DEVELOPMENT OF A MODEL TO OPTIMIZE THE ENERGY EFFICIENCY OF RESIDENTIAL BUILDING AND THEIR IMPACT ON THE LOW VOLTAGE GRID

Michael DAHMS
amperias GmbH – Germany
Michael.dahms@amperias.com

Torsten SOWA
amperias GmbH – Germany
Torsten.sowa@amperias.com

ABSTRACT

The consequences of the energy transition are causing a far-reaching change in the electricity supply system. To analyse these changes, load scenarios for future power grids are essential. The largest energy consumption is within the fast-growing building sector, which has therefore a large impact on these scenarios. This paper presents an approach to optimize load demand scenarios and possible expansions regarding the energy efficiency of buildings. The optimization problem is modelled bottom-up and solved using a neural network, which is trained with reinforcement learning. Two different training methods, iterative or incremental, are evaluated. Exemplary results demonstrate that the evaluated reinforcement learning agents are able to achieve a significant cost reduction of the considered buildings through efficiency measures. The evaluated training methods show that Proximal Policy Optimisation results in a higher cost reduction than Advantage Actor Critic optimisation and show a more robust performance over the different complexities.

INTRODUCTION

The current climate change as well as the latest political gas conflict demonstrate the high impact of the energy sector and point out how indispensable a change towards more sustainability and renewable energies is. The consequences of the energy transition are causing a profound change in the power supply system, making future grid scenarios difficult to predict. The sustainability of the energy supply and the improvement of energy efficiency in the building sector, where the integration of sustainable energy technologies is required, are key factors. The building sector is responsible for about 40% of end-use energy consumption and 36% of all CO2 emissions in Europe [1].
European climate targets include a near-climate-neutral building stock by 2050, with an 80% reduction in primary energy demand and the generation of at least 50% of energy from renewable sources to achieve this [2]. With a projected doubling of the current international building stock up to that point as well as the current EU renovation wave of three percent, meeting these targets is a major challenge [3]. As part of the Climate Protection Program 2030, funding programs for energy-efficient construction and renovation have therefore been increased, which has led to a significant increase in demand for subsidized energy consulting for buildings in recent years [1].
New ideas and methodologies are needed to realize this change in the building sector for maintaining climate policy goals. In addition to structural changes, building independent upgrades such as the use of photovoltaic (PV) systems and heat pumps increase the energy efficiency of a building and thus represent an opportunity for overall improvement in the energy performance of the building. The consequences resulting from these changes in the building sector have a large impact on the electricity grid, especially on the low and medium voltage grids [4]. To estimate the consequences for the power infrastructure more accurately, the buildings are modelled with a bottom-up approach. The model optimizes the investment decision by reducing energy consumption of buildings through structural changes and building-independent upgrades. This model is solved with a neural net, which is trained by reinforcement learning.

METHODOLOGY

This paper presents a modelling approach for a bottom-up-modelling of building to determine their impact on the low and medium voltage grid (*Figure 1*). The following terminologies are used:

- A component *B* is a structural part of a building with a constant heat transfer coefficient (e.g. a wall).
- An object *N* consists of at least one component.
- A virtual building is the combination of several objects to form a complete representation of a building.
- A building independent expansion is the installation of DER such as PV.

In this approach, a building consists of several objects (e.g. walls, doors), which can be modified with e.g. insulation. The investment decision on the expansions is based on the minimalization of the overall costs of heating, electricity and expansion (see chapter *Virtual building with extension options*). The determination of the optimal solution for each building is solved as a combinatorial optimisation problem (see chapter *Defining the expansion decisions as a combinatorial optimization problem*), which are solved using neural networks (NN). The NN are trained with reinforcement learning (RL) (see chapter *Reinforcement learning*). In contrast to a probability-based approach, the chosen method does not require the determination of probabilities of possible grid states. Therefore, the chosen method can estimate the impact on distribution grids for a wide variety of scenarios for e.g. different technology and cost developments under great uncertainty.

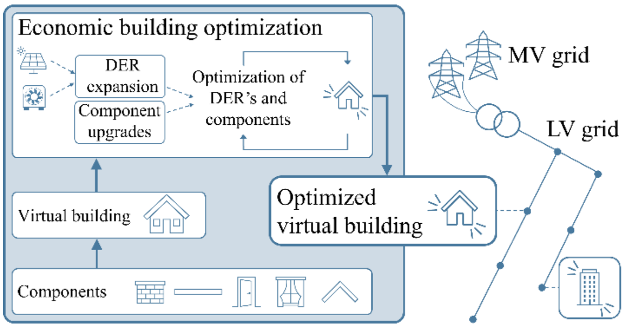

Figure 1: Overview of the bottom-up model of a virtual building which is divided into different components and optimized considering decentralized energy resources (DERs).

Energy Efficiency Optimization Model

The model developed in this paper aims to identify expansion decisions that optimise the energy costs of a building. Data of the building and its environment (e.g. weather data) are used as input parameters for the building model. To optimise buildings with regards to energy efficiency, the optimal extension is depicted as a combinatorial optimization problem. The optimization problem is solved with a trained NN.

Figure 2 shows an overview of the model. The building is modelled using a bottom-up design. Therefore, this model is flexible with regards to input data (e.g. weather conditions, insulation on door and windows).

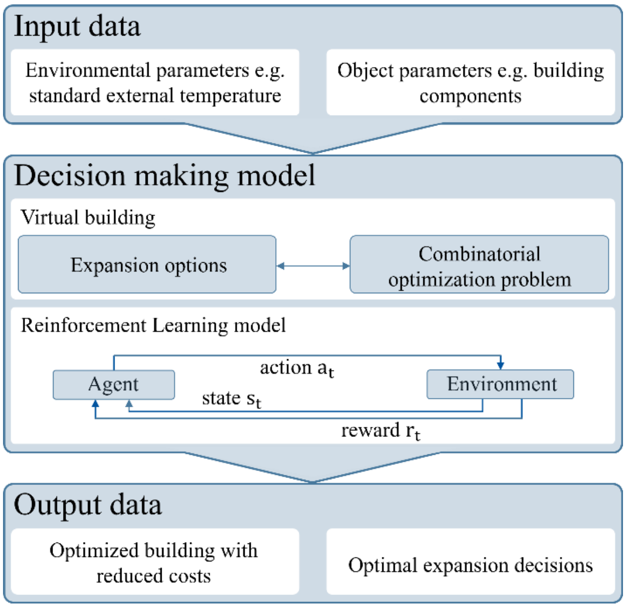

Figure 2: Model overview: With the input data set, the decision model consisting of different expansion options formulated as a combinatorial optimization problem generates results using a reinforcement learning approach. The results represent optimized buildings with reduced energy consumption and the associated optimal expansion decisions.

To reduce the complexity of the bottom-up optimization, a clustering of the building components is used for a dimensional reduction, which is advantageous with increasing building complexity. The neural network learns how to optimise the building by varying structural changes (e.g. additional insulation, replacement of windows) and building-independent expansions (use of decentralized energy resources (DER), e.g. new construction of a PV system or heat pump). The results represent optimised buildings with reduced energy consumption and the associated optimal expansion decisions.

Virtual building with extension options

Component upgrades (e.g. windows, walls, doors) exist for each component B. Changes in the heat transfer coefficient and thus the transmission heat loss (THL) can be made in the building components through expansion. The THL is directly related to the energy demand of a building component, so changes have a direct impact on the absolute energy demand. The expansion of a component causes investment costs $C^{INVEST_{COM}}$, which depend on the properties of the component, such as the surface area, and the selected expansion option. In the case of a combination of several components, the THL and the costs result from the sum of the THL and upgrade costs of the individual components. For building modeling, the building shell method is used, according to which the heating load can be determined from the sum of transmission and ventilation heat losses [5] [6]. After combining several components into buildings, DER expansion options can be perceived. In contrast to the component upgrades, no influence is exerted on the absolute energy demand, but the self-consumption is changed. This expansion causes further investment costs $C^{INVEST_{DER}}$. As a result, the total costs of an optimized object N can be calculated as

$$C^{INVEST}(N) = \sum_{e \in E} C_e^{INVEST_{DER}} + \sum_{b \in B} C_b^{INVEST_{COM}} \quad (1)$$

The energy costs $C^{TIME}(N)$ determined for a fixed period enables a comparison of different components. Due to the direct correlation of the required electrical energy to heat energy costs, a cost reduction also leads to an increased energy efficiency of an object. Consequently, the search for a combination of measures for maximum cost reduction represents the objective function.

Defining the expansion decisions as a combinatorial optimization problem

Let $J = \{j_1, j_2, \ldots, j_n\}$ be the set of all permissible combinations of expansion measures. The perception of a combination j determines the condition of an object and its costs. The total costs of an object are defined as follows:

$$C(N_j) = C^{INVEST}(N_j) + C^{TIME}(N_j) \quad (2)$$

The aim is to find an optimal combination $j^{opt} \in J$ such that $C(O_{j^{opt}}) \leq C(O_j)$ for all $j \in J$. For that the set of all

combinations must first be determined. The cardinality of this set $|J|$ depends on the complexity of the object. Each component $b \in B$ is the instance of a class $k \in K$. Let B_k be the subset of all components B of type k. A limited number of upgrades is defined for each class.
Consequently, for each component instance $b_k \in B_k$ there exists a finite set of upgrades I_k. For each of these options $i \in I_k$ there is a set of potential actions M_i. The number of combinations for the component upgrades is calculated as follows:

$$\prod_{k \in K} \prod_{i \in I_k} |M_i|^{|B_k|} \tag{3}$$

The DER expansion E can be modelled analogously. For each expansion $e \in E$, a limited number of expansion options is given. The set M_e describes these available measures, so that the number of combinations results in:

$$\prod_{e \in E} |M_e| \tag{4}$$

The product of the number of combinations for the component upgrades and DER expansion gives the total number of all combinations $|J|$.

$$|J| = \prod_{k \in K} \prod_{i \in I_k} |M_i|^{|B_k|} * \prod_{e \in E} |M_e| \tag{5}$$

For each set M a measure $m_0 \in M$ is defined, which describes the original characteristic of an object. For all expansion options $|M| \leq 2$.
The complexity of the number of possible combinations therefore grows exponentially with the number of components and DER expansion. Let $n = |E| + |B|$ be this number. The use of a brute-force algorithm f to solve this problem leads to a complexity of $f \in O(2^n)$, which is consequently not solvable in polynomial time. In the context of this paper, reinforcement learning will therefore be used to solve the problem [7].

Reinforcement learning

To use reinforcement learning, the problem must be formulated as a Markov Decision Problems (MDP). An agent interacts with an object that change the state of the object and receives a reward that depends on the costs incurred. The environment describes the object itself and its environmental parameters (*Figure 2*).
There are different methods to solve combination problems with RL [8]. In the incremental method, the agent carries out individual expansion measures one after the other. An action is therefore equivalent to the choice of exactly one expansion measure [9] [10]. The action space $|A^{INK}|$ is discrete.
With the iterative method, an action represents a complete expansion decision for the entire object. In this case, a measure is assigned to each expansion option. One possible combination $j \in J$ is selected per timestep. The action space A_{IT} is therefore identical with the set of all combinations J.

The agent's reward R is identical in the incremental and iterative methods and is defined by the negative cost of an object regarding demand and invest. Due to the discrete solution spaces, the algorithms that can be considered for this optimisation problem are limited. Algorithms that operate exclusively in continuous action spaces (e.g. Deep Deterministic Policy Gradient (DDPG)) cannot be applied to the problem. Furthermore, the cardinality of the action spaces is decisive for the selection of the methods and policed-based methods are well adapted for proportionally increasing decision amounts. This increases proportionally for both incremental and iterative methods.
Therefore, the use of policy-based methods promises the greatest success. Compared to value-based methods, the exact value of all actions does not need to be known to identify the best action. In related research [11] [12] [13] [14], policy gradient methods have already been successful in solving combinatorial optimisation problems. Due to high stability and generalisation possibilities for further development steps, which are essential in the implementation of bottom-up design, a Proximal Policy Optimisation (PPO) algorithm is used for training the agents and an Actor Critic method with an advantage function (A2C) is used, as shown in *Figure 3* [15].

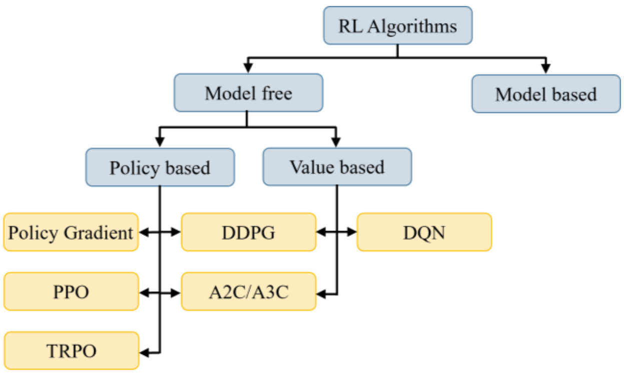

Figure 3: Overview and categorization of selected RL algorithms.

RESULTS

Several agents are trained in incremental (Inc) and iterative (It) design with PPO and A2C. During training, an agent goes through several decision processes for which actions must be selected. The length L of a decision process corresponds to the number of expansion decisions to be made. For each decision process, a new randomised building is generated. For training all variable and constant parameters of the object and its environment are randomly chosen within predefined limits. Several different architectures of neural networks are trained for the actor π_θ and critic ϑ_φ. The feedforward neural networks differ in their number and size of layers. After agent training, they are tested on generated buildings, where the evaluation objects are created using the same principle as the training objects with randomly chosen parameters.
The costs of the individual components are summed up and

compared for different agents and network architectures and serve as the main evaluation criterion. In the following, different training and test designs are presented, which differ in the complexity (K1-K3) of the considered objects:

- Optimisation of single object (K1)
- Optimisation of two individual objects (K2)
- Optimisation of assemblies consisting of six objects (K3)

The objects are considered within a constant time period with a constant electricity price*.
The following metadata were used:

- $|J|$: Number of possible combinations
- $|A^{INC}|$ or $|A^{IT}|$: Size of the action spaces for incremental or iterative design
- L: Length of a decision process
- n: Number of actions executed in the training
- n_O: Number of objects for evaluation
- KNN: Network architectures of the actor and the critic specifying the layers

The results of the objects with $|B| = 2$ are shown below in Table 1. The number of possible combinations of expansion options is 3,2e^5. The neural networks of the actor and critic were trained in 20.000 timesteps and evaluated on 10.000 components. Within a decision process, six actions are executed.

Table 1: Algorithm Metadata of the executions in K2.

$\|J\|$	$\|A^{INC}\|$	$\|A^{IT}\|$	L	n	n_O	KNN
3,2e^5	50	3,2e^5	6	2e^5	1e^4	32x32 64x64 128x128 512x512

For objects with $|B| \leq 2$, the calculation of the optimal costs is still feasible. *Figure 4* compares the results of different network architectures, aggregating the costs of 10 components. The average value of all trained agents is 1.05e^7 €, of the random agent 6.86e^7 € and the optimal value 3.97e^6 €. Across all networks, the PPO agent has shown the most significant cost reduction.

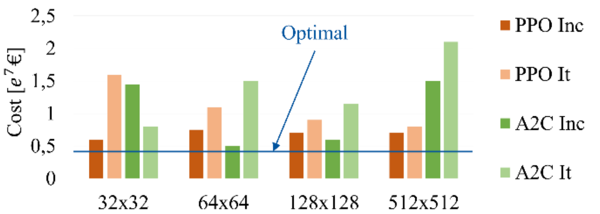

Figure 4: Overview of the results of the different KNNs in K2.

To illustrate the sum of the individual object costs, *Figure 5* shows the costs of the three complexity categories (K1, K2, K3) as an example. In addition, the calculated optimal costs are given as a comparison. The training results are agents with a network architecture with 64x64 layers.
As an example, K2 can be mentioned here again: With this architecture, the best result is achieved with an A2C agent in incremental design. The cost of the incremental A2C agent, shown in *Figure 5* (top right), is 8,5 % of the cost obtained by random actions of an untrained agent. The optimal cost is 5,8 % of the cost of the random agent. This corresponds to 1,9 % (A2C Ink) and 1,3 % (Optimal) of the original cost.

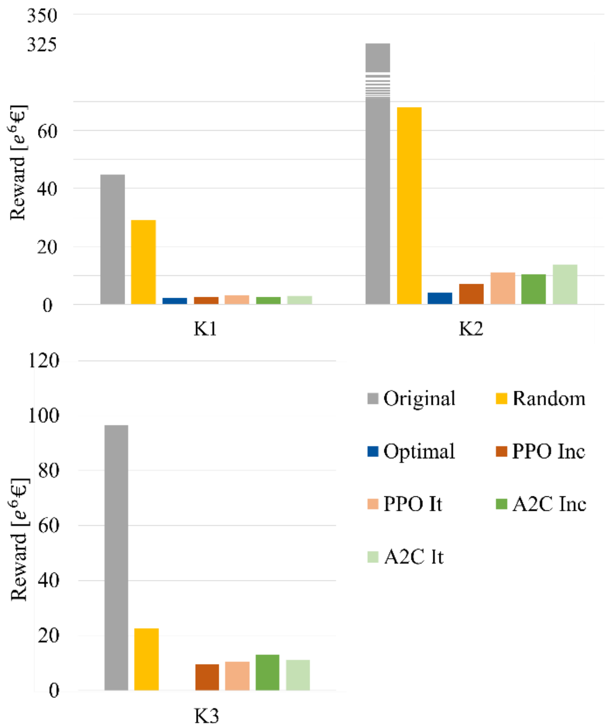

Figure 5: Cost and average reward from the agents (64x64 KNN) compared to original and optimal values for K1, K2 (top) and K3 (bottom).

For all complexity categories, reducing costs and thus increasing energy efficiency can be achieved in the objects. All agents can achieve a learning outcome and better results in each training cycle than a random agent The selected reinforcement learning methods are therefore suitable for solving the combination problem described. Across all training and test iterations, the PPO agent trained in incremental design shows the most stable and reliable results for reducing costs. It achieves both the lowest mean value and the lowest absolute values in two out of three complexity categories. In general, agents using the PPO algorithm are more successful in optimising the objective function than agents trained with A2C.
For both agents, lower costs can be achieved using an incremental design. The reason for this is the high cardinality of the action spaces in the iterative design, whereby an agent only gains experience for a limited number of actions while training, which is a cause for the reduced performance of the iterative agents.

* 10 years and 0,32 €/kWh

CONCLUSION

The aim of the work is to identify and evaluate reinforcement learning methods for improving the energy properties of buildings through structural changes. For this purpose, a combination problem was formulated as an MDP regarding optimal extension decisions in order to test different reinforcement learning architectures. The presented experiments are related to structural expansions of different complexity. In the evaluation, A2C and PPO algorithms were used to train RL agents of different designs to minimise a cost function. The results can be summarised as follows:

- The evaluated RL agents were able to achieve a significant cost reduction of the considered objects.
- Agents using the PPO algorithm are more successful in minimising the objective function than A2C agents and are able to show more stable behaviour across complexities and designs.

Therefore, the chosen methods are suitable for decision making to improve the energy efficiency of the building.
This model depicts the basics for a comprehensive approach to analyse the impact of different technologies on the distribution grids. Further research will focus on continuing the model with load flow calculations.

Furthermore, a combination of RL methods with other algorithms will be developed. These algorithms can be developed that exclude actions and combinations to reduce the action sparce that would, for example, lead to an increase in costs. By clustering similar components, a general dimension reduction can also be made.

ACKNOWLEDGMENTS

This project received funding from the German Federal Ministry of Economic Affairs and Climate Action under the funding No. 03EI6053B.

Federal Ministry for Economic Affairs and Climate Action

REFERENCES

[1] S. Becker, J. Hagen, R. Krüger et A. Exner, DENA-Gebäudereport 2023, Berlin: Deutsche Energie-Agentur GmbH (dena), 2022.

[2] D. V. Bürger, D. T. Hesse, D. A. Palzer, B. Köhler, S. Herkel, D. P. Engelmann et D. D. Quack, Klimaneutraler Gebäudebestand 2050: Energieeffizienzpotenzialeund die Auswirkungen des Klimawandels auf den Gebäudebestand, Freiburg im Breisgau: Umweltbundesamt, 2017.

[3] . e. V. Bundesverband der Energie- und Wasserwirtschaft, «BDEW Energie. Wasser. Leben.,» 09 Januar 2023. [En ligne]. Available: https://www.bdew.de/service/daten-und-grafiken/bdew-strompreisanalyse/.

[4] R. Jackson, E. Zhou et J. Reyna, «Building and grid system benefits of demand flexibility and energy efficiency,» *Joule Previews,* pp. 1921 - 1933, 18 08 2021.

[5] DIN e. V. (Hrag.), DIN TS 12831-1: Verfahren zur Berechnung der Raumheizlast, Berlin: Berlin, 2020.

[6] DIN e. V. (Hrag.), DIN EN 12831-1: Energetische Bewertung von Gebäuden –Verfahren zur Berechnung der Norm-Heizlast, Berlin: Berlin, 2017.

[7] P. Gritzmann, Grundlagen der Mathematischen Optimierung: Diskrete Strukturen, Komplexitätstheorie, Konvexitätstheorie, Lineare Optimierung, Simplex-Algorithmus, Dualität, Wiesbaden: Springer Vieweg, 2013.

[8] P. D. E. Alpaydin, Maschinelles Lernen, Berlin, Boston: De Gruyter Oldenbourg, 2019.

[9] N. Mazyavkina, S. Sviridov, S. Ivanov et E. Burnaev, Reinforcement Learning for Combinatorial Optimization: A Survey, Lisboa: Computers & Operations Research Volume 134, 2021.

[10] J. Schulman, F. Wolski, P. Dhariwal, A. Radford et O. Klimov, «Proximal Policy Optimization Algorithms,» *Computer Science,* p. eprint arXiv:1707.06347, 2017.

[11] S. Kim et H. Lim, «Reinforcement Learning Based Energy Management Algorithm for Smart Energy Buildings,» School of Electrical Engineering and Computer Science, Gwangju, 2019.

[12] K. Mason et S. Grijalva, «A review of reinforcement learning for autonomous building energy management,» *Computers & Electrical Engineering,* pp. 300-312, 09 2019.

[13] A. Perera et P. Kamalaruban, «Applications of reinforcement learning in energy systems,» Urban Energy Systems Laboratory, Dübendorf, 2021.

[14] M. Rose, L. Lenz, T. Sowa et I. Hebbeln, «Planning LV grids by predicting residual loads of households via methods of machine learning,» chez *CIRED 2020 Berlin Workshop*, Berlin, 2020.

[15] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver et K. Kavukcuoglu, «Asynchronous Methods for Deep Reinforcement Learning,» chez *The 33rd International Conference on Machine Learning*, New York, 2016.