A Reinforcement Learning Demand Response Model Considering Demand Elasticity during High Electricity Prices in Germany

Jawana GABRIELSKI[1], Eduardo SALAZAR[2], Ulf HÄGER[3], Mauricio SAMPER[4]

[1]TU Dortmund, ie3, Martin-Schmeißer-Weg 12 44227 Dortmund, +49 231 755 4451, jawana.gabrielski@tu-dortmund.de, ie3.etit.tu-dortmund.de/institute/team/jawana-gabrielski/, Nachwuchsautorin

[2]National University of San Juan (UNSJ), Libertador General San Martin Avenue 1109, esalazar@iee.unsj.edu.ar

[3]TU Dortmund, ie3, Martin-Schmeißer-Weg 4-8 44227 Dortmund, +49 231 755 2394, ulf.haeger@tu-dortmund.de, ie3.etit.tu-dortmund.de/institute/team /dr-ing-ulf-haeger/

[4]National University of San Juan (UNSJ), Libertador General San Martin Avenue 1109, msamper@iee-unsjconicet.org

Abstract: This article introduces a new demand response model based on consumer prices, approached through the lens of reinforcement learning. In this case, the proposed methodology, leveraging elasticity analysis and reinforcement learning, is adept at identifying optimal times (hours of the day) when users are inclined to alter their demand. This leads to a dual benefit: a reduction in peak loads and a decrease in energy bills for consumers. Furthermore, considering that the reinforcement learning algorithm aims to maximize the benefits for both users and service providers, a time-of-use pricing system is derived. This system adapts to the fluctuating prices in the wholesale market as well as the elasticity of demand. Hence, it is intriguing to compare the obtained values against the prices currently offered to users. It is demonstrated that by having a consumer price that accounts for user behavior, drastic price increases for consumers can be avoided. This enables users to actively participate in modifying their demand while simultaneously preventing profit losses for electricity service providers. In a nutshell, a powerful tool for consumer price formulation is presented, taking into account uncertainties in both prices and demand.

Keywords: demand response, electricity prices, reinforcement learning.

1 Introduction

German electricity market has experienced sharp price fluctuations over the last years. Especially the Ukrainian war had a significant impact on electricity prices; moreover, the outage of nuclear power plants in France also affected them, leading to higher and more volatile electricity market prices [1]. Consequently, user prices also increased, e.g. residential prices more than doubled between 2021 and the second half of 2022 [2]. This surge has not only enhanced the awareness for electricity consumption but even provoked behavioral changes to avoid higher costs. In response to the challenges resulting from such price volatility, time-dependent price models such as time-of-use or dynamic prices have emerged as attractive solutions, to align energy consumption patterns with market dynamics. Implementations in other countries have shown several benefits of time of use prices, including save on retail electricity sales as well as demand side flexibility [3]. Starting from the year 2025 it is mandated that every German energy supplier must offer these prices as an alternative to their costumers [4]. For this, innovative tools capable of formulating new prices according to both price and demand behavior are required. These tools will be able to maximize benefits for both users and energy providers. This paper proposes a method to use behavioral changes resulting from higher prices as a reference in order to obtain denotive elasticity values. Therefore, an intelligent Demand Response (DR) model based on prices, allowing the determination of both real-time consumer prices and a time-of-use pricing schemes is suggested. With these consumer prices, a quantitative comparison is made between the current option versus the proposed model. For the price formulation, a Reinforcement Learning (RL) method is used, which obtains prices that maximize benefits for both users and energy marketers, considering an environment of uncertainties. Thus, it is demonstrated that the proposed pricing based on artificial intelligence is suitable for price formulation compared to other options.

2 Literature Survey

The electric systems are constantly evolving due to new technologies being added to the supply chain, both physically and computationally. The inclusion of unconventional renewable energies has caused price variations in the market due to the unpredictable nature of the primary resource. Additionally, the emergence of electric vehicles and distributed generation within distribution systems has created the need for new transactional mechanisms that can understand the behavior of these new actors. The scientific community is currently focused on finding intelligent tools to predict and understand the behavior of these actors in order to improve consumer price formulation and maximize the benefits for all parties involved. This research approach has been decided to be continued with, as promising results have been shown in addressing current challenges such as uncertainties in electricity supply prices caused by external factors like wars or long-term events. In this particular study, specific events that have caused ideal scenarios for price variation and changes in user behavior have been considered.

The evolving consumption patterns in the electric energy market, influenced by price variations, have raised significant concerns. A strategic solution to achieve energy balance and enhance market planning both in the short and long term is DR, a key component of Demand Side Management. DR involves sending economic signals to consumers to encourage adjustments in electricity usage. It can be broadly categorized into two types: price-based DR and incentive-based DR. Price-based DR utilizes fluctuating price signals to affect user behavior, while incentive-based DR employs rewards or penalties to induce specific demand changes, such as fines for exceeding a certain consumption limit or rewards for reducing consumption like tariff category adjustments. Long-term strategies primarily focus on price-based DR. Numerous studies effectively address the challenges of price-based DR. For example, research identified as [5] developed a model that considers the impact of significant penetration of unconventional renewable energies on wholesale prices. This model, employing RL known for its effectiveness in uncertain scenarios, successfully correlates pricing with consumer behavior.

Another study, referred to as [6], proposes a combined price and incentive-based DR model. Its goal is to smooth out demand peaks by offering variable pricing and incentives under a direct load control scheme. This approach highlights the potential of artificial intelligence algorithms in managing price-related uncertainties in the wholesale market. Additionally, in study [7], a price-based DR model utilizing RL demonstrates the ability to identify and create optimal signals for various user types. This is achieved through a satisfaction function that balances user benefits. Conversely, research [8] explores an incentive-based DR model. Here, through consumer modeling, incentives are provided to reduce demand, aiming to maximize benefits for both consumers and utility companies or marketers.

2.1 Research Gaps and Contributions

From the current advancements in the field, this work addresses certain critical areas. Despite progress in DR program research, a significant gap remains in their practical application in real-world scenarios. Real data, such as that from a smart grid project in Argentina as referenced in [5], has been used as input for various models. However, accurately identifying user behaviors remains a challenge. There is an ongoing need for methodologies that effectively model demand elasticity to provide valuable input for these models.

Furthermore, while state-of-the-art models have proposed promising real-time pricing solutions, their large-scale applicability remains distant from reality. This is primarily due to the prevalence of fixed and flat rate plans among users and the low penetration of smart meters capable of receiving pricing signals. Therefore, proposed schemes and solutions need to incorporate pricing options feasible over longer time scales.

To address these challenges, this work presents significant contributions to enhance DR models and advance the state of the art:

- A methodology for analyzing demand elasticity is introduced, using consumption and price data from Germany.
- Transition pricing solutions are developed, focusing not only on real-time pricing but also on identifying critical peaks and implementing real-time adjustments.

3 Problem Formulation

To address the problem of DR, it is essential to model user behavior based on prices, and therefore it becomes necessary to obtain enough signals that represent elasticity (e) and behavior that allow having the appropriate signals for the model. This becomes all the more relevant considering for a Service Provider (SP) or a marketer, planning and formulating long-term prices for them would be a difficult task, while maximizing their utility. Therefore, in this work, a scenario is considered where a group of users (d) are purchasing electricity from a SP (b). The core of its business is, in turn, buying electricity from the wholesale energy market (a) and reselling it to the users while generating profit. From this reasoning two very important considerations arise. On one hand, the SP should be an operatively functioning operator that is capable of bringing wholesale market prices to users effectively in order for them not to witness pronounced price fluctuations. On the other hand, it is important to develop pricing strategies that the users remain able to contribute actively with their varying demand whilst it is not affected on their satisfaction that is under a management of the RL algorithm (c). Based on these premises, each of the actors involved in this work will be formulated. Figure 1 demonstrates the explained interactions.

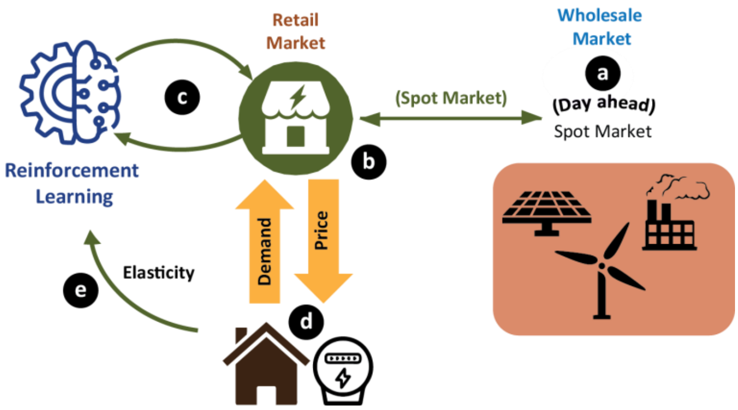

Fig. 1. Market Scheme

3.1 User Model

In this section, the user modeling is formulated. This model is based on a user whose elasticity has been obtained from an analysis of historical demands and prices. The data used are described in more detail in the application section. The demand and prices before the price increase are called "Old," and after the increase, "New." Taking the price increase and the variation in demand, the elasticity (1) and the auto-elasticity (2) were calculated to understand the consumer behavior pattern. Then, these elasticities are added into the total elasticity that will serve as input in the RL algorithm.

$$\xi = \frac{\Delta d_u / d_u}{\Delta p_u / p_u} \tag{1}$$

$$a\xi_{i,j} = \frac{p_j}{d_i} \times \frac{d_{i+1} - d_i}{p_{i+1} - p_i} \tag{2}$$

For the RL algorithm, it is necessary to create objective functions based on the elasticity profiles of users. These functions will serve as rewards that the algorithm seeks to maximize. The user's benefit in this context is defined as the price paid for electrical service in relation to the cost of satisfaction (3) [9].

$$Uben_{u,h} = \sum_{u=1}^{U} \sum_{h=1}^{H} \left[(1 - \rho_u) \cdot \left(\phi_{u,h}(d_{u,h}) - \rho_u \cdot \left(\Delta d_{u,h} \cdot p_{u,h} \right) \right) \right] \tag{3}$$

$$\phi_{u,h}(d_{u,h}) = d_{u,h} \cdot \beta_{u,h} \left(\frac{\sum_{h=1}^{H} d_{u,h}}{d_{u,h}} \right)^3 - \sum_{h=1}^{H} d_{u,h} \tag{4}$$

$$\Delta d_{u,h} = d_{u,h} + (\xi_u \cdot a\xi_u) \cdot \frac{p_{u,h} - p_{min}}{p_{min}}, \; p_{min} \le p_{u,h} \le p_{max} \tag{5}$$

Here, $Uben_{u,h}$ stands for the utility or benefit function for users $\{u \in U\}$ in each hour $\{h \in H\}$, where U is the number of considered users and H the number of considered hours. ρ_u is a weighting factor that balances the user's preference between price and satisfaction. $\Delta d_{u,h}$ indicates the change in consumption, whether it is reduced or increased, as a part of DR actions. As described in (5), ξ_u and $a\xi_u$ refer to the previously mentioned elasticities. The term $p_{u,h}$ is the price paid by users for their energy consumption. Additionally, $\phi_{u,h}$ represents the user satisfaction factor, as defined in (4), and this factor aligns with the concept of decreasing marginal utility. In economic terms, decreasing marginal utility occurs when the perceived utility decreases with each additional unit of consumption. Equation (4) introduces a β term, representing the consumer's inclination towards DR; a higher β implies a more conservative stance towards reducing consumption. Moreover, both the increase and decrease in consumption must be considered, along with their impact on consumer satisfaction. Therefore, a model that accommodates both scenarios is required.

3.2 Service Provider Model

In this research, the model developed portrays the electric SP as a market agent in the commercial electric market, involved in buying energy from the wholesale market and selling it to users in the retail energy market. The SP is not responsible for maintaining and operating the distribution networks; its activities are limited to the commercialization of energy. The aggregator's profit in this business model is derived from energy trading activities, and the profit function is described in section (6).

$$SPBen_{u,h} = \sum_{u=1}^{U} \sum_{h=1}^{H} \left[\left(\Delta d_{u,h} (p_{u,h} - z_{u,h} \right) \right] \tag{6}$$

In this context, $z_{u,h}$ symbolizes the purchase price of energy in the wholesale market for SP.

3.3 Objective Function

The objective function of the model, aimed at maximizing both social and commercial benefits for users and the aggregator/marketer, is presented in (7).

$$FO = \max \sum_{u=1}^{U} \sum_{h=1}^{H} \left(SPben_{u,h} + Uben_{u,h} \right) \tag{7}$$

3.4 Reinforcement Learning

In this work, RL is utilized for managing the demand of electric distribution users. This has been implemented in a simulated environment; however, the approach aims for the algorithm to be informed by user behaviors in an actual scenario. RL relies on the interactive learning that occurs between agents and the environment. Here, the objective of a learning agent is to identify the action that yields the greatest reward through experimentation. Interactions between the agent (SP) and the environment (users) take place at discrete intervals t, with each agent action producing a reward (financial outcome) and transitioning to a new state s_{t+1}. Q-learning is the method employed for the discovery of rewards and actions. Consequently, each action chosen by the agent incurs a reward (positive or negative) and prompts a transition to a new state. The set of state, action, reward, and subsequent state is termed "policy" (s, a, r', s') and is recorded in what is known as the Q-matrix. Through trial and error, the agent strives to maximize the reward derived from actions. Following the accumulation of experience, the Q-matrix is organized to pinpoint the optimal actions. Specifically, in this study to curtail demand, actions of the agent are delineated as price adjustments, and the advantage is depicted by the objective function outlined in (7).

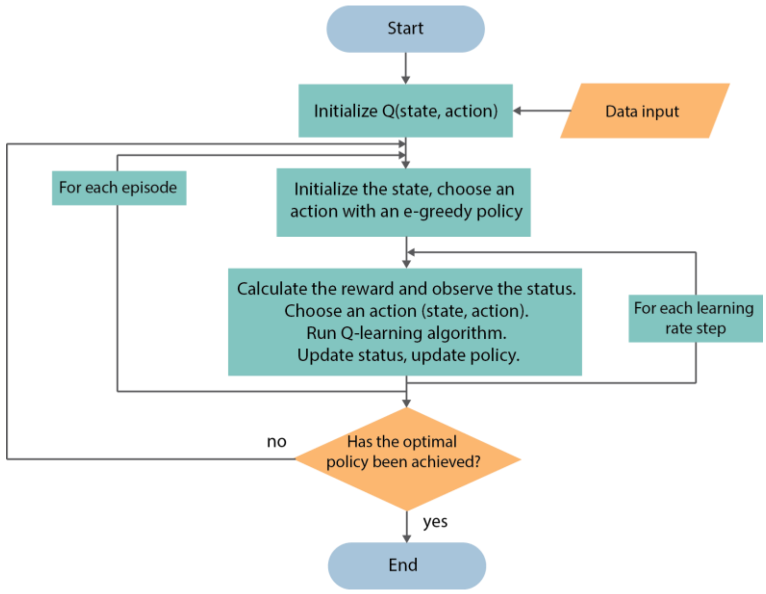

Fig. 2. Reinforcement Learning flow chart

4 Result Analysis

To identify the elasticity resulting from price increase, measured consumption time series are examined. Data from the openMeter platform [10] is used for this. The openMeter platform is an open data platform and contains measured consumption time series for different customers. In addition, it contains metadata, including location as well as area of the customer, however no information on the customer's tariffs nor on their installed devices are provided. For this paper, only residential consumption time series are considered.

In the process of identifying elasticity from customers, two problems were faced. On the one hand, different customers have different tariffs, which might have dissimilar conditions such as price guaranties for certain terms. Hence, it is unknown when the consumer received the price increases. On the other hand, there are also customers who increased their consumption during times of high prices, which might e.g. result from new devices.

Thus, consumers are searched, who reduced their load price-dependently by changing their behavior, i.e. not: reduced their load due to absence time or due to the installation of new more efficient devices (which in the long term consume price-independent less energy). For this, a multi-step approach is applied: the first step is to identify consumers who are generally eligible by selecting customers who lowered their yearly consumption in the year 2022 compared to average of the last two years before. In the second step, for the selected costumers from the first step, the course of the yearly time series is compared for different years. To do this, the yearly time series are smoothed using a rolling window. This enables to identify longer periods of low electricity consumption. The third step is to compare the course of the average daily consumption, separated by week and weekend. It allows recognize time series, which are reduced by a time-independent offset, resulting for example form new devices. The filtering is depicted in Figure 3.

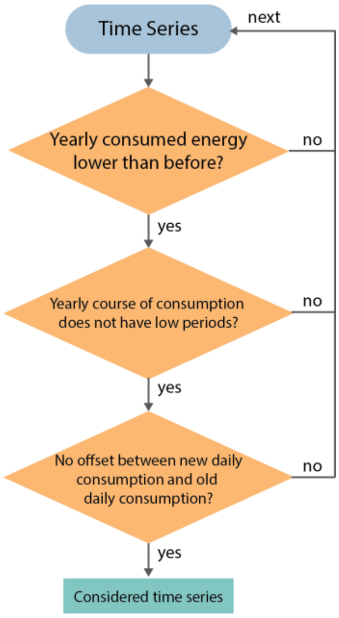

Fig. 3. Filtering flow chart

Employing the described method, 15 time series are identified, which are averaged and used to apply the price formulation. As mentioned, the consumer prices vary for different consumers, hence no common price can be considered, however, behavioral changes are not only influenced by the real customer price, but also by information from media etc. Thus, not the time of changing prices is detected, but the time of changing behavior. For this, the average of the yearly course of all selected time series is compared for different years and possible changepoints are identified, which are shown in Figure 4. Subsequent, the changepoint with the largest effect is selected.

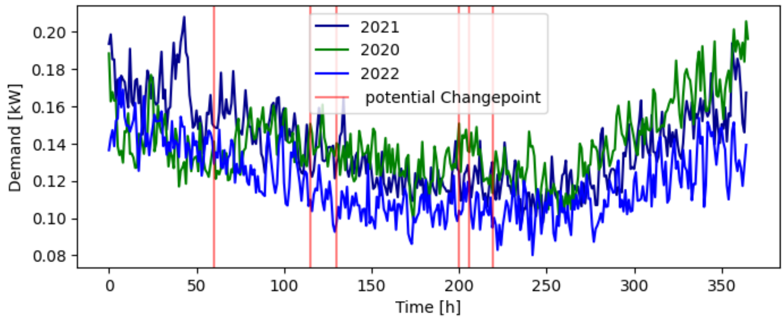

Fig. 4. Changepoint identification based on consumption change.

For the residential prices data from a German comparison portal for electricity prices is averaged for the time before and after the identified changepoint. Furthermore, the hourly market price average is calculated. To enable a comparison between market price and customer price, a factor is used, taking into account taxes and dues. The limits of the price range are calculated by applying a symmetric factor. Figure 5 shows the resulting prices. Once the values to identify the behavior of the users have been obtained, the necessary database is composed, which will be required as input to the algorithm, as is shown in Figure 2. A summary of the data used in this paper is presented in Figure 6.

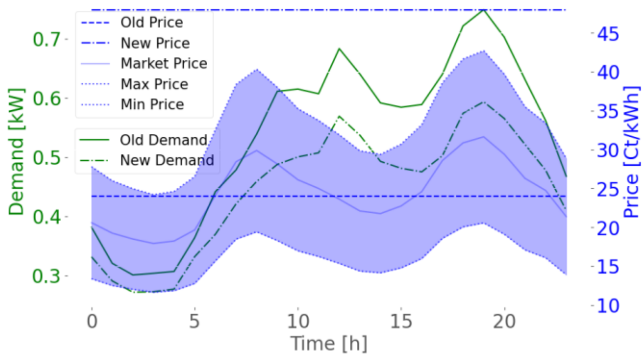

Fig. 5. Demand and Price Comparison

Given the relevance of elasticity in understanding user behavior, the elasticity values from equations (1) and (2) were calculated. Therefore, Figure 6 displays the auto-elasticity (2) as well as the result of ($\xi_u \cdot a\xi_u$). As can be seen in Figure 6(a), when auto-elasticity is obtained, it is possible to determine changes in consumption and price specifically with respect to the peak at 19:00 hours. However, once the elasticities are combined, it can be observed how the elasticity matrix also acquires the behavior of how users responded to the price increase as seen in Figure 6(b).

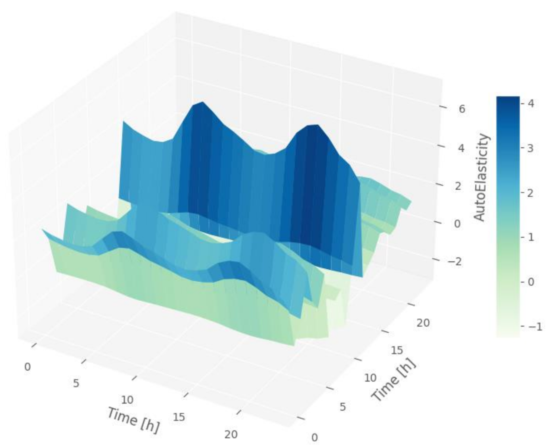

Fig. 6(a). Auto-elasticity

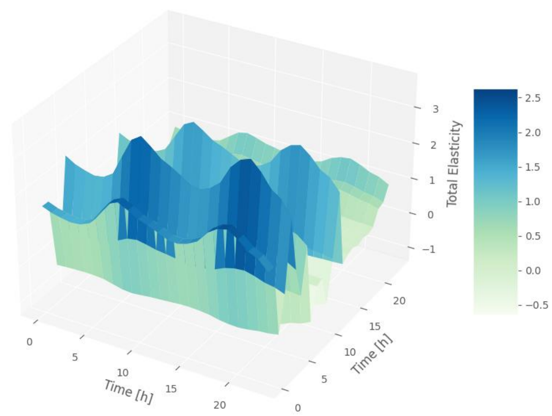

Fig. 6(b). Total Elasticity

The DR model is executed based on the diagram in Figure 2. Consequently, new prices are obtained, which will be sent to users under three different schemes (real-time pricing, time-of-use pricing, and critical peak pricing). In this regard, it can be observed in Figure 7(a) how the model effectively manages to provide users with prices according to their behavior. It is also noticeable that during peak times, higher prices are offered to encourage the reduction of peaks, and during off-peak hours, and depending on the user's elasticity, lower prices can also be obtained to encourage, on the other hand, increased consumption. Similarly, in Figure 7(b), it can be seen how the prices based on the time-of-use scheme are high during the two demand peaks, while during off-peak hours, for example from 24:00 to 06:00, the prices are reduced. Finally, in Figure 7(c), by offering a critical peak price, the model effectively sends a high price during the critical demand peak, which in this case occurs between 18:00 and 20:00.

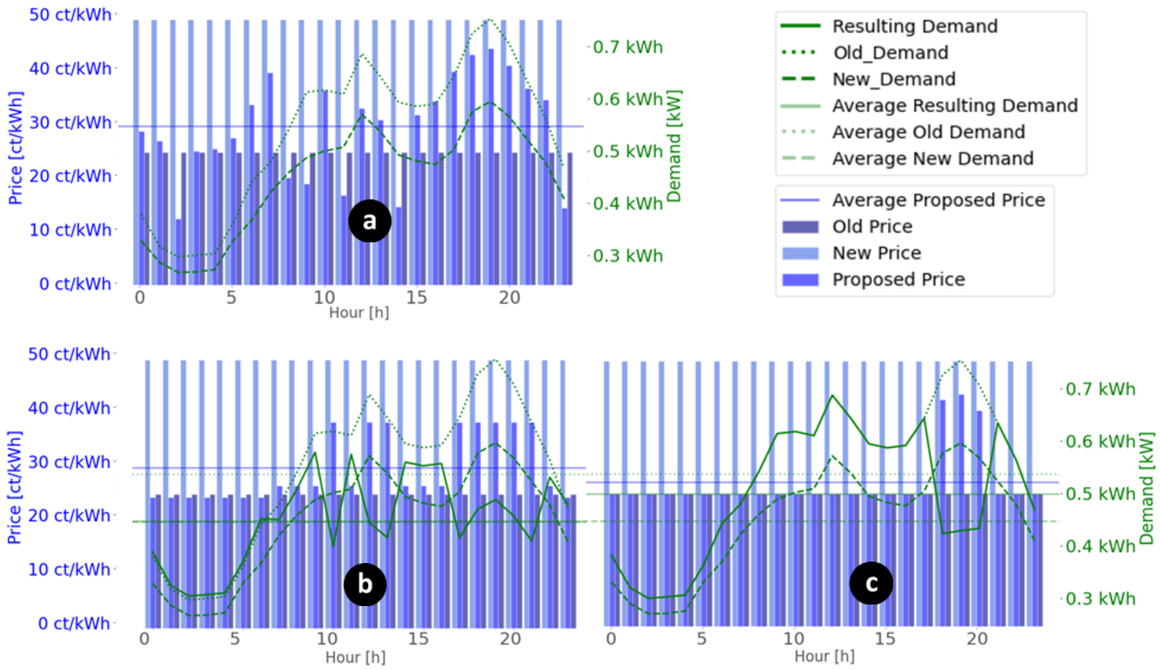

Fig. 7. Real-time pricing (a), time-of-use pricing (b), and critical peak pricing (c)

5 Conclusions and Future Works

This paper introduces a new demand response model based on price, approached through the lens of reinforcement learning. The proposed method leverages recent price increases in the German electricity market, to analyze elasticity. Based on this, in combination with typical patterns of market electricity prices, reinforcement learning is adept at identifying optimal times (hours of the day) when users are inclined to alter their demand. This leads to a dual benefit: a reduction in peak loads and a decrease in energy bills for consumers. Furthermore, considering that the reinforcement learning algorithm aims to maximize the benefits for both users and service providers, a time-of-use pricing system is derived. This system adapts to the fluctuating prices in the wholesale market as well as the elasticity of demand. Hence, it is intriguing to compare the obtained values against the prices currently offered to users. It is demonstrated that by having a price that accounts for user behavior, drastic price increases for consumers can be avoided. This enables users to actively participate in modifying their demand while simultaneously preventing profit losses for electricity service providers. In a nutshell, a powerful tool for price formulation is presented, taking into account uncertainties in both prices and demand.

The proposed method is applied to residential consumption and price time series of German costumers and exemplary resulting prices are presented. The presented work only considered costumers, which changed their behavior, neglecting customers, which behave price independently; hence the results tend to overestimate the elasticity and cannot be extrapolated. Further research based on a larger database could investigate the effect on the overall residential load. In addition, customer segmentation and identifying reasons for and against price independent behavior could be analyzed more in detail. As the considered customers pay their electricity based on a time-independent prices, load shifting is considered indirectly and should be investigated in further research based on demand time series with time dependent tariffs. Similar to this, moreover, it would be interesting to adapt the algorithm to take into account day ahead prices and build individual prices based on prices and elasticity.

Additional application of the proposed methods could be in the redispatch context, aiming at incentivizing to adapt the load to two separate zonal prices in order to reduce redispatch costs in Germany.

The research leading to this publication has received funding from Bundesministerium für Wirtschaft und Klimaschutz der Bundesrepublik Deutschland.

6 References

[1] "https://www.ffe.de/veroeffentlichungen/deutsche-strompreise-an-der-boerse-epex-spot-im-jahr-2022/."

[2] "https://www.bdew.de/service/daten-und-grafiken/bdew-strompreisanalyse/."

[3] T. International Renewable Energy Agency, *TIME-OF-USE TARIFFS INNOVATION LANDSCAPE BRIEF About IRENA*. 2019. [Online]. Available: www.irena.org

[4] "https://www.gesetze-im-internet.de/enwg_2005/__41a.html."

[5] E. J. Salazar, M. Jurado, and M. E. Samper, "Reinforcement Learning-Based Pricing and Incentive Strategy for Demand Response in Smart Grids," *Energies (Basel)*, vol. 16, no. 3, p. 1466, Feb. 2023, doi: 10.3390/en16031466.

[6] E. J. Salazar, M. E. Samper, and H. D. Patiño, "Dynamic customer demand management: A reinforcement learning model based on real-time pricing and incentives," *Renewable Energy Focus*, vol. 46, pp. 39–56, Sep. 2023, doi: 10.1016/j.ref.2023.05.004.

[7] D. Jang, L. Spangher, S. Nadarajah, and C. Spanos, "Deep reinforcement learning with planning guardrails for building energy demand response," *Energy and AI*, vol. 11, Jan. 2023, doi: 10.1016/j.egyai.2022.100204.

[8] I. Jendoubi and F. Bouffard, "Multi-agent hierarchical reinforcement learning for energy management," *Appl Energy*, vol. 332, Feb. 2023, doi: 10.1016/j.apenergy.2022.120500.

[9] R. Lu, S. H. Hong, and X. Zhang, "A Dynamic pricing demand response algorithm for smart grid : Reinforcement learning approach," *Appl Energy*, vol. 220, no. March, pp. 220–230, 2018, doi: 10.1016/j.apenergy.2018.03.072.

[10] "www.openMeter.de."