Machine Learning to improve indoor climate and building energy performance

Christoph Hopmann

Thesis to obtain the Master of Science Degree in

Energy Engineering and Management

Supervisors: Prof. Carlos Augusto Santos Silva

Eng. Christian Luft

Examination Committee

Chairperson: Prof. Jorge de Saldanha Gonçalves Matos

Supervisor: Prof. Carlos Augusto Santos Silva

Member of the Committee: Prof.ª Susana Margarida da Silva Vieira

October 2017

ACKNOWLEDGEMENTS

I would like to thank all the people who contributed in some way to the work described in this thesis.
First and foremost, I wish to thank my two supervisors, Professor Carlos Silva and Christian Luft, for their support during the creation of this thesis. Their regular inputs helped to drive this work towards its goal, while the freedom I was entrusted with enabled me to experiment with what works on what doesn't.
I would also like to thank Drees & Sommer for the opportunity to write my thesis with the company and all the material and idealistic support I received from the company.
Finally, I would like to thank InnoEnergy, the KIT and IST that allowed me to follow my studies in an incredibly inspiring environment.

ABSTRACT

Using model predictive control for energy management systems is becoming more popular. These systems control the building performance based on a forecast of the control based on a building model. Uncertainties in building operation are a major issue in the use of model predictive control, as they decrease prognosis precision. This includes user's varying thermal requirements. The aim of this thesis is to develop a comfort temperature predictor for individually controllable offices based on the weather conditions using machine learning. For this purpose, the fundamentals of indoor comfort and machine learning are presented. Recent trends in research on machine learning in buildings are reviewed. The methodology for developing and testing the comfort temperature predictor is explained. Afterwards the results of forecasting the comfort temperature and the energetic impact of the predictor are described.

The results show that supervised learning artificial neural networks and Gaussian Process Regression tools can predict comfort temperatures based on weather conditions with sufficient precision, better than the currently common temperature setpoints. Roughly one year of data is required to reach this performance. Real-time learning with reinforcement learning using artificial neural network value function approximation and sample reuse needs 50 days of learning to reach good precision. Supervised learning can reduce the heating load and reduce overheating, while there is no positive impact on the cooling load with either tool. The thesis closes with a review of the tool, an outlook towards its improvements and applications, and the overall impact of machine learning on the building sector.

Key-words: building energy systems, thermal comfort, machine learning, predictive control

RESUMO

A utilização de modelos de controlo preditivo para a sistemas de gestão energética é cada vez mais popular na gestão de energia em edifícios. Estes sistemas controlam o desempenho do edifício baseados na previsão de um modelo do edifício. A incerteza na operação de edifícios é um problema vital na utilização de modelos preditivos, uma vez que diminuem a precisão do prognóstico. Este problema inclui vários requisitos térmicos impostos pelo utilizador. O objetivo desta tese é o desenvolvimento de um preditor de conforto térmico para escritórios individualmente controlados, com base em condições meteorológicas utilizando Machine Learning. Com este propósito, são apresentados os fundamentos de conforto térmico interior e de Machine Learning. São ainda revistas as tendências recentes de investigação no campo de Machine Learning para edifícios e apresentada a metodologia para o desenvolvimento e teste do preditor para conforto térmico. Finalmente, são apresentados e discutidos os resultados da previsão de conforto térmico e o impacto energético do preditor.

Os resultados mostram que quer as redes neuronais artificiais, quer as ferramentas de regressão de processos gaussianos são capazes de prever as temperaturas de conforto baseadas em condições meteorológicas com precisão suficiente, e de forma melhor do que as temperaturas de referência utilizadas frequentemente. É necessário aproximadamente um ano de dados para atingir este nível de desempenho. A aprendizagem on-line baseada em reforço utilizando aproximações de valores de funções de redes neuronais artificiais e a reutilização de amostras necessita apenas de 50 dias de aprendizagem para alcançar um bom nível de precisão. A aprendizagem supervisionada é capaz de reduzir a carga de aquecimento e reduzir o sobreaquecimento, mas não tem um impacto positivo na carga de arrefecimento. A tese finaliza com uma revisão da ferramenta, uma discussão relativa a possíveis melhorias e aplicações, bem como o impacto futuro de Machine Learning no sector de edifícios.

Palavras-chave: sistemas energéticos de edifícios, conforto térmico, machine learning, controlo preditivo

Content

Figures VI

Tables VII

Symbols VIII

Indices VIII

Abbreviations IX

1. Introduction 1

2. Fundamentals 3

2.1. Indoor Thermal Comfort 3

2.1.1. Thermal Comfort Model according to Fanger 4

2.1.2. Adaptive Thermal Comfort 5

2.2. Machine Learning 6

2.2.1. Machine Learning Algorithms 8

2.2.2. Reinforcement Learning 14

3. Literature Review 16

3.1. Advanced HVAC Control 16

3.1.1. Recent developments in advanced HVAC Control 16

3.1.2. Advanced comfort control 17

3.2. Load Prediction 20

3.3. User Behaviour 22

3.4. Classification of Previous Work 24

3.5. Commercially Available Products 30

3.5.1. MeteoViva 30

3.5.2. BuildingIQ 30

3.5.3. Learning Thermostats 30

4. Methodology 32

4.1. Choice of Model Parameters 32

4.2. Data Preparation 34

4.3. Choice of Machine Learning Tools 34

4.4. Learning Approach, Testing, Validation 35

4.5. Energy Evaluation 36

5. Program Development 37

5.1. Used Data 37

5.2. Supervised Learning 38

5.3. Reinforcement Learning 40

5.4. Building Model 42

6. Results ... 45
6.1. Supervised Learning ... 46
6.1.1. Neural Networks ... 46
6.1.2. Regression Learners ... 49
6.1.3. Comparison of Supervised Learning Tools ... 51
6.2. Reinforcement Learning ... 52
6.2.1. ANN-based function approximation ... 53
6.3. Energy Evaluation ... 60
7. Conclusion ... 65
References ... XII
Annex ... XVII
1. PMV-Calculation [18] ... XVII
2. Values for decision Top 10 Neural Networks ... XVIII
3. Values for decision Top 10 Regression Learning Tools ... XVIII

Figures

Figure 1 Temperature-dependant shares of heat transfer mechanisms on human heat dissipation [24] 3
Figure 2 Primary, secondary and additional factors influencing thermal comfort [24] ... 4
Figure 3 Connection between PPD and PMV [25] ... 4
Figure 4 Types of Machine Learning, adopted from [11] ... 6
Figure 5 Perceptron ANN Scheme [11,31] ... 10
Figure 6 Feed-forward back propagation neural network [11,31] ... 11
Figure 7 Interaction between RL agent and Environment [39] ... 15
Figure 8 Block Diagram Intelligent Comfort Control System [42] ... 17
Figure 9 Scheme of Neural network for individual comfort evaluation [49] ... 22
Figure 10 Office Model ... 43
Figure 11 Neural Network Size and Weather Parameter Dependant Total MSE - Top 10 for eight sample rooms ... 46
Figure 12 Neural Network Top 3 Total RMS for all rooms, compared to reference scenario ... 47
Figure 13 Neural Network 2349-33 Total RMS over training data size ... 48
Figure 14 Neural Network 25679-33 Total RMS over training data size ... 48
Figure 15 Neural Network 245789-33 Total RMS over training data size ... 48
Figure 16 Regression Tool and Weather Parameter Dependant Total RMS - Top 10 for eight sample rooms ... 50
Figure 17 Regression Tool Top 3 RMS for all rooms, compared to reference scenario ... 50
Figure 18 Neural Network and Regression Tools Comparison ... 51
Figure 19 Supervised learning process ... 52
Figure 20 Reinforcement learning process ... 52
Figure 21 Reinforcement Learning: Error Function Evaluation - Simple Error ... 54
Figure 22 Reinforcement Learning: Error Function Evaluation - Half MSE ... 54
Figure 23 Reinforcement Learning: Error Function Evaluation – RMS ... 54
Figure 24 Reinforcement Learning: Discounted Learning - Discount Rate 0.985 ... 55
Figure 25 Reinforcement Learning: Discounted Learning - Discount Rate 0.995 ... 55
Figure 26 Reinforcement Learning: Sample Reuse - Learning Rate 0.01 ... 56
Figure 27 Reinforcement Learning: Sample Reuse - Learning Rate 0.05 ... 56
Figure 28 Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.01, Discount Rate 0.995 ... 57
Figure 29 Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.01, Discount Rate 0.985 ... 58
Figure 30 Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.05, Discount Rate 0.985 ... 58
Figure 31 Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.03, Discount Rate 0.985 ... 58
Figure 32 Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.03, Discount Rate 0.975 ... 59
Figure 33 Heating Loads per method and cardinal direction ... 61
Figure 34 Hours heating setpoints are not reached per method and cardinal direction ... 61
Figure 35 Cooling Loads per method and cardinal direction ... 62
Figure 36 Hours cooling setpoints are not reached per method and cardinal direction ... 62
Figure 37 Mean arithmetic error in Spring ... 63
Figure 38 Mean arithmetic error in summer ... 64
Figure 39 Mean arithmetic error in autumn ... 64
Figure 40 Mean arithmetic error in winter ... 64

Tables

Table 1 Groups of Machine Learning algorithms [11] 12
Table 2 Comparison of commonly used methods for load prediction, taken from [46] 20
Table 3 Classification of Papers on Artificial Intelligence in Building Energy Systems 25
Table 4 Available Rooms per floor 37
Table 5 Parameter Minima and Maxima for normalization.. 38
Table 6 Construction Materials [81]... 43
Table 7 Weather Parameters 45
Table 8 Matlab Standard Regression Learner Tools 49
Table 9 Annex 2: Values for decision Top 10 Neural Networks.. XVIII
Table 10 Annex 3: Values for decision Top 10 Regression Learning Tools XVIII

Symbols

Symbol	Meaning
a	Model parameter
c	Linear offset
e	Error
h	User input
P	probability
p	Placeholder for parameter
Q	Total error
T	Temperature in °C
w	Weight
x	Input
y	Prediction
α	Discount rate
β	Model parameter
δ	Neuron responsibility for error
ε	Model bias
λ	Adaptive coefficient

Indices

Index	Meaning
(h)	Hidden layer
(o)	Output layer
0,5MSE	Half mean squared error
comf	Comfort
i, j, k	Counting indices
max	Maximum value
min	Minimum value
real	Real value
RMS	Root mean squared error
RMT	Running mean temperature
simple	Arithmetic error
stand	Standardized value

Abbreviations

Abbreviation	Meaning
ANN	Artificial Neural Network
aPMV	Adaptive Predicted Mean Vote
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BNMI	Best Network after Multiple Iterations
CDC	Climate Data Centre
DP	Dynamic Programming
DWD	German Weather Service (Deutscher Wetterdienst)
EN	European Standard
EnEv	German Energy Saving Regulation (Energieeinsparverordnung)
ETH Zurich	Swiss Federal Institute of Technology in Zurich (Eidgenössische Technische Hochschule Zürich)
FCM	Fuzzy C-Means
FFNN	Feed Forward Neural Network
GPR	Gaussian Process Regression
HME	Hierarchical Mixture of Experts
HVAC	Heating, Ventilation, Air Conditioning
ICCS	Intelligent Comfort Control System
IMBPC	Intelligent Model-based Predictive Control
ISO	International Organization for Standardization
KDD	Knowledge Discovery in Databases
LBMPC	Learning-based Model Predictive Control
LED	Light-emitting diode
LS-SVM	Least-Squares Support Vector Machine
MC	Monte Carlo
MDP	Markov Decision Process
MPC	Model-based Predictive Control
MSE	Mean Squared Error
NNC	Neural Network Controller
NNEM	Neural Network Evaluation Model
OEM	Original Equipment Manufacturer
PEO	Predictive Energy Optimization
PID	Proportional-Integrate-Derivative
PMV	Predicted Mean Vote
PPD	Predicted Percentage of Dissatisfied

RBF	Radial Basis Function
RL	Reinforcement Learning
RMS	Root Mean Squared Error
SVM	Support Vector Machine
SVR	Support Vector Regression
TD	Temporal Difference
VAV	Variable Air Volume

1. Introduction

The built environment plays a significant role in modern societies, as humans in western countries spend 90 % of their time indoors [1,2]. Indoor comfort can have an important impact on someone's well-being, not only short-term while occupying a specific building, but also in the long-run. Thermal conditions in a building affect workers' productivity and may even influence the way users act and think. Workers in Great Britain spend around two percent of their office hours adjusting the thermal environment, equalling losses in productivity of 15 billion Euro [3]. In Australia overly warm offices reduce productivity by 5,9 billion Euro per year [3]. Creating a comfortable environment is therefore not only elementary for personal well-being but also has an economic upside.

A major issue that arises when choosing the environmental conditions within an office are personal preferences that may vary, for example, by several degrees centigrade. Even the same individual may have changing requirements depending on diverse influencing factors [4]. Consequently, a trend towards more individual control options can be observed [3]. Apart from making the personal thermal environment more comfortable, individual control options also make occupants more forgiving towards the conditioning system compared to systems centrally controlled by the building manager, i.e. wider condition bands are considered comfortable [5,6]. Individual control however entails a different issue: occupant behaviour is in general difficult to predict and is thus one of the largest sources of uncertainties during building operation. On average, users tend to endure uncomfortable conditions until a "crisis of discomfort" is reached, at which point they over-compensate even minor annoyances. Over-compensation may be boosted through a potential system inherent delay between thermostat input and system reaction. Generally, occupant behaviour can multiply the energy consumption of offices by a factor of two or more [4,6,7].

Regarding energy consumption, the building sector is the biggest consuming sector with a share of 40 %, followed by transport [8,9]. A significant percentage is used for room air conditioning. Considering climate change and related energy efficiency targets, for example the EU 2030 climate and energy framework, a subsistent burden falls onto the real estate industry [10]. With an ever-increasing demand towards more reliable indoor conditions in buildings and uncertainties connected with user behaviour there is a demand for the use of new technologies to both decrease energy consumption and increase user comfort.

Over the last years machine learning has found its way into a broad field of applications, ranging from the classic example of spam filtering to autonomous driving, smart personal assistants like Apple's *Siri* or Google *Now*, to load predictions in the field of energy [11], as it enables the use of large amounts of data for various purposes. In buildings, its main application has been in load prediction, control optimization, and occupation and comfort prediction. Commercial implementation however is scarce. Some companies use machine learning for model predictive control of HVAC systems, among them the German company MeteoViva and US-American BuildingIQ [12,13]. "Smart" Thermostats like the *Nest* or *tado* use machine learning to determine occupation patterns, setpoint preferences and the interaction between the conditioning system, the building and the environment [14,15]. Model predictive control for larger buildings and learning thermostats have shown potential energy savings of

up to 40 % compared to traditional control systems [16]. They are however applied in very different fields.
Model predictive control, in principle, is the idea of using a dynamic system model to forecast system behaviour. The forecast supports the choice of the control move at the current time. In an optimal scenario, the forecast, together with past data that is used to estimate the state value, enables the determination of the best possible control move [17]. In the context of building automation, this could mean using a dynamic building model to optimize, for example, the supply flow to HVAC components to reach the desired temperature setpoints, with a minimal energy demand. However, uncertainties in building operation are a major issue in the use of model predictive control, as they decrease prognosis precision. This includes weather forecast inaccuracy as well as user behaviour, comprising varying thermal requirements. The basic concept of using predictions to effectively control HVAC systems gets undermined by spontaneous changes in user requirements. If one could predict these varying needs, i.e. efficiently implement the concept behind smart thermostats on a bigger scale, uncertainty caused by user behaviour could be reduced, increasing the effectiveness of different types of predictive control algorithms.
Fanger developed the Predicted Mean Vote (PMV) with the purpose of evaluating thermal conditions in buildings depending on several factors, which could in turn be used to predict comfortable conditions [18]. However, the focus of PMV lies on the behaviour of groups, not individuals. Furthermore it includes highly personal and local parameters, that are difficult to measure, and is thus not easily applicable on a broad scale [4,19–21]. A method needs to be developed that uses parameters that are readily available in most buildings. Adaptive comfort theory suggests that users react to the climate. The aim of this thesis is to develop a comfort temperature predictor for individually controllable offices based on the environmental weather conditions using machine learning.
For this purpose, the fundamentals of indoor comfort and machine learning are presented. Recent trends in research on machine learning in buildings are reviewed and classified, and exemplary commercial applications are laid out. In the following chapter, the methodology for developing and testing the comfort temperature predictor is explained, followed by development of the proposed model. Afterwards the results of forecasting the comfort temperature and the energetic impact of the predictor are described. The thesis closes with a review of the tool, an outlook towards improvements and applications of the tool and the overall impact of machine learning on the building sector.

2. Fundamentals

To build a comfort temperature predictor it is necessary to understand both thermal comfort and machine learning. In the following chapter, the basics of human indoor thermal comfort are presented, along with the comfort approach developed by *Fanger* and the adaptive comfort approach. A general introduction to machine learning is also given, with a focus on supervised learning and reinforcement learning. Unsupervised learning methods are not presented in detail, as they are not applied in the development of the comfort predictor.

2.1. Indoor Thermal Comfort

Thermal comfort is achieved when a human does not feel the need to change the state of the thermal environment or to adapt to it. This condition is reached when the heat balance of a human is neutral, i.e. the heat produced by metabolism is equal to the heat dissipated to the environment [22,23]. Dissipation of heat takes place through convection, radiation and evaporation, with their shares depending on the ambient temperature, as depicted in Figure **1** [24].

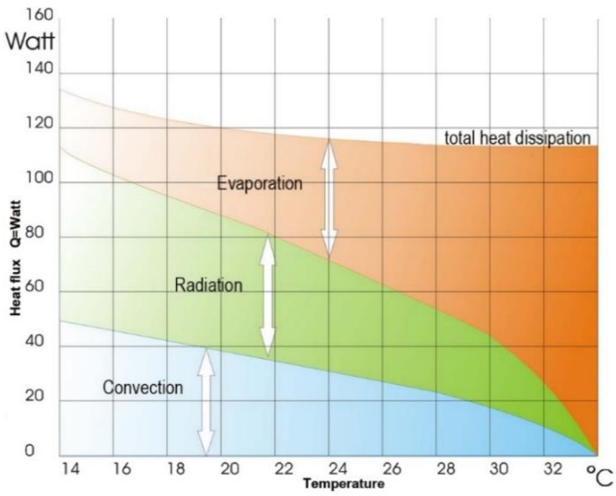

*Figure **1** Temperature-dependant shares of heat transfer mechanisms on human heat dissipation [24]*

The perception of the thermal environment is mainly influenced by six factors: air temperature, velocity and relative humidity, temperature of surrounding surfaces, clothes and a person's activity. Additionally, there are further physical, physiological and other factors, that influence one's thermal comfort, as depicted in Figure **2** [24]. The thermal environment can have a significant impact on productivity and health of the user and shows an influence on the way humans think. Studies show that warm temperatures suit to creative work, while colder environments help to keep people alert during monotonous activities. It is suggested that unfit thermal conditions lead to annual productivity losses of around 15 billion Euro in Great Britain and 5,9 billion Euro in Australia [3]. Suiting ambient conditions to occupants is thus an important task when dimensioning conditioning systems.

In this chapter, engineering approaches to evaluate indoor thermal comfort - the predicted mean vote (PMV) and the predicted percentage of dissatisfied (PPD) developed by *Fanger*, which are common practice in industry - will be presented, as well as more recent methods based on the adaptive comfort approach [5,18].

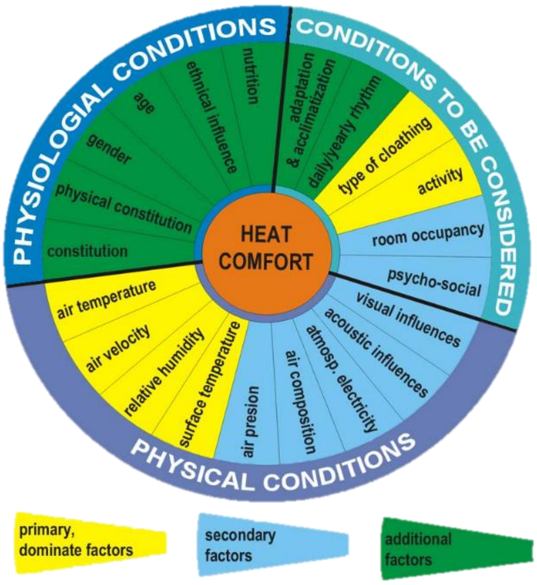

*Figure **2** Primary, secondary and additional factors influencing thermal comfort [24]*

2.1.1. Thermal Comfort Model according to *Fanger*

Fanger developed a thermal comfort model based on a heat balance model of the human body. Within this model, the body is in thermal exchange with the surroundings only as a “passive recipient of thermal stimuli” [23]. It was tested and confirmed through experiments in climate chambers, with subjects going through different thermal conditions and answering a questionnaire on their thermal perception. *Fanger* used the results of these studies to develop an equation that can be used to predict the thermal comfort of a group of people based on the indoor conditions, and is especially suited for artificially climatized spaces. The input parameters for the PMV are the air temperature and velocity, the relative humidity, the mean radiant temperature, the clothing level and the activity level. The detailed calculations can be found in Annex **1**. The PMV is a scale ranging from - 3 (cold) to + 3 (hot), with a 0-vote being neutral. *Fanger* defined the range from - 0,5 to + 0,5 as acceptable conditions. The PPD is used to evaluate the percentage of occupants that will be dissatisfied under certain ambient conditions. The PPD is related to the PMV: for a PMV of ± 3, 90 % of the occupants will be unhappy with their environment, while at a PMV of 0 5 % of occupants will be discontent [24]. Both indices are depicted in Figure **3** [25]. It has to be noted that, due to the way the PMV has been developed, it can only be used under steady-state conditions. The PMV is used in several national and international standards, among them the ASHRAE standard for indoor comfort [26], the European standard EN 15251 [27] and the ISO 7730 [28].

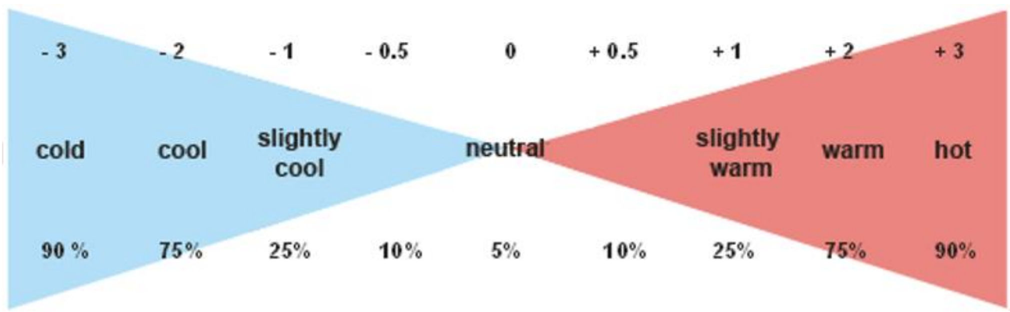

*Figure **3** Connection between PPD and PMV [25]*

2.1.2. Adaptive Thermal Comfort

In contrast to the assumption that humans are passive towards their thermal environment, people usually adapt if given the chance. Adaptation happens both consciously and unconsciously, for example by changing clothes or behaviours. This forms the basis for the adaptive thermal comfort principle, according to which “people react in ways which tend to restore their comfort if a change occurs such as to produce discomfort” [5]. There are several variables building occupants react to, among them the climate and the nature of the building they are in and the services it offers. This leads to a constant change in comfort temperature, which is not well depicted by the PMV. Going more into detail, the outdoor temperature plays the most important role of all the weather factors, as it influences clothing, posture, the metabolic rate for given activities and the way in which occupants use building services. It seems that both past outdoor temperature, dating back to up to a week, and the weather forecast for the present day mainly influence people’s choice of clothes – assuming there is no strict dress code [6]. Comfort may also be influenced by sociological and geographical preferences, which is not included in the *Fanger* model [19]. Another important factor is the occupants’ expectation towards the building’s thermal performance, which highly varies with the building equipment. Interestingly enough, more control options usually make occupants more forgiving, i.e. widening their comfort range [5,6]. Due to psychological adaptation, the comfort range proposed by the adaptive thermal comfort indices usually exceeds the range suggested by the PMV, implying over-conditioning when using the PMV [19]. Adaptive comfort standards have generally been developed with naturally ventilated buildings in mind, since the connection between outdoor temperature and comfort temperature is clearer than for heated and cooled buildings. They are still applicable to mechanically ventilated buildings.

Humphreys and *Nicol* and *Nicol* and *Raja* suggest an adaptive comfort temperature standard using the running mean outside temperature

$$T_{RMT,n} = (1-\alpha)\cdot\sum_{i=0}^{j}\alpha^{i}T_{n-1-j}. \tag{1}$$

With

$T_{RMT,n}$ … Running mean temperature for n days

α … discount rate

T_n … Temperature on day n.

Their comfort temperature is calculated defined as

$$T_{comf} = \begin{cases} 0{,}302T_{RMT,n} + 19{,}39\ for\ T_{RMT,n} > 10°C \\ 22{,}88\ °C\ for\ T_{RMT,n} < 10°C \end{cases} \tag{2}$$

The numeric values used were gathered through empirical field studies [23,29]. *Yao et al*. suggest a combination of the PMV and adaptive approaches, called the Adaptive Predicted Mean Vote, aPMV. An adaptive coefficient λ is added to the PMV calculation to compensate for the PMV’s problem in warm and cold climates, leading to

$$aPMV = \frac{PMV}{1+\lambda\cdot PMV} \tag{3}$$

With the coefficients according to equation 4 [30].

$$\lambda = \begin{cases} 0{,}293\ for\ warm\ conditions \\ 0{,}125\ for\ cold\ conditions \end{cases} \tag{4}$$

2.2. Machine Learning

The goal of machine learning is to build computer systems that automatically improve with experience [11]. A machine is thus given the ability to learn without being explicitly programmed [31]. Present data is used to predict or to respond to future data [32]. As *Kwok et al*. put it: "Fundamentally, the emphasis of machine learning is on the system's ability to adapt or change. […] After learning or adaptation, the system is expected to have better future performance on the same or a related task" [33]. Apart from "self-improving" systems, machine learning also helps to understand and collect useful information from large or complicated sets of data [33]. The field has shown fast developments over the last years and is becoming more and more popular for a large variety of applications, ranging from the classic example of spam filtering to autonomous driving, smart personal assistants like Apple's *Siri* or Google *Now*, to load predictions in the field of energy [11].

Machine learning algorithms can be classified into three main learning types: supervised learning, unsupervised learning and reinforcement learning. There are also hybrids between the learning types, like semi-supervised learning. Figure **4** gives a graphical representation of these types including their typical output data type, the used method and an application example.

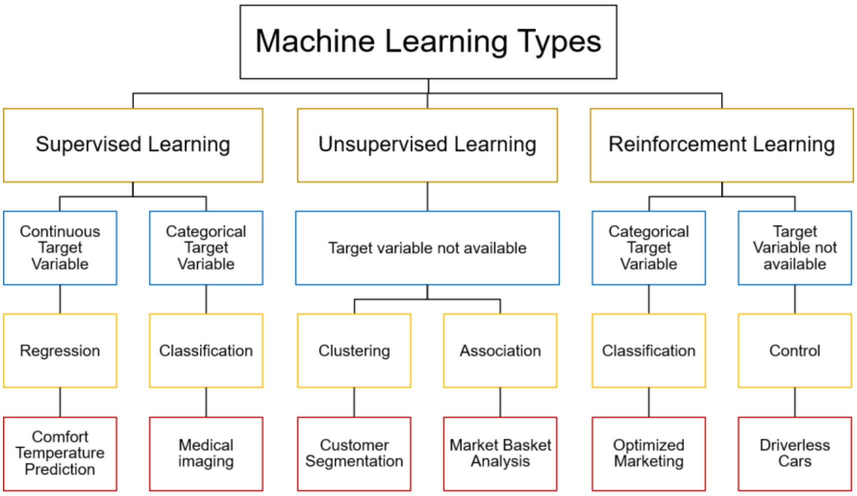

*Figure **4** Types of Machine Learning, adopted from [11]*

In supervised learning, the learning system is fed with inputs and the corresponding desired outputs, so called labelled data or training data. Depending on the type of label the learning approach differs. For continuous output data, for example housing prices, one uses regression, while for categorical outputs, like medical diagnosis, classification is used. Training sets are usually human-made and have a significant impact on the quality of the learning process. Pre-processing of training data is thus essential in supervised learning; inconsistencies, contradictions and other errors have to be filtered. Furthermore, training sets should be sufficiently big and, ideally, range over all potential inputs and the corresponding outputs. Semi-supervised learning uses the same principle as supervised learning, with the main difference being that labels only exist for parts of the data set. Semi-supervised learning is the preferable method for model building [11,31–33].

In unsupervised learning, no labelled data exists. The aim is to gain additional insights into the data set through clustering and association. It generally uses distance or similarity measures between data points to find patterns [11,31,33].

Reinforcement learning does not rely on existing data but rather learns from interaction with its environment. Through feedback from the environment the learning machine (or agent) receives an evaluation (reward or punishment) about a chosen action. Using a mix of exploiting existing knowledge and exploring the unknown the agent tries to maximize its reward or minimize the punishment, respectively. It is very suitable in situations where no learning data is available or updates occur in very short intervals [11,33–35].

2.2.1. Machine Learning Algorithms

Machine learning algorithms can be sorted into 12 groups with similar approaches and outputs. A summary of those and their respective typical learning method, learning approach, and application along exemplary algorithms are given in Table **1**. Some groups will be described in more detail in this chapter.

Regression Analysis

A model is fit to data to produce a model that can be used to predict future data. This model contains the relationships between input and output data that are statistically significant. These relationships do, in general, not point out any causation, though. There are three key sets of variables in regression learning: the dependent or response variables (y), the independent or predictor variables (x) and the model parameters that shall be estimated by the regression model. Most regression learning algorithms assume a known probability distribution behind the data set, for example a normal distribution. These algorithms are called parametric. Non-parametric methods construct the probability distribution from data instead, increasing the amount of data they require for proper learning.

The simplest kind of regression is Linear Regression, where a linear predictor function is fit to the supplied training data. A regression with p independent variables has the shape

$$y_i = \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i = x_i^T \beta + \varepsilon_i \tag{5}$$

With

- i = 1 … n
- β … Model parameter
- ε … Model bias

or alternatively in matrix representation

$$y = x \cdot \beta + \varepsilon_i\ . \tag{6}$$

Data fitting highly depends on the used error function. One example is the Least Squares Method

$$min_{c,\beta} Q(c,\beta) \text{ for } Q(c,\beta) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - c - \beta x_i)^2 \tag{7}$$

With

- Q … total error
- C … Linear offset
- β … linear slope

which can be used to fit the linear parameters α and β to the training data. [11,31,33]

A more detailed regression methodology is the Polynomial Regression, which models the dependency between the dependent and independent variables as a m-th degree polynomial,

$$y_i = a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_m x_i^m + \varepsilon_i \tag{8}$$

With

- i = 1 … n
- a … Model parameter
- ε … Model bias

Higher order polynomial regression, while still being a linear regression problem, is better suited to fit non-linear relationships between predictor and response variables. Higher order polynomials tend to overfitting, i.e. the model works very well on the training data set but does not perform well with other data sets. Caution is needed when working with fourth degree or higher order polynomials [11,31].

One type of non-parametric regression is the Gaussian process regression (GPR). IN GPR, the response variable is developed from latent variables from a Gaussian process (GP) and an explicit basis function h. A GP is, at this, a set of random variables in which any finite number of elements have a joint Gaussian distribution. The make-up of a GP, consisting of a mean function and a covariance function, impacts the performance of a GPR, and must be chosen accordingly. Examples of covariance functions are the Matern 5/2 or Rational quadratic, described in detail with other examples in the *Matlab* documentation in the chapter “Kernel (Covariance) Function Options” [36].

Support Vector Machines (SVM)

Support vector machines are a tool used for classification. SVM look for the optimal hyperplane to divide two classes from each other. This plane is situated in a way that maximizes its distance to the closest points of the two classes. These closest points are the Support Vectors [11,35].

Decision Trees

Decision trees are tree-like graphs with two types of nodes: leaves indicating the class or region defined by the response variable and decision nodes specifying a test on a single attribute. Decision trees are non-parametric models that can be used for both regression and classification, and are therefore divided into these two categories. Their main advantage over other machine learning approaches is the simple interpretability. Following the graphical decision tree from top to bottom makes decisions traceable and allows the formation of rules from the trees [11,32,35].

Naive Bayes Method

Several machine learning techniques are based on the *Bayes Theorem*

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{9}$$

With

P(A) … Prior Probability (Probability of an event before some evidence is considered)

P(A|B) … Posterior Probability (Probability of event A happening given event B)

P(B) … Marginal Likelihood

P(B|A) … Likelihood,

with the Naive Bayes method being the simplest of these techniques. As with all Bayesian methods the machine does not learn iteratively but uses inference from distributions of variables. It requires that all variables involved are fully independent, highly limiting its fields of application [11].

Artificial Neural Networks (ANN)

Artificial neural networks are based on the functionality of biological neural networks, and mostly on the working principle of the human nervous system. The smallest unit of a nervous system, both artificial and biological, is a neuron. Neurons can process and transmit information through electrochemical signals. Electrical signals are used for the transfer of continuous information while chemical signals need to breach a certain threshold to enable transfer. A neural network is then constructed of several neurons.

This concept can be transferred into a learning machine. The simplest ANN is the Perceptron, depicted in Figure **5**. The perceptron takes several inputs x_n that are multiplied with a neuron-specific weight w_n to produce a binary output. A very similar approach are sigmoid neurons. Using a sigmoid function

$$\frac{1}{1+e^{-t}} \tag{10}$$

instead of a step function enables a continuous output. There are also other functions that can be applied to produce continuous outputs.

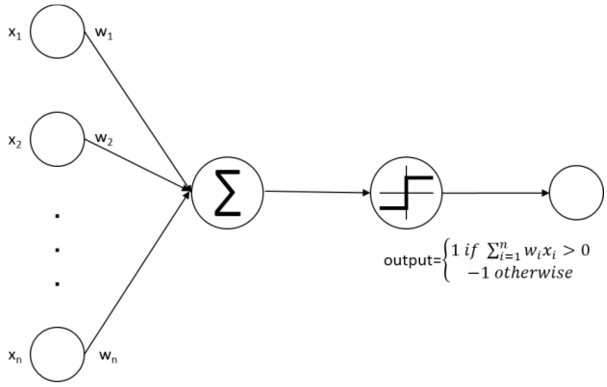

*Figure **5** Perceptron ANN Scheme [11,31]*

Single perceptron or sigmoid neurons can only solve linear problems; however, a network is able to solve nonlinear problems as well. Figure **6** illustrates a Feedforward Neural Network with Back Propagation. It is called feedforward as inputs only move in one direction, without any loops. Back propagation is a supervised learning concept for neural networks. Errors are found according to the desired output and then propagated back into the previous layers. It generally works on a gradient descent principle, so the neuron function (activation function) should be differential. The number of input and output neurons is predefined by the application of the neural network, however finding an optimal number of hidden layers as well as neurons in those hidden layers is rather difficult. It is important to meet a good fit between the task at hand and the hidden neurons, because too many neurons will lead to overfitting, while too few might make the algorithm too general and not applicable to the problem at hand. The main advantage of ANN is its ability to work with any type of data. Moreover, it is very suited to solve nonlinear problems and is highly scalable. By increasing the number of layers and neurons it is possible to achieve a high level of abstraction on billions of data points. On the other hand ANNs should not be applied to problems that have a linear underlying structure, since ANNs tend to overfitting for linear systems [11,31–33,35].

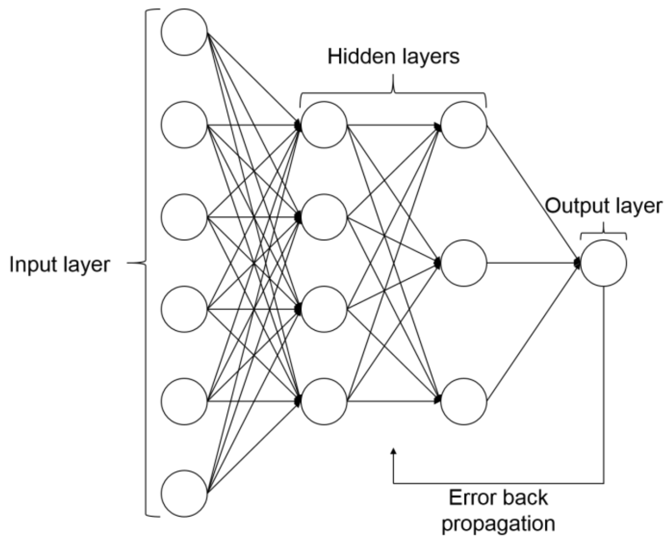

*Figure **6** Feed-forward back propagation neural network [11,31]*

*Table **1** Groups of Machine Learning algorithms [11]*

Group	Learning Method	Approach	Application	Examples
Regression Analysis	Supervised, Reinforcement	Relationship between dependent and independent variable is estimated by probabilistic method or error function minimization		Linear Regression, Polynomial Regression, Least Squares Regression
Distance Based Algorithms	Unsupervised, Supervised	Use of distance between features	Classification	k-Nearest Neighbour method
Regularization Algorithms	Supervised, Reinforcement	Extension of Regression analysis that introduces penalty term to balance complexity and precision		
Tree-based algorithms	Supervised	Use of sequential conditional rules	Decision making and classification	C4.5/C5, CART, Random Forest
Bayesian Algorithms		Use inference from distributions of variable	Classification and inference testing	Naïve Bayes
Clustering	Unsupervised	Maximize intracluster similarities and minimize intercluster similarities	Marketing, Demographic studies	k-Means
Association rule mining		Relationship among variables is quantified for predictive and exploratory objectives		Apriori, Eclat
Artificial Neural Networks (ANN)	Unsupervised, Supervised	Inspired by biological neural networks	Learning nonlinear relationships	Perceptron, Back-Propagation, Radial Basis Function Network

Group	Learning Method	Approach	Application	Examples
Deep Learning	Unsupervised, Supervised	Multiple hidden layer neural nets	Extraction of higher level of information from huge datasets	
Dimensionality reduction	Supervised	Use of various transformations and supervised learning approaches	Amplify signal in data prior to modelling	
Ensemble Learning	Supervised, Unsupervised, Reinforcement	Combination of multiple machine learning algorithms	Requiring results superior to single algorithms	Boosting, Bagging, Stacking
Text Mining	Supervised		Create insight from unstructured textual data	

2.2.2. Reinforcement Learning

Reinforcement Learning (RL) is a machine learning methodology in which the learning agent improves by interacting with its environment. Environment, in this specific case, describes any component that the learning agent connects with, and whose behaviour it cannot directly control. A RL agent interacts with the environment by performing an action according to a predefined policy. The agent receives feedback from the environment with a so-called reward that depends on an unalterable reward function. The learner aims at maximizing its total reward, and thus improves through discovering which actions yield the highest reward by trying them. Figure **7** depicts the interaction between the RL agent and the environment. One realises that there is usually a time delay between the action taken and the reward, as indicated by the time index t+1 on the reward that follows the action [37]. The total reward is calculated according to a value function, which gives the total amount of reward that can be expected in the long-run. Reinforcement learning is generally a trade-off between exploration and exploitation. Exploitation, or performing a so called greedy action, means using existing knowledge of the value of actions and choosing the action with the highest reward according to the current value function. Exploration on the other hand means using an action that is currently considered sub-optimal to improve the value function and the value-estimate of the non-greedy action. Properly balancing exploration and exploitation is a key challenge in RL. In contrast to other machine learning methods, a RL agent does not learn what a correct action is, but rather how correct its actions are [34]. One can separate the RL agent into four main sub elements:

1. Policy: The policy defines the agent's behaviour at a given time depending on the perceived state of the environment
2. Reward function: The reward function defines the goal of the learning problem and cannot be altered by the RL agent. It may be used to improve the policy.
3. Value function: The value function specifies what good behaviour of the agent is in the long run, it defines the total amount of reward that can be accumulated in the future.
4. Model of the environment (optional): A model of the environment may be used to predict the state of the environment after performing an action, and thus also the reward received for an action.

The agent needs to know about the state of its environment in order to properly choose an action. It should therefore receive immediate sensations such as sensory measurements. It may also get processed sensory measurements or even complex structures built from the original data. The learning agent does not need to be omniscient towards its environment, it only needs as much information as is necessary for the learning process. Many RL problems are modelled as so-called Markov decision processes (MDP). A MDP has the Markov property, i.e. its state signal contains all relevant information to make an educated decision. Therefore "The best policy for choosing actions as a function of a Markov state is just as good as the best policy for choosing actions as a function of complete histories" [38]. Even if a state does not have the Markov property one should try to approximate it as such, since any state should be a good basis to predict future rewards and actions.

RL methods can be divided into three main categories: Dynamic Programming Methods (DP), Monte-Carlo Methods (MC) and Temporal Difference Methods (TD). In classical DP so called full backups are performed on each state. This means that the value of one state is updated according to every possible successor state and its likelihood. More effective methods do not perform full backups but rather rely on partial backups. Since DP uses the value of successor states, the method requires a precise model of the environment. MC does not require a model of the environment. It updates state values by averaging a number of returns received in the state. Learning is performed according to episodes, a series of states and actions. TD combines DP and MC. TD methods learn on a step by step basis like DP, but do not require a model of the environment. Value updates are performed based on the observed reward and an estimate of the value of the next state [34,38].

In simple cases policies can be stored in state-action tables, directly correlating an action to a state. For more complex learning problems, for example when state variables have continuous values, this is not practical. Function approximation is required in these cases. The policy is translated into a function with state variables as parameters. Supervised machine learning tools have proven to be practical to achieve function approximation, with the two most used tools being linear gradient descent and multilayer ANNs with backpropagation [38].

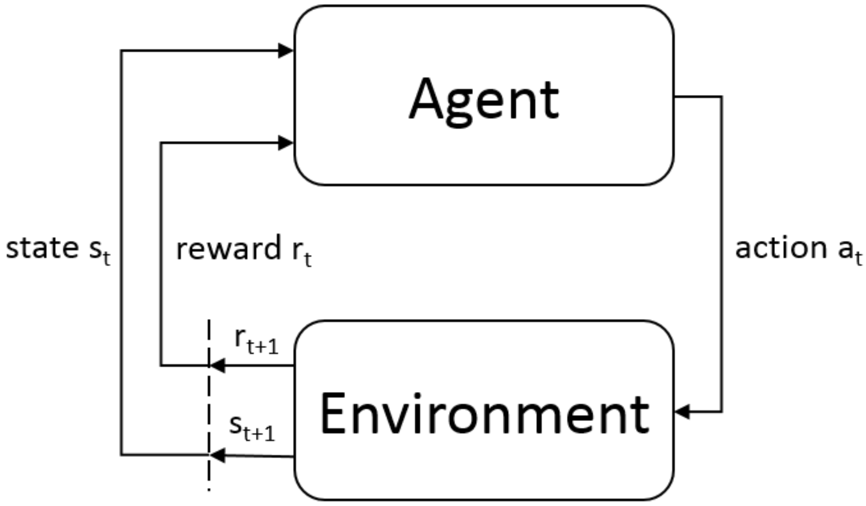

*Figure **7** Interaction between RL agent and Environment [39]*

3. Literature Review

In this chapter previous relevant activities in the field of artificial intelligence in building energy systems are presented. In more detail, the presented topics are HVAC control, user behaviour, and commercially available products in the field. Where available, reviews are presented first, followed by selected papers in order of publishing.

3.1. Advanced HVAC Control

3.1.1. Recent developments in advanced HVAC Control

Dounis and *Caraiscos* [40] present classic control systems for buildings and show the potential of intelligent control systems with a focus on agent-based intelligent control systems. Thermostats were originally used for feedback control of temperature. To avoid constant on and off switching of the thermostat a dead zone was applied. This type of control is called bang-bang control with dead zone. To avoid over-conditioning that is common with bang-bang control Proportional-Integrate-Derivative (PID) controllers were introduced. They are now the most common type of controllers in use, albeit requiring proper design to avoid instability. The focus in research has been on optimum, predictive and adaptive control. All three methods require models, which are usually nonlinear. Due to implementation issues, industrial development has not followed scientific studies. Predictive control allows the integration of estimated future events, enabling proper utilization of active building components, storage systems and night cooling. Adaptive controllers are able to self-change their operational parameters according to climate conditions and building types. Since the 1990s artificial intelligence methods have found application in control techniques. Neural Networks with their ability to solve nonlinear problems have proven to be especially useful. Among other things, it enables the real-time use of non-simplified PMV calculations for thermal comfort evaluation. Fuzzy control is another promising development for building management systems and has shown a large potential. Using genetic algorithms, fuzzy controllers can be optimized during operation. Intelligent systems also enable the use of model-free controllers. Hybrid controllers, combining for example neural networks and fuzzy controllers have seen development as well.

Sun and *Huang* [41] review the developments in HVAC control and demand management between 2011 and 2016. They find that for HVAC control development, the main goals have been an improvement of the robustness of control and of the efficiency of system level real-time optimization. For demand management, the trend was in coordinated control of building-groups rather than a focus on individual buildings.

The idea behind more robust HVAC control is to improve the controller performance when facing uncertainty. Three main types of uncertainty are pointed out: model-inherent uncertainty, process-inherent uncertainty (for example sensor bias) and scenario forecast uncertainty (for example weather forecast). In most cases, uncertainties are integrated into the design by using upper and lower bounds for the respective parameters. Other ways to integrate uncertainty include normal distributions and stochastic models. Uncertainties can also be integrated into model-based predictive control approaches. Real-time optimization consists of fitting the set points of the HVAC system to prevalent and predicted conditions that may affect the system’s operation. It can be prediction-based or non-prediction based, with both requiring models to evaluate the impact of set point changes. Especially model-based predictive control (MPC) has gained popularity and has shown effectiveness in improving indoor climate and energy efficiency. The main difficulties with MPC lie within uncertainties and the need for building or system models. Furthermore, there is a need to improve the computational efficiency, especially since an increasing number of components with growing complexity is integrated into the HVAC control.

3.1.2. Advanced comfort control

Liang and *Du* [42] developed an intelligent comfort control system (ICCS) that combines human learning and minimum power control strategies. Learning is used to adjust the PMV index to specific user preferences. A direct neural network controller is designed to solve the nonlinearities of the PMV model. The energy consumption is reduced by controlling the air volume fed into the room (VAV method). At the start of the learning process *Fanger's* PMV model is used which is then adapted to fit user inputs. Figure **8** shows the block diagram of *Liang* and *Du*'s concept. The NN controller for the ICCS is based on a back-propagation algorithm. The user input into the ICCS is limited to three values - colder, neutral, and warmer – that are used to adjust the setting of the PMV model. The ICCS is able to create a comfortable indoor thermal environment through learning and the NN controller, while the VAV control enables energy savings.

The use of PMV as a control index for thermal comfort is problematic, as it requires a set of sensors that are not standard to gather: the mean radiant temperature, the air velocity, which is a highly localized parameter that cannot be estimated precisely in most cases, and the human factors clothing and metabolic have to be considered constant. *Liang* and *Du* concluded that improved and cheaper sensor technology will enable an actual application of their intelligent comfort control system.

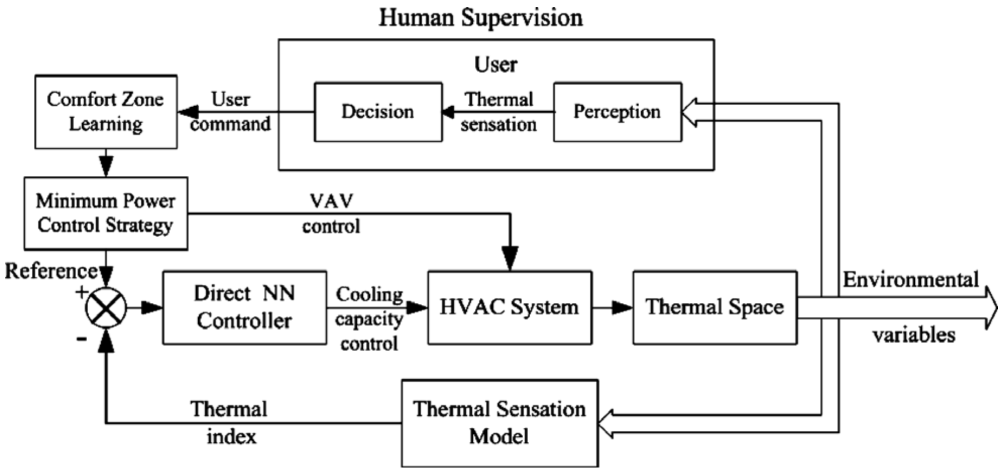

*Figure **8** Block Diagram Intelligent Comfort Control System [42]*

Data mining, especially decision trees, is used by *Gao* and *Menzel* [43] to determine comfort constraints and the influence of external conditions. Data collection, and with it large sets of data, are becoming more and more available as prices of sensors decrease. The use of data mining on these data sets might enable the prediction of the energy demand of a building. Data mining helps to discover patterns and correlations within large datasets. *Gao* and *Menzel* apply data mining to two rooms after cleansing the measured data from faulty data and obvious noise. The indoor comfort is analysed based on internal and external environmental conditions. It can be observed that data mining, with sufficient input data, is able to correctly correlate the environmental conditions to the thermal comfort, and thus evaluate the energy requirements under different conditions. The results from data mining in one specific building may also be applicable to optimize similar existing buildings.

Smart thermostats from companies such as Nest Labs and Honeywell Inc. are among the few commercial products bringing artificial intelligence into the control of HVAC systems. However, their specific learning approaches are not publicly available. *Barrett* and *Linder* [44] propose an open learning architecture based on Bayesian Learning to predict room occupancy and Reinforcement Learning (Q-Learning) for a thermostat control policy. The control policy aims at providing a comfortable environment to the user at the lowest possible cost. Q-Learning is chosen as a reinforcement learning method since it does not require a full model of the environment and is suitable to handle uncertainties. It has furthermore shown a broad applicability. One issue is that Q-Learning often requires a long learning period in a set environment to achieve good performance. The Bayesian learning approach combines Bayesian inference with agent learning which allows for the estimation of the likelihood of occupancy at certain times based on previous occupancy patterns. An occupancy sensor is used to gather the occupancy data. The state space of *Barrett* and *Linder*'s Q-learning agent consists of the room temperature, the time to occupancy and the outside temperature. Its action space is turning on and off the heating and cooling, respectively. A reward system depending on the state space and the action taken is developed, in which not using energy (i.e. neither heating nor cooling) either due to the room being at set point or not being occupied is neutral. Energy use is punished. The value of the punishment depends on whether using energy was correct based on the relation between room temperature and set point temperature. The implemented occupancy prediction shows a good performance after approximately 40 days of learning during system operation, and is able to recover from changes in the occupancy pattern in a similar timeframe. *Barrett* and *Linder* note that learning time could most likely be drastically reduced if days were grouped, for example into workdays and weekends. In their study, grouping is not used. Using an offline learning approach, the proposed control scheme is able to achieve a good indoor comfort after a few trials. It is furthermore able to outperform an "Always-on"-thermostat as well as a programmable thermostat in the used testing scenario. The authors note that a well set programmable thermostat may perform just as well as their smart thermostat.

In an attempt to get closer to the commercialisation of model based predictive control (MPC) *Ruano* et al. [45] installed a control solution named Intelligent MBPC (IMBPC) in a building at the University of Algarve in Faro, Portugal. The system requires weather data and forecasts that are provided by an energy autonomous weather station. It measures and forecasts the air temperature, air relative humidity and global solar radiation. Inside the controlled rooms a set of sensors measures the air temperature and relative humidity, movement, the state of windows/doors, wall temperatures and light. The IMBPC requires an existing building management system that allows bidirectional communication. The predictive models used are Radial Basis Function Neural Networks (RBF NN) designed by a Multi-Objective Genetic Algorithm. Occupation is schedule-based. The PMV is used as a thermal comfort index, with fixed clothing, metabolic rate and air velocity. The system operation is optimized according to a cost function based on the electricity contract of the university. Experiments performed during the exam period at the University of Algarve suggest that IMBPC can achieve energy savings compared to scheduled regular control schemes, while guaranteeing thermal comfort in times of occupation. Potential savings depend on the weather conditions, building characteristics and occupancy patterns.

3.2. Load Prediction

Zhao and *Magoulès* [46] review recent methods of predicting the energy consumption in buildings, including their main advantages and disadvantages. Because of the variety of influencing factors on a building's energetic performance, among them non-deterministic features such as the weather and the user behaviour, precise predictions are hard to obtain.

The first category of tools described are "engineering methods" which use physical principles for forecasting. A large variety of software tools is available on the market, such as *DOE-2*, *EnergyPlus* and *ESP-r*. These tools have proven to be useful for precise predictions, however they require detailed input data to achieve this. Some of this input data may be difficult to collect, as for example the thermal behaviour of building components over the whole range of indoor and outdoor conditions they go through during operation. The level of detail applied in models can vary depending on the task at hand.

The second category are statistical regression models. From historical data, a function connecting energy consumption or other impacting performance indices to influencing factors is generated. This method does require the collection of historical data before application. Its main usages so far have been the prediction of energy usage according to certain parameters, the prediction of energetic performance indices and the estimation of influential parameters on a building's energetic performance such as a total heat loss coefficient. *Zhao* and *Magoulès* also present two of the most used artificial intelligence tools used, Neural Networks and Support vector machines. Neural networks are good for solving nonlinear problems, making them useful for an application within buildings. They are used for prediction purposes, ranging from electricity production of renewable energy sources to hourly load prediction, control and operation optimization and estimation of usage parameters. Neural networks require the collection of training data prior to their application. Support vector machines, too, are useful for solving nonlinear problems. Their main application so far has been in load prediction, so far most of them have been rather specific, so that a general assessment of their performance cannot be given. Table **2** gives a comparative analysis of the previously presented tools. Depending on the application, one tool or another may be useful. Artificial intelligence based methods may find more utilization in the future, as the field is developing quickly.

*Table **2** Comparison of commonly used methods for load prediction, taken from [46]*

Method	Model Complexity	Easy to use	Speed	Inputs	Accuracy
Elaborate engineering	Fairly high	No	Low	Detailed	Fairly high
Simplified engineering	High	Yes	High	Simplified	High
Statistical	Fair	Yes	Fairly high	Historical data	Fair
ANNs	High	No	High	Historical data	High
SVMs	Fairly high	No	Low	Historical data	Fairly high

Edwards, *New* and *Parker* [47] evaluate seven different machine learning algorithms applied to hourly electricity consumption prediction in a residential building. Energy modelling can generally be divided into two general types: forward modelling, in which input data is run through engineering models to calculate the desired output, and inverse modelling, where input, output, and a general mathematical relationship between them is known and a statistical method is used to optimize the model parameters. Sensor-based modelling methods can be seen as a hybrid between the two general methods. The sensor data provides a model for the entire building, which can be seen as the forward part, while machine learning helps to improve the parameters of the underlying engineering model for the sensor data. The machine learning algorithms compared in the paper are Linear Regression, Feed Forward Neural Network (FFNN), Support Vector Regression (SVR), Least Squares Support Vector Machine (LS-SVM), Hierarchical Mixture of Experts (HME), Fuzzy C-Means (FCM) with FFNN and Temporal dependencies. Linear regression, being the simplest tool, serves as a baseline benchmark. The methods are applied to three residential buildings equipped with a sensor network. Depending on the building, the methods show varying performance, as the three buildings exhibit rather different consumption patterns. While FFNN has shown to perform best for commercial buildings, its performance on residential buildings is not as good. For one building, it performed only slightly better than linear regression. The more advanced methods LS-SVM, HME and FCM with FFNN show a good performance for all analysed cases, with LS-SVM being the statistically best technique for predicting residential electricity consumption over the next hour.

3.3. User Behaviour

It has been proven that neural networks are a good tool to approximate nonlinear relationships between input and output. *Atthajariyakul* and *Leephakpreeda* [48] thus suggest the use of Neural Networks for a computationally more efficient and more precise calculation of the PMV index. A multi-layer feedforward network is fed with the wet bulb temperature to estimate the relative humidity, the clothing value, the metabolic rate, the air temperature, the globe temperature to estimate the mean radiant temperature and the air velocity. The neural network is trained to deliver results equalling the static *Fanger* PMV model. *Atthajariyakul* and *Leephakpreeda* show that the Neural Network is capable of producing real time PMV calculations that can be used for HVAC control in good agreement with *Fanger's* PMV model.

Liu, *Lian* and *Zhao* [49] use a back propagation neural network to create an evaluation model for individual thermal comfort (Neural Network Evaluation Model, NNEM). In contrast to other works, the basis for their thermal comfort model is not the PMV, but rather a model based on air temperature, air velocity and humidity, and mean radiant temperature as inputs, with outputs between 0 and 1. 0 indicates a cold sensation, 1 warm, and 0,5 a neutral thermal sensation. Figure **9** shows a scheme of the principle neural network used. Utilizing the NNEM, a neural network controller (NNC) for a room air conditioner is presented. Based on the input parameters of the NNEM the compressor frequency, the rotational speed of the fan and the angles of the vanes are adjusted to achieve the desired thermal conditions. The combination of NNEM and NNC is able to create thermal conditions close to what is required from the user. The authors point out that there are improvements that need to be performed: the most energy efficient combinations of the environmental factors that are considered comfortable need to be determined, as well as the most efficient control of the air conditioner.

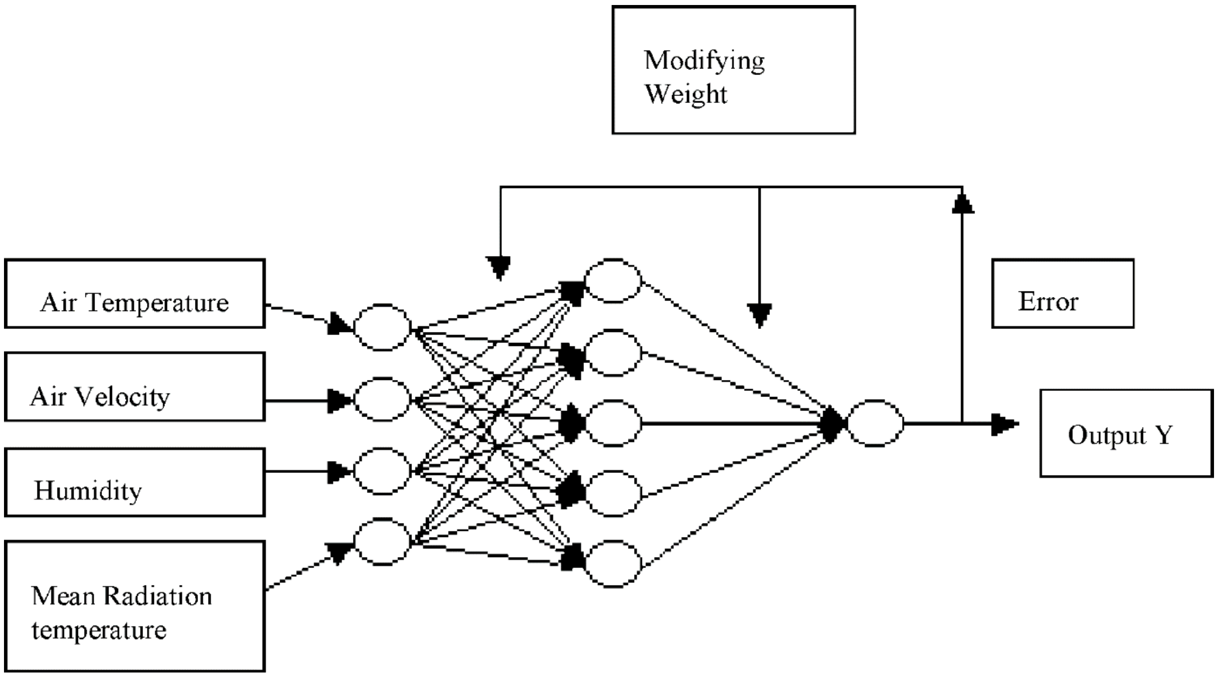

*Figure **9** Scheme of Neural network for individual comfort evaluation [49]*

Precise prediction of occupant numbers can be used for a more efficient control of building equipment. *Mamidi*, *Chang* and *Maheswaran* [50] use multiple statistical learning methods like linear regression, logistic regression, multi-layer perceptron (a neural network method) and support vector machines to build an estimation and a prediction model. Data from a self-made sensor, incorporating sound, wide-field motion detection, narrow-field motion detection, ambient light, temperature, humidity, carbon dioxide and door state in combination with ground truth data, is used for training. Through experiments, it is shown that different machine learning tools can perform occupancy estimation with simple sensors. Multilayer Perceptron shows the best performance for this task. Furthermore, it is possible to perform occupancy prediction using similar machine learning tools with decent accuracy.

Buratti, *Vergoni* and *Palladino* [51] use artificial neural networks to fit PMV calculation with less inputs that are easier to gather than the original data. The use of ANN makes it possible to link the PMV to both the indoor and outdoor environment. Training data for the ANN was collected during eleven experiments at the University of Perugia, Italy. The data acquired were indoor air temperature, air velocity, globe thermometer temperature, air pressure and air relative humidity, outdoor air temperature and relative humidity and thermal comfort votes from users based on the PMV scale as well as users' positions in the room during the experiments. The ANN is programmed using *Matlab* and is a two-layer feedforward NN. It is trained using nine input parameters: clothing, gender, metabolism, age, position (X-Y), indoor and outdoor air temperature. The number of neurons used in the hidden layer was varied to find the optimum. The trained ANN showed good fit to the questionnaires used as validation data and was thus tested on other, unrelated case studies. It proved to correlate better with thermal votes than the *Fanger* static model approach

3.4. Classification of Previous Work

In Table **3** a collection of papers on artificial intelligence in building energy systems is classified. Bold lines mark the end of a group of papers. The first group contains reviews, followed by papers on neural networks, decision tree models, reinforcement learning, individually created methods, comparative studies and mixed applications, and finally a paper on user interaction. The main findings of the classification are that the integration of advanced control systems into existing HVAC systems is feasible [52], with all advanced control systems outperforming classic controllers, concerning both thermal comfort and energy efficiency [34,40,45,53–55]. However, in some cases it proved to be difficult to properly estimate the real performance of classic controllers, since only few of them are configured correctly [44,54]. It may well be that less advanced methods are more effective than the most recent tools, especially in cases where correlations are simple [56]. While the potential of intelligent controllers has been proven to be high in a wide scale of papers, the commercial employment is still very low [40,57,58]. It can be observed that the implementation of the control systems is difficult: a proper implementation requires a high level of expertise and the demand on computational power for real time application is still high, and some of the methods used require extensive sensor networks [40,41,45–47,50,55,59].

The learning process is another issue hindering the spread, as these learning periods are still fairly long, making large sets of learning data a necessity. Complete real-time learning is not yet an option, as the impact on the system performance would be too high [34,37,45,46,50,51,55,57,59]. For a successful implementation it would be necessary to create user acceptance, i.e. the user needs to be able to understand how the system is working, with the interaction being as attractive as possible [60]. Neural networks are the dominant method applied, either by themselves or in conjunction with other tools. The use of neural networks is mostly due to their strength in solving nonlinear issues, which are prevalent in building energy systems.

*Table **3** Classification of Papers on Artificial Intelligence in Building Energy Systems*

Ref.	Title	Issue tackled	Method used	Key findings
[41]	Recent Developments in HVAC System Control and Building Demand Management	Review of development in HVAC control and Demand Management from 2011-2016	[1]	Improvement of robustness of control and efficiency of real-time optimization Coordinated control in building-group-level Increased system complexity keeps control improvements challenging Trend towards MPC that proved to be quite effective; can be auto-regressive moving-average model, resistance-capacitance mode, or NN model Genetic algorithms and Game theory for multi-building coordination
[46]	A review on the prediction of building energy consumption	Review of work related to modelling and predicting building energy consumption	[1]	ANNs most used AI tool for energy prediction SVMs find increasing use, needing less training data Both require data for learning and expertise in use to create valid results
[40]	Advanced control systems engineering for energy and comfort management in a building environment – A review	Review of control systems for energy management and comfort in buildings, in particular multi-agent systems	[1]	Industrial Development has not followed research due to implementation issues Adaptive Fuzzy controllers are seen as the most promising adaptive controllers for buildings Main issues: Need for a building model, nonlinearities, required computational power for real time parameter estimation All advanced control systems outperform classic controllers
[42,61]	Thermal Comfort Control Based on Neural Network for HVAC Application	PMV-based comfort zone learning, optimization of system operation	Back-propagation direct **Neural Network** controller for HVAC	Clothing values and activity level for PMV cannot be measured and have to be set to a constant
[45,54]	Energy Savings in HVAC Systems Using Discrete Model-Based Predictive Control	Integration of a MPC into an existing HVAC system to guarantee thermal comfort and reduce energy consumption	Radial Basis Function **Artificial Neural Networks** plus cost function and optimization function, inverse model for MPC	MPC achieves significant energy savings Traditional controllers also suffer from bad settings, sensor misplacement and other issues that improve the relative performance of experimental controllers

[49]	A neural network evaluation model for individual thermal comfort	Prediction of individual thermal comfort using NN	Back propagation **neural network**	Optimal combination of different environmental factors for thermal comfort not yet estimated
[48]	Neural computing thermal comfort index for HVAC systems	Real-time estimation of thermal comfort through the PMV index	Feed forward **Neural Network Models** using back-propagation	Good agreement between NNM and Fanger's PMV model Using a lot of input data, some which are difficult to measure in real life conditions, such as clothing
[57]	Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems	Review on ANN-MPC and development of an ANN-MPC for a residential building	**Artificial Neural Network** Model Based Predictive Control Best Network after multiple Iterations (BNMI) method	ANN-MPC have been researched for several building types, but are not in commercial application yet Issue: MPC require rich data sets, containing all potential working conditions One has to carefully consider the control objective, e.g. minimize energy consumption against minimize operating costs
[51]	Thermal Comfort Evaluation Within Non-residential Environments	Prediction of thermal comfort using an adaptive comfort model	Two-layers feedforward **Artificial Neural Network**	The ANN, after training, gave results closer to actual comfort votes than the static Fanger comfort prediction
[52]	Neural networks based predictive control for thermal comfort and energy savings in public buildings	Integration of a MPC into an existing HVAC system to guarantee thermal comfort and reduce energy consumption	Radial Basis Function **Artificial Neural Networks** plus cost function and optimization function, inverse model for MPC	Good integration into existing system for both summer and winter (Portugal) possible
[45]	PVM-based intelligent predictive control of HVAC systems	Implementation of a commercializable MPC in an existing HVAC system	Radial Basis Function **Neural Networks**	Energy savings up to 50 % over schedule based control expected Extensive sensor network required: environmental condition, several indoor sensors for thermal conditions and occupation needed Requirement of historic data for learning
[43]	Using data mining in optimisation of building energy consumption and thermal comfort	Prediction of room temperature through data mining	Data Mining using Knowledge Discovery in Databases (KDD), C4.5 **decision tree**	External weather conditions are necessary to keep the error sufficiently low Data mining in one specific building may serve as a basis for similar buildings

	management		learner	
[58]	Sustainability Through Innovation in Product Life Cycle Design	Energy saving potential of a smart, occupancy learning thermostat in an office (only occupancy prediction, no learning of HVAC system behaviour)	**Decision Tree**	Only two really smart thermostats on the market: Google's *Nest* and *Heat Genius* [62,63] (authors did not know about Honeywell *evohome* [64] and *ecobee3* [65] and did not consider *tado* "smart") Potential savings highly influenced by climate zone and building type
[59]	Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools	Evaluate effect of eight structural building parameters on the heating and cooling load	**Decision Trees**, iteratively reweighted least squares	
[37]	Online tuning of a supervisory fuzzy controller for low-energy building system using reinforcement learning	Tune a supervisory fuzzy rule based controller through reinforcement learning	**Reinforcement Learning**	Expert rule based control is still dominant in practice due to the complexity and time-demand of more recent technologies Fully real-time based reinforcement learning may take too long to converge A denser discretization increases learning space and time Division of inputs and outputs into several layers to improve computation time Long learning period may have a negative impact on performance
[44]	Autonomous HVAC Control, A Reinforcement Learning Approach	Learning environment for thermostats to improve indoor comfort and energy efficiency	**Bayesian Learning** for occupancy prediction and **Reinforcement Learning** (Q-Learning) for control policy	A learning thermostat may show no advantages over a carefully programmed programmable thermostat, in other cases however it might be beneficial for energy consumption and thermal comfort

[34]	Reinforcement learning for energy conservation and comfort in buildings	Development of a controller for thermal comfort, air quality and energy efficiency optimisation	**Reinforcement Learning**	Three main areas of activity in building controller development: NN, fuzzy systems and predictive control, as well as their combination Developed controllers always showed reduced energy consumption, even if they solely aimed at improving comfort Main benefit of RL: continuous learning, allowing adaptation to a changing environment, e.g. due to equipment deterioration Main issue: Need for exploratory action costly, both comfort- and energy-wise, i.e. a long offline learning period is needed for industry application
[53]	Energy-Efficient Building HVAC Control Using Hybrid System LBMPC	Model-based control for system identification and control optimization	Learning-based **model predictive control** (LBMPC) using two parallel models for optimization	Significant energy savings can be achieved over PID-control at comparable comfort levels High need for simplifications to keep computational time reasonable Modell creation and learning requires detailed knowledge of the system and as much input data as possible
[55]	Occupancy Based Demand Response HVAC Control Strategy	Occupancy based control strategy	**Markov Chain**	Necessity to gather ground truth data for learning is problematic Energy savings 10 % and more, depending on utilization and climate zone
[21]	Learning User Preferences to Maximise Occupant Comfort in Office Buildings	Predicting user comfort based on temperature and a distance measure	Occupant-specific database	Non-individual standard (PMV) is personalised by a learning approach

[47]	Predicting future hourly residential electrical consumption	Comparison of 7 Machine Learning techniques for hourly electricity consumption prediction in residential buildings	Linear Regression, Feed Forward NN, Support Vector Regression, Least Squares Support Vector Machine, Hierarchical Mixture of Experts, Fuzzy C-Means with Feed Forward NN	Issue with forward model based prediction: time and skill for model development rare/too expensive for broad scale application Sensor-based modelling requires existing sensor data (can also come from a similar building) “Machine learning allows asymptotic approximation to the ‘true’ model of the data” Costs for sensoring too high as of 2012 Least Squares SVM performs statistically best to predict electricity consumption in residential buildings
[50]	Improving building energy efficiency with a network of sensing, learning and prediction agents	Occupancy tracking and prediction for anticipatory HVAC control	Neural Network, Linear Regression, Gaussian processes, Support Vector Machine	Requirement of at least several weeks of learning data Very accurate occupancy prediction using off-the-shelf sensors Broad sensor network required Non-intrusive method for occupancy measurement and prediction
[56]	Predicting electricity energy consumption – A comparison of regression analysis, decision tree and neural networks	Electricity load prediction for grid stability	Regression Learning, Decision Trees, Neural Networks	In simple cases with linear correlations, simpler tools may prove to be more effective than more advance methods
[60]	Learning from a learning thermostat: lessons for intelligent systems for the home	Analyse user interaction with smart devices based on the Nest (First Version)	[1]	Users did not appreciate extreme changes to their schedules performed by the Nest Changes did not need to be too detailed, but should rather make overall sense Back-Up when key elements of “smartness” fail Exception Flagging and conveying how the “intelligence” of the system works could improve learning and user interaction

[1] No specific machine learning method use due to nature of article

3.5. Commercially Available Products

So far there has been little commercial application of the presented methods which aim to improve indoor comfort and energy efficiency. MeteoViva is a German company providing model-based predictive control services, that operates mostly in Germany as of now. BuildingIQ is a US-American company providing similar services based on black box models developed through machine learning tools. The most prominent application of machine learning in the building sector may be the *Nest Intelligent Thermostat* by Nest Labs, belonging to the Google parent company Alphabet Inc. A diverse range of smart thermostats have been introduced to the market, among them *Heat Genius*, *ecobee*, *tado*, and Honeywell's *evohome*. The next sections describe these products in more detail.

3.5.1. MeteoViva

MeteoViva provides model-based predictive control services. A modular model of the building or component is created, with the intent of finding the thermal influence of specific components, such as walls or machinery. MeteoViva requires a historic load curve for the initial setup. High-resolution regional weather forecasts are externally bought and fed into the control scheme in order to pre-emptively change settings. To find the optimal settings, a set of potential actions is run through the model, out of which the best performing ones concerning a cost function are chosen. By gathering further load curves during operation, the system can improve the parameters of individual components to get a better behavioural fit between model and reality. The keys to energy savings using this system are the prevention of over-conditioning and an optimized use of active building components. A more detailed description of the working principle is given in MeteoViva's patent EP 1 134 508 B1 [12,66]. In a set of projects the software has shown energy savings between 20 % and 40 % [16]

3.5.2. BuildingIQ

BuildingIQ is based in California, USA, and provides cloud-based optimization software for large and complex buildings. The tool called Predictive Energy Optimization (PEO) utilises data measured from different parts of the HVAC which are retrieved through communication with the existing building management system, weather forecasts and energy cost forecasts to optimize HVAC operation while maintaining indoor comfort. A building model is used for optimizing the processes. This model is created using supervised machine learning, incorporating a variety of data retrieved from the BMS. It is continuously improved during operation. BuildingIQ's main regions of work so far have been the USA and Australia, with a focus on complex, single-user buildings that have shown most benefit from PEO, ranging from 10 % to 20 % savings in energy costs, going up to 40 % for more extreme cases. As one of their latest acquisitions BuildingIQ is starting to incorporate group votes into their cloud service to guarantee thermal comfort while performing their efficiency measures. [13,67–70]

3.5.3. Learning Thermostats

The *Nest*, the *Heat Genius*, *ecobee*, Honeywell *evohome* and *tado* are all examples of programmable and self-learning thermostats mainly aimed at homes and partially at small businesses, with a growing diversity of products that came out into the market over the last years. The *Nest Learning Thermostat* has been on the US market since 25th October 2011, with a third generation of the thermostat being released 1st September 2015. The Nest is considered to have opened the field for other smart thermostats, not so much through its advanced technology, but rather through its design and simple user interface [60]. It is currently sold in the United States, Canada, the United Kingdom, Belgium, France, Ireland, the Netherlands, Italy and Spain [62]. A release for Germany and Austria is planned for late 2017 [71]. Using motion sensors and smartphone GPS, its machine learning algorithm learns the occupancy schedule and a preferred set point schedule. It uses its internet connection to collect weather data that is fed into a double layer system to optimize the HVAC operation. The interior layer is learning the house's inner dynamics, while the outer layer studies the house's reaction to the external environment, e.g. towards the weather conditions. This inverse modelling of a house should enable forecasting the house's behaviour to external conditions, allowing appropriate reactions by the HVAC to create the required thermal set point without the risk of over-conditioning [15]. The *Nest* does not have multi-zoning, i.e. every room that it controls must have the same temperature schedule. Depending on the size of the controlled house and the use, this may well lead to an increase in energy consumption compared to "non-intelligent" programmable thermostats. The other four presented thermostats are able to perform zoning, giving them a technological edge over the *Nest*. The *Heat Genius* however is not able to learn the thermal behaviour of houses. It uses weather data to adjust preheating times without taking the thermal characteristics of individual buildings into account [63]. Honeywell's *evohome* does not support schedule learning, as it relies on programmed occupation schedules [64]. The *ecobee3* and the *tado* both learn a buildings thermal behaviour as well as occupation schedules, and the building may be divided into several zones, making them the most advanced thermostats available on the market thus far. *Tado* uses the occupant's smartphone for location tracking, adjusting the temperature set points according to all users' distances towards the building [14,65].

It has to be noted that, for trade reasons, none of the OEMs have made the working principle of their learning process public, so that an independent evaluation of their functionality is difficult to perform. It may well be that energy savings are caused by the users' increased awareness of their energy consumption rather than the learning process of the thermostat [44].

4. Methodology

In this chapter, the methodology to develop a comfort temperature predictor using machine learning is explained. Model parameters will be chosen based on *Fanger's* PMV-model and its shortcomings as has been described in previous works using the PMV for comfort prediction. To balance these shortcomings and add some depth to the predictor, adaptive comfort approaches will be added onto the PMV. According to the chosen parameters and the desired output appropriate machine learning tools are determined along with fitting learning, testing and validation approaches. The simplifications that have been made are also summarised.

4.1. Choice of Model Parameters

The PMV is widely applied and accepted for designing HVAC-systems and evaluating thermal comfort. Its main parameters are the air temperature, mean radiant temperature, air velocity, relative humidity, occupant's clothing and occupant's activity level. Several authors have applied the PMV for thermal comfort prediction [19,45,48,57,61], all facing the same issue: only the air temperature and relative humidity are generally available. The mean radiant temperature and air velocity are difficult and expensive to measure, and the clothing and activity level are highly personal. *Liang* and *Du* and others use (seasonally) constant estimations for clothing and activity [20,48,49,61] and *Atthajariyakul* and *Leephakpreeda* replace the mean radiant temperature with the globe temperature, since it is easier to measure [48].

Another issue using the PMV is its restricted applicability due to the assumption made by *Fanger* during development. Humans are viewed as thermally passive towards the environment, i.e. the interaction between body and environment is only determined by the physics of heat and mass transfer, while in reality, the body adapts through its thermo-regulation system. It is furthermore only valid under steady state conditions [23]. As a consequence, the perceived warmth in warmer environments and cold in colder environments is overestimated [5]. While it has been proven to work decently well for large groups in moderate climates, it has not been developed for individual comfort prediction. Individual preferences may vary up to one PMV-scale unit (equalling 3 K), which led *Wyon* to state that "When individual variation is so large, […] the practical value of estimating group mean neutral temperatures is very limited. It is more important to provide the practical means for individuals to adjust their own heat loss, […]." [72]. For these reasons, the PMV's main parameters are only used as an orientation to develop the comfort temperature predictor.

Nicol and *Humphrey* showed that preferred temperatures largely vary and need to be adapted frequently. Adaptive thermal comfort theory suggests that one key influencing factor on the thermal comfort temperature is the outdoor air temperature. It gains a high importance due to its influence on the parameters that are used in the PMV, for example the clothing level and the metabolic rate [5]. It has been suggested that not only the outdoor temperature of the present day but also the running mean outside temperature for up until a week influence the thermal comfort. The running mean outside temperature for a day n can be calculated using

$$T_{RMT,n} = (1-\alpha) \cdot \sum_{i=0}^{j} \alpha^i T_{n-1-j} \qquad (11)$$

With

α … Discount factor

j … Number of days considered relevant

While not being highly influential, the best correlation between comfort temperatures and the running mean temperature has been found for α = 0,8, which is thus used [29]. The running mean temperature is assumed as a replacement for the clothing value and the metabolic rate used in the PMV. In order to keep predictions reasonable even under cold conditions *Holopainen* et al. suggest using a minimum running mean temperature [23]. This will be considered, with a first minimum set to 10 °C. Since the aim is to predict the comfort temperature, relative humidity measurements from within the room cannot be directly used. As indoor air temperatures usually lie within a limited range of temperatures it is fair to assume a relatively direct link to the outdoor vapour pressure if no humidification is used. In case of humidification, one may neglect the humidity as an influencing factor overall. In moderate climates, there may also be a direct link between outdoor and indoor relative humidity, which is therefore also considered as an alternative to the vapour pressure. Radiation data and air speed are difficult to implement since they are highly local parameters. Radiation data is furthermore rarely measured. As an approximation, the use of cloudiness, supplied by a nearby public weather station, as an indicator for the incoming irradiation and as an indicator for (perceived) likelihood of rain and the respective choice of clothes is suggested. Alternatively, daily sunshine hours are considered. The air velocity is not implemented. It is assumed that in general air speeds will not be high enough to cause a user discomfort. From intuition, the daily maximum and minimum temperature may also have an influence on the comfort temperature and are thus considered for testing.

Summarizing, the thermal comfort temperature shall be predicted using:

- the outdoor air temperature,
- the running mean outdoor air temperature,
- the daily minimum and maximum temperature,
- the outdoor vapour pressure or relative humidity
- and the cloud cover or sunshine hours.

In accordance with the adaptive comfort standard, the outdoor air temperature and the running mean temperature are taken as fixed parameters, with all others being seen as optional. All sensible combination, i.e. no two parameters shall describe the same concept, are tested.

4.2. Data Preparation

Machine learning tools' performance is highly dependent on the quality of the data fed into the learner. Hence, it is necessary to filter out noise and illogical data points as good as possible. Furthermore, a data point, consisting of all input and output parameters for a given point in time, has to be complete, i.e. all parameters need to have sensible values, to be considered for use in learning.

As a first step, all illogical parameter values are taken out. For this purpose, all temperature setpoints below 16 °C and above 28 °C are considered faulty values and are removed. As weather data from a service provider is used, it is assumed that the weather data has already been cleansed of noise and illogical data. Next, data points with missing parameters are removed from the total data set. Since the numerical values of the input and output parameters take on different orders of magnitude they are normalized as a final step.

4.3. Choice of Machine Learning Tools

Three to six input variables are chosen to predict the individual comfort temperature. The output value is a continuous variable. Depending on the availability of learning data, this means that two main groups of learning tools are applicable: Supervised Regression Learning and Reinforcement Regression Learning. *Matlab* will be used to develop the learners. It includes a regression learning and a neural network toolbox which will be used [36]. While standard adaptive thermal comfort theories imply a linear correlation between the comfort temperature and the outdoor air temperature and humidity [23], studies by *Nicol* and *Humphreys* suggest non-linearity in artificially climatized buildings for at least the outdoor temperature [5]. There is no clear indication of which regression tool might perform superior to others, so that all available standard tools will be tested. The unclear nature of the correlations suggests that artificial neural networks may be very useful.

Since the availability of training data before implementation is not always given, reinforcement learning, as a means of learning during system operation, needs to be considered as well.

4.4. Learning Approach, Testing, Validation

Due to the unclear behaviour of the comfort temperature in relation to the available weather parameters it is first necessary to find the optimal combinations of parameters for the tools. Furthermore, different settings within the machine learning tools must be optimized.

For regression learning, this includes among other things the error function. By using diverse standard tools from the *Matlab* Regression Learner app this variation can be covered.

For the neural networks, it is necessary to adjust the number of nodes and hidden layers. In a first approach, the number of nodes is varied from nine to sixty nodes in steps of three. If deemed necessary due to unsatisfying results, the number of hidden layers is increased in a second step.

Testing and validation is performed internally by the machine learning tools. The performance indicator to evaluate the tools is the root mean squared error (RMS) between the predictor and the target data. A tool is considered viable if it fulfils these two targets:

1. A baseline scenario is created. When the outdoor temperature is below 21 °C, the heating setpoint is 21 °C. If it is above 23 °C, the cooling setpoint is 23 °C. Between those two temperatures, the setpoint temperature is equal to the outdoor air temperature. The tool should perform better (i.e. have a lower RMS) than this baseline scenario.
2. There is usually not one fixed comfort temperature but rather a range of temperatures that are comfortable, which is usually considered 3 °C to 4 °C wide. A mean deviation of 2 °C from the setpoint temperature is therefore considered feasible. The RMS of a tool must be below 2 °C.

If a tool fulfils both requirements, the tools are further analysed. Should an extensive number of tools be viable, only the three best performing tools from each machine learning category would be considered. A key performance metric for machine learning tools is the required amount of training data for good performance. The behaviour of a tool over a range of dataset sizes is therefore analysed.

In newly constructed buildings and buildings without any long term measured data available, real-time learning, may be a viable method to implement a comfort temperature learner, using reinforcement learning. As supervised learning methods are, when possible, generally more reliable, reinforcement learning is only applied if at least one of the supervised methods fulfils the aforementioned conditions.

If applied, reinforcement learning is used based on artificial neural networks for function approximation of the value function. Function approximation is useful for the task at hand as it enables a continuous output and the processing of a wide range of continuous input. The alternative, look-up tables, are impractical mainly due to the wide range of potential input combinations and would require substantial simplifications [38]. Function improvement is based on gradient descent.

The ANN's parameters are based on the results from the supervised learning approach, as these are considered to be optimized to a certain degree, once Reinforcement Learning is applied. In a first step, the error function is chosen between half of the MSE and the RMS. All tests are run over varying learning rates from 0.005 to 0.05 in increments of 0.005.

The chosen method has two improvements tested on it. First is a discounted learning rate in order to reduce potential overfitting and reusing samples multiple times to speed up training. The discount rate and the number of times a data point is used are varied. Due to limitations in the time available to develop this thesis, as well as the need for a building to perform real training in operation, the data already used for supervised training is again used for a kind of “artificial real-time training”. This means that the gathered data is fed into the learning agent as a time series, with the improvement of the learning agent being controlled after every “timestep”. 365 data points, i.e. data points representing a year, are used as training data, the remainder is used as the testing data set. The target for Reinforcement Learning is to achieve an RMS below 2 °C within a half a year, with small overfitting, so that learning can continue after the optimum is reached without too drastic increases, minimizing the required level of supervision.

4.5. Energy Evaluation

Building operation requires a balance between user comfort and building operational cost. While the focus of this paper is to investigate the general feasibility of the comfort temperature learning based on environmental factors, a shift into application will require at most a minimal increase in energy consumption compared to standard control. Building models can help to develop an idea of the energy consumption under varying circumstances with relatively low effort. A single user office in a modern office building is used to evaluate the energy performance of the best performing supervised and reinforcement learning tools, compared to a baseline scenario of 21 °C heating setpoint and 23 °C cooling setpoint. The orientation of the office is changed to cover the four main cardinal directions. A weakness of this model is that it does not reflect potential energy savings and more efficient operation of active building components used for room air conditioning. Developing a model able to reflect these behaviours would exceed the frame of this work, though.

The office model is constructed using the 3D modelling software *SketchUp* from Trimble Navigation Limited, extended with the *OpenStudio-PlugIn* from the Alliance for Sustainable Energy, LLC [73,74]. *EnergyPlus* from the United States Department of Energy is used for the actual energetic analysis [75].

5. Program Development

Using supervised and reinforcement learning in *Matlab*, a tool to predict the comfort temperature of individual users in office spaces, dependant on weather conditions, shall be developed. Data from a German office building is used to train and test the tools. The setpoint predictions provided by the tools are used in a building simulation of a simple office model to evaluate the energy impact of the developed tools.

5.1. Used Data

All temperature setpoints used have been collected at the central building of the savings bank *Kreissparkasse Göppingen* in Germany (48,7 ° N, 9,6 ° E). The building went into use in the fall of 2011, and hourly temperature setpoints are available from 01.04.2012 to 02.12.2014, summing up to a total of 976 days. The data derives from 35 rooms with individual temperature control spread over the first to fourth floor of the building, however no data was available for the ground floor. Table **4** shows the number of rooms available per floor.

*Table **4** Available Rooms per floor*

Floor	**Number of Rooms available**
1	5
2	14
3	2
4	14

Weather data collected directly on site is small and only includes the outdoor temperatures and relative humidity. As it does not contain all data deemed necessary, the use of onsite data is discarded. The German weather service (DWD) provides historical weather data from weather stations all around Germany through the Climate Data Centre (CDC). The closest station to Göppingen that provides the required data is located at Stuttgart-Echterdingen (48,7 ° N, 9,2 ° E DWD-station ID 4931). [76]

Since complete weather data sets close enough to the building are only available as daily averages the temperature setpoints are averaged over a day. A day is only considered suitable for use if it includes at least eight hourly temperature setpoints, else the day is excluded from the dataset.

As the parameters have varying orders of magnitude, normalization of the data has been done to a range of 0 and 1 to help to improve the learner's performance. Normalization is achieved using

$$p_{stand} = \frac{p_{real} - p_{min}}{p_{max} - p_{min}} \tag{12}$$

with

p_{stand} … Standardized parameter value

p_{real} … Real parameter value

p_{min} … Minimum parameter value

p_{max} … Maximum parameter value.

The minima and maxima for all parameters are presented in Table **5**. They are gathered by analysing the actual minima and maxima of the parameters in the measured data and adding a small buffer for the temperatures, the minimum of the vapour pressure and the maximum of the sunshine hours.

*Table **5** Parameter Minima and Maxima for normalization*

Parameter [Unit]	**Minimum**	**Maximum**
Setpoint temperature [°C]	16	30
Outdoor daily mean temperature [°C]	-20	35
Outdoor daily minimum temperature [°C]	-25	25
Outdoor daily maximum temperature [°C]	-5	45
Running mean temperature [°C]	10	35
Vapour pressure [mbar]	0	25
Relative Humidity [%]	0	100
Cloud cover [-]	0	8
Sunshine hours [h]	0	16

5.2. Supervised Learning

The Mathworks' *Matlab 2017a* contains two toolboxes suited for machine learning applications: the Statistics and Machine Learning toolbox and the Neural Network toolbox. The Machine Learning toolbox contains several tools that enable the set-up of, among other things, regression learning tools based on input and output parameters. It comes with a total of 19 standard tools which can be divided into linear regression, decision trees, support vector machines, bagged and boosted trees, and Gaussian process regression (GPR). As a first step, all tools are run with standard conditions to estimate their overall feasibility. The Neural Network toolbox enables a quick set-up of artificial neural networks. Different learning methods can be applied that vary in computational effort and performance.

For a first evaluation Levenberg-Marquardt backpropagation is used as it is a good compromise between learning speed and performance. It stops training once the MSE of validation samples starts increasing [77]. One hidden layer is used first, with the number of neurons varying between nine and 60 in increments of three in order to find the best, non-overfitting number of neurons. The performance of all tools is analysed using the RMS of the output compared to a testing data sample.

In order to decide on which tools and which weather parameters to focus on learning is performed for a sample of eight rooms, with two rooms from each floor. The two rooms are chosen randomly among all rooms on the floor. The three best performing regression tools and the three best performing neural networks are chosen according to their respective boxplots, i.e. according to their median, furthermore their lower and upper quartile, and their minima and maxima. They are then trained with the remaining rooms and reanalysed in comparison to a baseline scenario of 21 °C heating setpoint and 23 °C cooling setpoint. Out of the six fully analysed tools the top three are chosen to vary the amount of training data fed into them. For this purpose, the whole data set is still fed into the trainer, but the amount used for training is varied between 5 % and 80 % in steps of 5 %, with the remainder being equally split into use for validation and testing.

It has to be noted that the Matlab tools automatically perform the data normalization, so the aforementioned normalization is not manually applied before training with the *Matlab* tools [36].

5.3. Reinforcement Learning

Matlab has no standard tools for reinforcement learning so it is necessary to develop a specific algorithm for this purpose. Since a continuous output is required from a set of up to six continuous inputs a lookup table for the value function is deemed impractical. Rather, a function approximation using artificial neural networks is performed. The neural network is based on the results from the supervised learning approach, as it is already optimized in size for the problem at hand. The error between the prediction and the desired setpoint is used as a penalty function and is directly used to improve the value function through gradient descent based error backpropagation for the ANN. The following process is used to train and evaluate the reinforcement learning agent:

1. Standardize inputs and target values
2. Initialise all weights in range [0;0,1]
3. Forward propagate data point
4. Calculate error
5. Use error to correct value function through gradient descent or back propagation
6. Repeat 3.-5. until full training set is used

A year worth of data, i.e. 365 data points, is used for training, while the remainder is used for testing. After each step of training, the test data is run through the current model to calculate its RMS. This is done to get an insight on the development of the RMS over time. After baseline analysing the simple agent, improvements are tested. A discounted reward, or in this case rather a discounted learning rate, is tested for reducing overfitting, while the (repeated) reuse of samples is tested to decrease the amount of data needed for the targeted performance. The discount rate for learning is varied between 0.5 % to 2.5 %, the number of sample reuses is varied between one and ten repetitions. After individual testing both methods are combined to evaluate their combined performance improvement. Three target values are analysed: Time until RMS below 2 °C, time until RMS below 1.5 °C and minimum RMS.

The artificial neural network used for the function approximation is based on the results from supervised learning and uses the same weather parameters as input as well as the same number of nodes in the hidden layer as the best-performing supervised ANN. The output layer has one node, as there is only one output value. The desired output is continuous. All nodes are sigmoids. Error back propagation is performed using a gradient descent approach, with the error, half of the mean squared error and the root mean squared error tested as error functions. Using only sigmoids, error backpropagation is performed as follows [35]:

1. Forward propagate input through neural network to receive output y

2. Calculate error e between target t and output per error function

3. Calculate output neurons responsibility for error $\delta^{(o)} = y \cdot (1 - y) \cdot e$

4. Calculate hidden neurons responsibility with h_j being the output of the j-th hidden neuron and w_{jo} the weight of the j-th neuron in the output layer

5. Update the weights with learning rate α

a.a. Output layer: $w_{jo} = w_{jo} + \alpha \cdot \delta^{(o)} \cdot h_j$

a.b. Hidden layer: $w_{kj} = w_{kj} + \alpha \cdot \delta_j^{(h)} \cdot x_k$ with x_k being the k-th input

6. Repeat 1-5 until training is terminated according to termination criterion

In the given case training will be continued until the training data set is exhausted to get an insight on the behaviour of the learning agent.

5.4. Building Model

A model of a single office is used to evaluate the energetic performance of a supervised and a reinforcement learning tool. The model is constructed with a room from the office building of Drees & Sommer at Obere Waldplätze 11, Stuttgart, Germany, serving as an orientation. The exterior constructions are according to the German energy saving regulation (*Energieeinsparverordnung*, EnEv) [78]. Figure **10** shows the office as modelled with the software *SketchUp*. As depicted, the office has one window, placed on the only exterior wall. The remaining surfaces are interior walls. The office is 4.0 m deep and 2.5 m wide, with a height of 2.8 m. The window is 2.1 m wide and 1.7 m high. It is placed 0.8 m above the floor and 0.2 m from the side walls. It has an U-Factor of 1.3 W/m²K, a solar heat gain coefficient of 0.6 and a visible transmittance of 0.78. For the model, it is assumed that all surrounding rooms are conditioned in a similar fashion as the model office, so that interior surfaces are considered to be adiabatic. As with most modern office buildings interior walls, ceilings and floors are light constructions. The details for the construction are given in Table **6**. An exterior venetian blind is used for shading. The blind has horizontal slats with a width of 0.08 m and a separation of 0.04 m. The blind is controlled by the incoming solar radiation and starts shading at an incoming solar radiation of 50 W/m² on the window surface. Infiltration is set to 0.7 air changes per hour, in accordance with the EnEv [78]. The three internal loads are one occupant, electric equipment and lighting. The occupant performs work at the desk, leading to an activity level of 115 W [79]. The office is equipped with a laptop and two additional screens accounting for a load of 120 W. Lighting is supplied by LEDs, adding up to 14 W [80]. The office is occupied from 8 am to 12 pm and from 1 pm to 6 pm on weekdays, there is no occupation during weekends. The electric equipment and lights are turned on during these periods as well. Air conditioning is modelled using the *EnergyPlus* Ideal Loads

Air System template, which supplies conditioned air under idealised conditions. The system's heating and cooling capacity is limited to 400 W to represent the limitations of state-of-the-art active ceilings used for individual temperature control in offices. The HVAC system is available two hours before and after occupation starts.

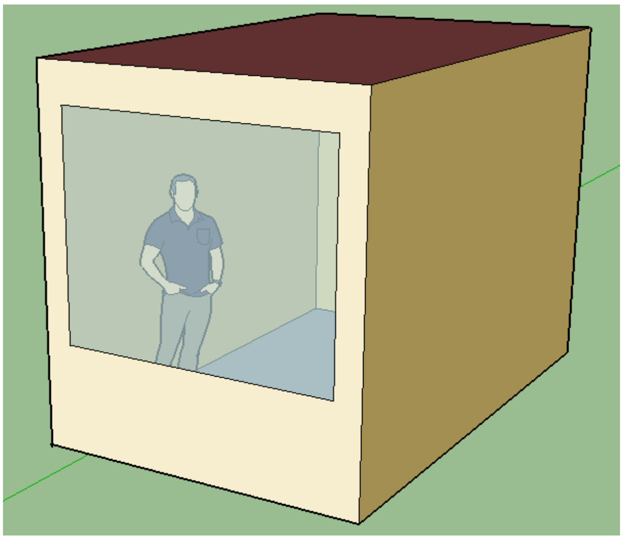

*Figure **10** Office Model*

The simulation is performed using setpoints from the eight rooms previously used for optimizing the neural network. Only setpoint data created with the best performing weather parameter combination is used. It can be assumed, however, that the other combinations would yield very similar results in the building simulation. The office is simulated with the exterior wall facing in the four main cardinal directions, i.e. east, south, west and north to incorporate potential differences caused by a varying solar radiation. A run time of one year is chosen to cover a range of potential weather conditions. Since the reinforcement learning tools are trained using data from April 1st 2012 until April 1st 2013, this period is also used for the building simulation.

*Table **6** Construction Materials [81]*

<table>
<tr><th>Construction element</th><th>Material</th><th>Thickness [m]</th><th>Specific heat [J/kgK]</th><th>Conductivity [W/mK]</th><th>U-Value [W/m²K]</th><th>Other information</th></tr>
<tr><td rowspan="2">Exterior Wall</td><td>light concrete</td><td>0.2</td><td>1000</td><td>1.3</td><td rowspan="2">0.28</td><td></td></tr>
<tr><td>fibre insulating board from mineral wool</td><td>0.115</td><td>100</td><td>830</td><td></td></tr>
<tr><td>Interior Wall</td><td>drywall</td><td>0.12</td><td>680</td><td>0.25</td><td>1.54</td><td>adiabatic</td></tr>
<tr><td>Ceiling</td><td colspan="4">massless construction</td><td>1.54</td><td>adiabatic</td></tr>
</table>

6. Results

The main goal of this thesis is to evaluate the feasibility of machine learning tools to predict indoor comfort temperatures for individually controllable conditioning systems. Supervised learning is used first to gain a general insight on the performance of machine learning tools in the given context. If this deviation, measured through the root mean squared error, is within the desired range of 2 °C, the feasibility of reinforcement learning is evaluated.

All machine learning tools specific settings, such as the number of nodes for a neural network, have to be optimized. For all tools, the input data has to be optimized as well. Table **7** shows all available weather parameters and a respective code number that is used for quick referencing and in graphics. Code number 1 is internally used for the time stamp and thus not used for external coding.

*Table **7** Weather Parameters*

Weather Parameter	Code Number
Average outdoor air temperature	2
Vapour pressure	3
Cloud cover	4
Relative humidity	5
Sunshine hours	6
Maximum outdoor air temperature	7
Minimum outdoor air temperature	8
Running mean air temperature	9

6.1. Supervised Learning

To evaluate the general applicability of machine learning to predict the temperature setpoint of individually conditioned rooms supervised learning methods – neural networks and regression learning tools – are developed using the *Matlab* Neural Network toolbox and the Statistics and Machine Learning toolbox. In a first step, the best-performing combinations of weather parameters and neural network sizes are evaluated, then the optimal weather parameter combination and regression learning tools are sought. In a last step, the minimum training data size for the best performing tool is analysed. The key parameter used for decision making are the Mean Squared Error (MSE) and the Root Mean Squared Error (RMS) of the tool's output compared to the measured data.

6.1.1. Neural Networks

To evaluate the performance of different combination of weather parameters and network sizes eight rooms of the building, two from each floor, are randomly chosen. For these rooms, all 33 sensible weather parameter combinations are used to teach a neural network with sizes from nine to 60 nodes, in steps of three. Each network size-weather parameter combination is run ten times, as the initial conditions used for the network trainer change. The arithmetic mean of all ten runs is taken. Figure **11** shows a boxplot for the ten best-performing combinations, ranked per the median of the total MSE of the eight rooms. All ten combinations have a neural network size of 33 nodes. The three weather parameter combinations with the lowest median MSE are air temperature-relative humidity-sunshine hours-maximum air temperature-running mean air temperature (2-5-6-7-9), air temperature-cloud cover-relative humidity-maximum air temperature-minimum air temperature-running mean air temperature (2-4-5-7-8-9) and air temperature-vapour pressure-cloud cover-running mean air temperature (2-3-4-9). The values used for decision making can be found in Annex **2**. Any further analysis will be performed on these three combinations using all available setpoint datasets.

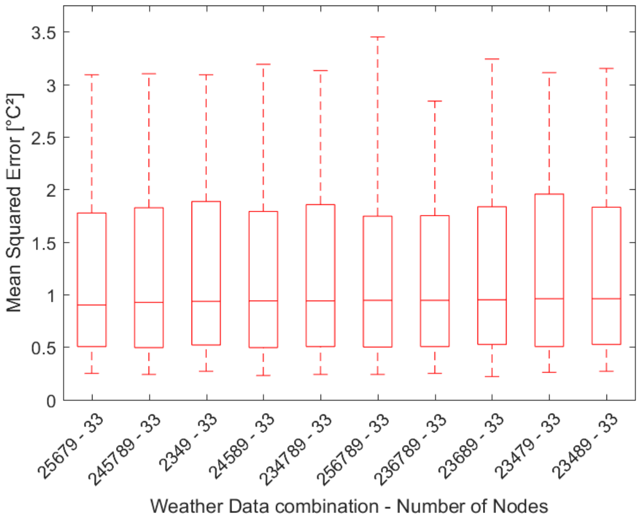

*Figure **11** Neural Network Size and Weather Parameter Dependant Total MSE - Top 10 for eight sample rooms*

The top three combinations are used to train further networks for the remaining 27 rooms. The resulting boxplots of the RMS are shown in Figure **12**, including a boxplot for the reference scenario. All neural networks have a median RMS of 1.12 °C and a maximum of 1.93 °C, 1.99 °C and 1.95 °C from left to right. All three outperform the reference scenario (21 °C heating setpoint, 23 °C cooling setpoint) with its median RMS of 2.11 °C and a maximum of 2.75 °C.

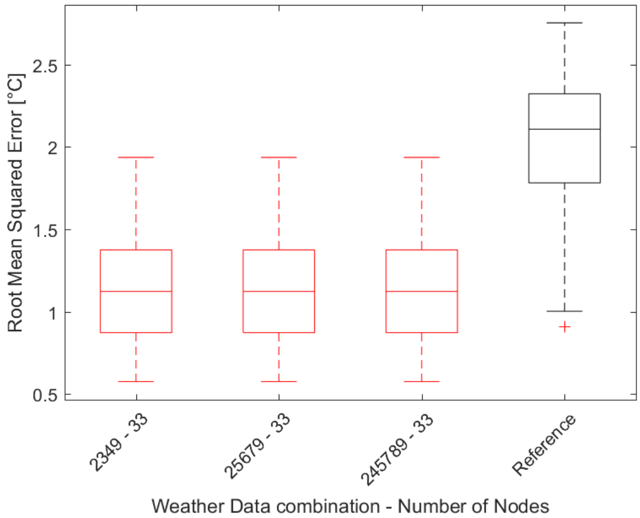

*Figure **12** Neural Network Top 3 Total RMS for all rooms, compared to reference scenario*

The datasets used have an average size of 870 data points. In the used setting, with 60 % of all available data being used for training, the neural network can outperform the reference scenario and even the maxima are below the target value of 2 °C for the RMS. It is thus reasonable to perform an analysis of the required amount of data to reach the target value. For the three combinations, the size of the training set is therefore varied from 5 % to 80 % of the available dataset, in increments of 5 %. Since all runs are new, including the one with 60 % of the data as training data, the performance slightly changes compared to the previous results. Figure **13** to Figure **15** show the results for the three weather parameter combinations. The medians show a logarithmic decay, with a median RMS of less than 2 °C being achieved with 15 % of the available data being used for training, equalling 131 data points. At roughly 30 %, or 261 data points, the medians go below 1.5 °C. It can be observed that the more weather parameters are available, the lower the RMS at 5 %. This gain from having more parameters is not visible from 10% onwards anymore. The decay of the medians starts to become small at around 50 % to 60 %. As a conclusion, at 15 % the use of neural networks becomes generally possible, and it starts to become feasible at 40 %, equivalent to 348 data points. It can also be concluded that using an overall larger data set improves the functionality of a neural network to predict indoor temperature setpoints.

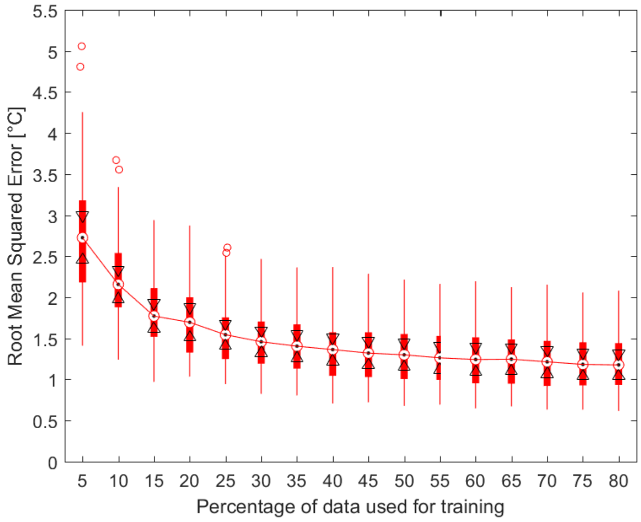

*Figure **13** Neural Network 2349-33 Total RMS over training data size*

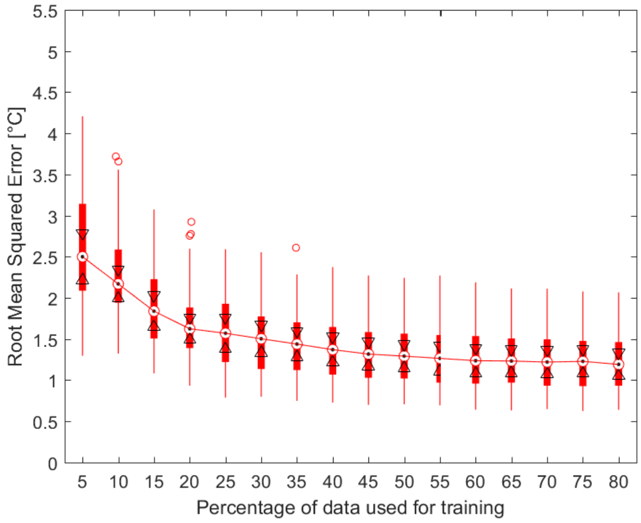

*Figure **14** Neural Network 25679-33 Total RMS over training data size*

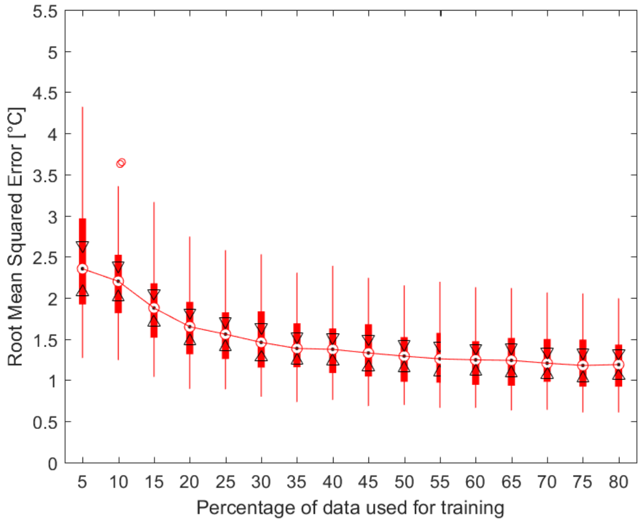

*Figure **15** Neural Network 245789-33 Total RMS over training data size*

6.1.2. Regression Learners

Similarly to the neural networks all potential weather parameter combinations are used to train regression learners. The Matlab Statistics and Machine Learning toolbox includes a regression learner with a total of 19 tools, as listed in Table **8**. For referencing in graphics, a code system is used, the code for each tool can be found in Table **8**.

*Table **8** Matlab Standard Regression Learner Tools*

Regression Tool	Code Number
Linear Regression (LR) **– Linear**	1
Linear Regression **– Interaction**	2
Linear Regression **– Robust**	3
Linear Regression **– Stepwise**	4
Tree **– Complex**	5
Tree **– Medium**	6
Tree **– Simple**	7
Support Vector Machine (SVM) **– Linear**	8
Support Vector Machine **– Quadratic**	9
Support Vector Machine **– Cubic**	10
Support Vector Machine **– Fine Gaussian**	11
Support Vector Machine **– Medium Gaussian**	12
Support Vector Machine **– Coarse Gaussian**	13
Ensemble **– Boosted Trees**	14
Ensemble **– Bagged Trees**	15
Gaussian Process Regression (GPR) **– Squared Exponential**	16
Gaussian Process Regression **– Matern 5/2**	17
Gaussian Process Regression **– Exponential**	18
Gaussian Process Regression **– Rational Quadratic**	19

The eight sample rooms are used for a first evaluation, out of which the three best performing tools are used for further analysis. With the regression learner tool, the initial conditions for training do not change, and a single run per set-up is sufficient. Figure **16** shows the result of the top ten regression learning tools. The nine best performing tools are all Gaussian Process Regression (GPR), the tenth best tool is the Bagged Tree Ensemble tool. Since the tool ranked first and third are the same, and the weather parameters of the second and third are the same, the fourth ranked combination is used instead, to guarantee some variation in the further analysis. The tools used are GPR – Exponential with 2-4-5-7-8-9, GPR – Matern 5/2 with 2-3-4-7-8-9 and GPR – Rational quadratic with 2-4-5-7-8-9. The exact values used for decision making can be found in Annex **3**.

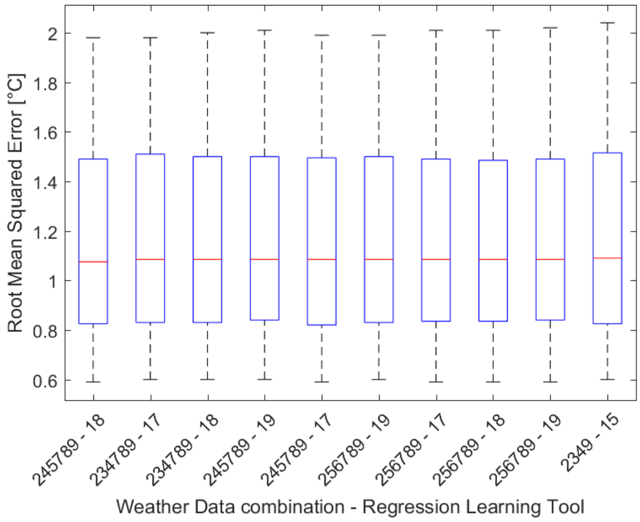

Figure 16 *Regression Tool and Weather Parameter Dependant Total RMS - Top 10 for eight sample rooms*

Figure **17** shows the RMS for the chosen tool-parameter combinations in comparison to the reference scenario. The median for GPR Exponential is 1.15 °C and the maximum at 1.98 °C, for GPR Matern 5/2 the values are 1.16 °C and 2 °C respectively and for GPR Rational Quadratic 1.15 °C and 2 °C respectively. The medians of all tools are below the threshold value of 2 °C and the maximum RMS are roughly equal to the threshold. All analysed tools clearly outperform the reference scenario. Overall, the use of regression learning tools is viable with the amount of data available.

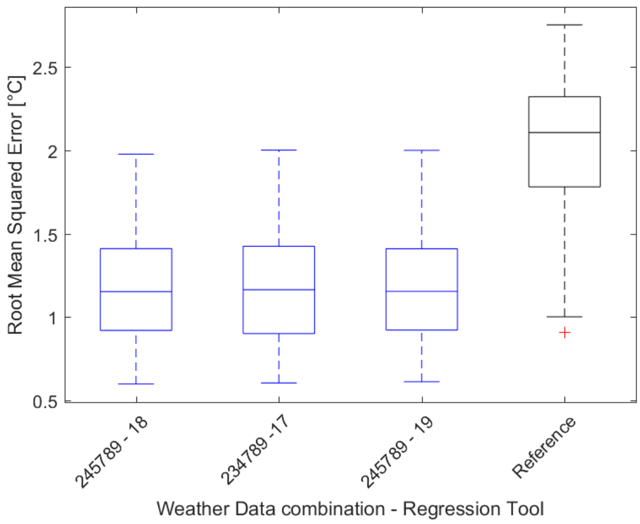

*Figure **17** Regression Tool Top 3 RMS for all rooms, compared to reference scenario*

6.1.3. Comparison of Supervised Learning Tools

The previous analysis has shown that the six further developed tools are viable to predict indoor temperature setpoints based on outdoor weather data. All of them outperform the reference scenario and have both median and maximum RMS below the target threshold value of 2 °C. As can be seen in Figure **18**, the performance of the tools is similar. The neural networks have a median of 1.12 and outperform the regression tools by approximately 0.03 °C. The maxima vary only slightly as well, with the best performing tool, the neural network 2349-33, having a maximum RMS of 1.93 °C and the worst, regression tools 2-4-5-7-8-9 – GPR Rational Quadratic and 2-3-4-7-8-9 – GPR Matern 5/2, having a 0.07 °C higher maximum. Taking also the quartiles into account the neural network tool 245789-33 shows the most desirable performance, as the central 50 % of results are the lowest of all tools, again though with only some hundredths of a degree deviation between the tools.

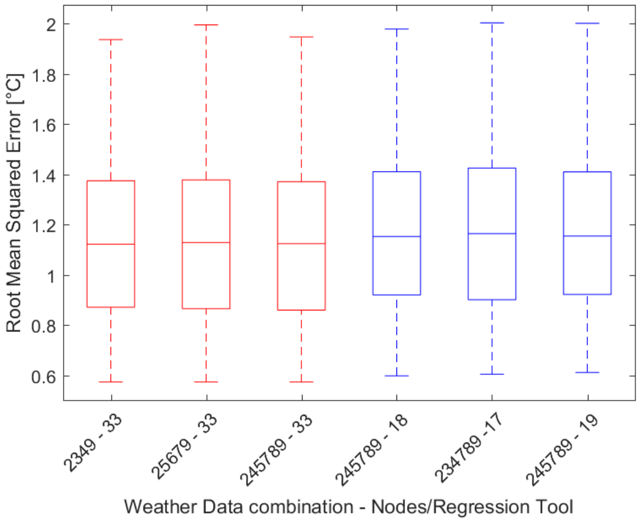

*Figure **18** Neural Network and Regression Tools Comparison*

6.2. Reinforcement Learning

Since the supervised learning tools comply with the conditions of a RMS below 2 °C and outperform the baseline scenario, the performance of reinforcement learning tools is evaluated. To clarify the difference between reinforcement learning and the previously performed supervised learning: supervised learning is a very linear process, as depicted in Figure **19**. First, training data is collected for a predefined period of time, which is then used as a complete batch to train the comfort predictor. In a final step, the predictor is put into operation. In reinforcement learning on the other hand, as depicted in Figure **20**, the predictor is put into operation in the beginning. On its first day of operation it uses the weather forecast to predict the comfort temperature, and then uses the user input to correct the model parameters to improve its prediction for day two. This daily learning continues, until a terminal condition for learning is reached.

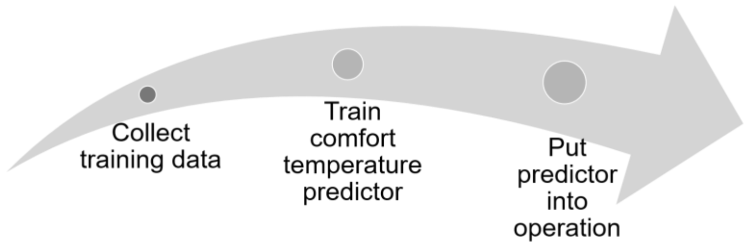

*Figure **19** Supervised learning process*

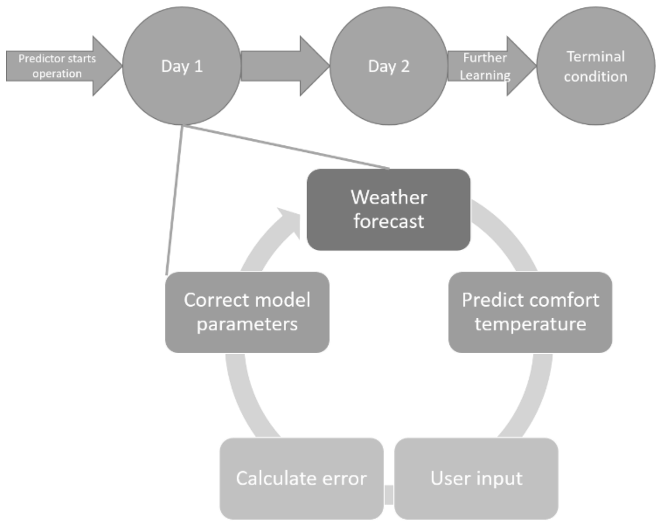

*Figure **20** Reinforcement learning process*

6.2.1. ANN-based function approximation

Supervised learning methods are often used for function approximation of the value function in reinforcement learning [38]. ANNs have shown good potential for comfort temperature learning in the supervised learning approach, so that their application in reinforcement learning is only logical. In the following section results of a multilayer-perceptron ANN with error backpropagation for function approximation of a reinforcement learning agent are presented. The shown graphs depict the results for the weather combination 2-4-5-7-8-9 with 33 nodes in the hidden layer. The results for the weather combinations 2-3-4-9 and 2-5-6-7-9 are like the presented ones.

Error Function Evaluation

Three error functions for the difference between the estimate of the value function, *y*, and the actual user input, *h*, are evaluated. The first is the simple error between the input and the estimate, depicted in Figure **21**

$$e_{simple} = h - y, \tag{13}$$

second is half the mean squared error, depicted in Figure **22**,

$$e_{0,5MSE} = -\frac{1}{2} \cdot (h - y)^2, \tag{14}$$

the third the root mean squared error, depicted in Figure **23**,

$$e_{RMS} = -\sqrt{(h - y)^2}. \tag{15}$$

For the given learning rates e_{simple} and e_{RMS} show a steeper reduction in the RMS of the testing data. Their optima are lower than for the half MSE as well, reaching values of around 1.5 °C for all learning rates starting from 0.015 and higher. However, the RMS also shows a strong tendency to overfitting after the optimum is reached, while the simple error shows some early fluctuations with increasing learning rates. $e_{0,5MSE}$ does not show a strong overfitting behaviour. It can on the other hand be assumed that the optimum has not been reached within the learning period of 365 days. Since the RMS-error function outperforms the half-MSE for the used learning rates which are generally standard in machine learning, and seems slightly more stable than the simple error, the RMS-error function is used for further improvements on the reinforcement learning tool.

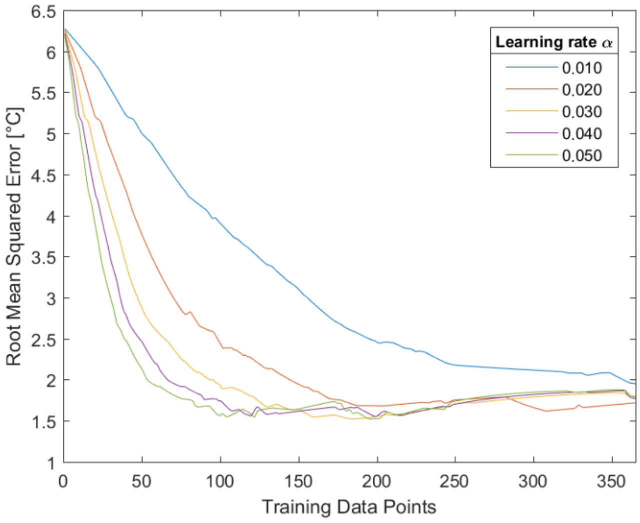

*Figure **21** Reinforcement Learning: Error Function Evaluation - Simple Error*

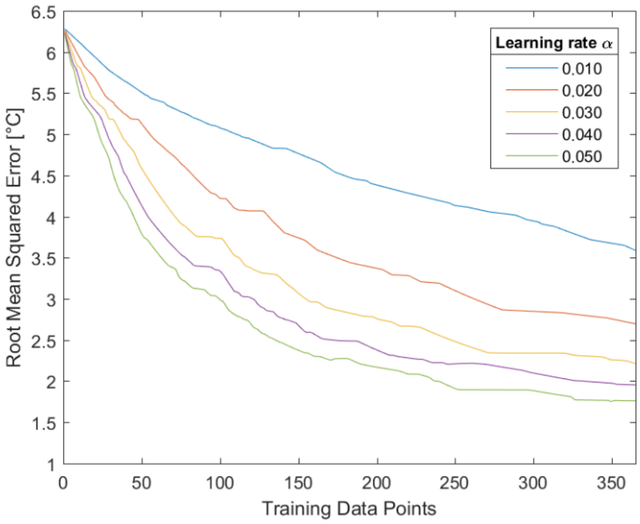

*Figure **22** Reinforcement Learning: Error Function Evaluation - Half MSE*

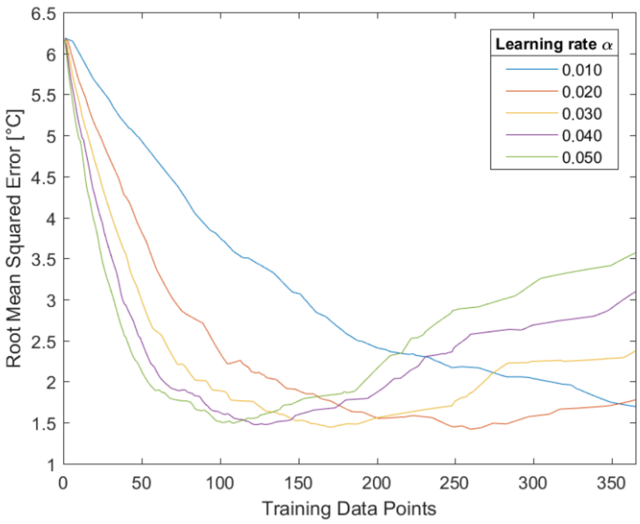

*Figure **23** Reinforcement Learning: Error Function Evaluation – RMS*

Discounted Learning Rate and Sample Reuse

Discounting the learning rate reduces the influence of new inputs later into the learning process and may help to reduce overfitting. Sample reuse is using the same data points for training more than once and may help to reduce the required number of training data points to reach an optimal behaviour. A combination of both methods might have the potential to improve both factors of a Reinforcement Learning tool. For discounted learning the learning rate is reduced by between 0.5 %, see Figure **25**, to 1.5 %, see Figure **24**, with each new data point used. The higher the discounting, the later the optimum is reached. For lower learning rates, this leads to a lower total optimum that can be reached, as the applied learning is getting closer to zero before the optimum can actually be achieved. For higher learning rates a clear reduction in overfitting can be observed. While the RMS for the learning rate of 0.05 at the end of the learning period is almost double its minimum for a discount rate of 0.995, it is barely higher than its minimum for a discount rate of 0.985. It does however require 100 days to reach the minimum, compared to roughly 50 days it takes with less discounting.

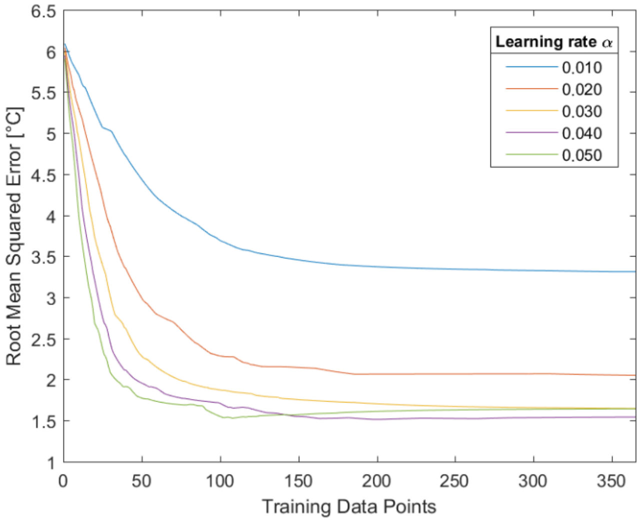

*Figure **24** Reinforcement Learning: Discounted Learning - Discount Rate 0.985*

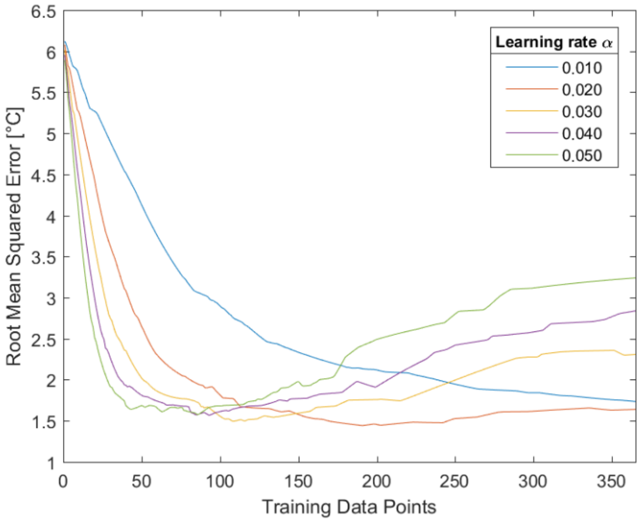

*Figure **25** Reinforcement Learning: Discounted Learning - Discount Rate 0.995*

In sample reuse, the same data point is used for several training sessions. One repetition equals to the previously used solution without discounting. In Figure **26** and Figure **27** sample reuse is applied with learning rates of 0.01 and 0.05. An increase in repetitions reduces the number of data points needed to reach a local minimum but also increases the tendency to overfitting. For higher learning rates, it is possible to reach the local minima within less than 20 days, however, drastic overfitting starts relatively early, so that the RMS increases fast after having reached its minimum. Learning rates above 0.05 further increase this effect.

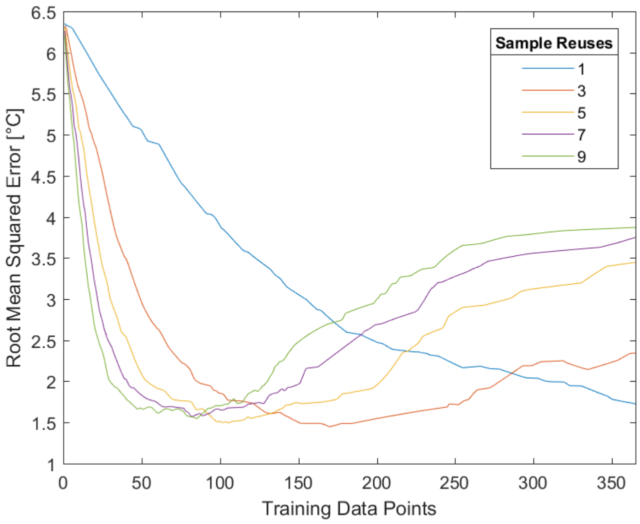

*Figure **26** Reinforcement Learning: Sample Reuse - Learning Rate 0.01*

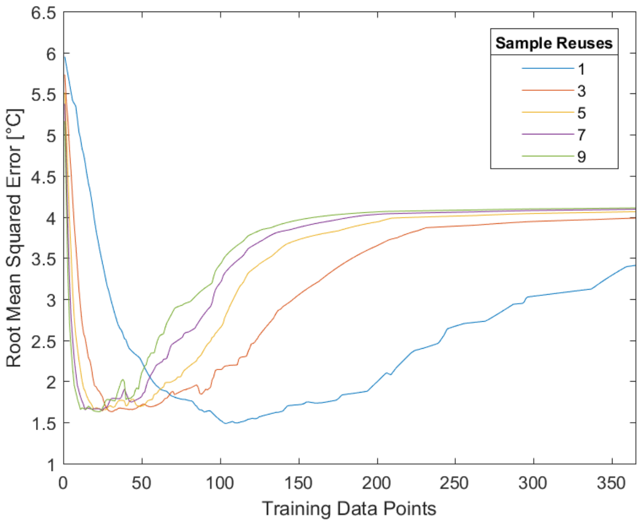

*Figure **27** Reinforcement Learning: Sample Reuse - Learning Rate 0.05*

Discounting helps to reduce the tendency to overfit but reduces the learning speed, while sample reuse increases the learning speed but increases the tendency for overfitting. Uniting both methods may combine their benefits, thus leading to an overall improved performance. Including the learning rate, this leads to a total of three parameters that need to be optimized. Figure **28** and Figure **29** show the clear influence of the discount rate on learning, with a higher discount reducing overfitting from 100 data points onwards. A higher discount rate does on the other hand slow down learning, and leads to the optimum being reached at a later point. For further improvement, higher learning rates are tested with varying discount rates. The results are shown in Figure **30** to Figure **32**. As expected, a higher learning rate increases the learning speed, resulting in a steeper slope at the start of the learning phase. A learning rate of 0.05 shows comparatively high fluctuations, as seen in Figure **30**, especially around 50 training data points. It is therefore considered too unstable for further use. A learning of 0.03 is more stable than 0.05 while still being visibly faster than 0.01. A discount rate of 0.985 does not give the same smooth outrun for 0.03 as it does for 0.01, so a higher discount rate of 0.975 is tested. For five or more sample reuses a RMS minimum is reached before fifty data points are used. The behaviour when reusing a sample five times is very desirable. It shows a significant steep reduction of the RMS in the beginning before flattening out at around 50 data points. It shows a slight decrease up until 125 data points, after which the learning stops with an RMS of approximately 1.55 °C. While slower setups achieve lower final RMS, the combination of speed, final RMS and stability achieved with the setup of five sample reuses, a learning rate of 0.03 and a discount rate of 0.985 looks the most promising for real life application.

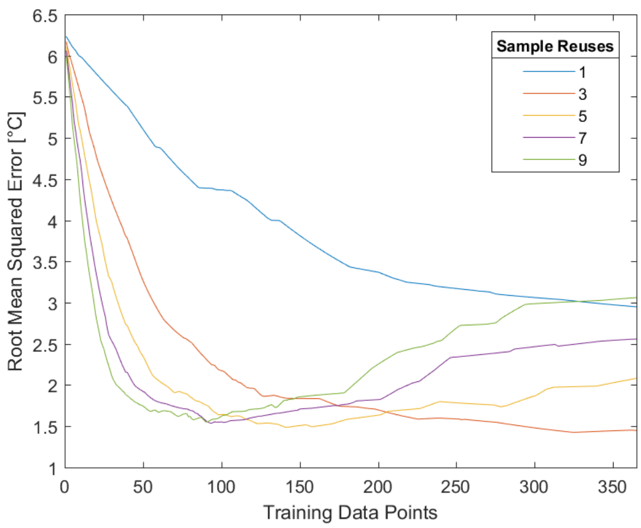

*Figure **28** Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.01, Discount Rate 0.995*

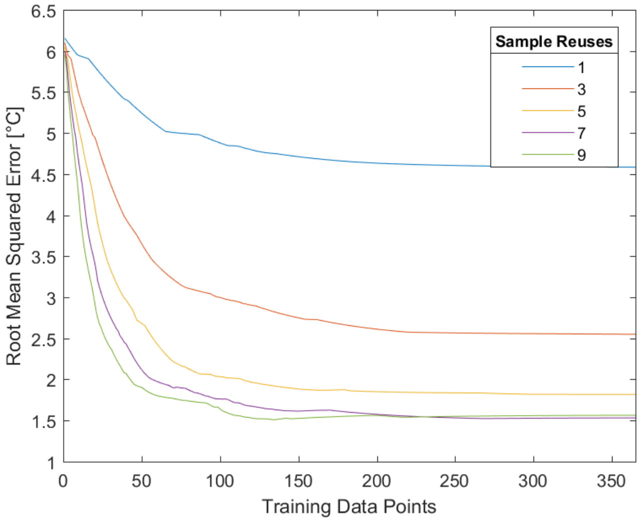

*Figure **29** Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.01, Discount Rate 0.985*

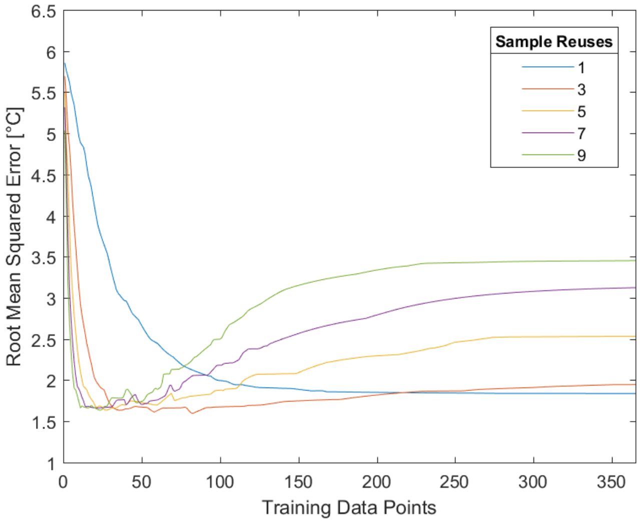

*Figure **30** Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.05, Discount Rate 0.985*

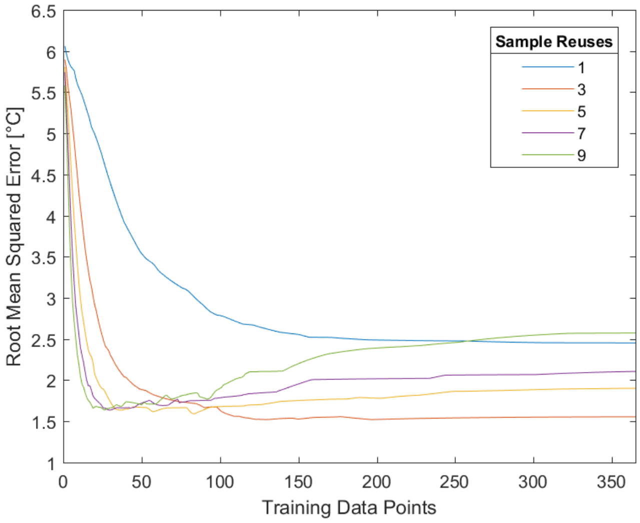

*Figure **31** Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.03, Discount Rate 0.985*

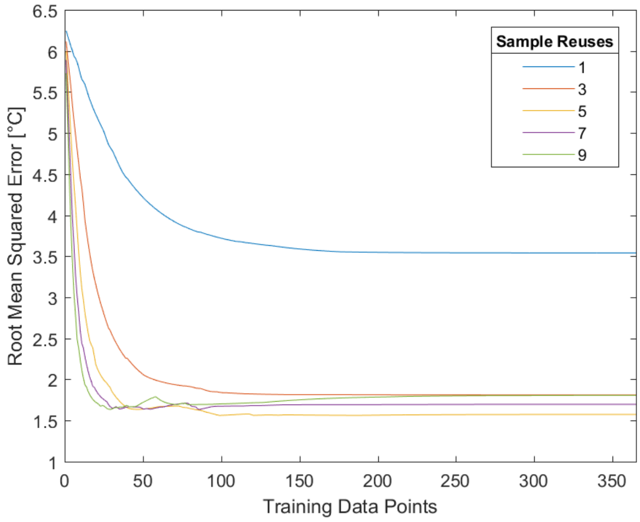

*Figure **32** Reinforcement Learning: Discounted Sample Reuse - Learning Rate 0.03, Discount Rate 0.975*

6.3. Energy Evaluation

It has been shown that the comfort temperature for individuals can be predicted with sufficient precision using supervised and reinforcement machine learning tools. The tested tools outperform the currently common heating setpoint of 21 °C and cooling setpoint of 23 °C and could thus help to reduce over conditioning. For a potential real-life application, the tools should, apart from precisely predicting the user's comfort temperature, at most marginally increase the energy consumption compared to the baseline scenario. For this purpose, a simple office model is used, in which a 400 W idealised air conditioning system aims at reaching the setpoints provided by the machine learning tools and the baseline scenario. As an example, for the supervised learning method "artificial neural network with 33 nodes" and the weather parameter combination "daily average outdoor air temperature, cloud cover, relative humidity, daily maximum and minimum air temperature, and running mean air temperature" is used. It has shown the best results in the prediction performance tests and would thus also be the most likely candidate for later application. For the reinforcement learning value function approximation with the same neural network setup is used. The office's exterior wall is rotated consecutively to face the four main cardinal directions.

The sensibility of the used office model can be analysed by comparing the received heating and cooling loads to the energy performance certificate of the actual office building. The useful energy demand for heating is given at 55.5 kWh/m²a, or 555 kWh/a for the used 10 m², and 222 kWh/a for cooling. The model values are in a range of 200 kWh/a for heating and 200 kWh/a to 500 kWh/a for cooling. These values are opposite of the performance certificate, but they are however considered plausible, as the general order of magnitude is similar, and the order and direction of deviation can be explained by the relatively high internal loads of roughly 21 W/m² in the office model compared to a whole building with on average lower internal loads of 14 W/m². [80]

Figure **33** shows the median sensible heating loads per year of the eight sample rooms for each used method and cardinal direction. As expected the heating load is lowest when the exterior wall is facing south, and highest when facing north, due to varying solar gains. The heating load for reinforcement learning are the lowest for all cardinal directions, with 174.5 kWh/a for the south facing exterior wall and 229.7 kWh/a when facing north. The supervised learning approach and the baseline scenario require similar amounts of heating. The baseline scenario requires 185 kWh/a when the office is facing south and 252.3 kWh/a when facing north, compared to 192.6 kWh/a and 250.3 kWh/a respectively for the supervised artificial neural network. Since the power of the HVAC system is limited to 400 W it is necessary to also evaluate whether or not the system is actually powerful enough to supply the heating and cooling required. Figure **34** depicts the time that the system did not reach the needed heating setpoints depending on the used method to find the setpoints, and cardinal direction. The heating setpoints from reinforcement learning are not reached 15 h to 27 h in the sample year, while the baseline setpoints are not reached between 28 h to 46 h. This implies that, with a more powerful HVAC unit, the heating load of the baseline scenario should increase the most.

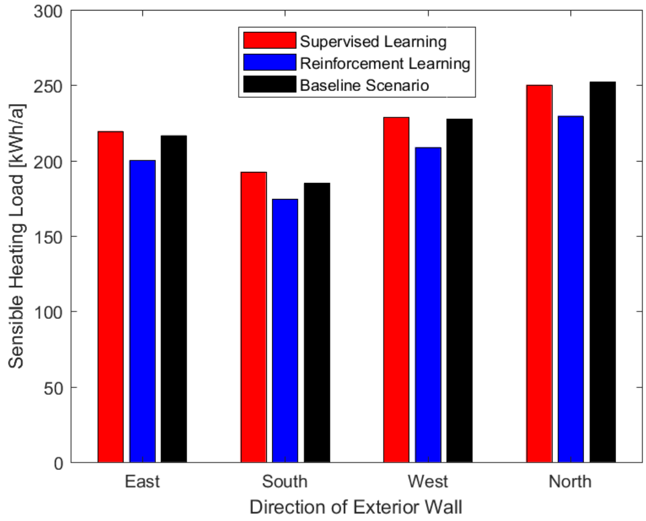

*Figure **33** Heating Loads per method and cardinal direction*

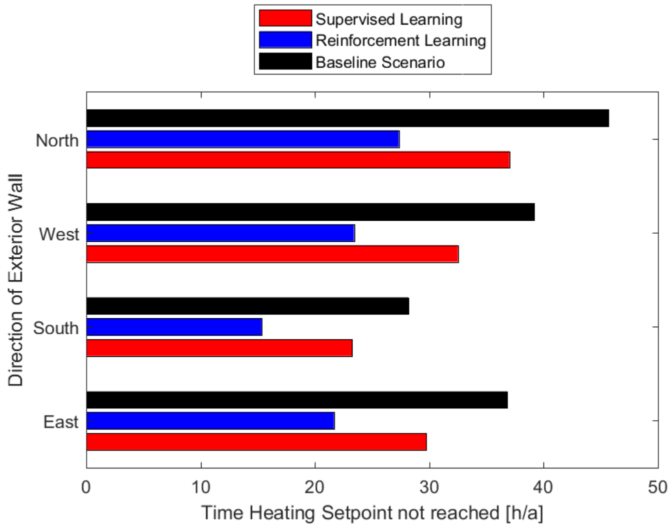

*Figure **34** Hours heating setpoints are not reached per method and cardinal direction*

Second, the cooling loads and the hours during which the cooling setpoint is not reached are evaluated. The cooling loads, depicted in Figure **35**, per cardinal direction are reversed to the heating loads, again due to varying solar heat gains. Using the setpoints provided by the two machine learning tools leads to higher cooling loads than the baseline scenario. When the exterior wall faces north, the loads for supervised and reinforcement learning are 348.3 kWh/a and 364.3 kWh/a respectively, roughly double the baseline scenario load of 176.7 kWh/a. With a south facing exterior wall the loads are 513.7 kWh/a, 532.2 kWh/a and 315.9 kWh for supervised learning, reinforcement learning and the baseline scenario. The fulfilment of the setpoint shows a similar trend, with the setpoints from the baseline scenario being fulfilled roughly 60 h to 200 h more a year than for the machine learning tools, as shown in Figure **36**. In line with the heating loads, this means that a more powerful heating unit would lead to a higher increase in energy consumption for the machine learning tools than for the baseline scenario.

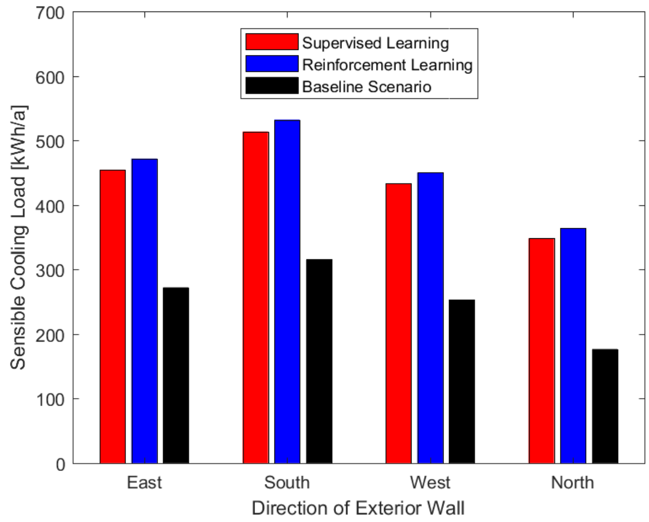

*Figure **35** Cooling Loads per method and cardinal direction*

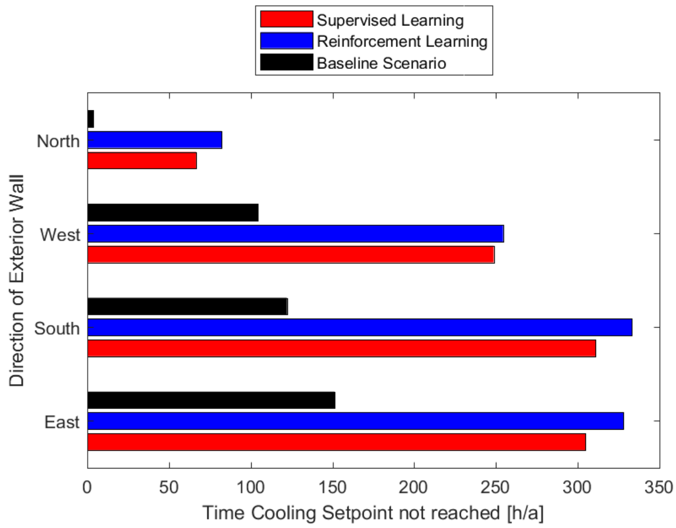

*Figure **36** Hours cooling setpoints are not reached per method and cardinal direction*

It is important to notice that these results would not depict the actual energy consumption within the rooms, as there is a deviation between the used setpoints and the user requirements. As has been analysed, the machine learning tools perform better when it comes to predicting the user requirements, which implies that the models using machine learning based setpoints are closer to the actual consumption of the offices than the baseline scenario. Previously, the root mean squared error was used to evaluate the performance of the different methods. While it is useful to evaluate the magnitude of the deviation, it has no implication about their general direction. To develop and insight into which direction the energetic values from the simulations tended if there was actual user input the mean arithmetic error between each the baseline scenario, the supervised learning tool and the reinforcement learning tool, and the actual user inputs is calculated. It is necessary to have four separate calculations for the baseline scenario, as the used setpoint is dependent, among other things, on the incoming solar radiation. To get a separate insight on cooling and heating the error is calculated separately for the calendar seasons, spring from March to May, summer from June to

August, autumn from September to November and Winter from December to February. The predicted setpoints are subtracted from the actual user input, a negative error thus means that the predicted setpoints are too high, and vice versa. Figure **37** through Figure **40** depict the boxplots for the mean arithmetic error for the eight sample rooms for European meteorological spring, summer, autumn and winter, respectively. From spring to autumn, the baseline setpoints exceed the user input. In the transition seasons the median deviations between the baseline and the user input are at around 2 °C for all setups, whereas the machine learning tools provide setpoints that, in median of the averages, do not deviate from the user input. Since cooling is dominant in the two seasons, it is implied that the cooling loads from the baseline scenarios are too low compared to a situation with user input, while the model for the machine learning tools is relatively precise. Summer shows a similar relationship between the used setpoints and the user input, the range of deviations for the machine learning tools is bigger, ranging from 1 °C to – 2 °C for supervised and 2.5 °C to – 2 °C for reinforcement learning. The mean arithmetic error of the baseline cases has a median of around – 3.5 °C and is in a range of - 2 °C to – 4.5 °C. Again, this implied an increase in the cooling load if user input was possible. In winter, the baseline scenarios have mean arithmetic errors in a range of + 1 °C to – 3 °C, with medians at around – 0.5 °C. Supervised learning is in a range of ± 0.5 °C with a median close to zero, and reinforcement learning ranges from + 3 °C to – 1 °C with a median at around 1.3 °C. The negative median for the baseline scenario implies that overheating takes place, whereas the positive median of the reinforcement learning tool implies that the room is too cold and the overall heating load of the reinforcement learning tool would be higher given user input.

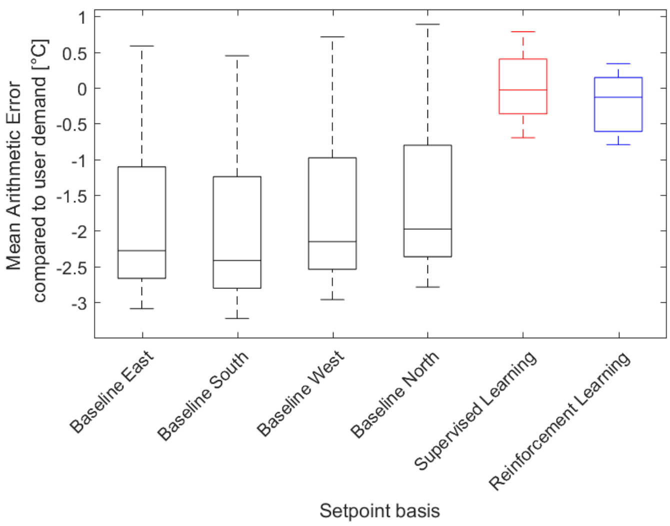

*Figure **37** Mean arithmetic error in Spring*

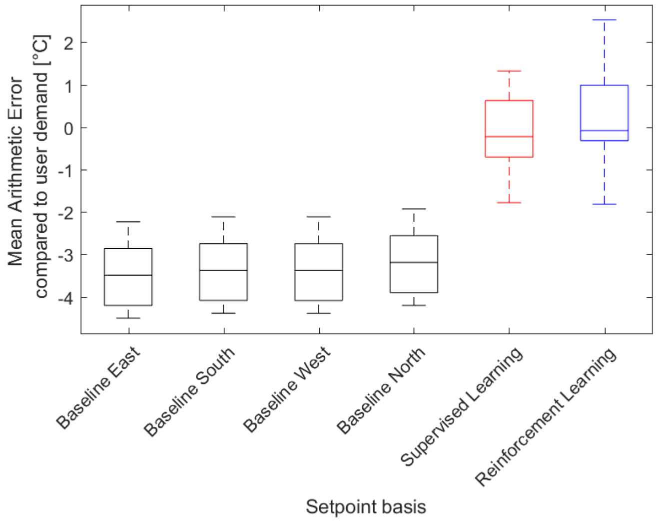

*Figure **38** Mean arithmetic error in summer*

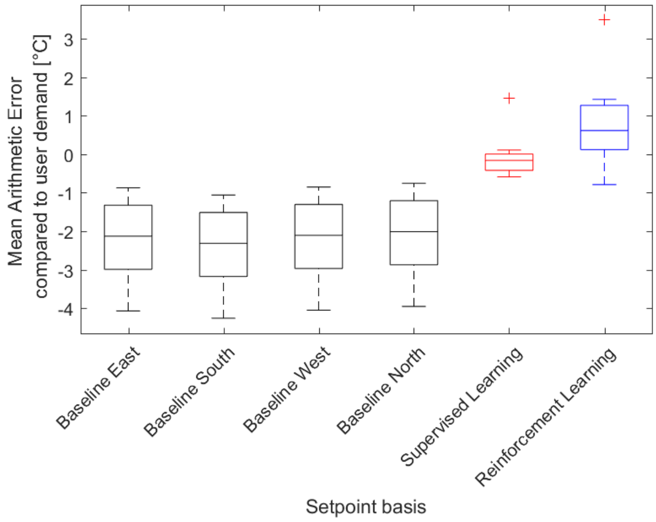

*Figure **39** Mean arithmetic error in autumn*

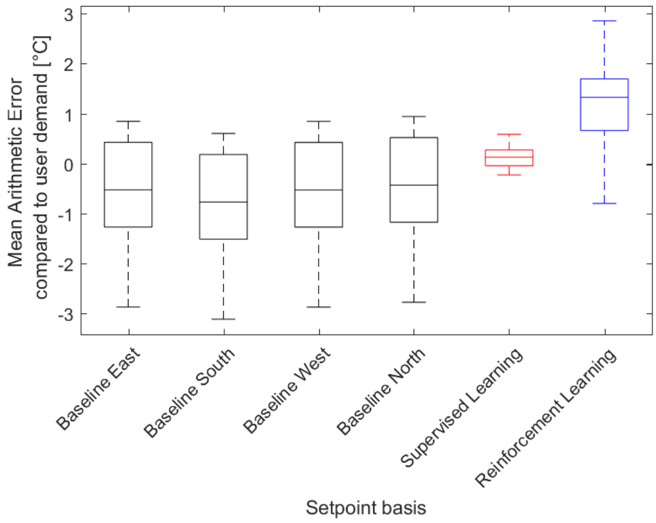

*Figure **40** Mean arithmetic error in winter*

7. Conclusion

User behaviour is a large factor of uncertainty in building operation. While occupancy tracking and prediction are becoming more and more available through various tools, such as learning thermostats [14,62–65], the prediction of individual comfort temperatures with simple, non-intrusive means is still lacking. Using *Fanger's* PMV model is problematic for individual prediction as it was developed with groups of people in mind, while inter-person preferences may differ by up to 3 K [4]. It furthermore requires data that is difficult to measure or highly specific, making its broad-scale application complicated and expensive [20]. According to the adaptive comfort approach the comfort temperature is influenced, among other things, by the weather, mostly the outdoor air temperature, as it influences factors such as clothing and metabolic rate for given activities [5]. Machine learning has the potential to find the relation between weather parameters and an individual's comfort temperature, as it enables an efficient analysis of large amounts of data. If measured data is available, supervised learning can draw connection between input data and the provided, desired output to make future predictions. If no data is available, reinforcement learning can be used for real-time learning, as the learning agent gathers its input by interacting with the environment [11,31–35].

To find the best performing tools and weather parameters to predict an individual's comfort temperature, different machine learning tools and weather parameter combinations were tested. The machine learning tools examined were supervised regression learning, artificial neural networks and reinforcement learning using artificial neural networks for function approximation. The tested weather parameters were the daily average temperature, daily minimum and maximum temperature, the running mean outdoor air temperature for one week, the outdoor vapour pressure or relative humidity, and the cloud cover or sunshine hours. The target for all tools is to outperform a baseline scenario using a heating setpoint temperature of 21 °C and a cooling setpoint temperature of 23 °C as well as having a root mean squared error between user input and prediction of less than 2 °C. The Mathworks' *Matlab* has a Statistics and Machine Learning toolbox and a Neural Network Toolbox that were used to develop and evaluate the supervised learning methods [77]. Apart from the performance when predicting comfort temperatures, it is important that the used tools do not increase energy consumption compared to the baseline scenario to be generally considered for application. A simple office model was therefore created with SketchUp, *OpenStudio* and *EnergyPlus* to compare the energetic performance of the baseline scenario compared to using the setpoints predicted by the reinforcement learning tools [73–75]. The used temperature setpoints were gathered at the central building of the savings bank *Kreissparkasse Göppingen* in Germany from the period of 01.04.2012 to 02.12.2014, at 35 rooms with individual temperature control spread over the first to fourth floor of the building. Weather data was obtained from the German weather service's closest station with sufficient data, located in Stuttgart-Echterdingen [76]. The same data was used for supervised learning and reinforcement learning, the latter being artificial, by feeding the learning agent with the data as a time-chain.

The best performing supervised neural networks had 33 nodes in one hidden layer and used the three weather parameter combinations air temperature - relative humidity - sunshine hours - maximum air temperature - running mean air temperature, air temperature - cloud cover - relative humidity -

maximum air temperature - minimum air temperature - running mean air temperature and air temperature - vapour pressure - cloud cover - running mean air temperature. They had median root mean squared errors over the 35 used rooms of 1.12 °C and a maximum RMS below 2 °C and furthermore outperformed the baseline scenario thus all being considered feasible for predicting individual comfort temperatures. To achieve this performance setpoint and weather data for approximately one full year was required. The three best performing regression learning tools were Gaussian Process Regression (GPR) – Matern 5/2 with air temperature-vapour pressure-cloud cover-maximum air temperature-minimum air temperature-running mean air temperature, GPR- Exponential with air temperature-cloud cover-relative humidity-maximum air temperature-minimum air temperature-running mean air temperature and GPR – Rational Quadratic with air temperature-cloud cover-relative humidity-maximum air temperature-minimum air temperature-running mean air temperature. All three showed a performance similar to the ANN, with their medians being 0.03 °C higher. They were thus also considered feasible to predict comfort temperatures. When using reinforcement learning with ANN for value function approximation it was possible to reach an RMS of roughly 1.5 °C, which was considered sufficiently low, with the speed depending on the used learning rate and error function. Using the RMS as the error function and a learning rate of 0,05 led to the minimum being reached after around 100 days. Using simple reinforcement learning with RMS for error back propagation led to considerable overfitting after reaching the minimum, which could be reduced using a discounted learning rate. For learning rates of 0.03 and higher, a discounting factor of 0.985 stabilized the RMS at roughly 1.6 °C after reaching the minimum. For speeding up the learning samples can be reused. Again, a compromise between speed, stability and overfitting had to be found. With a learning rate of 0.03, a discount factor of 0.975 and reusing the samples five times it was possible to reach an RMS of 1.7 °C after only 50 days of training and a final RMS of 1.6 °C. This learning time, as well as the RMS, were considered sufficient to consider the implementation of reinforcement learning.

The energetic performance of the tools was separated into heating and cooling. The reinforcement learning tool was able to outperform the baseline scenario for heating, while the performance of the supervised leaning was on par with the baseline scenario. Given the differences between the setpoints used in the building simulation and the actual user inputs it could be concluded that the reinforcement learning tool would require more energy in a real application than the simulation implies, as the setpoint temperatures were on average too low compared to actual user input. The baseline scenario on the other hand had too high setpoints, implying overheating. A very well-tuned supervised artificial neural network may be able to reduce overheating and therefore reduce the heating load. The cooling load for the machine learning tools gathered from the building simulations was approximately twice as high as the load for the baseline scenarios. However, the cooling setpoints of the machine learning tools matched the user inputs decently well, while the baseline scenario was roughly 2 °C too warm in the transition seasons spring and autumn, and 3.5 °C too warm in summer, compared to user inputs. This implied a drastically higher energy consumption in a real application for the baseline scenario as compared to the simulation. As there was no over-conditioning when using the baseline scenario, it can be concluded that the machine learning tools cannot provide any energetic benefit for cooling.

Especially when using activated building components which are thermally slow, they should help to increase user comfort due to the increased precision in predicting comfort temperatures.
In summary, both supervised and reinforcement learning tools can predict individual comfort temperatures based on weather parameters with the desired precision and their predictions being closer to user desires than the currently common heating and cooling setpoint temperatures. Artificial neural networks may be able to reduce overheating, thus decreasing the overall heating load, while machine learning tools in general might help to increase indoor comfort during the cooling period, however, potentially at the cost of increasing the cooling load, depending on user behaviour.
The results of this thesis show that it is possible to predict individual comfort temperatures using machine learning. A deeper analysis of the topic seems worthwhile, considering the impact setpoint prediction may have on the flexibility and efficiency of building energy management systems, especially in combination with other advanced control techniques such as occupancy prediction and (model) predictive control. Further tests need to be performed using a wider range of data, from varying locations to be able to give a statement about the general applicability. As the data basis for this work was coming from a bank the comfort temperatures throughout the year, while varying, lay within a certain range, especially since the choice of clothing is highly limited. Predicting comfort temperatures in workplaces with a larger possibility of clothing options may prove to be more challenging, and would thus prove to be a good benchmark for the tools. Apart from more theoretical analyses real life implementations are an important next step. They would enable insights into actual changes in energy consumption and more importantly create user feedback. As shown by *Yang* and *Newman* user satisfaction and willingness to interact with learning machines is an important factor towards their performance [60]. An advanced implementation would later also test the interaction of comfort prediction, occupation prediction and model predictive control. Implementing the tools within an existing framework would require an interface that enables docking into it. Considering that more and more companies are moving into the market a common standard would be useful [82].
While the use of machine learning fits the theme of the fourth industrial revolution by using data analysis and putting a closer relation between the building energy system's capacity and the use, there is also criticism about the rise in the variety of technologies used for building systems [83]. The higher level of complexity due to an increased number of components and the further specialised knowledge required for these individual parts make it more and more difficult for all involved users, from facility managers to occupants, to understand where a building system's behaviour is coming from. Early research on the Nest thermostat showed users switching back to classic scheduled thermostats since they were unable to understand why the system was acting the way it was [60]. Overall, the implementation of high tech solutions into building management systems leads to increased complexity and system vulnerability, in times where progress on building energy systems' hardware is low and the energy demand still needs steady decreases, tools such as machine learning offer interesting new potential. As *Dietmar Eberl*, professor of architecture at the ETH Zurich, and proponent of low-tech solutions, stated, it is necessary to "[…] use the technology required in buildings

more intelligently […]” [2][84], which leaves some room for interpretation and gives some food for thought about the future of building energy systems.

[2] Translated from German. Original: „[…] Die für den Bau notwendige Technik intelligenter nutzen […]“

References

[1] European Commision. Indoor air pollution: new EU research reveals higher risks than previously thought. Brussels, Belgium; 2003.

[2] Klepeis NE, Nelson WC, Ott WR, Robinson JP, Tsang AM, Switzer P et al. The National Human Activity Pattern Survey (NHAPS): A Resource for Assessing Exposure to Environmental Pollutants. Journal of Exposure Analysis and Environmental Epidemiology 2011;11:231–52.

[3] Gorvett Z. The never-ending battle over the best office temperature. BBC 2016, 20 June 2016; Available from: http://www.bbc.com/capital/story/20160617-the-never-ending-battle-over-the-best-office-temperature. [March 03, 2017].

[4] van Hoof J. Forty years of Fanger's model of thermal comfort: comfort for all? Indoor Air 2008;18(3):182–201.

[5] Nicol JF, Humphreys MA. Adaptive thermal comfort and sustainable thermal standards for buildings. Energy and Buildings 2002;34(6):563–72.

[6] O'Brien W, Gunay HB. The contextual factors contributing to occupants' adaptive comfort behaviors in offices – A review and proposed modeling framework. Building and Environment 2014;77:77–87.

[7] Haldi F, Robinson D. Modelling occupants' personal characteristics for thermal comfort prediction. Int J Biometeorol 2011;55(5):681–94.

[8] Economidou M. Europe's buildings under the microscope: A country-by-county review of the energy performance of buildings; 2011.

[9] Pinheiro MD. Introduction. Lisbon; 2016.

[10] European Commision. 2030 climate & energy framework. [September 09, 2017]; Available from: https://ec.europa.eu/clima/policies/strategies/2030_en.

[11] Ramasubramanian K, Singh A. Machine Learning Using R. Berkeley, CA: Apress; Springer Science+Business Media; 2017.

[12] MeteoViva GmbH. Funktionsweise. [March 14, 2017]; Available from: http://www.meteoviva.com/kennenlernen/leistungen/meteovivar-climate/funktionsweise/.

[13] BuildingIQ. Predictive Energy Optimization. [March 23, 2017]; Available from: https://buildingiq.com/products/predictive-control/buildingiq-platform/.

[14] tado. So steuert tado° die Heizung. [March 28, 2017]; Available from: https://www.tado.com/de/thermostat-heizung.

[15] Nest Labs. Thermal Model and HVAC Control White Paper; 2016; Available from: https://nest.com/downloads/press/documents/thermal-model-hvac-white-paper.pdf. [March 06, 2017].

[16] MeteoViva GmbH. Praxisberichte. [March 14, 2017]; Available from: http://www.meteoviva.com/kennenlernen/referenzen/praxisberichte/.

[17] Rawlings JB, Mayne DQ. Model predictive control: Theory and design. Madison, Wis.: Nob Hill Publ; 2009.

[18] Fanger PO. Thermal comfort: analysis and applications in environmental engineering. New York: McGraw-Hill; 1970.

[19] Gao PX, Keshav S. SPOT: A Smart Personalised Office Thermal Control System. UWSpace 2013.

[20] Erickson VL, Cerpa AE. Thermovote: Participatory Sensing for Efficient Building HVAC Conditioning. BuildSys 2012;'12.

[21] Schumann A, Wilson N, Burillo M. Learning User Preferences to Maximise Occupant Comfort in Office Buildings. Berlin, Heidelberg: Springer; 2010.

[22] Schmidt F. Nutzerkomfort. Karlsruher Institut für Technologie; 2016.

[23] Holopainen R, Tuomaala P, Hernandez P, Häkkinen T, Piira K, Piippo J. Comfort assessment in the context of sustainable buildings: Comparison of simplified and detailed human thermal sensation methods. Building and Environment 2014;71:60–70.

[24] Wagner A. Thermal comfort and air quality requirements - impact on design strategies and HVAC concepts. Karlsruher Institut für Technologie; 2014.

[25] Rietschel H. Raumklimatechnik. 16th ed. Berlin, Heidelberg: Springer; 2008.

[26] American Societey of Heating, Refigerating and Air-Conditioning Engineers. Physiological principles and thermal comfort(19938.1-8.32). Atlanta: ASHRAE.

[27] Deutsches Institut für Normung. Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics(EN 15251:2012-12). Berlin: Beuth Verlag GmbH; 2012.

[28] ISO International Organization for Standardization. Ergonomics of the thermal environment - Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria(ISO 7730:2006-05).

[29] Nicol F, Humphreys M. Derivation of the adaptive equations for thermal comfort in free-running buildings in European standard EN15251. Building and Environment 2010;45(1):11–7.

[30] Yao R, Li B, Liu J. A theoretical adaptive model of thermal comfort – Adaptive Predicted Mean Vote (aPMV). Building and Environment 2009;44(10):2089–96.

[31] Ng A. Machine Learning; 2017.

[32] Paluszek M, Thomas S. MATLAB Machine Learning. 1st ed. Berkeley, CA: Apress; Springer Science+Business Media LLC; 2017.

[33] Kwok JT, Zhou Z-H, Xu L. Machine Learning. Springer Handbook of Computational Intelligence 2015:495–522.

[34] Dalamagkidis K, Kolokotsa D, Kalaitzakis K, Stavrakakis GS. Reinforcement learning for energy conservation and comfort in buildings. Building and Environment 2007;42(7):2686–98.

[35] Kubat M. An Introduction to Machine Learning. Cham: Springer; Springer International Publishing; 2015.

[36] The MathWorks Inc. Matlab; 2017.

[37] Yu Z, Dexter A. Online tuning of a supervisory fuzzy controller for low-energy building system using reinforcement learning. Control Engineering Practice 2010;18(5):532–9.

[38] Sutton RS, Barto AG. Reinforcement Learning: An Introduction. 2nd ed. Cambridge, Massachusetts: The MIT Press; 2012.

[39] Gomes C. Intelligent Energy Management Systems in Office Buildings. Lisbon: Instituto Superior Técnico, Universidade de Lisboa.

[40] Dounis AI, Caraiscos C. Advanced control systems engineering for energy and comfort management in a building environment—A review. Renewable and Sustainable Energy Reviews 2009;13(6-7):1246–61.

[41] Sun Y, Huang G. Recent Developments in HVAC System Control and Building Demand Management. Curr Sustainable Renewable Energy Rep 2017;4(1):15–21.

[42] Liang J, Du R. Design of intelligent comfort control system with human learning and minimum power control strategies. Energy Conversion and Management 2008;49(4):517–28.

[43] Gao Y, Menzel K (eds.). Using data mining in optimisation of building energy consumption and thermal comfort management; 2010.

[44] Barrett E, Linder S. Autonomous HVAC Control, A Reinforcement Learning Approach. Lecture Notes in Computer Science 2015;9286:3–19.

[45] Ruano A, Pesteh S, Silva S, Duarte H, Mestre G, Ferreira PM et al. PVM-based intelligent predictive control of HVAC systems. IFAC-PapersOnLine 2016;49(5):371–6.

[46] Zhao H-x, Magoulès F. A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews 2012;16(6):3586–92.

[47] Edwards RE, New J, Parker LE. Predicting future hourly residential electrical consumption: A machine learning case study. Energy and Buildings 2012;49:591–603.

[48] Atthajariyakul S, Leephakpreeda T. Neural computing thermal comfort index for HVAC systems. Energy Conversion and Management 2005;46(15-16):2553–65.

[49] Liu W, Lian Z, Zhao B. A neural network evaluation model for individual thermal comfort. Energy and Buildings 2007;39(10):1115–22.

[50] Mamidi S, Chang Y-H, Maheswaran R (eds.). Improving building energy efficiency with a network of sensing, learning and prediction agents; 2012.

[51] Buratti C, Vergoni M, Palladino D. Thermal Comfort Evaluation Within Non-residential Environments: Development of Artificial Neural Network by Using the Adaptive Approach Data. Energy Procedia 2015;78:2875–80.

[52] Ferreira PM, Ruano AE, Silva S, Conceição E. Neural networks based predictive control for thermal comfort and energy savings in public buildings. Energy and Buildings 2012;55:238–51.

[53] Aswani A, Master N, Taneja J, Krioukov A, Culler D, Tomlin C (eds.). Energy-Efficient Building HVAC Control Using Hybrid System LVMPC; 2012.

[54] Ferreira PM, Silva SM, Ruano AE (eds.). Energy Savings in HVAC Systems Using Discrete Model-Based Predictive Control; 2012.

[55] Erickson VL, Cerpa AE (eds.). Occupancy Based Demand Response HVAC Control Strategy; 2010.

[56] Tso GK, Yau KK. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. Energy 2007;32(9):1761–8.

[57] Afram A, Janabi-Sharifi F, Fung AS, Raahemifar K. Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system. Energy and Buildings 2017;141:96–113.

[58] Matsumoto M, Masui K, Fukushige S, Kondoh S (eds.). Sustainability Through Innovation in Product Life Cycle Design. Singapore: Springer Japan; 2017.

[59] Tsanas A, Xifara A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy and Buildings 2012;49:560–7.

[60] Yang R, Newman MW (eds.). Learning from a learning thermostat: lessons for intelligent systems for the home. New York, NY, USA: ACM; 2013.

[61] Liang J, Du R (eds.). Thermal Comfort Control Based on Neural Network for HVAC Application; 2005.

[62] Nest Labs. Nest Thermostat. [March 14, 2017]; Available from: https://nest.com/thermostat/meet-nest-thermostat/.

[63] Heat Genius Ltd. Genius home. [March 14, 2017]; Available from: https://www.geniushub.co.uk/how-it-works/.

[64] Honeywell International Inc. evohome Wi-Fi. [March 14, 2017]; Available from: https://getconnected.honeywell.com/de/thermostate/evohome.

[65] ecobee. ecobee3. [March 28, 2017]; Available from: https://www.ecobee.com/ecobee3/.

[66] Werner M. Verfahren zur Steuerung des Klimas in einem wetterabhängigen Gebäude- oder Anlagenbereich(EP 1134508 B1); 2006.

[67] St. John J. BuildingIQ's CLoud Takes an 'All of the Above' Approach to Energy Management. [March 23, 2017]; Available from: https://www.greentechmedia.com/articles/read/buildingiqs-cloud-offering-takes-all-of-the-above-approach-to-energy-manage.

[68] Korporaal G. Sky-high hopes for cloud-based BuildingIQ's ASX listing. The Australian 23/17/2015, 23/17/2015; Available from: http://www.theaustralian.com.au/business/companies/skyhigh-hopes-for-cloudbased-buildingiqs-asx-listing/news-story/b824906361b3853ab1032eb534cfafac. [March 23, 2017].

[69] The MathWorks Inc. BuildingIQ Develops Proactive Algorithms for HVAC Energy Optimization in Large-Scale Buildings. [March 23, 2017]; Available from: https://de.mathworks.com/company/user_stories/buildingiq-develops-proactive-algorithms-for-hvac-energy-optimization-in-large-scale-buildings.html.

[70] Welch S, CrunchBase Staff. BuildingIQ. [March 23, 2017]; Available from: https://www.crunchbase.com/organization/buildingiq/contributors.

[71] Donath A. Google Nest kommt in deutsche Wohnzimmer. [March 14, 2017]; Available from: https://www.golem.de/news/hausautomatisierung-google-nest-kommt-in-deutsche-wohnzimmer-1701-125609.html.

[72] Wyon DP. Assessment of human thermal requirements in the thermal comfort region. Proceedings Thermal Comfort, Past, Present and Future. Garston:144–56; 1994.

[73] Trimble Navigation Ltd. SketchUp; 2016.

[74] Alliance for Sustainable Energy LLC. OpenStudio; 2017.

[75] National Renewable Energy Laboratory. EnergyPlus; 2017.
[76] Deutscher Wetterdienst. CDC (Climate Data Center). Offenbach; Available from: ftp://ftp-cdc.dwd.de/pub/CDC/. [06.13.2017].
[77] The MathWorks Inc. Produkte & Dienstleistungen. [April 28, 2017]; Available from: https://de.mathworks.com/products.html?s_tid=gn_ps.
[78] Bundestag of the Federal Republic of Germany. Verordnung über energiesparenden Wärmeschutz und energiesparende Anlagentechnik bei Gebäuden: Energieeinsparverordnung (EnEv); 2002.
[79] National Renewable Energy Laboratory. Energy Plus Version 8.7 Documentation: Input Output Reference; 2016; Available from: https://energyplus.net/sites/all/modules/custom/nrel_custom/pdfs/pdfs_v8.7.0/InputOutputReference.pdf. [July 12, 2017].
[80] Tritschler M, Hahn J, Krockenberger B, Grob R, Luft C. Abschlussbericht des Forschungsvorhabens Energieoptimiertes Bauen: Monitoring und Betriebsoptimierung der Kreissparkasse Göppingen. Esslingen, Germany; 2014.
[81] McKenna R. Energy Efficiency in Buildings. Karlsruhe; 2016.
[82] Lobo S. Keynote: Künstliche Intelligenz. Classic Remise Berlin; 2017.
[83] Domingos T. MEET 2030: Energia, Clima e Crescimento Económico: Oportunidades de Negócio em Portugal; 2016.
[84] Eberle D. Das Haus ohne Heizung (2226). Classic Remise Berlin; 2017.

Annex

1. PMV-Calculation [18]

$$PMV(M, W, p_a, t_a, \overline{t_r}, t_{cl}, f_{cl}, h_c) = (0.303 \cdot e^{-0.036 \cdot M} + 0.028) \cdot \big[(M-W) - 3.05 \cdot 10^{-3} \cdot (5733 - 6.99 \cdot (M-W) - p_a) - 0.42 \cdot \big((M-W) - 58.15\big) - 1.7 \cdot 10^{-5} \cdot M \cdot (5867 - p_a) - 0.0014 \cdot M \cdot (34 - t_a) - 3.96 \cdot 10^{-8} \cdot f_{cl} \cdot \big((t_{cl}+273)^4 - (\overline{t_r + 273})^4\big) - f_{cl} \cdot h_c \cdot (t_{cl} - t_a)\big] \quad \text{(A1)}$$

With M … Metabolic Rate

W … Effective Mechanical Power (0 for most indoor activities)

p_a … Humidity level

t_a … Air temperature

$\overline{t_r}$ … Mean radiant temperature

t_{cl} … Clothing surface temperature

f_{cl} … Clothing surface area factor

h_c … Convective heat transfer coefficient

The clothing surface temperature is calculated according to

$$t_{cl} = 35.7 - 0.028 \cdot (M-W) - I_{cl} \cdot (3.96 \cdot 10^{-8} \cdot f_{cl} \cdot [(t_{cl}+273)^4 - (\overline{t_r} + 273)^4) + f_{cl} \cdot h_c \cdot (t_{cl} - t_a)] \quad \text{(A2)}$$

With I_{cl} … clothing insulation

The convective heat transfer coefficient is calculated according to

$$h_c = \begin{cases} 2.38 \cdot |t_{cl} - t_a|^{0,25} \; if \; 2.38 \cdot |t_{cl} - t_a|^{0.25} > 12.1 \cdot \sqrt{v_{ar}} \\ 12.1 \cdot \sqrt{v_{ar}} \; if \; 2.38 \cdot |t_{cl} - t_a|^{0.25} < 12.1 \cdot \sqrt{v_{ar}} \end{cases} \quad \text{(A3)}$$

With v_{ar} … Air velocity

The clothing surface area factor is calculated according to

$$f_{cl} = \begin{cases} 1.00 + 1.290 I_{cl} \; if \; I_{cl} \leq 0.078 \frac{m^2 \cdot K}{W} \\ 1.05 + 0.645 I_{cl} if \; I_{cl} > 0.078 \frac{m^2 \cdot K}{W} \end{cases} \quad \text{(A4)}$$

The metabolic rate and clothing insulation are taken from tables in application. The clothing insulation, Equation A4, is then used to calculate the clothing surface temperature Equation A2 and the convective heat transfer coefficient Equation A3 through iteration.

2. Values for decision Top 10 Neural Networks

*Table **9** Annex 2: Values for decision Top 10 Neural Networks*

Parameter Nodes Combination	Median	Minimum	Maximum	Lower Quartile	Upper Quartile
25679 - 33	0.94	0.50	1.76	0.71	1.33
245789 - 33	0.95	0.49	1.76	0.70	1.35
2349 - 33	0.96	0.52	1.76	0.72	1.37
24589 - 33	0.96	0.48	1.79	0.70	1.34
234789 - 33	0.96	0.49	1.77	0.71	1.36
256789 - 33	0.97	0.49	1.86	0.70	1.32
236789 - 33	0.97	0.50	1.69	0.71	1.32
23689 - 33	0.97	0.47	1.80	0.72	1.35
23479 - 33	0.97	0.51	1.76	0.71	1.40
23489 - 33	0.97	0.52	1.77	0.72	1.35

3. Values for decision Top 10 Regression Learning Tools

*Table **10** Annex 3: Values for decision Top 10 Regression Learning Tools*

Parameter Tool Combination	Median	Minimum	Maximum	Lower Quartile	Upper Quartile
245789 - GPR Exponential	1.075	0.59	1.98	0.825	1.49
234789 - GPR Matern 5/2	1.085	0.6	1.98	0.83	1.51
234789 - GPR Exponential	1.085	0.6	2	0.83	1.5
245789 - GPR Rational Quadratic	1.085	0.6	2.01	0.84	1.5
245789 - GPR Matern 5/2	1.085	0.59	1.99	0.82	1.495
256789 - GPR Rational Quadratic	1.085	0.6	1.99	0.83	1.5
256789 - GPR Matern 5/2	1.085	0.59	2.01	0.835	1.49
256789 - GPR Exponential	1.085	0.59	2.01	0.835	1.485
234789 - GPR Rational Quadratic	1.085	0.59	2.02	0.84	1.49
2349 - Ensemble Bagged Trees	1.09	0.6	2.04	0.825	1.515