Applied Energy 156 (2015) 577–586

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Reinforcement learning for optimal control of low exergy buildings

Lei Yang [a], Zoltan Nagy [a,*], Philippe Goffin [b], Arno Schlueter [a]

[a] Architecture and Building Systems, ETH Zürich, John-von-Neumann-Weg 9, 8093 Zürich, Switzerland
[b] BuSy, Building Systems, ETH Zürich, John-von-Neumann-Weg 9, 8093 Zürich, Switzerland

HIGHLIGHTS

- Implementation of reinforcement learning control for LowEx Building systems.
- Learning allows adaptation to local environment without prior knowledge.
- Presentation of reinforcement learning control for real-life applications.
- Discussion of the applicability for real-life situations.

ARTICLE INFO

Article history:
Received 19 February 2015
Received in revised form 25 June 2015
Accepted 15 July 2015

Keywords:
Low exergy building systems
Zero net energy buildings
Reinforcement learning control
Energy efficient buildings
Sustainable building systems

ABSTRACT

Over a third of the anthropogenic greenhouse gas (GHG) emissions stem from cooling and heating buildings, due to their fossil fuel based operation. Low exergy building systems are a promising approach to reduce energy consumption as well as GHG emissions. They consists of renewable energy technologies, such as PV, PV/T and heat pumps. Since careful tuning of parameters is required, a manual setup may result in sub-optimal operation. A model predictive control approach is unnecessarily complex due to the required model identification. Therefore, in this work we present a reinforcement learning control (RLC) approach. The studied building consists of a PV/T array for solar heat and electricity generation, as well as geothermal heat pumps. We present RLC for the PV/T array, and the full building model. Two methods, Tabular Q-learning and Batch Q-learning with Memory Replay, are implemented with real building settings and actual weather conditions in a Matlab/Simulink framework. The performance is evaluated against standard rule-based control (RBC). We investigated different neural network structures and find that some outperformed RBC already during the learning phase. Overall, every RLC strategy for PV/T outperformed RBC by over 10% after the third year. Likewise, for the full building, RLC outperforms RBC in terms of meeting the heating demand, maintaining the optimal operation temperature and compensating more effectively for ground heat. This allows to reduce engineering costs associated with the setup of these systems, as well as decrease the return-of-invest period, both of which are necessary to create a sustainable, zero-emission building stock.

1. Introduction

Residential and commercial buildings account for nearly 40% of energy consumption and for 36% of greenhouse gas emissions [1]. Particularly, the energy needed for heating, which takes almost half of the energy used in building systems [2], is commonly supplied as electricity/fossil energy carrier with low efficiency and high greenhouse gas emissions [3]. *Low exergy (LowEx) buildings* are one promising approach to tackle both energy conservation and CO_2 reduction challenges [4]. By *LowEx* buildings, we understand buildings whose ratio of consumed energy to distributed energy, such as heat, is small [5]. This can be achieved using heat pumps with sufficiently high coefficients of performance (COP). It follows that buildings with low energy consumption e.g., passive or net-zero-energy buildings, but powered using a burning process cannot be considered as LowEx buildings. Combining the heat pump with photovoltaic-thermal (PV/T) modules allows to heat the building exclusively from renewable energy sources, making it emission-free.

In this research, we study the LowEx building shown in Fig. 1, located in a residential area in Zurich, Switzerland, in an oceanic climate (Köppen *Cfb*) [5,6]. The window facade ratio is 25%, and the construction is a Misapor facade with a *U*-value of 0.24 $W/m^2/K$. The windows have *U*-values of 1 $W/m^2/K$, *g*-values of 0.3, and light transparency of 0.58. The building has a heated

* Corresponding author. Tel.: +41 44 633 29 26.
E-mail address: nagyz@ethz.ch (Z. Nagy).

http://dx.doi.org/10.1016/j.apenergy.2015.07.050
0306-2619/

Fig. 1. The LowEx building studies in this research.

surface of approximately 900 m^2, and distributes the heat through floor heating. The roof is never shadowed, whereas the first floor is partially shadowed from November to February.

The schematic plot of the LowEx building systems is given in Fig. 2. Its main components are hybrid photovoltaic-thermal (PV/T) panels (Fig. 2(a)), geothermal boreholes (Fig. 2(c)), and high COP heat pump (Fig. 2(d)). There are other possible variations with different combinations of these components [7].

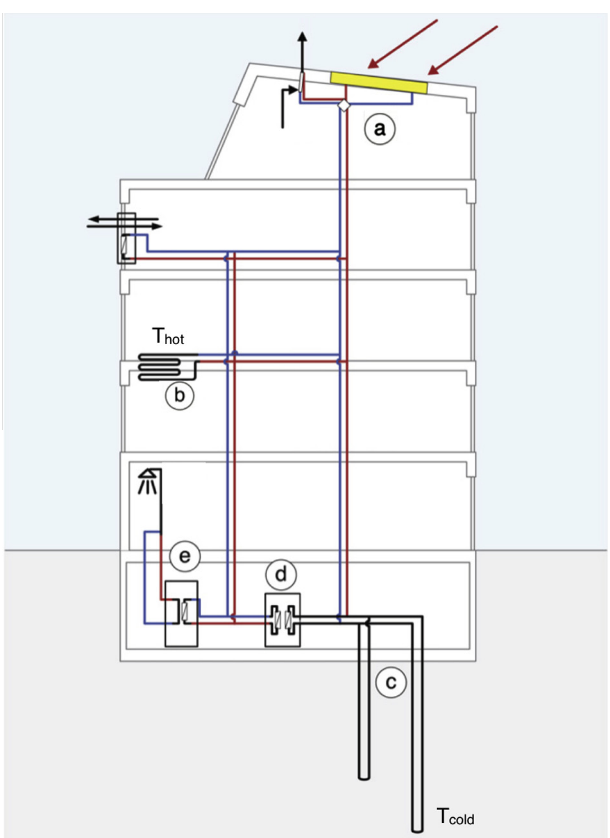

Fig. 2. Schematic of a low exergy building: (a) Hybrid photovoltaic-thermal (PV/T) panels with heat exchanger, (b) floor heating, (c) dual zone boreholes, (d) high COP heat pump, and (e) low temperature hot water storage. (blue: supply line, red: return line) *Adapted from [5]*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

LowEx building systems create more flexibility and generate new possibilities for the design of high performance buildings [5]. They maximize the connection to the freely available dispersed energy in the environment (e.g. solar, geothermal, or waste heat recycling). The core module is a heat pump with a high coefficient of performance (COP) [5]

$$COP = g \cdot COP_{\text{ideal}} = g \cdot \frac{T_{\text{hot}}}{T_{\text{hot}} - T_{\text{cold}}} \tag{1}$$

where g is the machine characteristic factor, its typical range is from 0.4 to 0.6, and T_{hot} and T_{cold} are sink and source temperature (in K), respectively. The difference, $T_{\text{hot}} - T_{\text{cold}}$, is the *temperature lift*. High COP is achieved through low temperature lift, i.e., by increasing T_{cold} and decreasing T_{hot}. With the arrangement of large heat exchange surface, like floor/wall/ceiling heating (Fig. 2(b)), T_{hot} can be easily reduced to 301–305 K (28–32 °C). Natural soil temperature increases proportionally with increased depth (e.g. 3 °C/100 m in Switzerland); thus, a deep borehole (Fig. 2(c)) is suitable as heat pump source [4,5]. PV/T panels (Fig. 2(a)) provide ground heat compensation besides electricity in summer times to maintain the initial T_{cold} level at the beginning of next heating period [4].

Generally, the primary goal of controlling the building environment (HVAC system) is to maintain the comfort level of the occupants, and to achieve good energy efficiency [8–10]. Classic controllers, such as ON/OFF and PID control, are the most widely used ones due to their simple structure and low initial cost [8]. Yet, HVAC systems are non-linear, multiple-input, multiple-output (MIMO) systems with large time delays and high order dynamics [8]. This substantially limits the performance of the classical control techniques. Advanced control techniques, such as model predictive control (MPC) present promising results [11,8], especially incorporating dynamic weather [12] and occupancy forecasting [13]. For instance, the MPC approach in [14], uses weather predictions to control building heating systems and achieves savings between 15% and 28% depending on the systems and weather conditions. However, the accuracy of these methods depends heavily on a good quality model of the building dynamics, which is difficult to construct given the complexity of HVAC systems [14,15].

Another potential approach is using intelligent or soft controllers that do not require a model, and are based on human perception of comfort [9]. Given that human sensation is subjective, intelligent controllers offer a good balance between occupant comfort and energy conservation in a dynamic environment. Several control strategies, e.g., variations of artificial neural networks (ANN) [16–18], and other methods have been proposed. For instance, to optimize air conditioning setback scheduling in public buildings, in [17] two years of weather data is used to train the neural network. A neural network is also used in [19] as a predictive controller for heating that adapts itself to real conditions (climate, building and user behavior) by solving the optimal control problem using dynamic programming. Further studies have demonstrated promising results using intelligent techniques alone or combined with other techniques [20], see also [9] for a review of learning algorithms. The drawbacks of these methods are that they typically require a large learning set in order to perform optimally.

Apart from the above mentioned techniques, reinforcement learning control (RLC) has also been considered for HVAC systems. The advantage of this approach is that the controller continuously learns from different operating conditions through interaction with the building, thereby improving its control policy online [21]. Thus, RLC offers a model-free approach capable of adapting to the local environment and operation conditions, such that the resulting controller is robust to time-varying disturbances.

Although numerical analysis and simulation study show that pure reinforcement learning can direct the controller to a