94% | 24.85% |

Table 54: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Commercial Retail with insulation according DECRETO 11 marzo 2008.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant | 18.20% | 22.62% | 52.67% | 40.56% |
| slack_01 | 23.05% | 29.31% | 3.61% | 1.47% |
| slack_1 | 32.23% | 40.94% | 5.34% | 2.96% |
| slack_10 | 23.09% | 28.84% | 19.95% | 9.10% |

Table 55: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Commercial Retail with insulation according DECRETO 11 marzo 2008.

| | saving | |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant | 0.00% | 0.00% |
| slack_01 | -11.79% | -17.99% |
| slack_1 | 17.23% | 3.44% |
| slack_10 | 38.02% | -4.15% |

Figure 108: Relative savings with respect of constant set-point, Commercial Retail with insulation according DECRETO 11 marzo 2008.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant | 187782.862 | 125.439 |
| Slack 0.1 | 216888.747 | 144.882 |
| Slack 1 | 173875.624 | 116.149 |
| Slack 10 | 163012.357 | 108.893 |

Table 56: Total primary energy consumption for HVAC, Commercial Retail with insulation according DECRETO 11 marzo 2008.

Last modification is an outdoor air control system that introduce the possibility for the HVAC system to shut down heating or cooling when outdoor air temperature reach set-point threshold. As results below show, when increasing quality of control system, Reinforcement Learning is not able to take advantage of this possibility but worse its performance.

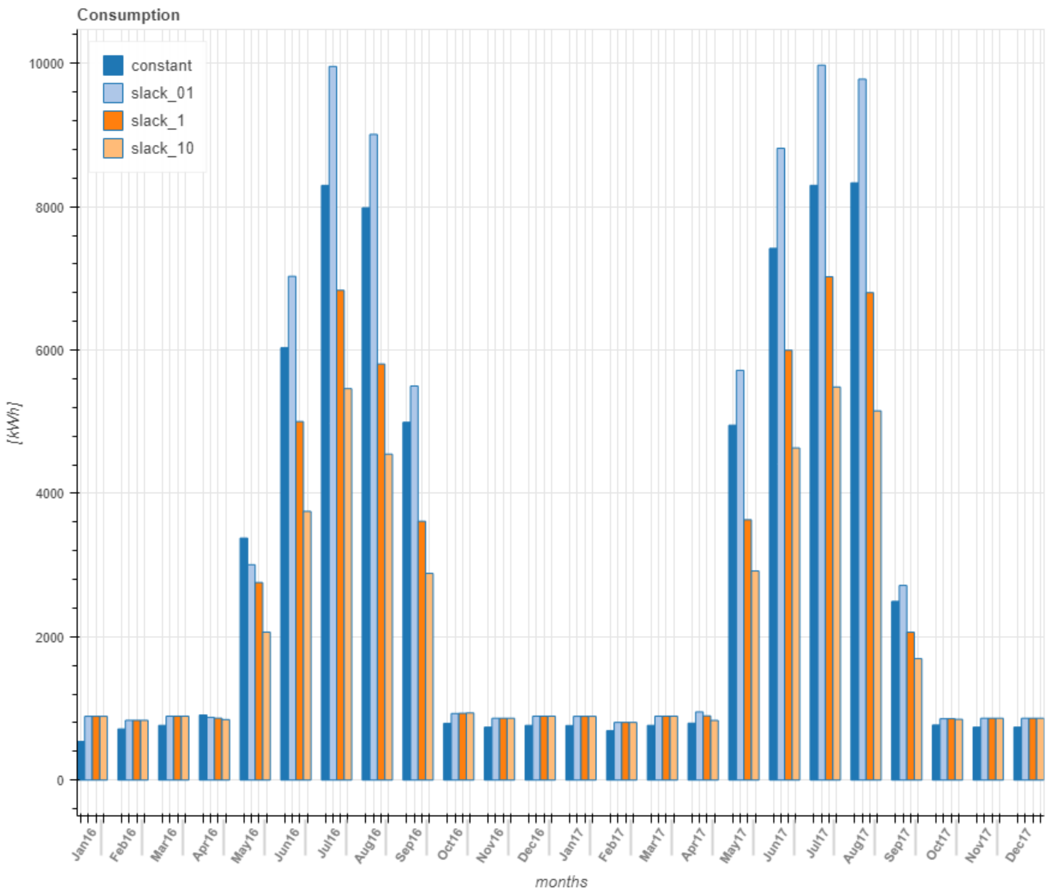



Figure 109: Primary Energy Consumption for HVAC system, Commercial Retail with outdoor air control.

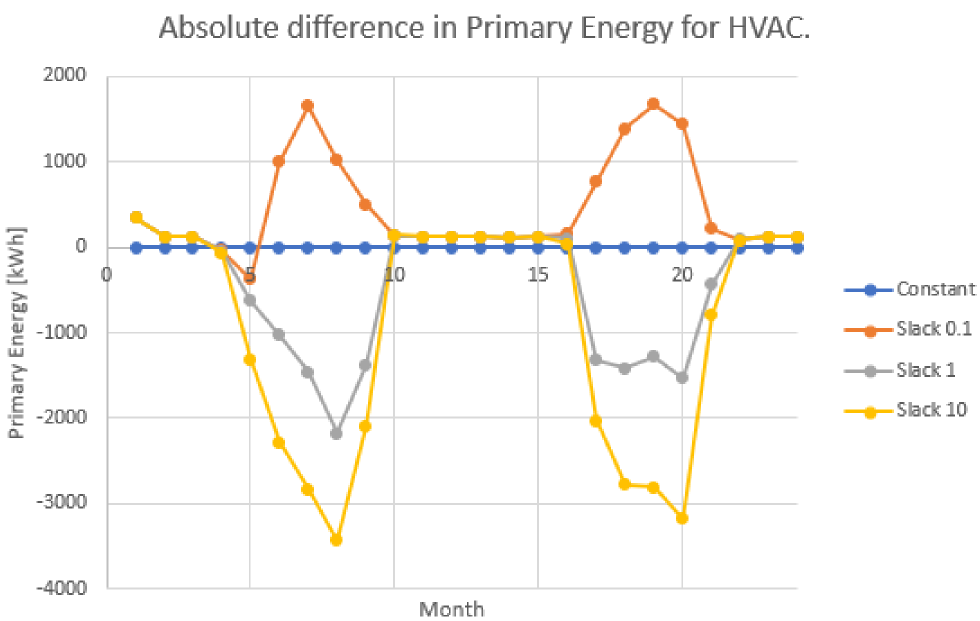



Figure 110: Absolute difference in Primary Energy Consumption for HVAC system, Commercial Retail with outdoor air control.

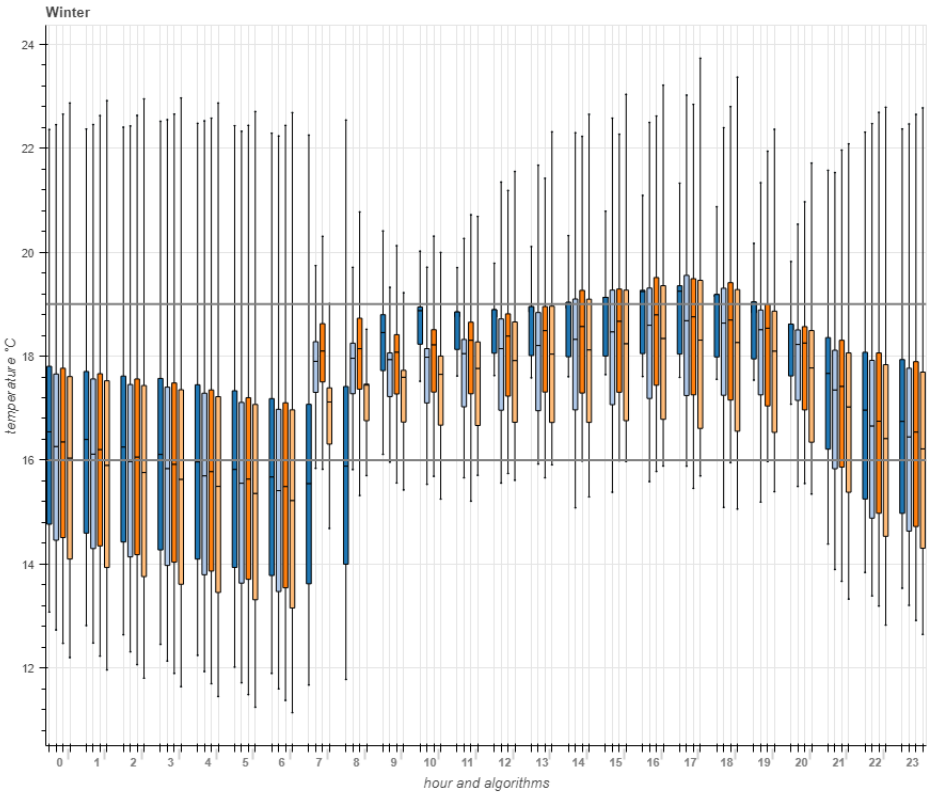



Figure 111: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Commercial Retail with outdoor air control.

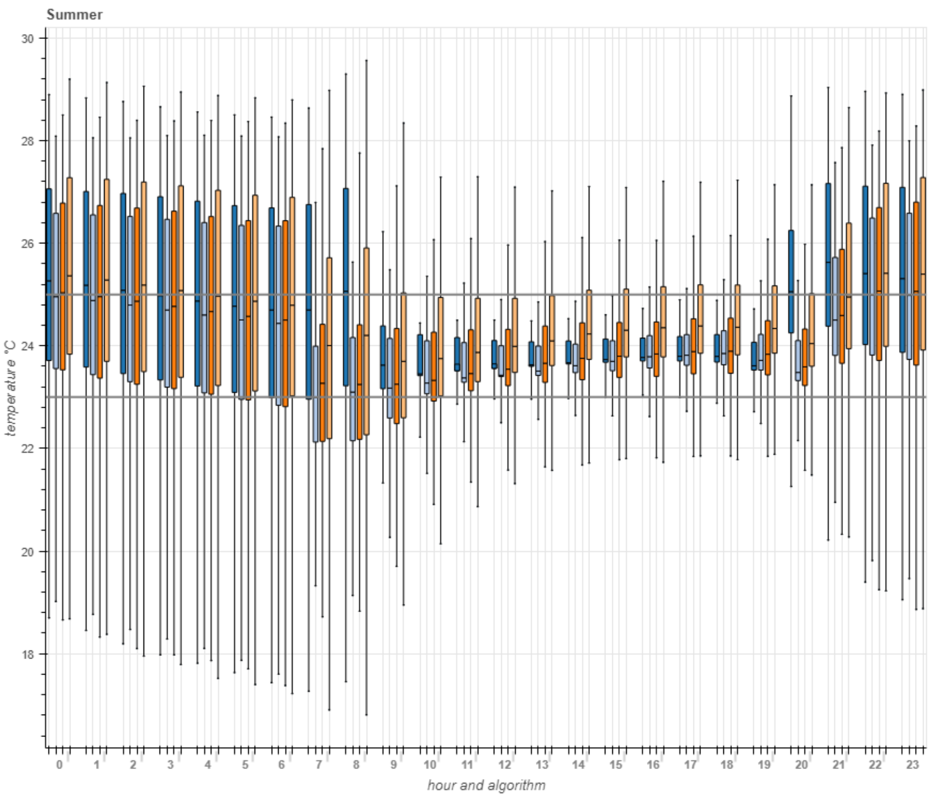



Figure 112: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Commercial Retail with outdoor air control.

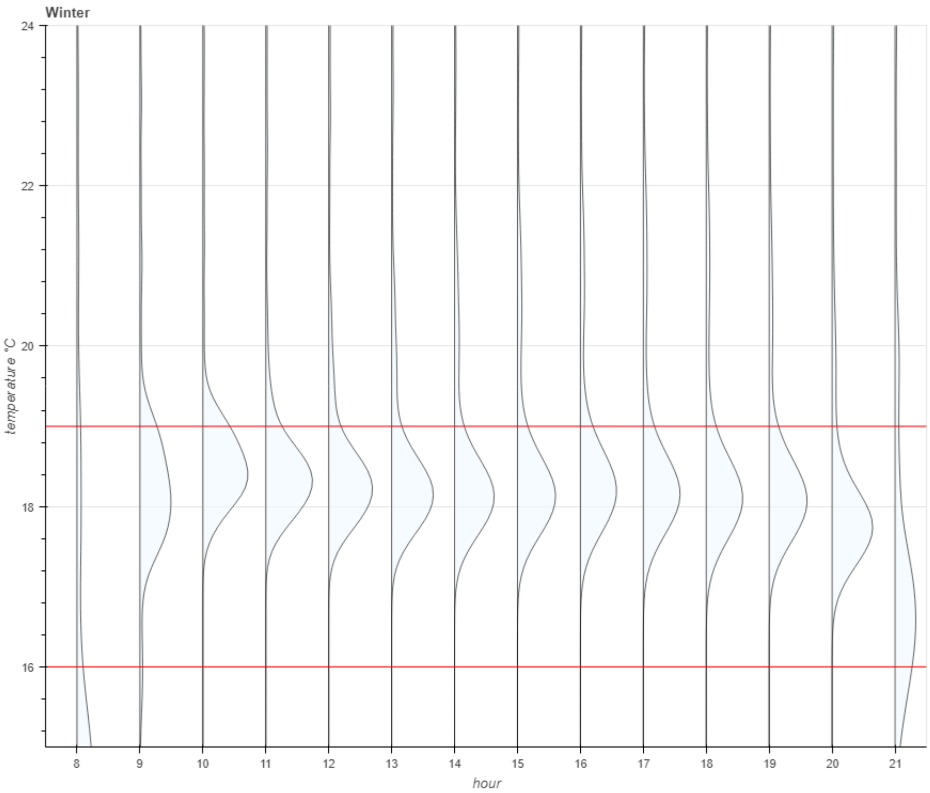



Figure 113: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with outdoor air control.

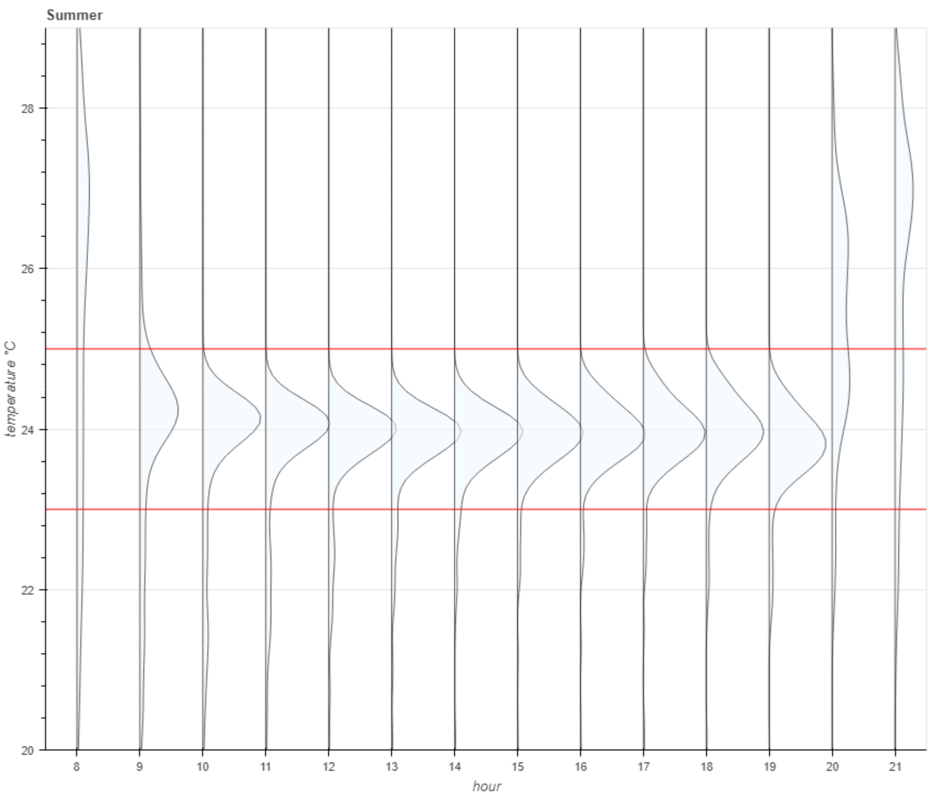



Figure 114: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Commercial Retail with outdoor air control.

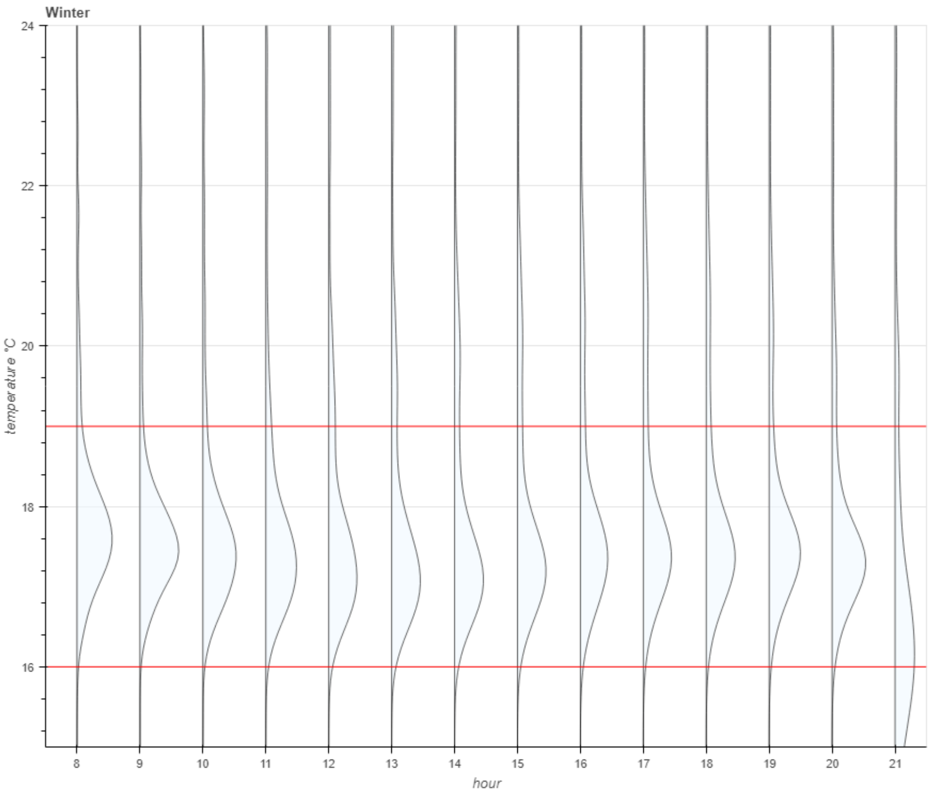



Figure 115: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with outdoor air control.

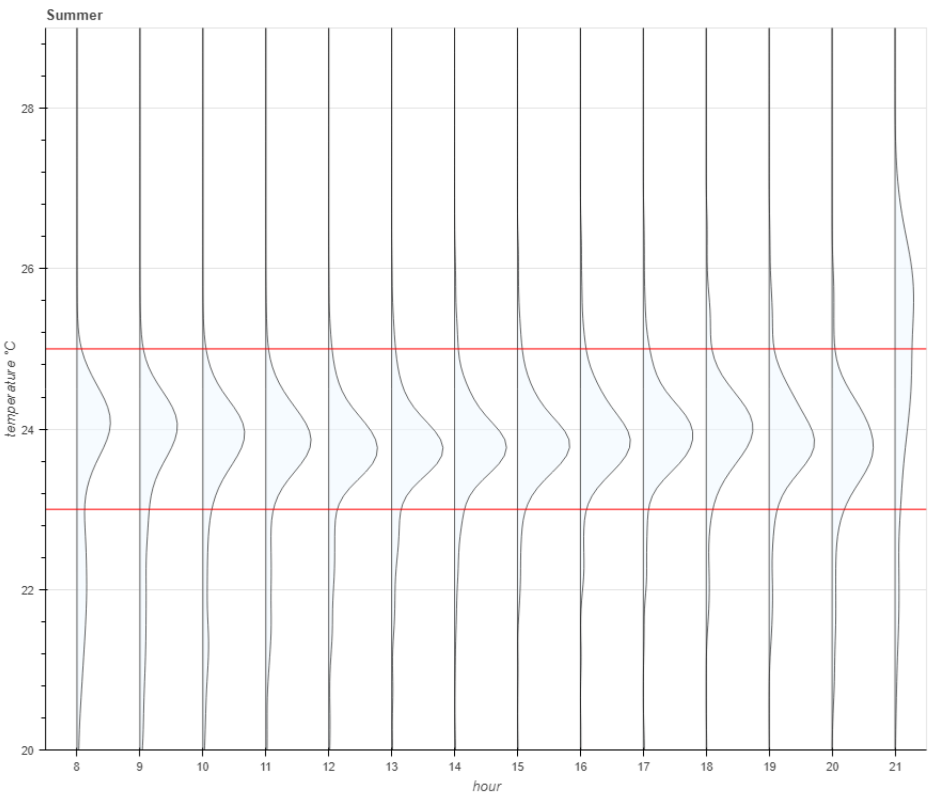



Figure 116: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 0.1, Commercial Retail with outdoor air control.

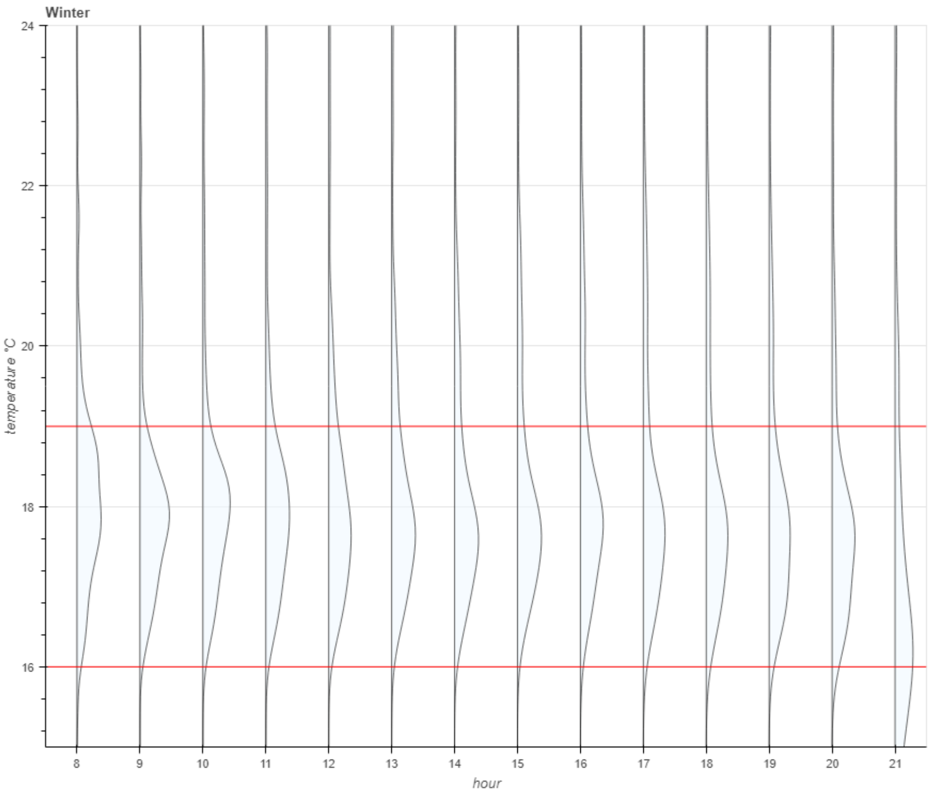



Figure 117: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with outdoor air control.

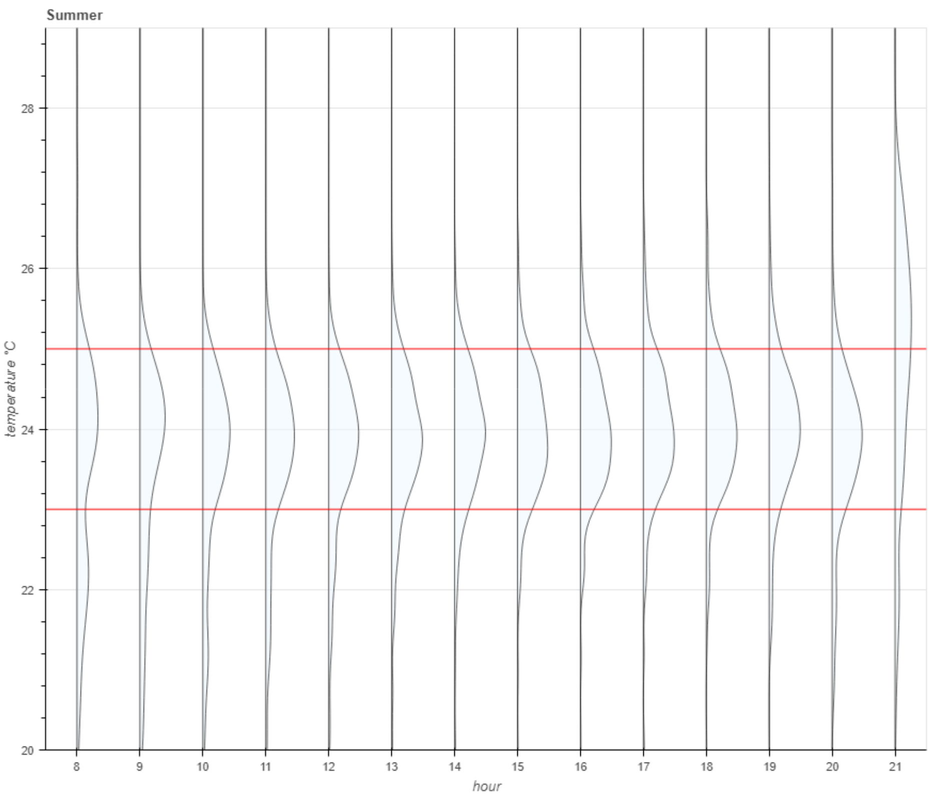



Figure 118: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Commercial Retail with outdoor air control.

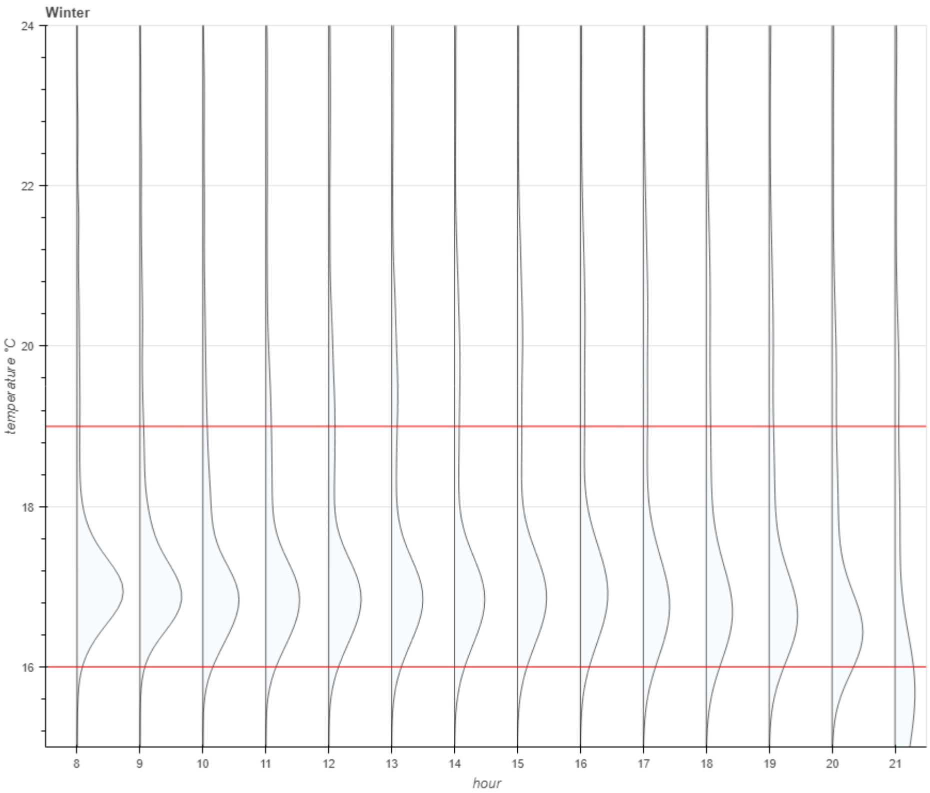



Figure 119: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with outdoor air control.

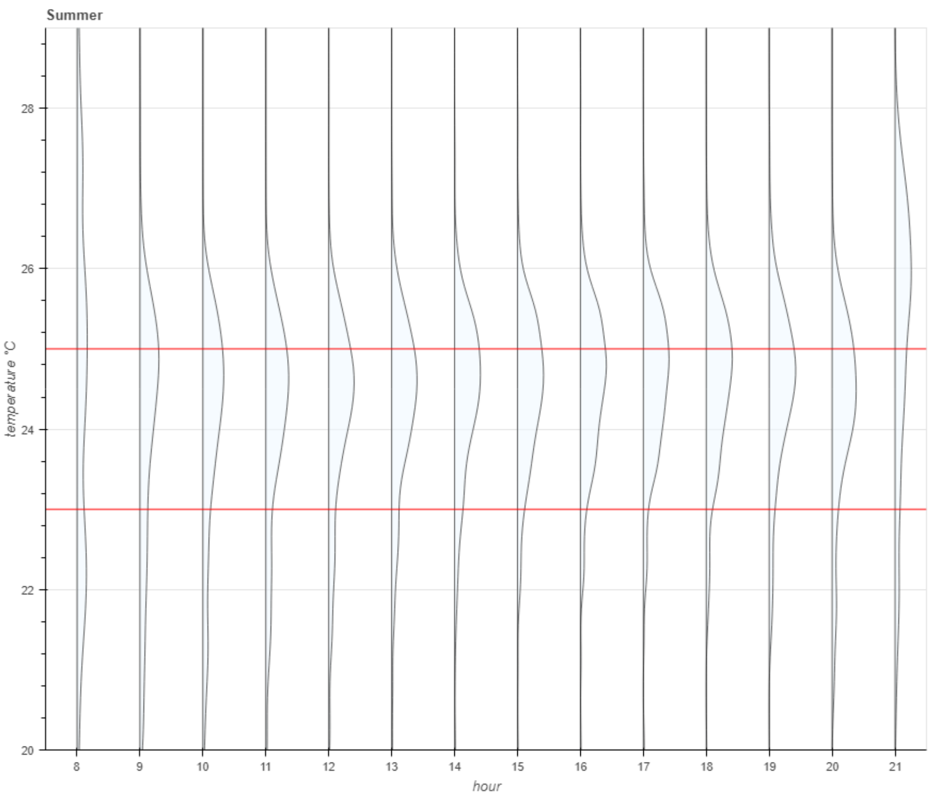



Figure 120: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Commercial Retail with outdoor air control.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant | 1.13% | 2.95% | 27.19% | 23.12% |
| slack_01 | 0.52% | 3.29% | 27.00% | 22.51% |
| slack_1 | 2.15% | 7.17% | 27.22% | 22.77% |
| slack_10 | 3.66% | 12.14% | 25.47% | 21.61% |

Table 57: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Commercial Retail with outdoor air control.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant | 19.30% | 23.76% | 49.88% | 37.84% |
| slack_01 | 24.48% | 30.85% | 3.23% | 1.54% |
| slack_1 | 25.62% | 33.49% | 4.93% | 3.09% |
| slack_10 | 24.28% | 30.34% | 17.88% | 8.22% |

Table 58: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Commercial Retail with outdoor air control.

| | | saving |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant | 0.00% | 0.00% |
| slack_01 | -14.73% | -19.82% |
| slack_1 | 19.78% | -19.82% |
| slack_10 | 37.03% | -19.82% |

Figure 121: Relative savings with respect of constant set-point, Commercial Retail with outdoor air control.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant | 72653.801 | 48.533 |
| Slack 0.1 | 83799.675 | 55.978 |
| Slack 1 | 61761.218 | 41.257 |
| Slack 10 | 50739.010 | 33.894 |

Table 59: Total primary energy consumption for HVAC, Commercial Retail with outdoor air control.

### 5.2 Reference Building: Supermarket.

For the Supermarket model, Reinforcement Learning algorithm has been introduced only in Sales area to reduce computational cost. Comparison has been made with constant set-point with HVAC system on for all the openings hour and constant set-point with HVAC system that shut down 30 minutes before closing that is a condition that allows the internal temperature to stay within the bounds in most days during last opening hour.

Results below shows that is more difficult for the Reinforcement Algorithm to achieve energy saving in comparison to a well sized HVAC system and well tuned control system.

The RL algorithm with two different slack values (1 and 10) were compared with a baseline of constant set-point with HVAC always on when Supermarket is open and constant set-point with HVAC shut down 30 minutes before the closing hour that was seen sufficient to guarantee that air temperature respect air temperature bound at closing hour in most cases.

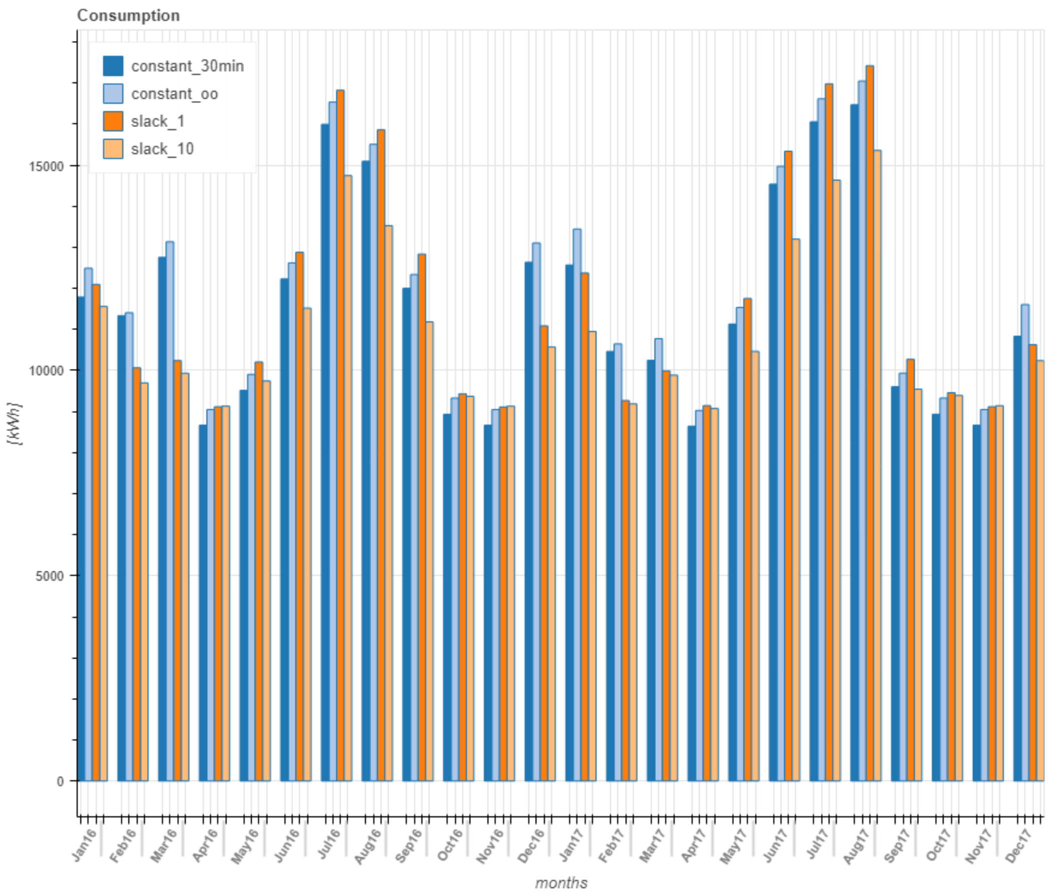



Figure 122: Primary Energy Consumption for HVAC system, Reference Building: Supermarket.

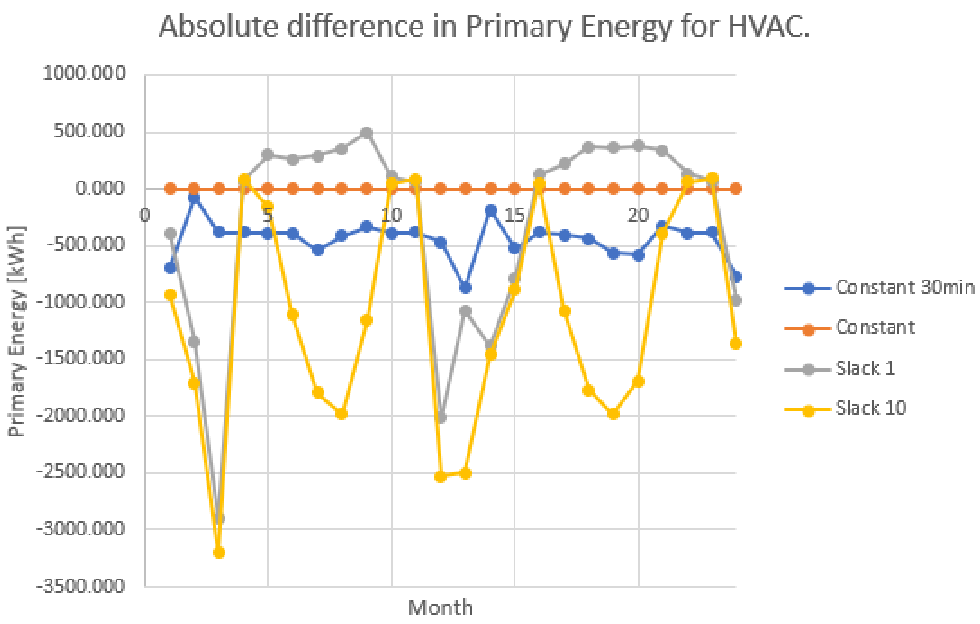



Figure 123: Absolute difference in Primary Energy Consumption for HVAC system, Reference Building: Supermarket.

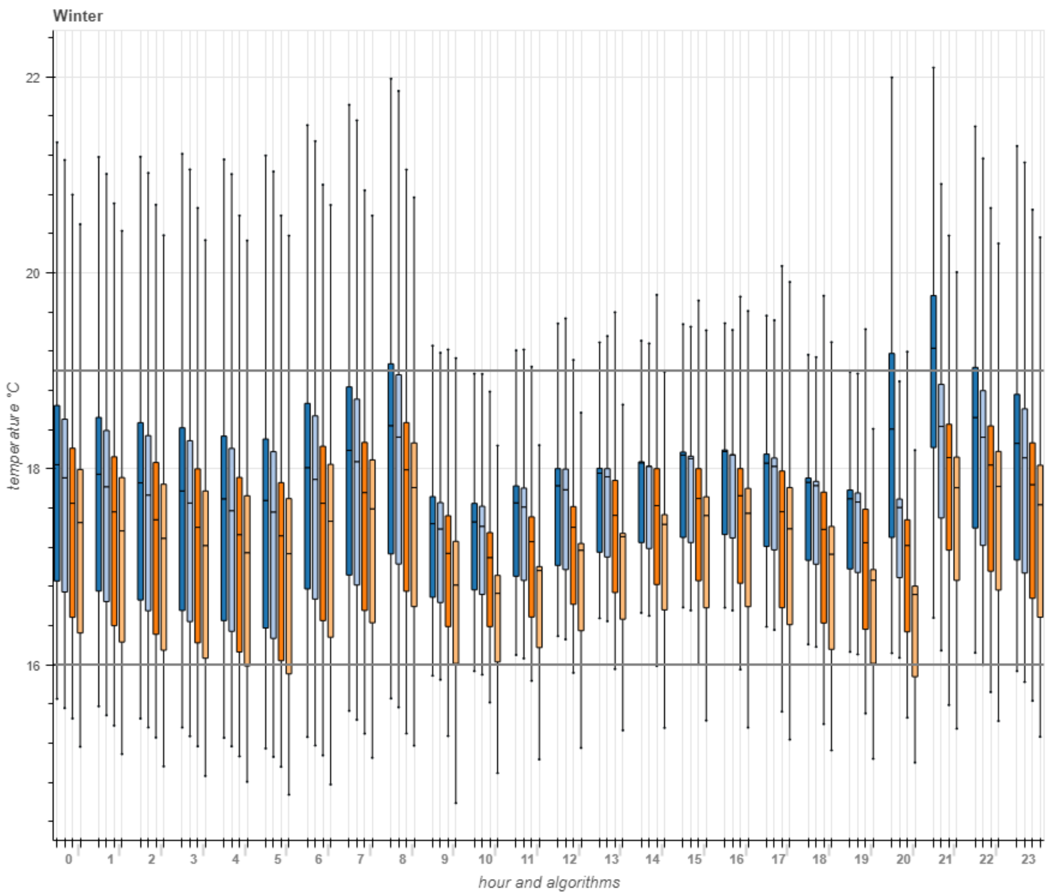



Figure 124: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Reference Building: Supermarket.

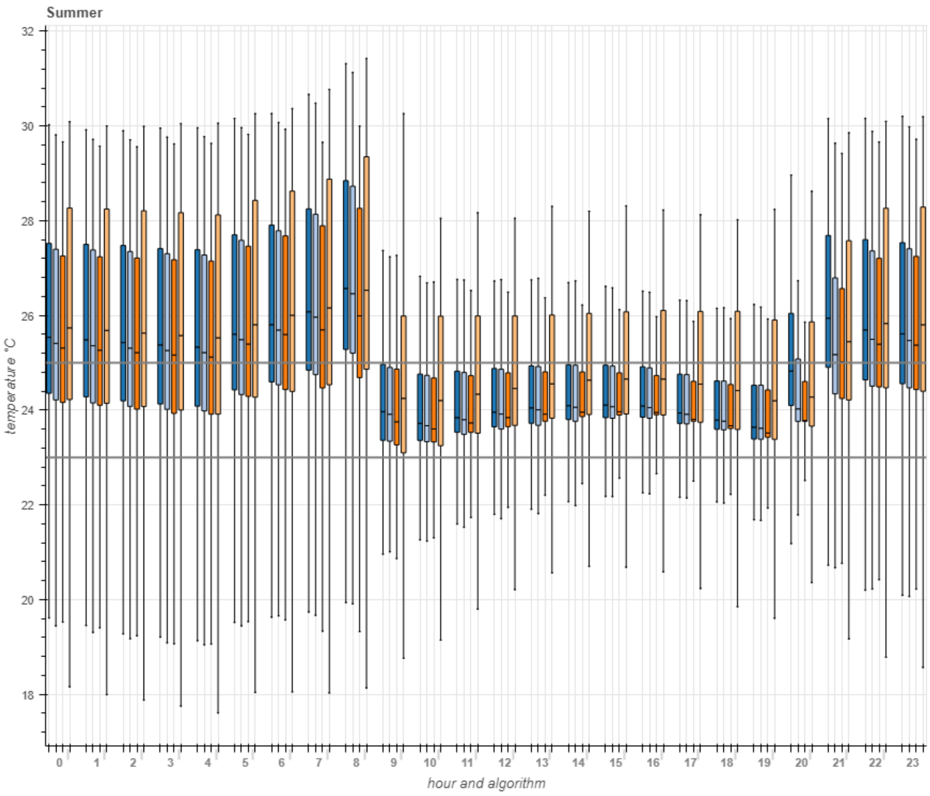



Figure 125: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Reference Building: Supermarket.

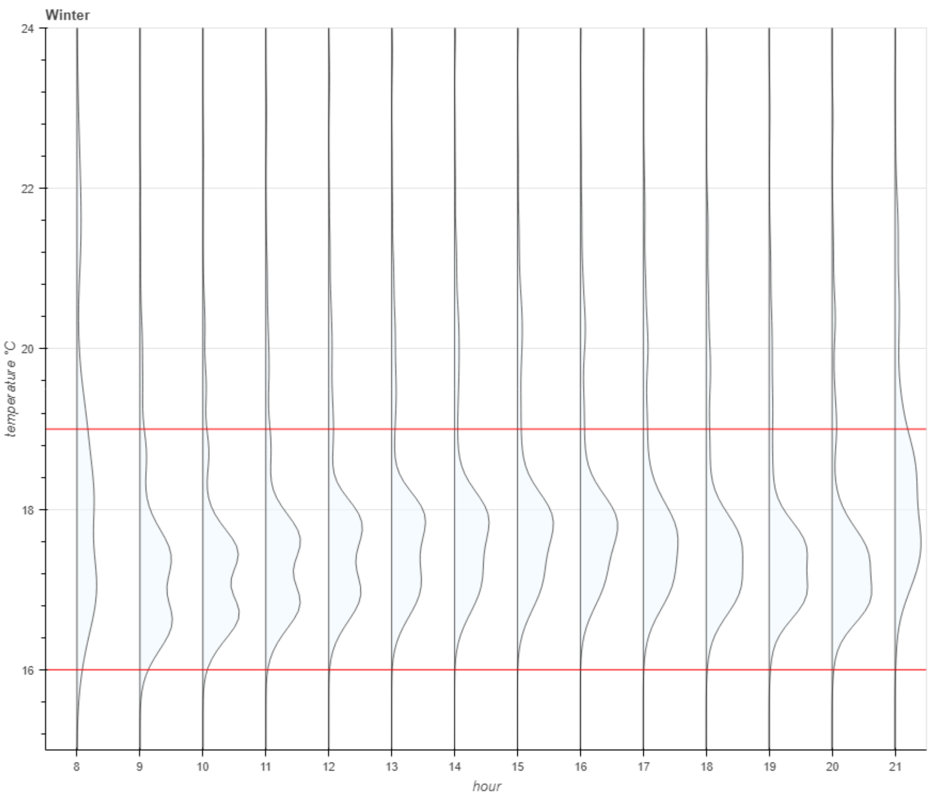



Figure 126: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket.

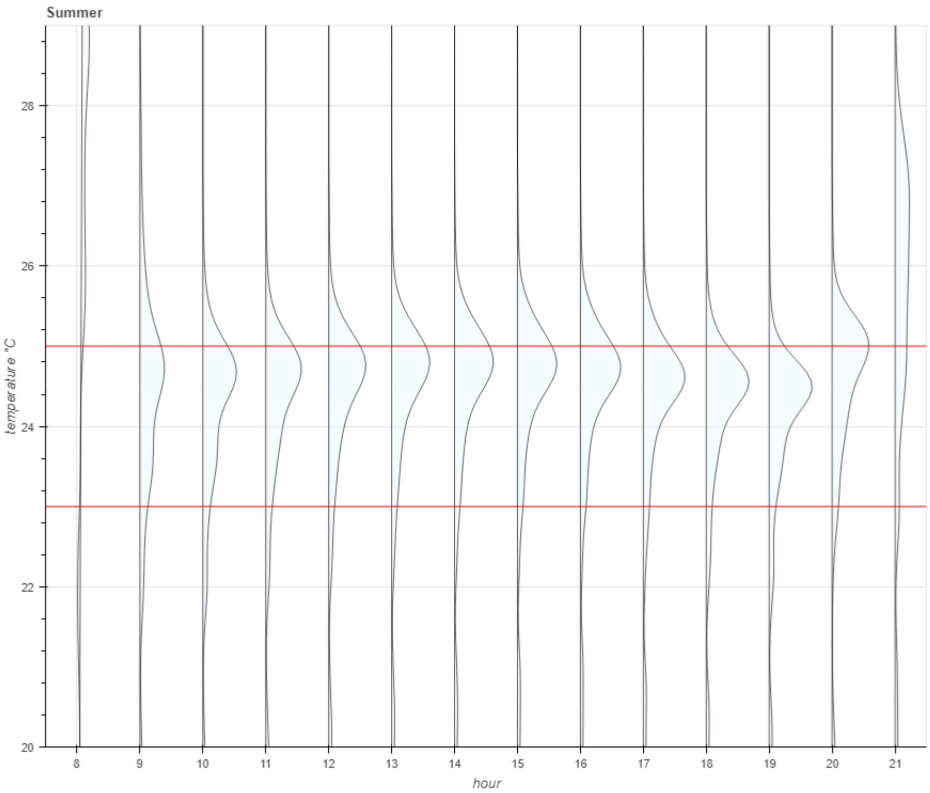



Figure 127: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket.

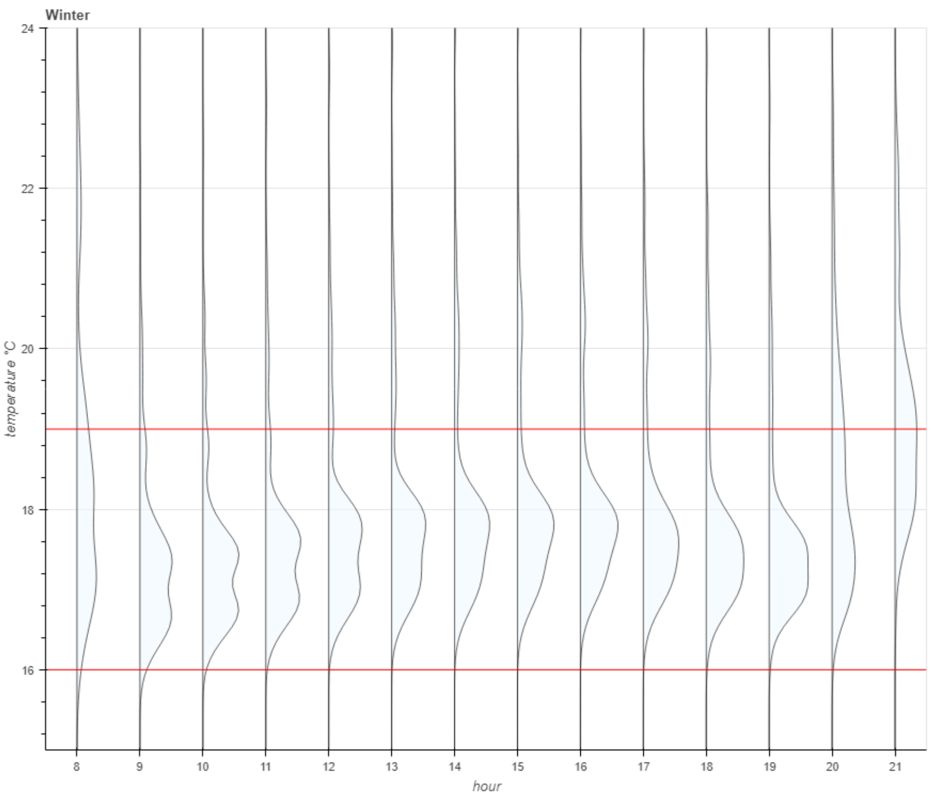



Figure 128: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket.

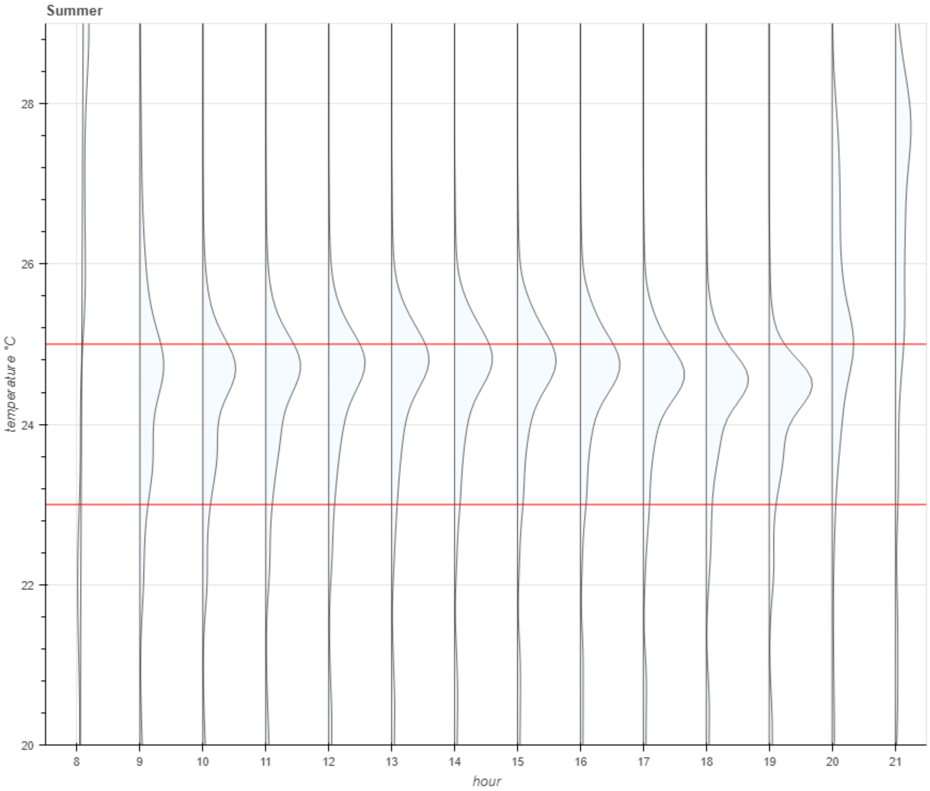



Figure 129: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket.

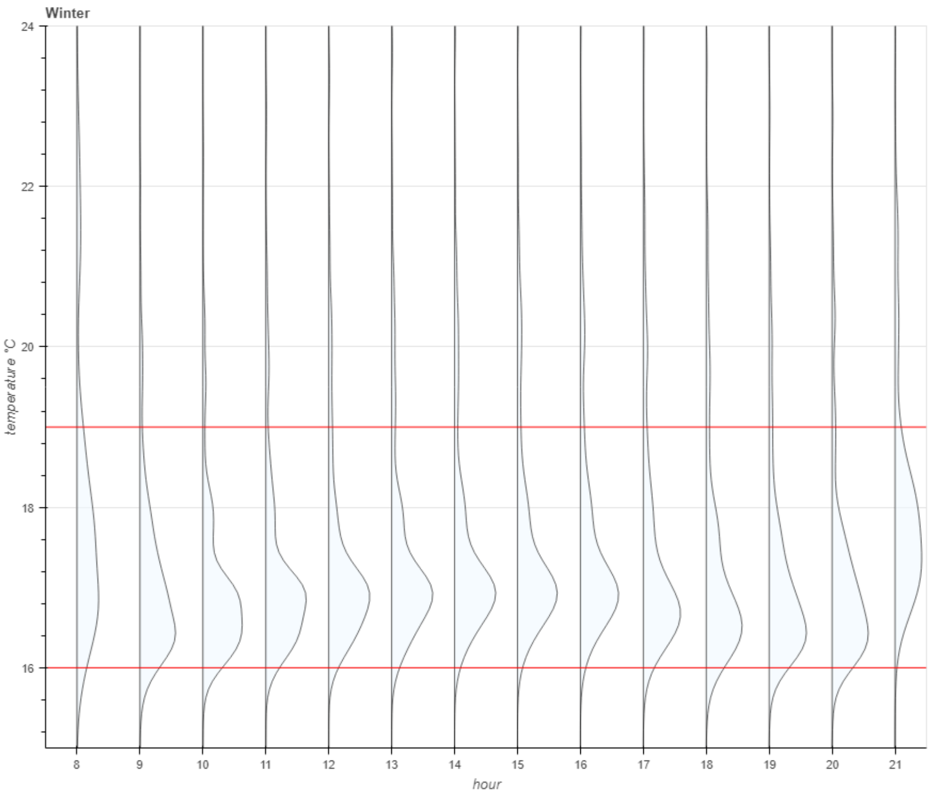



Figure 130: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket.

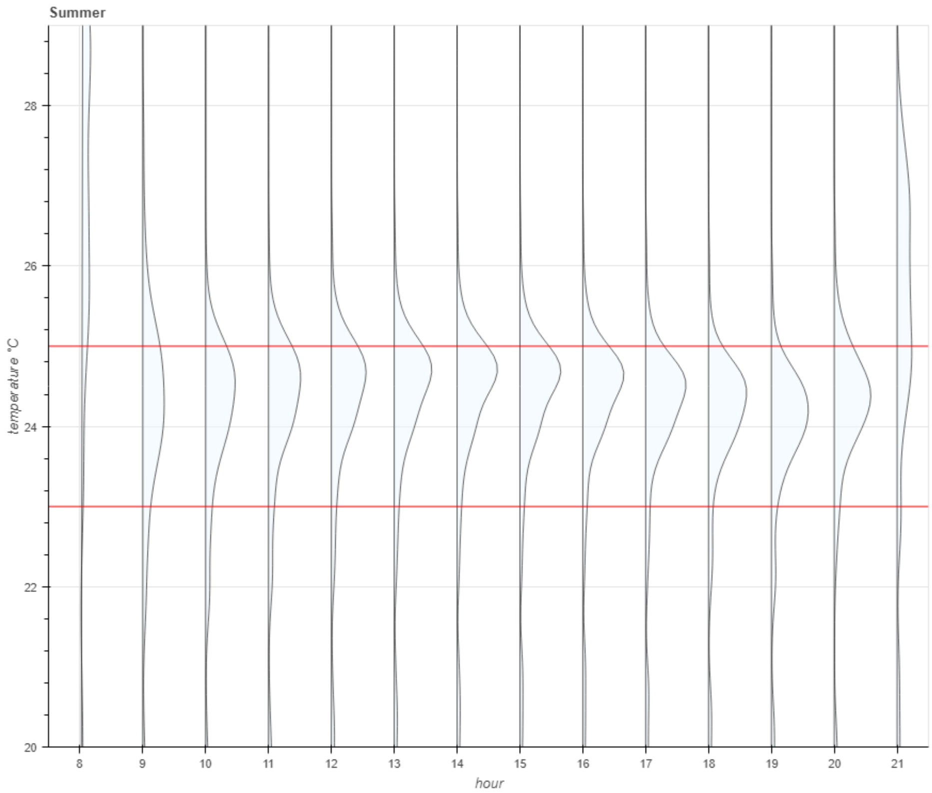



Figure 131: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket.

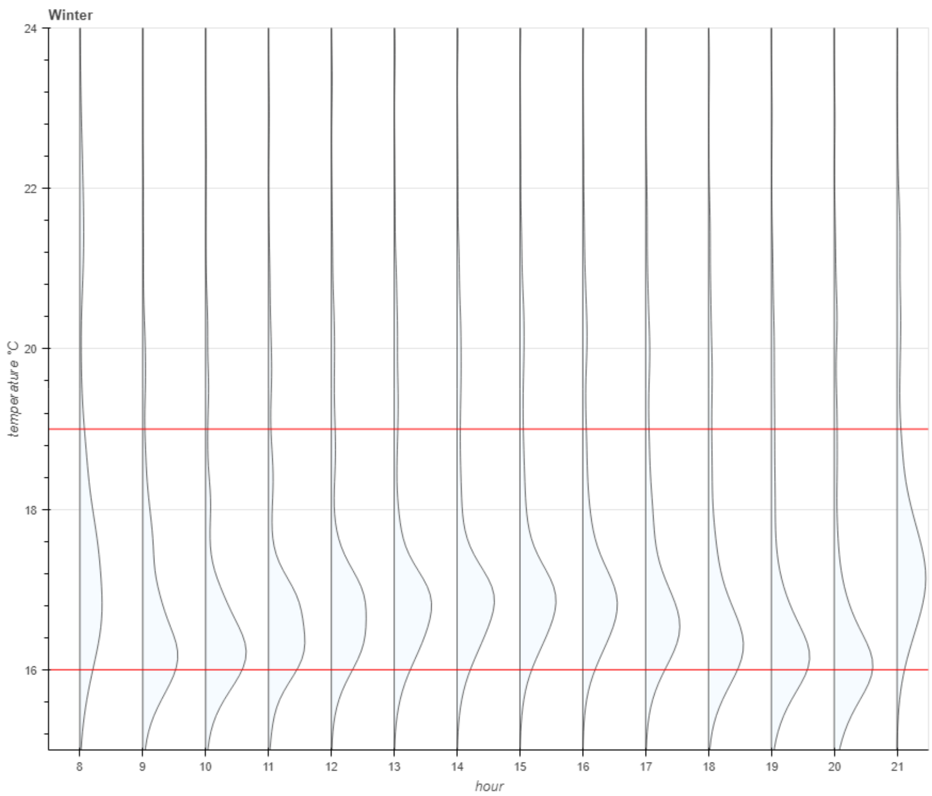



Figure 132: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket.

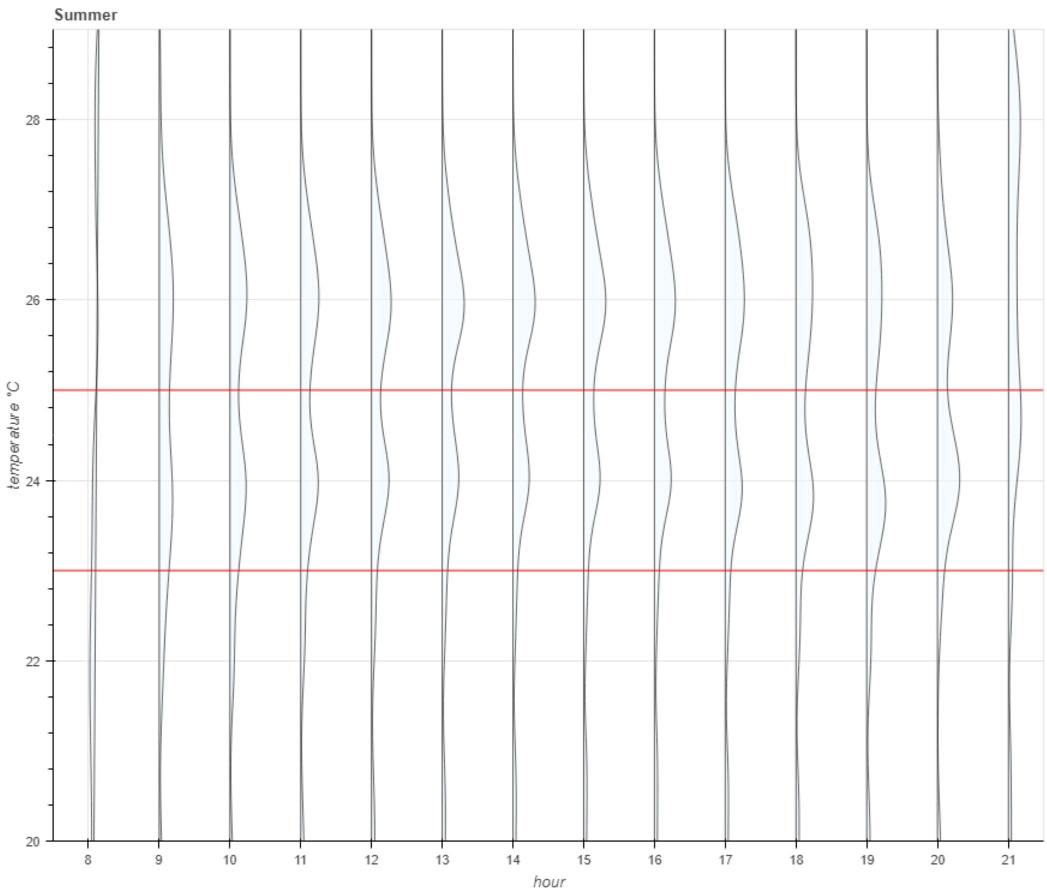



Figure 133: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant_30min | 0.89% | 9.77% | 20.19% | 13.83% |
| constant_oo | 1.17% | 11.57% | 14.07% | 11.35% |
| slack_1 | 2.46% | 15.04% | 13.44% | 10.68% |
| slack_10 | 20.86% | 43.11% | 13.27% | 10.52% |

Table 60: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Reference Building: Supermarket.

| season | below_%_22.5 summer | below_%_23.0 summer | above_%_25.5 summer | above_%_26.0 summer |
|---|---|---|---|---|
| experiment | | | | |
| constant_30min | 18.88% | 23.76% | 35.14% | 25.85% |
| constant_oo | 19.80% | 26.42% | 13.94% | 8.29% |
| slack_1 | 20.40% | 27.64% | 11.01% | 5.85% |
| slack_10 | 21.80% | 30.03% | 42.54% | 28.06% |

Table 61: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Reference Building: Supermarket.

| | saving | |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant_30min | 3.49% | 4.19% |
| constant_oo | 0.00% | 0.00% |
| slack_1 | -2.04% | 3.11% |
| slack_10 | 6.61% | 4.23% |

Figure 134: Relative savings with respect of constant set-point, Reference Building: Supermarket.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant 30 min | 277764.905 | 121.759 |
| Constant | 288442.393 | 126.439 |
| Slack 1 | 281501.962 | 123.397 |
| Slack 10 | 261188.994 | 114.493 |

Table 62: Total primary energy consumption for HVAC, Reference Building: Supermarket.

Then the Reinforcement Learning algorithm has been tested having the possibility to chose a set-point in the temperature bounds with a step of $0.5°C$ instead of 1. This increase complexity of the algorithm adding computational cost and increasing the time for the algorithm to learn by the environment but it may increase energy savings. Results are shown below:

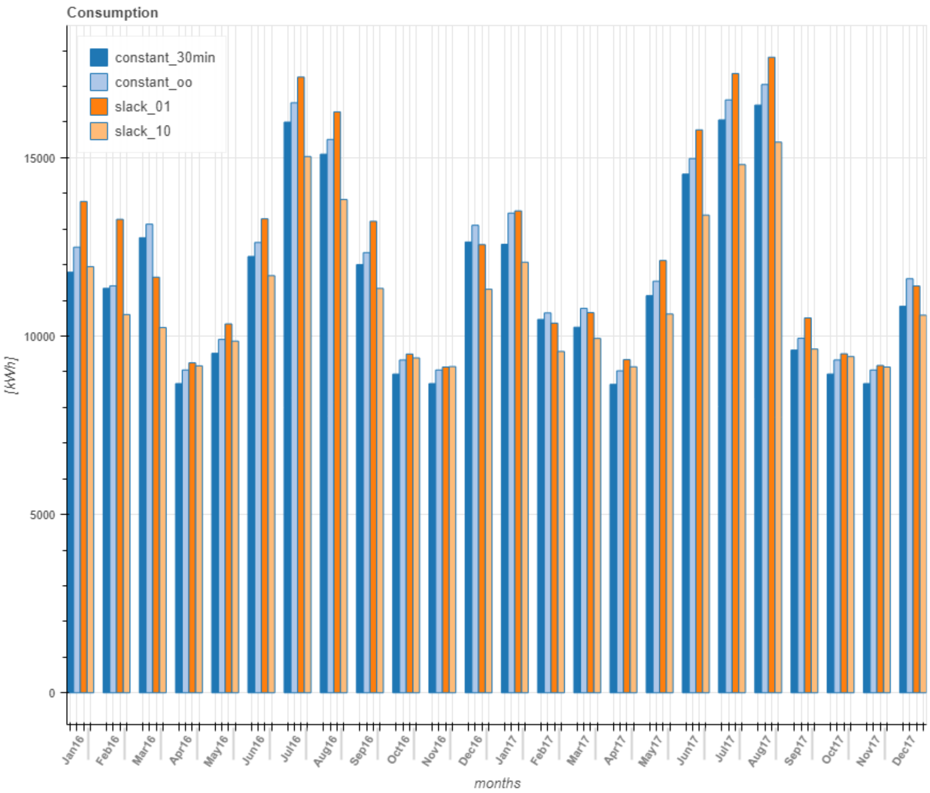



Figure 135: Primary Energy Consumption for HVAC system, Reference Building: Supermarket with $\Delta T = 0.5°C$ in the actions of the agent.

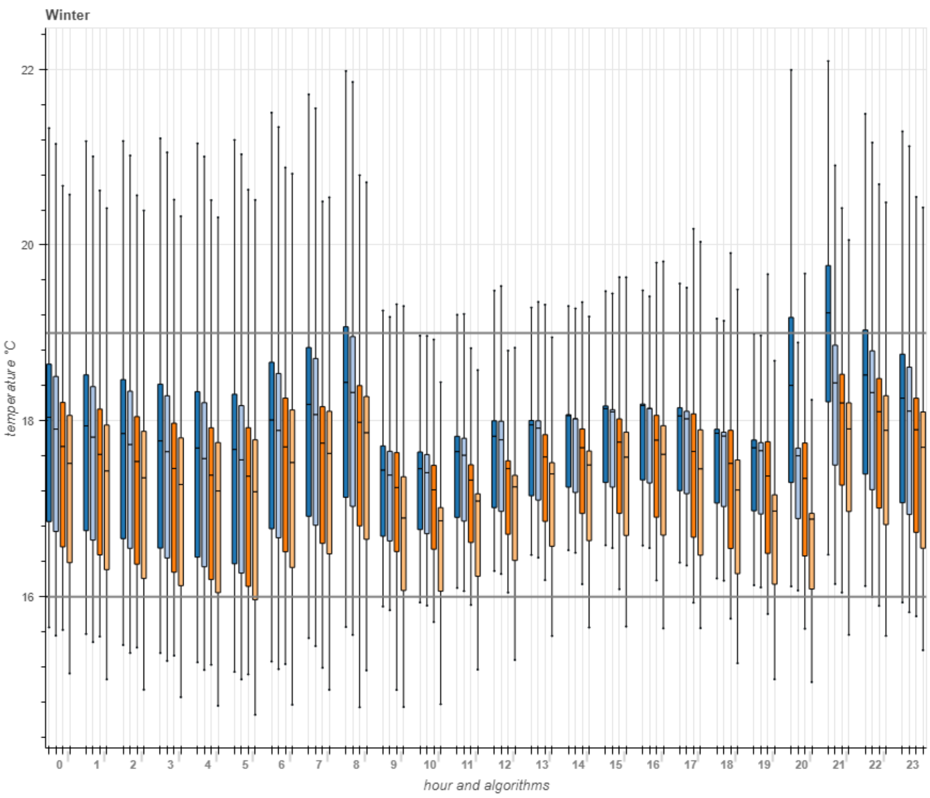



Figure 136: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Reference Building: Supermarket with $\Delta T = 0.5°C$ in the actions of the agent.

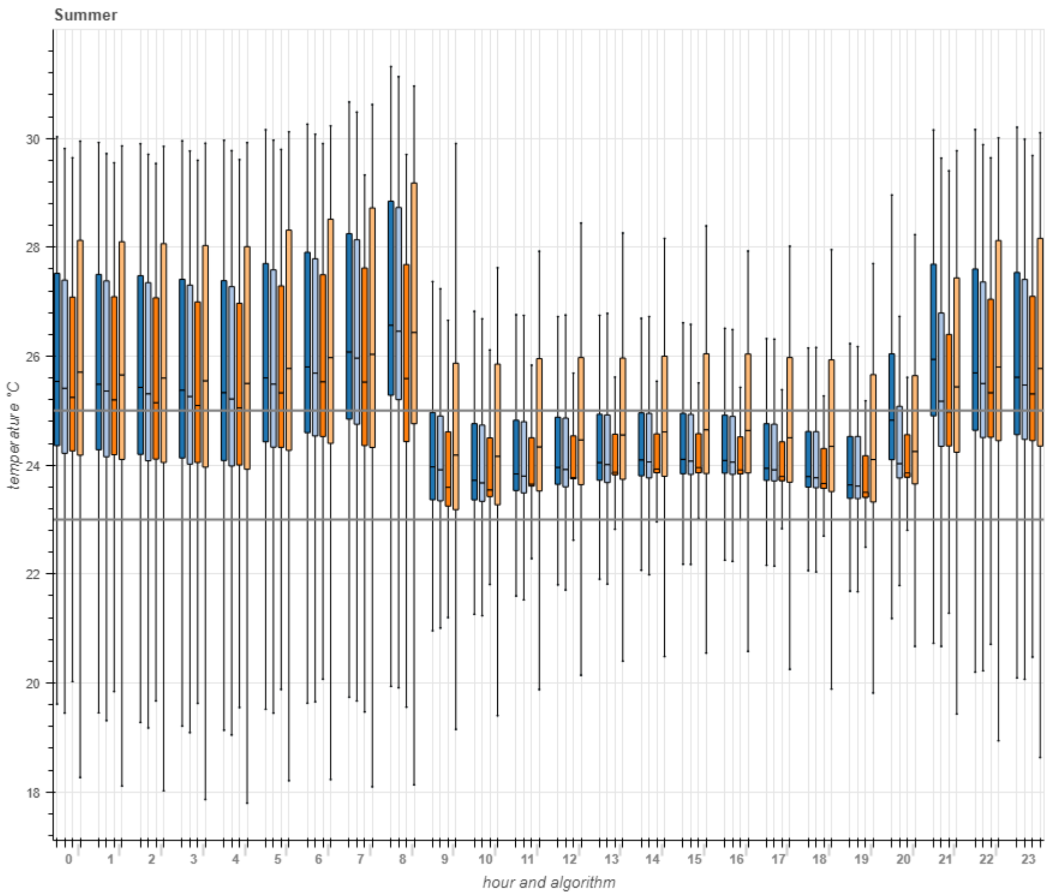



Figure 137: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Reference Building: Supermarket with $\Delta T = 0.5°C$ in the actions of the agent.

| | saving | |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant_30min | 3.49% | 4.19% |
| constant_oo | 0.00% | 0.00% |
| slack_10 | 5.65% | 2.61% |
| slack_5 | 0.44% | 2.06% |

Figure 138: Relative savings with respect of constant set-point, Reference Building: Supermarket with $\Delta T = 0.5°C$ in the actions of the agent.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant 30 min | 277764.905 | 121.758 |
| Constant | 288442.393 | 126.439 |
| Slack 5 | 279004.867 | 122.302 |
| Slack 10 | 267240.625 | 117.146 |

Table 63: Total primary energy consumption for HVAC, Reference Building: Supermarket with $\Delta T = 0.5°C$ in the actions of the agent.

Same simulation has been conducted for the Supermarket in climate zone **B**.

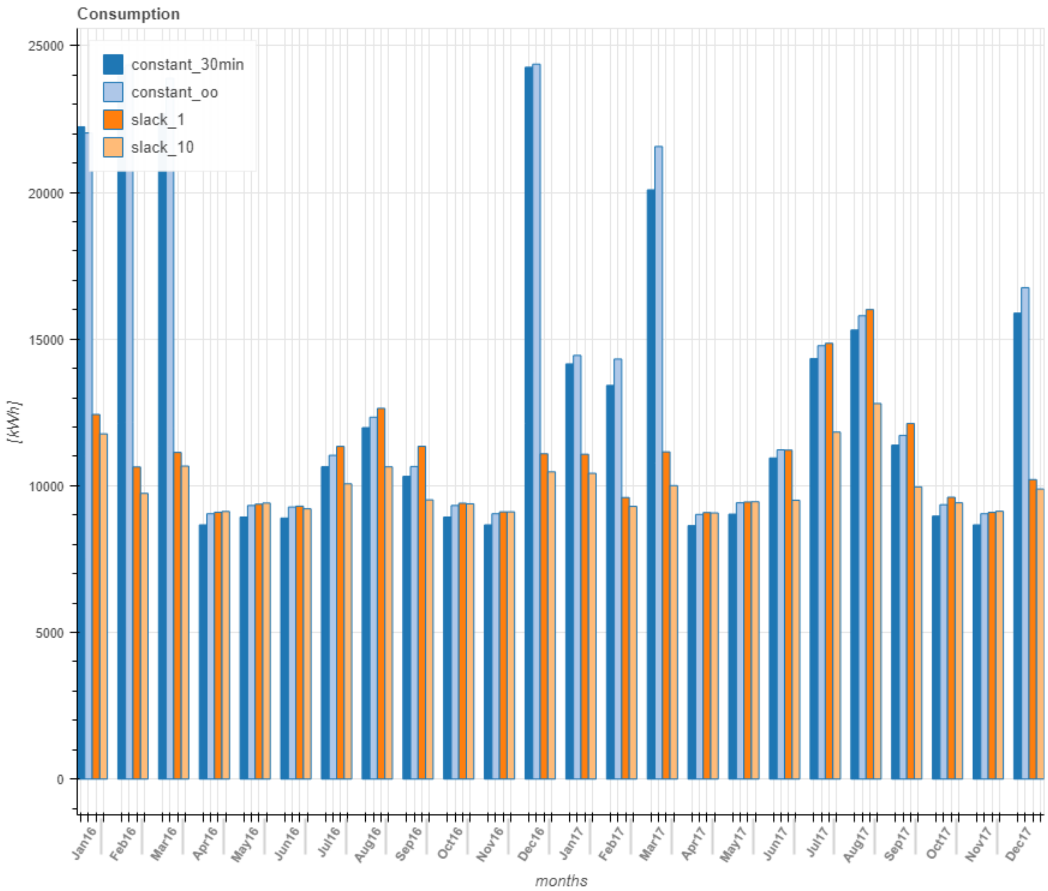



Figure 139: Primary Energy Consumption for HVAC system, Reference Building: Supermarket in climate zone **B**.

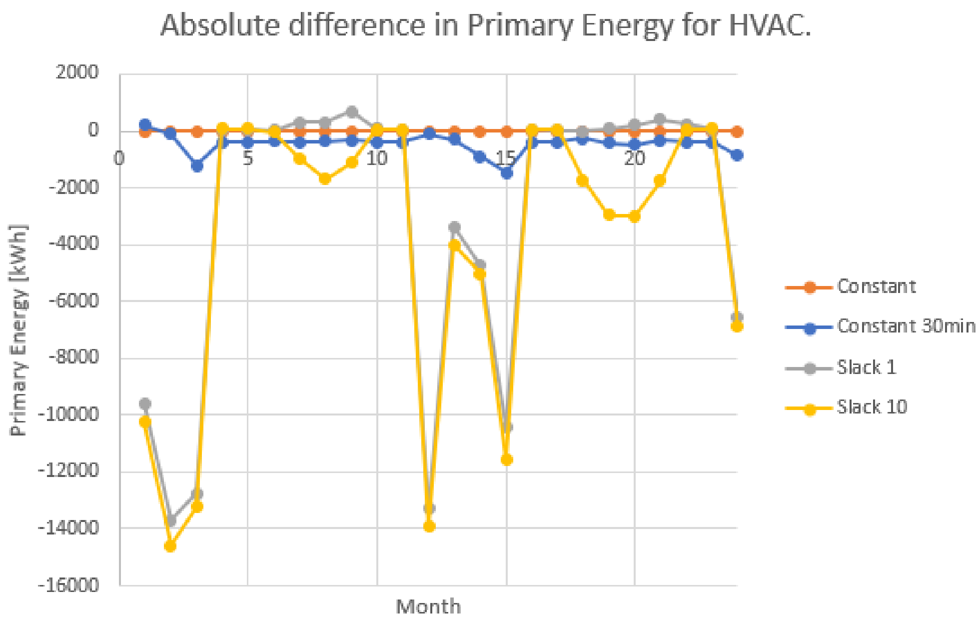



Figure 140: Absolute difference in Primary Energy Consumption for HVAC system, Reference Building: Supermarket in climate zone **B**.

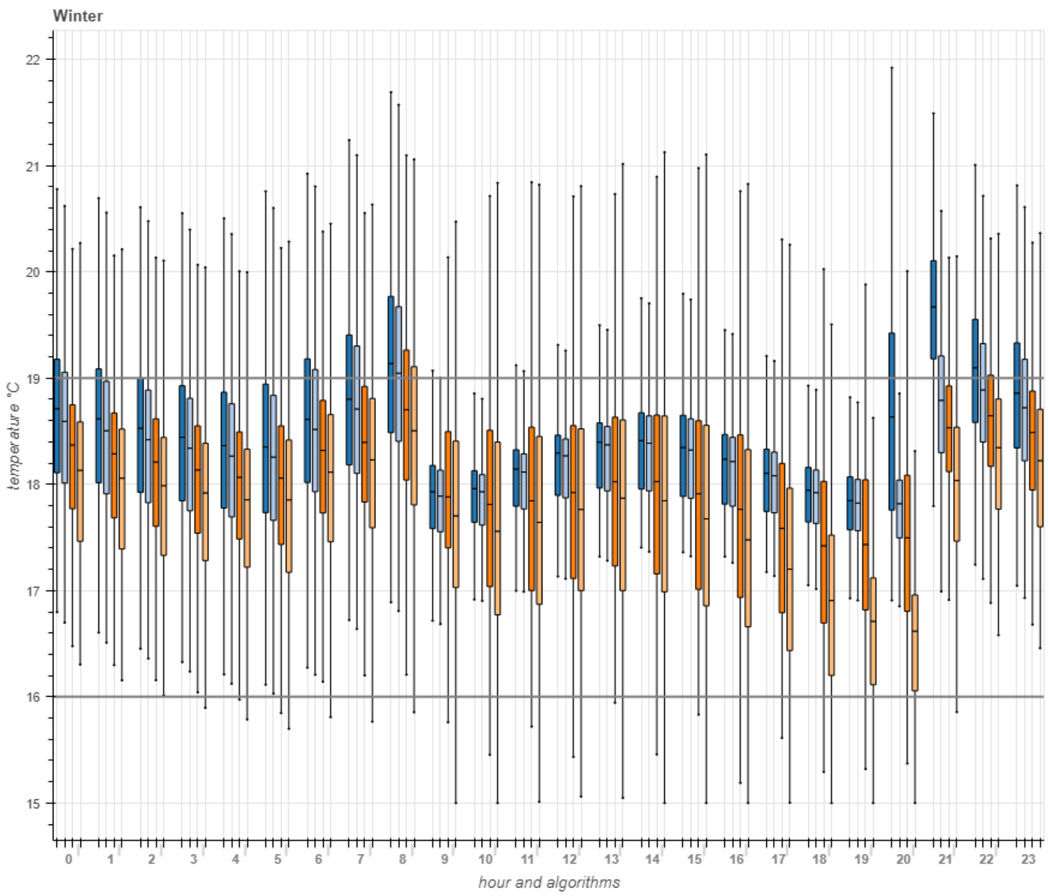



Figure 141: Heating season (1 Dec. to 31 Mar.) hourly mean internal air temperature, Reference Building: Supermarket in climate zone **B**.

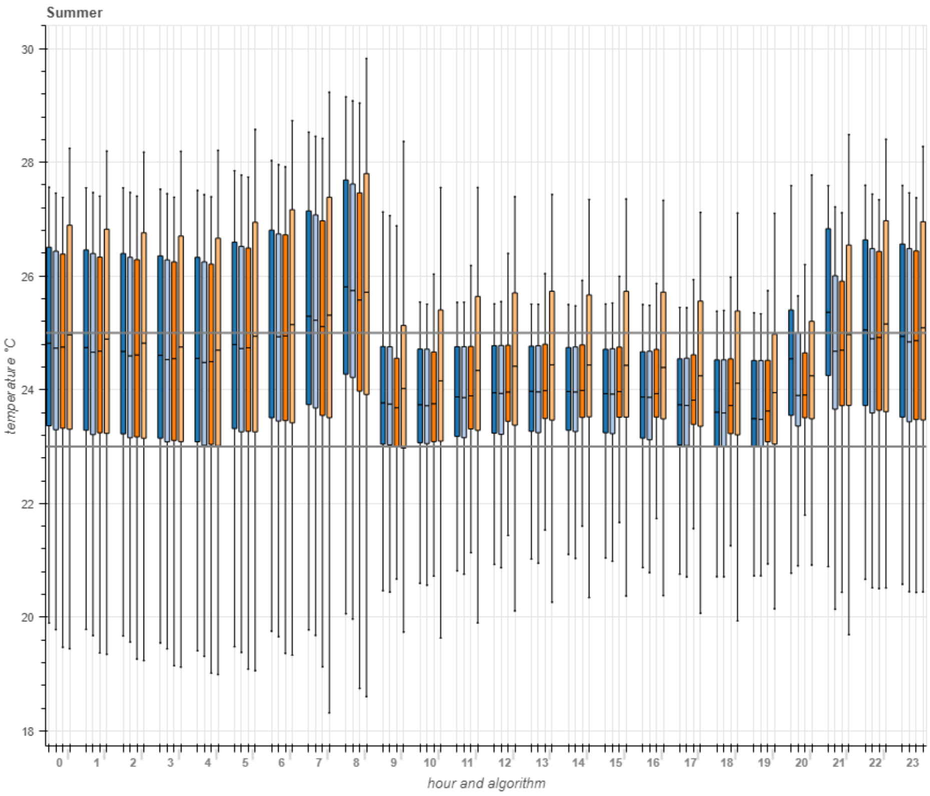



Figure 142: Cooling season (31 Mar. to 1 Dec.) hourly mean internal air temperature, Reference Building: Supermarket in climate zone **B**.

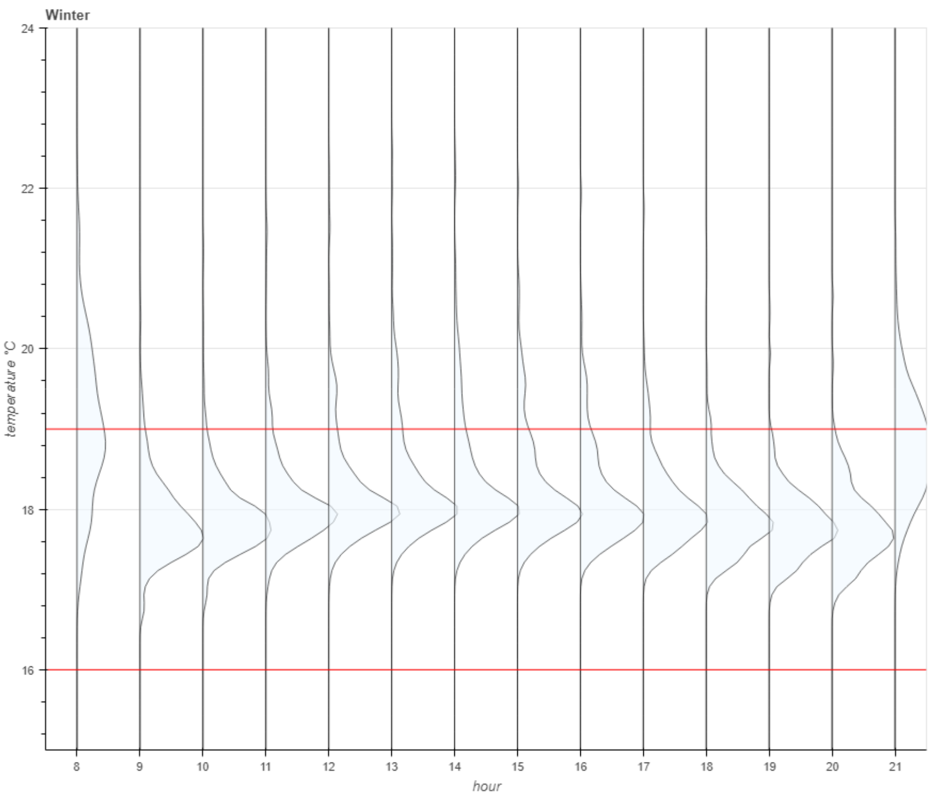



Figure 143: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket in climate zone **B**.

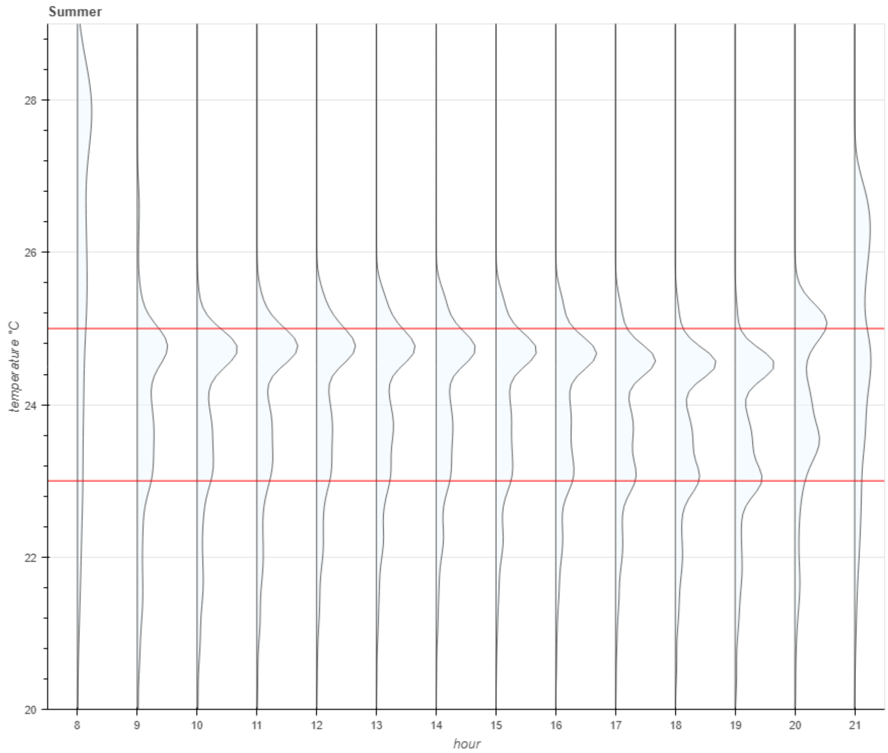



Figure 144: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket in climate zone **B**.

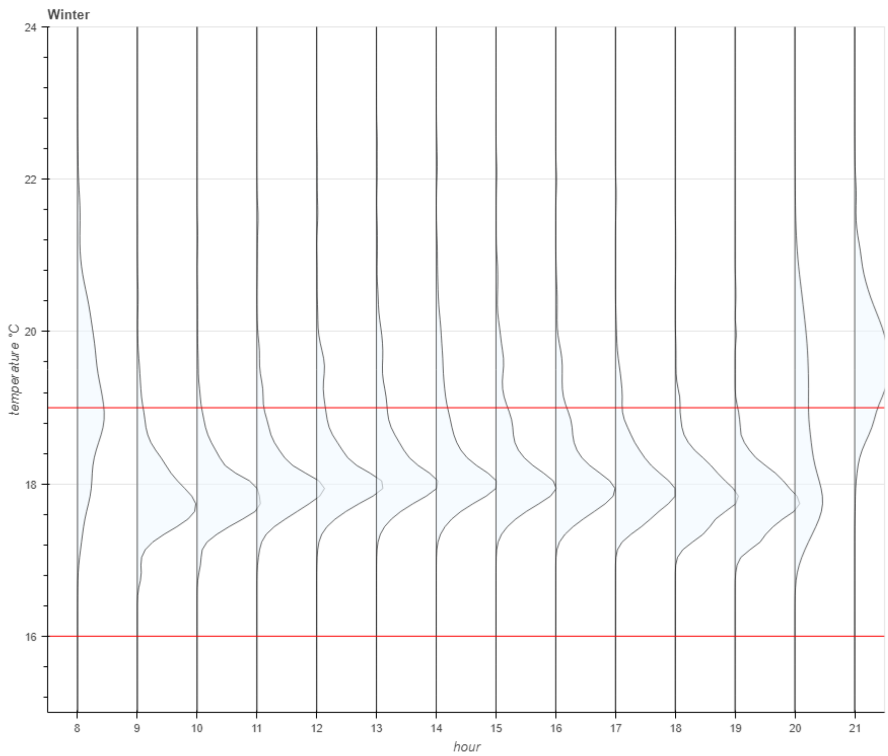



Figure 145: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution for Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket in climate zone **B**.

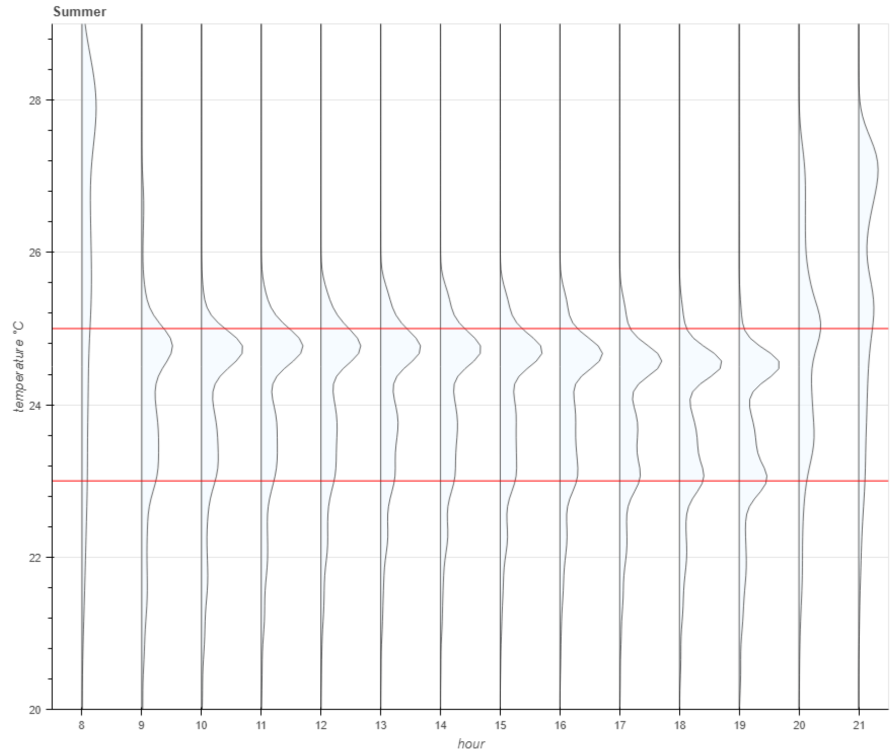



Figure 146: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket in climate zone **B**.

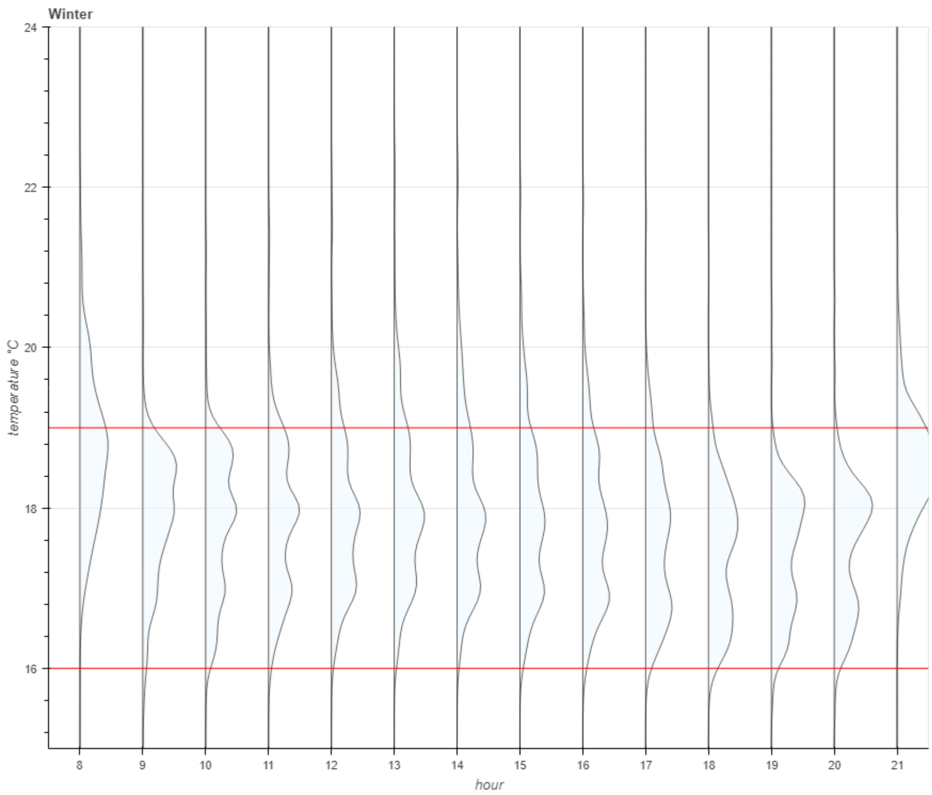



Figure 147: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket in climate zone **B**.

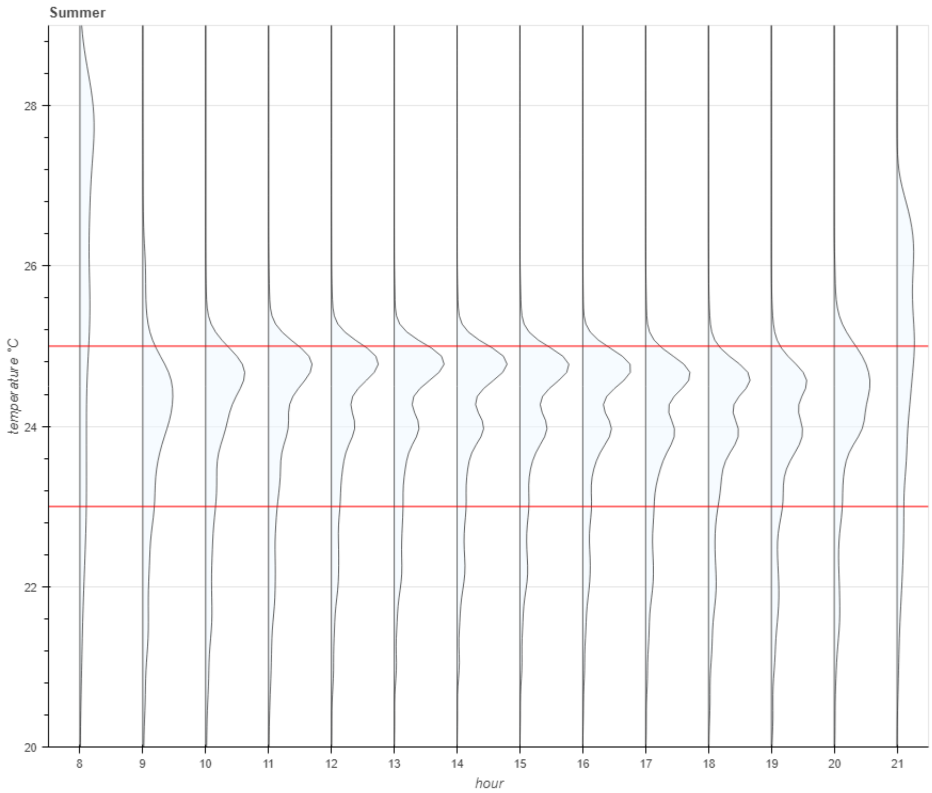



Figure 148: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket in climate zone **B**.

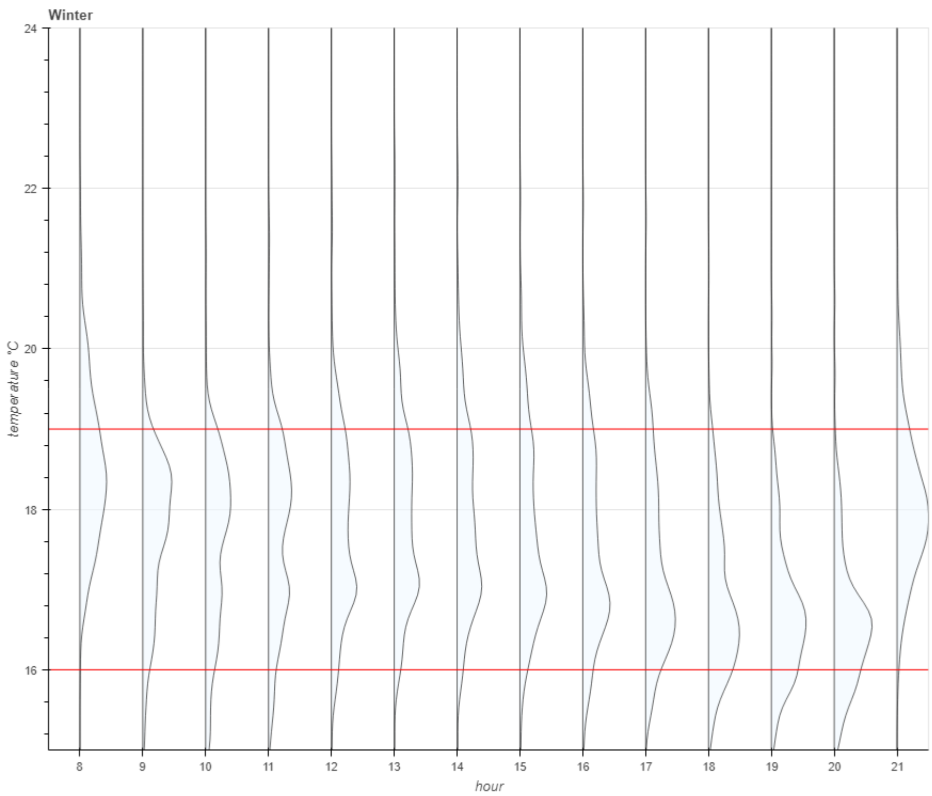



Figure 149: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket in climate zone **B**.

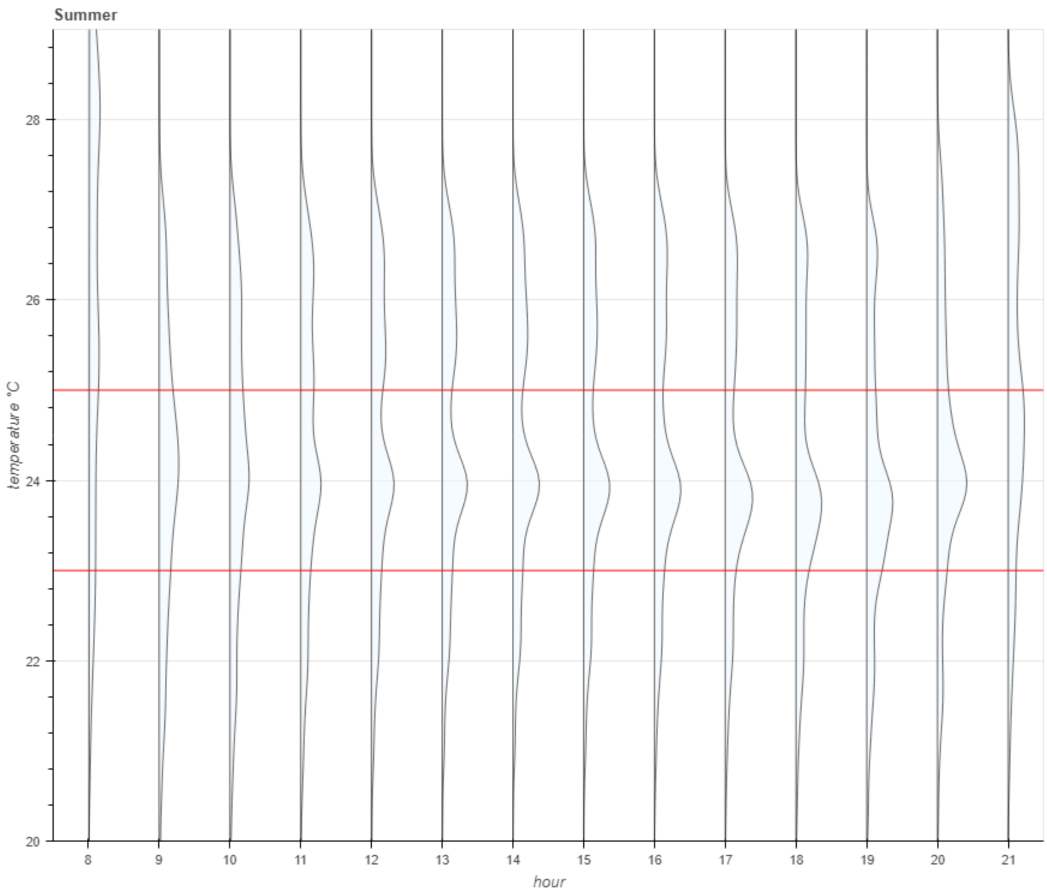



Figure 150: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket in climate zone **B**.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| **season** | **winter** | **winter** | **winter** | **winter** |
| **experiment** | | | | |
| **constant_30min** | 0.00% | 0.00% | 23.71% | 13.23% |
| **constant_oo** | 0.00% | 0.00% | 10.00% | 5.22% |
| **slack_1** | 0.76% | 5.73% | 9.49% | 4.88% |
| **slack_10** | 7.72% | 25.20% | 8.93% | 4.79% |

Table 64: Percentage of exceeding air temperature in Heating Season (1 Dec. to 31 Mar.), Reference Building: Supermarket in climate zone **B**.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| season | summer | summer | summer | summer |
| experiment | | | | |
| constant_30min | 24.11% | 35.34% | 23.12% | 15.28% |
| constant_oo | 26.56% | 38.12% | 4.08% | 2.67% |
| slack_1 | 25.86% | 34.14% | 4.02% | 1.68% |
| slack_10 | 26.49% | 35.85% | 29.23% | 19.65% |

Table 65: Percentage of exceeding air temperature in Cooling Season (31 Mar. to 1 Dec.), Reference Building: Supermarket in climate zone **B**.

| | | saving |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant_30min | 3.57% | 3.66% |
| constant_oo | 0.00% | 0.00% |
| slack_1 | -1.55% | 19.76% |
| slack_10 | 7.49% | 21.10% |

Figure 151: Relative savings with respect of constant set-point, Reference Building: Supermarket in climate zone **B**.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant 30 min | 321277.382 | 140.833 |
| Constant | 332109.408 | 145.581 |
| Slack 1 | 260382.887 | 114.139 |
| Slack 10 | 239920.764 | 105.170 |

Table 66: Total primary energy consumption for HVAC, Reference Building: Supermarket in climate zone **B**.

Then the random occupation profile has been implemented as in the Commercial Retail model [30] with the base occupation profile taken by the DOE model reported in Table 67

| Hour | Weekdays | Sat. | Sun. |
|---|---|---|---|
| 9 | 0.5 | 0.5 | 0.1 |
| 10 | 0.5 | 0.6 | 0.2 |
| 11 | 0.5 | 0.8 | 0.2 |
| 12 | 0.7 | 0.8 | 0.4 |
| 13 | 0.7 | 0.8 | 0.4 |
| 14 | 0.7 | 0.8 | 0.4 |
| 15 | 0.8 | 0.8 | 0.4 |
| 16 | 0.7 | 0.8 | 0.4 |
| 17 | 0.5 | 0.6 | 0.2 |
| 18 | 0.5 | 0.2 | 0.1 |
| 19 | 0.3 | 0.2 | 0.1 |
| 20 | 0.3 | 0.2 | 0.1 |

Table 67: Occupancy profile for Random Occupation Commercial Retail Model.

For climate zone **B** results are shown in the following figures:

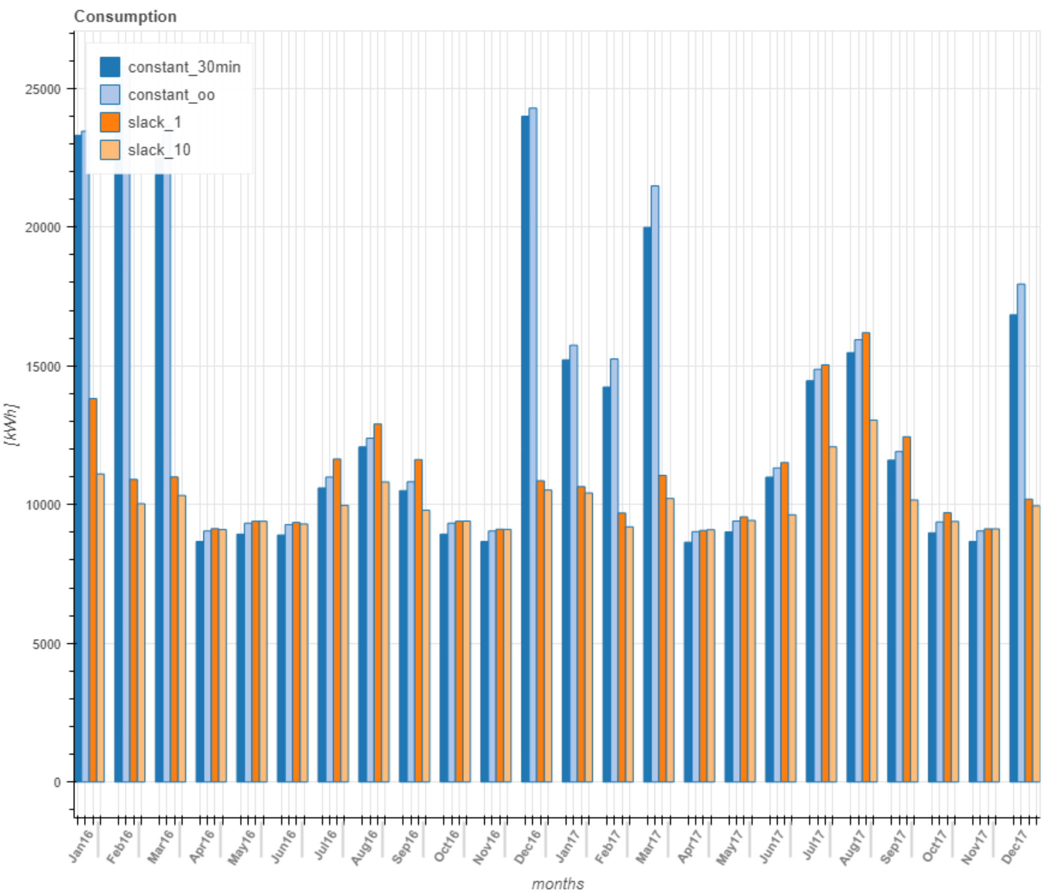



Figure 152: Primary Energy Consumption for HVAC system, Reference Building: Supermarket in climate zone **B** with random occupation model.

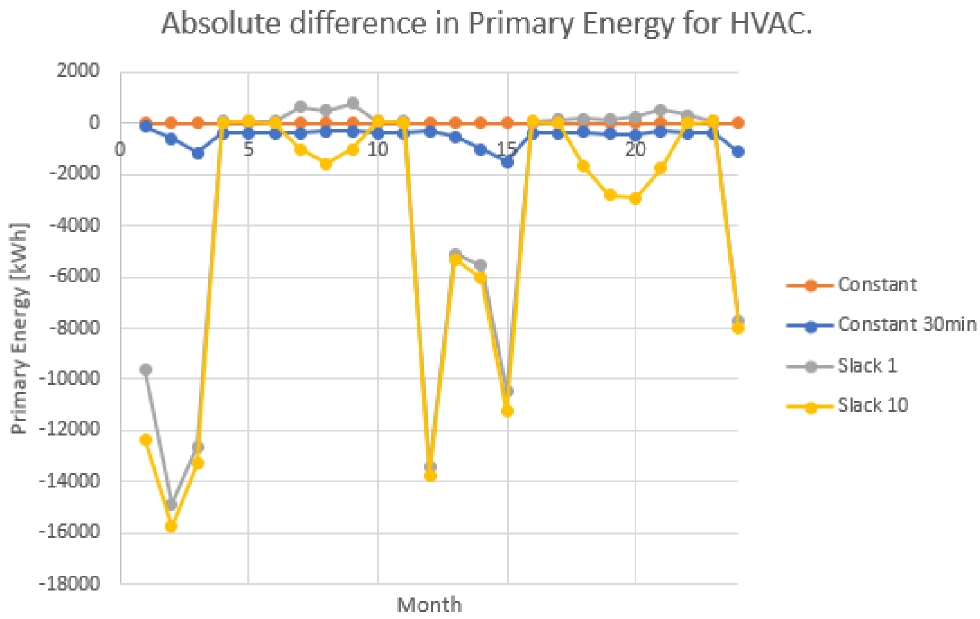



Figure 153: Absolute difference in Primary Energy Consumption for HVAC system, Reference Building: Supermarket in climate zone **B** with random occupation model.

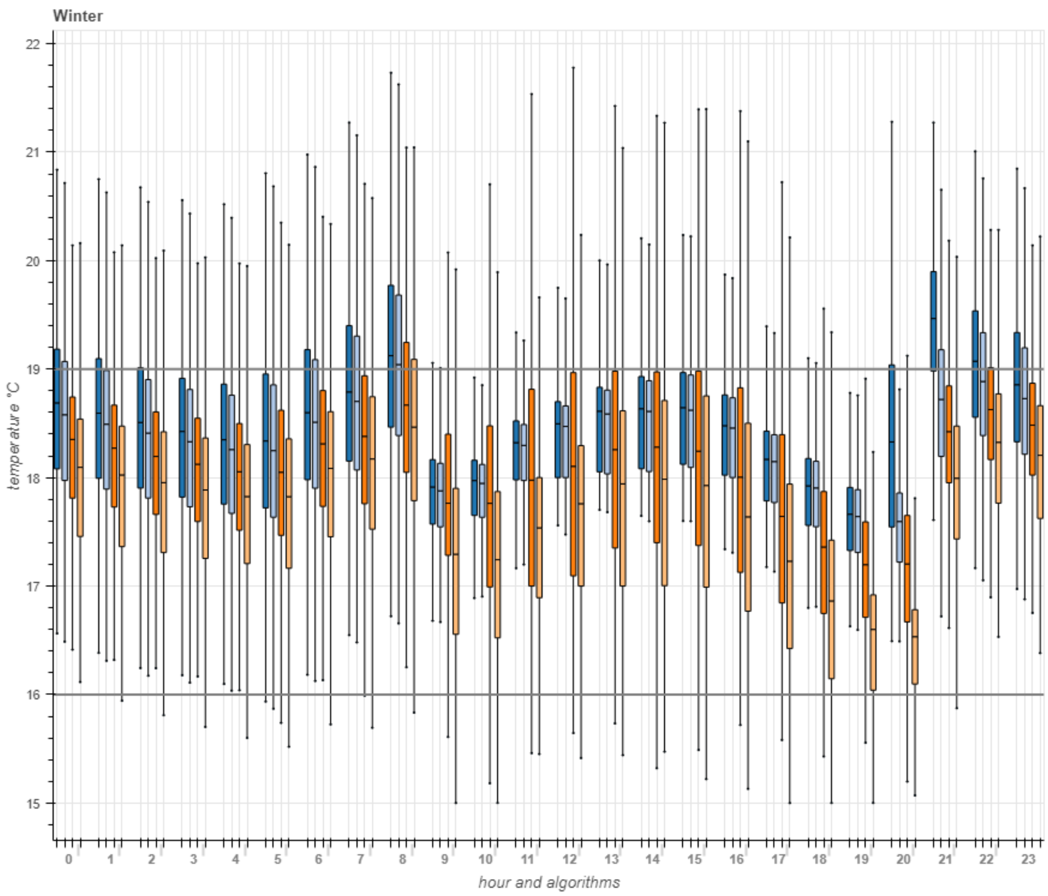



Figure 154: Heating season (1 Dec. to 31 Mar.) hourly mean internal air temperature, Reference Building: Supermarket in climate zone **B** with random occupation model.

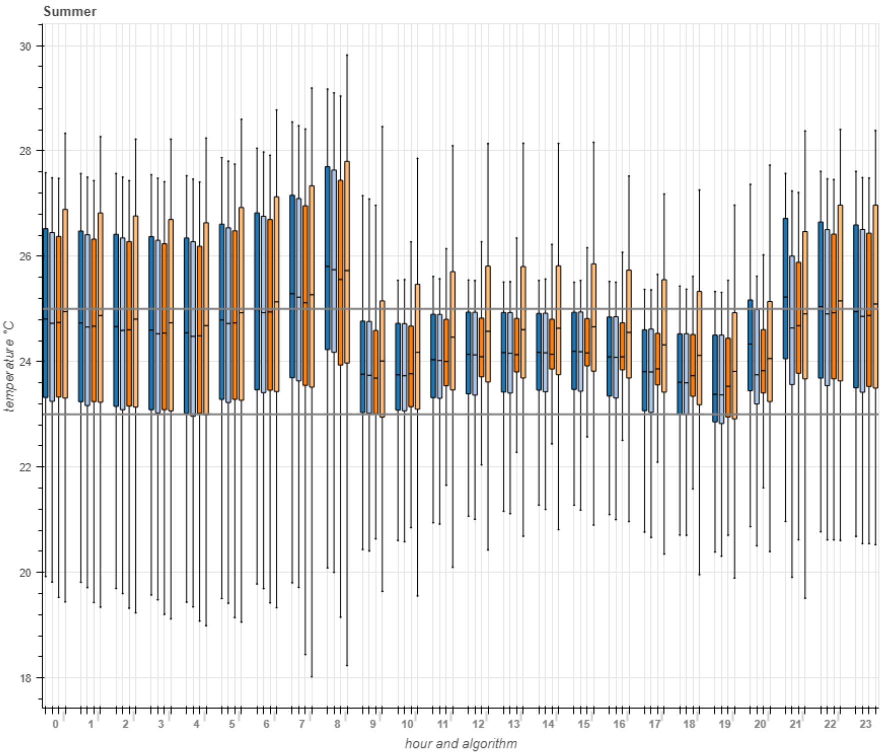



Figure 155: Cooling season (31 Mar. to 1 Dec.) hourly mean internal air temperature, Reference Building: Supermarket in climate zone **B** with random occupation model.

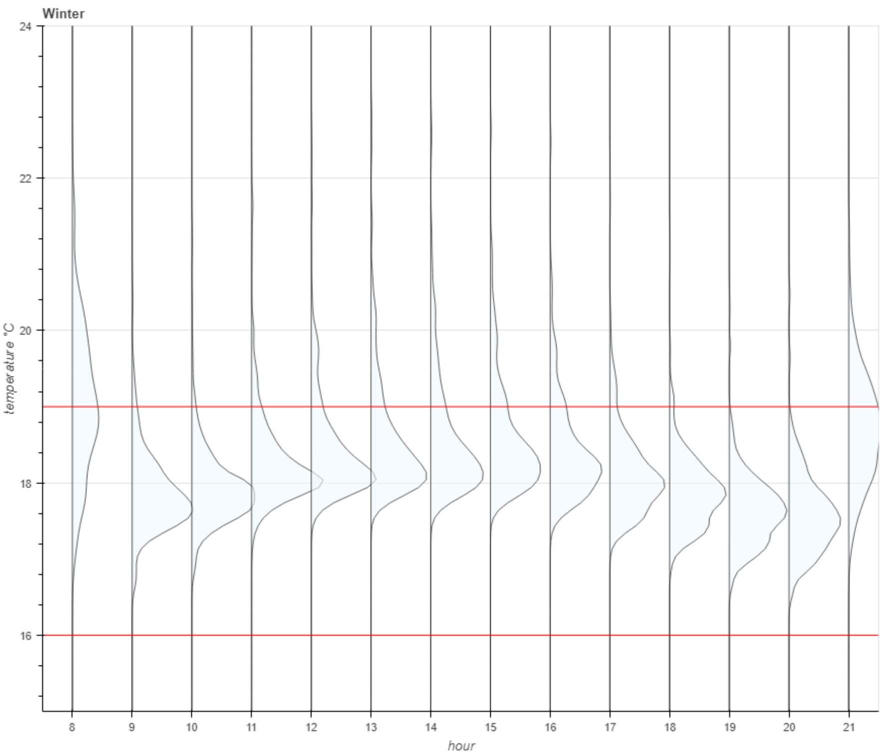



Figure 156: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket in climate zone **B** with random occupation model.

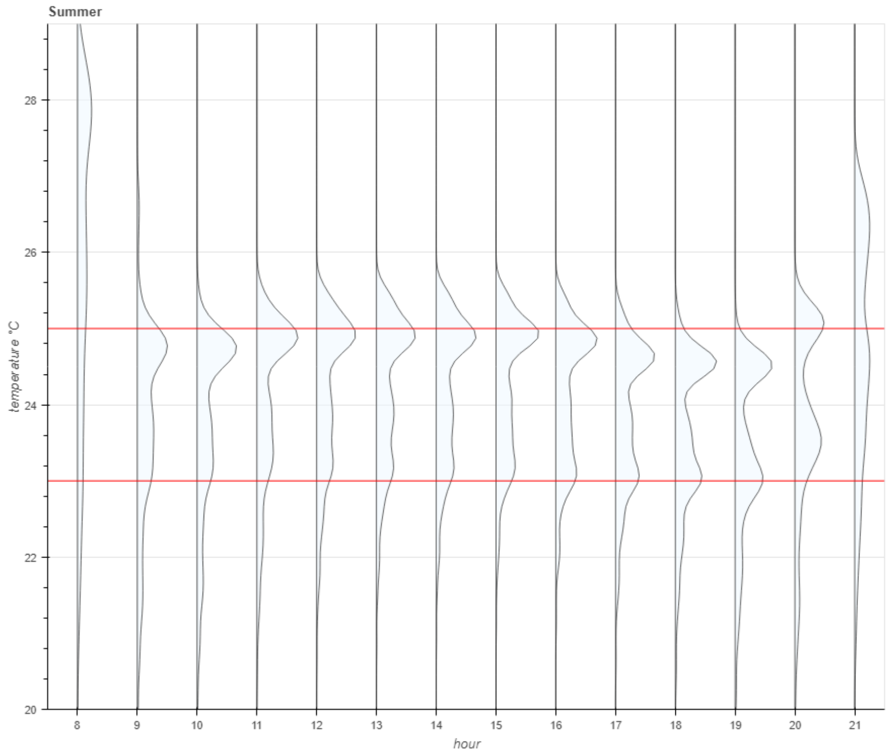



Figure 157: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket in climate zone **B** with random occupation model.

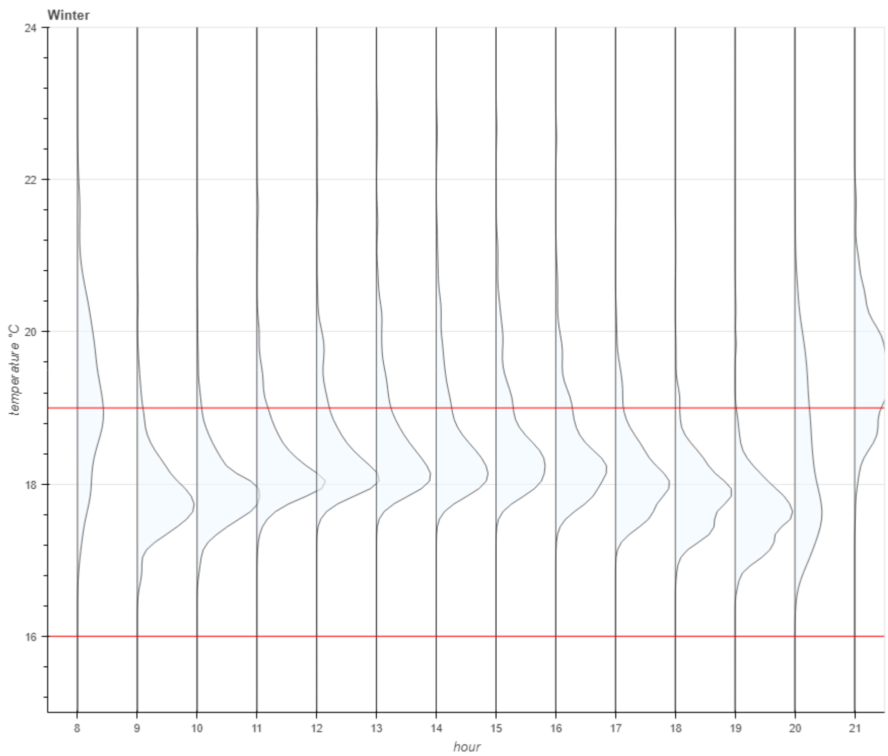



Figure 158: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket in climate zone **B** with random occupation model.

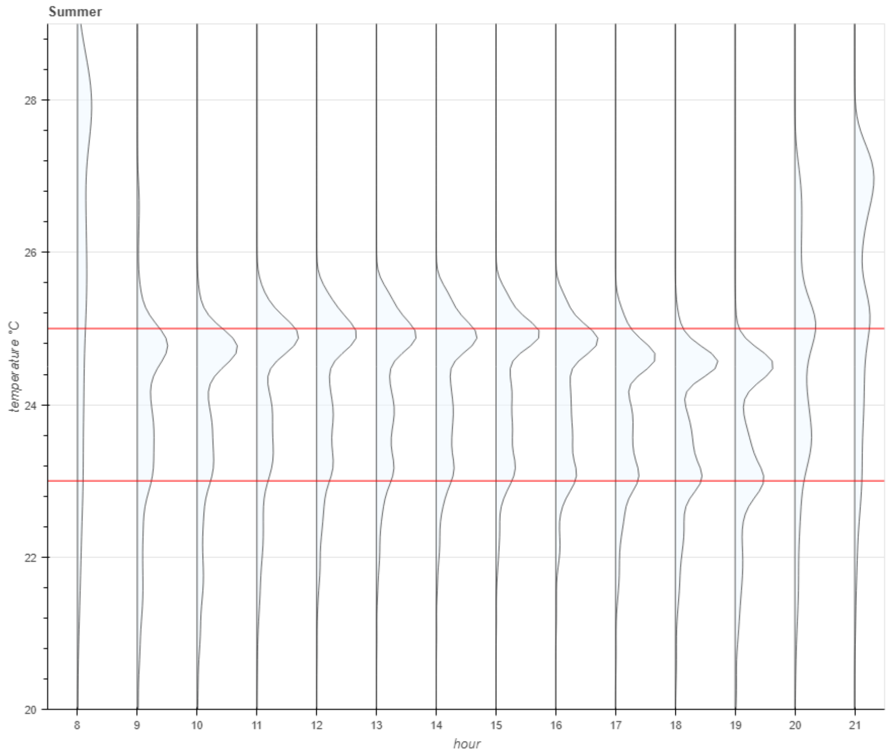



Figure 159: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution for Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket in climate zone **B** with random occupation model.

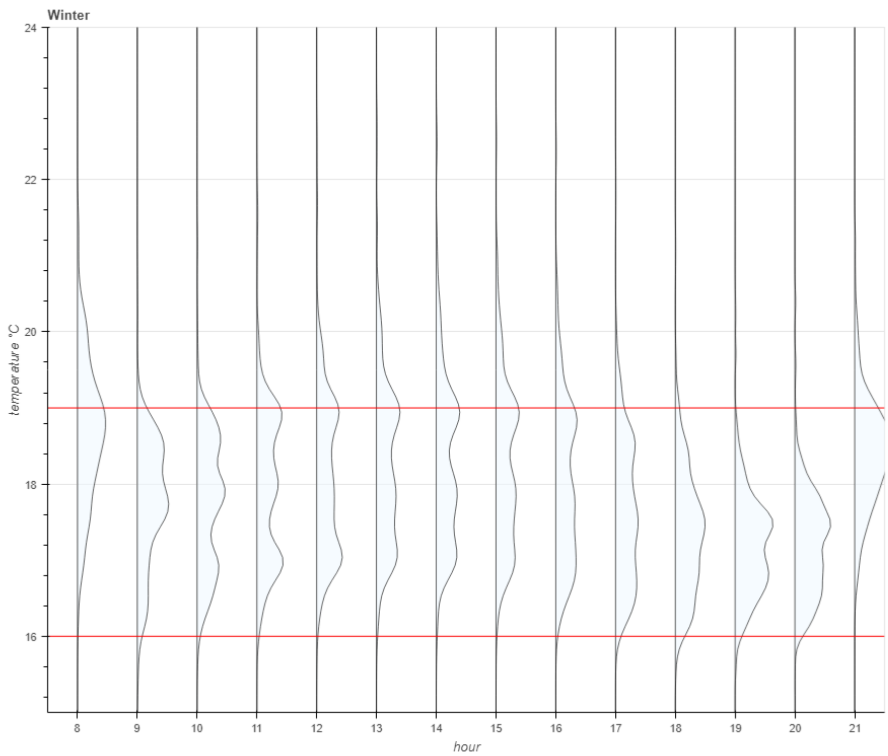



Figure 160: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket in climate zone **B** with random occupation model.

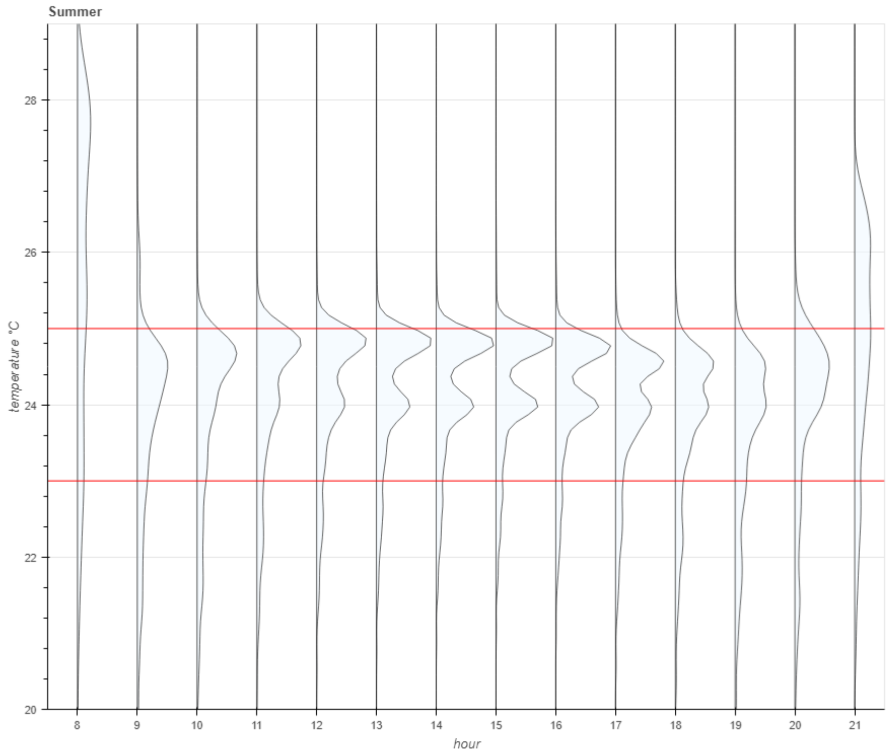



Figure 161: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket in climate zone **B** with random occupation model.

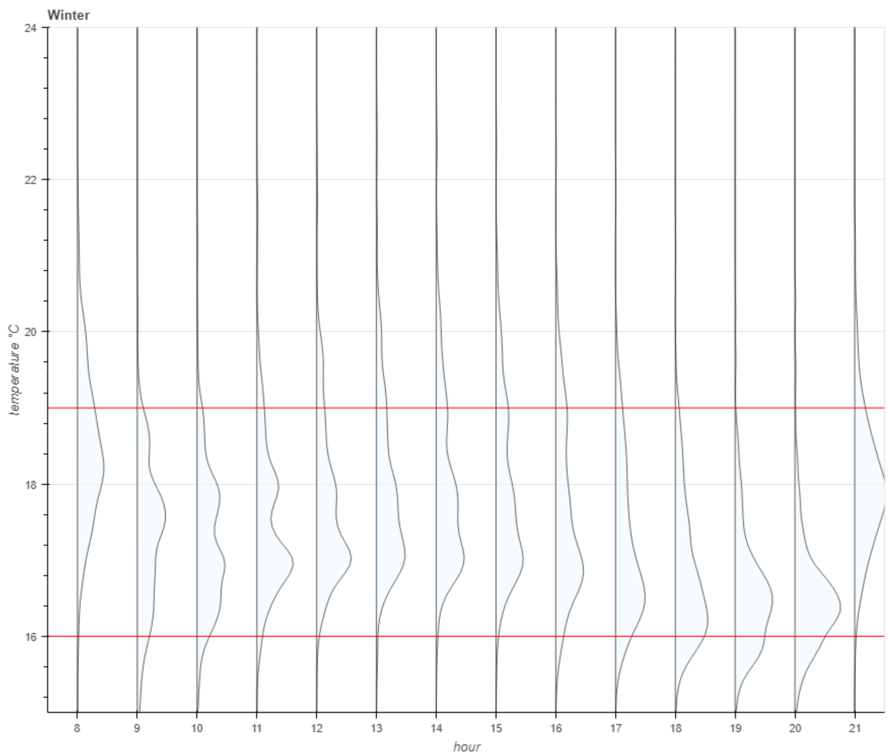



Figure 162: Heating season (1 Dec. to 31 Mar.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket in climate zone **B** with random occupation model.

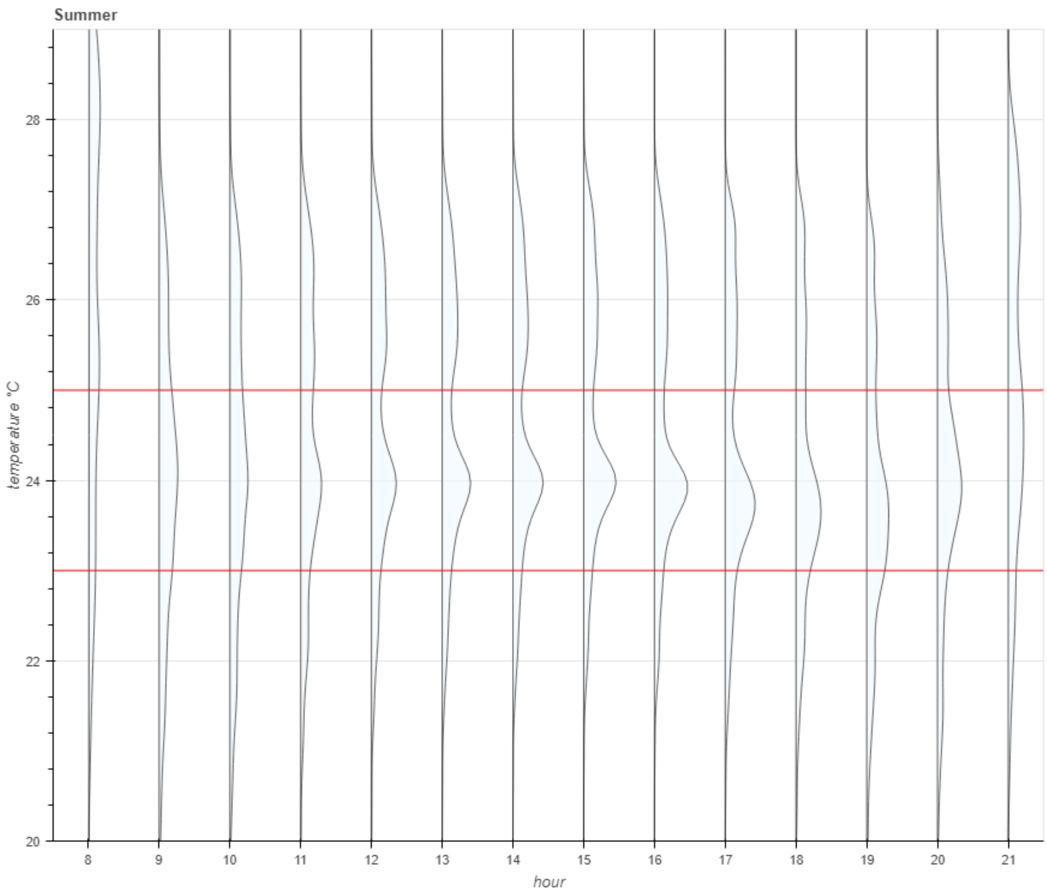



Figure 163: Cooling season (31 Mar. to 1 Dec.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket in climate zone **B** with random occupation model.

| | below_%_15.5 | below_%_16.0 | above_%_19.5 | above_%_20.0 |
|---|---|---|---|---|
| season | winter | winter | winter | winter |
| experiment | | | | |
| constant_30min | 0.01% | 1.06% | 14.95% | 7.58% |
| constant_oo | 0.00% | 0.92% | 12.94% | 7.37% |
| slack_1 | 1.29% | 6.26% | 12.51% | 6.95% |
| slack_10 | 8.19% | 30.38% | 11.55% | 6.58% |

Table 68: Percentage of exceeding air temperature in Heating Season (1 Dec. to 31 Mar.), Reference Building: Supermarket in climate zone **B** with random occupation model.

| | below_%_22.5 | below_%_23.0 | above_%_25.5 | above_%_26.0 |
|---|---|---|---|---|
| **season** | summer | summer | summer | summer |
| **experiment** | | | | |
| **constant_30min** | 34.86% | 45.30% | 19.57% | 13.11% |
| **constant_oo** | 38.77% | 48.96% | 4.05% | 2.65% |
| **slack_1** | 37.59% | 46.10% | 3.79% | 1.56% |
| **slack_10** | 37.95% | 46.08% | 31.20% | 21.11% |

Table 69: Percentage of exceeding air temperature in Cooling Season (31 Mar. to 1 Dec.), Reference Building: Supermarket in climate zone **B** with random occupation model.

| | saving | |
|---|---|---|
| **season** | summer | winter |
| **experiment** | | |
| **constant_30min** | 3.52% | 4.01% |
| **constant_oo** | 0.00% | 0.00% |
| **slack_1** | -2.38% | 20.70% |
| **slack_10** | 7.18% | 22.44% |

Figure 164: Relative savings with respect of constant set-point, Reference Building: Supermarket in climate zone **B** with random occupation model.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant 30 min | 326360.041 | 143.061 |
| Constant | 338674.079 | 148.459 |
| Slack 1 | 263330.302 | 115.431 |
| Slack 10 | 240596.894 | 105.466 |

Table 70: Total primary energy consumption for HVAC, Reference Building: Supermarket in climate zone **B** with random occupation model.

And for climate zone **E**:

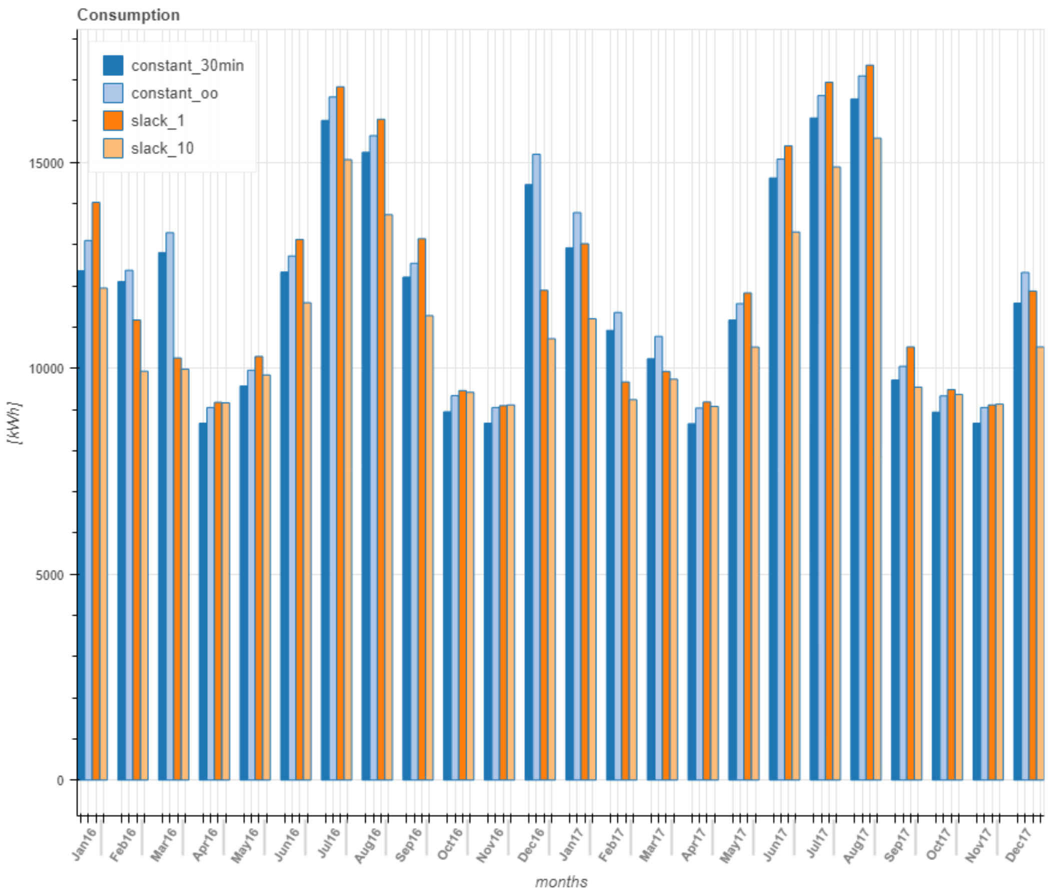



Figure 165: Primary Energy Consumption for HVAC system, Reference Building: Supermarket in climate zone **E** with random occupation model.

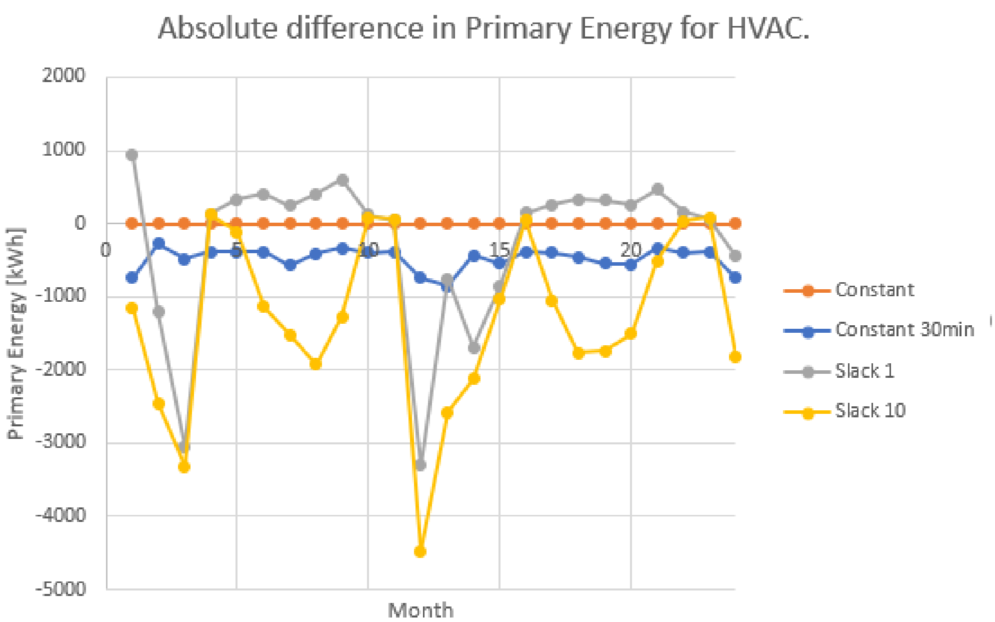



Figure 166: Absolute difference in Primary Energy Consumption for HVAC system, Reference Building: Supermarket in climate zone **E** with random occupation model.

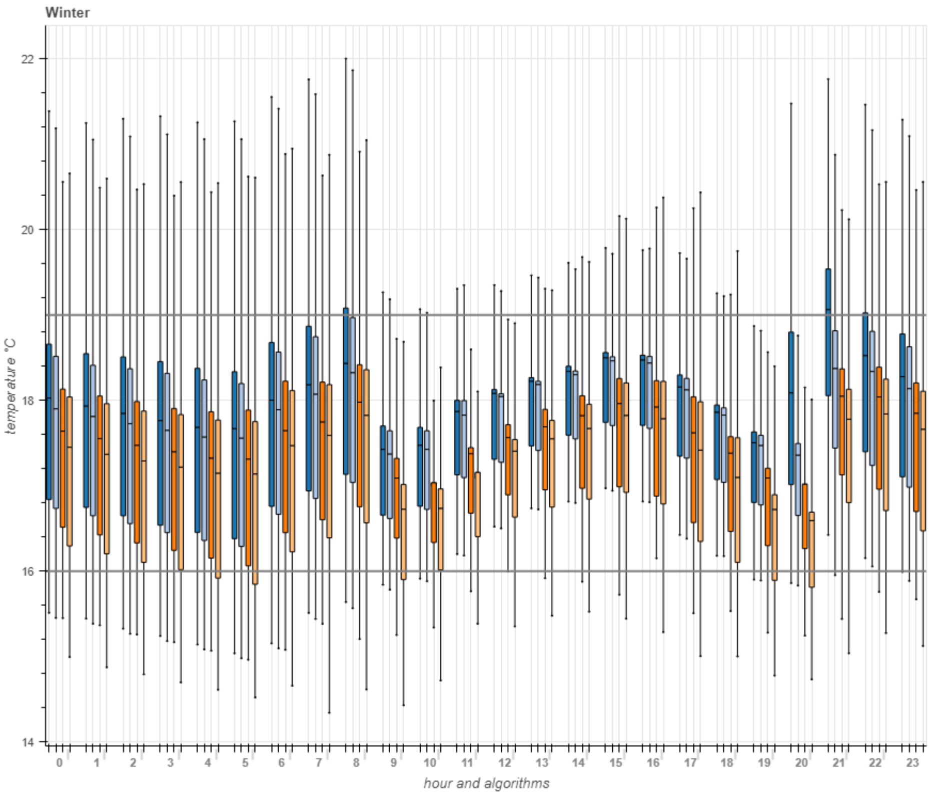



Figure 167: Heating season (15 Oct. to 15 Apr.) hourly mean internal air temperature, Reference Building: Supermarket in climate zone **E** with random occupation model.

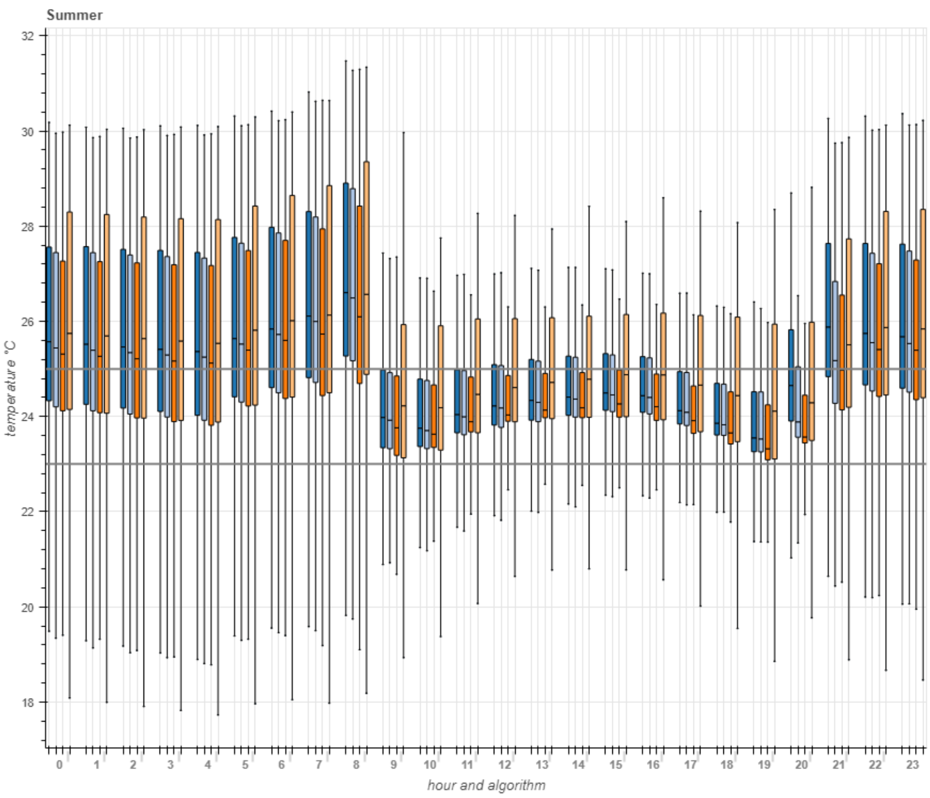



Figure 168: Cooling season (15 Apr. to 15 Oct.) hourly mean internal air temperature, Reference Building: Supermarket in climate zone **E** with random occupation model.

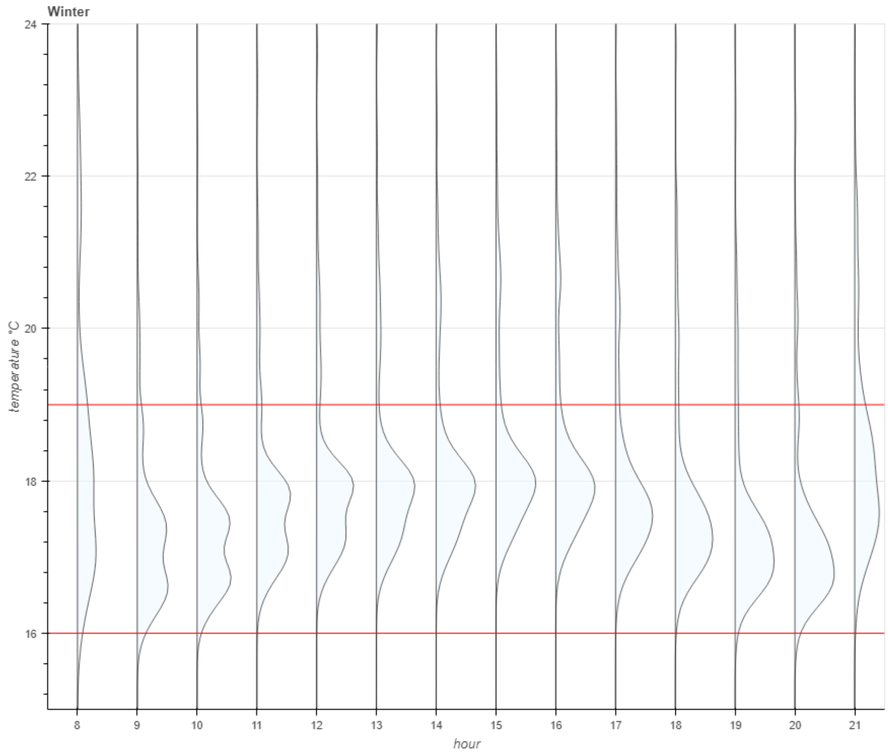



Figure 169: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket in climate zone **E** with random occupation model.

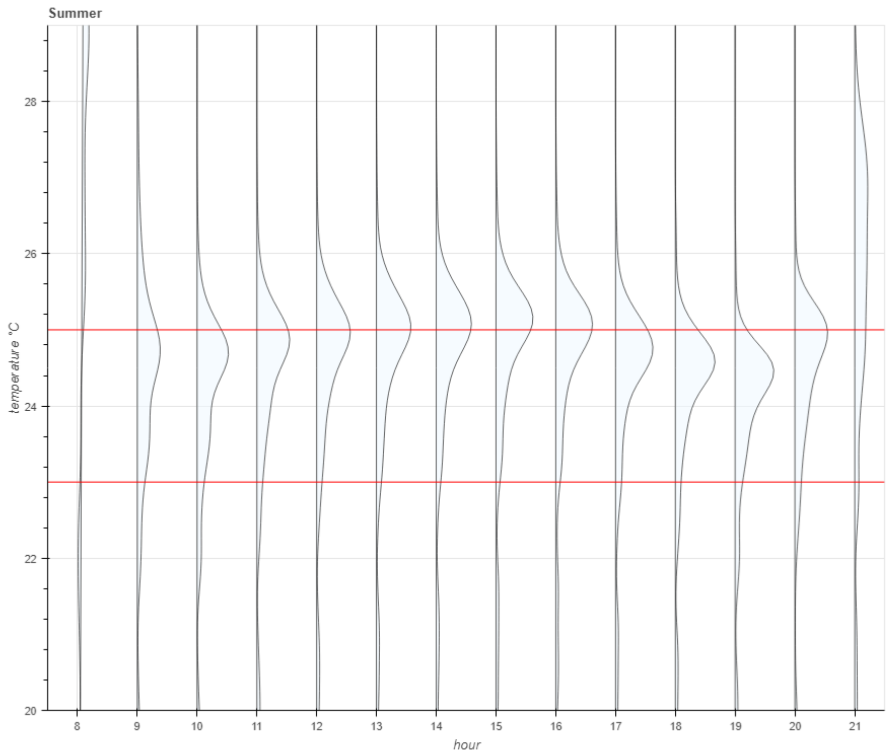



Figure 170: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm, Reference Building: Supermarket in climate zone **E** with random occupation model.

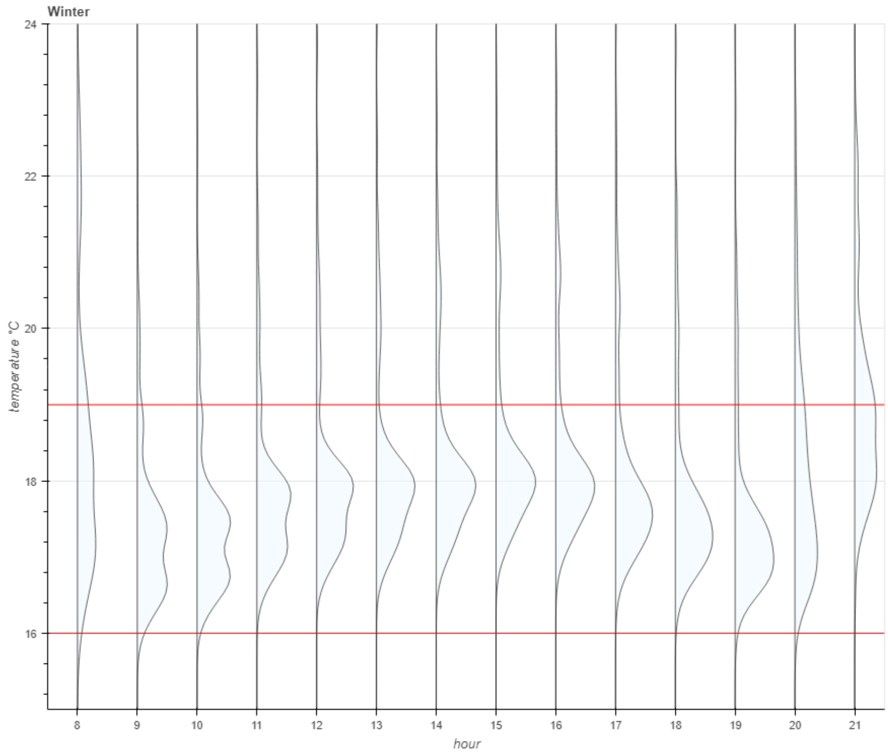



Figure 171: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket in climate zone **E** with random occupation model.

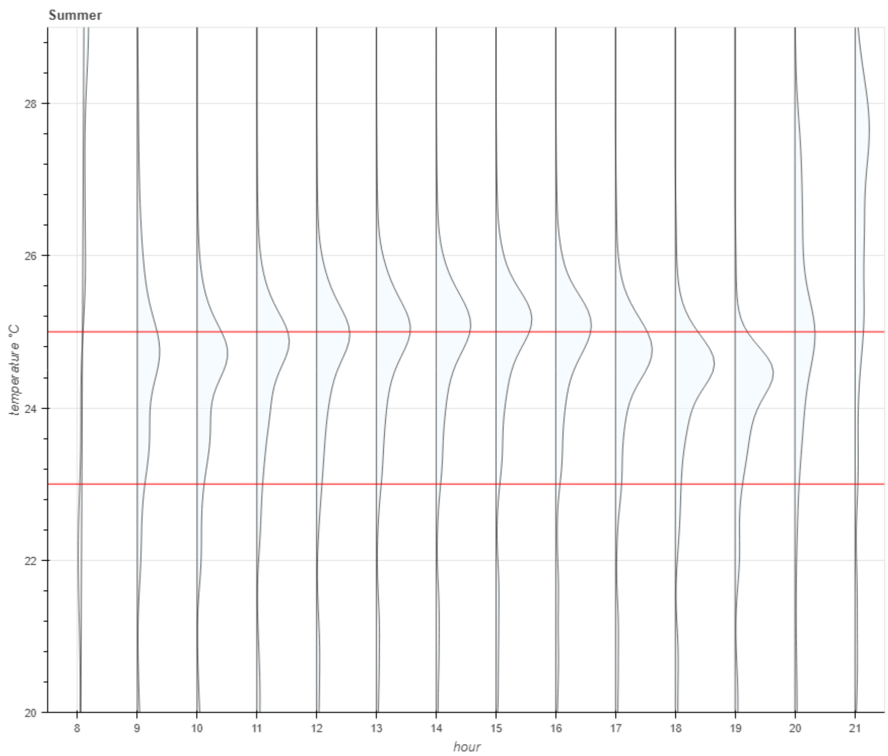



Figure 172: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for Constant algorithm with shut down 30 minutes before closure, Reference Building: Supermarket in climate zone **E** with random occupation model.

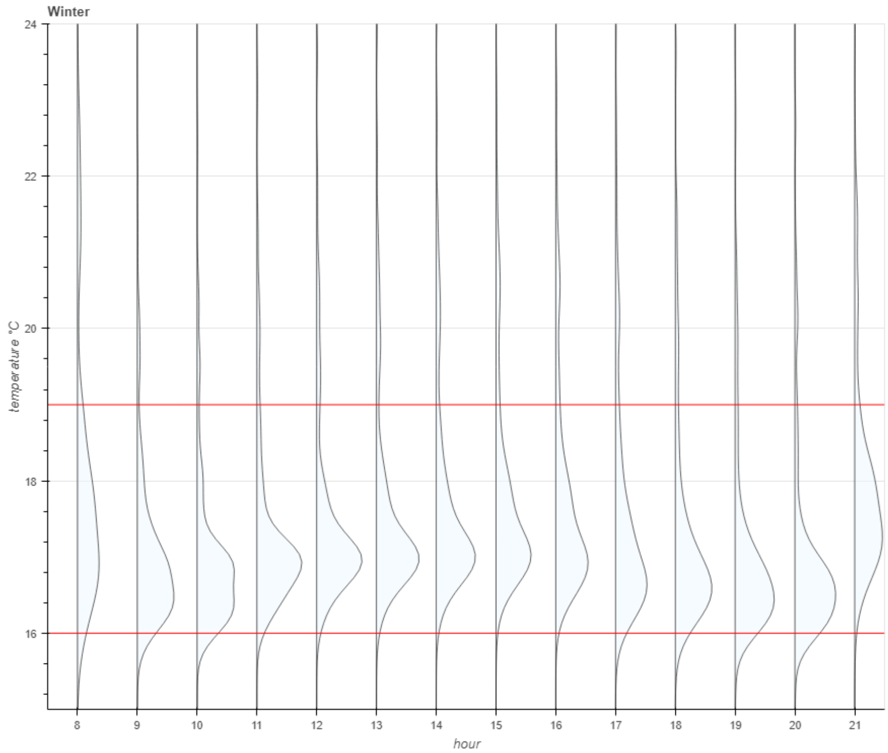



Figure 173: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket in climate zone **E** with random occupation model.

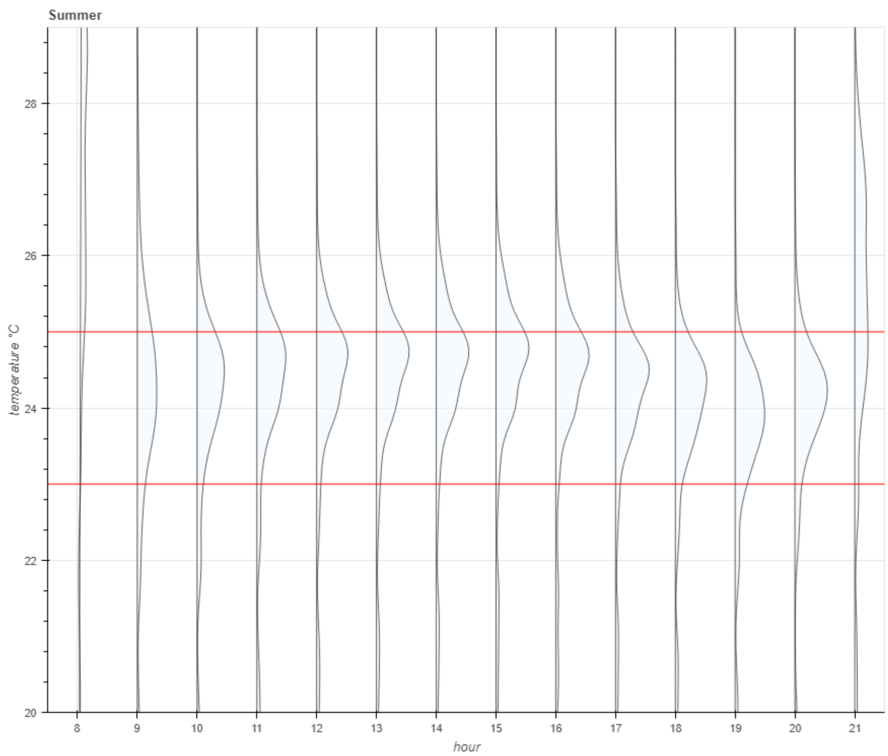



Figure 174: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 1, Reference Building: Supermarket in climate zone **E** with random occupation model.

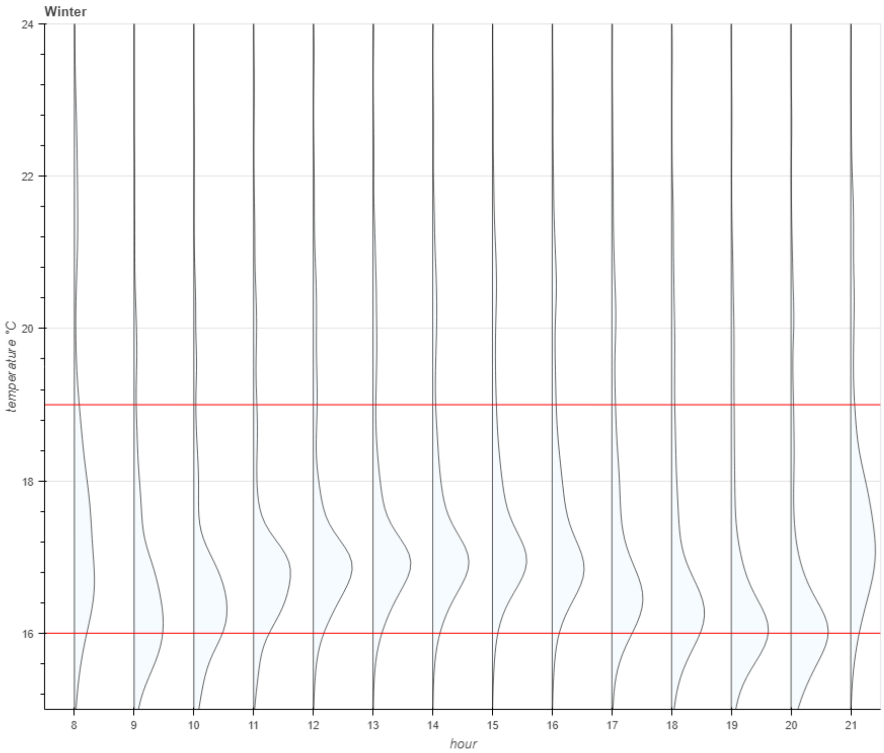



Figure 175: Heating season (15 Oct. to 15 Apr.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket in climate zone **E** with random occupation model.

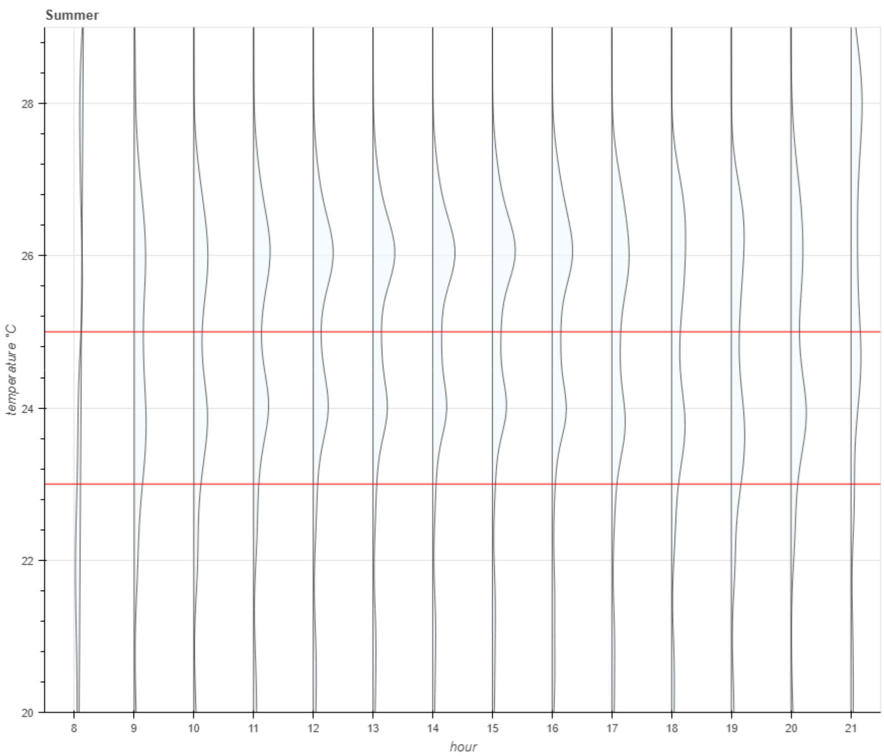



Figure 176: Cooling season (15 Apr. to 15 Oct.) hourly internal air temperature distribution for RL algorithm with slack = 10, Reference Building: Supermarket in climate zone **E** with random occupation model.

| season | below_%_15.5<br>winter | below_%_16.0<br>winter | above_%_19.5<br>winter | above_%_20.0<br>winter |
|---|---|---|---|---|
| experiment | | | | |
| constant_30min | 6.08% | 28.20% | 16.85% | 14.10% |
| constant_oo | 6.99% | 34.71% | 16.50% | 13.72% |
| slack_1 | 7.35% | 24.74% | 15.55% | 12.95% |
| slack_10 | 35.83% | 57.27% | 15.42% | 12.86% |

Table 71: Percentage of exceeding air temperature in Heating Season (15 Oct. to 15 Apr.), Reference Building: Supermarket in climate zone **E** with random occupation model.

| season | below_%_22.5<br>summer | below_%_23.0<br>summer | above_%_25.5<br>summer | above_%_26.0<br>summer |
|---|---|---|---|---|
| experiment | | | | |
| constant_30min | 25.19% | 32.74% | 31.91% | 23.27% |
| constant_oo | 30.72% | 38.20% | 15.69% | 8.53% |
| slack_1 | 32.11% | 41.04% | 12.51% | 7.20% |
| slack_10 | 32.40% | 40.40% | 47.23% | 31.47% |

Table 72: Percentage of exceeding air temperature in Cooling Season (15 Apr. to 15 Oct.), Reference Building: Supermarket in climate zone **E** with random occupation model.

| | | saving |
|---|---|---|
| season | summer | winter |
| experiment | | |
| constant_30min | 3.47% | 4.38% |
| constant_oo | 0.00% | 0.00% |
| slack_1 | -2.22% | 2.79% |
| slack_10 | 6.29% | 5.70% |

Figure 177: Relative savings with respect of constant set-point, Reference Building: Supermarket in climate zone **E** with random occupation model.

| Algorithm | kWh | $\frac{kWh}{m^2}$ |
|---|---|---|
| Constant 30 min | 283457.184 | 124.254 |
| Constant | 294949.570 | 129.292 |
| Slack 1 | 288856.413 | 126.621 |
| Slack 10 | 263893.046 | 115.678 |

Table 73: Total primary energy consumption for HVAC, Reference Building: Supermarket in climate zone **E** with random occupation model.

## 6 Conclusions and Future perspectives.

Through simulations conducted with EnergyPlus both for a real case study and for a Reference Building suggested by DOE the Reinforcement Learning based control for HVAC system has been able to achieve energy savings maintaining the same air temperature range both for cooling and heating seasons.

For the real case study located in Bergamo savings introduced by the algorithm are bigger because it can compensate some inaccuracies or changes in HVAC sizing or building destination where traditional controls lack in efficiency justifying the implementation of Reinforcement Learning algorithm as retrofit application for existent buildings.

Reference Building model can simulate application in new buildings where the HVAC system is well sized and tuned for the building destination. In this case the Reinforcement Learning algorithm is still able to achieve energy savings adjusting some parameters in the algorithm.

The performances of the algorithm are constant when varying model parameters such as location, destination or occupants behaviour without any change in the algorithm while traditional controls are not able to guarantee energy savings or acceptable air temperature for the occupants.

Reinforcement Learning algorithm learn by interaction with the environment so best savings can be obtained by implement the algorithm in a real HVAC system and monitor, through sensors, the actions chosen by the agent to identify mistakes and when it miss opportunity of savings to correct them to then apply in several buildings.

Next step is to monitor thermal comfort indicator such as **Predicted Mean Vote** or **Percentage of People Dissatisfied** [23] to control through operative temperature, that combine information of air and mean radiant temperature, to assess real occupants thermal comfort requirements. This procedure is not applicable in all HVAC system as there are more sensors required, such as for mean radiant temperature.

To further improve performance of Reinforcement Learning based control algorithm for HVAC system a possibility is to implement it directly in the inner-loop control of HVAC system and allow the algorithm to control directly actuators of valves and so the water flow in cooling and heating decks.

## List of Figures

## List of Tables

## References


[1] "2017 ASHRAE Handbook-Fundamentals(SI)", Chapter 7 Fundamentals of Control.

[2] "2007 ASHRAE Handbook-HVAC Application(SI)", Chapter 41 Supervisory Control Strategies and Optimization.

[3] G. Pattarello, "Model Predictive Control of HVAC Systems: Design and Implementation on a Real Case Study.", 2013.

[4] "Riscaldamento degli Edifici", P. Andreini, D. Pitimada, Hoepli 1995.

[5] "Pattern Recognition and Machine Learning", Cristopher M. Bishop, Springer 2006.

[6] "Reinforcement Learning: An Introduction", Richard S. Sutton and Andrew J. Barto, The MIT Press Cambridge 1998.

[7] "Fundamentals of Reinforcement Learning", Yann-Michael De Hauwere.

[8] "An Introduction to Statistical Learning", G. James, D. Witten, T. Hastie, R. Tibshirani, Springer 2013.

[9] "Neural Networks and Deep Learning", M. Nielsen, 2017.

[10] "Autonomous HVAC Control, A Reinforcement Learning Approach", E. Barret, S. P. Linder, Conference Paper 2015.

[11] Learning agent for a heat-pump thermostat with a set-back strategy using model-free reinforcement learning, F. Ruelens, S. Iacovella, B. J. Claessens, and R. Belmans, Energies, vol. 8, no. 8, pp. 83008318, 2015. [Online]. Available: http://www.mdpi.com/1996-1073/8/8/8300

[12] Deep reinforcement learning for building hvac control, T. Wei, Y. Wang, and Q. Zhu, in Proceedings of the 54th Annual Design Automation Conference 2017, ser. DAC 17. New York, NY, USA: ACM, 2017, pp. 22:122:6. [Online]. Available: http://doi.acm.org/10.1145/3061639.3062224

[13] "Building Control Virtual Test Bed User Manual", Lawrence Berkeley National Laboratory, 2015.

[14] "Impianti di climatizzazione. Tipologie applicative" L. Stefanutti, Tecniche Nuove.

[15] "EnergyPlus Version 8.5 Documentation - Getting Started", U.S. Department Of Energy, 2016.

[16] "EnergyPlus Version 8.5 Documentation - Engineering Reference", U.S. Department Of Energy, 2016.

[17] "EnergyPlus Version 8.5 Documentation - Input Output Reference", U.S. Department Of Energy, 2016.

[18] "Commercial Reference Building" - U.S. Department of Energy (DOE).

[19] "U.S. Department of Energy Commercial Reference Building Models of the National Building Stock" - NREL, National Renewable Energy Laboratory, 2011.

[20] "Elementi di Termofisica Generale e Applicata" - A. Cocchi, Esculapio, 1990.

[21] "Thermal bridging atlas of steel construction for improved energy efficiency of buildings (TABASCO)" - Directorate-General for Research and Innovation (European Commission)

[22] "UNI EN 1634-1:2018, Prove di resistenza al fuoco e di controllo della dispersione del fumo per porte e sistemi di chiusura, finestre apribili e loro accessori costruttivi".

[23] "UNI EN 15251:2007, Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics".

[24] "UNI 10339:1995, Impianti aeraulici al fini di benessere".

[25] "Impianti Idrosanitari", M. Doninelli, Quaderni Caleffi.

[26] "D.P.R. 26 agosto 1993, n. 412. Regolamento recante norme per la progettazione, l'installazione, l'esercizio e la manutenzione degli impianti termici degli edifici ai fini del contenimento dei consumi di energia, in attuazione dell'art. 4, comma 4, della L. 9 gennaio 1991, n. 10" con s.m.i.

[27] "DECRETO 11 marzo 2008", MINISTERO DELLO SVILUPPO ECONOMICO.

[28] "UNI EN 15232-1:2017, Prestazione energetica degli edifici".

[29] "UNI 5364, Impianti di riscaldamento ad acqua calda. Regole per la presentazione dell' offerta e per il collaudo."

[30] "A generalised stochastic model for the simulation of occupant presence.", J. Page, D. Robinson, N. Morel, J.-L. Scartezzini, LESO-PB, Batiment LE, Station 18, E cole Polytechnique Federale de Lausanne, CH-1015 Lausanne, Switzerland, Received 25 September 2006; received in revised form 15 January 2007; accepted 22 January 2007.