Theory of Hybrid Systems Informatik 2

RWTH AACHEN UNIVERSITY

Diese Arbeit wurde vorgelegt am
Lehr- und Forschungsgebiet Theorie der hybriden Systeme

Automatisierung und Analyse von Multi-Output Modellen des maschinellen Lernens bei Gebäudesteuerung

Automation and investigation of multi output Machine learning control in buildings

Bachelorarbeit
Informatik

August 2021

Vorgelegt von Presented by	Laura-Sophie Kirchner Matrikelnummer: 394393
Erstprüfer First examiner	Prof. Dr. rer. nat. Erika Ábrahám Lehr- und Forschungsgebiet: Theorie der hybriden Systeme RWTH Aachen University
Zweitprüfer Second examiner	Prof. Dr. rer. nat. Thomas Noll Lehr- und Forschungsgebiet: Theorie der hybriden Systeme RWTH Aachen University
Betreuer Supervisor	M.Sc. + Dipl.-ing Ahmed Abida Lehr- und Forschungsgebiet: Theorie der hybriden Systeme RWTH Aachen University

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich diese Bachelorarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Stellen meiner Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, habe ich in jedem Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht. Dasselbe gilt sinngemäß für Tabellen und Abbildungen. Diese Arbeit hat in dieser oder einer ähnlichen Form noch nicht im Rahmen einer anderen Prüfung vorgelegen.

Aachen, im August 2021

LAURA-SOPHIE KIRCHNER

Abstract

Buildings account for 30-45% of global energy consumption [25]. Together with the building construction sector, they are responsible for almost 40% of all direct and indirect CO2 emissions [1]. Heating, ventilation and air conditioning systems play an important role in this context [25]. Hence, the efficient use and control of HVAC systems has a major impact on minimizing energy consumption. Studies show that better control of HVAC systems can lead to energy savings of 13% to 28% [12]. To improve the efficiency of the HVAC system control, this thesis will consider Model Predictive Control and Neural Networks.

In order to minimize the computational cost and working effort, this work addresses the questions to what extent it is possible to adopt multi-output models for HVAC control instead of single-output models and whether one model structure is applicable for different zones of different buildings. To this end, we also try to automate the process of data preparation and parameter selection as much as possible.

In the course of this thesis, we consider one model and different preprocessing methods. This includes testing the response of the individual zones to different amounts of data, as well as to varying inputs, depending on the correlations.

In order to evaluate the achievement of the objectives, a graphical user interface is developed and five tests are evaluated.

The first test is to determine whether one preprocessing method in combination with the chosen model can successfully simulate the data of all zones, or at least outperform the others.

We select a preprocessing method in the second test and compare the MO and SO model losses.

During the third test, we examine the connections between one zone with particularly good results and one zone with particularly poor results.

In the fourth test, we evaluate the extent to which it is possible to convert from SO to MO models. For this purpose, we determine the minimum losses of each zone that can be achieved with our preprocessing methods and then compare the minimum achievable MO and SO losses independent of the respective preprocessing selection.

Finally, in test five, we want to discover patterns in the data that reveal the extent to which certain factors influence model losses.

Contents

List of Figures

List of Tables

1. Introduction

1.1. Motivation

Buildings are responsible for 30-45% of the world's energy consumption [25]. Combined with the buildings construction sector, they account for almost 40% of all direct and indirect CO2 emissions [1]. In recent years, energy consumption has increased significantly, in part due to increased heating and cooling needs resulting from extreme weather conditions [1]. Heating, Ventilation and Air Conditioning (HVAC) systems account for a large share of global energy consumption [25]. Therefore, efficient use and control of HVAC systems plays a major role in minimizing energy consumption. Studies show that better control of HVAC systems can result in 13% to 28% energy reductions [12]. Smart buildings are a way to meet human needs such as comfort and wellbeing while reducing CO2 emissions and energy consumption.

1.2. Related work

1.2.1. Model Predictive Control

The 2018 revision of the EU Directive states that all non-residential buildings with an HVAC output over 290kW and where "technically and economically feasible" should be equipped with building automation and control systems by 2025 [2]. While rule-based control (RBC) is still frequently used, model predictive control (MPC) has become much more popular in recent years due to the increase in thermal comfort and the simultaneous reduction in energy consumption by 15% to 50% [12]. MPC is used to predict the optimal future behaviour depending on a building model, selected objectives and weather forecasts.
Tarragona et al. [37] used MPC to control a heating system consisting of an air-to-water pump, thermal energy storage and solar panels to reduce energy costs. By taking into account all system-relevant inputs, such as heating demand, electricity prices and outdoor temperatures, testing different working horizons and MPC settings, and designing the optimization problem based on branch-and-bound optimization, the energy costs were reduced by about 50%.
Gholamzadehmir et al. [16] compare MPC with adaptive predictive control and conclude that MPC is not suitable for systems that rely on unpredictable data.
Since, according to Pippia et al. [32], deterministic MPC strongly depend on the quality of the predictions of disturbances and linear models might not be able to grasp the behaviour of a building under control, they created a stochastic scenario-based MPC (SBMPC) controller with a non-linear Modelica model.
Pippia et al. conclude that this controller outperforms an SBMPC with a linear model, as well as a deterministic controller with a Modelica model.
Nagpal et al. [29] introduce a new controller based on MPC with parametric uncertainties and time-varying constraints. They attempt to achieve robust control using a min-max optimization problem by minimizing a reasonable worst-case cost function and thus manage to keep the room temperature in the preferred range, even under

large model uncertainties.
Carli et al. [9] present an Internet of Things (IoT) based architecture for the implementation of MPC for HVAC systems. MPC is designed to optimize both indoor thermal comfort and energy consumption. The subsystems for sensing, control and actuating are connected to the internet and users can remotely set the control mode and set-points of the system. The architecture from Carli et al. was tested for a campus building at the Polytech of Bari.
As the use of MPC involves a lot of development effort, communication infrastructure and computer capacity, according to Drgoňa et al. [13], the building model and the MPC optimization problem are often simplified, which can lead to non-optimal performance. A suitable white box model can bring the performance of an MPC system closer to the theoretical limits. Therefore, Drgoňa et al. present a computationally efficient white box MPC model that uses first-principle physical forms and has been tested in a building in Belgium.
Ma et al. [23] deal with MPC with invariant sets, moving block strategy, dual stage optimization and predictions of building loads and weather forecasts for cooling systems that have a water tank for the storage of cold water produced by chillers. Their results show a 24.5% reduction in electricity expenditure.
Mantovani and Ferrarini [24] focus on the use of MPC in shopping centres to optimize the room temperature. They use a non-linear model as a basis for different MPC configurations and time-varying constraints. Their simulations achieve energy savings of about 4.5%.
Halvgaard et al. [17] want to control HVAC systems, blind positioning and electric lighting concurrently so that temperature, CO2 and lighting satisfy the occupants' comfort zone and maximize energy efficiency. For this, they have developed a stochastic MPC (SMPC) that uses weather forecasts to calculate how much energy would be consumed to maintain a certain comfort level and what costs this would entail. Compared to deterministic MPC (DMPC), SMPC is better in terms of comfort violations and non-renewable primary energy usage, as it incorporates uncertainty in weather forecasts. For heat pumps of residential buildings with water-based floor heating, Halvgaard et al. [17] use a linear state-based model and an economic MPC controller as a linear program, based on weather and electricity price forecasts, to exploit the thermal capacity of the building in such a way that energy consumption is shifted to times of low electricity prices. Their approach brings 25-35% electricity savings.
Moroşan et al. [27] use the intermittently operating mode of almost all types of buildings to achieve good control results and low computational effort with a distributed MPC control strategy with one information exchange per time unit. Their approach increases the thermal comfort level by about 36.7% while reducing the energy expenses by about 13.4%.

1.2.2. Artificial Neural Networks

In recent years, the use of Artificial Neural Networks (ANN) for the prediction of room temperature has also gained popularity. In this approach, relevant factors such as

building structure, weather forecasts and outdoor temperature are given to an ANN as input parameters and the future internal room temperature is calculated [15].
Egilegor et al. [14] use fuzzy control adjusted by a neural network with temperature and humidity as inputs and the thermal comfort index predicted mean vote as output to adjust the fan-coil air flow rate under different weather conditions to maximize comfort.
Since most studies deal with single-zone models, Huang et al. [19] present an ANN model-based system for multi-zone buildings that incorporates energy generation from mechanical cooling, ventilation, weather changes and effects of thermal interaction between adjacent zones, thus giving more accurate predictions than single-zone models and providing thermal comfort and reduced energy consumption.
Zhou et al. [39] use a Multi Output Long Short-Term Memory (LSTM) ANN to compensate for spatio-internal instability and time-lag effects in air quality forecasting.
Sendra-Arranz and Gutiérrez [34] propose several LSTM-based multi-step prediction models to estimate the energy consumption of an HVAC system one day ahead and test these models in the self-sufficient solar house MagicBox.
Mtibaa et al. [28] compare multi-output and single-output LSTM models based on multi-step prediction with direct sequence-to-sequence architecture to predict room temperature as accurately as possible for multi-zone buildings with different HVAC types.
To find the ANN that best reduces HVAC-related electricity expenditure, Deb et al. [11] test all subsets of 14 inputs of data from 56 office buildings in Singapore. They conclude that gross floor area, energy consumption by air conditioning, operational hours and chiller plant efficiency is the best combination of input variables.
In order to optimize a building's HVAC two-chiller system in terms of thermal comfort and energy consumption, Nasruddin et al. [30] combine an ANN with a multi-objective genetic algorithm.
Kusiak and Xu [22] also use multi-objective optimization to optimize an HVAC system in terms of energy consumption and thermal comfort with a dynamic ANN. With this approach, they were able to save about 30% energy compared to the standard control.
To address the problem that HVAC systems are nonlinear, complex and delayed, Shujiang Li et al. [35] develop a backpropagation neural network model controlled by a generalised predictive control method.
Afram and Janabi-Sharifi [6] develop several black-box models, including a multiple-input multiple-output (MIMO) ANN, for HVAC systems in a residential building and compare them with the corresponding grey-box models. They conclude that the ANN model outperforms all other models.

1.3. Research Questions

On the one hand, this thesis will address the question of how well different buildings of one HVAC type can be controlled with a single model. On the other hand, the question of how well a multiple-output system performs compared to single-output systems for these buildings will be considered. The goal is to minimize the computational power by not having to find a separate model for each building and train a seperate model for each output.

1.4. Outline

The background of the thesis such as the motivation, related work and the questions to be investigated were addressed in chapter 1. Chapter 2 deals with the theoretical foundations such as building modelling, MPC, optimization and ANNs. With regard to ANNs, special attention is paid to the possibilities of multi-output ANNs in order to be able to calculate several results with one ANN. In chapter 3, the data base is introduced and on this basis, a case study is carried out in chapter 4, that implements the methods proposed in chapter 2. In the end chapter 5 gives a short summary and an outlook on which aspects could be further explored in the future.

2. Theoretical foundations

2.1. Building modeling

From a cost-effective and sustainable perspective, the most efficient way to use HVAC systems is to improve the control mechanisms rather than upgrading the equipment [7]. Therefore, an accurate modelling of the building system is essential. HVAC modelling techniques are divided into three categories according to Afroz et al. [7] and Afram and Janabi-Sharifi[5]. In the following, these categories will be examined in more detail.

2.1.1. White Box Models

White Box models are also known as physics-based models, mathematical models or forward models [7]. These are primarily used to predict HVAC performance at the design stage. For this purpose, mathematical formulas are developed and solved depending on the fundamentals of Thermodynamics.

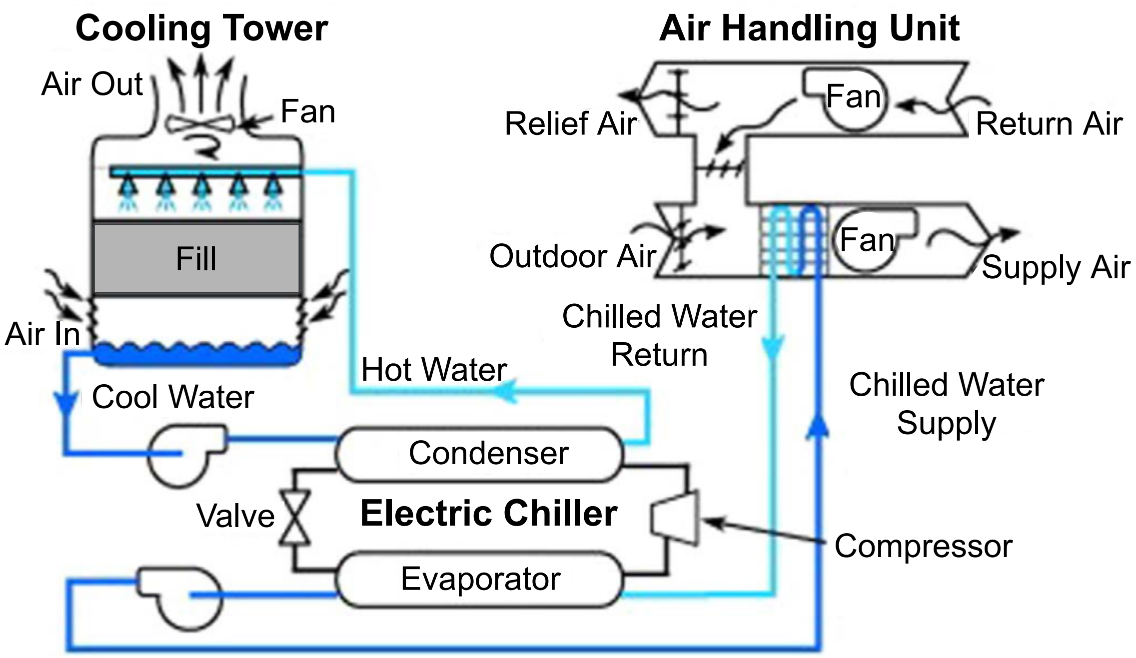


Figure 1: Example of a chilled water ventilation and air conditioning system, adopted from REF [7]

Figure 1 shows an example of an air conditioning and ventilation system as presented by Afroz et al. [7]. The main components are an air handling unit (AHU), chiller and cooling tower. In a zone model, an HVAC zone is defined as a cluster of adjacent spaces that often share a common AHU. The white-box modelling of temperature is often done with the heat conduction equation, heat balance, weighting factor and thermal network method. To calculate energy and mass balance, x specifies the following prerequisites:

- The air is completely mixed and the temperature is equal.
- The influence of opposite walls on the temperature in the zone is the same.
- The floor has no influence on the temperature.
- The air density is constant regardless of the room temperature or humidity.
- There are no pressure losses.

Under these conditions, the energy and loss balance equation, that models the energy changes of a zone, is defined by Afroz et al. [7] as:

$$\begin{aligned} C_z \frac{dT_z}{dt} =& f_{sa} \rho_a C_{pa} (T_{sa} - T_z) + 2U_{w1} A_{w1} (T_{w1} - T_z) + U_R A_R (T_R - T_z) \\ & + 2U_{w2} A_{w2} (T_{w2} - T_z) + q(t) \\ q(t) =& q_p + q_1 \end{aligned} \tag{1}$$

Another formula that can be established deals with the humidity level of the zone and is defined as:

$$V_z \frac{dW_z}{dt} = f_{sa}(W_{sa} - W_z) + \frac{P(t)}{\rho_a} \tag{2}$$

The variables are defined as follows:

Variable	**Meaning**
C_z	Overall thermal capacitance in kJ/°C
T_z	Zone temperature in °C
f_{sa}	Volumetric flow rate of the supply air
ρ_a	Air density of 1.25 kg/m^3
C_{pa}	Specific air heat of 1.005 kJ/kg °C
U_{w1}	Heat transfer coefficient of east and west walls
A_{w1}	East and west wall area in m^2
T_{w1}	East and west wall temperature in °C
U_R	heat transfer of the roof
A_R	Roof area in m^2
T_R	Roof temperature in °C
U_{w2}	Heat transfer coefficient of south and north walls
A_{w2}	South and north wall area in m^2
T_{w2}	North and South wall temperature in °C
$q(t)$	internal heat gain
V_z	Zone volume
W_z	Humidity ration of the zone in kg/kg
W_{sa}	Humidity ratio of supply air in kg/kg
$P(t)$	Evaporation rate of occupants

Table 1: 2.1.1 variable overview

For a complete white-box model, a large number of such equations is constructed.

2.1.2. Black Box Models

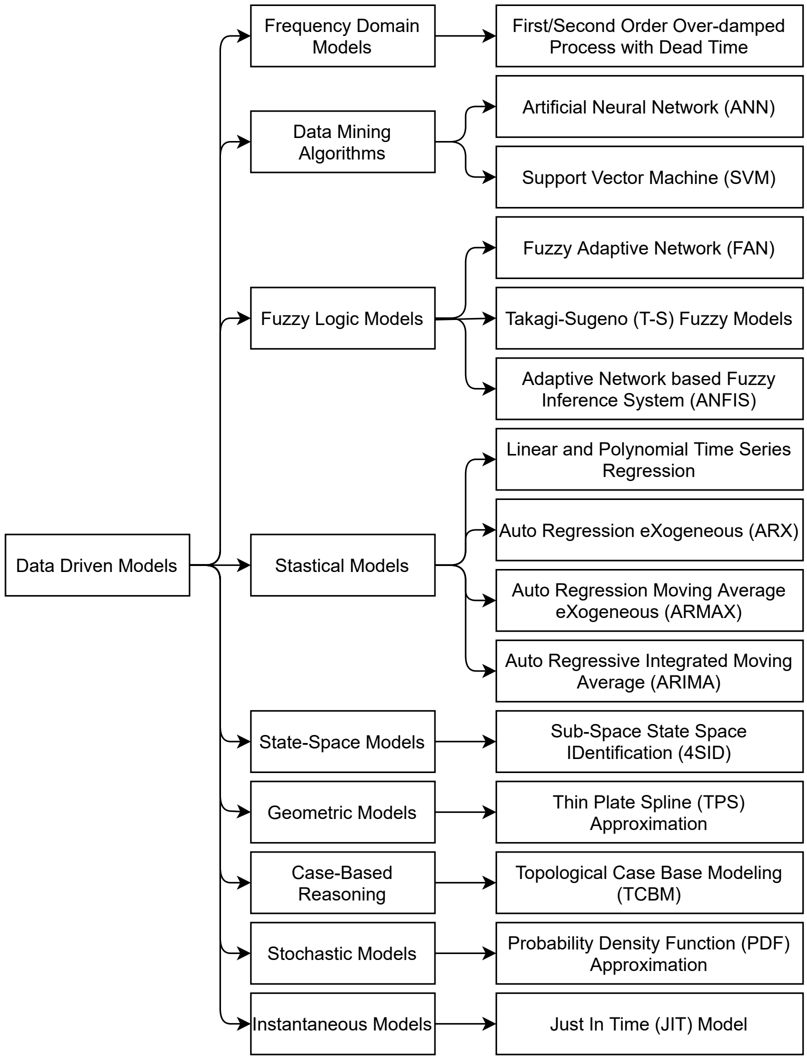


Figure 2: Data-driven modelling types, adopted from REF [5]

Black Box models, also known as data-driven models or inverse models, develop dependencies between input and output parameters based on system performance data [7]. The modelling capabilities of data-driven approaches have been investigated in various researches. They are well suited to improve existing HVAC systems. Figure 2 shows the different categories of data-driven models. Section 2.1.2.2 takes a closer look at the data-driven modelling technique ANN. Of the many categories, we will take a closer look at the stochastic and statistical models.

- **Statistical models**
  Statistical models have a high ease of tuning and can model noisy data [5]. Nevertheless, they have neither robustness for parameters nor for disturbances and a low prediction accuracy. Equation 3 shows a generalized structure of a statistical black box model with simple input and output.

$$\alpha(q^{-1})y(t) = \frac{\beta(q^{-1})}{\theta(q^{-1})}u(t) + \frac{\gamma(q^{-1})}{\delta(q^{-1})}w(t) \tag{3}$$

  Depending on the chosen polynomials, input, output and noise the different statistical models are created (see figure 2).

- **Stochastic models**
  Stochastic models are used to predict random HVAC processes, often using Probability Density Functions (PDF). A large amount of data is necessary for these to deliver reasonable results. Stochastic models have a low ease of tuning, no robustness for disturbances or parameters and a low generalization ability. However, they can model noisy data and, unlike statistical models, they have a high prediction accuracy.

2.1.2.1. Characterization

ANNs are conventionally divided into two categories: supervised and unsupervised learning. [33]

- **Supervised learning**
  The environment provides the desired output for the inputs and thus supervises the learning process as a kind of teacher.
  - **Reinforcement learning**
    The external environment does not give the desired output, but only the assessment for correctness. The learning process strengthens the connections between neurons that are used particularly often.
- **Unsupervised learning**
  The environment does not provide any information about the output. The learning process is based solely on the inputs and the internal updating of the weights, that are divided into clusters and thus get similar results when the input vectors are similar. A new cluster is created whenever a new pattern is discovered.

2.1.2.2. Artificial Neural Networks

The first artificial neuron, designed after neurons of the brain, was presented in 1943 [33]. Artificial neural networks are not based on direct rules and are not strictly programmed. Instead they learn with the help of trial and error processes. Neural networks are used in a wide range of disciplines nowadays, including weather forecasting, aviation and marketing [21].
They are designed to resemble certain properties of neurons in the brain [33]. There are four primary functions that an artificial neuron has adapted from a real neuron.

1. Take in information from an external environment
2. Decide whether this information is relevant or should be ignored
3. Process the data
4. Produce a result

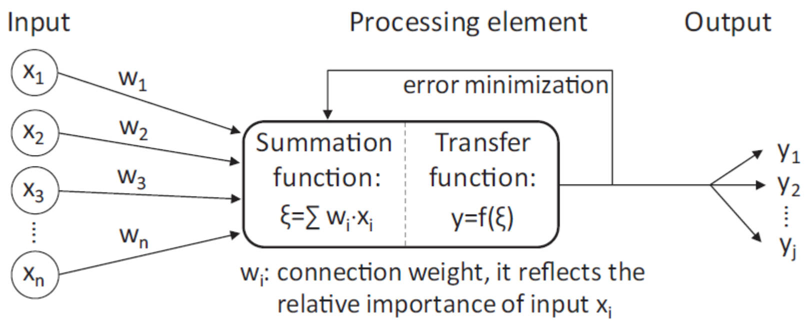


Figure 3: Structure of an artificial neuron. The transfer function represents the activation function. Adopted from REF [33]

Thus, as seen in figure 3, an artificial neuron has a multitude of inputs and a weight for each of these inputs. These are then multiplied and added up (equation 4).

$$\xi = \sum w_i \cdot x_i \tag{4}$$

ANNs are capable of arriving at appropriate results even with non-linear correlations, as well as small, large, incomplete or fluctuating data sets [33].

Activation functions
Activation functions receive the result of the summation function as input and are responsible for deciding whether the input is relevant or not. According to Profillidis and Botzoris [33], there are six standard activation functions, as can be seen in figure 5.
These are, on the one hand, linear activation functions [33]:

$$f(\xi) = a \cdot \xi + b \tag{5}$$

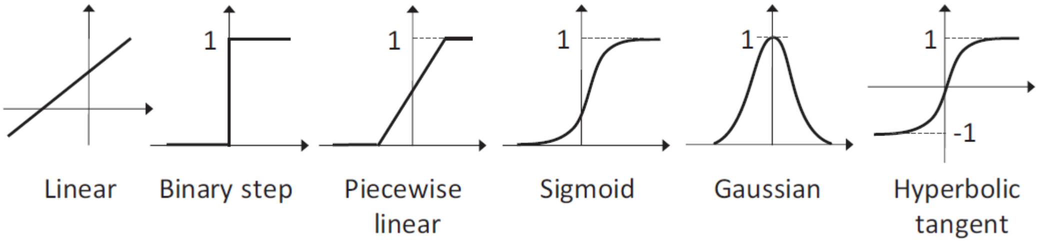


Figure 4: Standard activation functions, adopted from REF [33]

Binary step functions are also called threshold functions and are defined as follows [33]:

$$f(\xi) = \begin{cases} 1 & ,\xi \geq 0, \\ 0 & ,\xi < 0 \end{cases} \tag{6}$$

Another activation function is the piecewise linear function defined by the following formula [33]:

$$f(\xi) = \begin{cases} 1 & ,\xi \geq \xi_{\max}, \\ a \cdot \xi + b & ,\xi_{\min} > \xi > \xi_{\max}, \\ 0 & ,\xi \leq \xi_{\min} \end{cases} \tag{7}$$

The sigmoid function is defined over the interval (0,1) as

$$f(\xi) = \frac{1}{1 + e^{-b \cdot \xi}} \tag{8}$$

while the Gaussian function over the intervall (0,1] is defined as

$$f(\xi) = e^{-\xi^2} \tag{9}$$

Last but not least, the hyperbolic tangent function over the interval [-1, 1] is defined as

$$f(\xi) = \frac{2}{1 + e^{-2 \cdot \xi}} - 1 \tag{10}$$

Additionally there is the **Rectified Linear Units** activation function (ReLU), defined as [8]:

$$f(\xi) = \max(0, \xi) \tag{11}$$

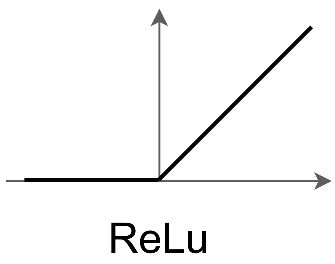


Figure 5: ReLU activation function, adapted from REF [8]

Layers

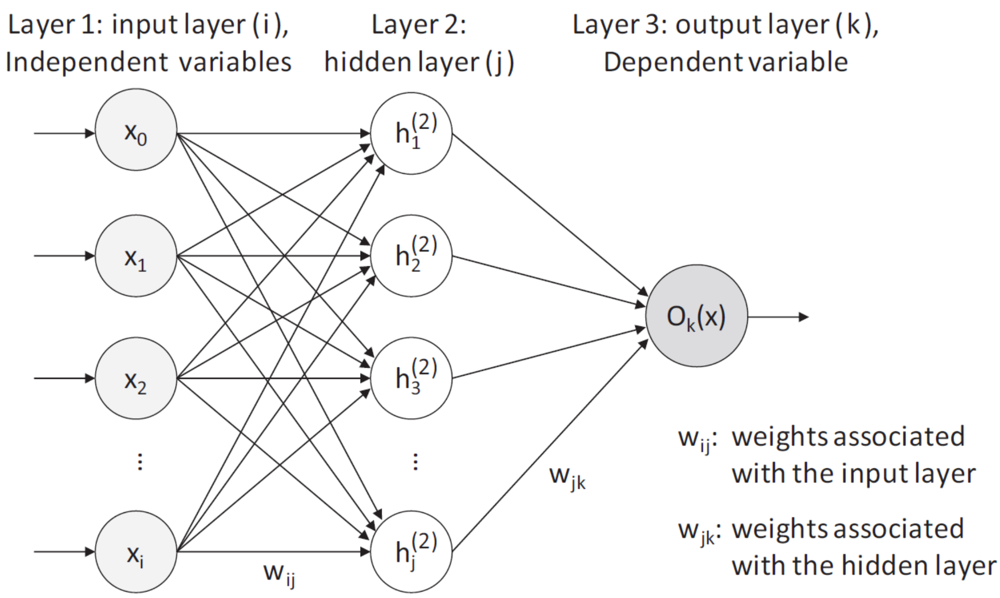


Figure 6: Layer structure of a Feedforward Artificial Neural Network, adopted from REF [33]

- A topology describes which nodes are connected to each other and can be represented by a matrix of the form $n \times n$, where n denotes the number of units [33].
  - The cells list the weights that pass from the unit with the index of the column to the unit with the index of the row.
- Units are sorted into groups called layers depending on certain characteristics.
- Each layer consists of several neurons, with each neuron of one layer connected to each neuron of the next layer [36].
- The structure of the layers of an ANN is shown in figure 6.

- Layers are classified as input, hidden and output layers [33].
  - Nodes of input layers receive data from the external world
  - Nodes of output layers send their results to the outside world
  - If a layer cannot be categorized as either an input or an output layer, they are referred to as hidden layers
- Too many neurons in hidden layers can lead to so-called over-fitting, which means too much adaptation to the training set and thus a lack of generalization. [20]
  - In contrast to the training set, the results of the test set do not give good results.
- Too few neurons lead to under-fitting, which means that the network is too generalized.
  - Neither the training nor the test set provide good results [20].

Architecture Types
There are many different types of ANNs. Generally they can be divided into feedforward and recurrent ANNs [33].

- **Feedforward**
  The coupling is unidirectional, which means that a node to which another node sends its output cannot send back any information [33]. This also applies to indirect transmission via one or more other units. Multi-layer feedforward neural networks (MLFNN) are the most widely used ANNs [36].
- **Recurrent Neural Networks**
  - Recurrent NNs (RNN) contain at least one cycle, which means that the transmission of outputs is bidirectional and a unit to which another unit sends data can send its output to this unit as well as to previous layers [33].
  - The resulting time delay can be specified and can be zero.
  - The structure can be defined via deterministic transitions by the following formula 12, where h defines a hidden state and the exponent represents the layer and the subscript the time unit.
    $$h_t^{l-1}, h_{t-1}^l \rightarrow h_t^l \tag{12}$$

  According to noa [3], there are two ways to build recurrency into MLFNNs:
  - Lead feedback from a hidden layers to the input layer
  - Lead feedback from the output layer to the input layer

  The difference is whether you want to achieve special focus on the input or output values.

- Long-Short-Term Memory Neural Networks (LSTM) are RNNs that are particularly well suited for maintaining dependencies of time sequences over a long distance [26].
- LSTMs have a memory block that stores the current state of the network and controls the flow of information with so-called gates [26].
  These are devided into input gates i_t, forget gates f_t and output gates o_t. Forget gates are responsible for segmentation into subsequences to prevent overloading of the sequences. At each time step, it is necessary to decide whether the stored information should be retrieved, retained or overwritten.
- The information is stored in memory cells $m_t^l \in \mathbb{R}^n$, where l and t are defined as in the transition definition of RNNs (equation 12) [26].
- In contrast to conventional RNNs, the transitions of LSTMs are defined as [26]:

$$h_t^{l-1}, h_{t-1}^l, m_{t-1}^l \rightarrow h_t^l, m_t^l \tag{13}$$

2.1.2.3. Multi-Output Routine

Multi-output (MO) ANNs are ANNs that have more than one neuron in the output layer. The network can therefore calculate more than one output of different types at the same time. [38]

Challenges
The challenges of multi-outputs are primarily divided into the four Vs: volume, velocity, variety and veracity, although these originally refer to inputs. [38]

- **Volume**
  This term denotes a large increase in output labels, that can lead to problems such as scalability problems, insufficient annotations or label imbalance.

- **Velocity**
  Fast acquiring of output labels and the associated problems such as concept drift and changes in output distributions.

- **Variety**
  Possibly different structure of the output labels, that leads to difficulties in modelling dependencies between the different outputs and finding a suitable loss function.

- **Veracity**
  Problem of quality differences in output labels, that is characterized by noise or incomplete data.

Goal
Given an input space of the form $\mathcal{X} = \mathbb{R}^d$ and an output space of the form $\mathcal{Y} = \mathbb{R}^m$, the goal of multi-output learning is to find a function $F : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ with the training set $\mathcal{D} = \{(x_i, y_i) \mid 1 \leq i \leq n\}$, where $F(x, y)$ denotes the compatibility of x and y. [38]

The output with the highest compatibility value is then calculated for the test set:

$$f(x) = \tilde{y} = argmax_{y \in \mathcal{Y}} F(x, y) \tag{14}$$

Subfields
According to Xu et al. [38] MO learning has the main subfields Multi-label Learning, Multi-target Regression, Label Distribution Learning, Label Ranking, Sequence Alignment Learning, Network Analysis, Data Generation Semantic Retrieval and Time-series Prediction.
For this report the most relevant subfields are Multi-target regression, that focuses on predicting multiple output values per instance, and Time-series Prediction, since it focuses on predicting future values based on previous observations.

- **Multi-target regression**
  The labels of each instance are represented by a vector that indicates the association of the instance to each label [38]. An unknown instance $x \in \mathcal{X}$ is then mapped to a vector $f(x) \in \mathcal{Y}$ using the regression function $f(\cdot)$.

- **Time-series Prediction**
  This MO method has as input data of a time period and calculates as output a data vector for a point in time in the future [38]. The output at time t is defined as $y_i^t \in \mathcal{Y} = \mathbb{R}^m$, with the output for the time span from $t = 0$ to $t = T$ defined as the vector $y_i = (y_i^0, ..., y_i^t, ..., y_i^T)$.

Evaluation
There are different approaches to assessing an MO-model. According to Xu et al. [38] these include Hamming loss, macro- and micro-averaging, one-error, ranking-loss, average precision, mean absolute error, mean squared error, average correlation coefficient and Intersection over union threshold. In the following we will focus on metrics for regression problems.

- **Mean Absolute Error (MAE)**
  MAE calculates the absolute difference between the calculated and the actual result [38]. Since MAE is usually defined for single outputs (SO), it must be extended to iterate over all outputs.

$$MAE = \frac{1}{m} \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{15}$$

- **Mean Squared Error (MSE)**
MSE is originally defined for SO ANNs as well, but in contrast to MAE it focuses on strong deviations and thus yields larger values [38]. As MAE it can also be extended to MO settings.

$$MSE = \frac{1}{m}\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{16}$$

- **Average correlation coefficient (ACC)**
This metric calculates the strength of the association of the calculated and actual output [38]. Let y_i^m and $\hat{y}_i^m$ be the real and calculated m outputs of input x_i and $\overline{y}^l$ and $\overline{\hat{y}}^l$ the real and calculated outputs for label l over all samples. Then ACC is defined as:

$$ACC = \frac{1}{m}\sum_{l=1}^{m}\frac{\sum_{i=1}^{n}(y_i^l - \overline{y}^l)(\hat{y}_i^l - \overline{\hat{y}}^l)}{\sqrt{\sum_{i=1}^{n}(y_i^l - \overline{y}^l)^2\sum_{i=1}^{n}(\hat{y}_i^l - \overline{\hat{y}}^l)^2}} \tag{17}$$

2.1.3. Grey Box Models

Grey Box models are a combination of White Box and Black Box models [7]. The structure is adopted from the physically-based models, while the parameters are data-based [5]. Due to the physical model, the grey-box model has more general results than a data-driven model. At the same time, it is more accurate than a physics-based model as the data-based parameters also capture the effects of non-modelled dynamics.

2.2. Model predictive control

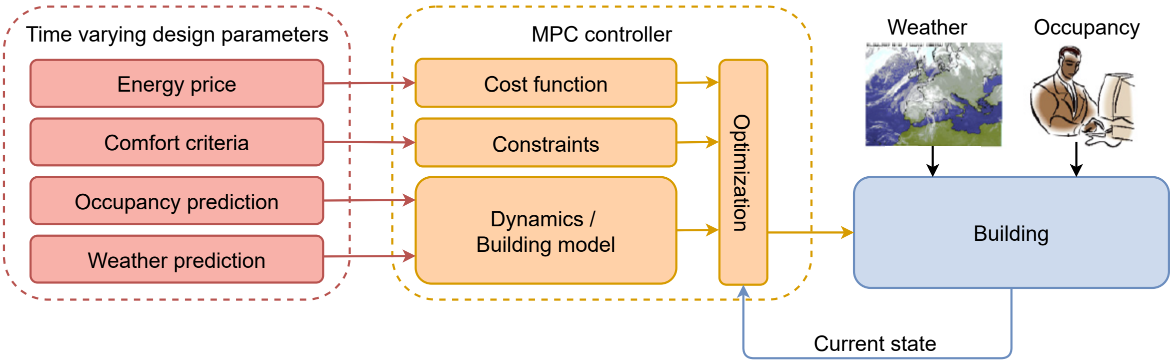


Figure 7: Basic principle of MPC according to REF Široký et al. [40]

- Model Predictive Control (MPC) is a class of control methods that emerged in the late seventies to early eighties [40].

- MPC predicts the future states of a system based on a mathematical model of that system [4].
- It generates a control vector to minimize a cost function over a finite prediction horizon, taking into account disturbances and constraints [4].

In the following we will take a closer look at the modeling and optimization in MPC. In section 2.2.2 we will deal with the optimization of buildings and thus with cost functions and optimization methods. Figure 7 depicts the basic principle of MPC.

2.2.1. Modeling

Modelling in MPC is divided into white box, grey box and black box modelling as in section 2.1. For MPC, however, a distinction must be made between the modelling of the building envelope, the disturbances and the system itself.

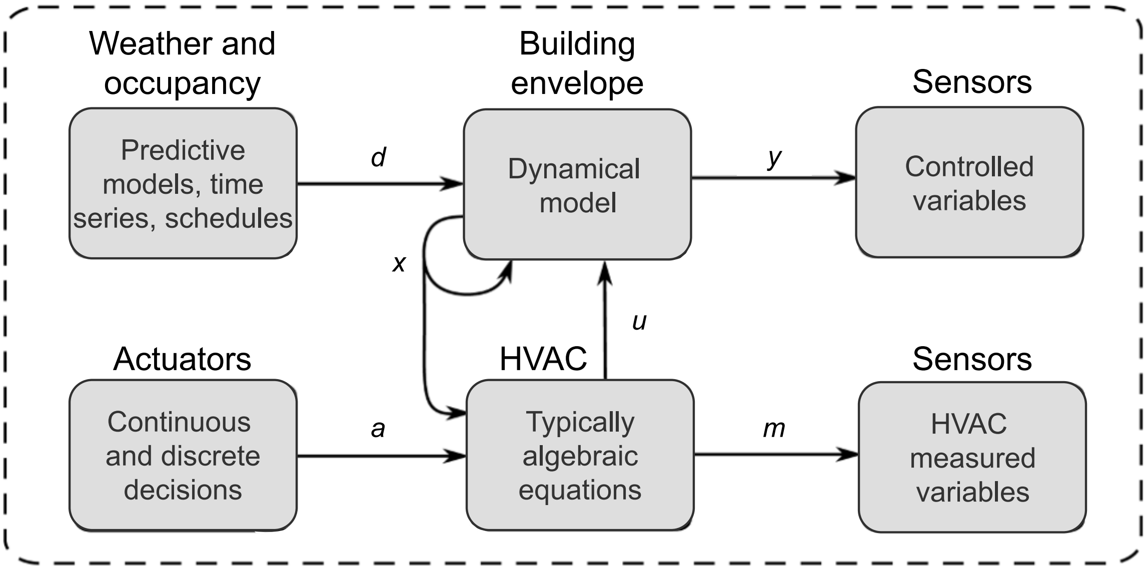


Figure 8: Structure of MPC building modeling, adopted from REF [12]

Figure 8 shows the interaction of the individual components to be modelled, where d represents the disturbances, y the outputs, x the state values, u the inputs of the building envelope, a the actuators and m additional relevant variables. These variables are considered in more detail in section 2.2.2 for the cost function.

Building envelope

- Envelope refers to all separations between different thermal zones or from the building to the outside, e.g. walls, windows, floors or ceilings [12].
- Building envelope models should model all heat transfers through the separations, and in particular take into account conduction, radiation and convection [12].

According to Drgoňa et al. [12], the conduction heat transfer is typically calculated with the 1D transient heat equation. This is defined in [18] as follows:

$$\dot{q} = kA\frac{dT}{dx} \tag{18}$$

Here $\dot{q}$ indicates the heat transfer rate in W, k the thermal conductivity in $\frac{W}{m-K}$, A the area normal to the heat flow in m^2 and $\frac{dT}{dx}$ the temperature gradient in $\frac{°C}{m}$.

- For radiation and convection models many different approaches are widely applied [12].
- In black-box models envelope is only included indirectly by calculating a function over indoor and outdoor disturbances.

Disturbance model
Another factor to be modelled are disturbances, i.e. all factors that cannot be actively controlled but have an influence on the result, e.g. weather forecasts or occupancy. [12]

- In order to include weather disturbances, either an existing service provider can be used or a machine learning (ML) model can be trained in this respect.
  - As weather stations are often located further away from the buildings under consideration, the evaluations in the first case may not reflect the actual weather conditions around the building.

For example, to model the uncertainty due to outdoor temperature, Mantovani and Ferrarini [24] present a model that includes three days of historical data:

$$T(t) = \alpha(t) \cdot \sin(\frac{2 \cdot \pi}{24 \cdot 60 \cdot 60} \cdot t - \frac{\pi}{2}) + T_{oa,avg}(t) \tag{19}$$

Here $\alpha(t)$ is the time-variant amplitude profile and $T_{oa,avg}(t)$ is the time-variant average value profile at time t.

Regulation System
The modelling of the HVAC system can be accomplished in different forms and has a high level of complexity. [12]

- In some cases, each component is controlled by the MPC controller, while in other cases MPC controls higher-level interfaces.
- Since HVAC systems can have a number of different components that vary in their properties to be modelled, the complexity is increased. For example, pumps and fans have non-linear characteristics.
- Dampers, amongst many others, are active components, while pipes are static components.

To resolve this complexity, an optimization problem must be solved.

2.2.2. Optimization

Cost functions

Cost functions or objective functions are used to specify the minimization constraints of the optimization problem. According to Drgoňa et al. [12], the cost function in the general MPC optimization problem is defined as equation 20a:

$$\min_{u_0,\dots,u_{N-1}} \ell_N(x_N) + \sum_{k=0}^{N-1} \ell_k(x_k, y_k, r_k, u_k, s_k) \tag{20a}$$

$$\text{s.t. } x_{k+1} = f(x_k, u_k, d_k) \tag{20b}$$

$$x_0 = \hat{x}(t) \tag{20c}$$

$$y_k = g(x_k, u_k, d_k) \tag{20d}$$

$$r_k = r(t + kT_s) \tag{20e}$$

$$u_k = f_{HVAC}(x_k, a_k, m_k) \tag{20f}$$

$$s_k = h(x_k, y_k, u_k, r_k) \tag{20g}$$

$$d_k = d(t + kT_s) \tag{20h}$$

$$x_k \in \mathcal{X}, u_k \in \mathcal{U}, a_k \in \mathcal{A}, s_k \in \mathcal{S}, k \in \mathbb{N}_0^{N-1} \tag{20i}$$

Here $\ell_N(x_N)$ stands for the omittable penalty of the cost function, $\ell(x_k, y_k, r_k, u_k, s_k)$ for the stage cost that assigns costs to the values, x_k stands for the values of states, y_k for the outputs, r_k for reference signals, u_k for the inputs of the building envelope, s_k for slack variables, d_k for the disturbances, a_k for the HVAC actuators and m_k for additional measured variable values in the kth step in the prediction horizon N with sampling time T_s. x_0 indicates the initial states of the state variables (20c). f predicts the state values (20b), g the outputs (20d), r the reference signals (20e) and d the disturbances (20h). h calculates violations of further criteria, such as thermal comfort (20g).

- Examples of minimization targets are minimization of costs or green house gases [12].
- Minimization problems can also be transformed into maximization problems
  - e.g. maximizing the use of renewable energy.
- All these factors can be objectives of the cost function.
- If one wants to fulfil several objectives at once, one speaks of so-called multi-objective optimization, for which there are several optimization approaches.

Optimization method

Optimization methods include goal attainment, minimax and Pareto front. [12]

- **Goal attainment**
  - For multi-objective problems, goal attainment is the most commonly used approach.
  - The objectives are each assigned a weight and the result is determined depending on these weights.
  - e.g. the cost function (equation 20a) can be reformulated as follows:

$$\min_{u_0,\dots,u_{N-1}} \sum_{k=0}^{N-1} (||Q_s s_k||_2^2 + Q_u k_k u_k) \tag{21}$$

  - Q_s and Q_u respectively give weighted matrices for the slack and envelope variables, while k_k is responsible for the time reference.
  - Taking time variation into account is relevant, for example, in the case of price changes.
- **Minmax**
  - In contrast to goal attainment, the focus of the minmax approach is on keeping the worst case as good as possible.
  - Cost functions that are established with the minmax method are particularly suitable for safe solutions under unpredictable effects.
- **Pareto front**
  - With the pareto front method, the relationship between several objectives is constructed in such a way that every minimization or maximization of one objective leads to the opposite in the other objective.

Constraints

MPC is also called constrained control because it can find a reasonable solution for many different types of constraints, both for input, output and states [4].
These constraints can be roughly divided into two categories [4, 12]:

- inequality constraints, such as control input range
- equality constraints, such as capacity limits and building model dynamics

Constraints are also divided into hard and soft constraints [12]:

- Hard constraints are constraints that have to be complied with at all times in the entire prediction horizon.

- e.g. control action bounds, given by the following equation:

$$\underline{u} \leq u_k \leq \overline{u} \tag{22}$$

- u_k are the building envelope inputs, while $\underline{u}$ and $\overline{u}$ depict the lower and upper limits.

- Soft constraints are constraints for which infringements are tolerated.
  - For this purpose slack variables s_k are penalised in the cost function.
  - e.g. thermal comfort inequality constraints, given in equation 23.

$$\underline{y_k} - s_k \leq y_k \leq \overline{y_k} + s_k \tag{23}$$

  - y_k depicts the output, while $\underline{y_k}$ and $\overline{y_k}$ define the lower and upper limits.

- Thermal comfort inequality constraints (equation 23) are also an example for time-varying constraints.
  - Change over time, unlike constant constraints, because $\underline{y_k}$ and $\overline{y_k}$ are defined as time-varying

- Slew rate constraints penalise the degree of a change in certain variables to avoid overshooting and peak behaviour.
  - Equation 24 restricts the scope for change of the building envelope inputs within one step.

$$\Delta u_k = u_k - u_{k-1} \tag{24a}$$

$$\underline{\Delta u} \leq \Delta u_k \leq \overline{\Delta u} \tag{24b}$$

Table 2 shows an overview of the different forms of constraints and their relevant distinctions.

Oldewurtel et al. [31] subdivide constraints into the categories linear, convex quadratic, chance, second order cone, switched and nonlinear.

- For linear constraints, which are the most commonly used, an upper and lower bound is defined, while convex quadratic constraints are bounded by an ellipsoid.
- Chance constraints set a satisfiability probability for constraints under uncertainty and can possibly result in second order cone constraints, of which the mapping space is cone-shaped.
- Switch constraints switch between constraints depending on fulfilled or unfulfilled conditions.
- Non-linear constraints are all constraints that do not fit into one of the other categories and are represented by any non-linear function.

Form	Violations	Time	Math	Variables	Meaning
Inequality	Soft	Varying	Linear	States	Model dynamics
Equality	Hard	Constant	Nonlinear	Outputs	Ranges
			Mixed-integer	Inputs	Slew rate
					Move blocking
					Terminal

Table 2: Overview of types of constraints, adopted from REF [12]

2.2.3. Conclusion

In summary, MPC provides a good opportunity for building control. It takes into account a variety of factors, such as the building structure, possible disturbances, weather and other constraints. To accommodate these as well as possible, there are many different modelling options, including both linear and non-linear functions. There are a variety of different modelling options to best fit the given conditions and specific objectives are established.

3. Data

3.1. Data description

To test the models, building data is provided as CSV files by a real commercial company that optimizes the building climate and reduces costs and CO2 emissions. Given is data from one zone, each of 23 different buildings, numbered from 0 to 18 and from 20 to 23. The data was measured 24 hours a day, seven days a week every 15 minutes for a different number of years, but in some cases only entered when changes occurred. The goal is to automatically train models for all control data of the data sets and compare the results in terms of differences between buildings and between single- and multi-output models. The following functions were implemented using *Python* (Version 3.8.5).

3.2. Parameter selection

Since the buildings provide different data, the parameter selection must be handled flexibly. For this, the column headers are scanned for keywords and used as output if one is included. These keywords are:

- Climate_sensor_RBG_basic_setpoint_room_temperature
- Climate_sensor_Presence
- CTL
- Flow_temperature
- Massflow
- ventilation_system_office_volume_flow_release
- Individual_room_control_Heating_circuit_Room_temperature_Basic_set_point
- Ventilation_system_supply_air_temperature
- Individual_room_control_heating_circuit_Operating_mode
- Cooling_circuit_Component_activation_Return_temperature_Switch-on_temperature
- ventilation_system_RLT_supply_air_temperature
- ventilation_system_volume_flow_release
- Ventilation_system_volume_flow_release
- Ventilation_system_supply_air_temperature

For some of the data, both the actual and the predicted values are provided. Since in practice only the predicted values are available, the actual values are not taken into account.
For other use cases, however, it could be useful to select whether the actual or the predicted value is used, depending on the correlation. The Pearson Correlation Coefficient (PCC) (r_{xy}) indicates the correlation between two variables and is defined as follows [10]:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{25}$$

For this purpose, the average correlation of both values can be compared for each building. For example, by calculating the difference of the absolute averages, when F is the set of features with the word forecast, N is the set of features of the associated real values and O is the set of outputs:

$$d = |r_{fo}| - |r_{no}| \tag{26a}$$
$$f \in F, n \in N, o \in O \tag{26b}$$

An exemplary illustration for zone 0 with the consideration of 1 year of data can be seen in figure 9.

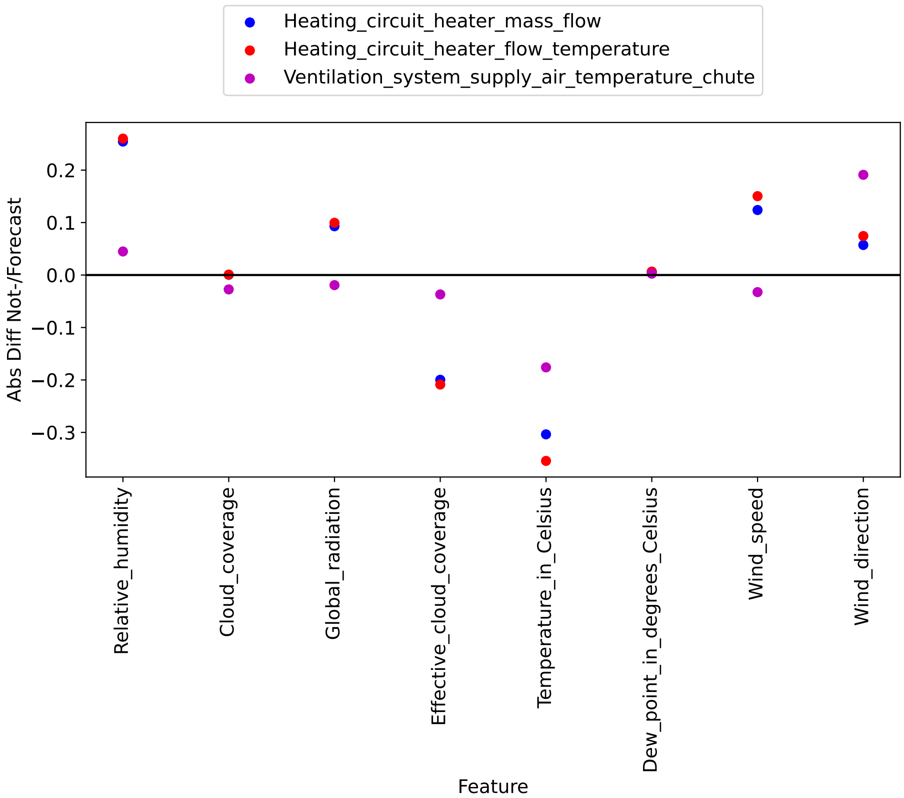


Figure 9: Comparison of correlation from weather -/forecasts to outputs in zone 0 with 1 year of data

Zone 0 has three outputs and real values as well as predictions on relative humidity, cloud coverage, global radiation, effective cloud coverage, temperature, dew point, wind speed and wind direction. If d (equation 26a) is positive, the absolute correlation between the output and the forecast is higher than between the output and the real value.
Thus, for Zone 0, it makes more sense to use the forecast data for relative humidity, whereas for temperature, the real weather data is more useful. This allows an assessment of whether forecasts or actual values should be used for all weather data, unless a standard is generally predefined.

3.3. Data preprocessing

The company provides a CSV file per building in which the data is separated by ','.

- Since in some cases the data sets were only updated when a change occurred, empty cells are replaced by the last previous filled cell.
- Columns that contain no value or only a constant value are deleted, as they have no influence on the result.
- Since changes may occur over time, data that is too old may worsen results. Therefore only relatively recent data should be considered and data that is at least 3 years older than the most recent data is deleted.
- The date, given in date time format, is divided into separate columns for year, month, day, hour and minute.
- All rows that still have an empty cell are deleted.

3.4. Data correlation

To visualize the correlation between the individual parameters, heatmaps are created. To create a heatmap, a correlation matrix is first generated from the data. A correlation matrix indicates in each cell how high the PCC is. A heatmap maps each PCC value to a colour. Figure 10 gives an example of a heatmap for zone 0. The darker the colour in a cell, the closer the correlation is to 1. Since the correlation of a column to itself is always 1, the diagonal is generally the darkest shade. Due to the fact that the correlation to the year is generally rather low, especially when looking at a few years, the column for the year was removed before the creation of the heatmap.
The colour scale for the evaluation can be found in appendix F in figure 100. Figures 101 - 122 illustrate heatmaps for all zones.

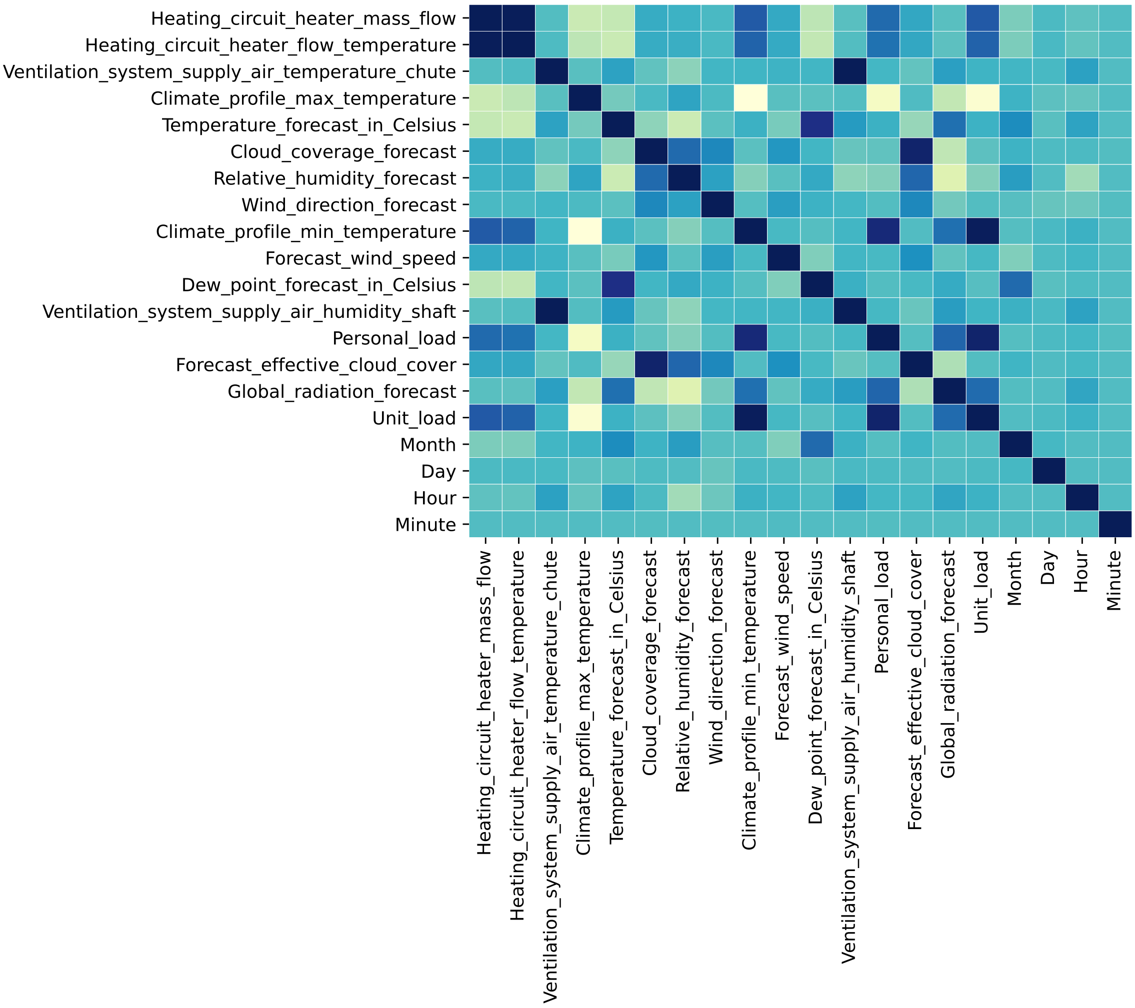


Figure 10: Visualization of Correlations between data for zone 0

- To improve the performance of the model, inputs with a low correlation to the outputs can be deleted.
- The inputs I that have a correlation greater than the minimum correlation r_{min} to at least one output O were retained.

$$i \in I \Rightarrow \exists o \in O : |r_{io}| \geq r_{min} \tag{27}$$

- To keep an sufficient amount of inputs, the minimum correlations 0.05 and 0.1 are tested in chapter 4 in the Case Study.

4. Case study

4.1. Configurations

Tensorflow (version 2.4.1) and especially *Keras* (version 2.4.0) were used to implement the models. The graphics were created using *Matplotlib* (version 3.4.1) and *Seaborn* (version 0.11.0). In addition, *Sklearn* (version 0.23.2), *Numpy* (version 1.19.2) and *Pandas* (version 1.1.3) were used.
The model used was a feedforward model with four hidden layers, where all layers are fully interconnected, as it was given by the company. The first layer has 12, the second layer 24, the third layer 12 and the fourth layer two nodes. As activation function ReLU (equation 11) is used and the loss is calculated with the MSE loss function (equation 16). This model structure is trained once as MO for all outputs (figure 25) and once per output as SO (figure 26). Examples of these structures are portrayed in appendix A in figures 25 and 26. These models are trained for each combination of the amount of years (one vs three) and the preprocessing method (Drop not-forecast data or additionally correlation less than 0.05 or less than 0.1) used.

4.2. Graphical User Interface

To make the selection of preprocessing and outputs as flexible and easy to use as possible, a Graphical User Interface (GUI) was developed (Figure 11).

Model settings

- Option to import an arbitrary CSV file.
- Specification options for the percentage of test data, number of years and minimal correlation to consider for the training.
- A dropdown menu to select how weather data is handled. The choice is between using the real or predicted weather data or keeping both.
- Checklist and manual input options to select the strings to be searched for as outputs.

Model training

- Choice between MO and SO model.
- Dropdown menu that lists all found outputs to select an output.
  - MO: text over dropdown menu: 'to plot results:'
    Used to select output to plot afterwards
  - SO: text over dropdown menu: 'Choose output to train:'
    Used to select output to train SO model

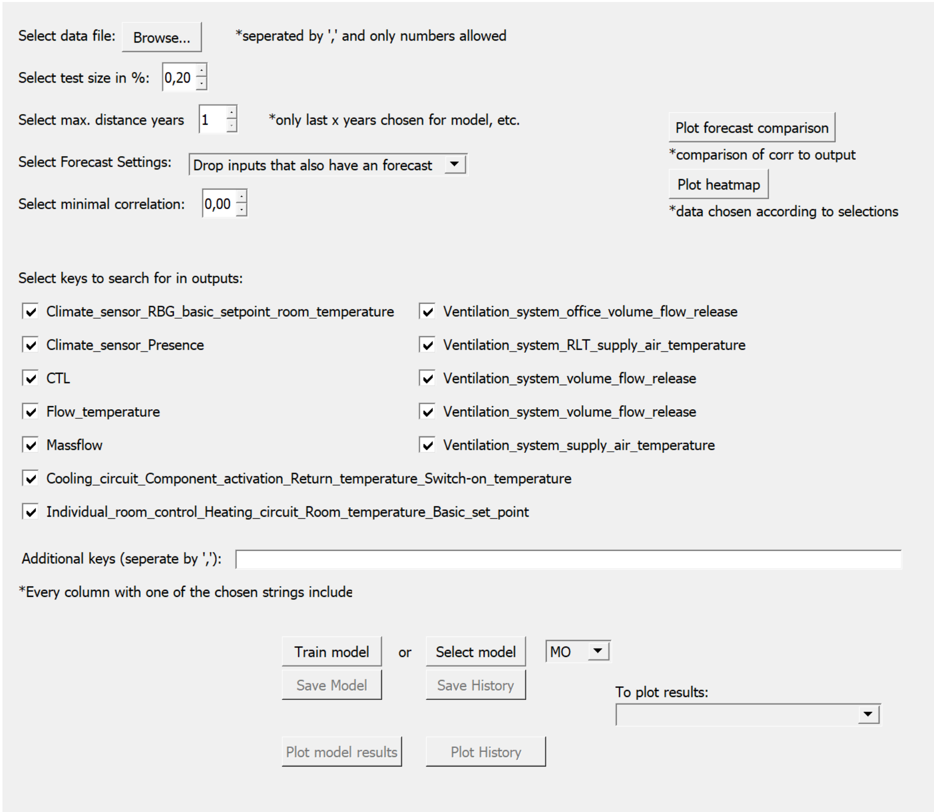


Figure 11: Graphical User Interface

- Options to train new or import existing model.
- When model is trained, button to save model and history is activated.

Model visualization

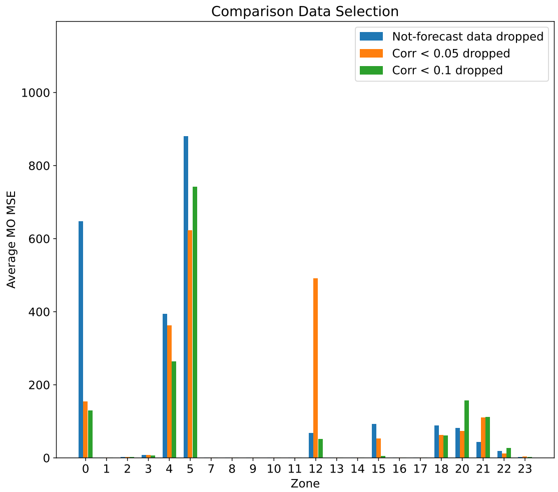
- If something is to be plotted, a new window opens with the figure and a save button.
- Buttons are provided to assist with the model settings. They are deactivated if no keys are selected.
  Heatmap to get an impression of the correlations.
  Graphics for comparing correlations of weather forecasts and -data (3.2).
- The button to plot results is activated after training or choosing a model.
- An output for the result plot is chosen from the dropdown menu.
- The button to plot the history is only activated when the model was trained.
- If the model settings are changed after a model has been selected or trained, the buttons are deactivated again because the GUI assumes that the model no longer matches the settings. After training or selecting another model, the buttons are reactivated.

4.3. Results

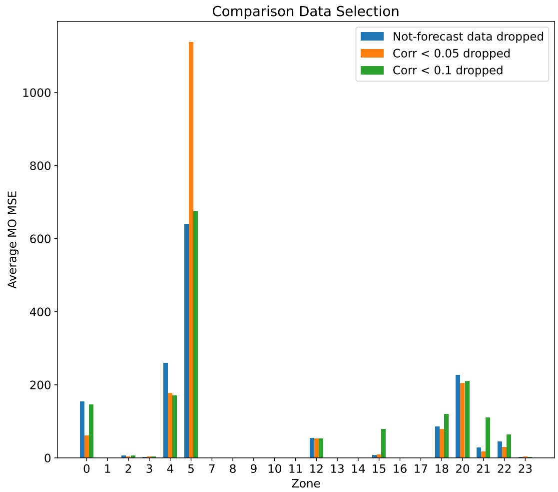
In the following sections we will compare different tests on the model results. All outputs of zone 6 are constant, thus this zone will not be considered for these evaluations.

4.3.1. Test 1: Preprocessing comparison

In order to test whether one model is suitable for different zones, it would be ideal if a preprocessing method can also produce optimal results with this model.
For this reason we first compare the losses of the different preprocessing options in graph 12.



(a) 3 years



(b) 1 year

Figure 12: Comparison of average MSE over all outputs of all zones of all preprocessing options

Description

- Figure 12 shows the average MO MSE values over all outputs with linear y-scale for one and three years of data.
- For each zone there are three bars, one per preprocessing method.
- Since a higher MSE value indicates a higher loss, the higher the bar is the worse are the model results.
- The bars vary considerably, with one bar being higher than the others in some cases. In figure 12a, the blue bar in zone 5 has a significant deflection, while in figure 12b, the orange bar of zone 5 deviates strongly from the others.

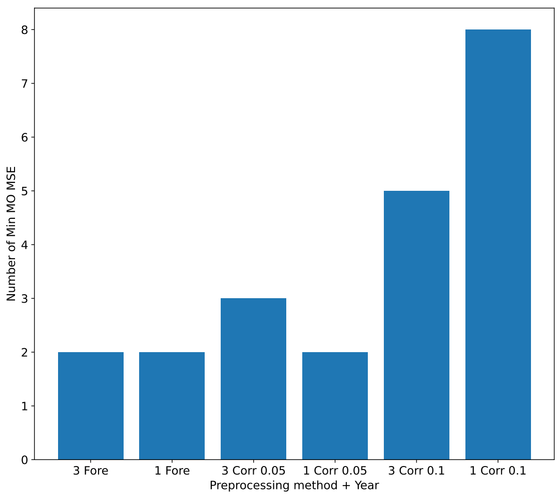
Analysis

- Especially zone 5 has rather high MSE values for every data selection type we use.
- There are zones, that produce really good results for each data selection type.
- Considering three years produces more high MSE values.
- Considering one year produces higher MSE values for especially zone 5.

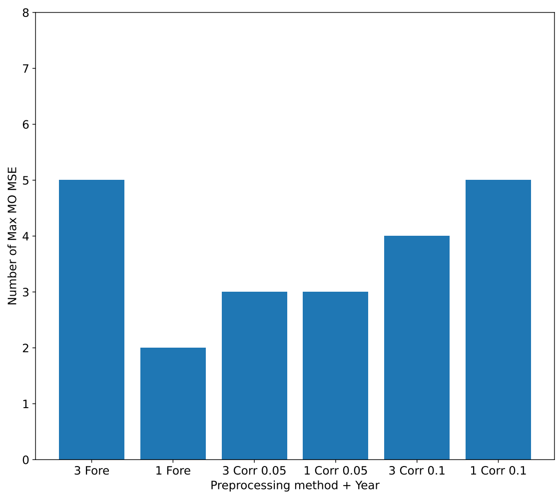
Conclusion

- The model is presumably not applicable for zone 5.
- There is no preprocessing method, that produces good results for every zone.

Since none of the preprocessing methods in graph 12 perform particularly better than the others, we will take a look at how often which preprocessing method produces the best or worst results (figure 13).



(a) Min



(b) Max

Figure 13: Number of times per preprocessing method the MSE value is minimal/maximal

Description

- The number of times each preprocessing method in combination with the amount of data produces the minimum MSE values is shown in figure 13a.
- In Figure 13b, the number of times each preprocessing method combined with the quantity of data produces the maximum MSE values is shown.

- Each preprocessing combination produces the best results at least two times and a maximum of eight times.
- The number of maximum MSE values lies between two and five.

Analysis

- Choosing the data of one year in combination of dropping data with a correlation less than 0.1 most often achieves the best results.
- One year with Corr < 0.1 and three years with dropping not-forecast data produce the worst results most of the time.
- Dropping a correlation less than 0.1 is the only preprocessing technique, that produces the best results more often than the worst.
- One year with dropping non-forecast data and three years with corr < 0.05 show the best results as often as the worst.
- Three years with dropping non-forecast data and one year with corr < 0.05 result in the worst MSE values for more zones than in the best values.

Conclusion

- The combination of one year and dropping data with a correlation less than 0.1 shows the best ratio of best and worst MSE values.
- In the following tests, the focus will be on models that use one year of data and dropped columns with less than 0.1 correlation to all outputs.

4.3.2. Test 2: MO vs SO model comparison

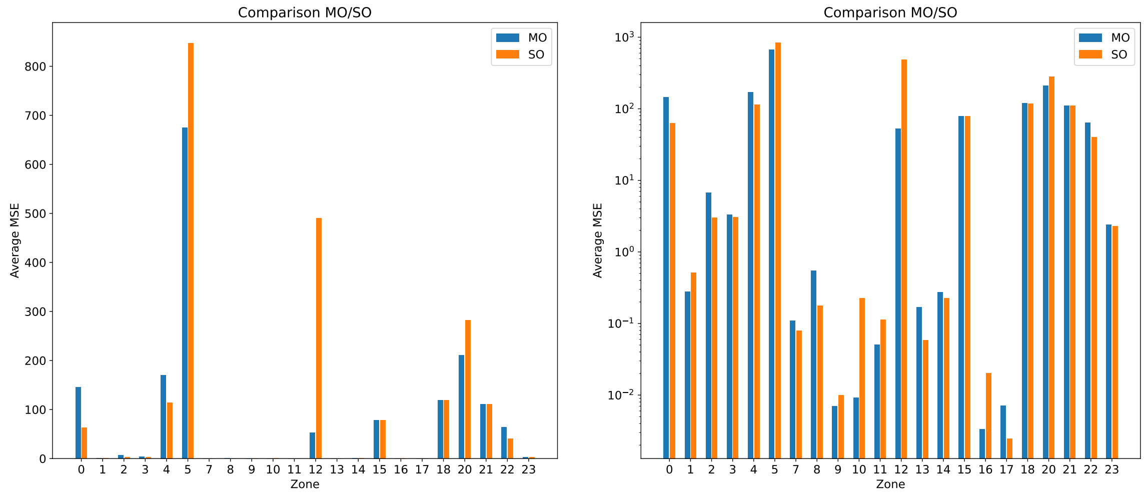
The assessment of the MO models depends on the results of the SO models. Test 2 aims to determine how the MO models perform compared to the SO models, given one year of data and a correlation ≥ 0.1 to at least one output as the result of test 1.

Description

- The average MSE values for the MO and SO models of all zones is shown in figure 14. For all other preprocessing options considered, there are equivalent figures in appendix C.
- Figure 14b shows the plot of 14a with a logarithmic y-scale, to visualize the differences between the MO and SO results, even for especially low MSE values.

Analysis

- For both MO and SO the worst results are visible in zone 5.



(a) Linear scale

(b) Logarithmic scale

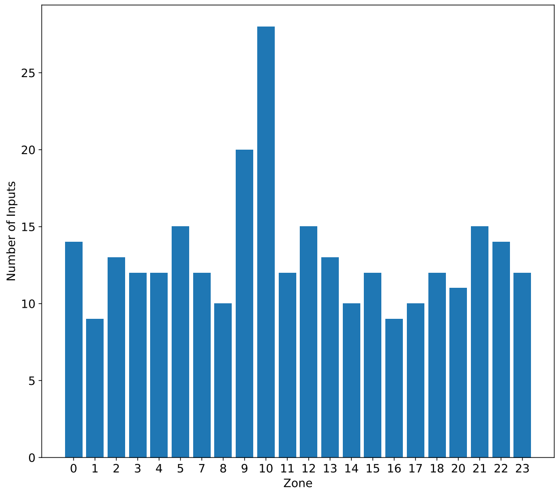
Figure 14: Comparison of MO and SO average MSE values over all outputs of all zones using 1 year of data and dropping columns with a correlation less than 0.1 to all outputs

- In zones 0, 2, 4, 8 and 22 the SO models achieved significantly better results than the MO models.
- In zones 1, 5, 10, 12 and 20 the MO models achieved significantly better results than the SO models.
- Zones 3, 7, 9, 11, 13, 14, 15, 16, 17, 18, 21 and 23 show similar results regarding MO and SO MSE values.
- Especially zone 12 shows a large improvement in the MO models.

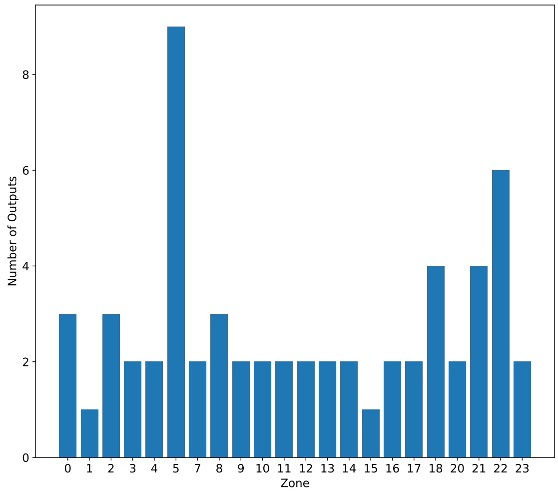
Conclusion
Since in most zones the losses are very similar, sometimes even better with MO, converting to MO would be useful for this preprocessing combination.

4.3.3. Test 3: Zone MO comparison

In order to investigate the cause of the strong deviation of the results, we compare in test 3 a zone with particularly good and a zone with particularly poor results in the models that require one year of data and a correlation over 0.1.
For this purpose, we first look at the number of inputs and outputs as well as the share of the outputs in the data. Equivalent figures for other preprocessing methods are shown in appendix B.



(a) Number of inputs



(b) Number of outputs

Figure 15: Number of in-/outputs when considering one year and dropping data with a correlation less than 0.1 to all outputs

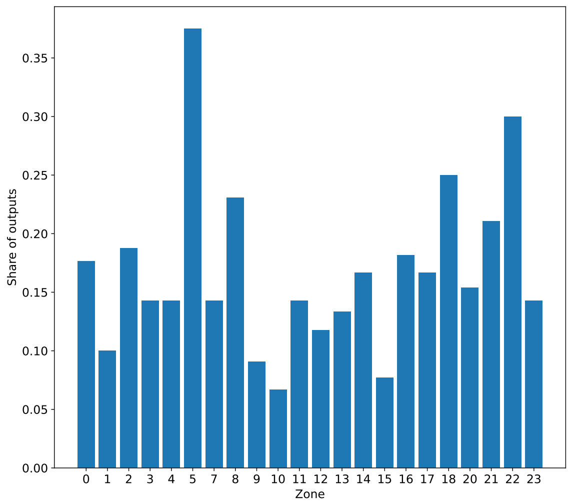


Figure 16: Percentage of outputs in data

Description

- Figure 15a depicts the number of inputs of each zone. In most cases, the number is between 8 and 15, with deviations in zones 9 and 10.
- The number of outputs is shown in figure 15b. The median is 2 with little variation in some zones. Especially zones 5 and 22 deviate strongly upwards.
- The share of outputs in the data is shown by figure 16. There is considerable variation in the height of the bars.

Analysis

- Zone 16 is one of the zones with the fewest inputs.
- Zone 5 is a medium zone in terms of the number of inputs.
- The number of outputs in zone 16 average.
- Zone 5 has by far the most outputs.
- About one fifth of the data in zone 16 are outputs.
- More than one third of the data in zone 5 are outputs.
- A theory could be that Zone 5 could address **Volume** of the challenges considered in section 2.1.2.3, but since the SO results are even worse for this zone, this theory is most likely not the underlying issue.

Conclusion

- Both the number of outputs and the percentage of outputs in the data seem to affect the outcome.
- It appears that the number of inputs has no significant influence as such.

Since we have only considered the average MSE values so far, we now examine the MSE values of the individual outputs to assess whether the high or low MSE value is produced by all or several or only one output.

Description

- Graphs 17 and 18 show the MSE value for each output of the selected zones and models. Equivalent plots for all other zones can be found in Appendix D.
- Zone 5 has nine outputs and for each of these outputs the MO and SO MSE values are represented by bars.
- In the case of zone 16, there are two outputs and the MSE value of the MO and the SO model is shown.

Analysis

- Some outputs of zone 5 belong to ventilation_system_volume_flow, but although these should therefore represent similar data, 1a and 2c give good results but 2b and 2a give poor results.
- In most cases of zone 5, SO and MO have either both high or both low MSE values. The exception is Ventilation_system_RLT_volume_flow_release, where the SO MSE is significantly less than the MO MSE.

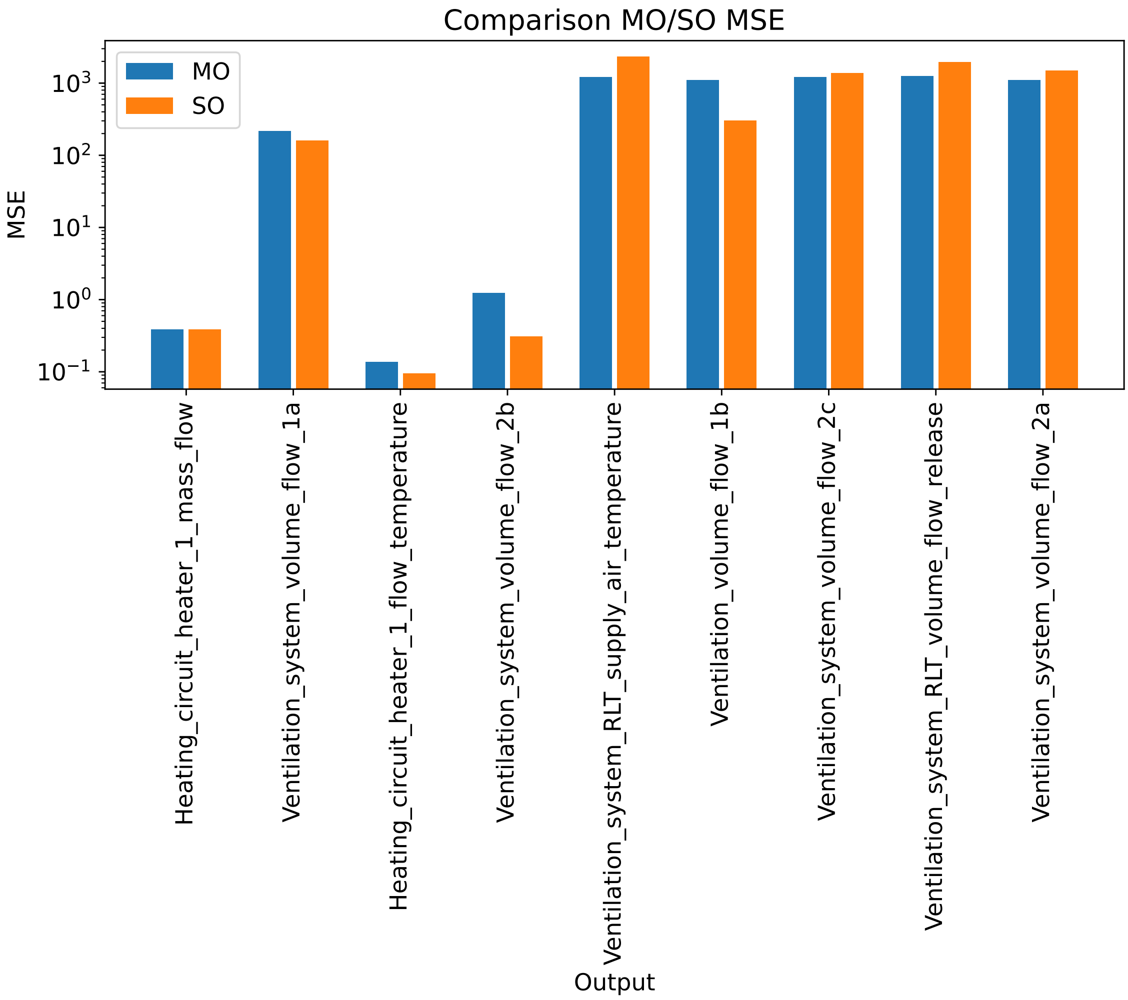


Figure 17: MSE values of zone 5 with logarithmic scale

- All outputs in zone 16 have very low MSE values ($\sim\leq 0.04$).
- All MO MSE values of zone 16 are below 0.01.
- One output has a higher MSE value for the MO model of zone 16, the other has a higher MSE value for the SO model.
- The difference in MSE values of zone 16 between MO and SO are more pronounced in the case where the MO MSE is better.

Conclusion

- The outputs of zone 5 do not deliver equally bad results.
- Losses are very different across similar data.

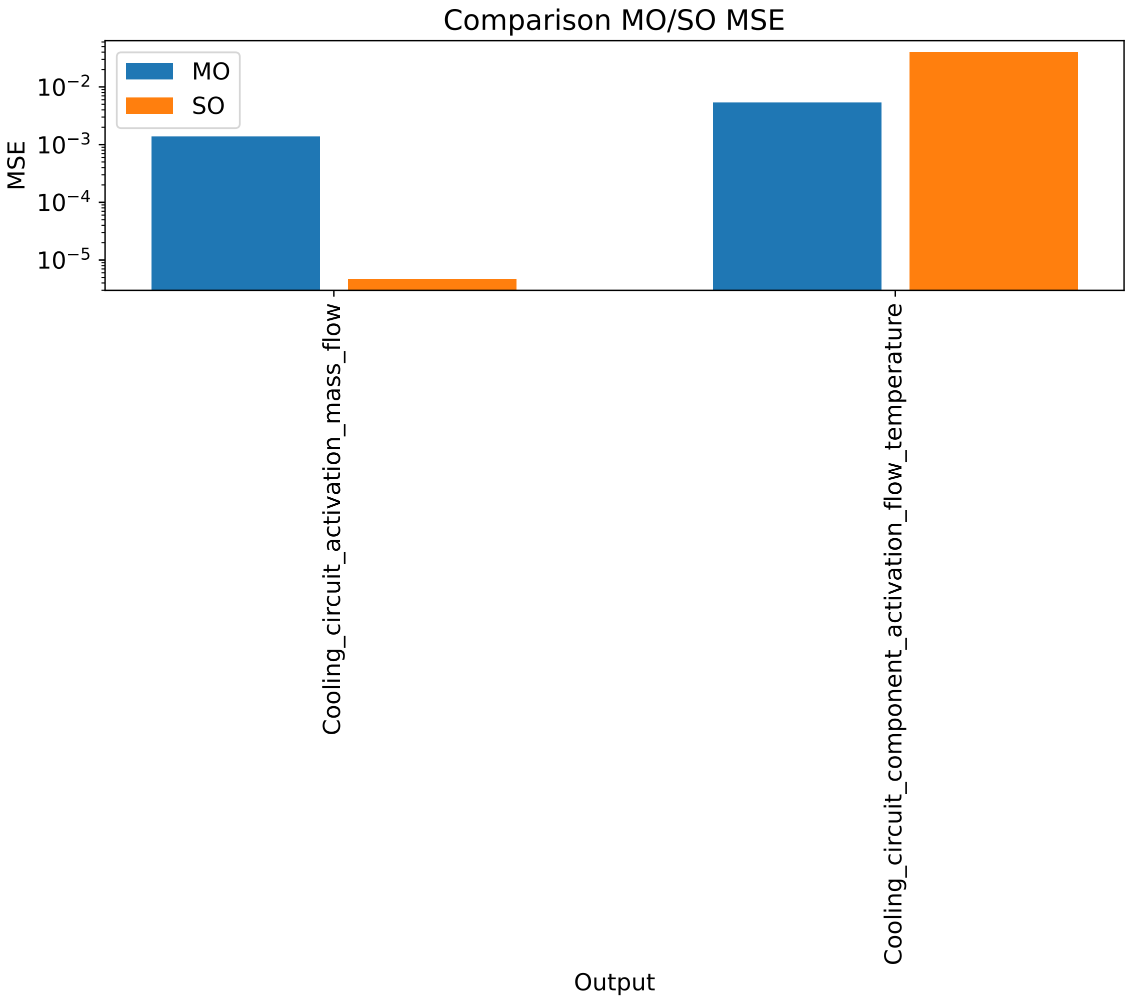
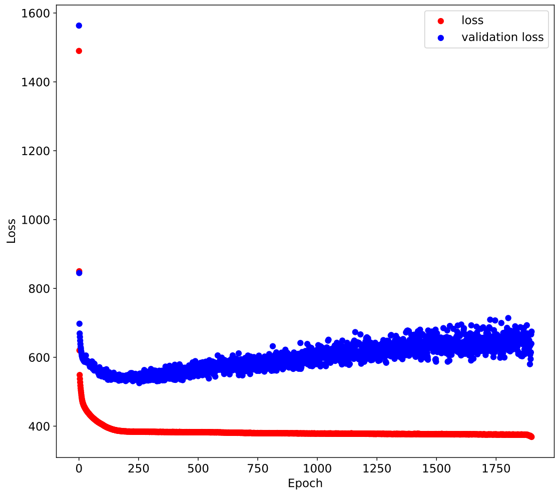


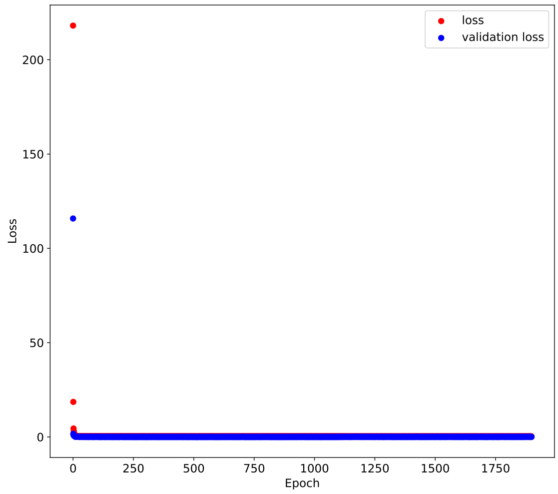
Figure 18: MSE values of zone 16 with logarithmic scale

- For zone 16, all outputs achieve good results.

One factor that has a major impact on the quality of the results is the assessment of the training process to determine any over- or underfitting (2.1.2.2). For this purpose, the loss of the training data is depicted for each training period (epoch). For MO models, this means that the loss is calculated on average over all outputs. In addition, there is validation data, in our case the test data, for which the loss was also calculated in each epoch. However, the validation data has no influence on the model, so it is not used for the training itself. The illustrations of all history can be found in Appendix E. In more detail, we examine the history for zones 5 and 16, which are located in figures 64 - 67 and 85 - 86. The different scales of the y-axes should be noted.



(a) Zone 5



(b) Zone 16

Figure 19: MO history for zones 5 and 16

Description

- The MO losses of the training and validation data over 1900 epochs is depicted in figure 19a.
- Graphic 19b visualizes the MSE values of zone 16 for all epochs.

Analysis

- For both the test and the validation data of zone 5 the final loss is high. This may indicate underfitting, which is not the case here.
- The validation data of zone 5 shows an upward trend. Together with the constant tendency of the training data, this rather speaks against underfitting and is a clear sign that the model has adapted too much to the training data and thus does not deliver good results for unknown data. This is therefore a case of overfitting.
- The fact that the loss of training data in zone 5 is so high despite overfitting suggests that the model may not be suitable for reproducing this data.
- In zone 16 it is visible that both data sets ultimately have low losses and develop a constant tendency. Hence there is neither overfitting nor underfitting.

Conclusion

- Overfitting with nevertheless high losses can be observed in zone 5.
- An ideal training process can be seen in zone 16.

An assessment of the associated SO models with regard to over- and underfitting, including a brief justification, can be found in table 3.

Zone	Output	Assessment	Reason
5	Ventilation_ ... _1b	Overfitting	Rising validation loss
	heating_ ... _mass_flow	Neither	Both constant low losses
	Ventilation_ ... _2c	Unsuitable model	Both constant high losses
	Ventilation_ ... _1a	Unsuitable model	Both constant high losses
	Ventilation_ ... _2b	Unsuitable model	Both constant high losses
	Ventilation_ ... _release	Neither	Both consistently low losses
	Ventilation_ ... _2a	Overfitting	Rising validation loss, decreasing traing loss
	Heating_ ... _temperature	Overfitting	Low training losses, high validation losses
	Ventilation_ ... _temperature	Neither	Both consistently low losses
16	Cooling_ ... _temperature	Tendency towards overfitting	Both low losses, validation loss rising trend
	Cooling_ ... _mass_flow	Neither	Both consistently low losses

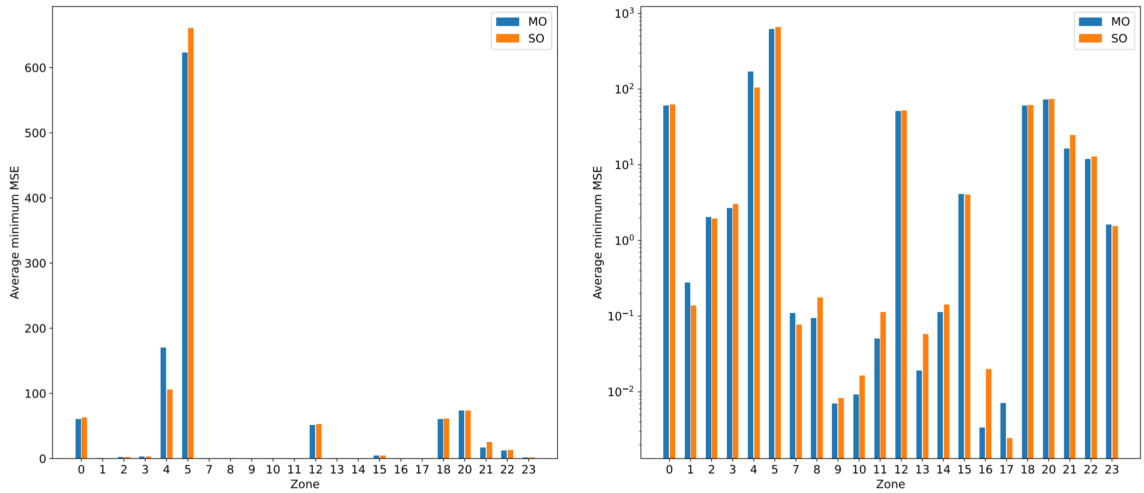
Table 3: Evaluation of over- and underfitting for zones 5 and 16

Conclusion

- The cause of high MSE values can usually be attributed to a subset of the outputs.
- There is a possibility that the magnitude of the loss is related to the number and proportion of outputs.
- In the case of zone 5, the results are poor due to an incompatibility with the model, which leads to overfitting with high MSE values.

4.3.4. Test 4: MO evaluation

So far, we have primarily looked at how well the overall combination of preprocessing and one model works for the different zones. In this test we want to take a look at how well the data can be represented by this one model, independent of the exact preprocessing method.
For this purpose, we first have illustrated the minimum MSE values (figure 20). Between the zones, as well as between MO and SO, there is no separation by preprocessing method.



(a) Linear scale

(b) Logarithmic scale

Figure 20: Minimal MSE values of all zones

Description

- Minimal MSE values from all of our selected preprocessing methods are shown in figure 20a with linear scaling.
- In order to allow the assessment of low MSE values, the minimum losses are presented with logarithmic scaling in figure 20b.
- There is one bar per zone for the minimum MO and for the minimum SO values.

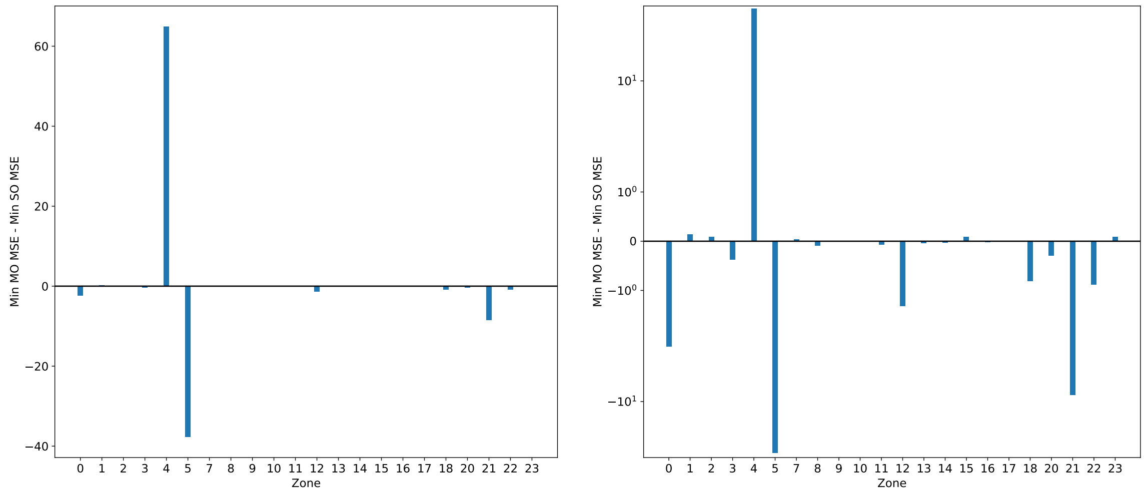
Analysis

- There are considerably higher minimum possible MSE values for zone 5 than for the other zones, which supports the thesis that the chosen model is not an adequate model for the data of the zone.
- The minimum possible MSE values are very similar for the MO and SO models.
- 14 out of 23 models have very low minimum losses.

Conclusion

- It can be seen that although for some zones the minimum possible MSE values of the model are high, most zones provide good MSE values by means of varying preprocessing options.
- Given the very similar minimal results, with the right choice of preprocessing MO models offer a good alternative to SO models.

In order to enable a reliable comparison of the minimum MSE values between the MO and SO models, we look at the difference between these values in more detail below.



(a) Linear scale

(b) Symmetrical logarithmic scale

Figure 21: Difference of MO and SO MSE values with different y-scales

Description

- Figure 21a calculates the difference of the minimal MO MSE and minimal SO MSE values.
- A bar pointing downwards indicates a higher SO MSE, whereby the MO model gives better results. Similarly, a bar pointing upwards indicates a better SO model.
- An equivalent to graph 21a is provided by graph 21b with logarithmic scaling.

Analysis

- Zones 4 and 5 have the strongest spikes, with zone 4 having a better SO and zone 5 a better MO model.
- Zones 0, 12, 18, 21 and 22 also still have relatively legible trends, which suggest better MO models.
- Zones 1, 2, 4, 7, 15 and 23 show a trend towards SO models, albeit often minimal.
- A trend towards MO models is shown by zones 0, 3, 5, 8, 11, 12, 13, 14, 18, 20, 21 and 22.
- Even with a logarithmic y-scaling, zones 9, 10, 16 and 17 do not show a well distinguishable difference between the MO and SO model results.

Conclusion

- Most zones achieve better results with MO models than with SO models.
- The difference between MO and SO is usually minimal when SO is better.

Since the peaks are not in relation to the MSE values, the difference of the MSE values seems greater in e.g. zone 5 than it does when comparing the MSE values themselves. Therefore we also visualized the proportion of MO MSE values in the SO MSE values (figure 22).

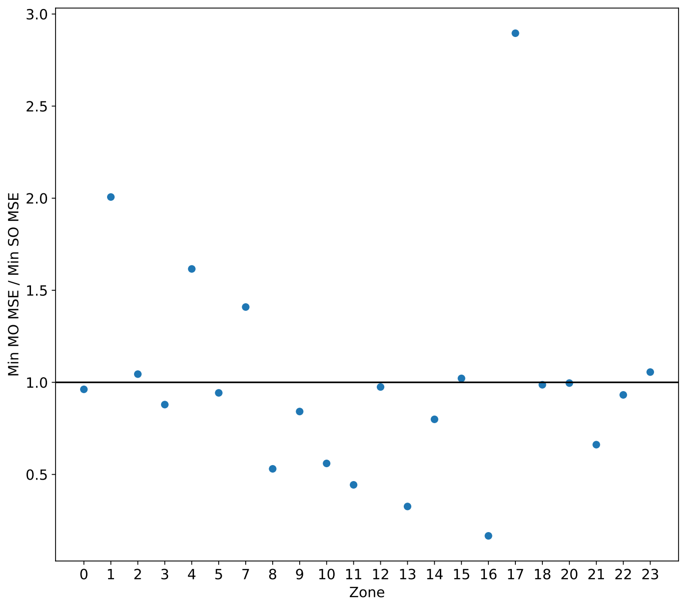


Figure 22: Share of MO MSE values in SO MSE values

Description

- Figure 22 displays points for the diversion of the MO by the SO MSE values for each zone.
- Values below one suggest better MO model results, while values higher than one express better SO MSE values.
- The farther away the dot is from one, the greater is the difference between MO and SO MSE values in relation to the level of the MSE values.

Analysis

- In comparison to figure 21 the peak of zone 5 is much less.
- The highest peak is produced by zone 17, which tends to SO.

- Zones 9, 10 and 16 show better MO MSE values than SO MSE values.

Conclusion

- Most zones show better results for MO Models.
- The relative difference of the results is mostly higher when the SO models produce better results.

4.3.5. Test 5: Pattern evaluation

In the final test, we want to look for patterns across all preprocessing methods, datasets and zones. To this end, we have created a heatmap, and the creation process of this heatmap is shown in figure 23.

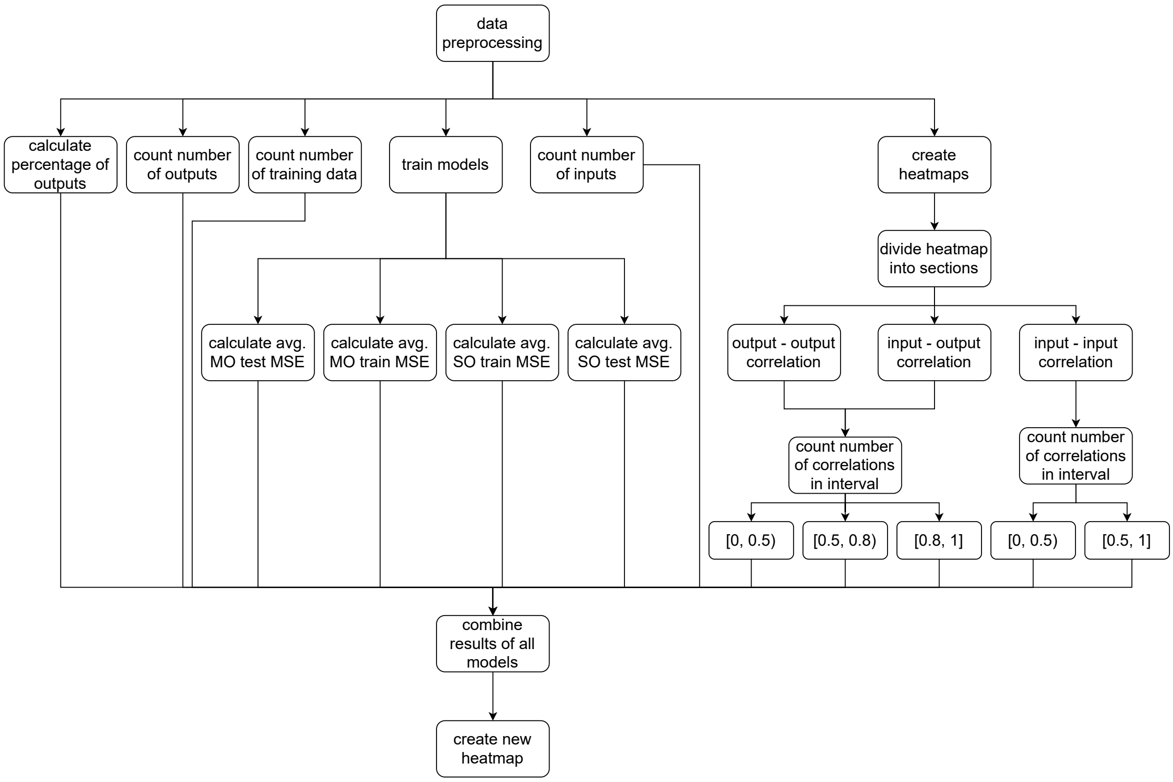


Figure 23: Procedure for creating the pattern heatmap

- General information such as the number of inputs, outputs and training data, as well as the share of outputs in the data have been summarised.
- Average MSE values of the training and test data were compiled for all MO and SO models respectively.
- To include the influence of correlations, we have

- divided each heatmap into the correlations of outputs to outputs, inputs to outputs and inputs to inputs.
- specified the intervals [0, 0.5), [0.5, 0.8) and [0.8, 1] for out-out and in-out correlations and the intervals [0, 0.5) and [0.5, 1] for in-in correlations.
- counted how often the correlations in the respective areas lie in the corresponding intervals. The diagonal was not integrated.

- In the end all collected data was combined for all preprocessing options.
- The corresponding heatmap is shown in figure 24.

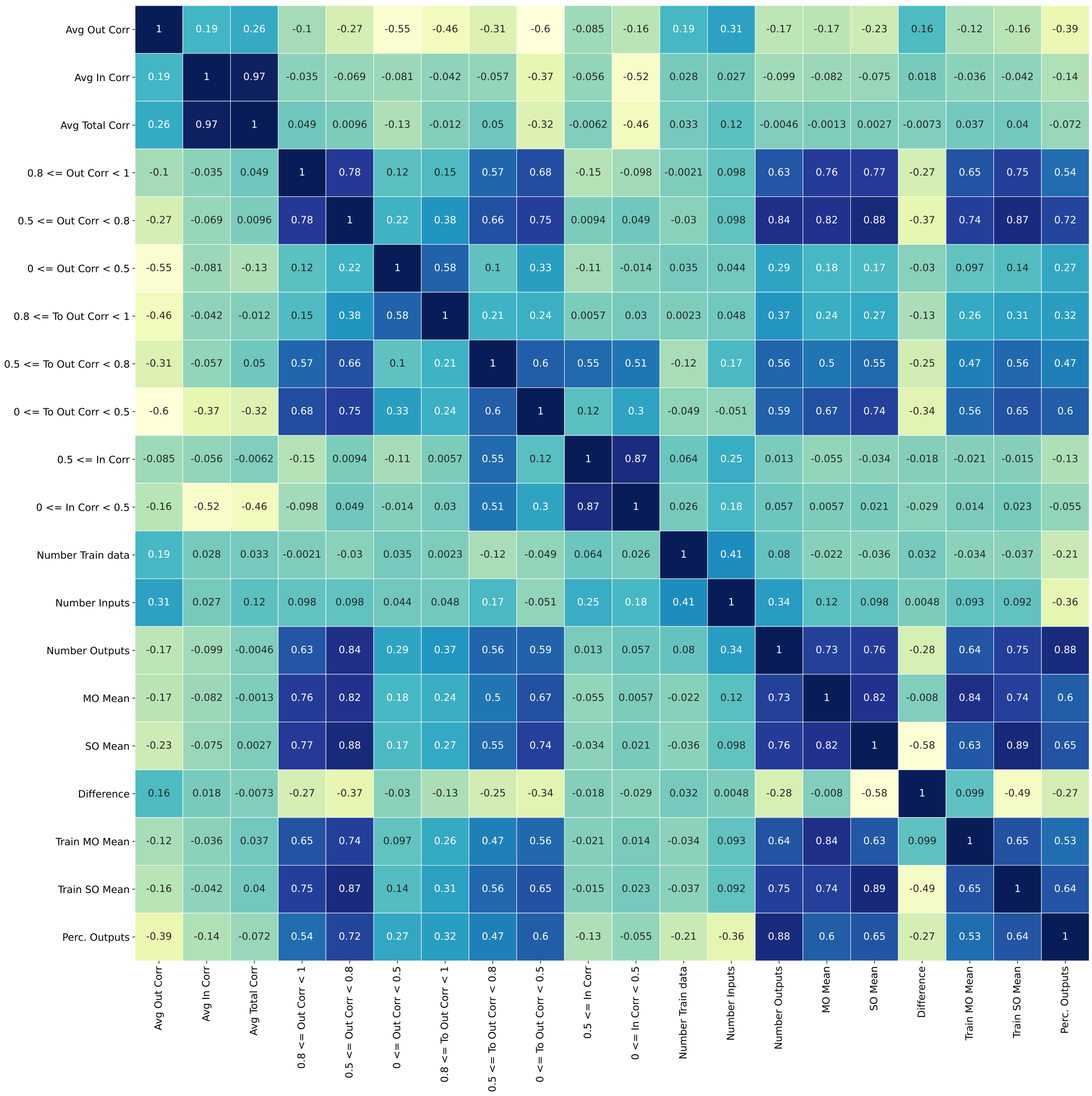


Figure 24: Pattern for all preprocessing methods

Description

- Figure 24 is a heatmap of the elements from which we want to derive a pattern.
- High negative correlations do not occur, so a focus on positive values is sufficient.
- Groups of high correlations are indicated by their dark blue colour.

Analysis

- There is a significant correlation between the MSE values and the number of very high correlations between outputs. It is important to note that the correlation with the SO MSE values (0.77, 0.88) seems to be slightly higher than with the MO MSE values (0.67, 0.82).
- The number of low correlations between inputs and outputs, also has a perceptible relationship with the results (0.67, 0.74). Medium-high correlations also have a relatively high correlation to the results (0.5, 0.55), only the number of very high correlations does not seem to have a significant influence, which could also be due to the possibly very low number of very high correlations.
- In general, the number of training data and inputs does not seem to have a perceptible effect on the average MSE values either, which fits to the results from 4.3.1 that none of the preprocessing methods provide ideal results for all zones.
- The correlations between the inputs do not seem to have an explicit influence on the MSE values.
- With a correlation between 0.6 and 0.76, the number of outputs, as well as the share of outputs in the total data, offer an aspect that seems to have an influence on the MSE values.
- The similarity between the MO and SO models is supported by the fact that the correlation between the MO and the SO MSE values is 0.82.
- A generally good sign for the reliability of the models is the correlation of 0.84 and 0.89 between the training and test MSE values of the models.

Conclusion

- There is a significant correlation between the MSE values of the models and the number of high correlations between outputs and between inputs and outputs.
- Another relevant factor seems to be the number of outputs, although there is also a high correlation with the SO MSE values, which reduces the weighting of this correlation.

5. Conclusion and Future Work

5.1. Conclusion

In the course of this work, we have addressed the questions of the extent to which a single model architecture is applicable to different building zones and whether SO models can be replaced by MO models.

- Using one preprocessing method and one model for all zones does not give ideal results and no preprocessing method is significantly better than the others. One year's data with forecasts instead of real weather data and a correlation of at least 0.1 from each input to at least one output still shows the best results most of the time, but also with one other preprocessing method the most poor results.

- It depends strongly on the zones how well the MO models represent the different outputs. There are zones that have better MO models as well as zones that have better SO models. For some zones, the results are quite similar.

- Comparing the best MO and SO models, it became clear that for some zones the model itself is not suitable for representing the data. The MO models most often provide the better models, but the relative difference is greatest when the SO models perform best.

- There is a significant correlation between the MSE values of the models and the number of high correlations between outputs and between inputs and outputs. Another relevant factor seems to be the number of outputs, although there is also a high correlation with the SO MSE values, which reduces the weighting of this correlation.

All in all, we have seen success in using MO models. Two thirds of the zones show better minimal results with MO models and apart from one zone the absolute differences between the other MO and SO results are marginal. Finding a preprocessing method that gives good results with the given model for all zones may not have had conclusive success, but the model itself seems to work for most of the zones.

5.2. Future Work

In the future, it would make sense to evaluate:

- How well a different model could represent the data.

- To what extent other correlation methods such as Spearman are suitable for preprocessing.

- Whether another assessment method such as cross-validation would be more appropriate.

- If other patterns can be discovered.

The GUI could be extended to include related functions, where one can select the type of correlation or the model itself, i.e. number of layers, nodes and type of model.

References

[1] Buildings – Topics. URL https://www.iea.org/topics/buildings.

[2] Directive (EU) 2018/ of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency. page 17.

[3] Recurrent Neural Networks: Design and Applications - Google Books. URL https://books.google.de/books?hl=de&lr&id=ME1SAkN0PyMC&oi=fnd&pg=PA1&dq=Recurrent+neural+networks+&ots=7ctufK2TRp&sig=Kx5R-p4ZZj6SnwJd2vLMHOAC1fg#v=onepage&q=Recurrent%20neural%20networks&f=false.

[4] Abdul Afram and Farrokh Janabi-Sharifi. Theory and applications of HVAC control systems – A review of model predictive control (MPC). *Building and Environment*, 72:343–355, February 2014. ISSN 03601323. doi: 10.1016/j.buildenv.2013.11.016. URL https://linkinghub.elsevier.com/retrieve/pii/S0360132313003363.

[5] Abdul Afram and Farrokh Janabi-Sharifi. Review of modeling methods for HVAC systems. *Applied Thermal Engineering*, 67(1-2):507–519, June 2014. ISSN 13594311. doi: 10.1016/j.applthermaleng.2014.03.055. URL https://linkinghub.elsevier.com/retrieve/pii/S1359431114002348.

[6] Abdul Afram and Farrokh Janabi-Sharifi. Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods. *Energy and Buildings*, 94:121–149, May 2015. ISSN 03787788. doi: 10.1016/j.enbuild.2015.02.045. URL https://linkinghub.elsevier.com/retrieve/pii/S0378778815001504.

[7] Zakia Afroz, Gm Shafiullah, Tania Urmee, and Gary Higgins. Modeling techniques used in building HVAC control systems: A review. *Renewable and Sustainable Energy Reviews*, 83:64–84, March 2018. ISSN 13640321. doi: 10.1016/j.rser.2017.10.044. URL https://linkinghub.elsevier.com/retrieve/pii/S1364032117314193.

[8] Abien Fred Agarap. Deep Learning using Rectified Linear Units (ReLU). *arXiv:1803.08375 [cs, stat]*, February 2019. URL http://arxiv.org/abs/1803.08375. arXiv: 1803.08375.

[9] Raffaele Carli, Graziana Cavone, Sarah Ben Othman, and Mariagrazia Dotoli. IoT Based Architecture for Model Predictive Control of HVAC Systems in Smart Buildings. *Sensors*, 20(3):781, January 2020. ISSN 1424-8220. doi: 10.3390/s20030781. URL https://www.mdpi.com/1424-8220/20/3/781.

[10] Erhard Cramer, Udo Kamps, and Ansgar Steland. *Grundlagen der Wahrscheinlichkeitsrechnung und Statistik*. Springer, 2008.

[11] Chirag Deb, Siew Eang Lee, and Mattheos Santamouris. Using artificial neural networks to assess HVAC related energy saving in retrofitted office buildings. *Solar Energy*, 163:32–44, March 2018. ISSN 0038092X. doi: 10.1016/j.solener.2018.01.075. URL `https://linkinghub.elsevier.com/retrieve/pii/S0038092X18300963`.

[12] Ján Drgoňa, Javier Arroyo, Iago Cupeiro Figueroa, David Blum, Krzysztof Arendt, Donghun Kim, Enric Perarnau Ollé, Juraj Oravec, Michael Wetter, Draguna L. Vrabie, and Lieve Helsen. All you need to know about model predictive control for buildings. *Annual Reviews in Control*, 50:190–232, 2020. ISSN 13675788. doi: 10.1016/j.arcontrol.2020.09.001. URL `https://linkinghub.elsevier.com/retrieve/pii/S1367578820300584`.

[13] Ján Drgoňa, Damien Picard, and Lieve Helsen. Cloud-based implementation of white-box model predictive control for a GEOTABS office building: A field test demonstration. *Journal of Process Control*, 88:63–77, April 2020. ISSN 09591524. doi: 10.1016/j.jprocont.2020.02.007. URL `https://linkinghub.elsevier.com/retrieve/pii/S0959152419306857`.

[14] B Egilegor, J P Uribe, G Arregi, E Pradilla, and L Susperregi. A FUZZY CONTROL ADAPTED BY A NEURAL NETWORK TO MAINTAIN A DWELLING WITHIN THERMAL COMFORT. page 9.

[15] Betul Bektas Ekici and U. Teoman Aksoy. Prediction of building energy consumption by using artificial neural networks. *Advances in Engineering Software*, 40(5): 356–362, May 2009. ISSN 09659978. doi: 10.1016/j.advengsoft.2008.05.003. URL `https://linkinghub.elsevier.com/retrieve/pii/S0965997808001105`.

[16] Maryam Gholamzadehmir, Claudio Del Pero, Simone Buffa, Roberto Fedrizzi, and Niccolo' Aste. Adaptive-predictive control strategy for HVAC systems in smart buildings – A review. *Sustainable Cities and Society*, 63:102480, December 2020. ISSN 22106707. doi: 10.1016/j.scs.2020.102480. URL `https://linkinghub.elsevier.com/retrieve/pii/S2210670720307009`.

[17] Rasmus Halvgaard, Niels Kjolstad Poulsen, Henrik Madsen, and John Bagterp Jorgensen. Economic Model Predictive Control for building climate control in a Smart Grid. In *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, pages 1–6, Washington, DC, USA, January 2012. IEEE. ISBN 978-1-4577-2159-5 978-1-4577-2158-8 978-1-4577-2157-1. doi: 10.1109/ISGT.2012.6175631. URL `http://ieeexplore.ieee.org/document/6175631/`.

[18] Jan LM Hensen and Roberto Lamberts. *Building performance simulation for design and operation*. Routledge, 2012.

[19] Hao Huang, Lei Chen, and Eric Hu. A neural network-based multi-zone modelling approach for predictive control system design in commercial buildings. *Energy and Buildings*, 97:86–97, June 2015. ISSN 03787788. doi: 10.1016/j.enbuild.2015.03.045. URL `https://linkinghub.elsevier.com/retrieve/pii/S0378778815002546`.

[20] S. A. Kalogirou. Artificial neural networks in energy applications in buildings. *International Journal of Low-Carbon Technologies*, 1(3):201–216, July 2006. ISSN 1748-1317. doi: 10.1093/ijlct/1.3.201. URL `https://academic.oup.com/ijlct/article-lookup/doi/10.1093/ijlct/1.3.201`.

[21] Sakshi Kohli, Surbhi Miglani, and Rahul Rapariya. BASICS OF ARTIFICIAL NEURAL NETWORK. page 7, 2014.

[22] Andrew Kusiak and Guanglin Xu. Modeling and optimization of HVAC systems using a dynamic neural network. *Energy*, 42(1):241–250, June 2012. ISSN 03605442. doi: 10.1016/j.energy.2012.03.063. URL `https://linkinghub.elsevier.com/retrieve/pii/S0360544212002666`.

[23] Yudong Ma, Francesco Borrelli, Brandon Hencey, Andrew Packard, and Scott Bortoff. Model Predictive Control of thermal energy storage in building cooling systems. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 392–397, Shanghai, China, December 2009. IEEE. ISBN 978-1-4244-3871-6. doi: 10.1109/CDC.2009.5400677. URL `http://ieeexplore.ieee.org/document/5400677/`.

[24] Giancarlo Mantovani and Luca Ferrarini. Temperature Control of a Commercial Building With Model Predictive Control Techniques. *IEEE Transactions on Industrial Electronics*, 62(4):2651–2660, April 2015. ISSN 0278-0046, 1557-9948. doi: 10.1109/TIE.2014.2387095. URL `http://ieeexplore.ieee.org/document/7000560/`.

[25] D Mariano-Herna and F Santos García. A review of strategies for building energy management system: Model predictive control, demand side management, optimization, and fault detect & diagnosis. *Journal of Building Engineering*, page 12, 2021.

[26] Miguel Martínez Comesaña, Lara Febrero-Garrido, Francisco Troncoso-Pastoriza, and Javier Martínez-Torres. Prediction of building's thermal performance using lstm and mlp neural networks. *Applied Sciences*, 10(21):7439, 2020.

[27] Petru-Daniel Moroşan, Romain Bourdais, Didier Dumur, and Jean Buisson. Building temperature regulation using a distributed model predictive control. *Energy and Buildings*, 42(9):1445–1452, September 2010. ISSN 03787788. doi: 10.1016/j.enbuild.2010.03.014. URL `https://linkinghub.elsevier.com/retrieve/pii/S0378778810000915`.

[28] Fatma Mtibaa, Kim-Khoa Nguyen, Muhammad Azam, Anastasios Papachristou, Jean-Simon Venne, and Mohamed Cheriet. LSTM-based indoor air temperature prediction framework for HVAC systems in smart buildings. *Neural Computing and Applications*, 32(23):17569–17585, December 2020. ISSN 0941-0643, 1433-3058. doi: 10.1007/s00521-020-04926-3. URL http://link.springer.com/10.1007/s00521-020-04926-3.

[29] Himanshu Nagpal, Andrea Staino, and Biswajit Basu. Robust model predictive control of HVAC systems with uncertainty in building parameters using linear matrix inequalities. *Advances in Building Energy Research*, 14(3):338–354, July 2020. ISSN 1751-2549, 1756-2201. doi: 10.1080/17512549.2019.1588165. URL https://www.tandfonline.com/doi/full/10.1080/17512549.2019.1588165.

[30] Nasruddin, Sholahudin, Pujo Satrio, Teuku Meurah Indra Mahlia, Niccolo Giannetti, and Kiyoshi Saito. Optimization of HVAC system energy consumption in a building using artificial neural network and multi-objective genetic algorithm. *Sustainable Energy Technologies and Assessments*, 35:48–57, October 2019. ISSN 22131388. doi: 10.1016/j.seta.2019.06.002. URL https://linkinghub.elsevier.com/retrieve/pii/S2213138818305629.

[31] Frauke Oldewurtel, Alessandra Parisio, Colin N. Jones, Dimitrios Gyalistras, Markus Gwerder, Vanessa Stauch, Beat Lehmann, and Manfred Morari. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45:15–27, February 2012. ISSN 03787788. doi: 10.1016/j.enbuild.2011.09.022. URL https://linkinghub.elsevier.com/retrieve/pii/S0378778811004105.

[32] Tomas Pippia, Jesus Lago, Roel De Coninck, and Bart De Schutter. Scenario-based Nonlinear Model Predictive Control for Building Heating Systems. *arXiv:2012.02011 [cs, eess]*, December 2020. URL http://arxiv.org/abs/2012.02011. arXiv: 2012.02011.

[33] V.A. Profillidis and G.N. Botzoris. Artificial Intelligence—Neural Network Methods. In *Modeling of Transport Demand*, pages 353–382. Elsevier, 2019. ISBN 978-0-12-811513-8. doi: 10.1016/B978-0-12-811513-8.00008-X. URL https://linkinghub.elsevier.com/retrieve/pii/B978012811513800008X.

[34] R. Sendra-Arranz and A. Gutiérrez. A long short-term memory artificial neural network to predict daily HVAC consumption in buildings. *Energy and Buildings*, 216:109952, June 2020. ISSN 03787788. doi: 10.1016/j.enbuild.2020.109952. URL https://linkinghub.elsevier.com/retrieve/pii/S0378778819312861.

[35] Shujiang Li, Shuang Ren, and Xiangdong Wang. HVAC Room Temperature Prediction Control Based on Neural Network Model. In *2013 Fifth International Conference on Measuring Technology and Mechatronics Automation*, pages

606–609, Hong Kong, January 2013. IEEE. ISBN 978-1-4673-5652-7 978-0-7695-4932-3. doi: 10.1109/ICMTMA.2013.151. URL `http://ieeexplore.ieee.org/document/6493803/`.

[36] Daniel Svozil, Vladimir Kvasnicka, and Jirí Pospichal. Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, 39(1):43–62, November 1997. ISSN 01697439. doi: 10.1016/S0169-7439(97)00061-0. URL `https://linkinghub.elsevier.com/retrieve/pii/S0169743997000610`.

[37] Joan Tarragona, Cèsar Fernández, Luisa F. Cabeza, and Alvaro de Gracia. Model predictive control applied to a heating system with PV panels and thermal energy storage. pages 2836–2843, Rome, Italy. doi: 10.26868/25222708.2019.210319. URL `http://www.ibpsa.org/proceedings/BS2019/BS2019_210319.pdf`.

[38] Donna Xu, Yaxin Shi, Ivor W. Tsang, Yew-Soon Ong, Chen Gong, and Xiaobo Shen. A Survey on Multi-output Learning. *arXiv:1901.00248 [cs, stat]*, October 2019. URL `http://arxiv.org/abs/1901.00248`. arXiv: 1901.00248.

[39] Yanlai Zhou, Fi-John Chang, Li-Chiu Chang, I-Feng Kao, and Yi-Shin Wang. Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. *Journal of Cleaner Production*, 209:134–145, February 2019. ISSN 09596526. doi: 10.1016/j.jclepro.2018.10.243. URL `https://linkinghub.elsevier.com/retrieve/pii/S0959652618332694`.

[40] Jan Široký, Frauke Oldewurtel, Jiří Cigler, and Samuel Prívara. Experimental analysis of model predictive control for an energy efficient building heating system. *Applied Energy*, 88(9):3079–3087, September 2011. ISSN 03062619. doi: 10.1016/j.apenergy.2011.03.009. URL `https://linkinghub.elsevier.com/retrieve/pii/S0306261911001668`.

A. Models

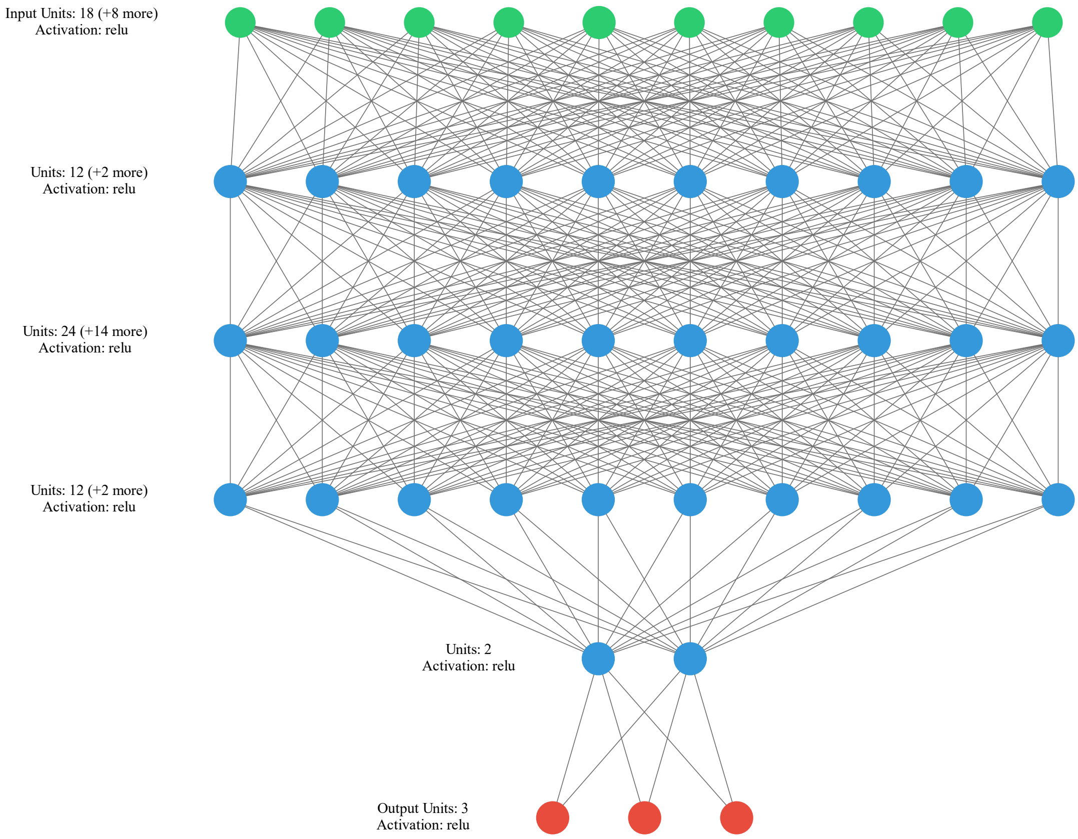


Figure 25: Exemplary MO model structure for zone 0 where too many nodes are shown in reduced number and noted in the margin

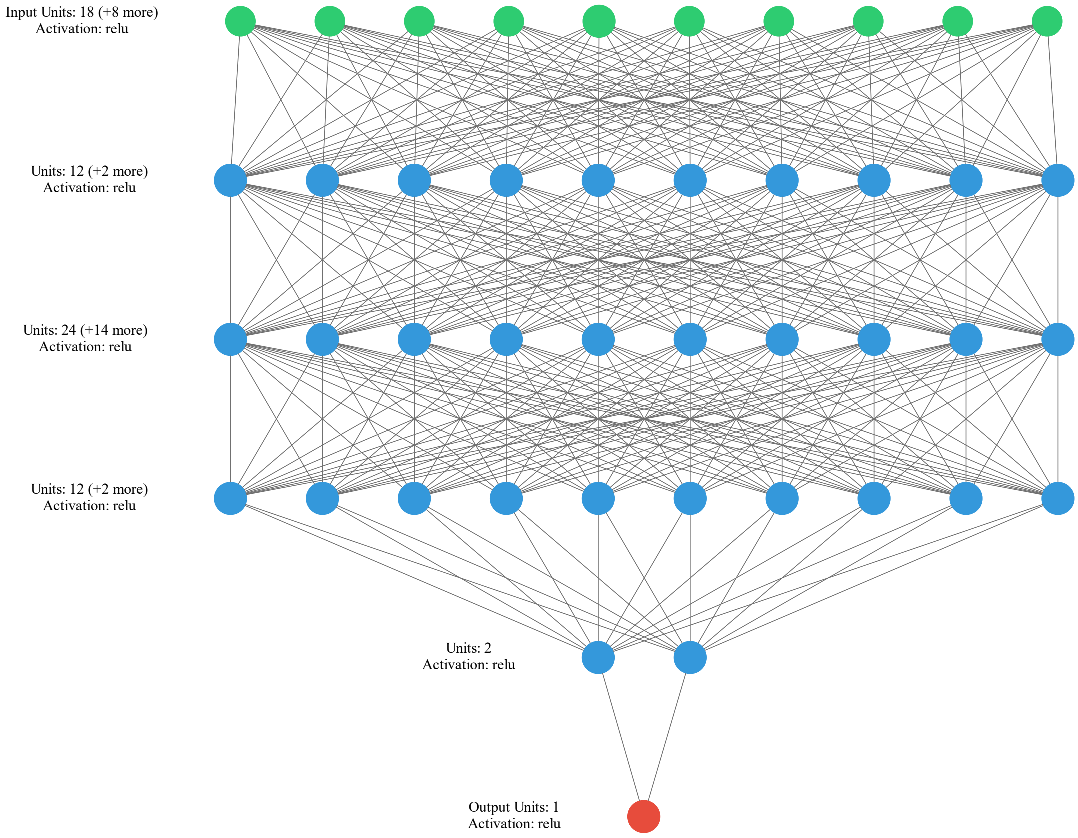
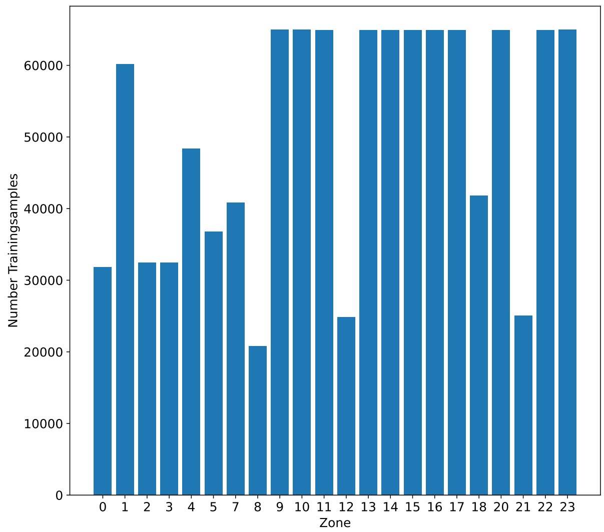
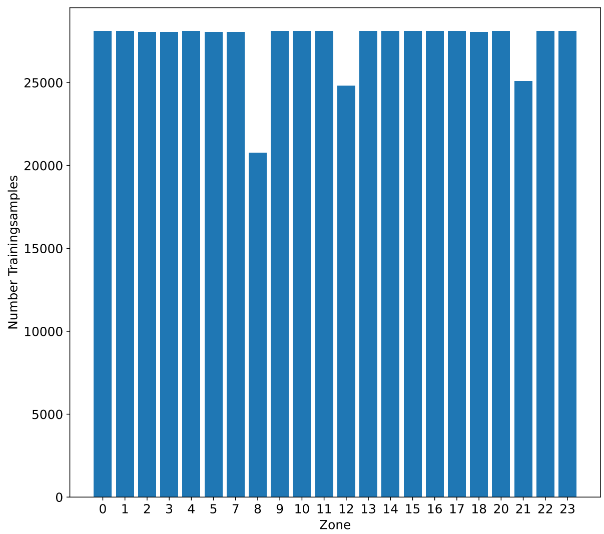


Figure 26: Exemplary SO model structure for zone 0 where too many nodes are shown in reduced number and noted in the margin

B. Data evaluation

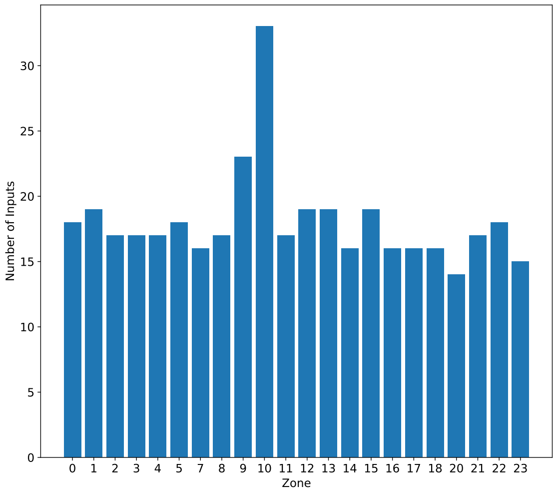


(a) Considering 3 years

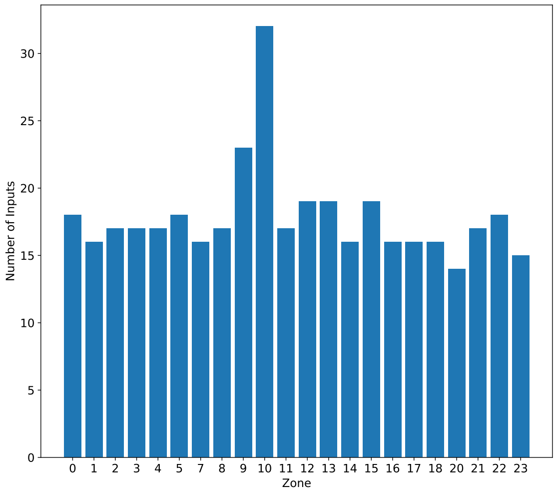


(b) Considering 1 year

Figure 27: Comparison number of training samples of the zones after dropping non-forecast data

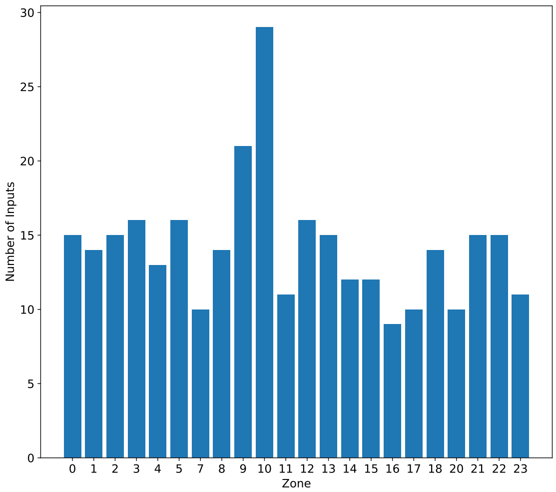


(a) Considering 3 years

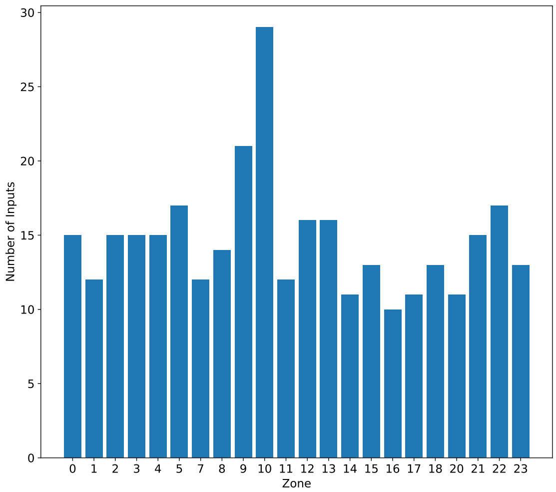


(b) Considering 1 year

Figure 28: Comparison number of inputs of the zones after dropping non-forecast data

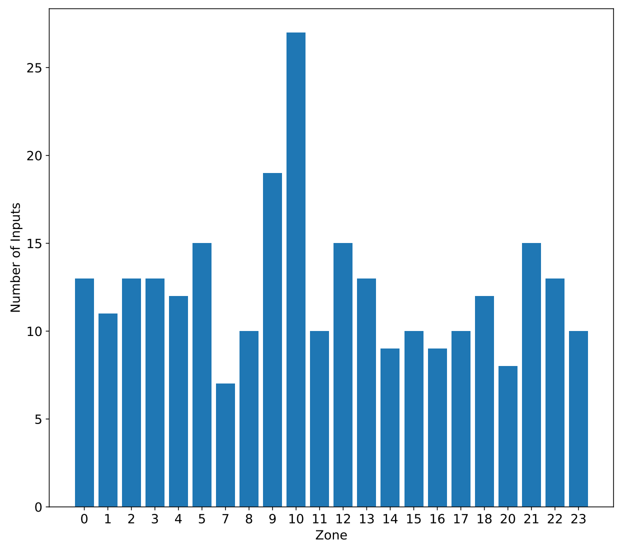


(a) Considering 3 years

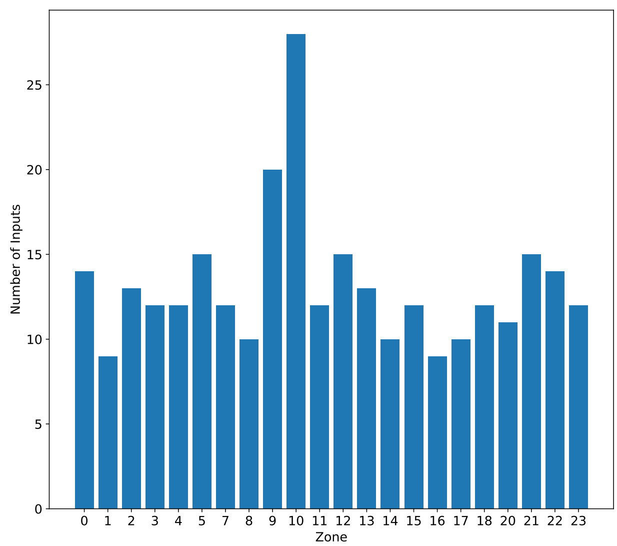


(b) Considering 1 year

Figure 29: Comparison number of inputs of the zones after dropping non-forecast data and data with no correlation >0.05 to at least one output

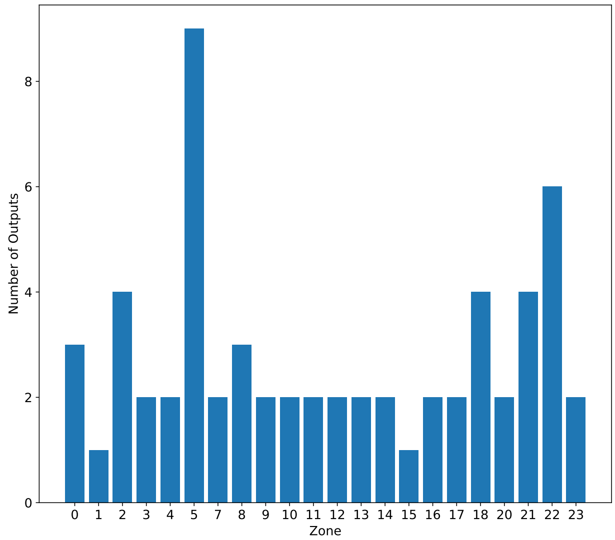


(a) Considering 3 years

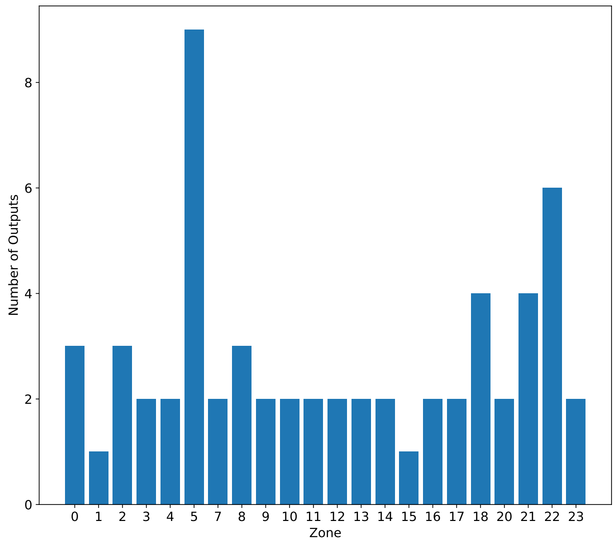


(b) Considering 1 year

Figure 30: Comparison number of inputs of the zones after dropping non-forecast data and data with no correlation > 0.1 to at least one output



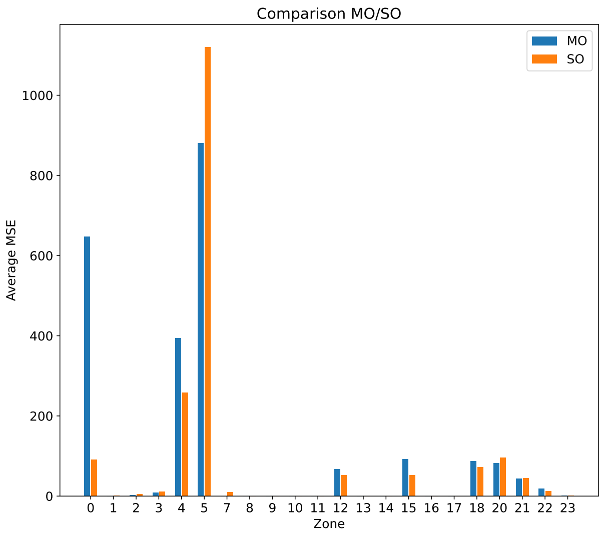
(a) Considering 3 years



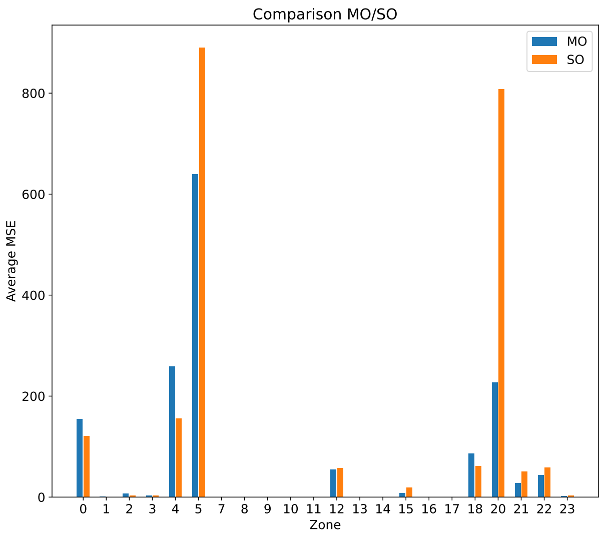
(b) Considering 1 year

Figure 31: Comparison number of outputs of the zones

C. MO SO evaluation

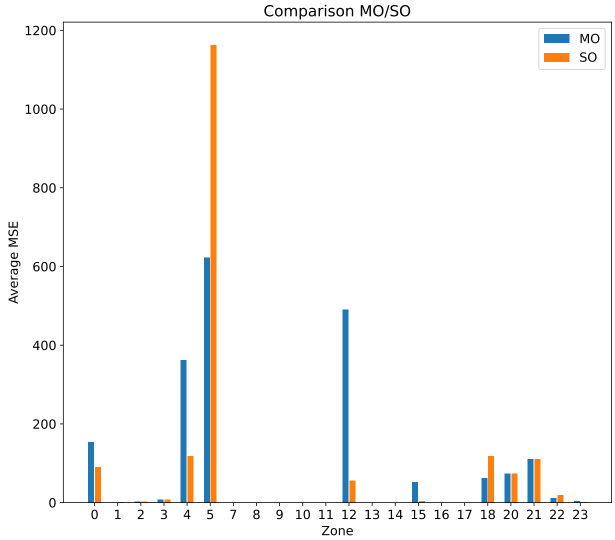


(a) Considering 3 years

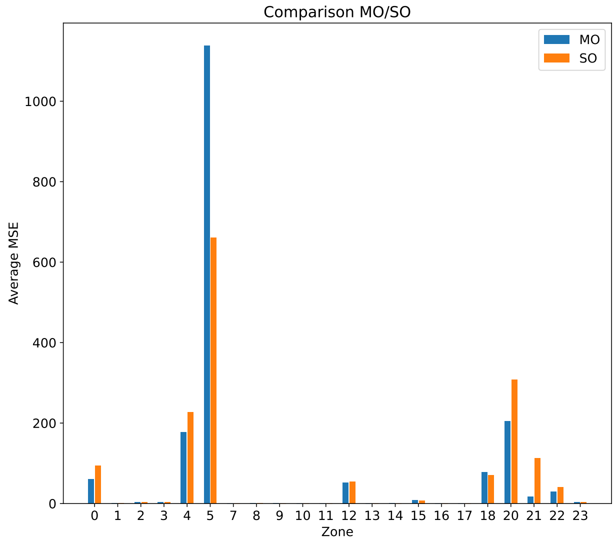


(b) Considering 1 year

Figure 32: Comparison of MO and SO average MSE values over all outputs of all zones after dropping non-forecast data

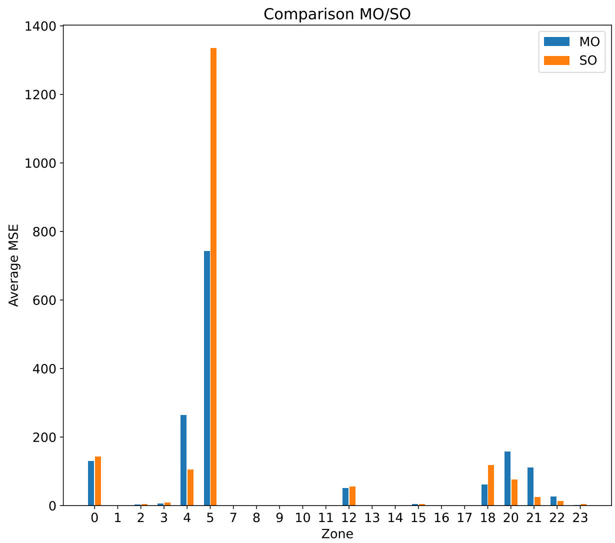


(a) Considering 3 years

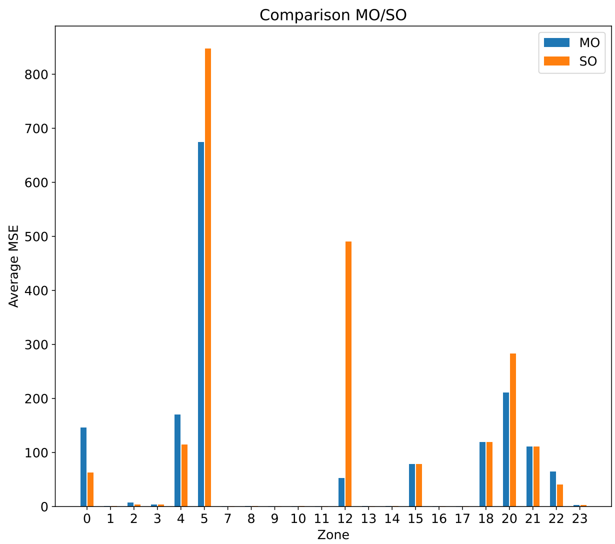


(b) Considering 1 year

Figure 33: Comparison of MO and SO average MSE values over all outputs of all zones after dropping non-forecast data and data with no correlation > 0.05 to at least one output



(a) Considering 3 years



(b) Considering 1 year

Figure 34: Comparison of MO and SO average MSE values over all outputs of all zones after dropping non-forecast data and data with no correlation > 0.1 to at least one output

D. Zone MSE evaluation

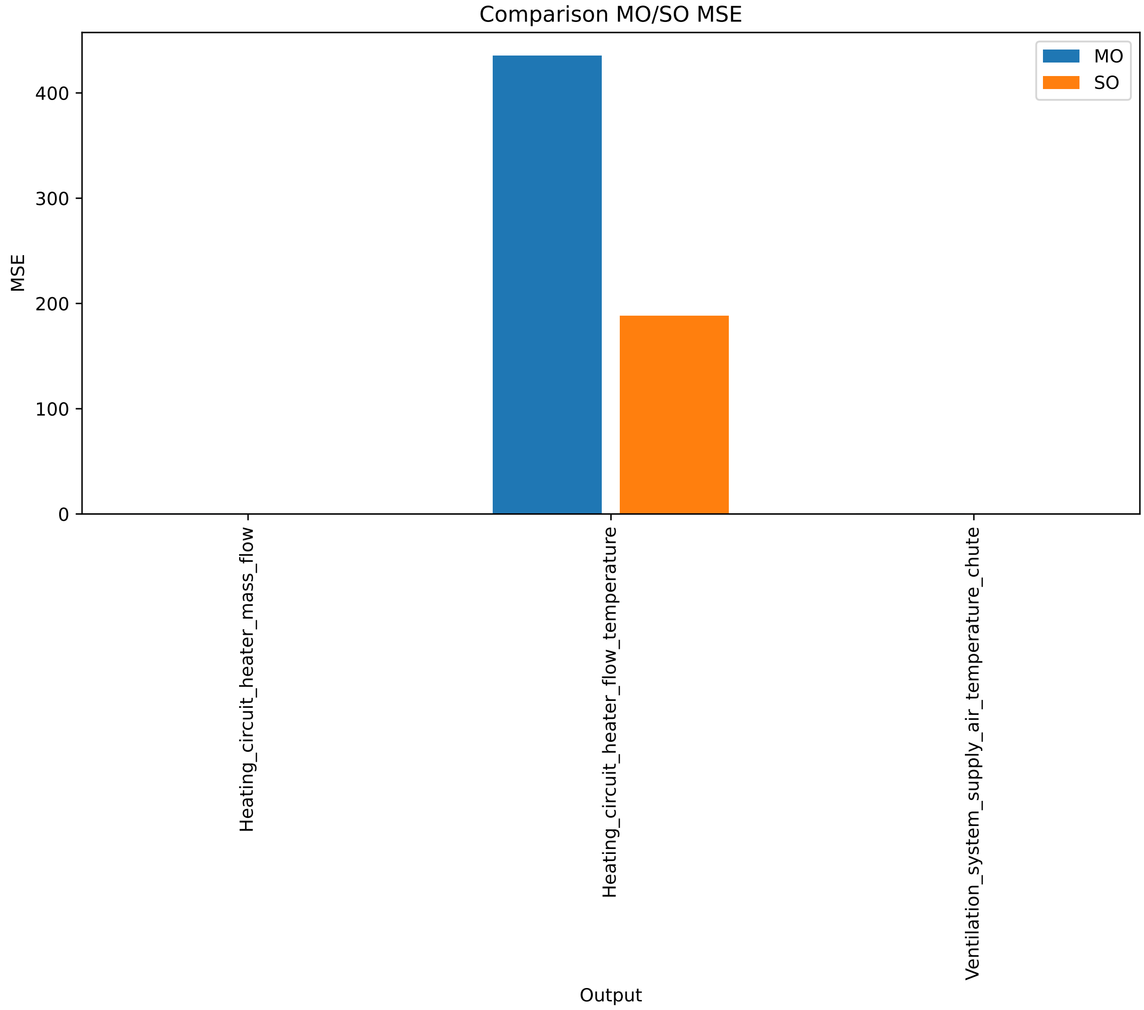


Figure 35: MSE values of each output of zone 0 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

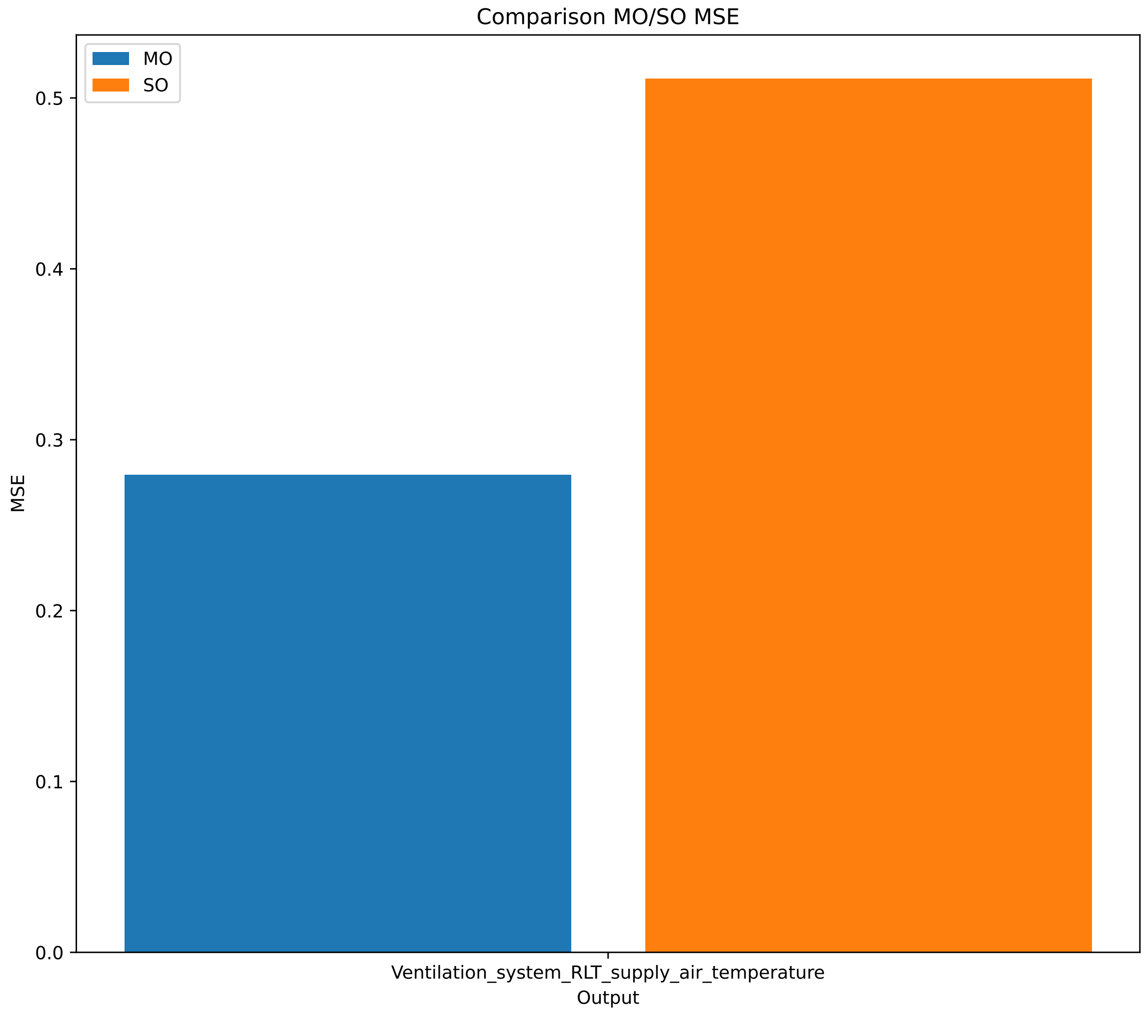


Figure 36: MSE values of each output of zone 1 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

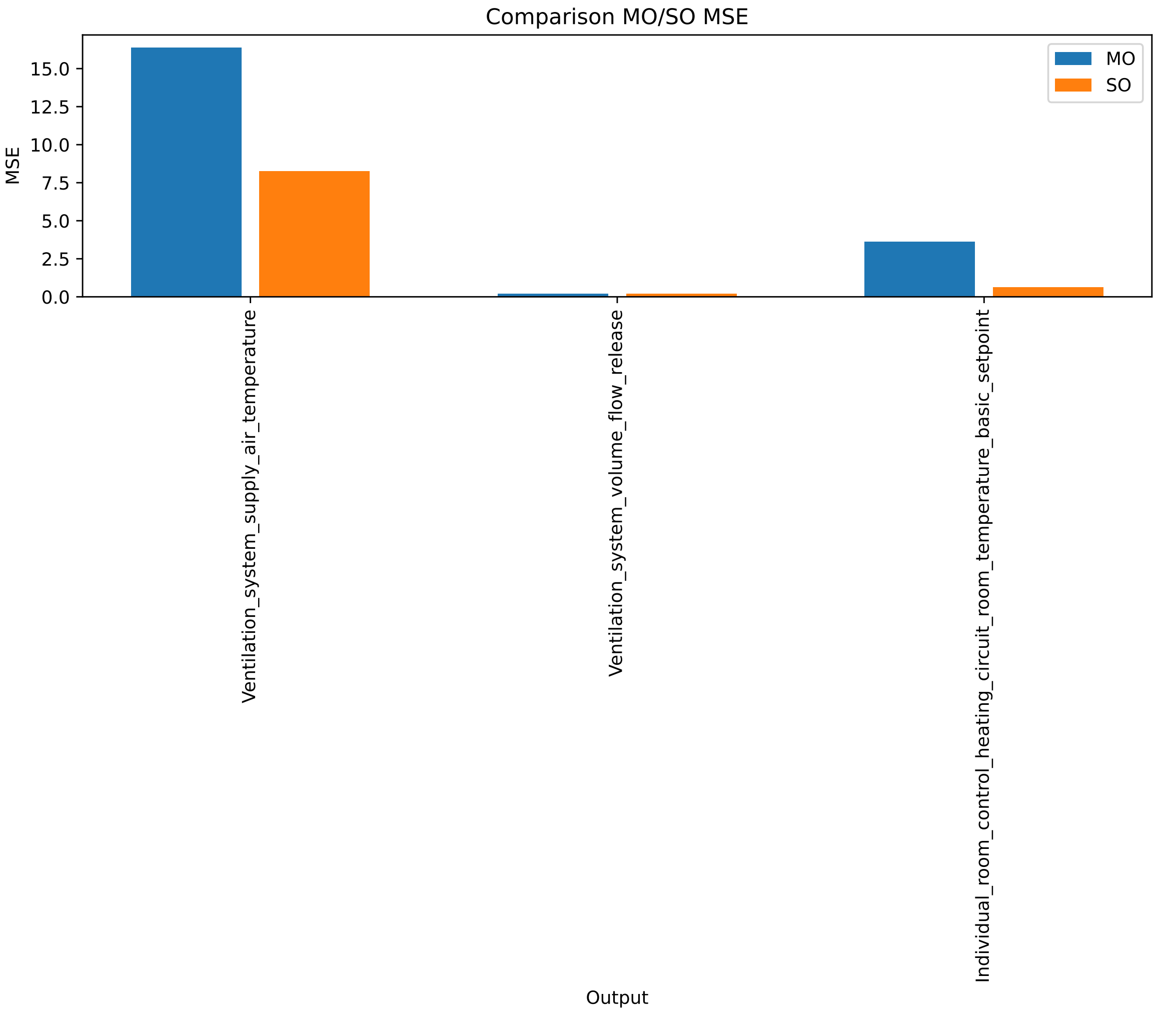


Figure 37: MSE values of each output of zone 2 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

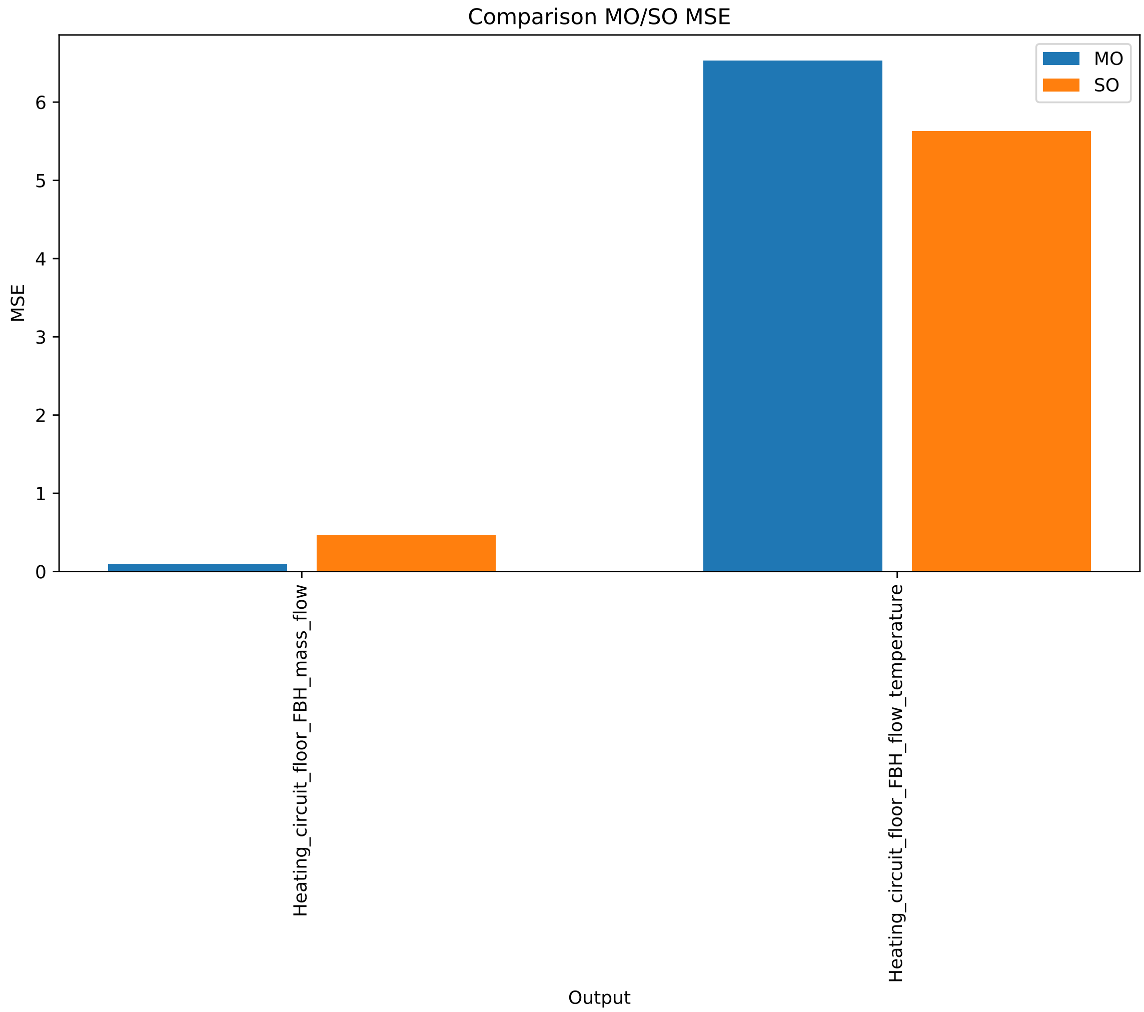


Figure 38: MSE values of each output of zone 3 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

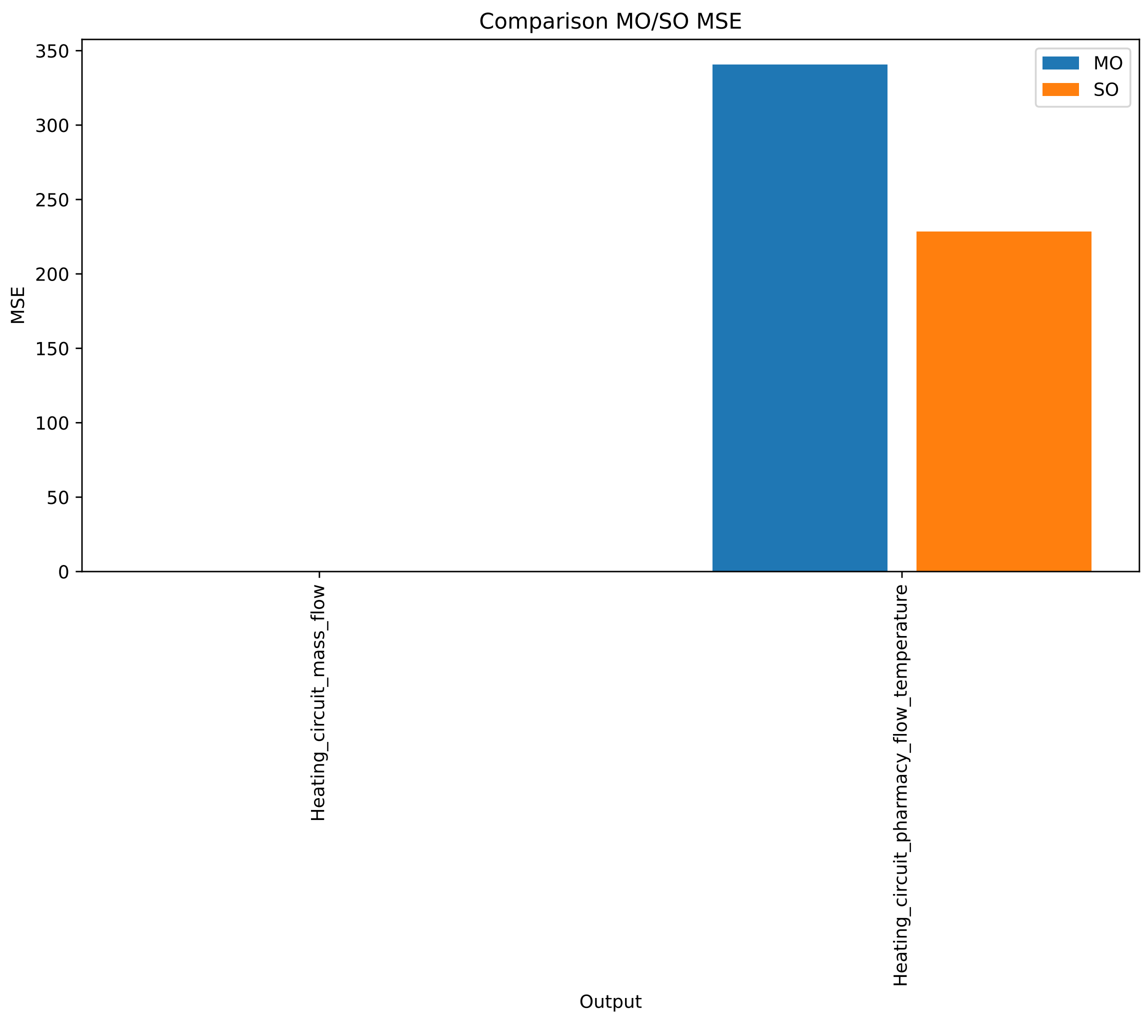


Figure 39: MSE values of each output of zone 4 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

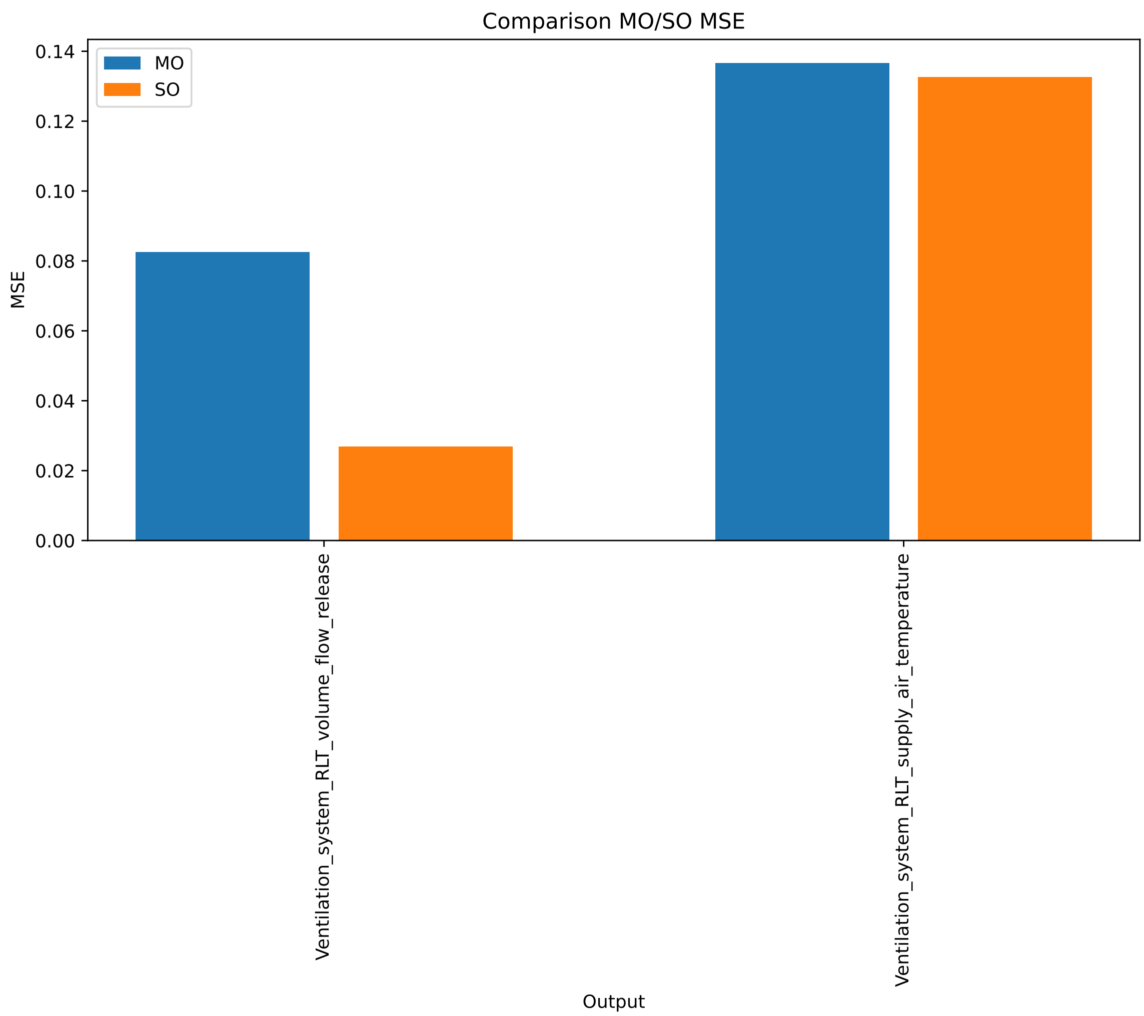


Figure 40: MSE values of each output of zone 7 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

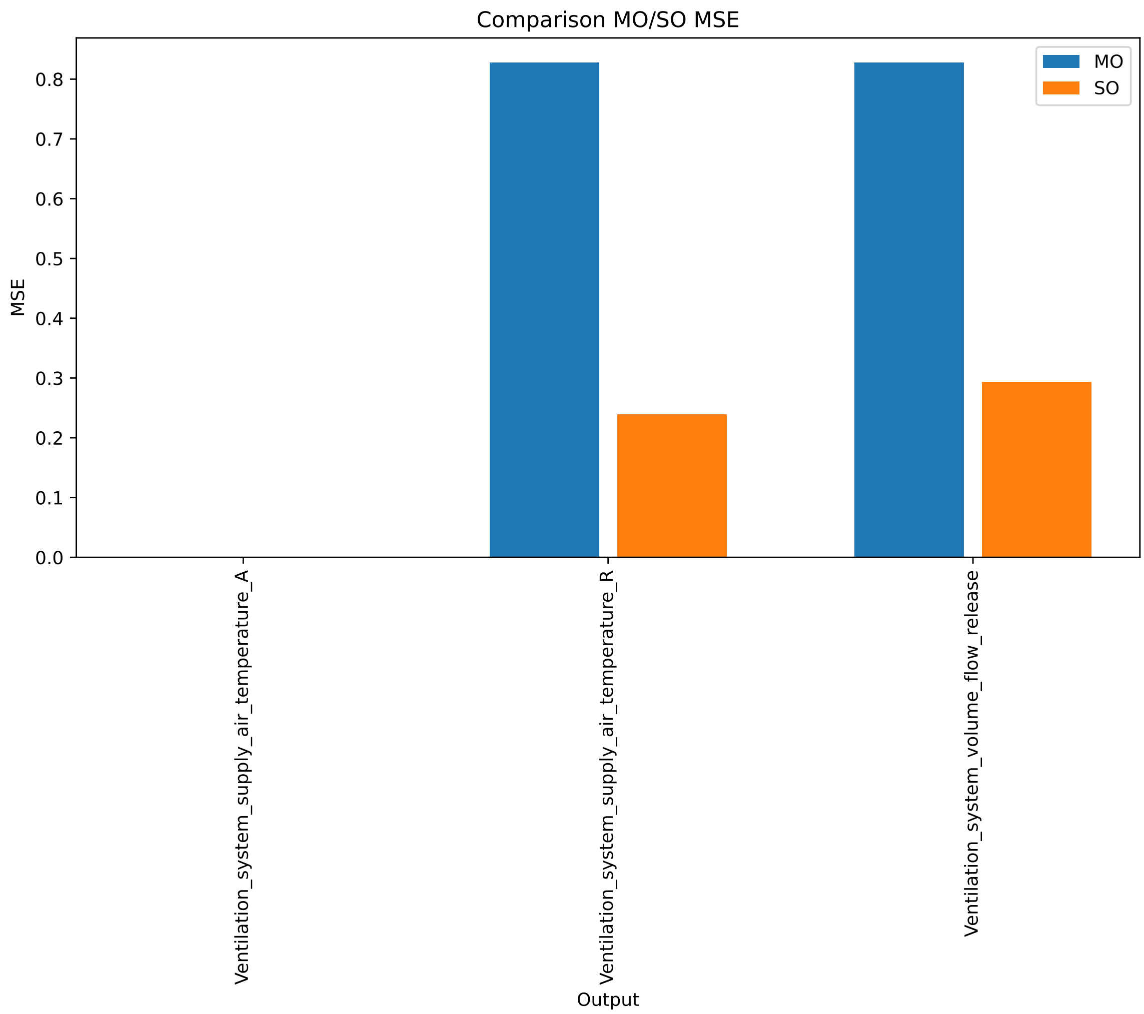


Figure 41: MSE values of each output of zone 8 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

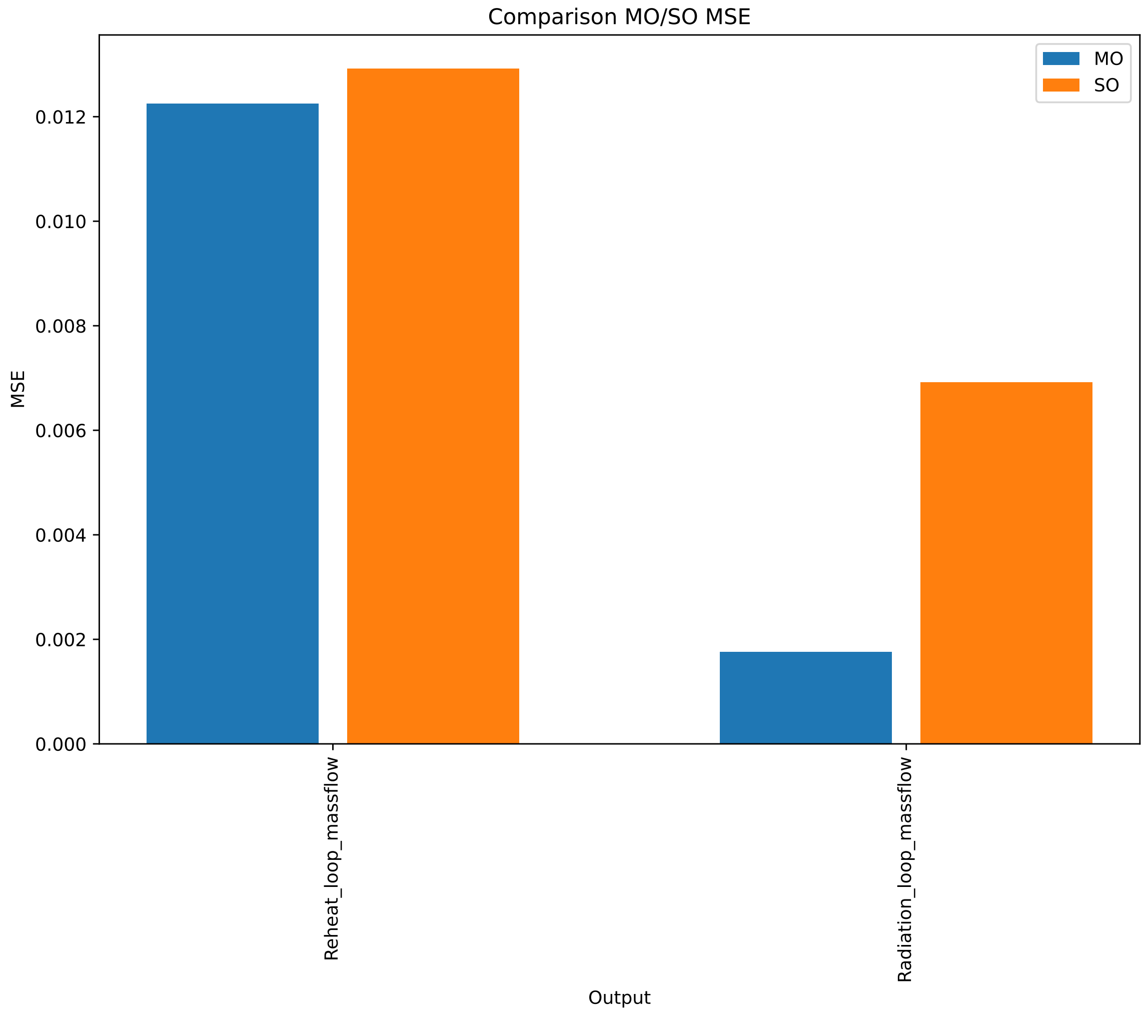


Figure 42: MSE values of each output of zone 9 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

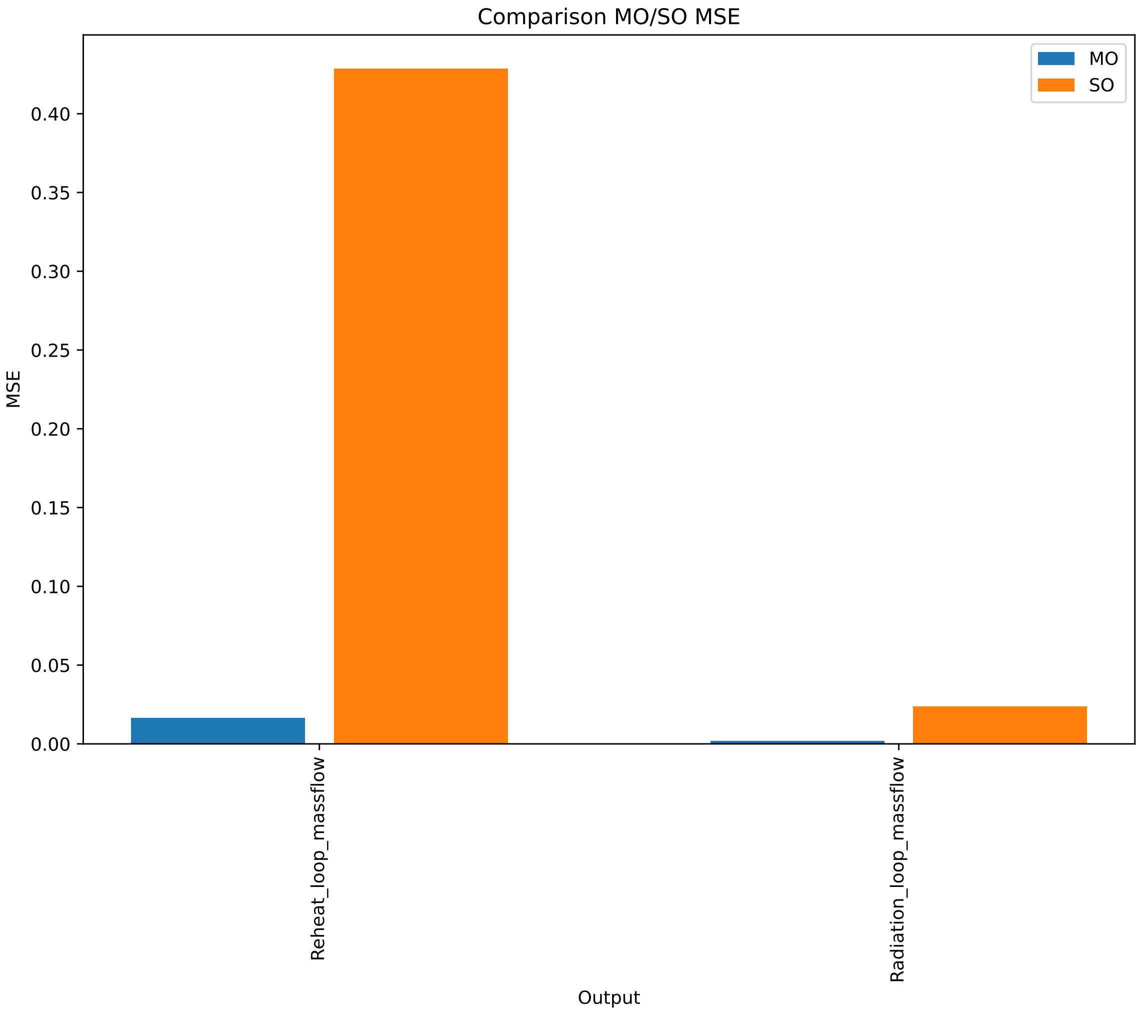


Figure 43: MSE values of each output of zone 10 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

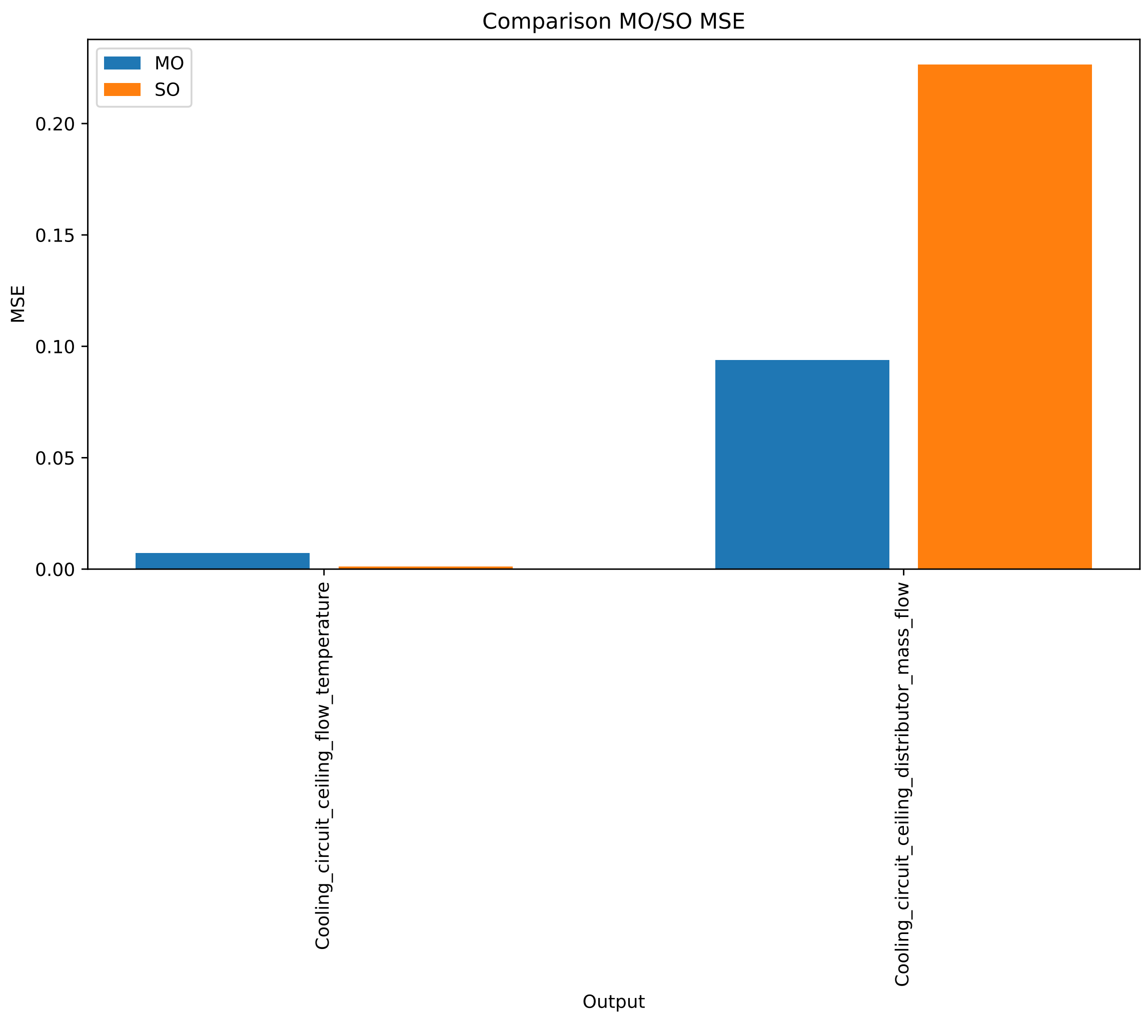


Figure 44: MSE values of each output of zone 11 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

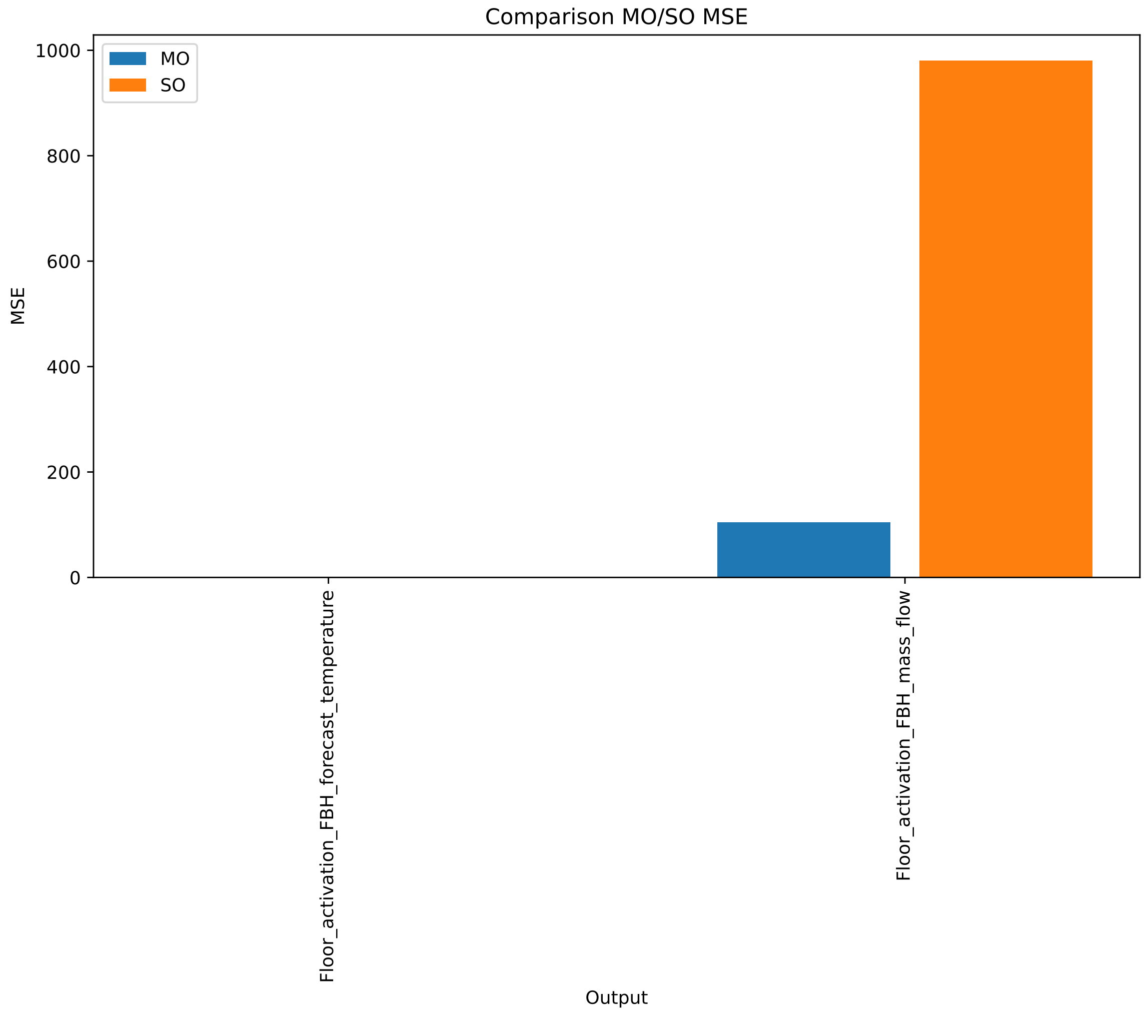


Figure 45: MSE values of each output of zone 12 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

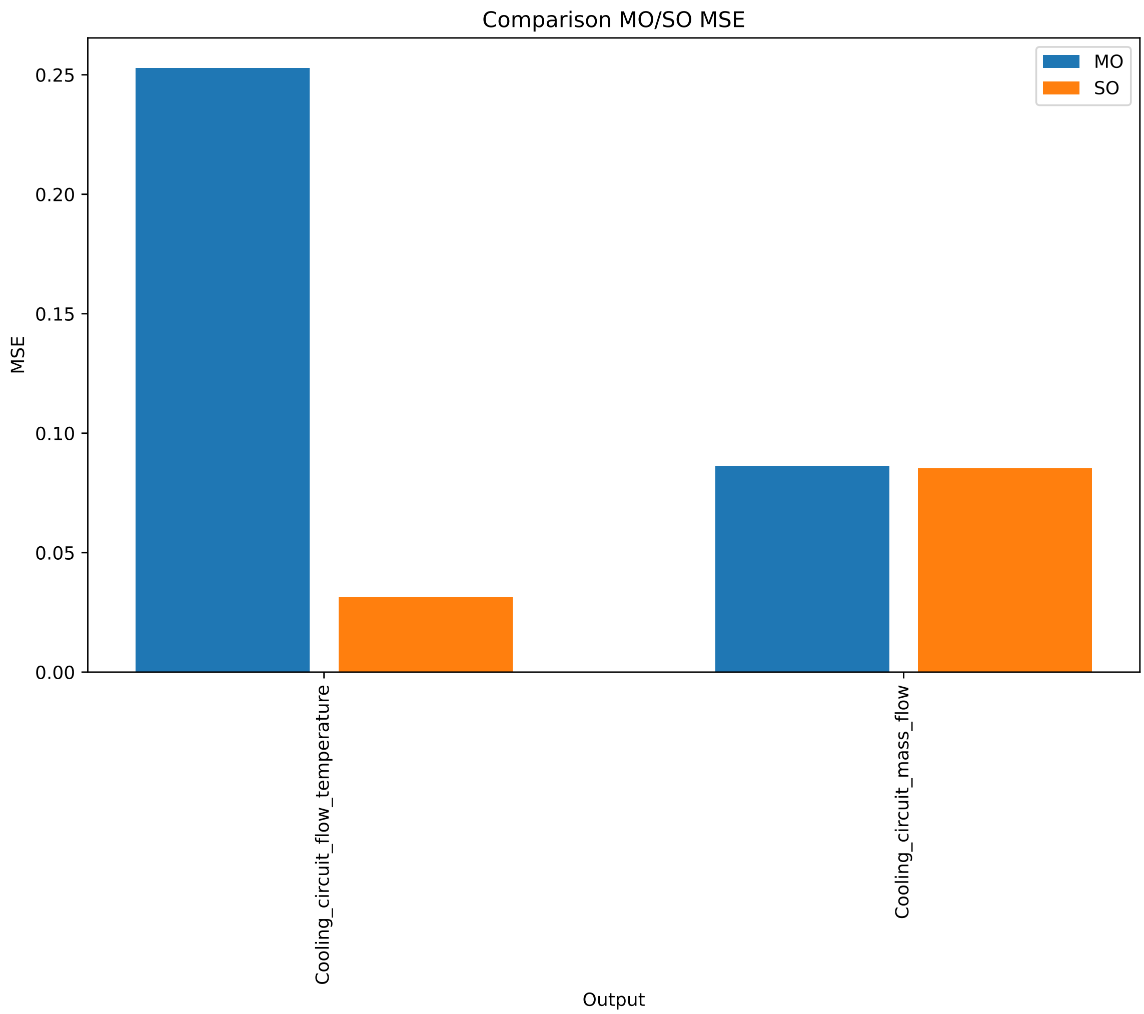


Figure 46: MSE values of each output of zone 13 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

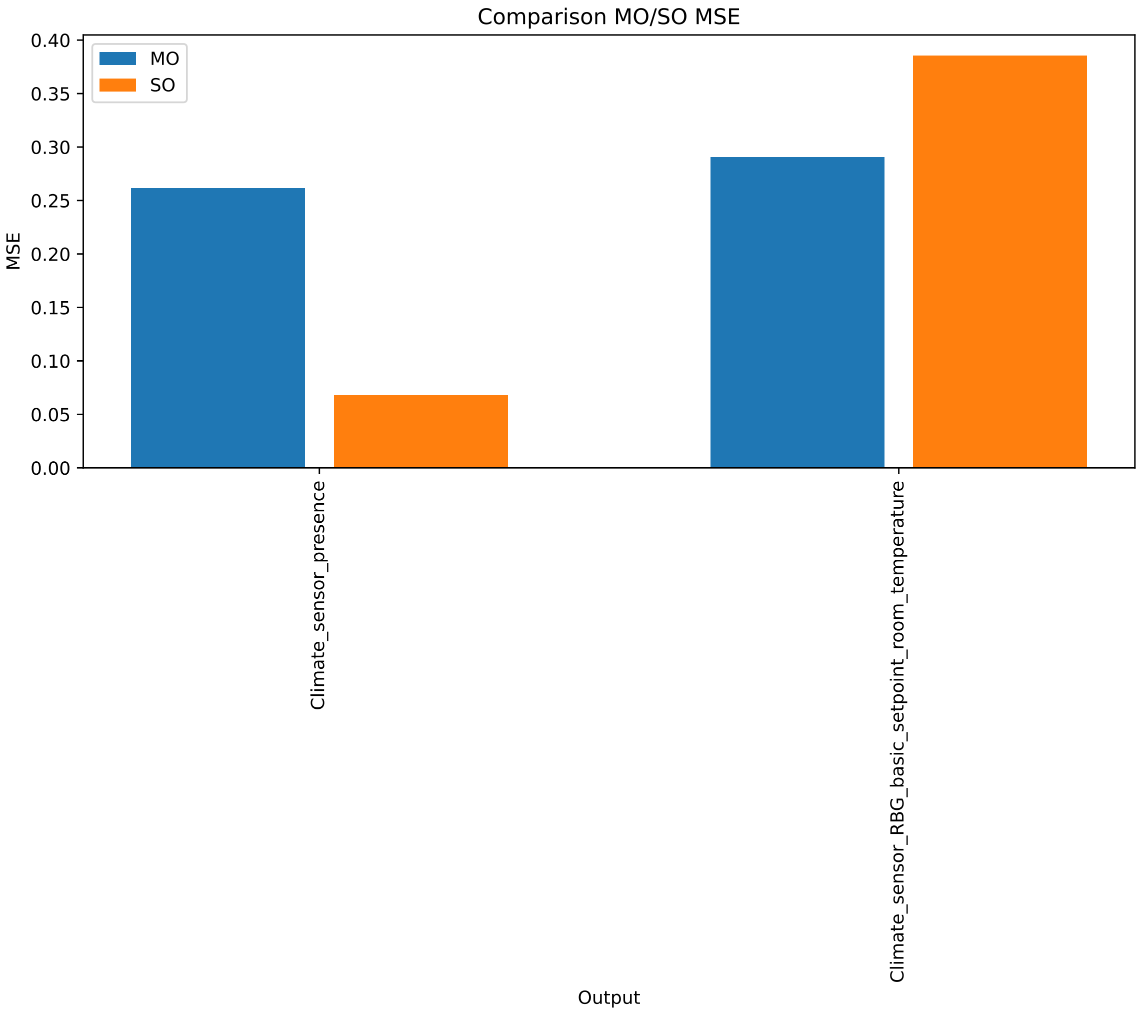


Figure 47: MSE values of each output of zone 14 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

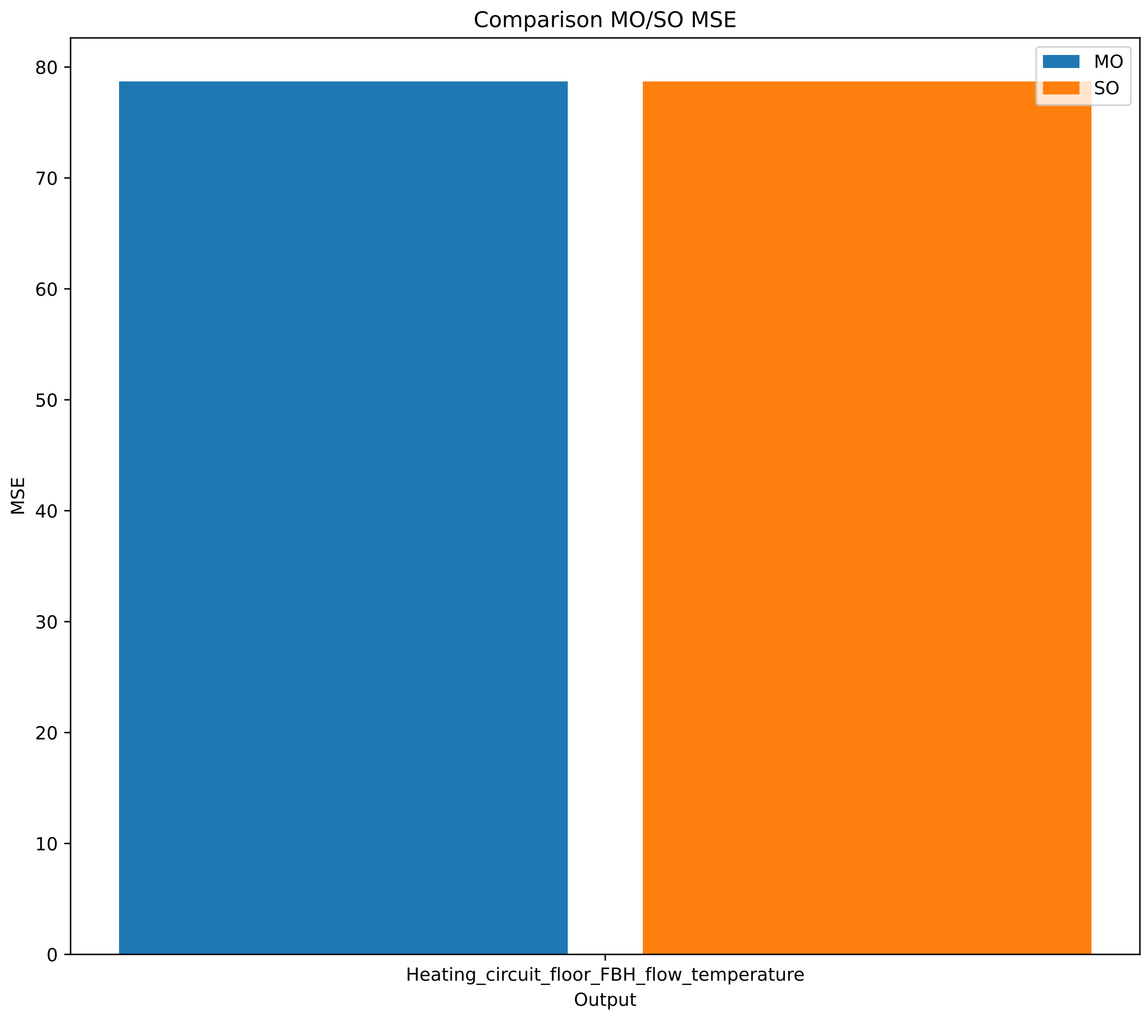


Figure 48: MSE values of each output of zone 15 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

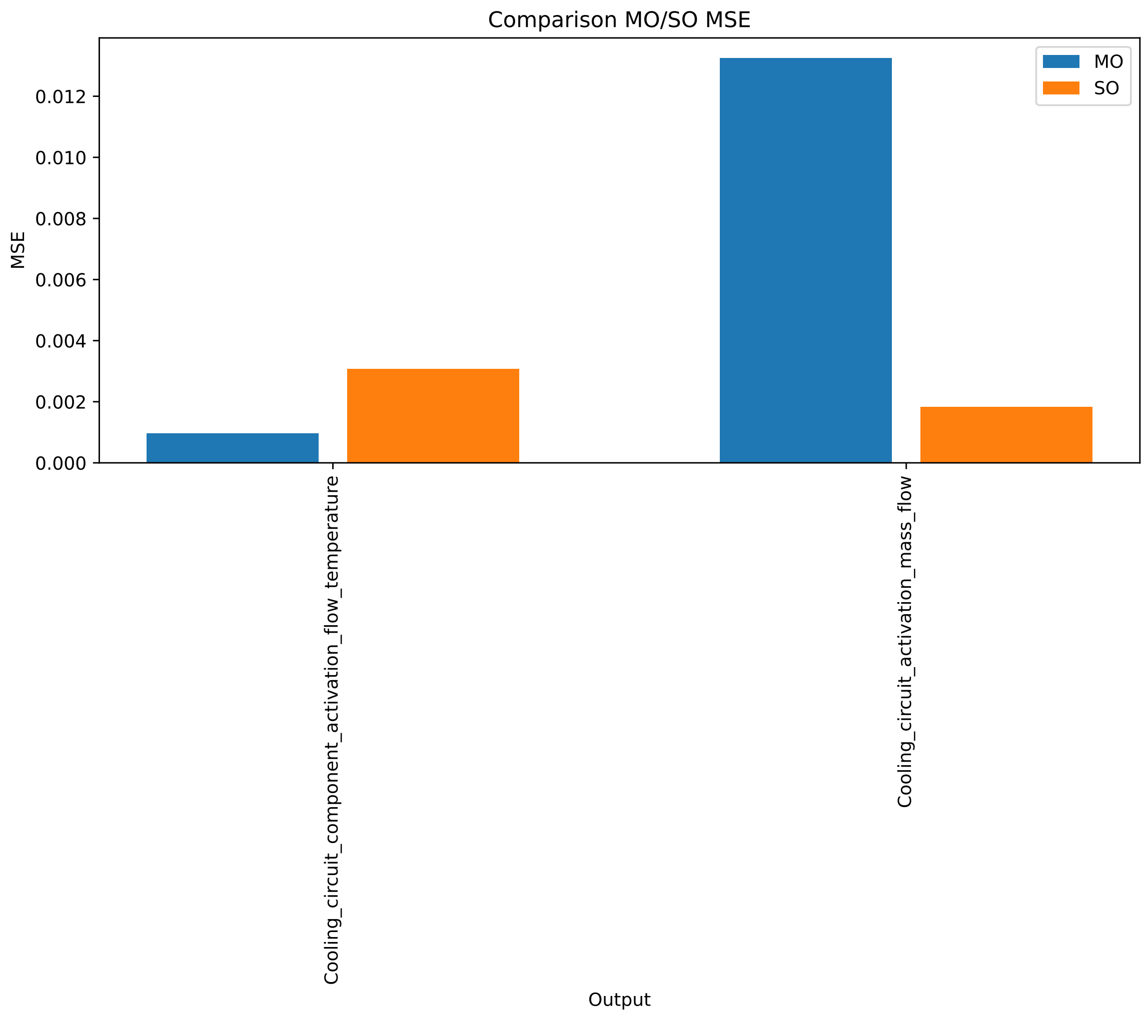


Figure 49: MSE values of each output of zone 17 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

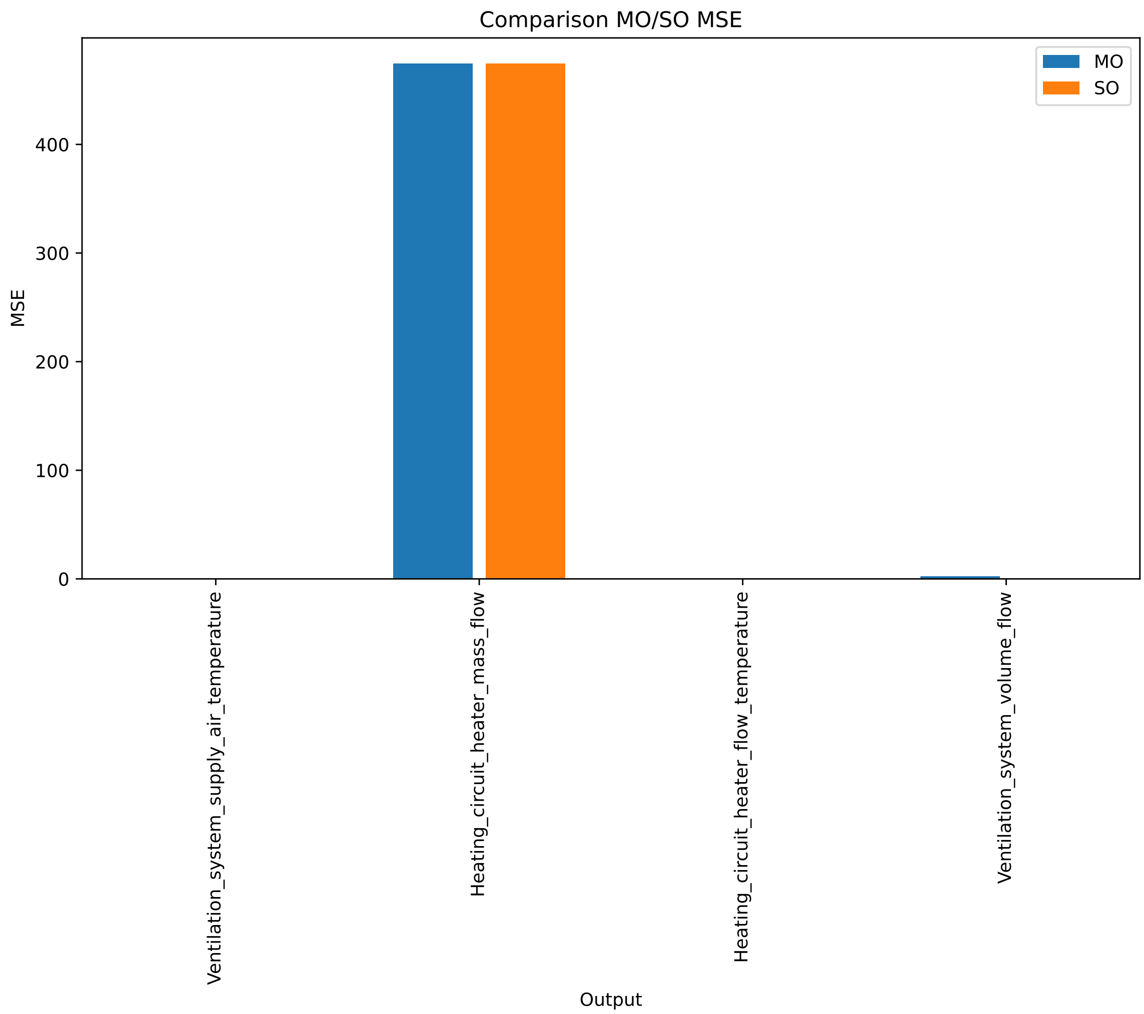


Figure 50: MSE values of each output of zone 18 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

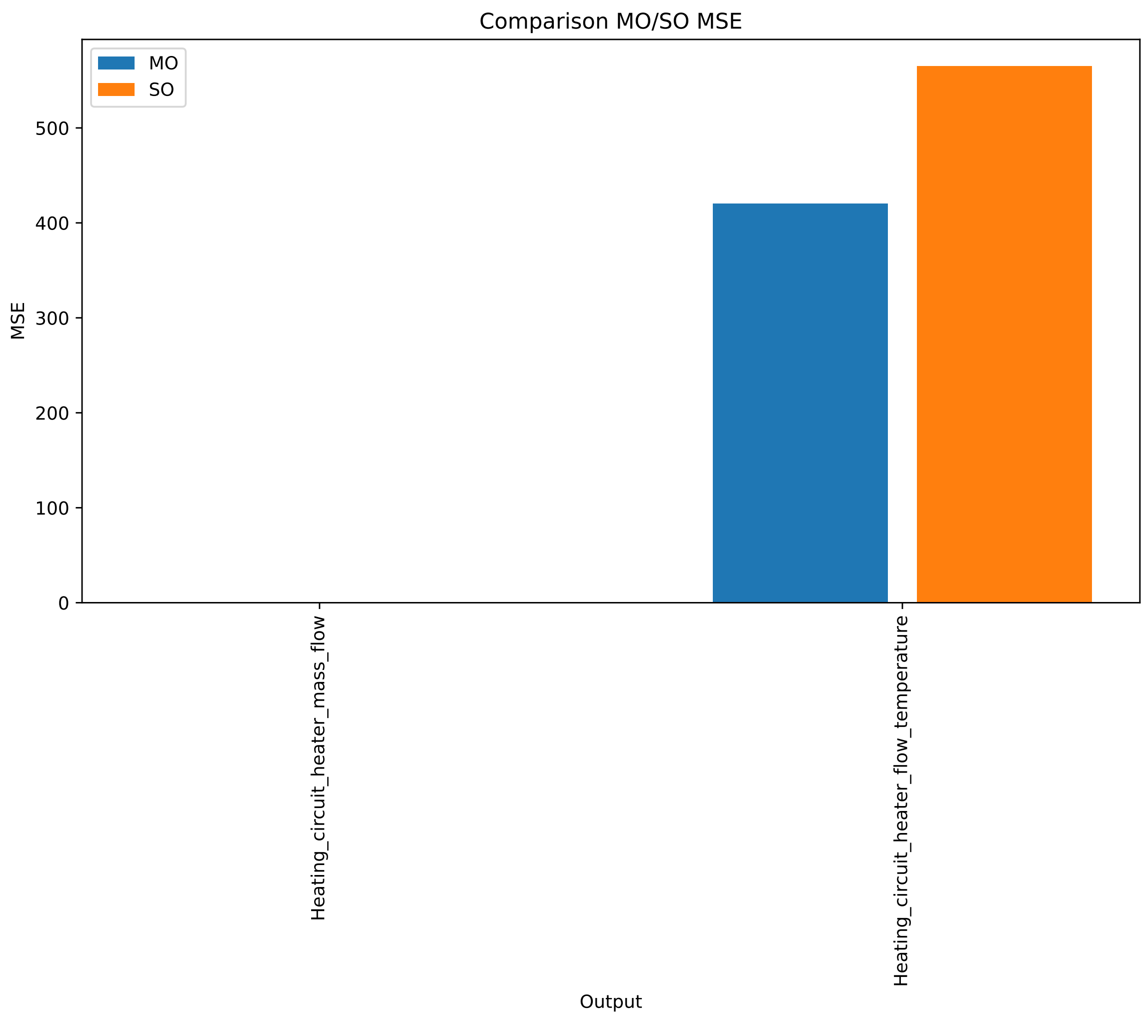


Figure 51: MSE values of each output of zone 20 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

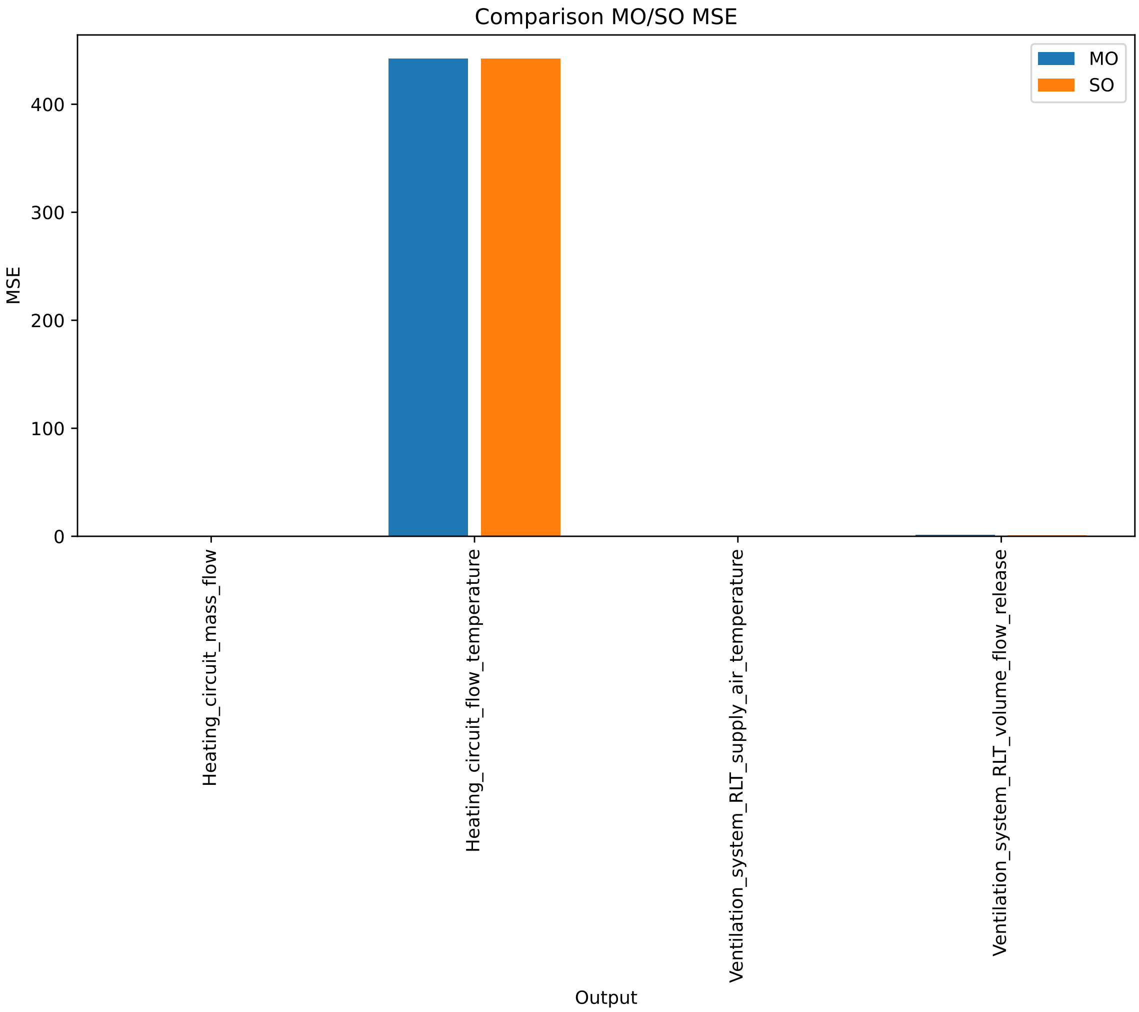


Figure 52: MSE values of each output of zone 21 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

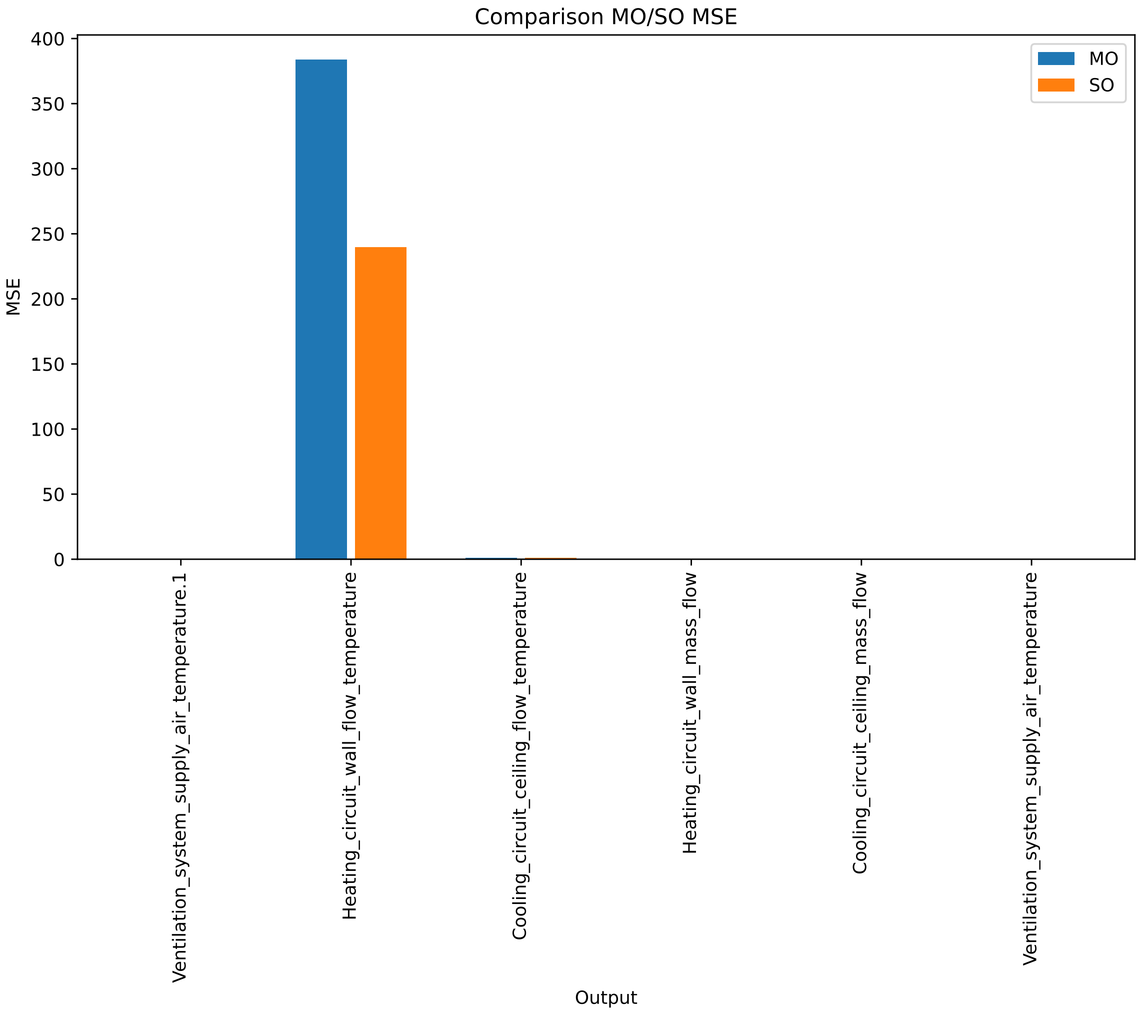


Figure 53: MSE values of each output of zone 22 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

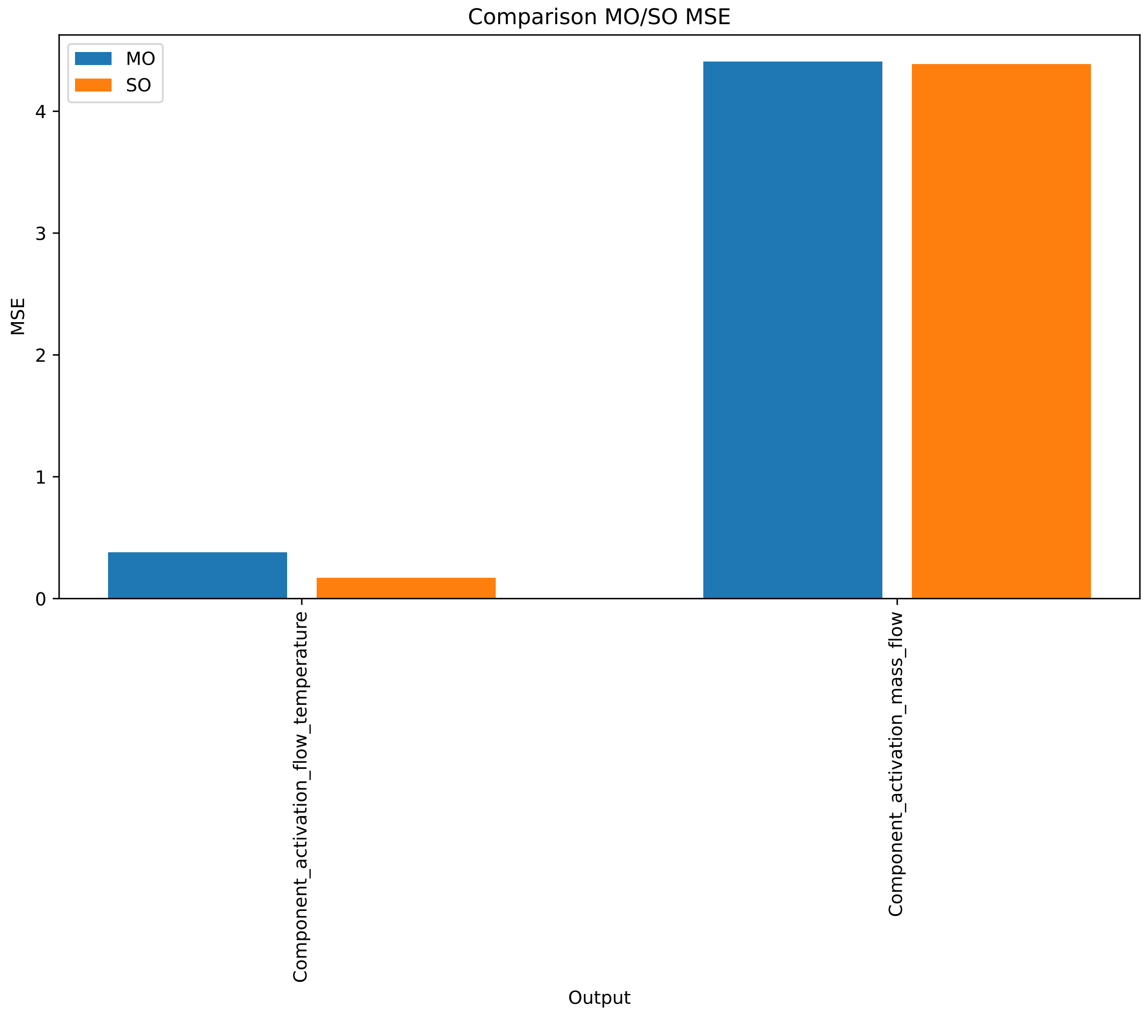


Figure 54: MSE values of each output of zone 23 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

E. History

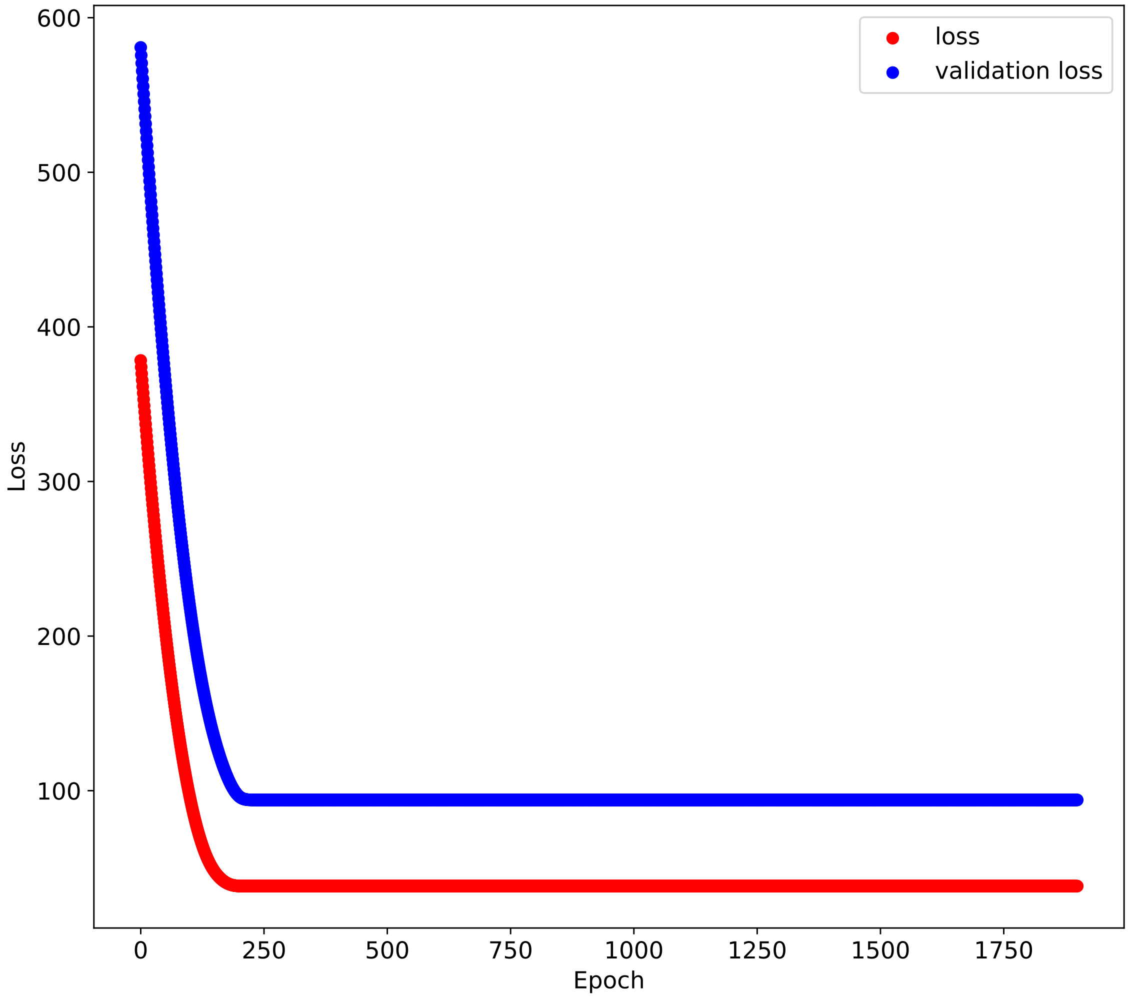
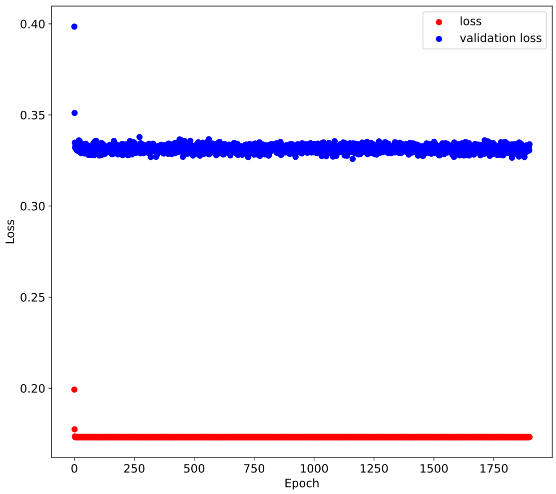
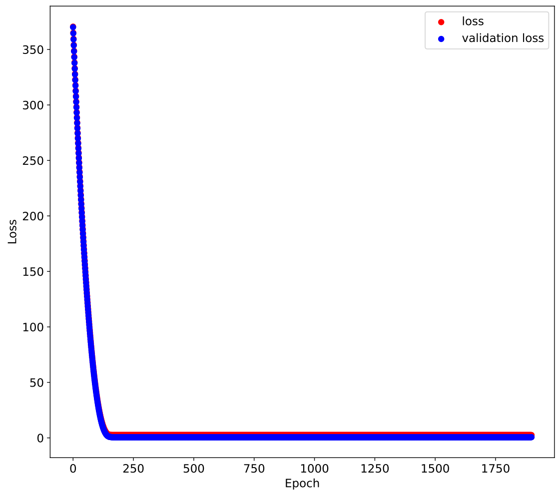


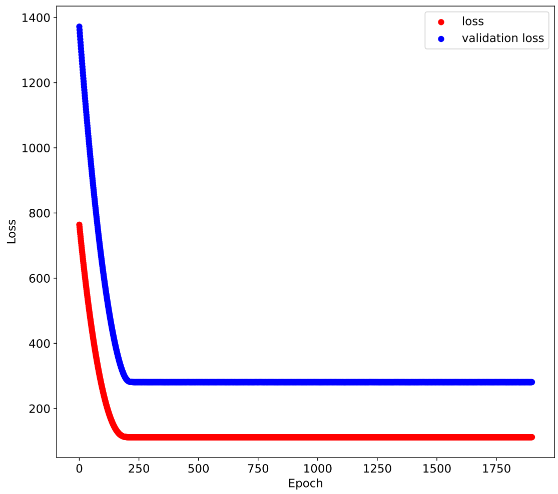
Figure 55: MO training history of zone 0 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Heating_circuit_heater_massflow



(b) Ventilation_system_supply_air_temperature_chute



(c) Heating_circuit_heater_flow_temperature

Figure 56: SO training history of zone 0 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

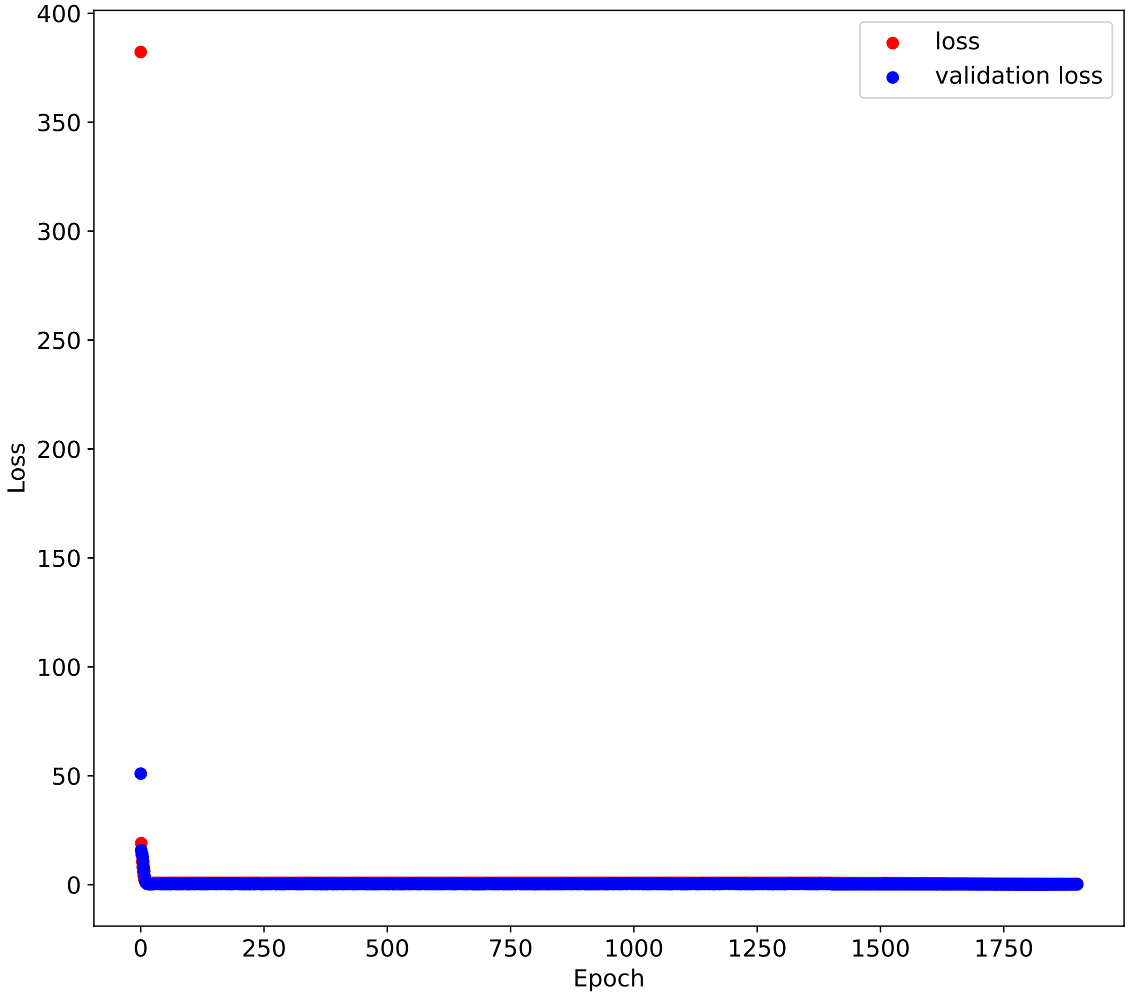


Figure 57: Training history of zone 1 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output (only one output)

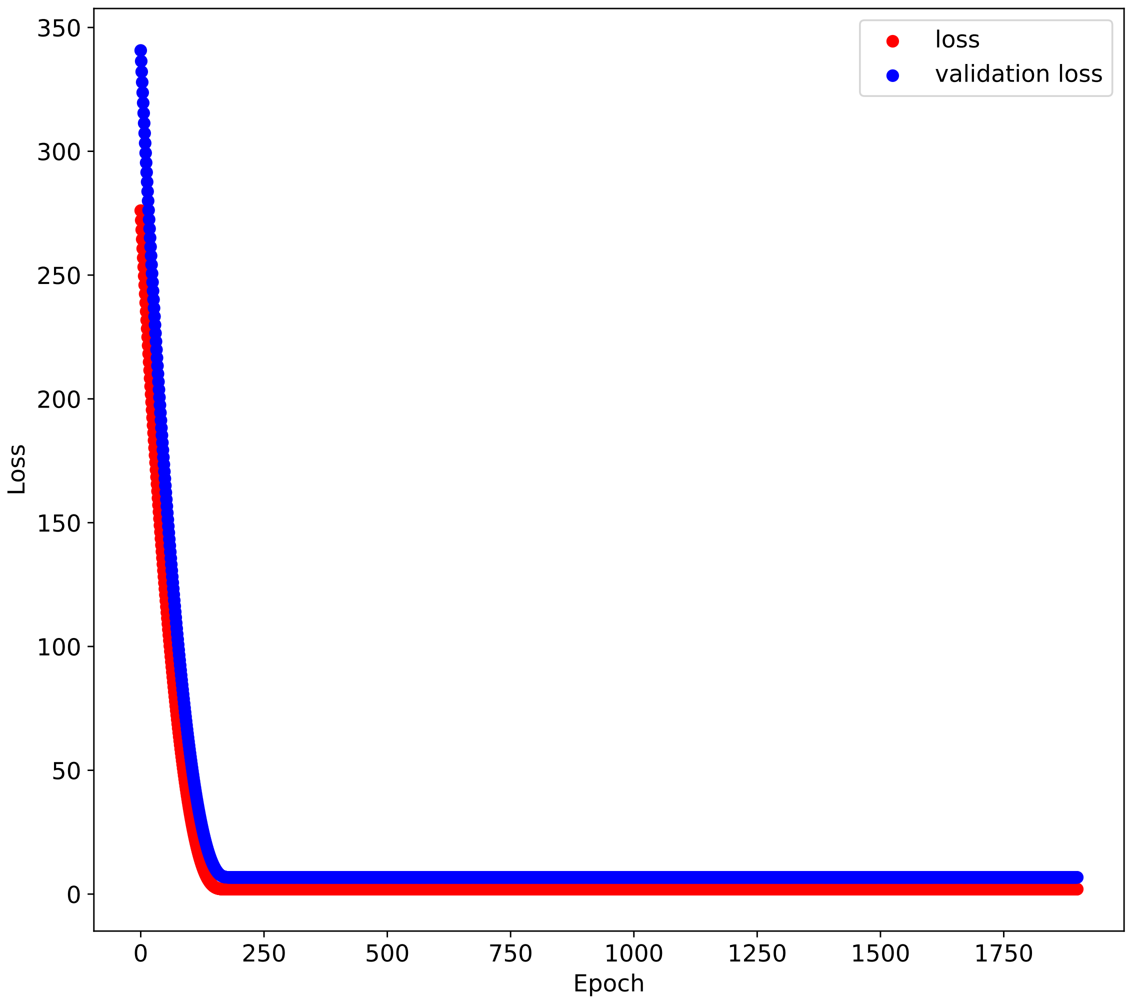
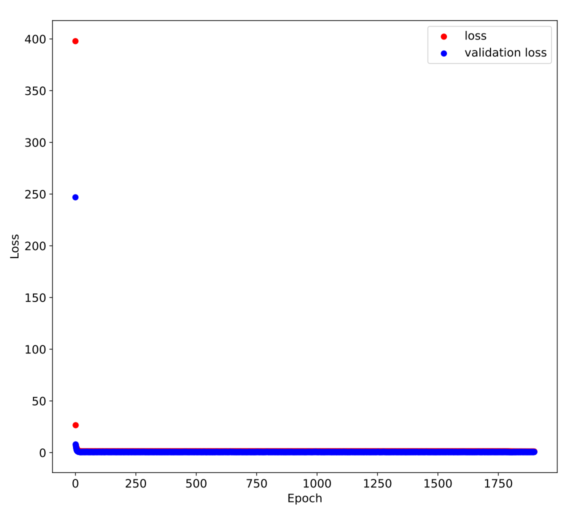
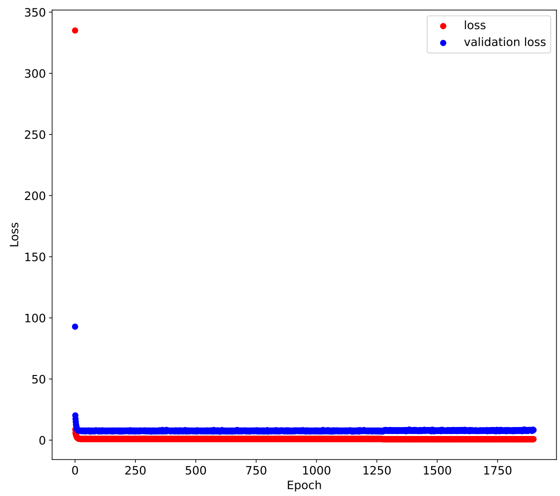


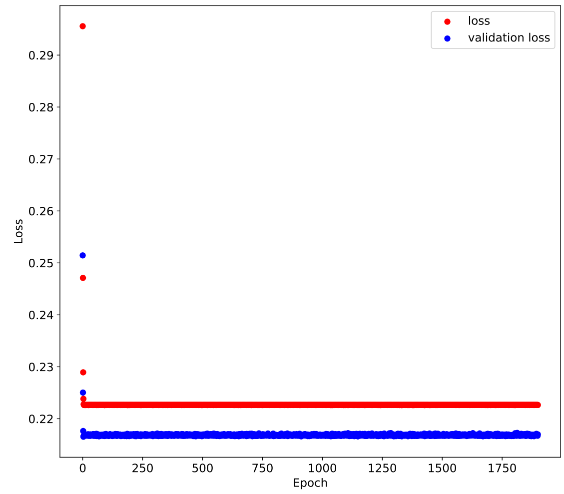
Figure 58: MO training history of zone 2 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Ventilation_system_supply_air_temperature



(b) Individual_room_control_heating_circuit_room_temperature_basic_setpoint



(c) Ventilation_system_volume_flow_release

Figure 59: SO training history of zone 2 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

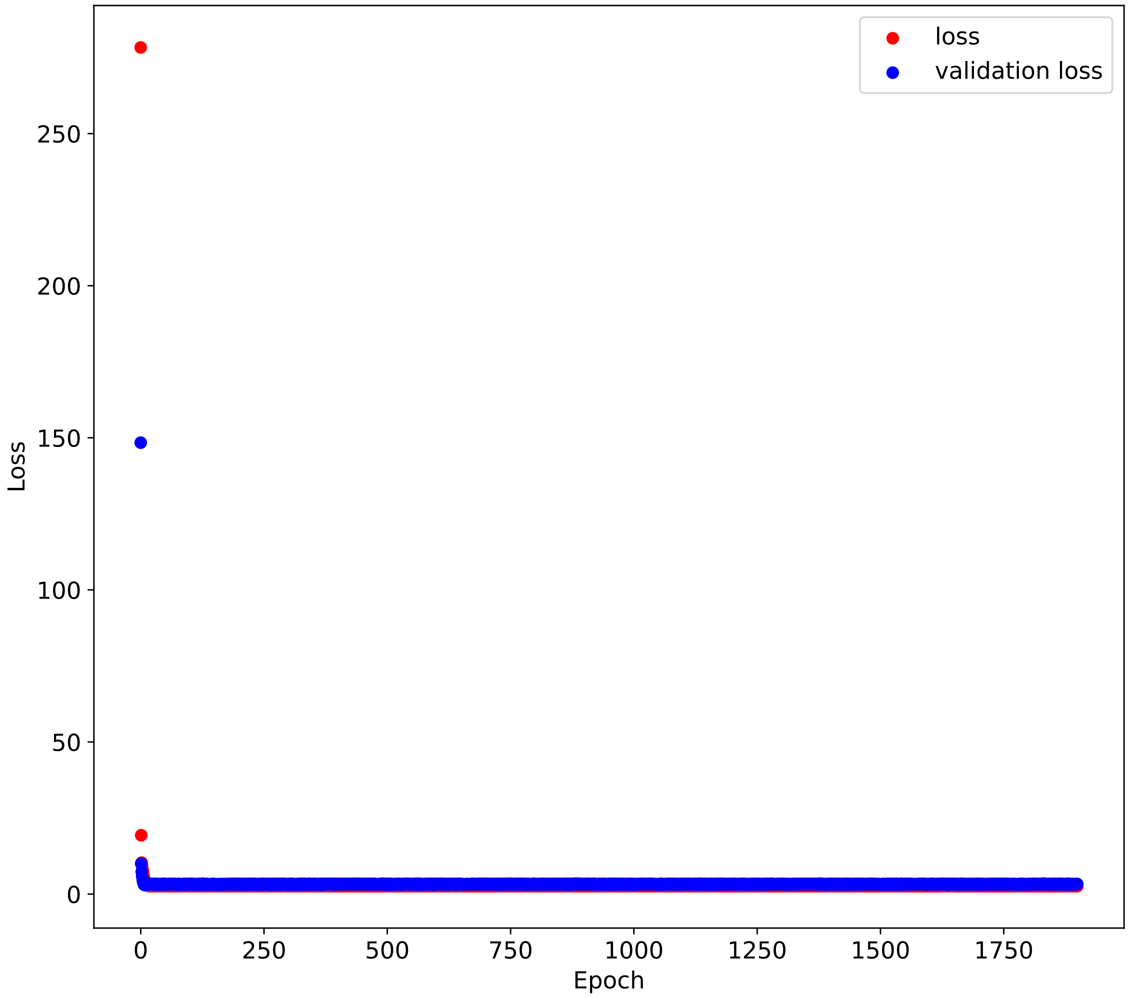
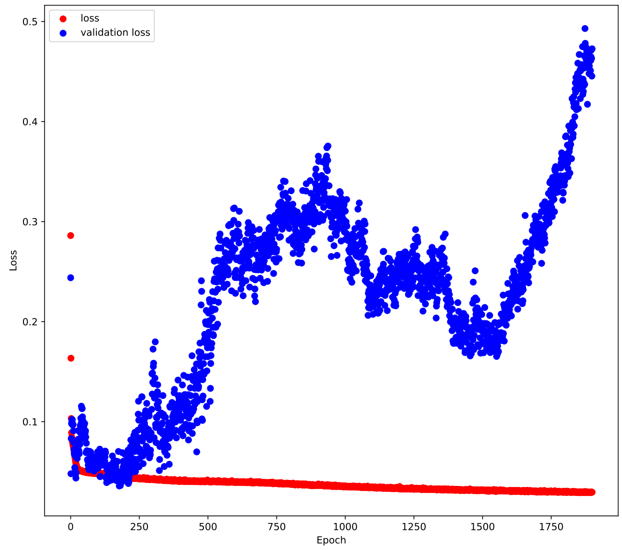
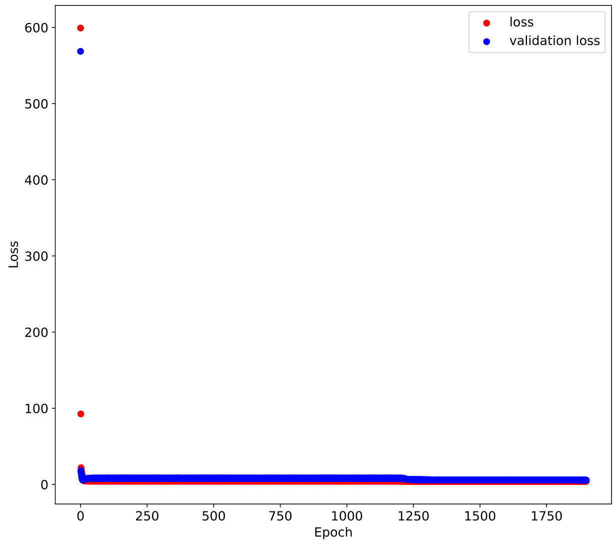


Figure 60: MO training history of zone 3 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Heating_circuit_floor_FBH_mass_flow



(b) Heating_circuit_floor_FBH_flow_temperature

Figure 61: SO training history of zone 3 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

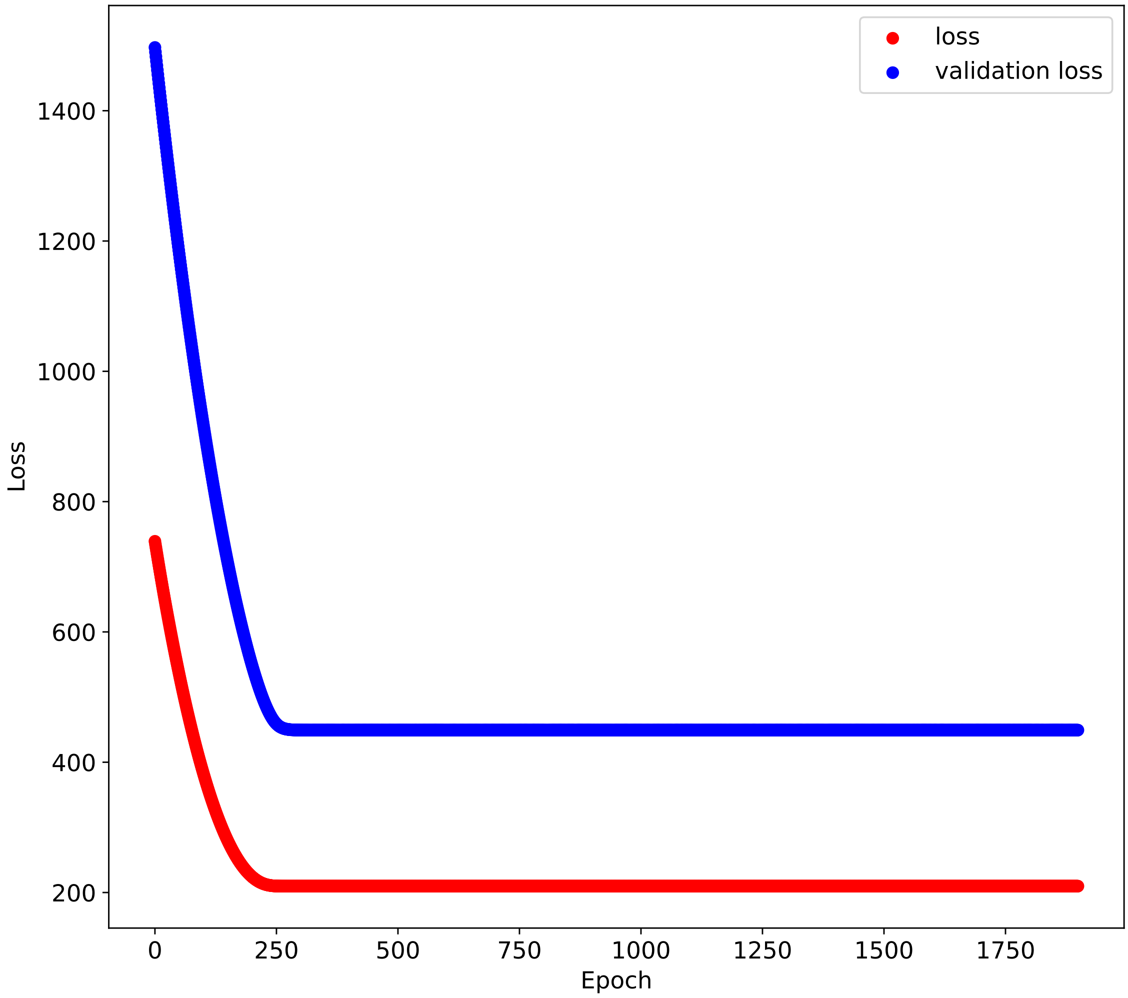
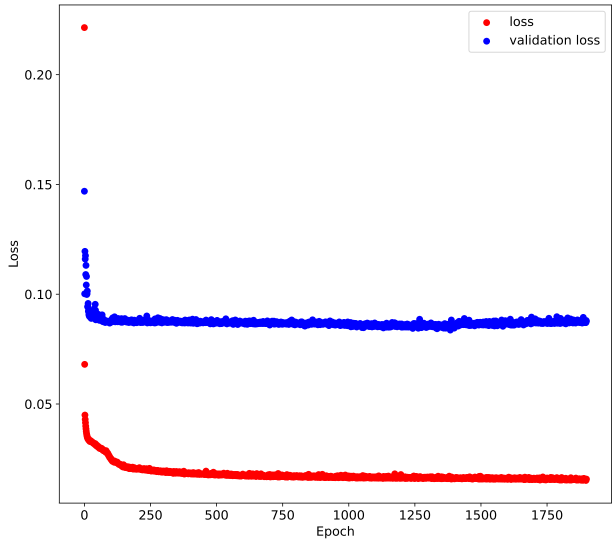
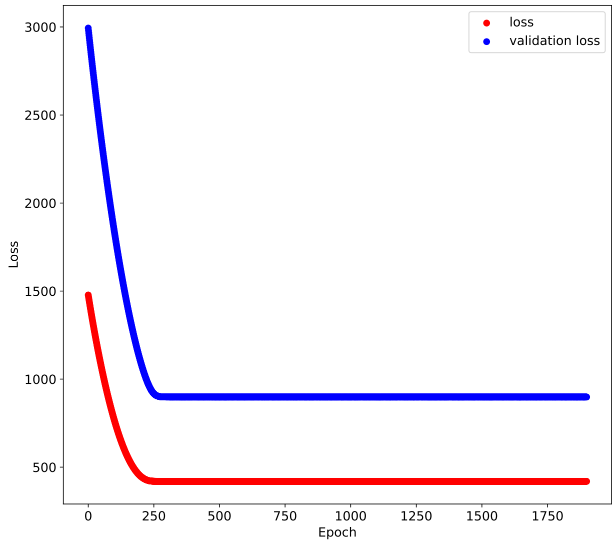


Figure 62: MO training history of zone 4 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Heating_circuit_mass_flow



(b) Heating_circuit_pharmacy_flow_temperature

Figure 63: SO training history of zone 4 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

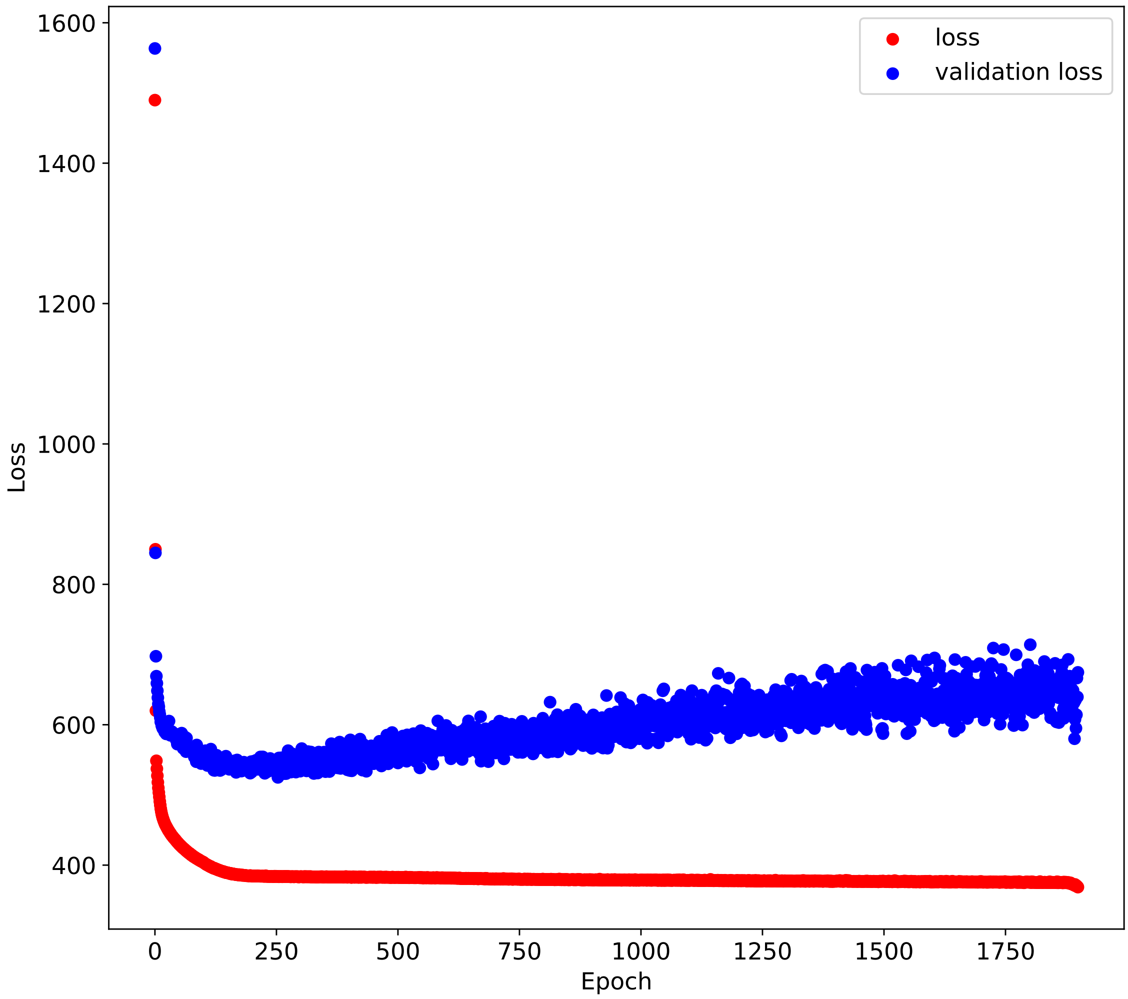
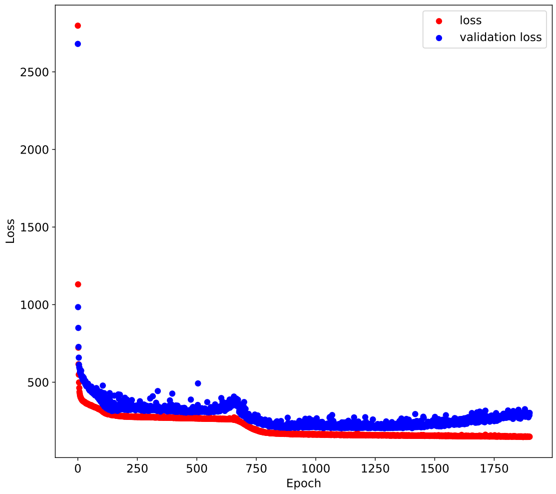
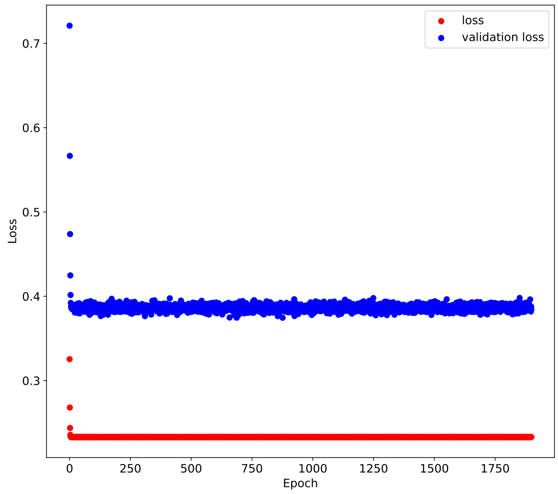


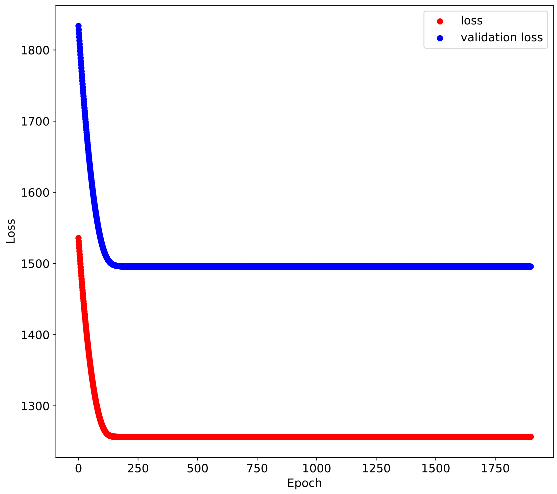
Figure 64: MO training history of zone 5 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Ventilation_volume_flow_1b

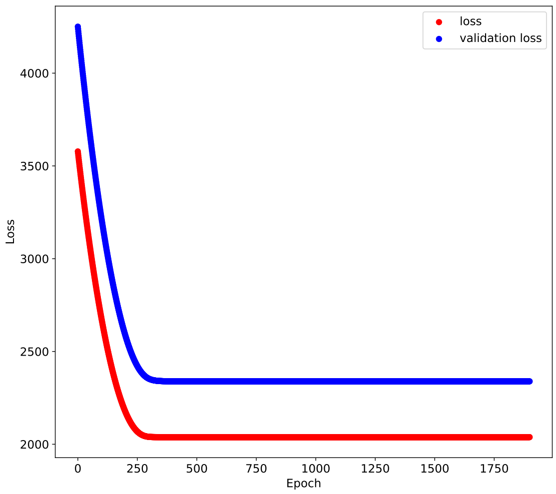


(b) Heating_circuit_heater_1_mass_flow

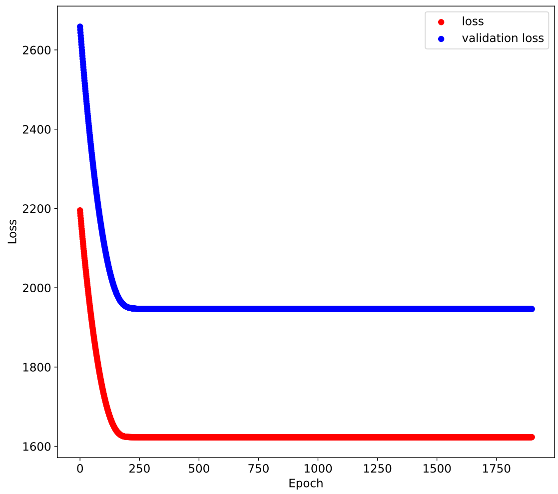


(c) Ventilation_system_volume_flow_2c

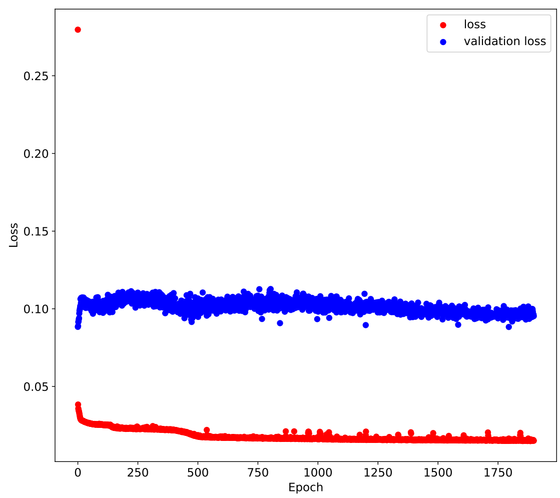
Figure 65: SO training history of zone 5 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Ventilation_system_volume_flow_1a

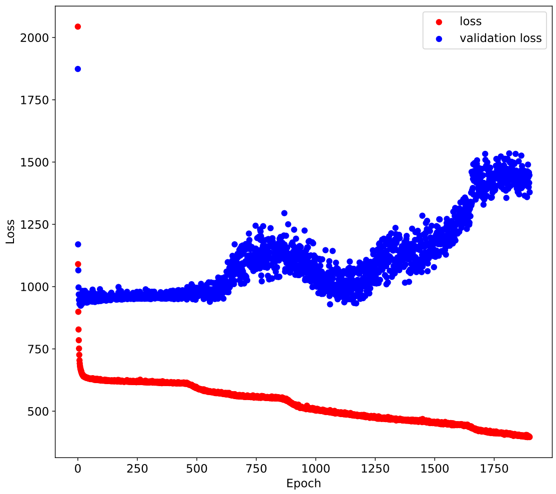


(b) Ventilation_system_volume_flow_2b

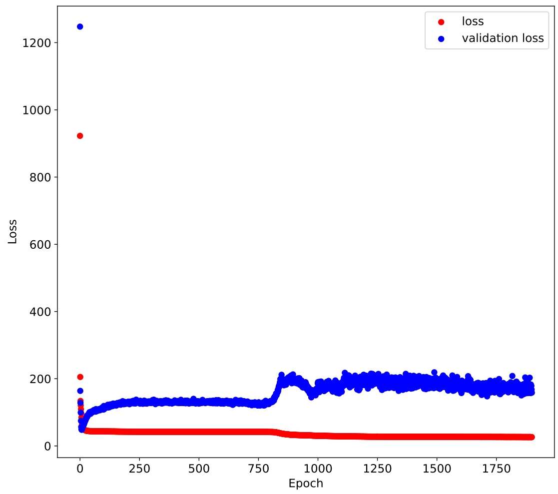


(c) Ventilation_system_RLT_volume_flow_release

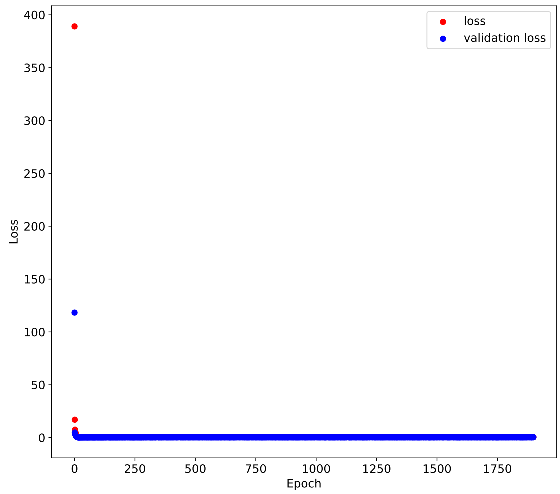
Figure 66: SO training history of zone 5 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Ventilation_system_volume_flow_2a



(b) Heating_circuit_heater_1_flow_temperature



(c) Ventilation_system_RLT_supply_air_temperature

Figure 67: SO training history of zone 5 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

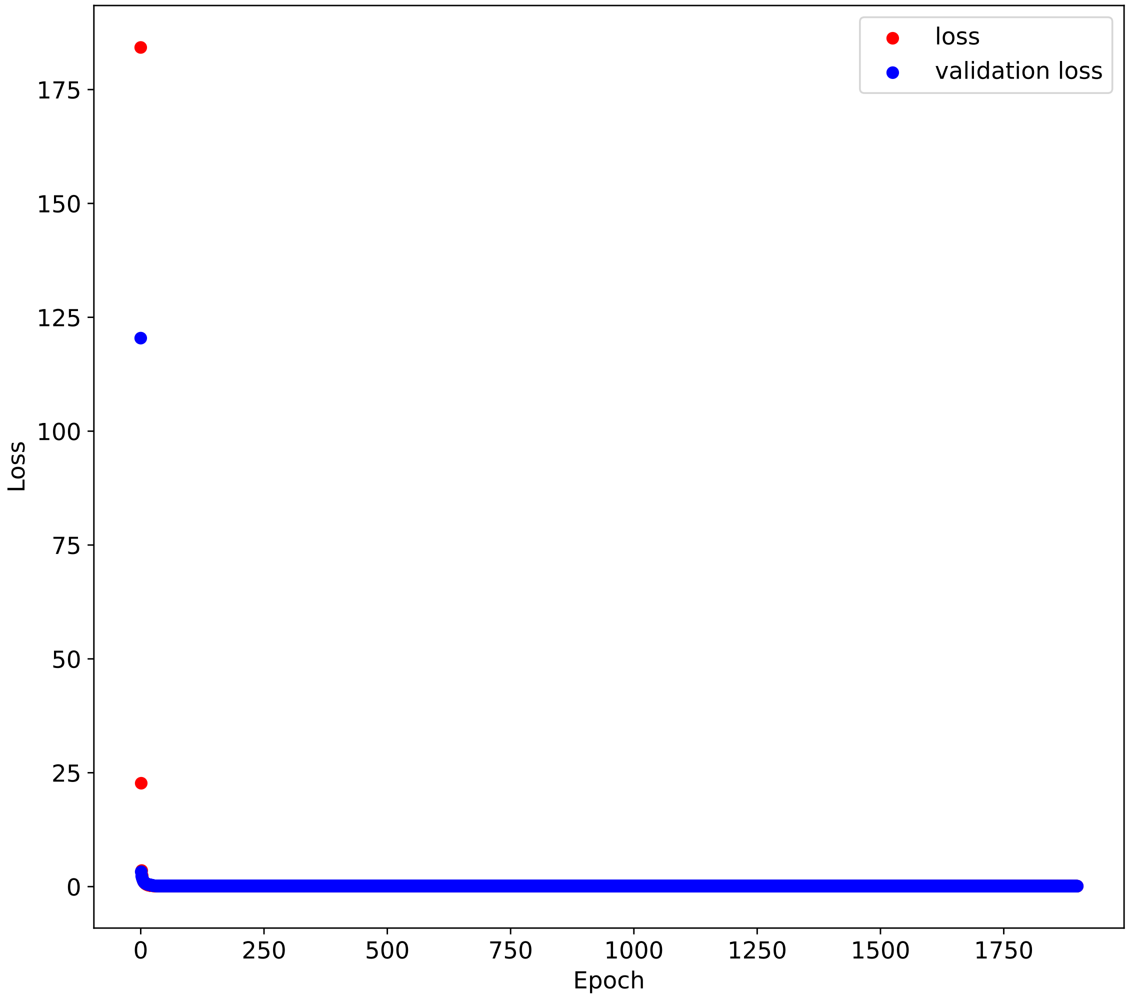
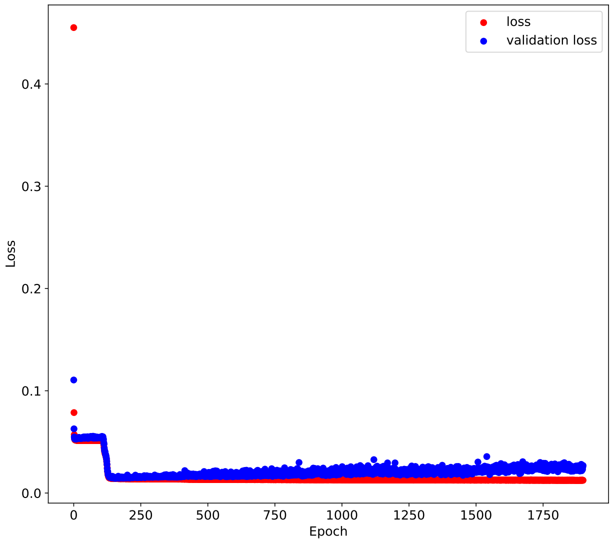
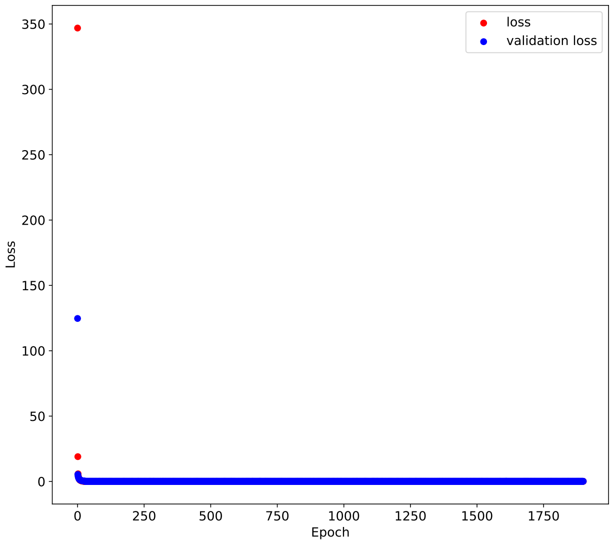


Figure 68: MO training history of zone 7 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Ventilation_system_RLT_volume_flow_release1



(b) Ventilation_system_RLT_supply_air_temperature

Figure 69: SO training history of zone 7 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

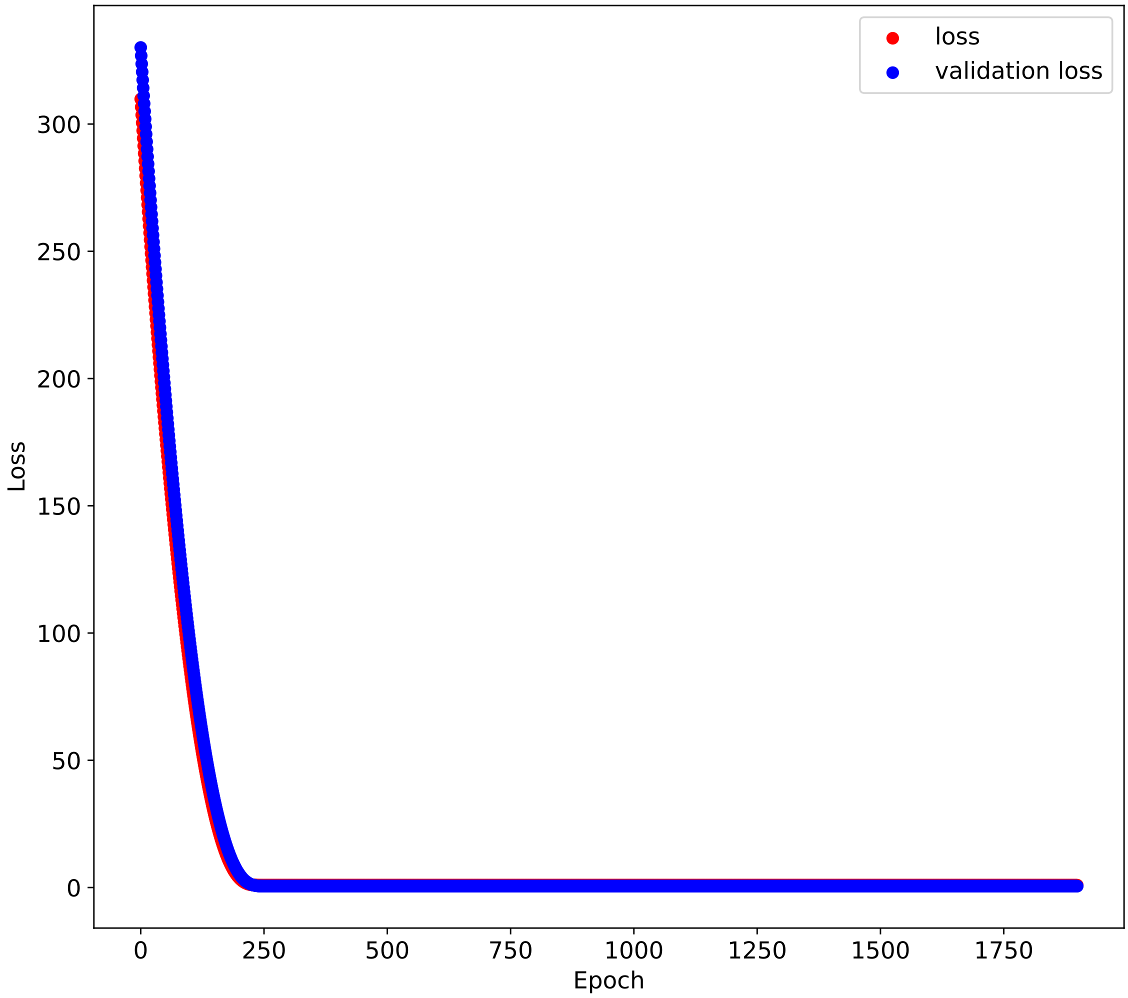
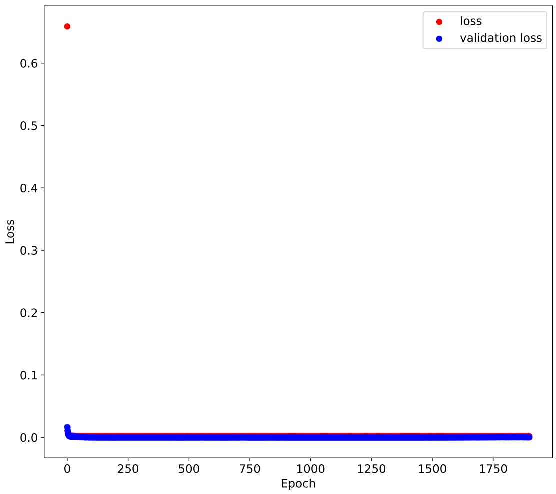
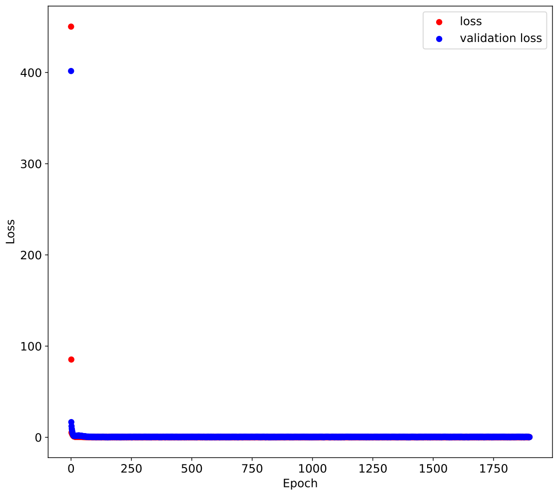


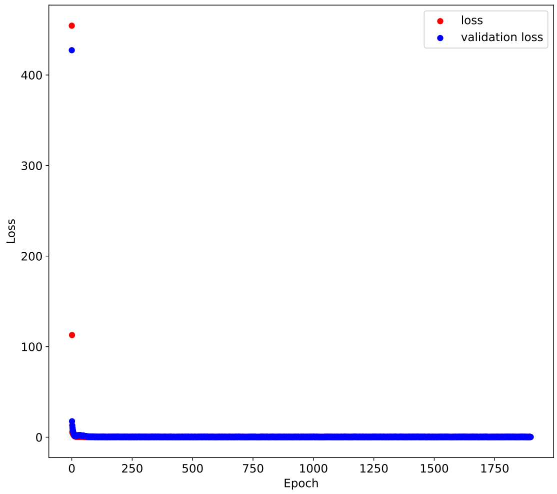
Figure 70: MO training history of zone 8 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Ventilation_system_volume_flow_release



(b) Ventilation_system_supply_air_temperature_R



(c) Ventilation_system_supply_air_temperature_A

Figure 71: SO training history of zone 8 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

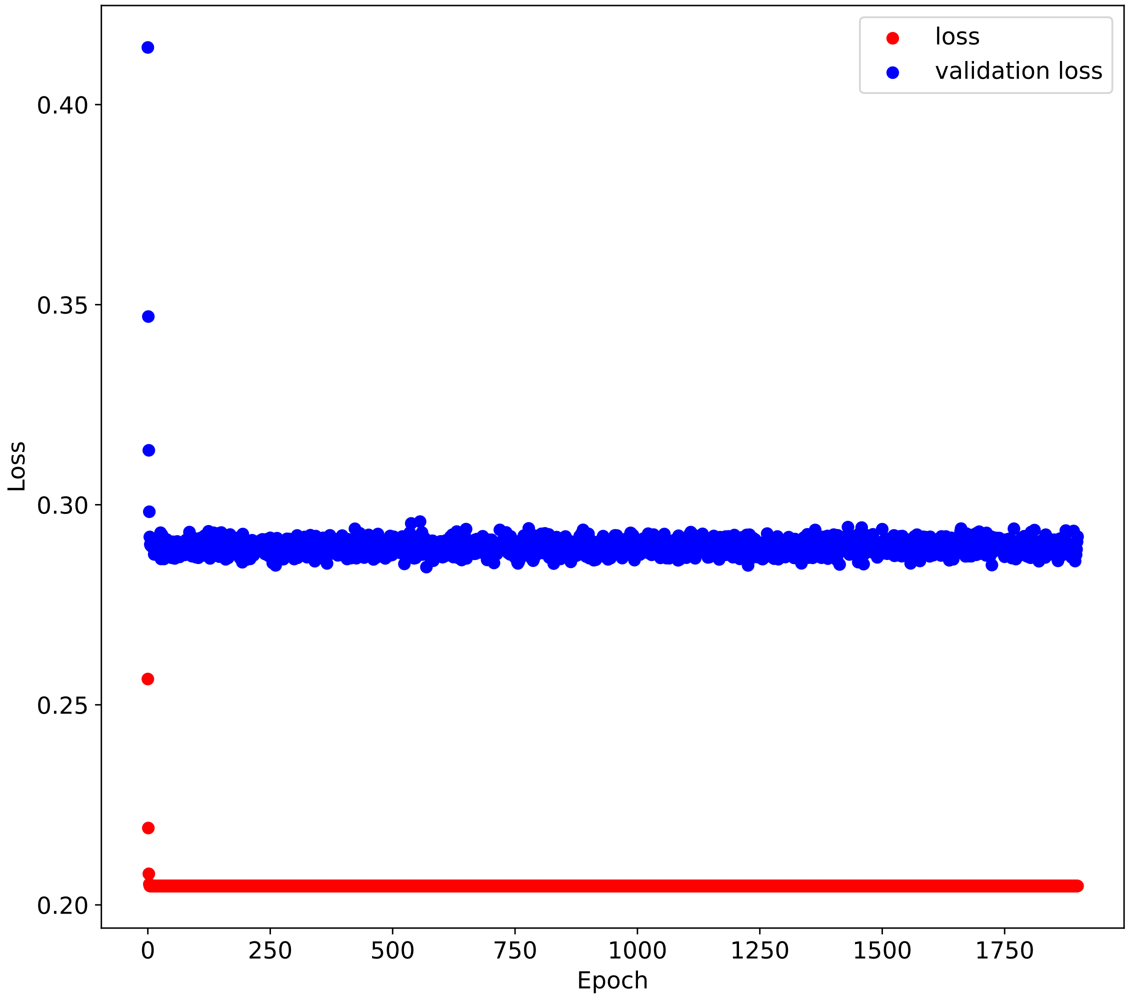
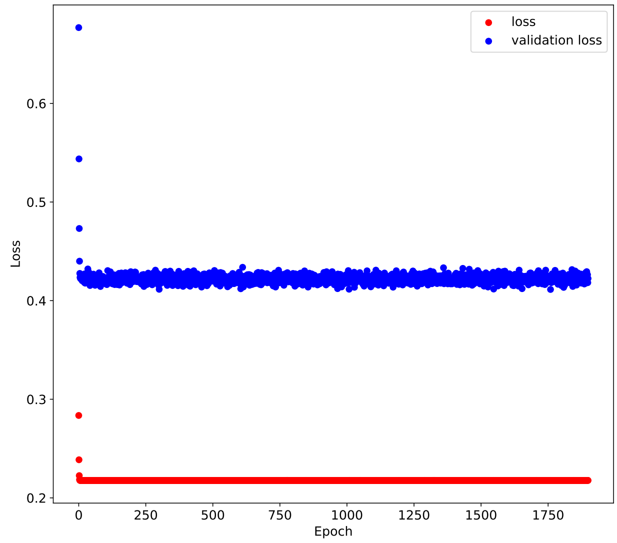
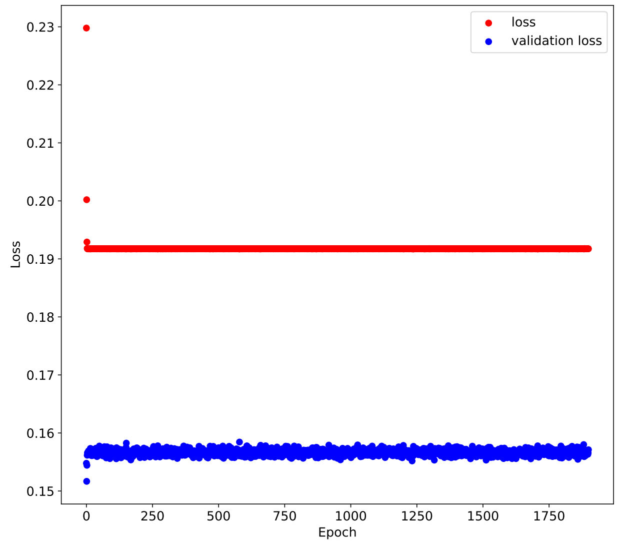


Figure 72: MO training history of zone 9 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Radiation_loop_massflow



(b) Reheat_loop_massflow

Figure 73: SO training history of zone 9 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

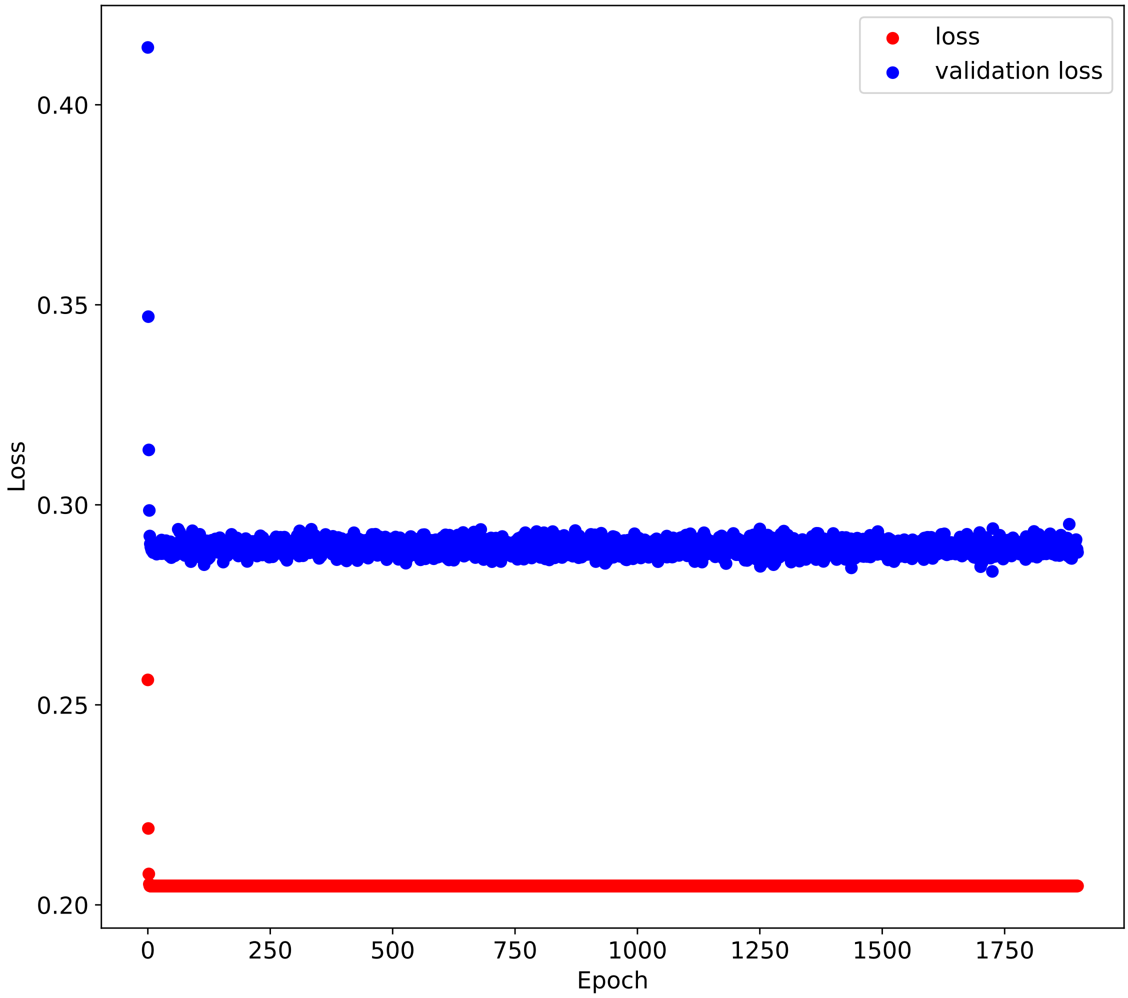
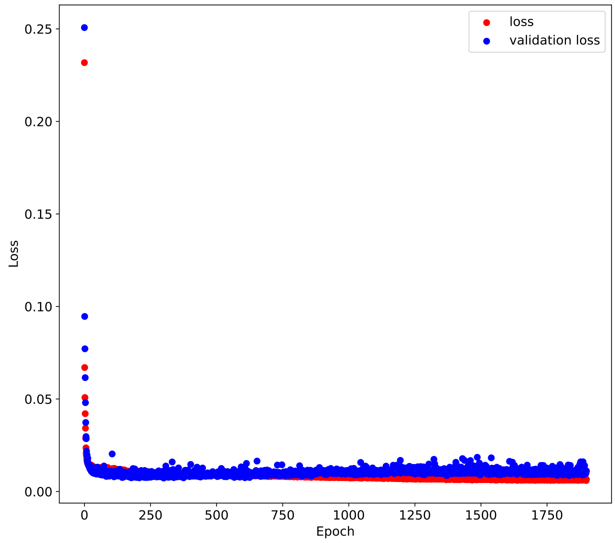
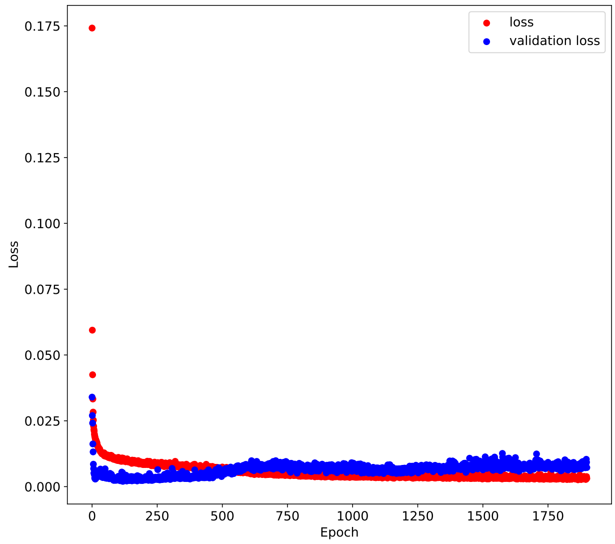


Figure 74: MO training history of zone 10 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Radiation_loop_massflow



(b) Reheat_loop_massflow

Figure 75: SO training history of zone 10 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

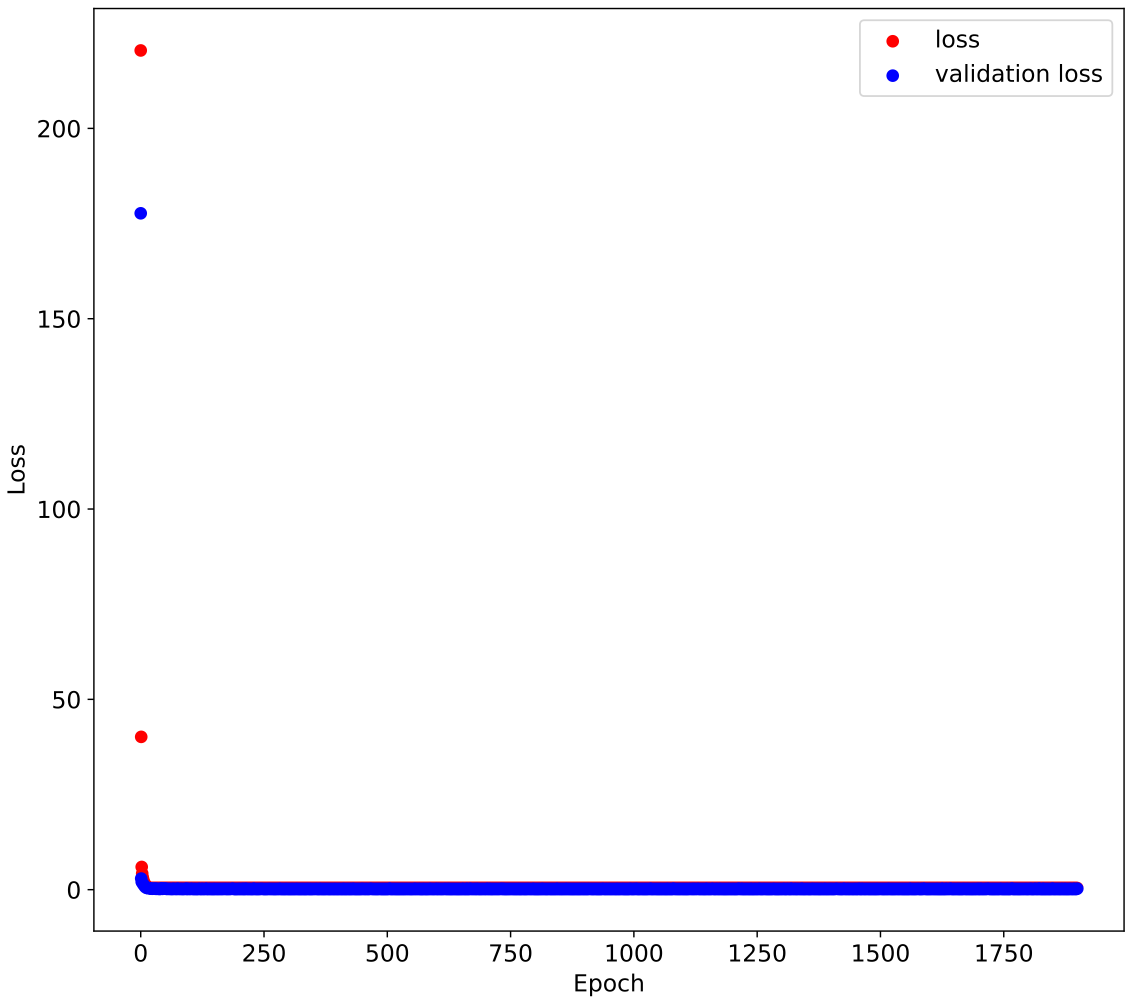
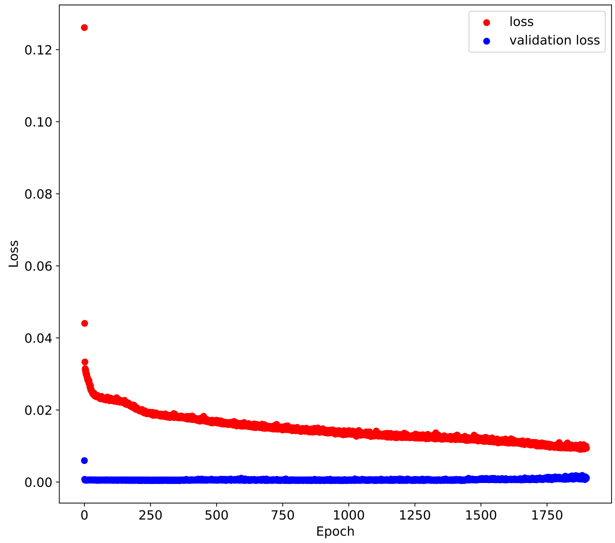
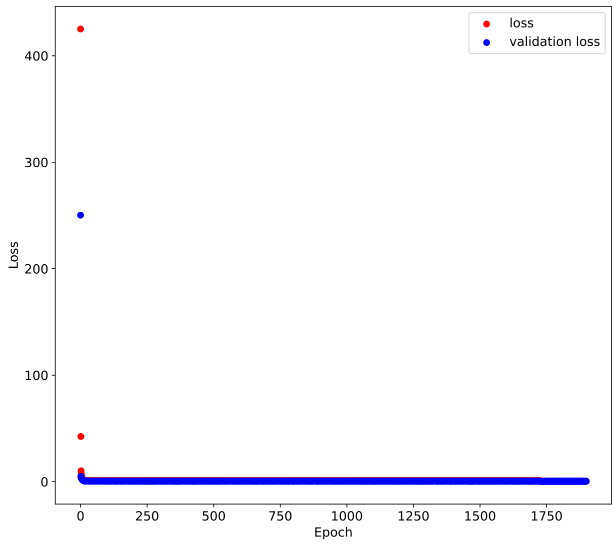


Figure 76: MO training history of zone 11 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Cooling_circuit_ceiling_distributor_mass_flow



(b) Cooling_circuit_ceiling_flow_temperature

Figure 77: SO training history of zone 11 when considering 1 year of data and after dropping non-forecast data and data with no correlation >0.1 to at least one output

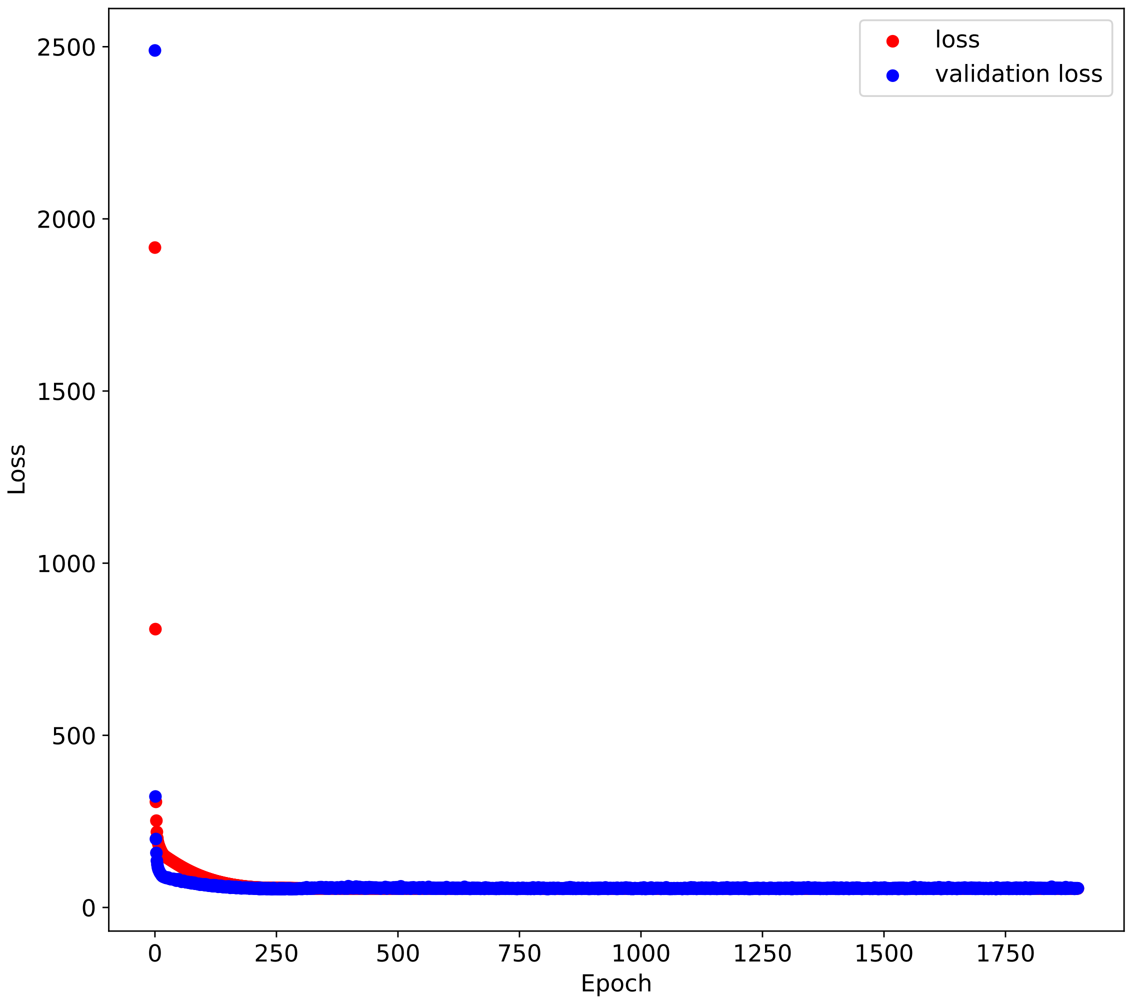
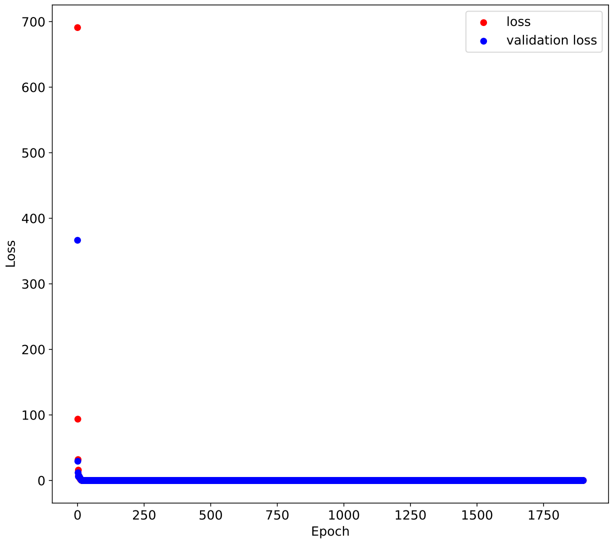
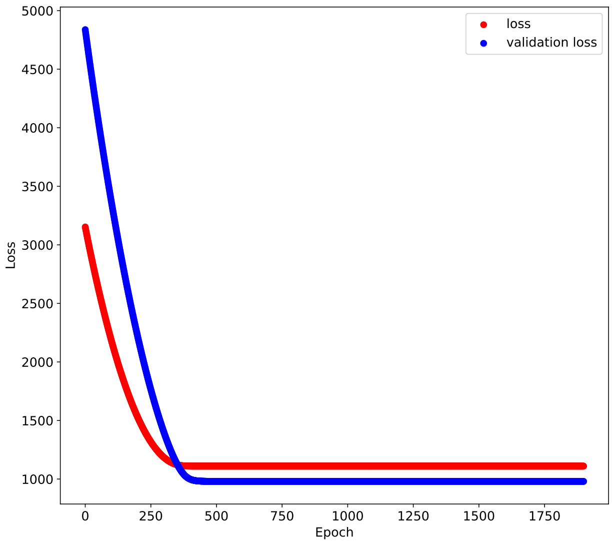


Figure 78: MO training history of zone 12 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Floor_activation_FBH_mass_flow



(b) Floor_activation_FBH_forecast_temperature

Figure 79: SO training history of zone 12 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

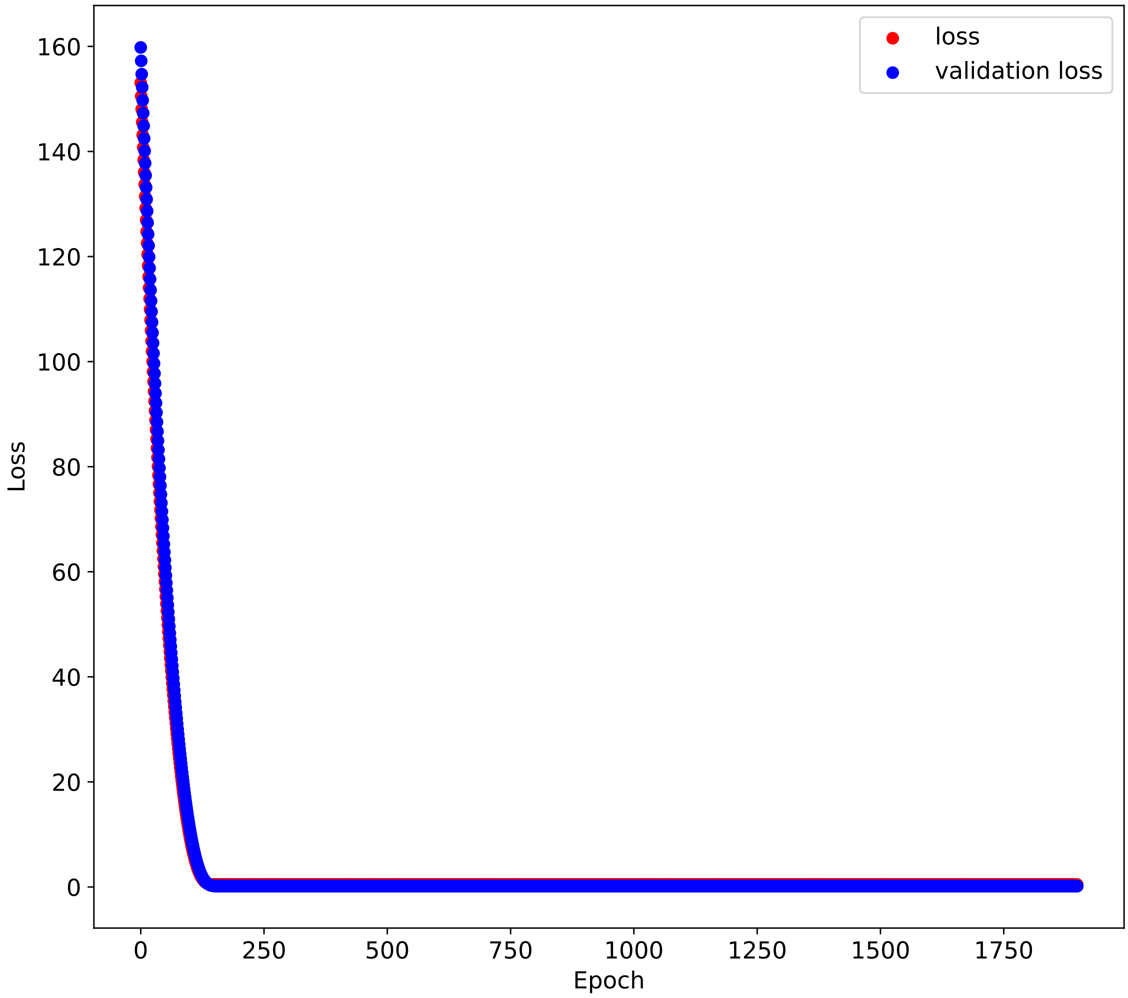
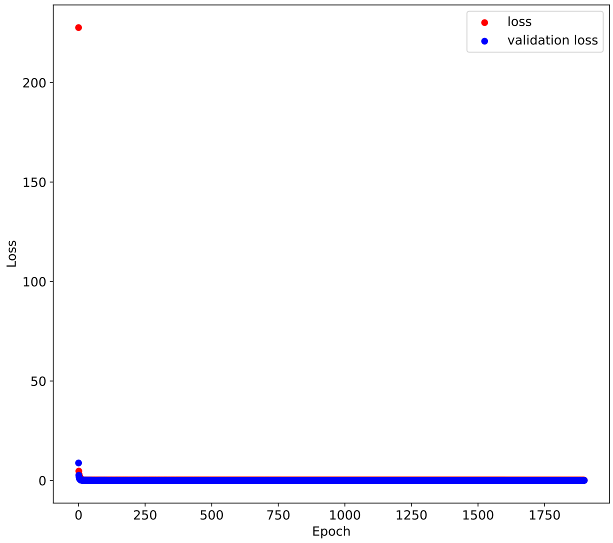
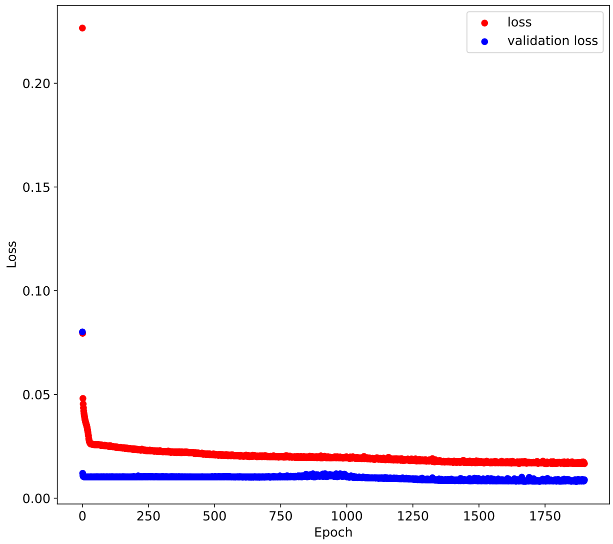


Figure 80: MO training history of zone 13 when considering 1 year of data and after dropping non-forecast data and data with no correlation >0.1 to at least one output



(a) Cooling_circuit_flow_temperature



(b) Cooling_circuit_mass_flow

Figure 81: SO training history of zone 13 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

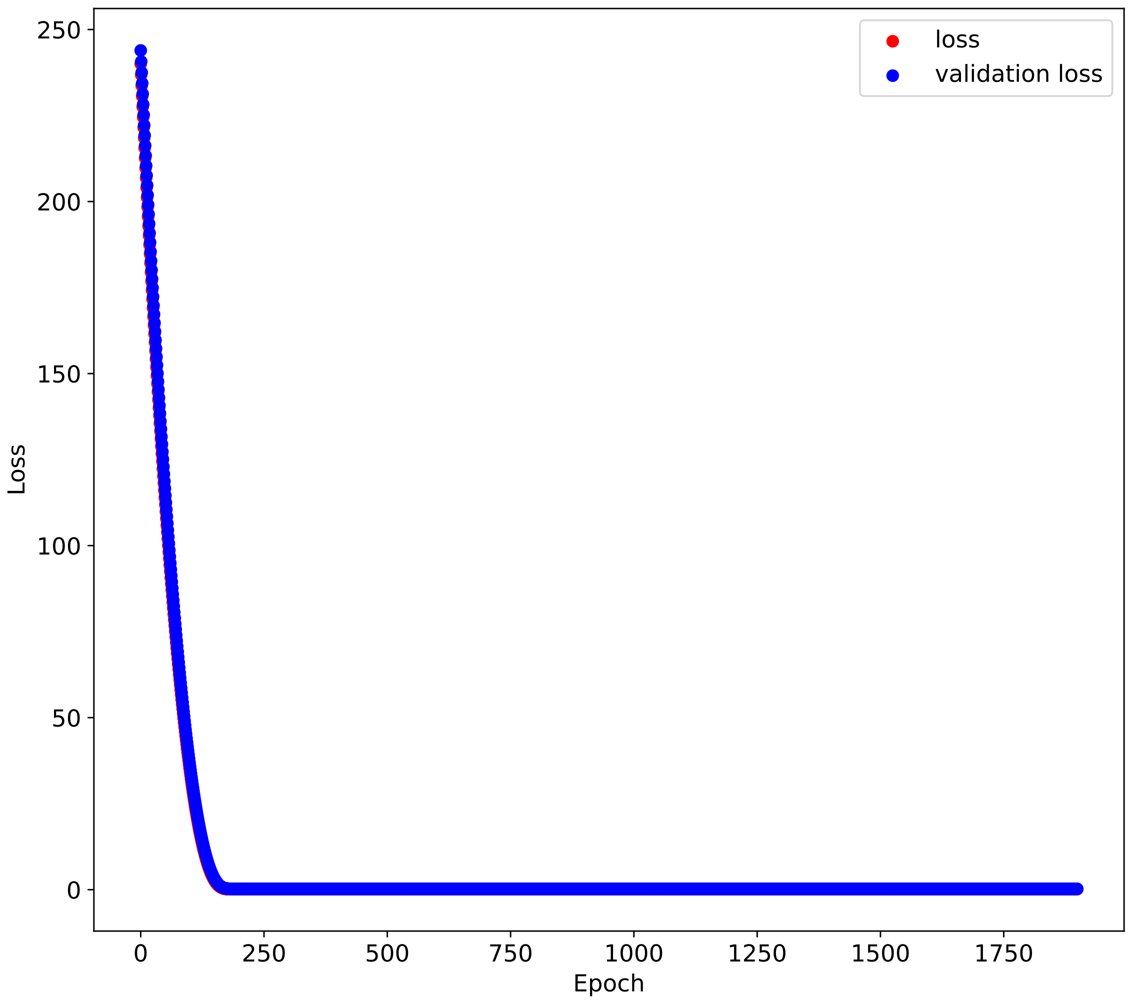
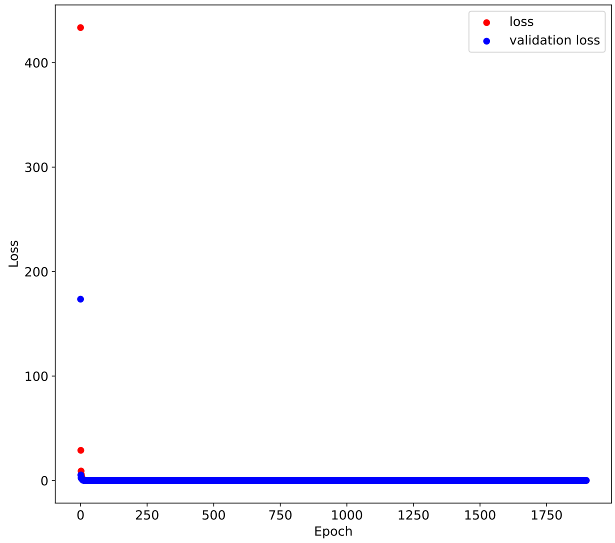
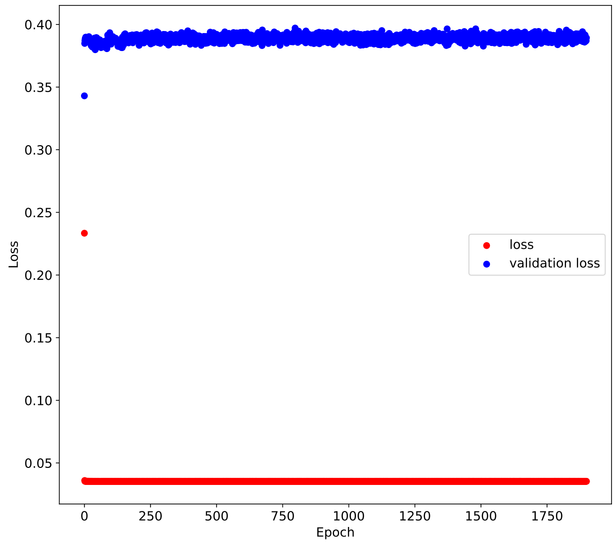


Figure 82: MO training history of zone 14 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Climate_sensor_RBG_basic_setpoint_room_temperature



(b) Climate_sensor_presence

Figure 83: SO training history of zone 14 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

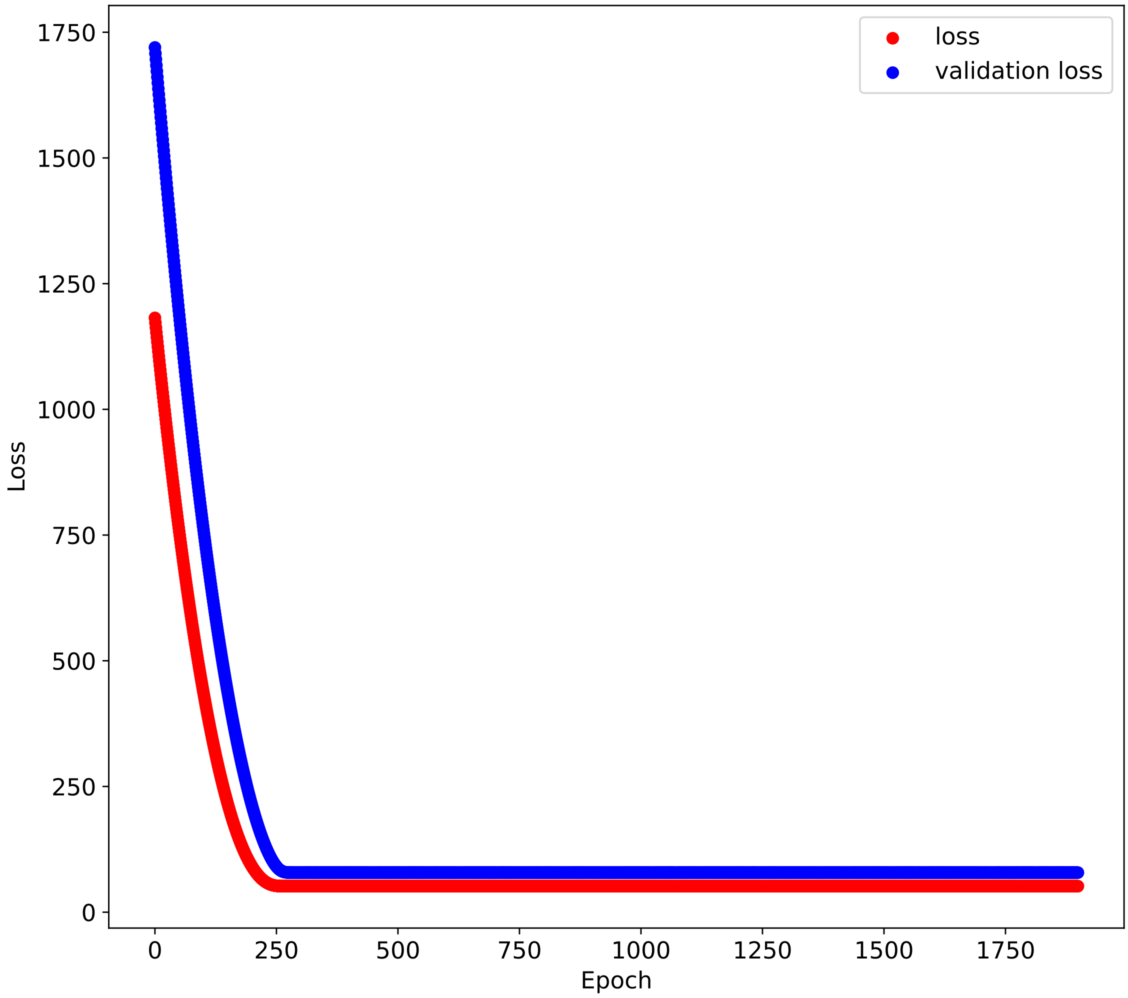


Figure 84: Training history of zone 15 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output (only one output)

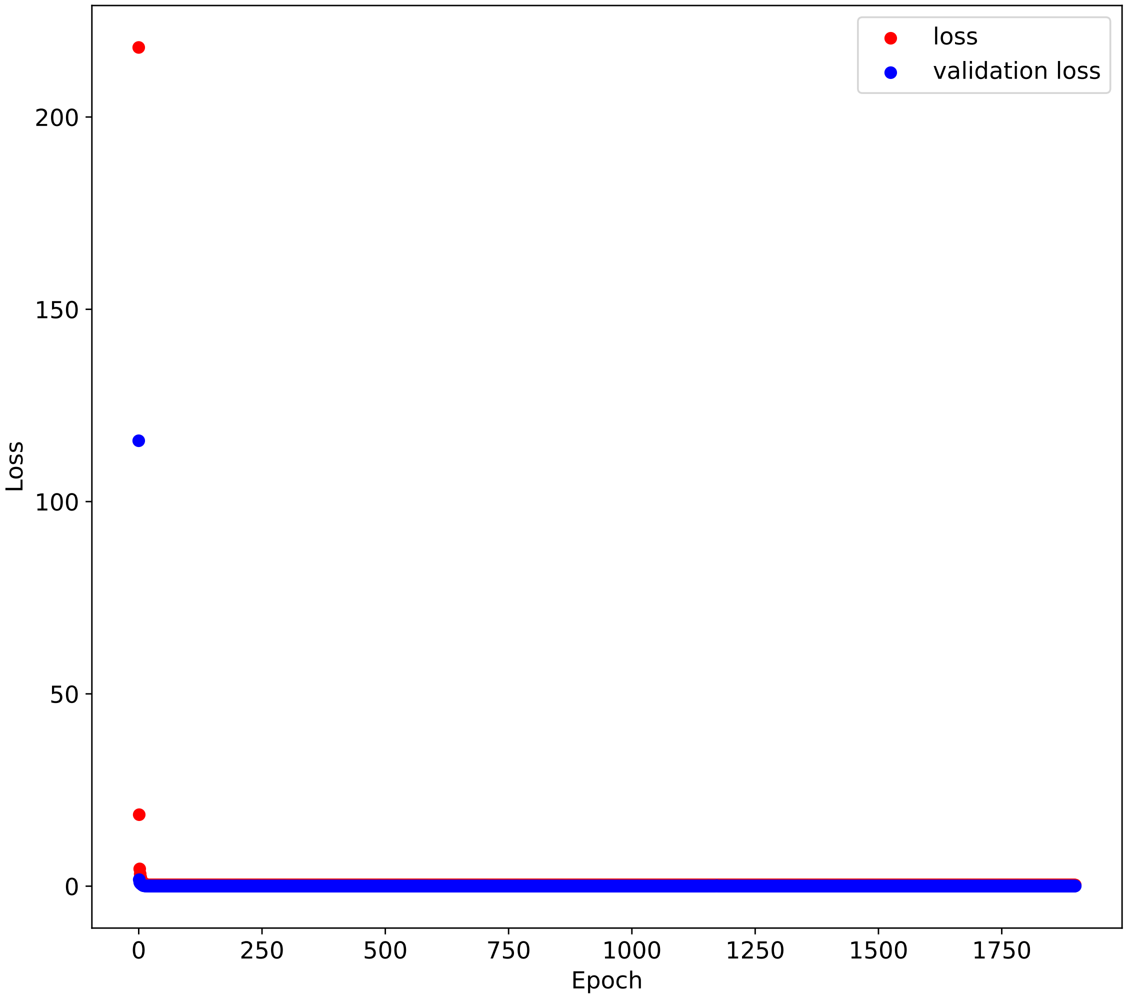
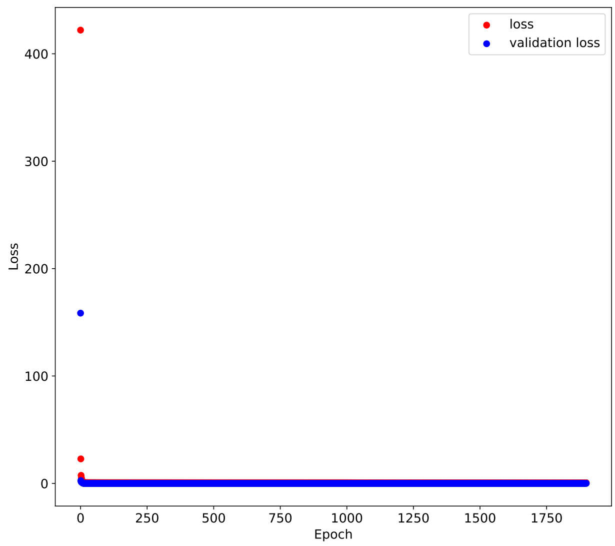
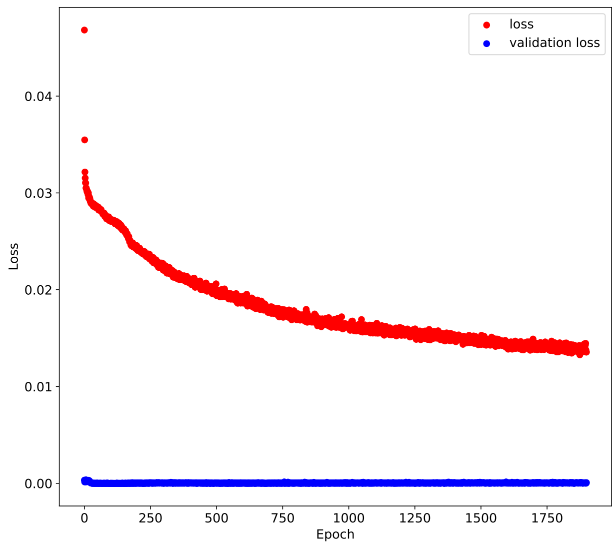


Figure 85: MO training history of zone 16 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Cooling_circuit_component_activation_flow_temperature



(b) Cooling_circuit_activation_mass_flow

Figure 86: SO training history of zone 16 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

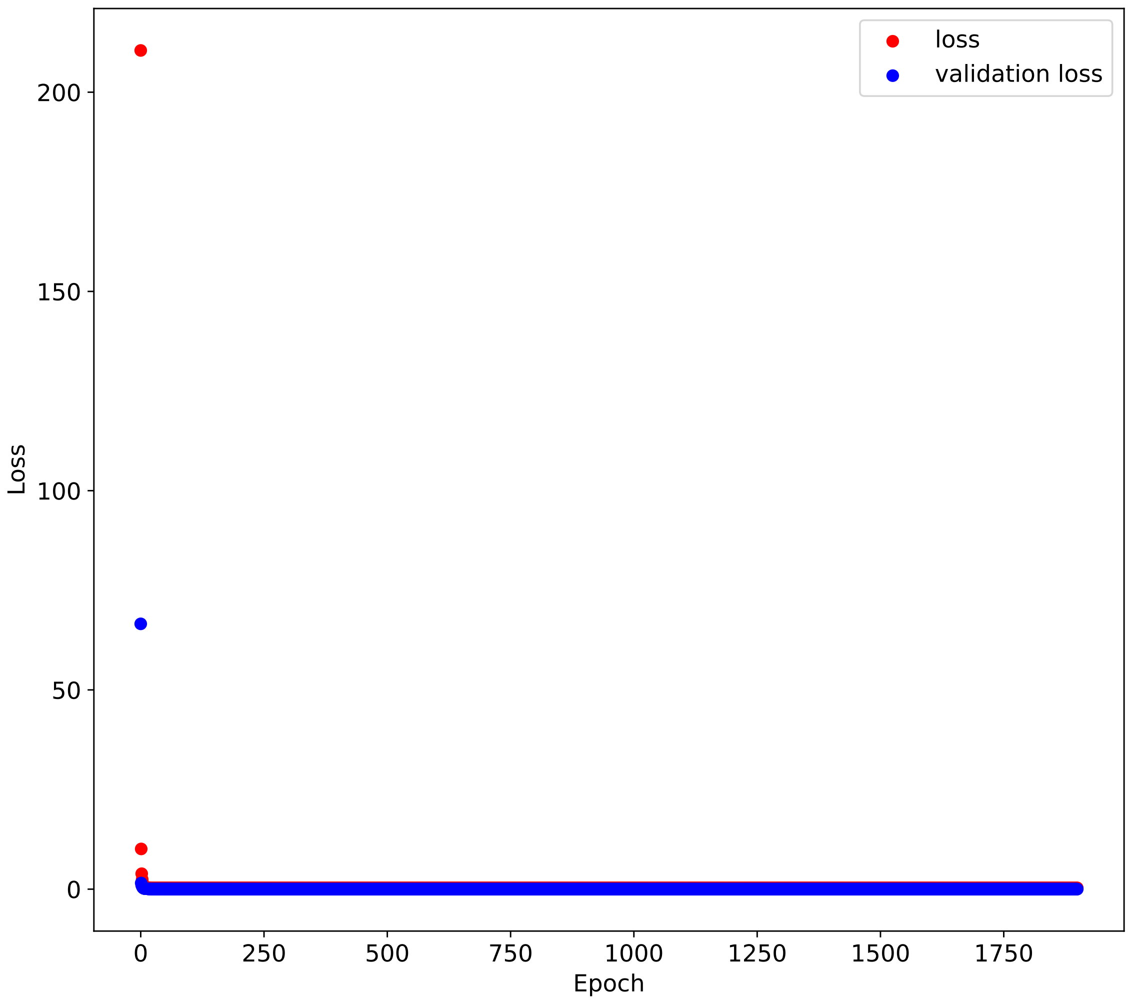
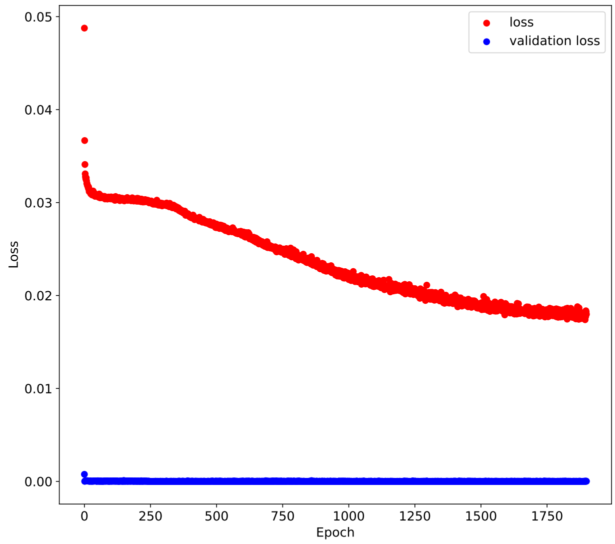
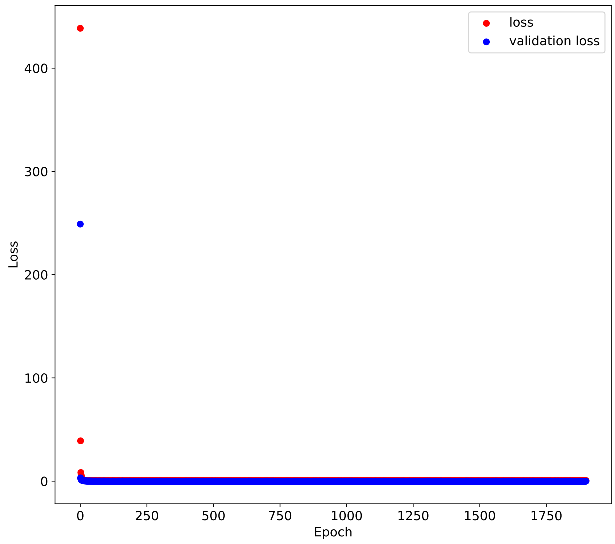


Figure 87: MO training history of zone 17 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Cooling_circuit_activation_mass_flow



(b) Cooling_circuit_component_activation_flow_temperature

Figure 88: SO training history of zone 17 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

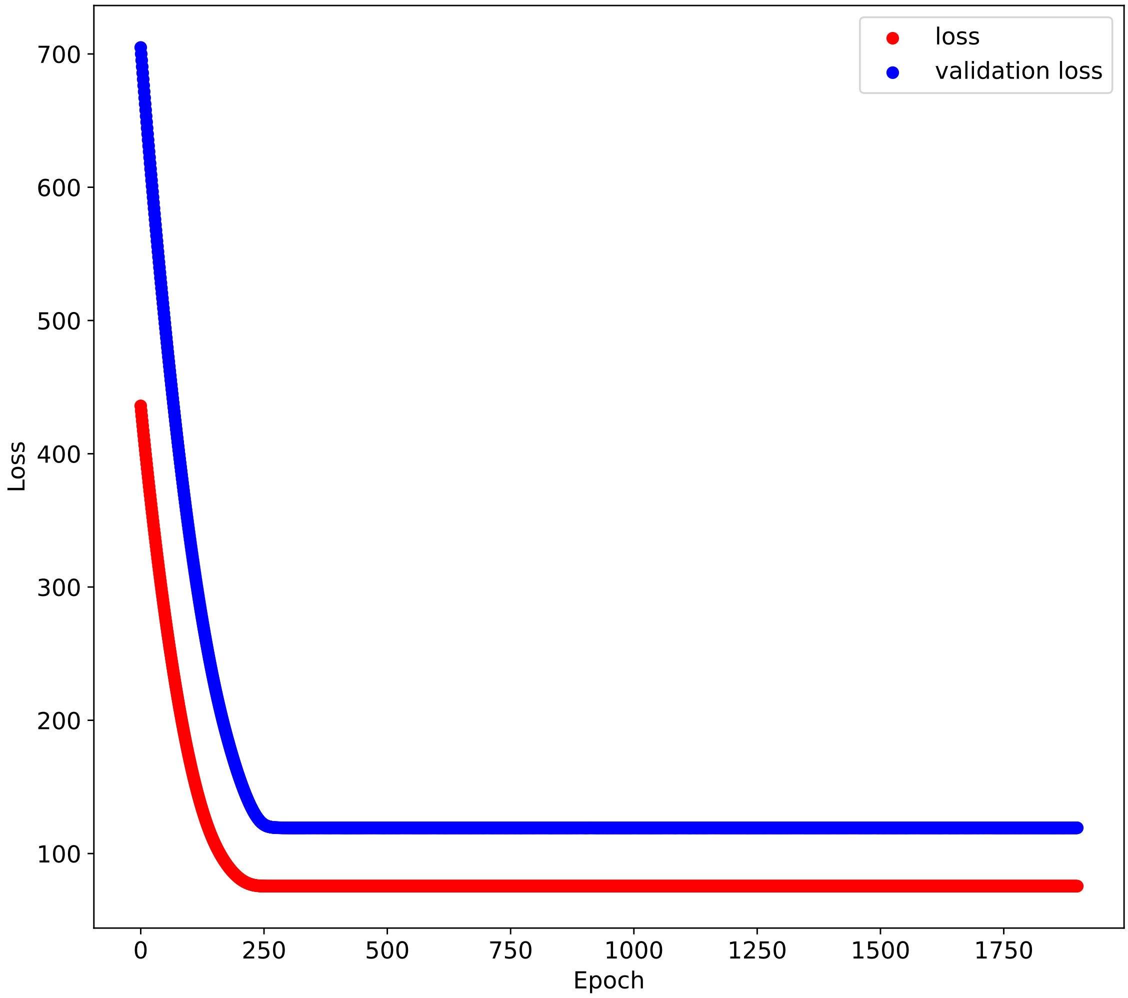
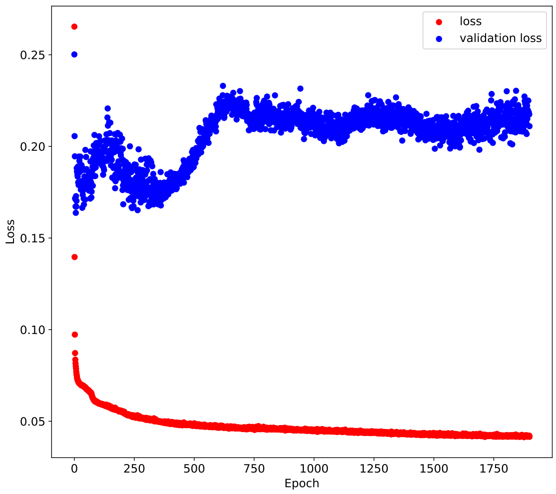
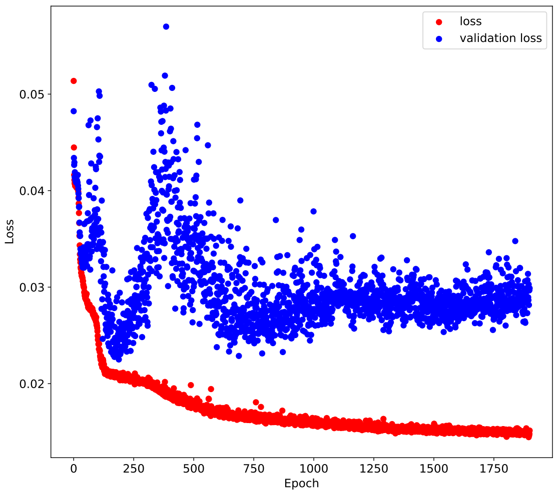


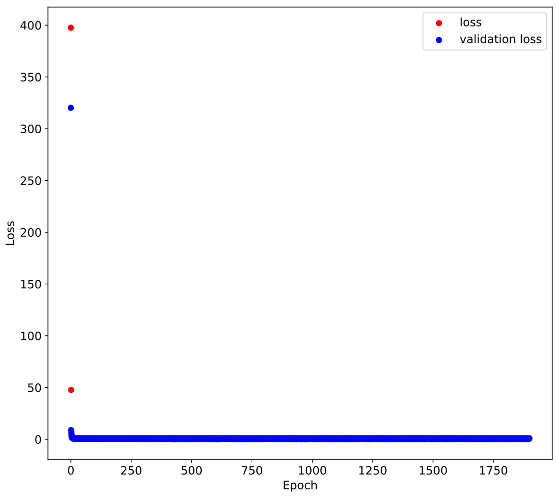
Figure 89: MO training history of zone 18 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



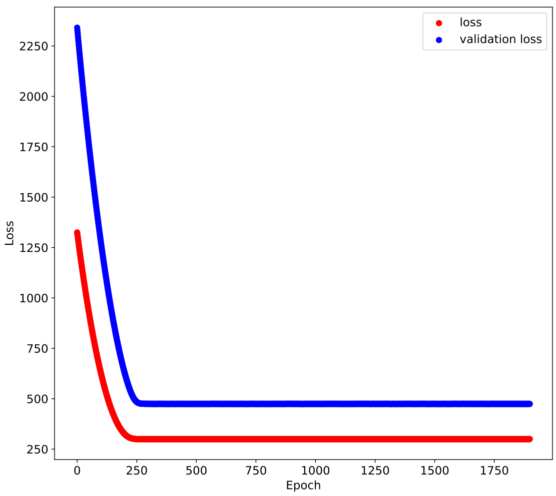
(a) Heating_circuit_heater_mass_flow



(b) Ventilation_system_volume_flow



(c) Ventilation_system_supply_air_temperature



(d) Heating_circuit_heater_flow_temperature

Figure 90: SO training history of zone 18 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

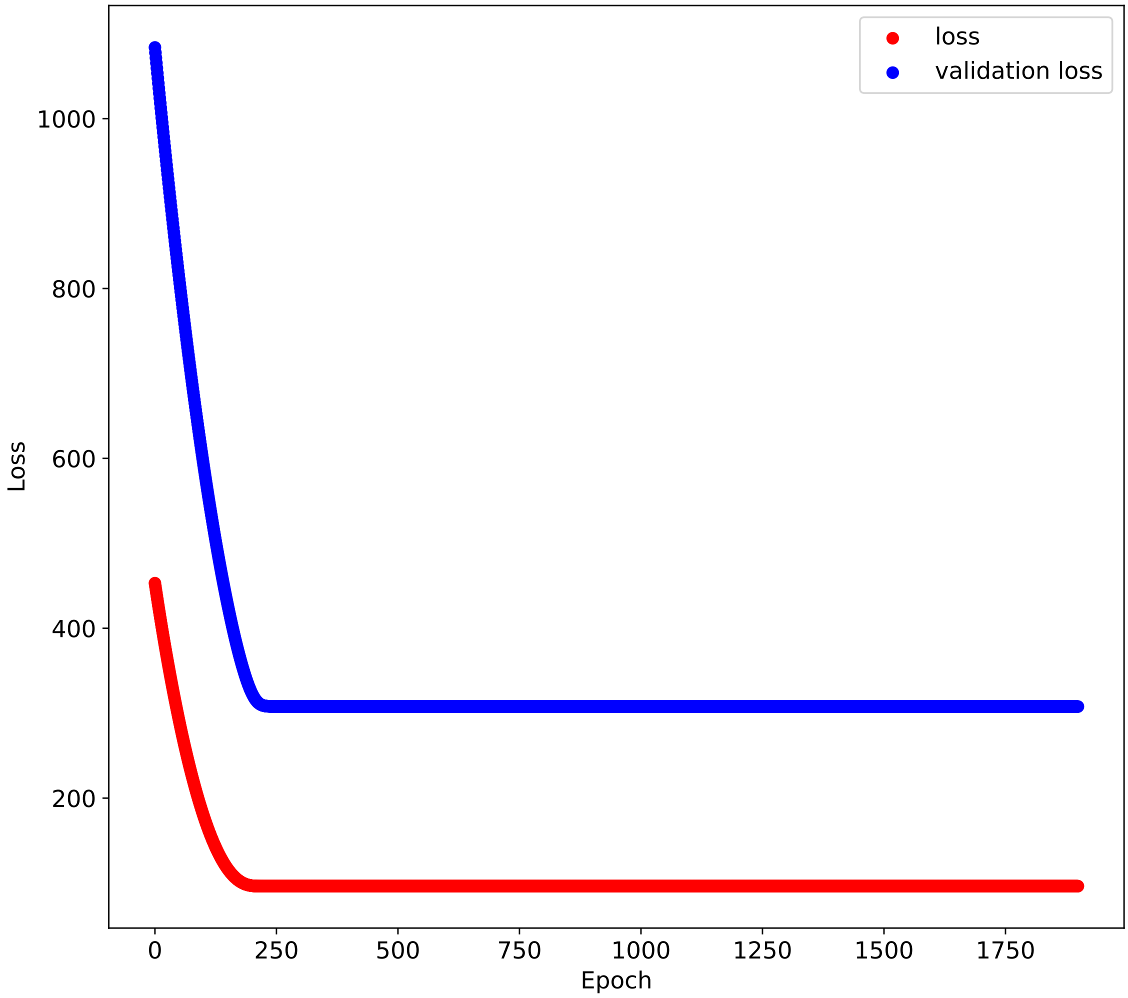
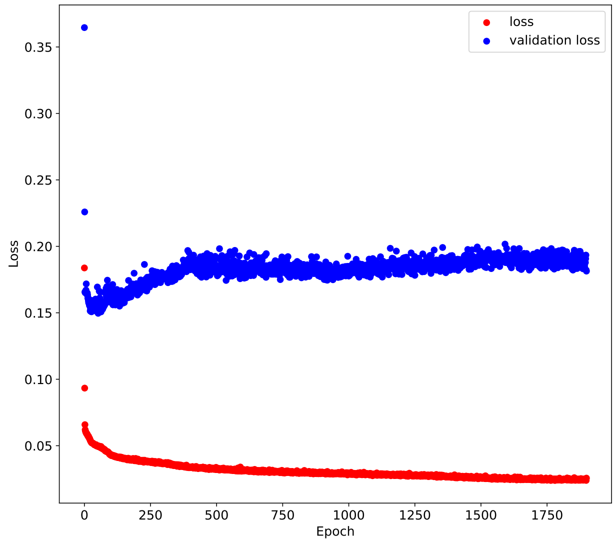
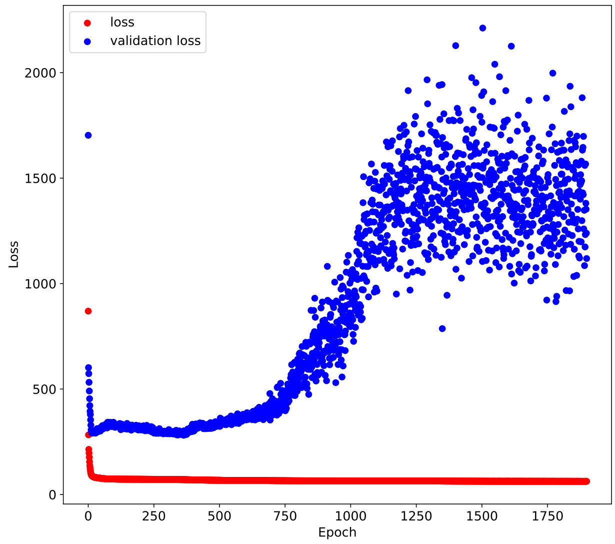


Figure 91: MO training history of zone 20 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Heating_circuit_heater_mass_flow



(b) Heating_circuit_heater_flow_temperature

Figure 92: SO training history of zone 20 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

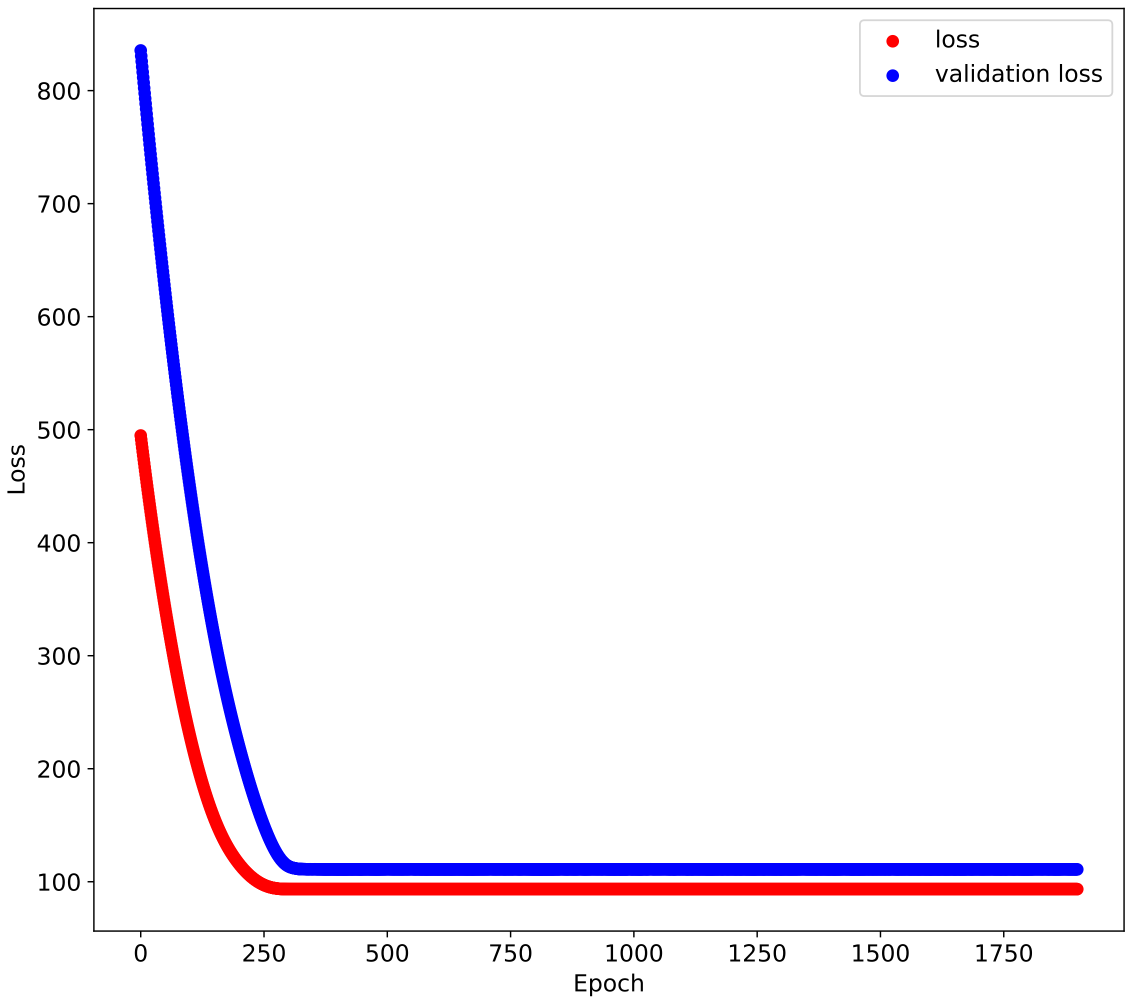
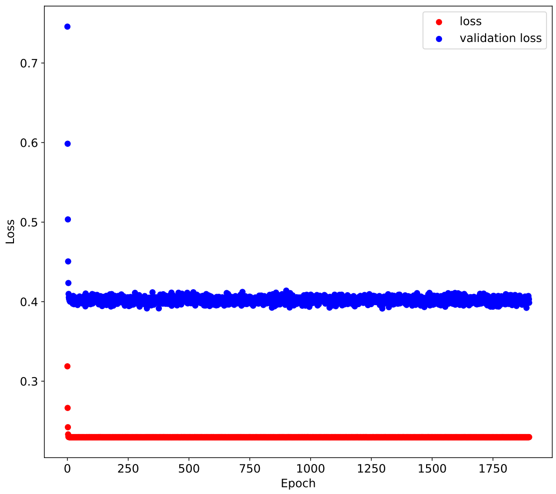
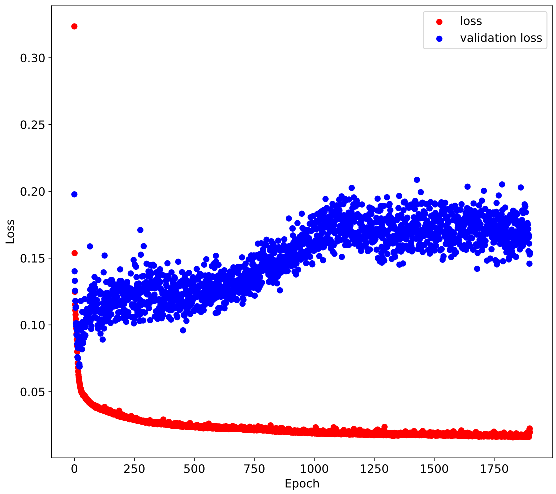


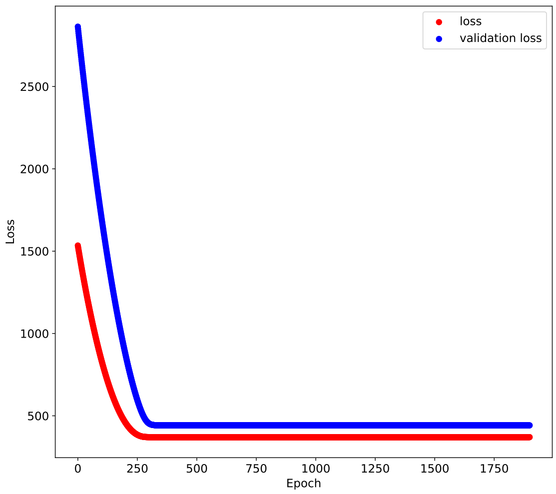
Figure 93: MO training history of zone 21 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



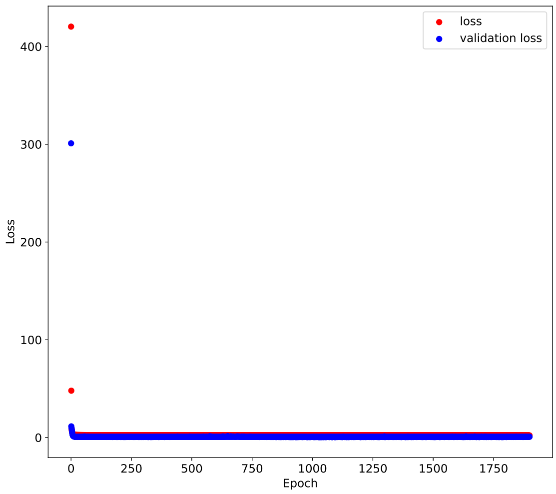
(a) Heating_circuit_mass_flow



(b) Ventilation_system_RLT_volume_flow_release



(c) Heating_circuit_flow_temperature



(d) Ventilation_system_RLT_supply_air_temperature

Figure 94: SO training history of zone 21 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

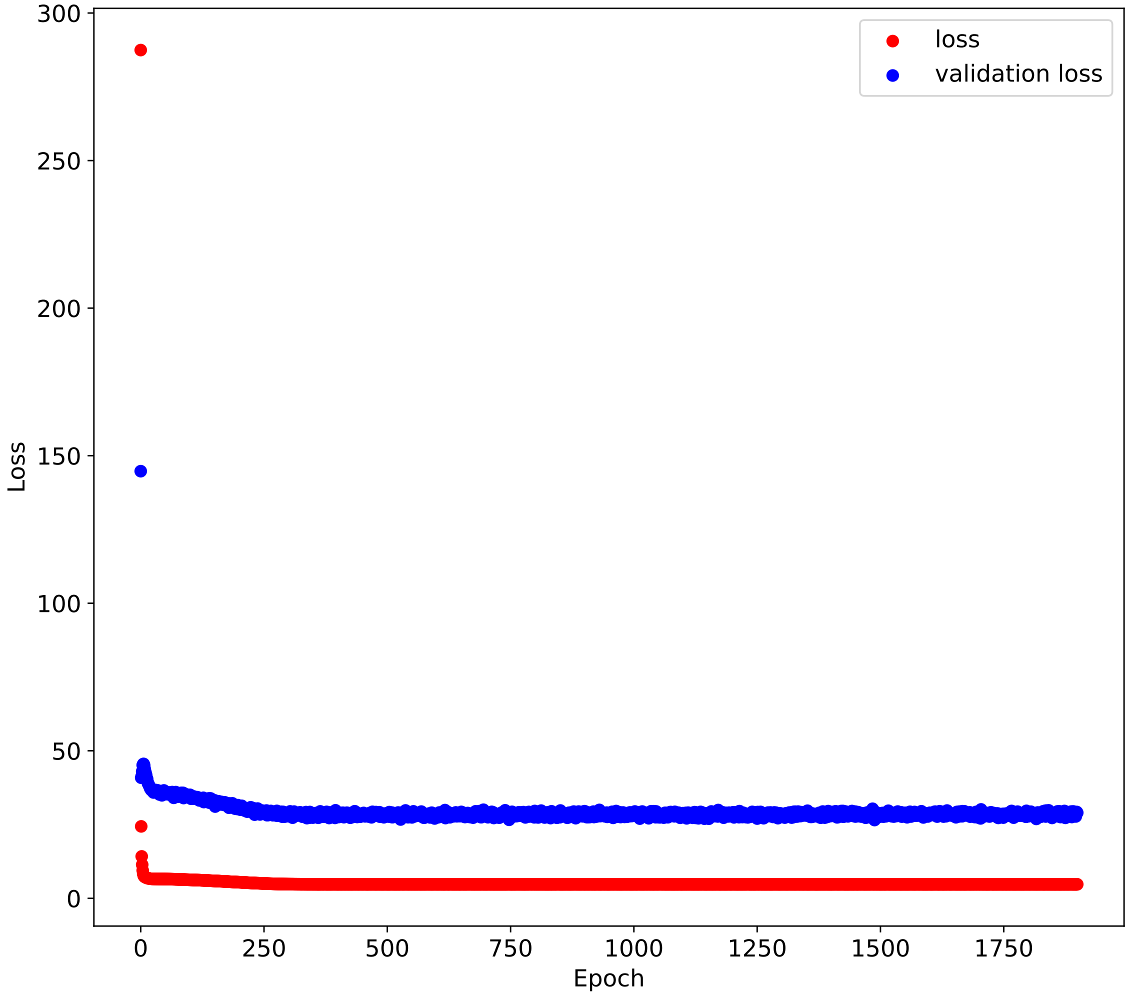
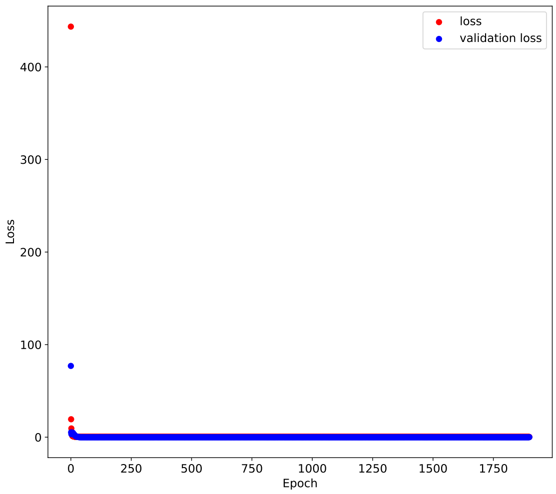
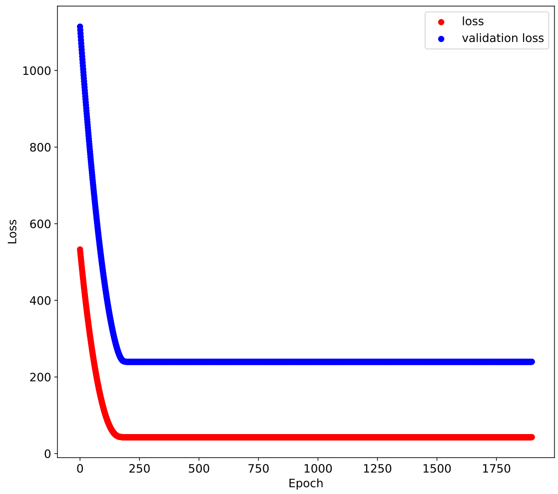


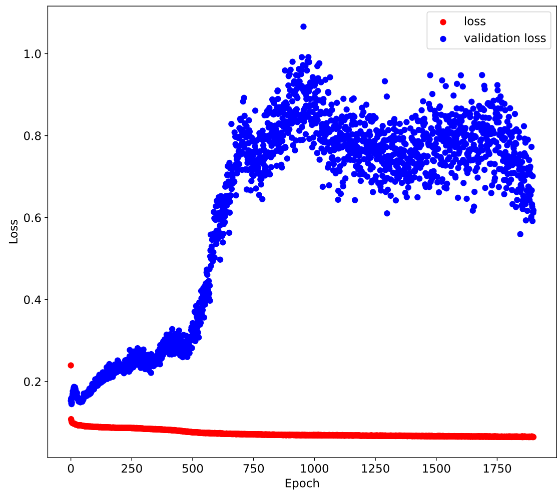
Figure 95: MO training history of zone 22 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Cooling_circuit_ceiling_flow_temperature

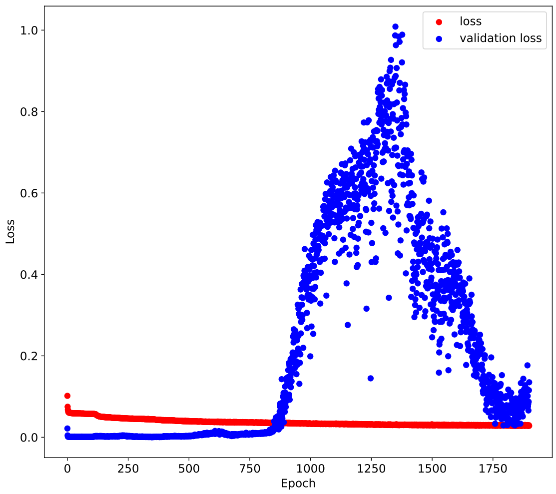


(b) Heating_circuit_wall_flow_temperature

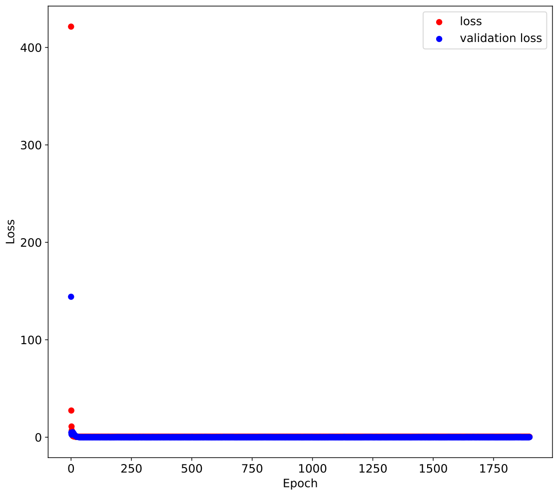


(c) Heating_circuit_wall_mass_flow

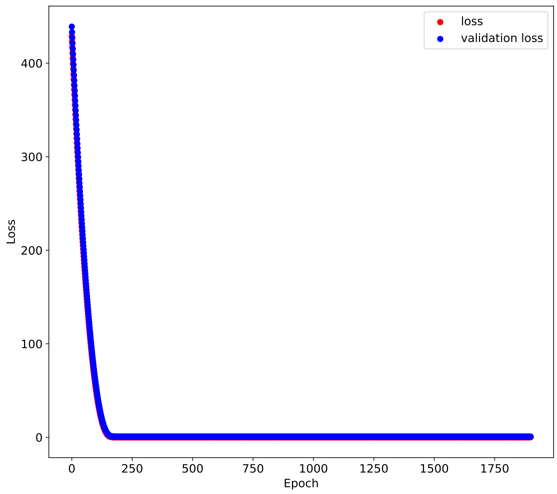
Figure 96: SO training history of zone 22 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Cooling_circuit_wall_mass_flow



(b) Ventilation_system_supply_air_temperature



(c) Ventilation_system_supply_air_temperature.1

Figure 97: SO training history of zone 22 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output

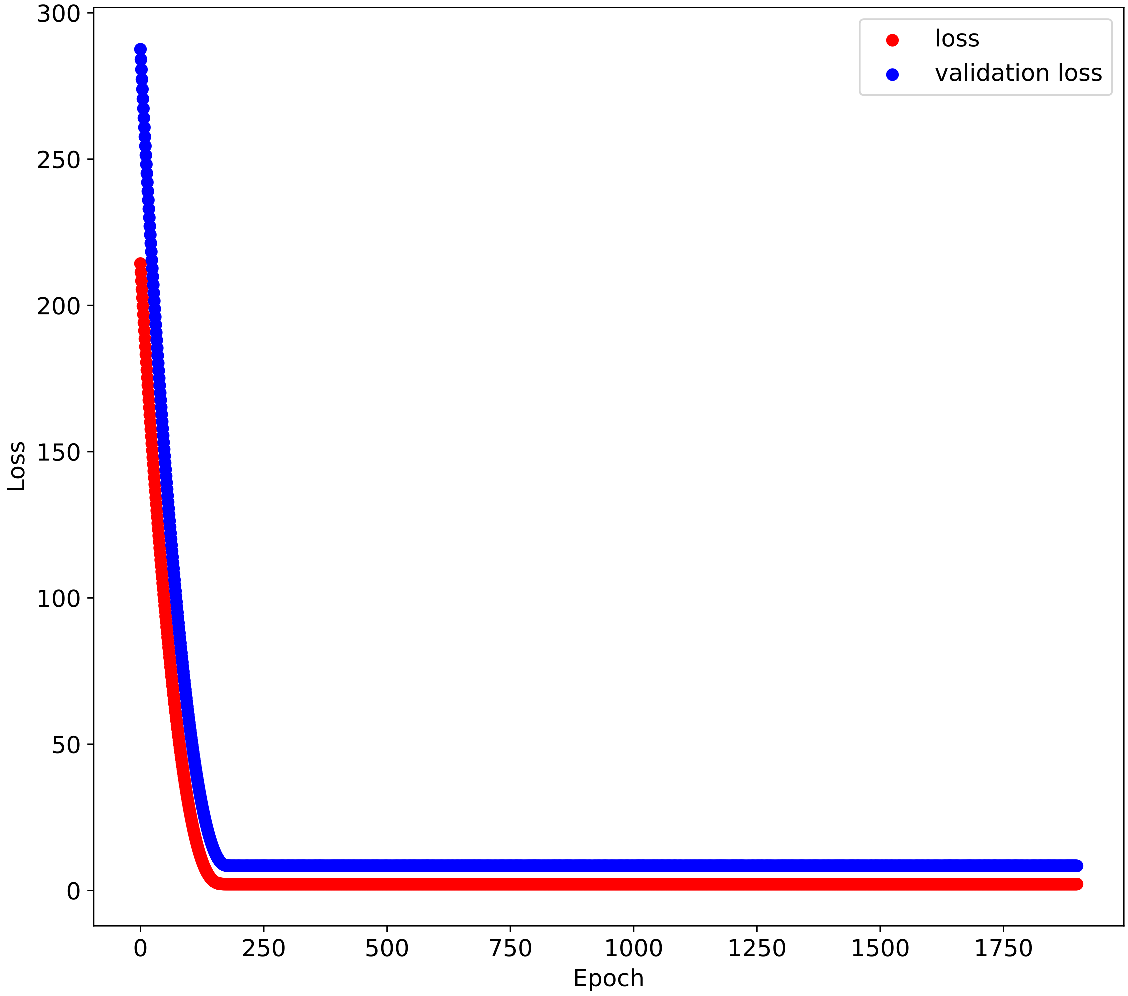
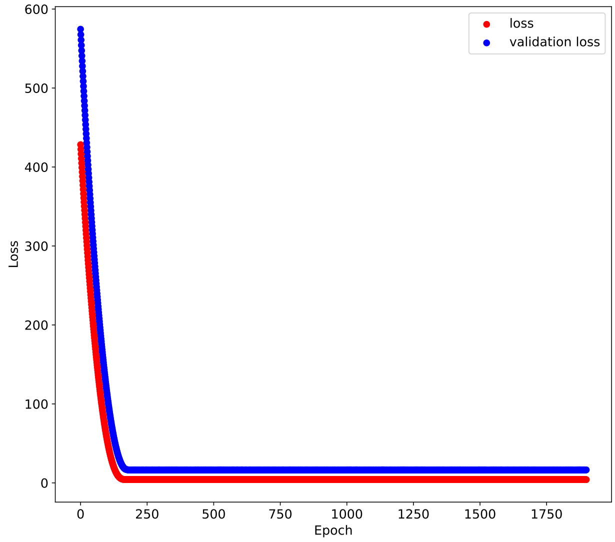
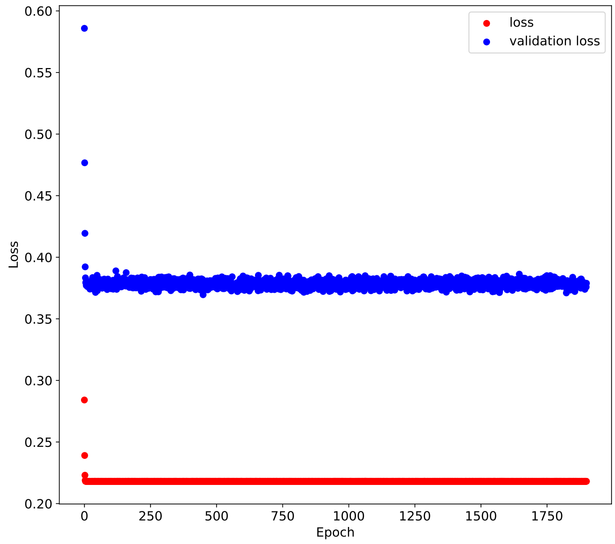


Figure 98: MO training history of zone 23 when considering 1 year of data and after dropping non-forecast data and data with no correlation > 0.1 to at least one output



(a) Component_activation_flow_temperature



(b) Component_activation_mass_flow

Figure 99: SO training history of zone 23 when considering 1 year of data and after dropping non-forecast data and data with no correlation $>$ 0.1 to at least one output

F. Heatmaps

Figure 100: Colorbar for all heatmaps

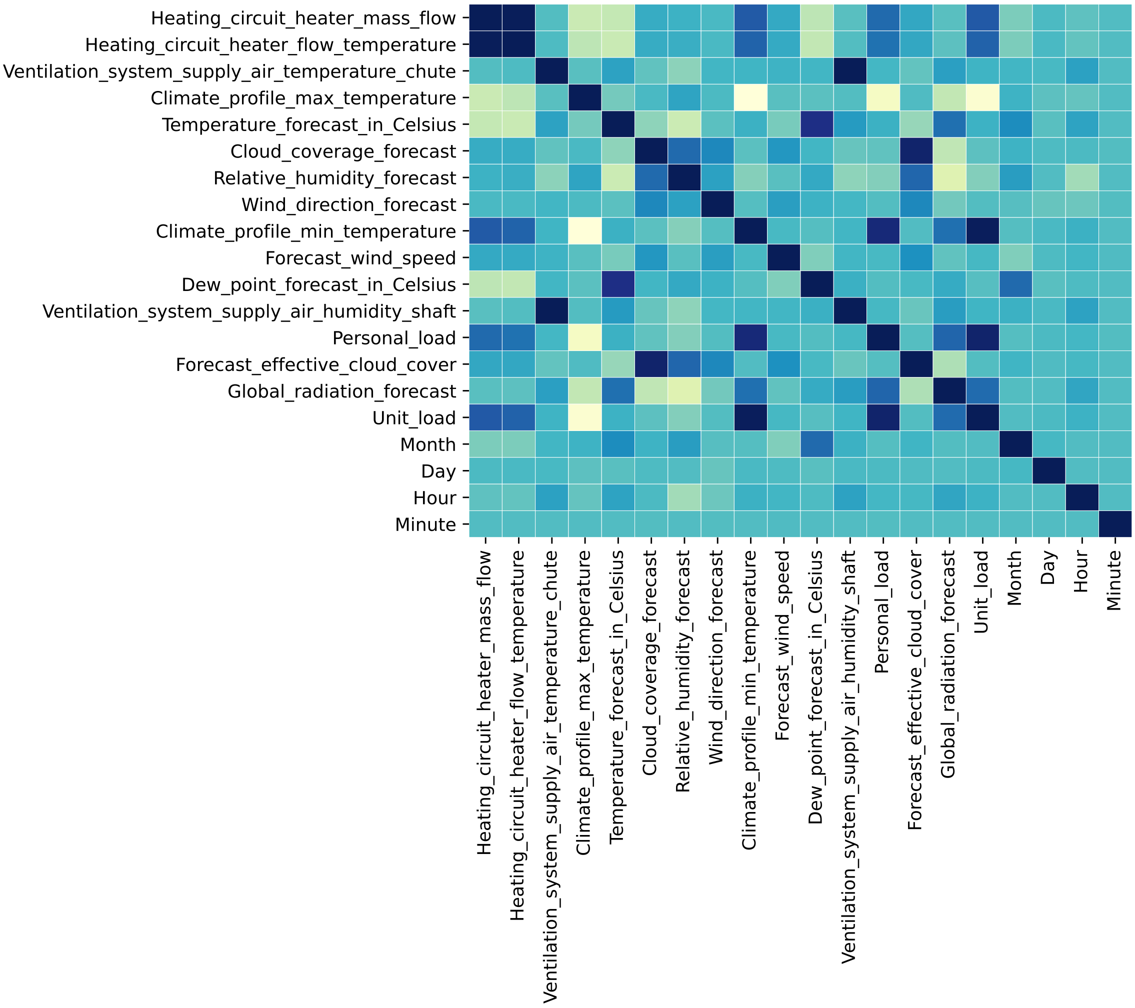


Figure 101: Visualization of Correlations between data for zone 0

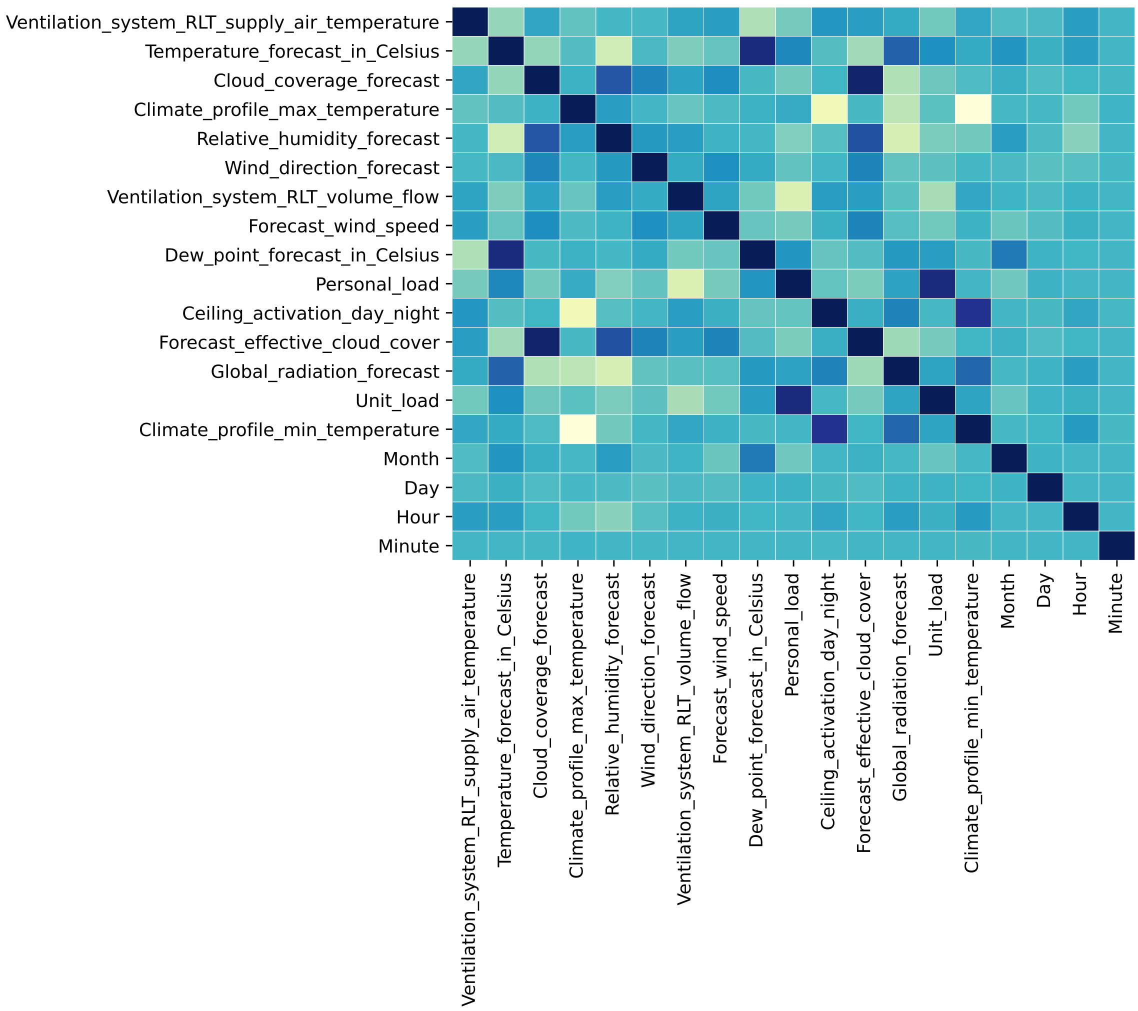


Figure 102: Visualization of Correlations between data for zone 1

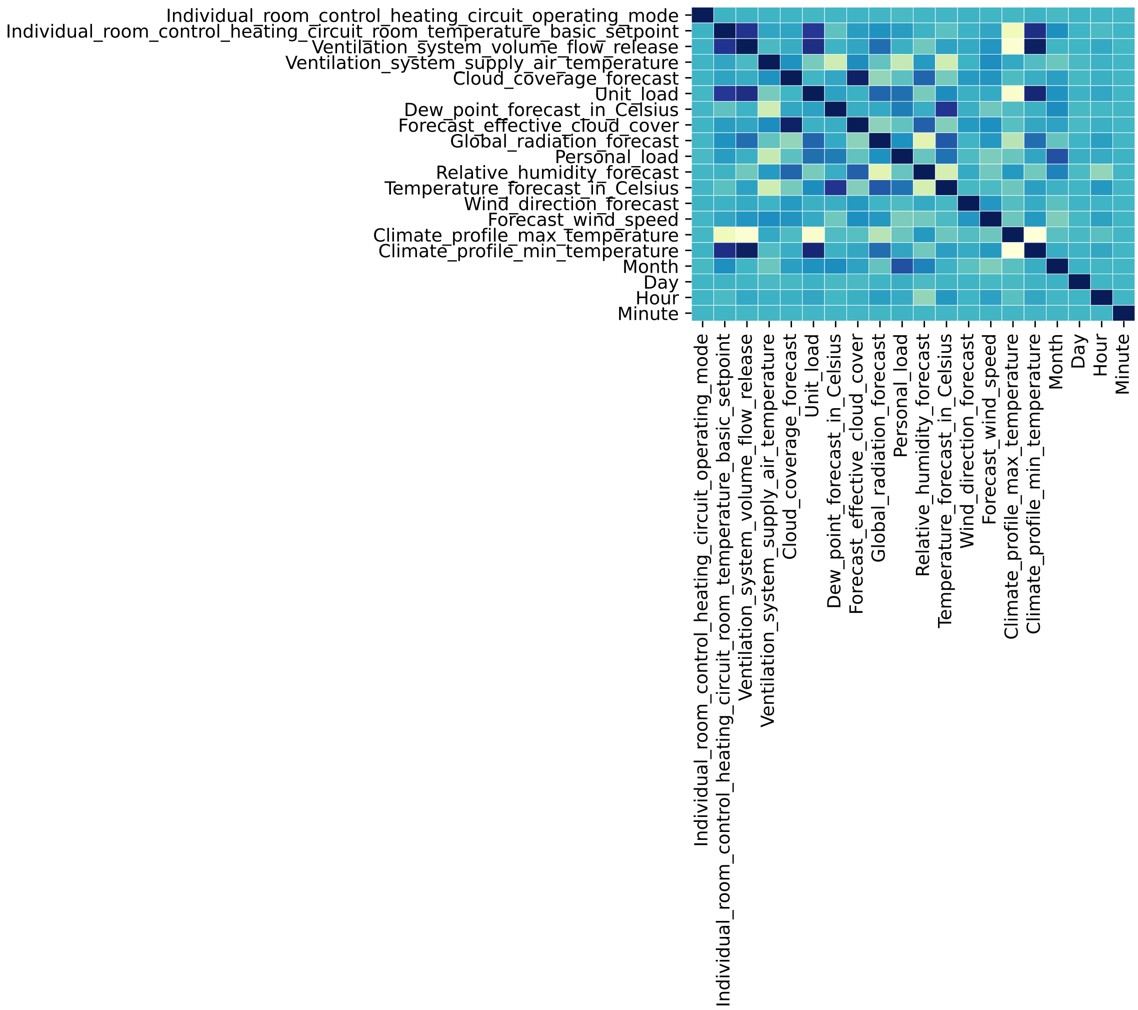


Figure 103: Visualization of Correlations between data for zone 2

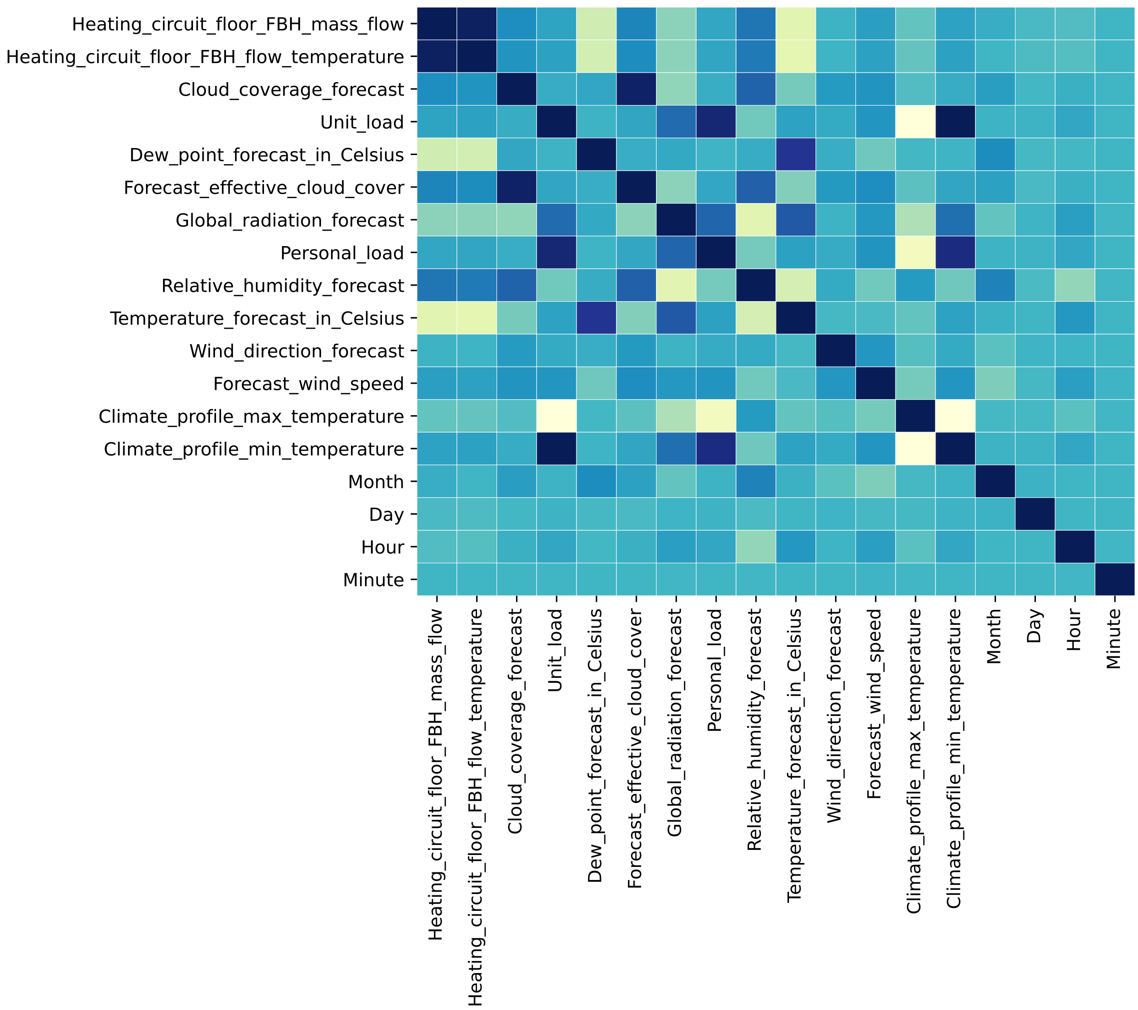


Figure 104: Visualization of Correlations between data for zone 3

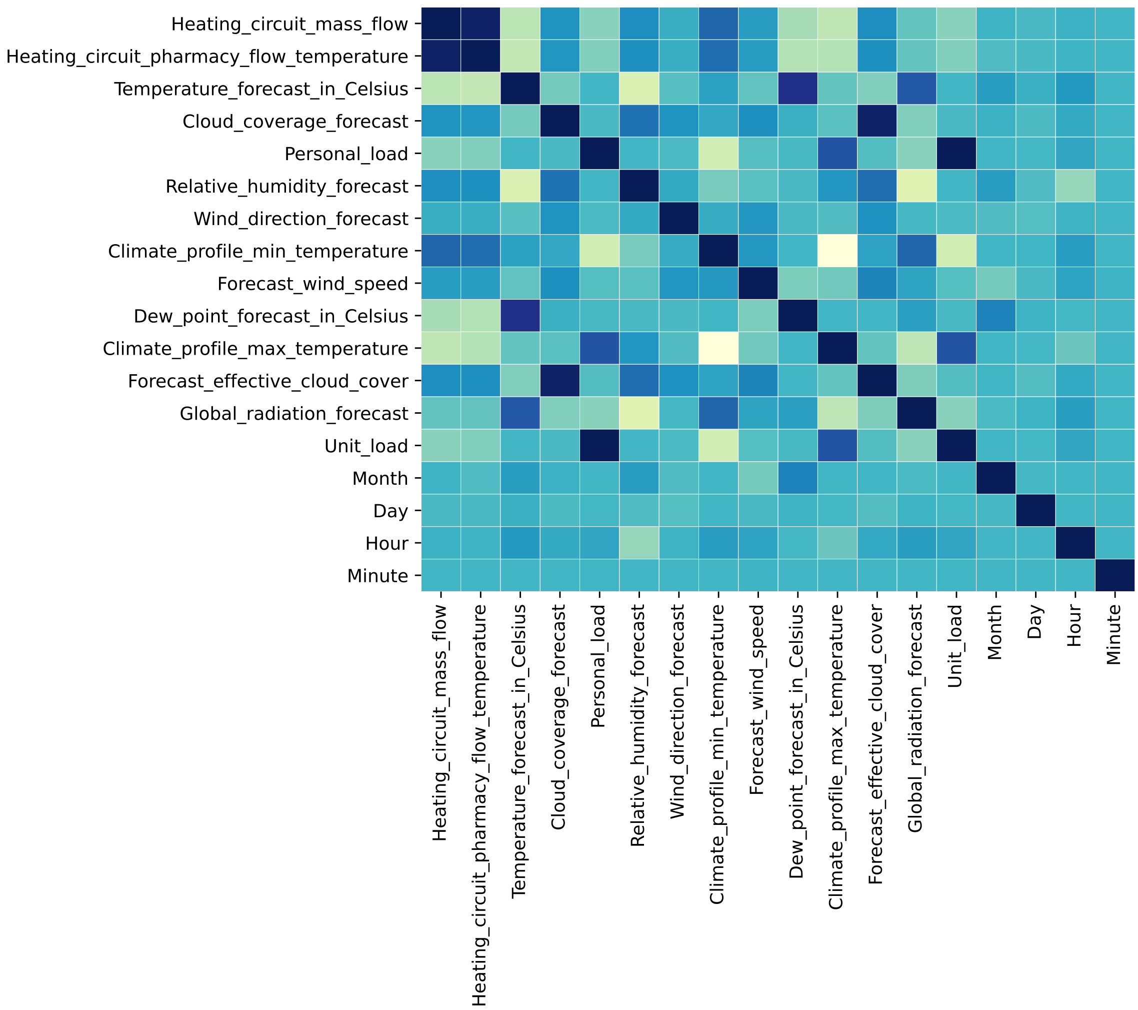


Figure 105: Visualization of Correlations between data for zone 4

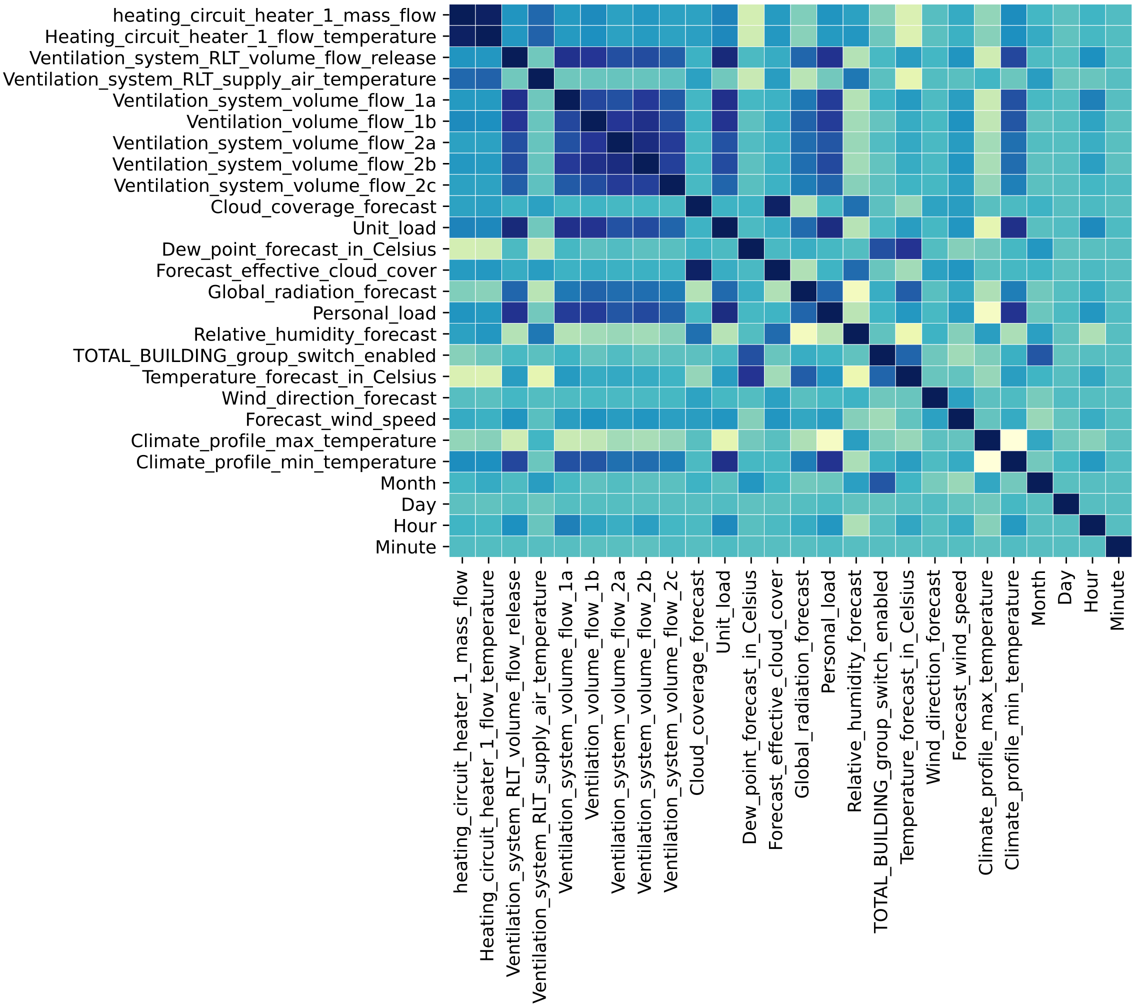


Figure 106: Visualization of Correlations between data for zone 5

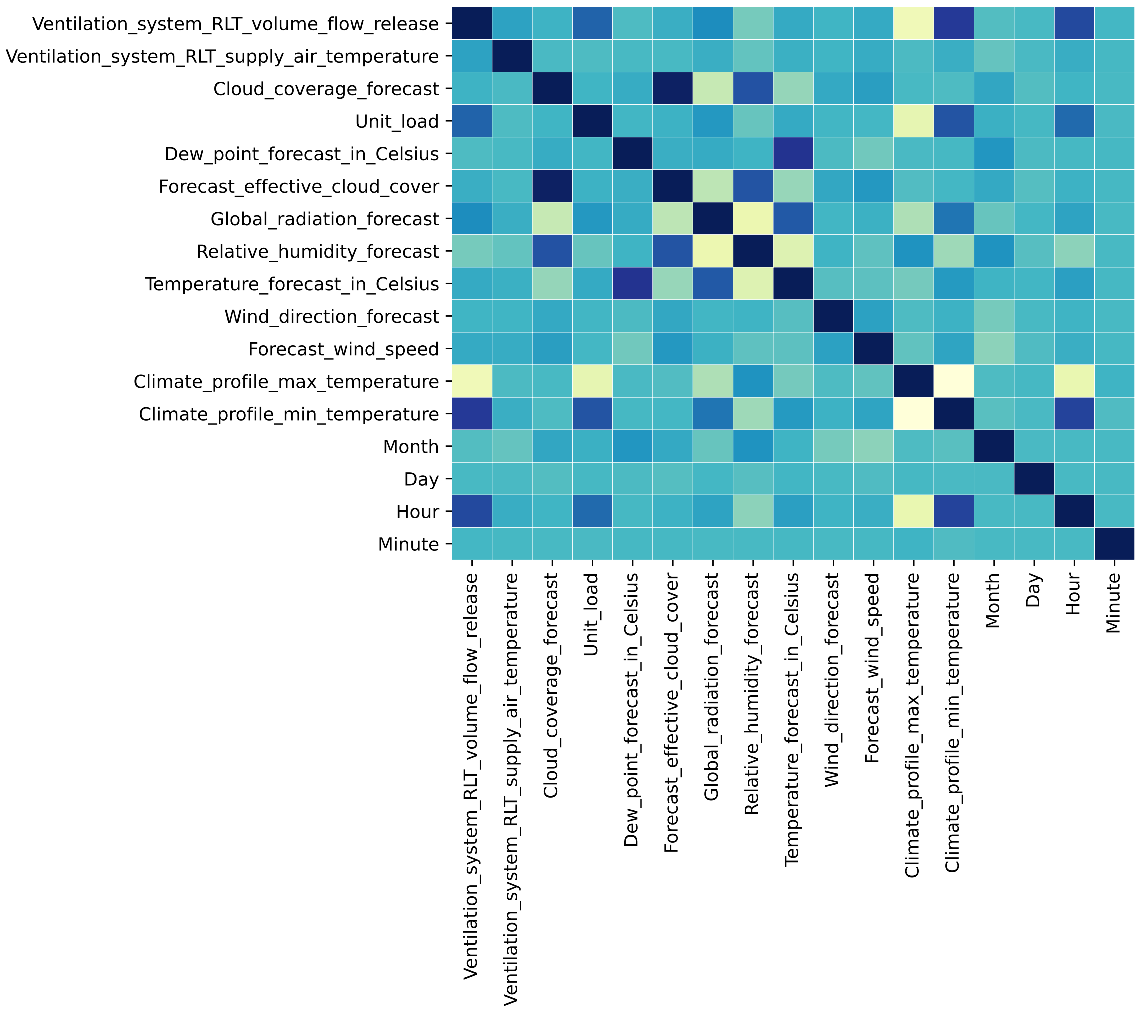


Figure 107: Visualization of Correlations between data for zone 7

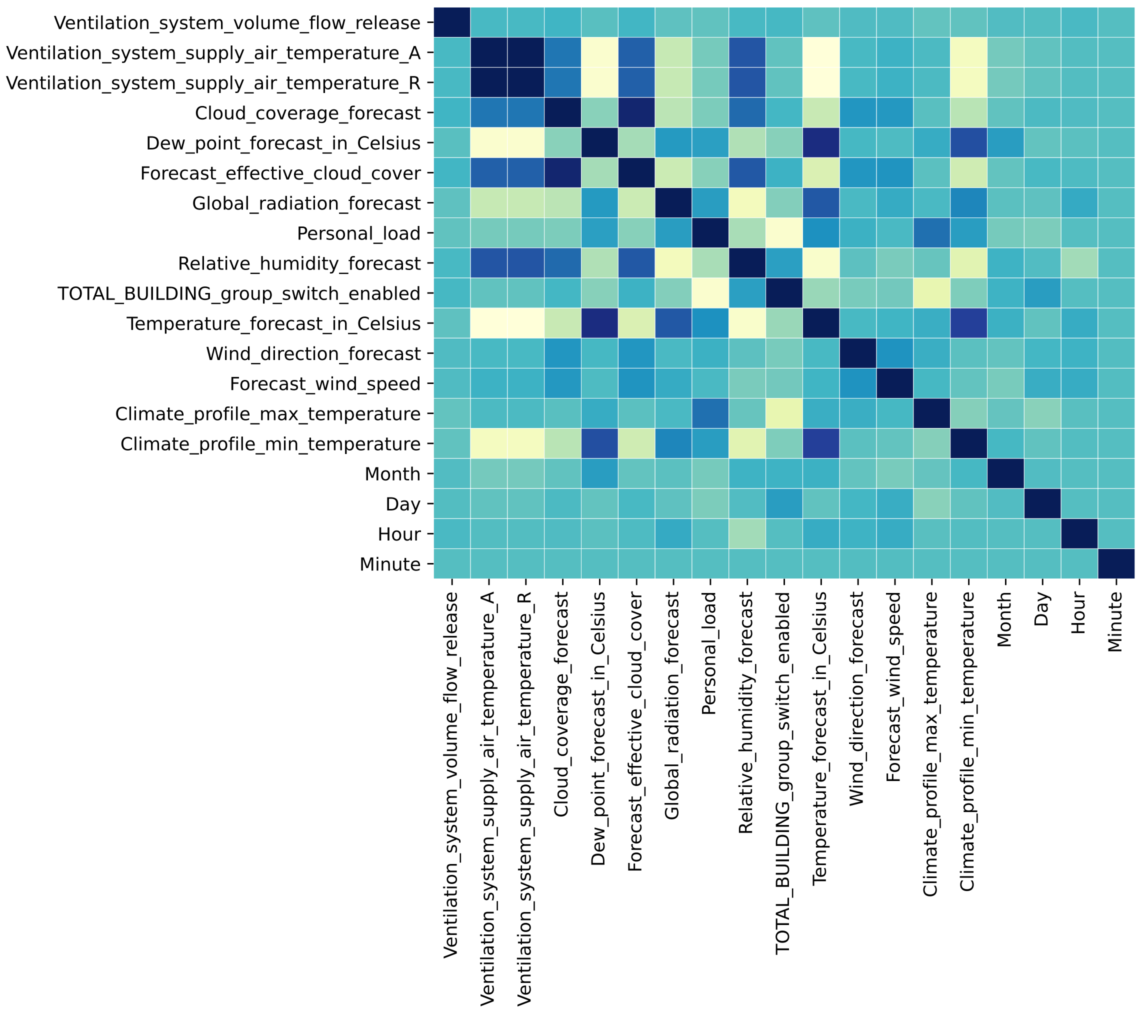


Figure 108: Visualization of Correlations between data for zone 8

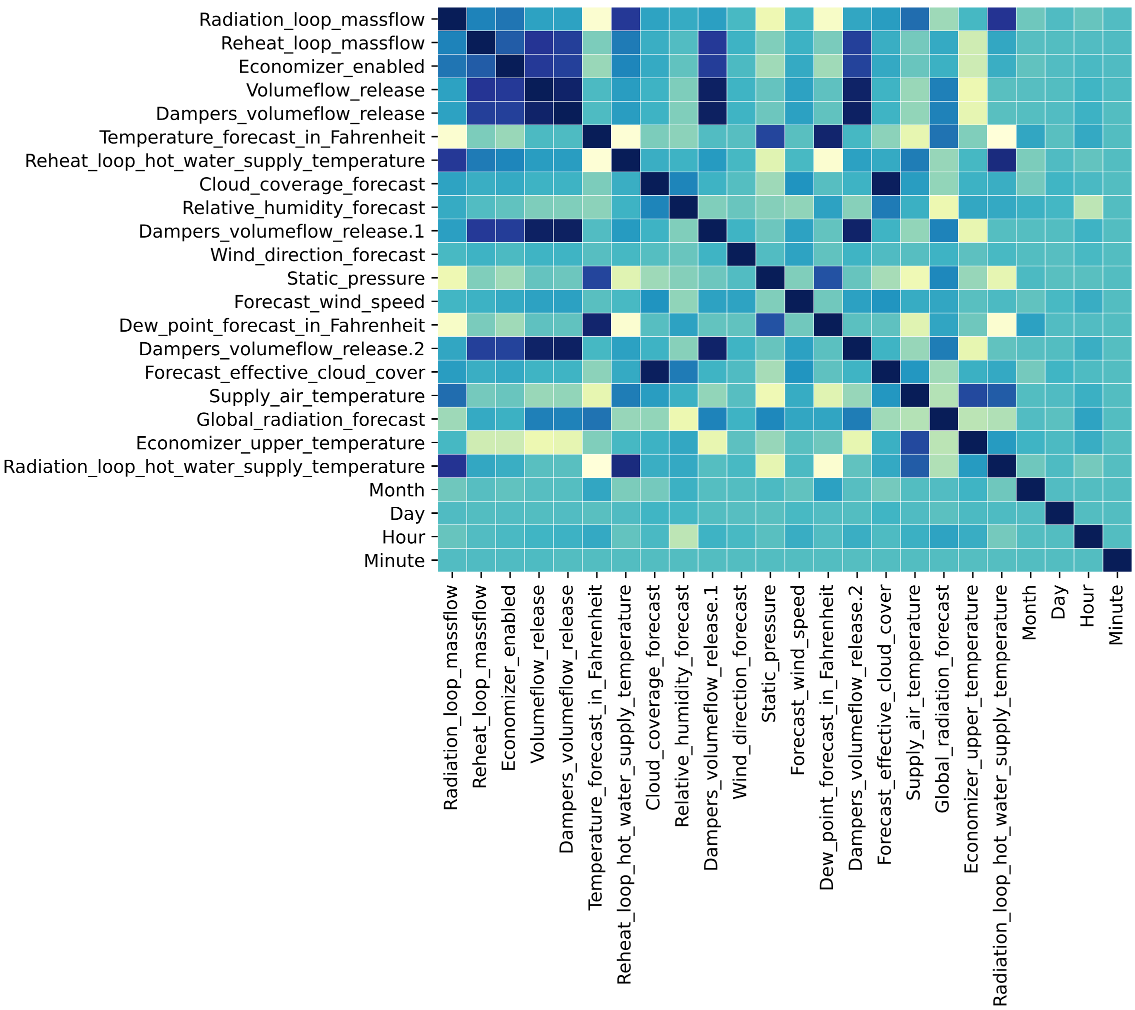


Figure 109: Visualization of Correlations between data for zone 9

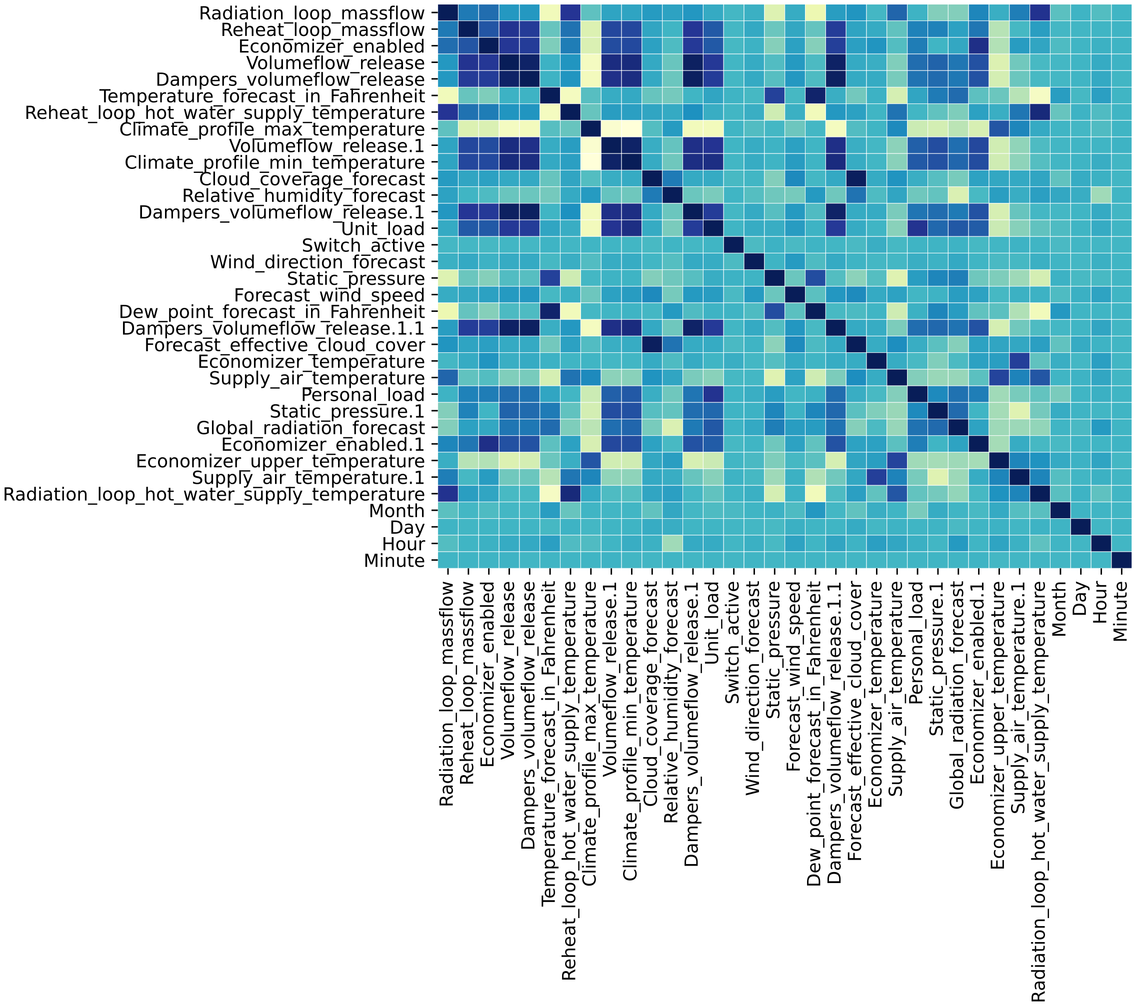


Figure 110: Visualization of Correlations between data for zone 10

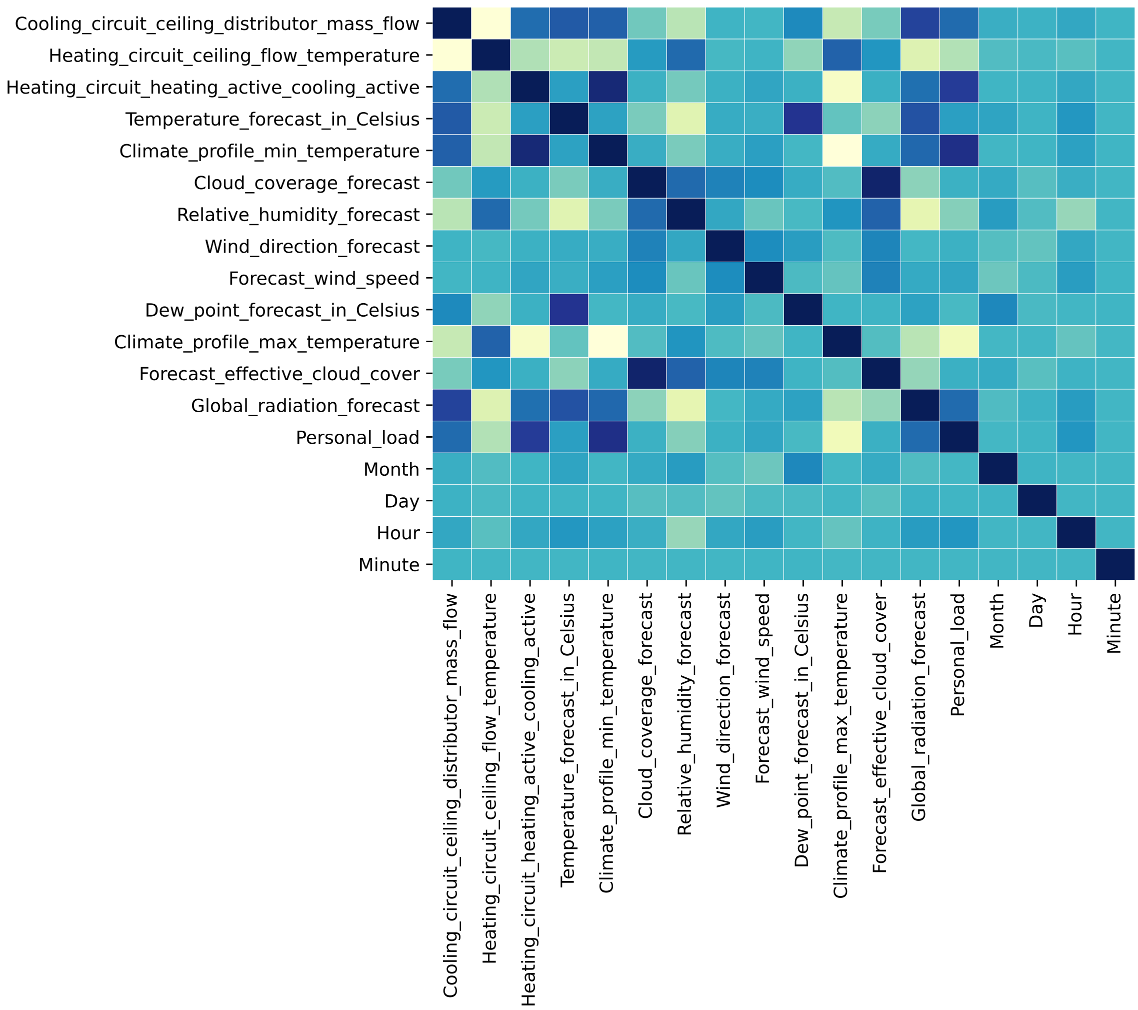


Figure 111: Visualization of Correlations between data for zone 11

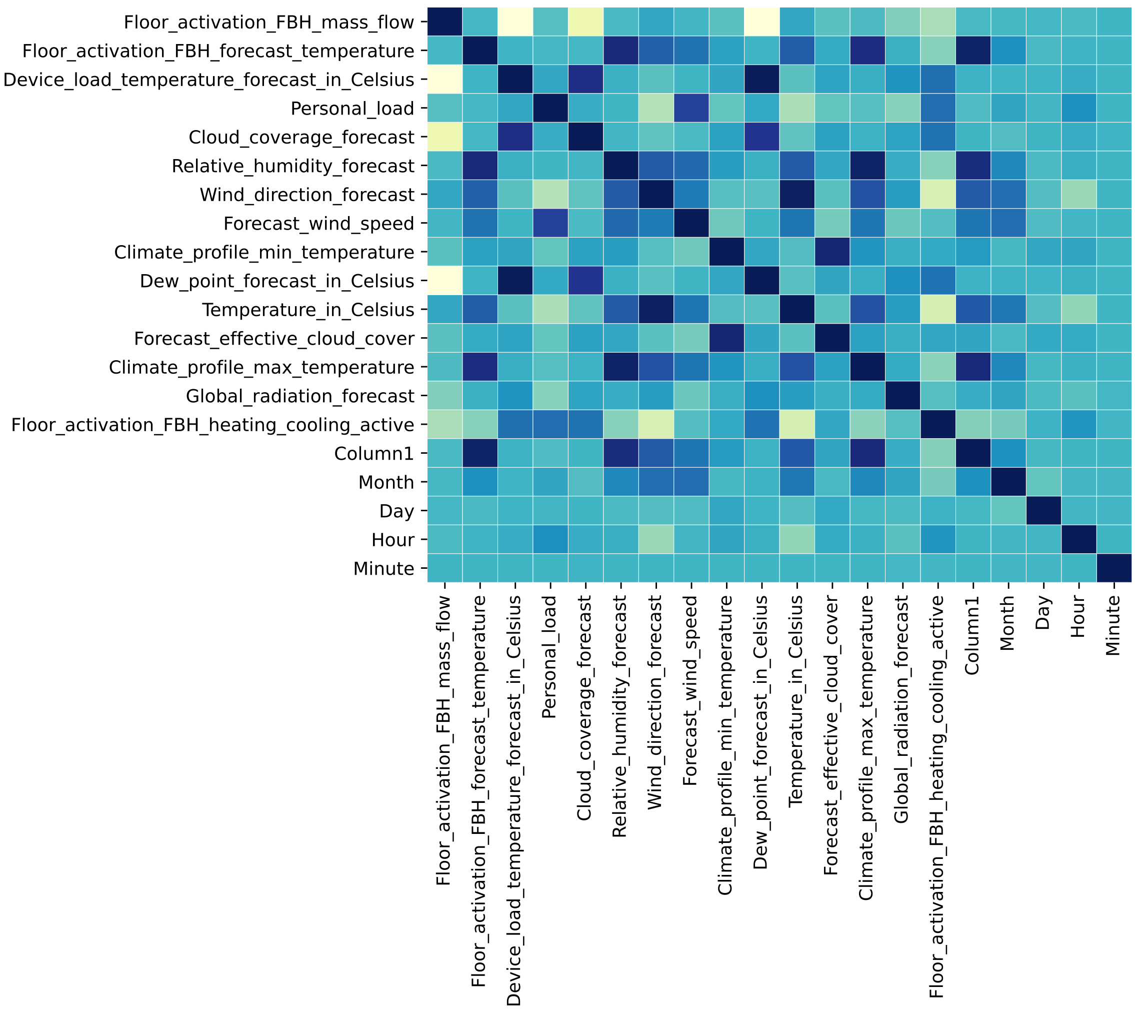


Figure 112: Visualization of Correlations between data for zone 12

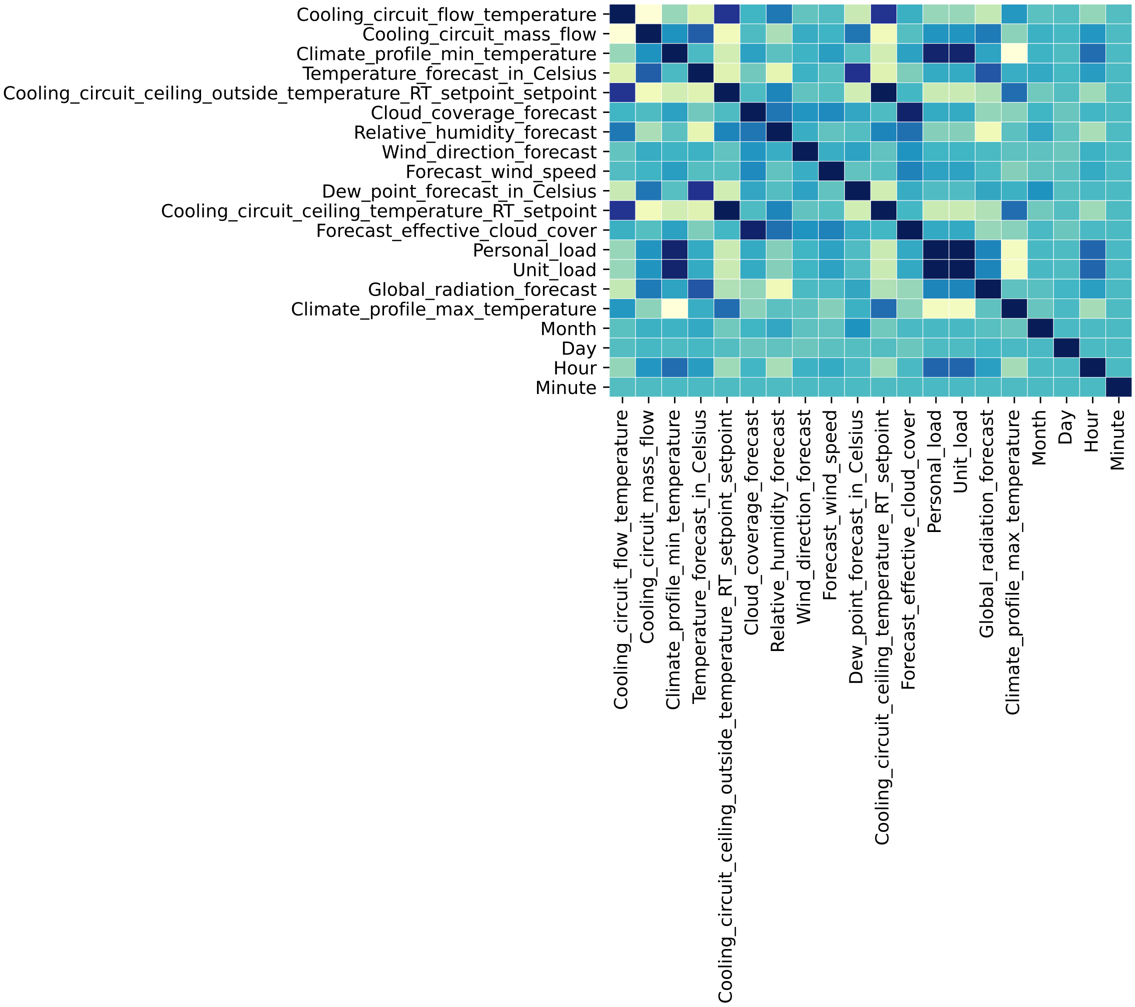


Figure 113: Visualization of Correlations between data for zone 13

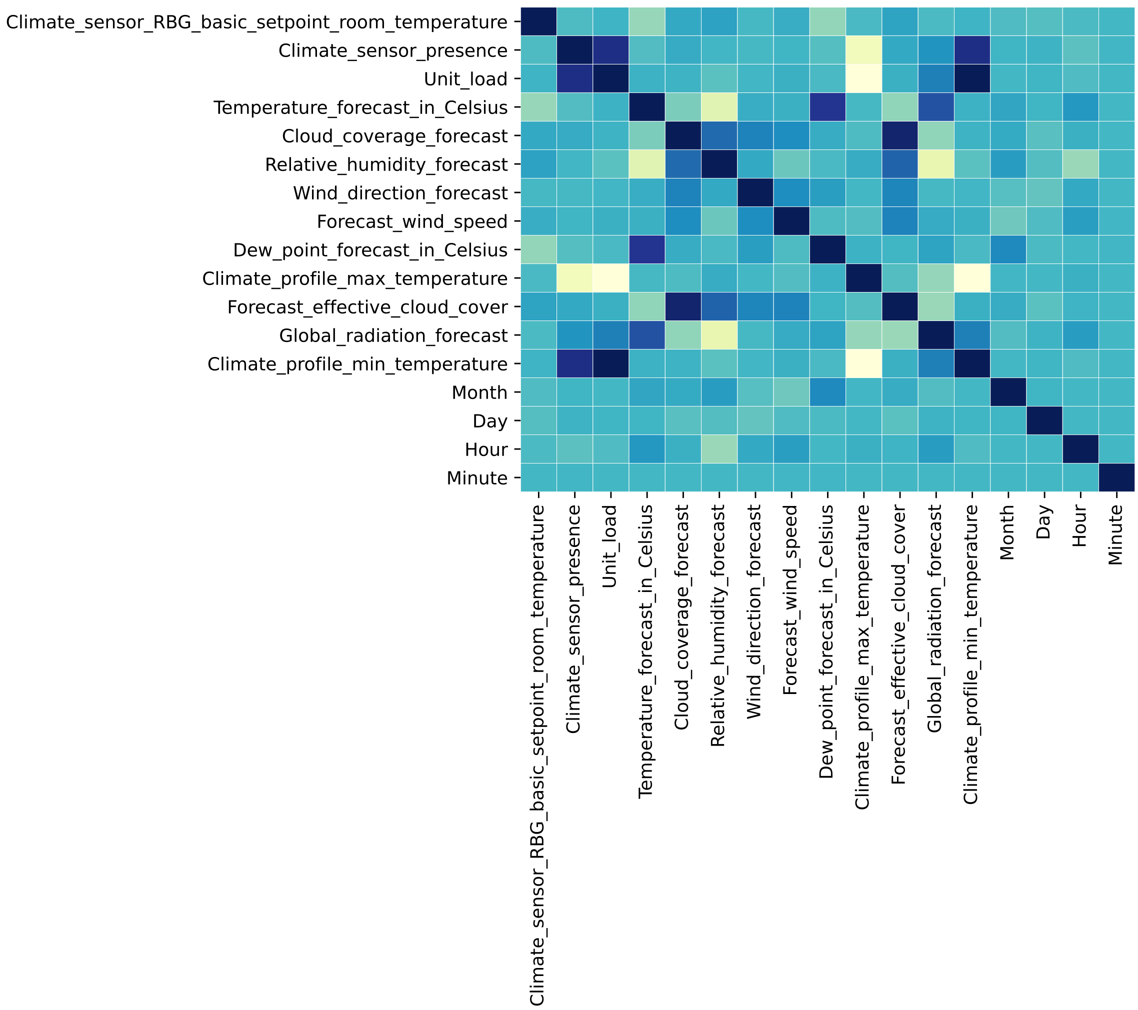


Figure 114: Visualization of Correlations between data for zone 14

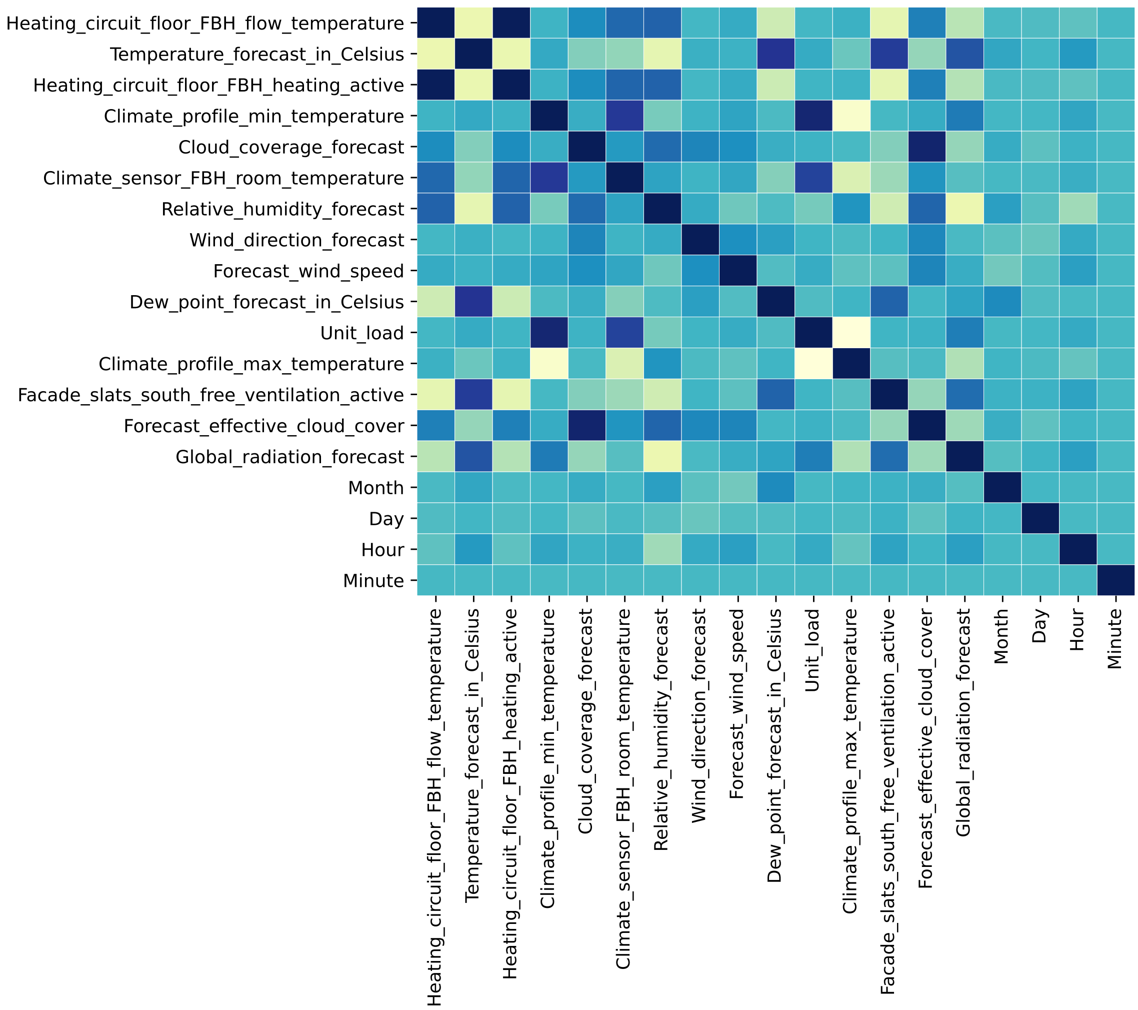


Figure 115: Visualization of Correlations between data for zone 15

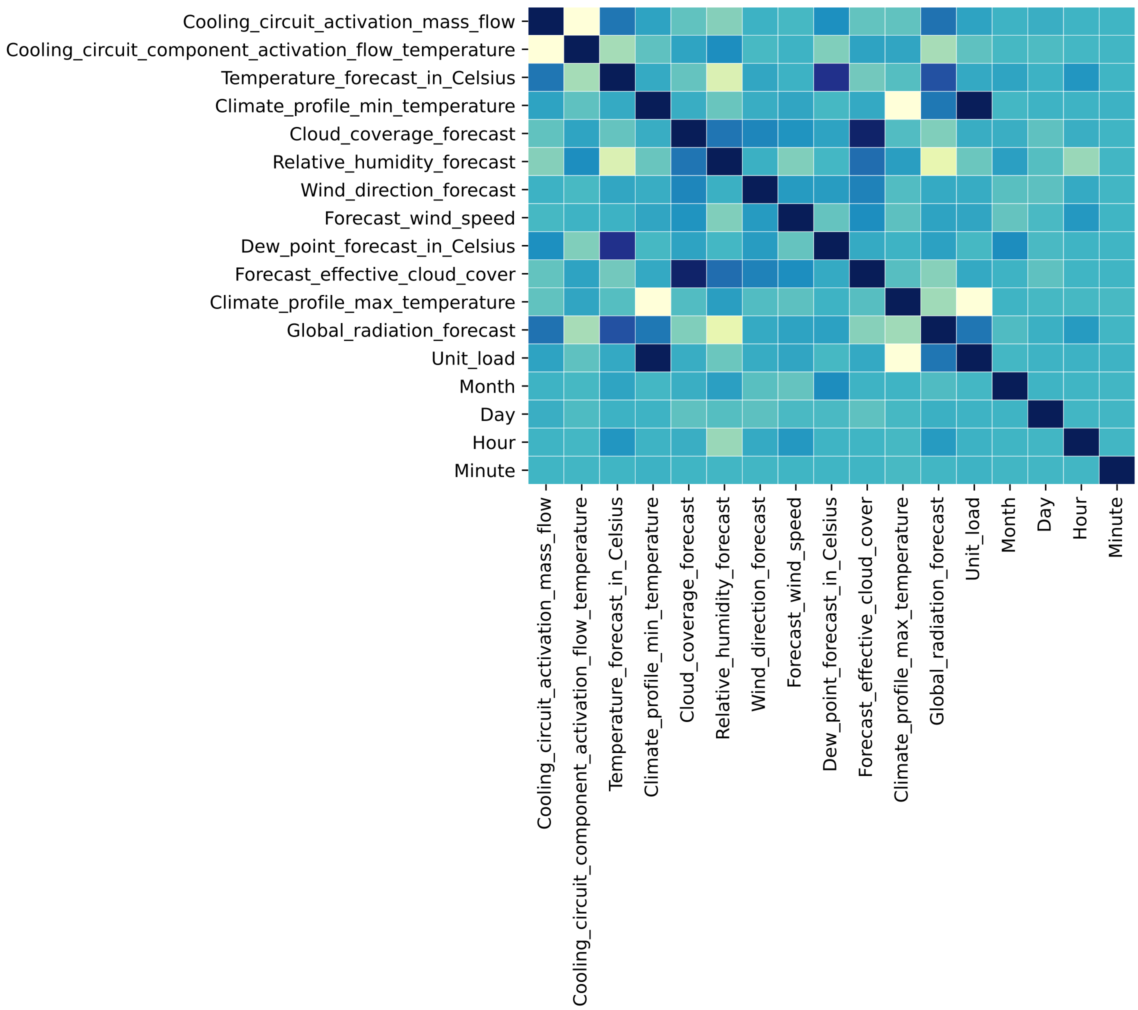


Figure 116: Visualization of Correlations between data for zone 16

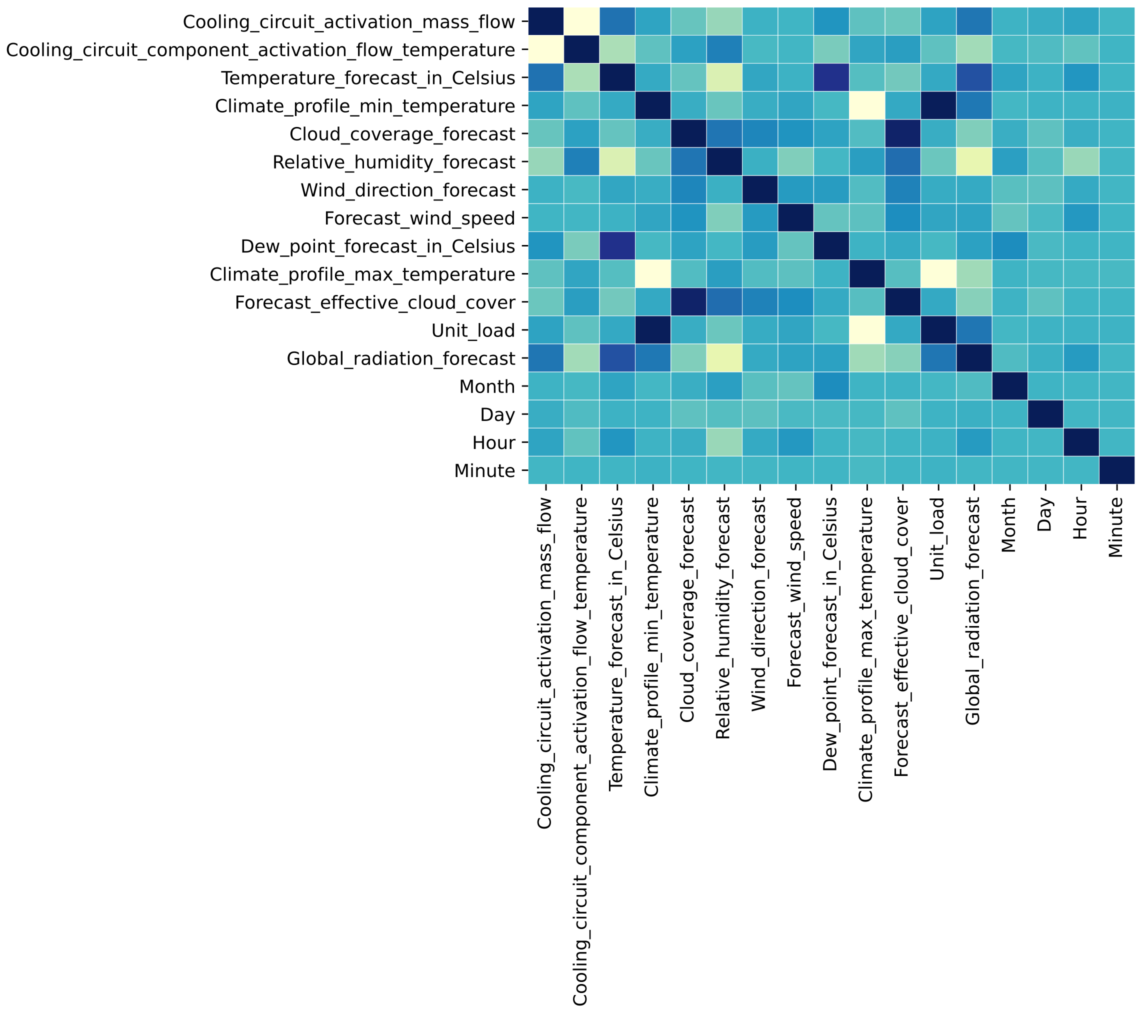


Figure 117: Visualization of Correlations between data for zone 17

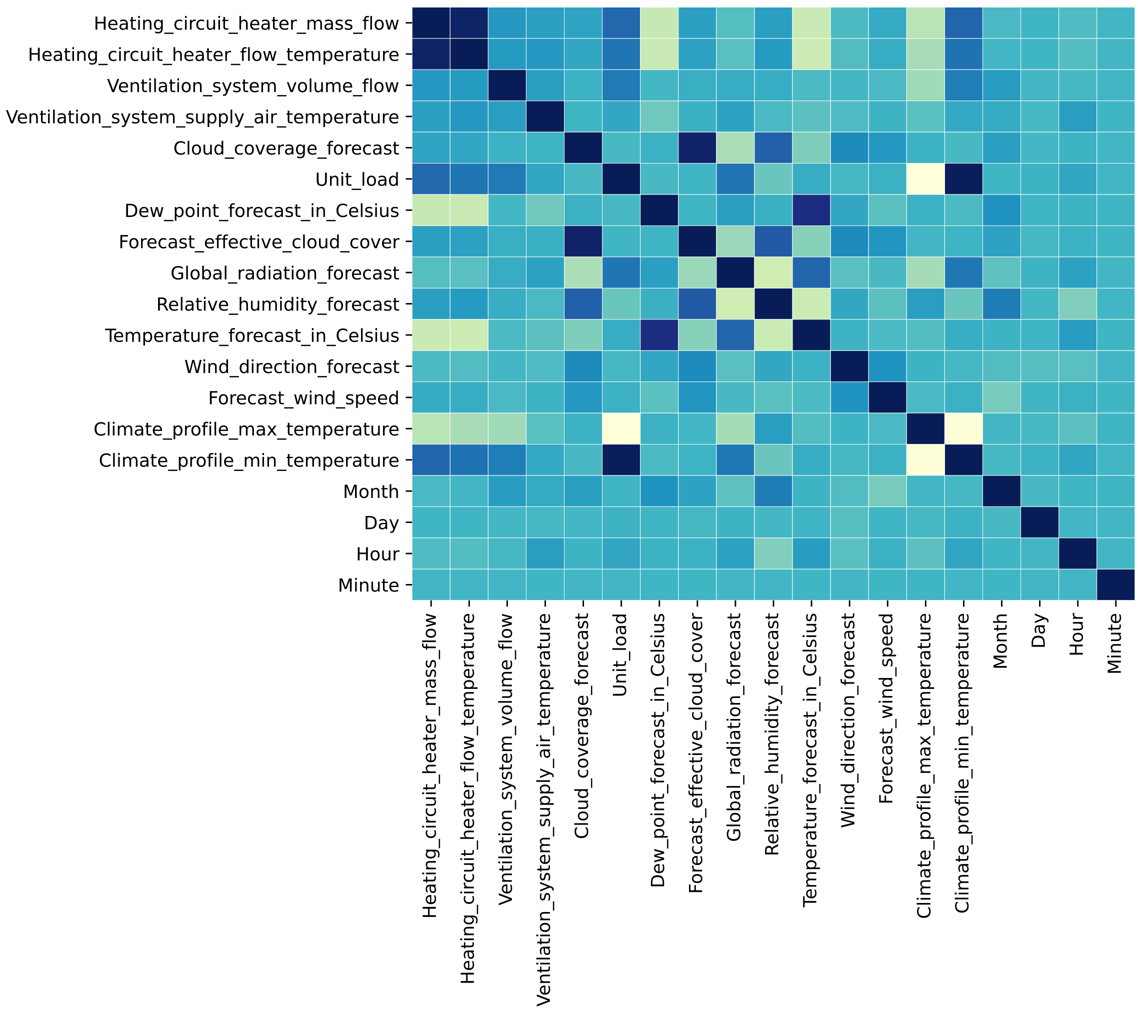


Figure 118: Visualization of Correlations between data for zone 18

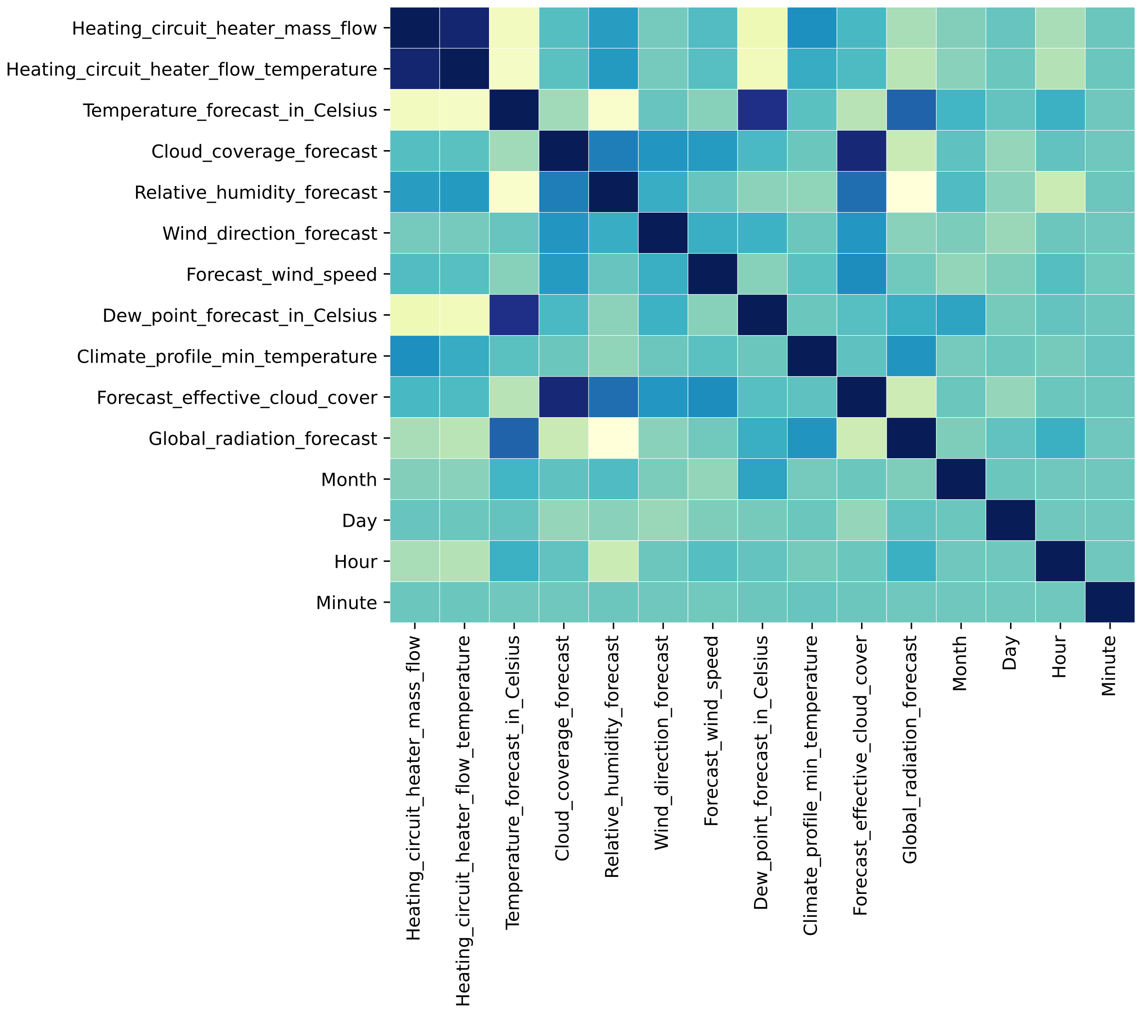


Figure 119: Visualization of Correlations between data for zone 20

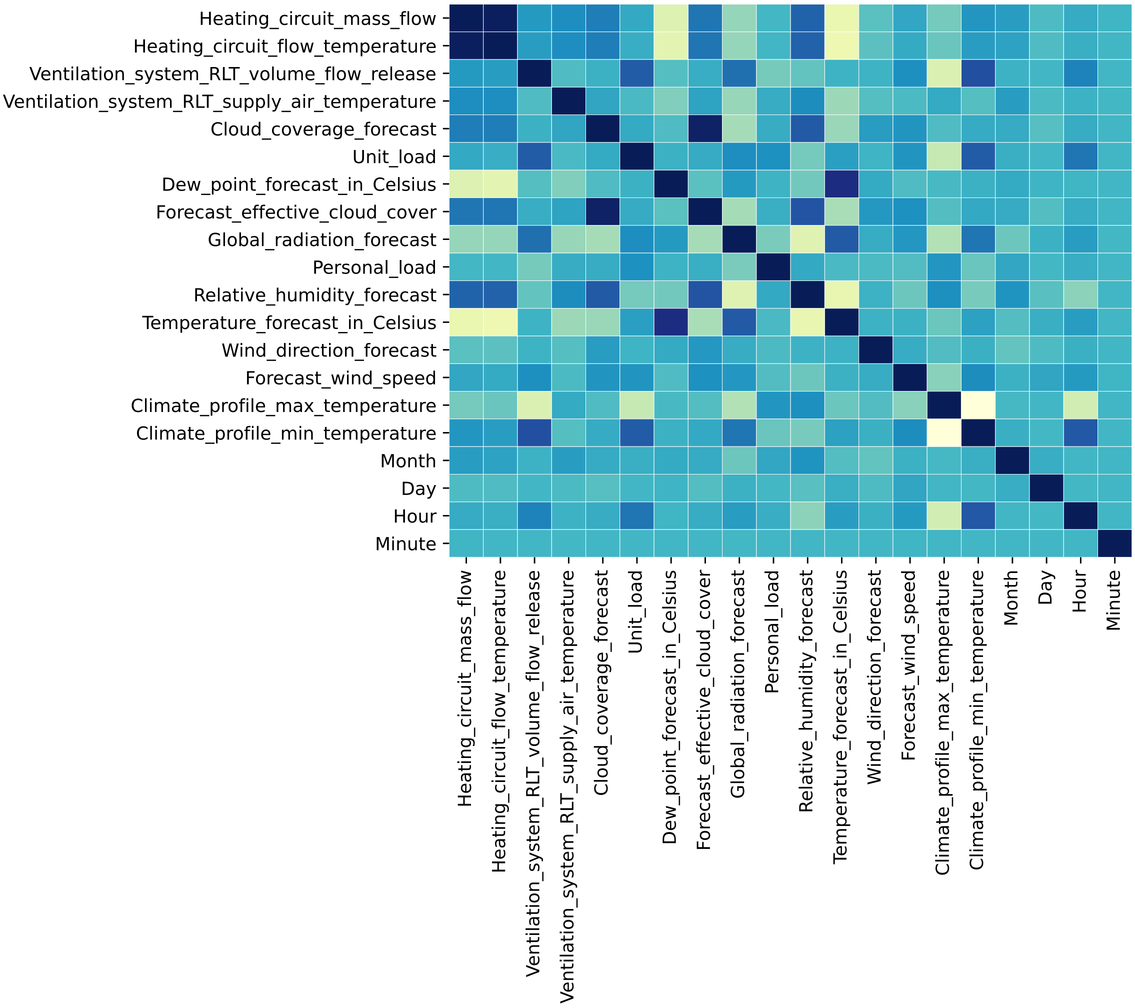


Figure 120: Visualization of Correlations between data for zone 21

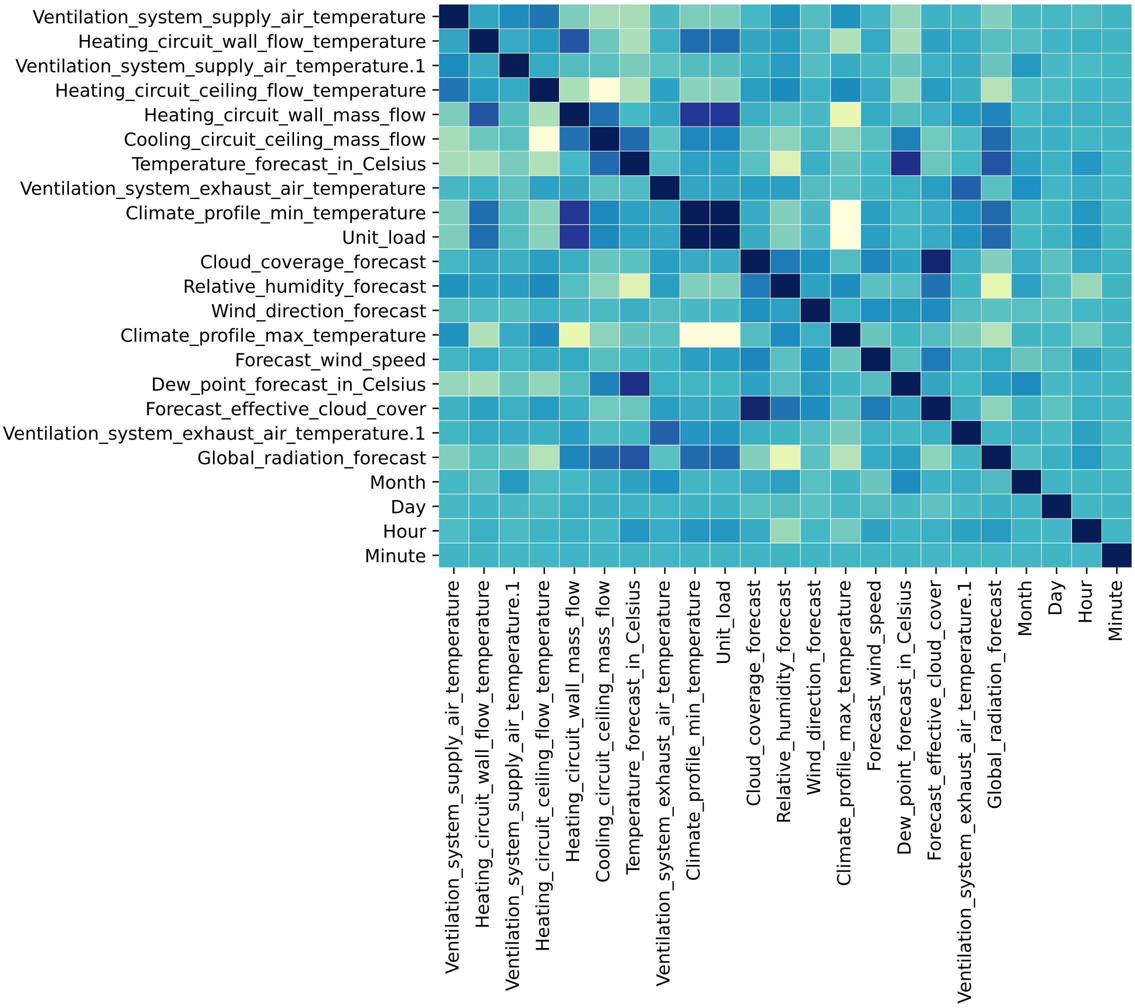


Figure 121: Visualization of Correlations between data for zone 22

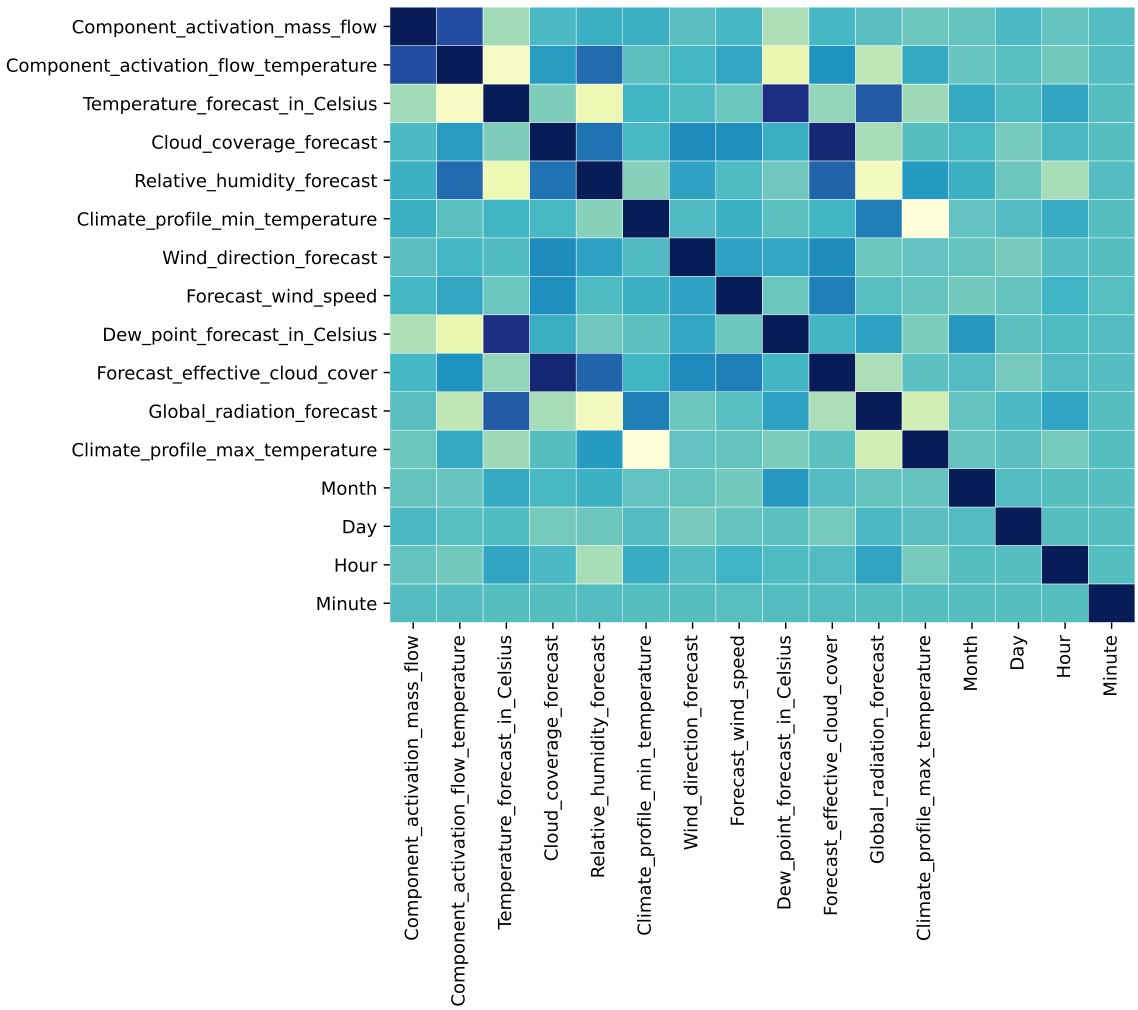


Figure 122: Visualization of Correlations between data for zone 23