Engineering 7 (2021) 1101–1114

Contents lists available at ScienceDirect

Engineering

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

Research
Active Support of Power System to Energy Transition—Article

Flexibility Prediction of Aggregated Electric Vehicles and Domestic Hot Water Systems in Smart Grids

Junjie Hu [a,*], Huayanran Zhou [a], Yihong Zhou [a], Haijing Zhang [a], Lars Nordströmd [b], Guangya Yang [c]

[a] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China
[b] Division of Electric Power and Energy Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm 10044, Sweden
[c] Center for Electric Power and Energy Department of Electrical Engineering, Technical University of Denmark, Kgs Lyngby 2800, Denmark

ARTICLE INFO

Article history:
Received 8 October 2020
Revised 16 January 2021
Accepted 29 March 2021
Available online 24 June 2021

Keywords:
Load flexibility
Electric vehicles
Domestic hot water system
Temporal convolution network-combined transformer
Deep learning

ABSTRACT

With the growth of intermittent renewable energy generation in power grids, there is an increasing demand for controllable resources to be deployed to guarantee power quality and frequency stability. The flexibility of demand response (DR) resources has become a valuable solution to this problem. However, existing research indicates that problems on flexibility prediction of DR resources have not been investigated. This study applied the temporal convolution network (TCN)-combined transformer, a deep learning technique to predict the aggregated flexibility of two types of DR resources, that is, electric vehicles (EVs) and domestic hot water system (DHWS). The prediction uses historical power consumption data of these DR resources and DR signals (DSs) to facilitate prediction. The prediction can generate the size and maintenance time of the aggregated flexibility. The accuracy of the flexibility prediction results was verified through simulations of case studies. The simulation results show that under different maintenance times, the size of the flexibility changed. The proposed DR resource flexibility prediction method demonstrates its application in unlocking the demand-side flexibility to provide a reserve to grids.

1. Introduction

Environmental concerns have largely promoted the use of renewable energy in recent decades [1]. This transition brings challenges to power system operation owing to the intrinsic uncertainty of renewable generation and distributed energy resources. The main challenge is the increasing power imbalance between demand and supply, which leads to an increasing demand for controllable resources to be deployed in the system, while the magnitudes of conventional ancillary service resources are decreasing [2]. In addition, the increasing number of connected distributed energy resources at medium and low voltage levels may cause congestion issues [3].

With the development of smart grids and electricity markets, demand response (DR) resources can be an integral part of the system operation. Thus, a promising solution to preserve the ancillary service provision resource is to aggregate the flexibility of DR resources [4]. Note that, according to the definition of the International Energy Agency (IEA), flexibility in the power system is the ability to maintain reliability by adjusting the generation or load during large disturbances [5]. Therefore, this study defined flexibility as the ability of DR resources to increase and decrease electricity demand. The flexibility of DR resources can be employed by the aggregator through price compensation or economic incentives, which allows the DR resources to modify their own electricity consumption behavior according to the needs of the system operation [6,7]. The aggregator can then trade the flexibility of DR resources in the ancillary service market to provide controllable resources for power system operation.

To date, many studies have been conducted on how to make full use of the flexibility of DR resources in providing ancillary services to grids. To achieve the peak load shifting effect, Li et al. [8] established a model for aggregated electric vehicles (EVs) to participate in price-based DR. Sanandaji et al. [9] employed the electricity consumption flexibility of aggregated residential heating, ventilation, and air conditioning loads to regulate reserve services with a certain ramping rate. Hu et al. [10] proposed the idea of using DR resources to achieve a power balance between supply and demand for a multi-area power system. The approach presented in Ref. [11] established reward rules for DR participants who provide flexibility

* Corresponding author.
E-mail address: junjiehu@ncepu.edu.cn (J. Hu).

https://doi.org/10.1016/j.eng.2021.06.008

and use the DR flexibility to realize load shifting as well as voltage improvement in supply feeder.

As can be seen in Refs. [8–11], the flexibility of aggregated distributed energy resources determines the effect of the implementation of the DR program [12]. However, considering the stochastic feature of end-user behaviors and the complex physical characteristics of DR resources, the aggregated flexibility of DR resources is usually uncertain before the implementation of the DR program. Therefore, several risk control strategies are widely adopted in DR programs to handle the uncertainties of the flexibility of DR resources. Ref. [13] introduced a conditional value-at-risk into EVs problems, thereby providing reserve in the ancillary service market to deal with the uncertainty of EV flexibility. Ref. [14] established a robust optimization problem for the optimal scheduling problem of aggregated EVs to obtain a conservative frequency regulation market participation strategy. To deal with the deviation between the actual and pre-scheduled flexibility of grid-connected EVs in the DR grogram, Ref. [15] formulated an optimization model based on the predictive control-based rolling horizon method. Han et al. [16] used a multiple-scenario-based stochastic programming method to handle the uncertainties of EV flexibility in the DR program. Although the risk control method used in Refs. [13–16] can help the aggregator consider the uncertainties in the day-ahead DR planning, the results may be highly conservative, which may lead to loss of economic benefits of the DR users and even prevent the full flexibility of DR resources in grid operations.

An accurate flexibility prediction model for aggregators is essential for dealing with uncertainty when participating in grid operations. However, only a few studies have been conducted on the flexibility prediction problem. Many studies have mainly focused on electricity demand prediction rather than directly predicting flexibility. Wang et al. [17] and Chen et al. [18] proposed a probabilistic load forecast method that is used in system planning and dispatch. The authors in Ref. [19] reviewed the current energy forecast method and provided future directions. However, they failed to discuss the direct representation and forecast of the flexibility method. At present, most of the research on flexibility is limited to the definition and evaluation of flexibility. Ref. [20] provided a representation method for EV flexibility, considering the participation of EVs in frequency regulation, and evaluated the flexibility of EVs in frequency regulation using three-year data. For residential DR load, the authors in Ref. [21] established a binomial probability distribution model of demand change and used maximum likelihood estimation to evaluate the flexibility of demand increase, demand decrease, or demand immobility.

In Ref. [22], a recurrent neural network (RNN) based flexibility prediction method was proposed to acquire the flexibility of aggregated domestic hot water system (DHWS). Using the flexibility prediction results of DHWS, the load shift scheduling problem in day-ahead and real-time can be carried out with certain accuracy. However, the proposed RNN-based flexibility prediction method can only predict the size of the aggregated flexibility, but cannot present the maintenance time of flexibility. Furthermore, normal RNN-based methods may show poor performance in complex problems. For example, they may suffer from gradient exploration, and the gradient vanishes when the input sequence is sufficiently long; thus, they cannot be applied to general cases [23,24].

Building on the previously developed RNN-based flexibility prediction method in Ref. [22], an advanced deep learning network, the temporal convolution network (TCN)-combined transformer, was adopted in this study to predict the real-time aggregated flexibility of two typical DR resources, EV and DHWS multi-step-ahead. This method considers the physical characteristics as well as the different power consumption strategies. The main contributions of this study are twofold. First, the study provided a general multi-step-ahead flexibility prediction method for DR resources. Through the proposed flexibility prediction method, the size of the aggregation flexibility and the maintenance time of flexibility can be obtained. Second, we applied the TCN-combined transformer model to address the flexibility prediction of aggregated EVs and DHWSs. The TCN-combined transformer-based prediction model is one of the most advanced models for modeling long-time-dependency problems. The accuracy of the flexibility prediction with a long predicted slot was verified through simulation of a case study. The DR planning and operation problem in, for example, offering reserve service to power grids, can be supported by the proposed flexibility prediction method.

The remainder of this paper is organized as follows: Section 2 provides the preliminaries of our work, including the description of the problem, system framework, and some important definitions. Subsequently, the TCN-combined transformer-based prediction method is described in detail in Section 3. Section 4 presents the simulation results and discussion. Finally, the conclusions are presented in Section 5.

2. Preliminaries

2.1. Definition of DR resources flexibility

In a broad sense, flexibility refers to the ability of the system to respond to internal and external uncertainties, that is, the response ability of the system when internal or external variables change [25]. However, in power systems, there is no uniform definition of flexibility. Currently, a widely accepted definition comes from IEA. It defines power system flexibility as the ability to maintain reliability by adjusting the generation or load during large disturbances [5]. According to IEA's definition of flexibility, for the DR resources on the load side, flexibility is reflected in the ability to increase or decrease electricity demand.

Fig. 1 illustrates the flexibility of DR resources. In terms of the base power demand (P_{base}) of DR resources, that is, the power consumption in normal use, the power demand can be increased to P_{max} or reduced to P_{min}, where $P_{min} < 0$ indicates that the DR resources can provide power to the grid. The values of P_{max} and P_{min} are constrained to the physical characteristics and user requirements of the DR resource, which are discussed in detail in Section 2.3. The distance between P_{base} and P_{min}/P_{max} is the capacity to reduce or increase the electricity demand, that is, flexibility.

This flexibility can be used as a reserve for power systems. In the traditional sense, reserves have an additional generation capacity above the expected load [16]. The reserve capacity setting protects the power system against the uncertain occurrence of future

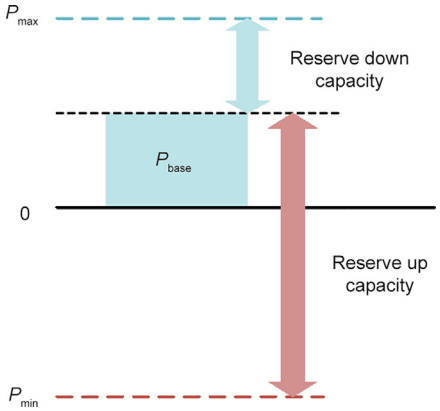

Fig. 1. The flexibility of DR resources.

operating events, including loss of energy or load forecasting errors. When such operating events occur, the DR resources can also cover the imbalance between the supply and demand by increasing or decreasing its electricity demand. Accordingly, the capacity of DR to reduce/increase electricity demand can be called the up/down reserve capacity, as shown in Fig. 1.

A reasonable way for DR resources to provide reserves is to allow them to declare the capacity of up and down reserves asymmetrically. However, the uncertainties and coupling in energy and power make the description of DR flexibility a complex problem. To better use the flexibility of DR resources in power systems, the flexibility prediction method needs to be explored.

2.2. Problem formulation and DR system framework

The main aim of our study was to predict the flexibility of two typical DR resources, that is, EV and DHWS, in multi-step-ahead by processing realistic electricity consumption data using the recorded historical end-users' behavior database.

In our prediction, we studied the DR flexibility in the form of aggregation as it can offset the uncertainty of internally distributed DR resources, and the existing electricity market rules do not allow the participation of individual customers. Considering the different electricity consumption characteristics of EVs and DHWSs, an EV aggregator (EVA) and a DHWS aggregator (DHWSA) were used to aggregate EVs and DHWSs, respectively.

$$\begin{cases} F_{\mathrm{EVA}}(t) = \hat{P}_{\mathrm{EVA}}(t) - P_{\mathrm{EVA}}(t) \\ F_{\mathrm{DHWSA}}(t) = \hat{P}_{\mathrm{DHWSA}}(t) - P_{\mathrm{DHWSA}}(t) \end{cases} \tag{1}$$

The flexibility of EVA and DHWSA at time step t can be denoted as $F_{\mathrm{EVA}}(t)$ and $F_{\mathrm{DHWSA}}(t)$, respectively, which can be derived from Eq. (1). $P_{\mathrm{EVA}}(t)$ and $P_{\mathrm{DHWSA}}(t)$ represent the base power demand of EVs and DHWSs at time step t, respectively, which are equivalent to P_{base} in Section 2.1. $\hat{P}_{\mathrm{EVA}}(t)$ and $\hat{P}_{\mathrm{DHWSA}}(t)$ represent the aggregated power of EVs and DHWSs at time step t when engaging in the DR program, respectively, which are further discussed in detail in the following section.

Fig. 2 shows the DR system framework. Note that we defined the power direction from the grid to the demand side as the positive direction. The EVA and DHWSA can participate in the electricity energy market and ancillary services market for reserve bidding. To assist the prediction and application of flexibility, the EVA and DHWSA can coordinate and schedule the power of all the distributed EVs and DHWSs, respectively, by sending a DR signal (DS, discussed in detail in Section 2.4) as an instruction. They can offer attractive discounts in the DR users' electricity bill when they change the electricity consumption behavior according to the DR signal instructions. To carry out the flexibility prediction of distributed EVs and DHWSs, both EVA and DHWSA use TCN-combined transformer-based technology. The TCN-combined transformer model is described in detail in Section 3.

2.3. Base power demand of EV and DHWS

As mentioned in Section 2.1, the flexibility of aggregated DR resources is closely related to the base power demand, and the base power demand of aggregated DR resources is influenced by the physical characteristics of DR resources and end-users' electricity consumption strategies.

2.3.1. Physical characteristics of each DR resource

Our study considered three operational modes of each EV, that is, fast-charging with rated fast-charging power P_{fast}, slow-charging with rated slow-charging power P_{slow}, and discharging with rated discharging power P_{dis}. We formulated the following constraints to represent the characteristics of a single EV, as shown in Eqs. (2)–(8):

$$P_{\mathrm{dis}} \le P_{\mathrm{EV}}(t) \le P_{\mathrm{fast}}, \tag{2}$$

$$E(t+1) = E(t) + \eta P_{\mathrm{EV}}(t)\Delta t, \ \text{if } t \in [t_{\mathrm{start}}, t_{\mathrm{dep}}), \tag{3}$$

$$E(t) = 0, \ \text{if } t \notin [t_{\mathrm{start}}, t_{\mathrm{dep}}], \tag{4}$$

$$E(t) = E_{\mathrm{start}}, \ \text{if } t = t_{\mathrm{start}}, \tag{5}$$

$$E(t) \ge E_{\mathrm{exp}}, \ \text{if } t = t_{\mathrm{dep}}, \tag{6}$$

$$E(t) \le E_{\mathrm{max}}, \tag{7}$$

$$P_{\mathrm{EV}}(t) \ge P_{\mathrm{slow}}, \ \text{if } E(t) < E_{\mathrm{ms}}, \tag{8}$$

where $P_{\mathrm{EV}}(t)$ represents the power of EV at time step t. $P_{\mathrm{EVA}}(t) = \sum_{\mathrm{EV}\in\mathrm{EVA}} P_{\mathrm{EVA}}(t)$ represents the relationship between $P_{\mathrm{EV}}(t)$ and $P_{\mathrm{EVA}}(t)$. The state of charge (SOC) of EV at time step t is denoted as $E(t)$. The energy transfer efficiency between the charging pile

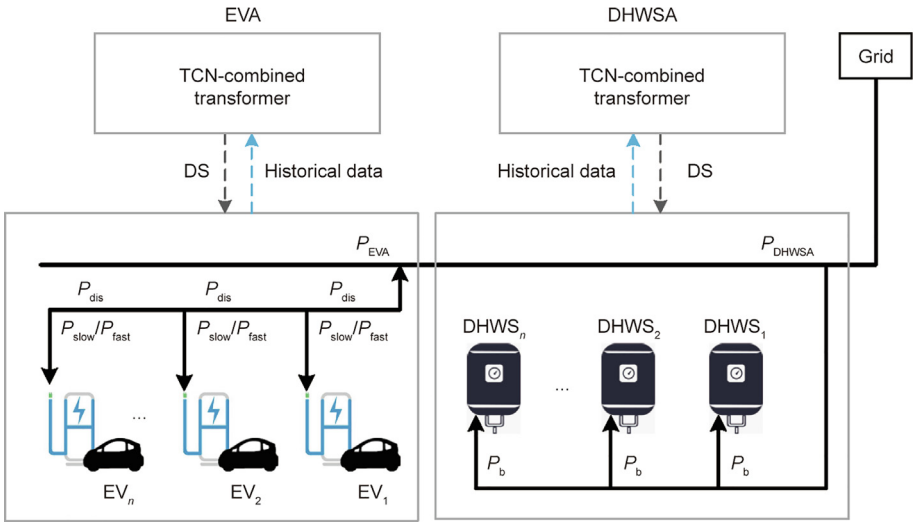

Fig. 2. System framework. P_{b}: the power consumption of the boiler in the DHWS; P_{dis}: discharging with rated discharging power; P_{fast}: fast-charging with rated fast-charging power; P_{slow}: slow-charging with rated slow-charging power.

(grid) and the EV battery is indicated as η. We used Δt to represent the sampling timescale. The initial SOC of a single EV when EV is plugged in at the charging start time t_{start} is represented as E_{start}. The expected SOC at the time of departure t_{dep} is denoted as E_{exp}. The maximum limitation of SOC is represented as E_{max}. The EV must be charged to a certain SOC to meet the driving demand, which is described in Eq. (6). The EV charging demand can be ensured when E_{ms} is set, which means that the EV must be in charging state when its SOC is lower than E_{ms}.

The DHWS user's requirement is to keep the water tank temperature T_{tank} close to a predefined temperature T_{ref}, with a maximum allowed deviation equal to the deadband T_{db}, as described in Eq. (9). The power consumption of the boiler in the DHWS is represented as P_b in this study, and the DHWS operation model is formulated as follows:

$$T_{ref} - T_{db} \le T_{tank}(t) \le T_{ref} + T_{db}, \tag{9}$$

$$T_{tank}(t+1) = T_{tank}(t) + \frac{Q_{heat}(t) - Q_{drain}(t) - Q_{loss}(t)}{C_p V_{tank}}, \tag{10}$$

$$Q_{heat}(t) = P_b \Delta t, \tag{11}$$

$$0 \le P_{DHWS}(t) \le P_b, \tag{12}$$

$$Q_{drain}(t) = V_{flow}(t) C_p (T_{outlet} - T_{inlet}), \tag{13}$$

$$Q_{loss}(t) = \Lambda_{tank}(T_{tank}(t) - T_{amb}(t)), \tag{14}$$

where $T_{tank}(t)$ is the temperature of the tank in the DHWS at time step t. The DHWS keeps the $T_{tank}(t)$ close to the set value T_{ref} by controlling the heat emitted $Q_{heat}(t)$ from the thermostatically controlled boiler. $Q_{drain}(t)$ and $Q_{loss}(t)$ are the heat losses from the hot water usage of the end users and the insulated tank to the surroundings at time step t, respectively. C_p and V_{tank} are the specific heat capacity of water and volume of the water tank, respectively. P_b refers to the power consumption of the boiler in the DHWS, and Δt represents the sampling time scale. The power of DHWS at time step t can be denoted as $P_{DHWS}(t)$, and $P_{DHWSA}(t) = \sum_{DHWS \in DHWSA} P_{DHWS}(t)$. T_{inlet} and T_{outlet} are the temperatures of the inlet and outlet water, respectively. $V_{flow}(t)$ is the total hot water flow used to serve the current DHWS activity of the household member. Λ_{tank} is the insulation coefficient of the water tank, and $T_{amb}(t)$ is the ambient air temperature at time step t.

2.3.2. Power consumption strategies for DR resources

In addition to the characteristics of DR resources, the user's preferred power consumption strategies also have a significant impact on the base power demand. In this study, two typical power consumption strategies are considered to investigate the impact on flexibility: Strategy 1 is the most primitive electricity consumption method, which is completely dependent on the users' electricity demand and strategy 2 is the most economical way to consume electricity. The user in this way would fully take the electricity tariff into consideration, and tend to consume more electricity when the tariff is low. In the period when the time of use (ToU) tariff is not favorable, strategy 1 is the main power consumption strategy of DR users. However, with the common use of ToU tariffs, strategy 2 becomes increasingly common among DR users.

In strategy 1, the EV starts charging as soon as it parks, and the water heater is only heated when the tank temperature reaches the lower limit.

However, as ToU tariff has been widely used in EV parking lots as well as in residential areas, the power consumption of EVs and DHWSs may be guided by the ToU tariff. Thus, in strategy 2, we consider an optimization power-consumption mode under the ToU tariff. The optimization objectives in strategy 2 for aggregated EVs and DHWSs are shown in Eq. (15).

$$\begin{cases} \min c_{ToU}(t) \cdot P_{EVA}(t) \cdot \Delta t \\ \min c_{ToU}(t) \cdot P_{DHWSA}(t) \cdot \Delta t \end{cases} \tag{15}$$

where $c_{ToU}(t)$ represents the ToU tariff.

Compared with strategy 1, some DR flexibility is activated by ToU at some time steps under strategy 2; however, the DR resources still have additional flexibility based on strategy 2. If the power grid has emergency reserve demand, predicting the DR flexibility in advance using strategy 2 would be helpful for DR resources to provide reserve for the power grid. This type of demand responds process can be motivated by additional subsidies.

Both power consumption strategies are constrained by the physical characteristics of each EV (Eqs. (2)–(8)) and DHWS (Eqs. (9)–(14)), and we assume that the fast-charging mode of the EV is only enabled when engaging in the DR program.

2.4. Demand response signal

In this section, we used the DS to predict DR flexibility and help in the implementation of DR management. As per the instructions in DR management, the DS is sent from the EVA or DHWSA to the distributed EVs or DHWSs. The different values of DS represent the different DR requirements, and the distributed EVs and DHWSs need to switch their electricity consumption according to the different values of DS they received. We defined three DS values to represent the corresponding responses of the EV and DHWS, as listed in Table 1. To better clarify the electric power and state changes of a single EV and DHWS corresponding to different DS values, a schematic diagram is shown in Fig. 3. Note that the same DS was sent from the EVA or DHWSA to all the distributed EVs and DHWSs in this study.

2.5. DS-based flexibility formulation

As we already set the DS in Section 2.4, it can be inferred that $\hat{P}_{EVA}(t)$ and $\hat{P}_{DHWSA}(t)$ under DS = 0 in Eq. (1) is equal to $P_{EVA}(t)$ and $P_{DHWSA}(t)$, respectively. Therefore, Eq. (1) can be rewritten as

$$\begin{cases} F_{EVA}(t) = \hat{P}_{EVA}(t)\Big|_{DS\neq 0} - \hat{P}_{EVA}(t)\Big|_{DS=0} \\ F_{DHWSA}(t) = \hat{P}_{DHWSA}(t)\Big|_{DS\neq 0} - \hat{P}_{DHWSA}(t)\Big|_{DS=0} \end{cases}, \tag{16}$$

where $\hat{P}_{EVA}(t)\Big|_{DS\neq 0}$ and $\hat{P}_{DHWSA}(t)\Big|_{DS\neq 0}$ represent the aggregated EV and DHWS power under DS ≠ 0, respectively. $\hat{P}_{EVA}(t)\Big|_{DS=0}$ and $\hat{P}_{DHWSA}(t)\Big|_{DS=0}$ represent the aggregated power of the EVs and DHWS under DS = 0, respectively.

Table 1
DS value and the corresponding reaction of single EV and DHWS.

DS value	Demand	EV response	DHWS response
1	Reserve down	Increase charging power or decrease discharging power	Increase the power consumption of the boiler
−1	Reserve up	Decrease charging power or increase discharging power	Decrease the power consumption of the boiler
0	Not engage in DR program	No need for state switching	No need for state switching

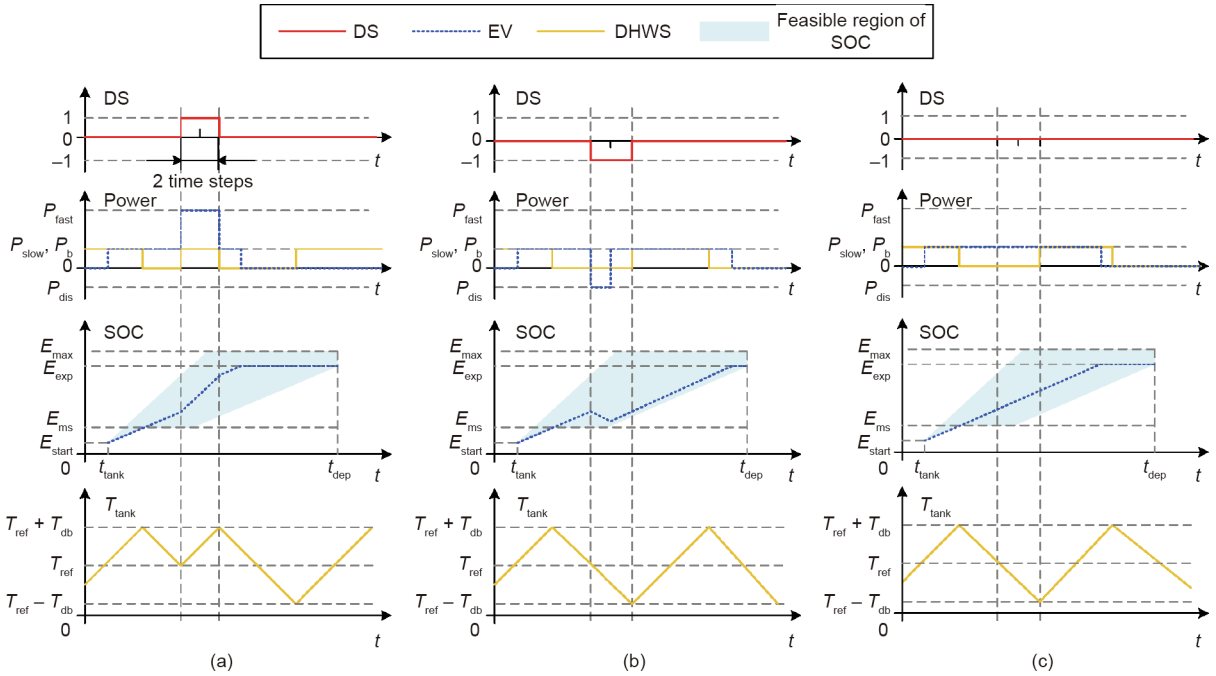

Fig. 3. The power consumption and state of a single EV and DHWS under different values of DS for two time steps: (a) DS = 1; (b) DS = −1; and (c) DS = 0.

DR programs usually require the flexibility of DR resources to provide a certain value that can be maintained for a period. Thus, a DS-based flexibility calculation method was proposed to derive flexibility under different maintenance times. First, we should fully activate the flexibility value of all the distributed EVs and DHWS for k time steps. Thus, the EVA and DHSWA can send $\mathrm{DS}(t)=1$, $\mathrm{DS}(t+1)=1$, …, $\mathrm{DS}(t+k)=1$ (or $\mathrm{DS}(t)=-1$, $\mathrm{DS}(t+1)=-1$, …, $\mathrm{DS}(t+k)=-1$) to all the distributed EVs and DHWS, which means increasing (or decreasing) electricity consumption at minimum remain k time steps. Then, the flexibility $F(t),F(t+1),\ldots,F(t+k)$ can be derived based on the DS signal. Based on the value of $F(t),F(t+1),\ldots,F(t+k)$, the flexibility remaining at minimum k time steps F^k can be obtained by finding the minimum value of $|F(t)|,|F(t+1)|,\ldots,|F(t+k)|$ as illustrated in Eq. (17), where the subscript indicates whether it is applied to EVA or DHWSA. Note that the physical meaning of finding the minimum value is to ensure that the flexibility can be presented in a block form with a constant value over several consecutive time steps. From Eq. (17), we can infer that, in addition to the base schedule, the maintenance time is also causes variation in flexibility as the energy of the DR resources is constrained. If the power causes the SOC reach the boundary of the EV battery capacity or the tank temperature of the DHWS to reach the boundary of the reference temperature, the DR resources cannot flow power to and from the grid. Thus, the flexibility decreases accordingly.

$$\begin{cases} \left|F^k_{\mathrm{EVA}}\right| = \min[|F_{\mathrm{EVA}}(t)|,|F_{\mathrm{EVA}}(t+1)|,\ldots,|F_{\mathrm{EVA}}(t+k)|] \\ \left|F^k_{\mathrm{DHWSA}}\right| = \min[|F_{\mathrm{DHWSA}}(t)|,|F_{\mathrm{DHWSA}}(t+1)|,\ldots,|F_{\mathrm{DHWSA}}(t+k)|] \end{cases} \tag{17}$$

Thus, the corresponding flexibility of the aggregated EVs and DHWSs can be obtained based on the above analysis and modeling method. To explain the proposed DS-based flexibility calculation process more intuitively, we draw an illustration in Fig. 3, which gives an example of a single EV and DHWS under different values of DS lasting for two time steps. The flexibility calculation results for the example shown in Fig. 3 can be found in Table 2.

Table 2
Flexibility calculation results corresponding to Fig. 3.

DS value and maintaining time	EV flexibility and maintaining time	DHWS flexibility and maintaining time
DS = 1, 2 time steps	P_{fast}−P_{slow}, 2 time steps	P_{b}, 2 time steps
DS = −1, 2 time steps	0	0

As defined in Eq. (16), when DS = 1, the flexibility of the example EV in Fig. 3 is P_{fast}−P_{slow}, and the flexibility can maintain two time steps. When DS = −1, the scheduling flexibility is P_{dis}−P_{slow}; however, considering the energy demand of EV users, the flexibility can only maintain one time step. Thus, if we calculate the flexibility of the example EV for two time steps, then its flexibility is 0 according to Eq. (16).

As for the example DHWS in Fig. 3, when DS = 1, it can provide a scheduling flexibility of P_{b} and maintenance time of two time steps. Owing to the tank temperature constraint, the DHWS cannot provide any flexibility when DS = −1. Considering that both EV and DHWS have time-coupling constraints, as shown in Eqs. (3) and (10), the flexibility of the next time step is affected by the flexibility of the previous time steps. Thus, if the state of the tank is heating water rather than losing heat when DS = −1, the DHWS can interrupt the heating to provide scheduling flexibility.

3. Flexibility prediction method and algorithm

3.1. Sequence to sequence (Seq2Seq)-based flexibility prediction method

To predict the flexibility of EVs and DHWSs multi-step-ahead, we first predict $\hat{P}_{\mathrm{EVA}}(t)$ and $\hat{P}_{\mathrm{DHWSA}}(t)$ under DS = 0 and DS ≠ 0 according to Eqs. (16) and (17). Here, we illustrate the prediction process. The analyses of $P_{\mathrm{EVA}}(t)$ and $P_{\mathrm{DHWSA}}(t)$ were similar. Therefore, to simplify, $P_{\mathrm{EVA}}(t)$ is analyzed only in the later paragraphs in Section 3.

We denote t as the first time step of the predicted slot. We denote $P_{\mathrm{EVA}}(t)$ as the base power for EVA (this would be different

in different strategies). As the first DS (DS(t)) changes $P_{\mathrm{EVA}}(t)$ to $\hat{P}_{\mathrm{EVA}}(t)$, we use $F(\cdot)$ to denote this transformation, and thus we have Eq. (18).

$$\hat{P}_{\mathrm{EVA}}(t) = F[P_{\mathrm{EVA}}(t), \mathrm{DS}(t), e_{\mathrm{EVA}}(t)]. \tag{18}$$

In Eq. (18), $e_{\mathrm{EVA}}(t)$ represents other external conditions that may contribute to the value of $\hat{P}_{\mathrm{EVA}}(t)$, which depends on the problem strategies. Note that the prediction of $P_{\mathrm{EVA}}(t)$ can be formulated as a time-series prediction model as in Refs. [26,27], which can also ensure the time coupling constraints of EV and DHWS. We use $f(\cdot)$ to denote the time-series predictor and the prediction of $P_{\mathrm{EVA}}(t)$ can be formulated as Eq. (19).

$$\begin{aligned} P_{\mathrm{EVA}}(t) = f[&P_{\mathrm{EVA}}(t-1), P_{\mathrm{EVA}}(t-2), ..., P_{\mathrm{EVA}}(t-N); \\ &u_{\mathrm{EVA}}(t), u_{\mathrm{EVA}}(t-1), ..., u_{\mathrm{EVA}}(t-N)] \end{aligned} \tag{19}$$

Similarly, $u_{\mathrm{EVA}}(t)$ is the external factor that may affect $P_{\mathrm{EVA}}(t)$, and N is the time dependency. Substituting Eq. (19) to Eq. (18), and applying the definition in Eqs. (20)–(22):

$$\begin{aligned} \mathrm{Encoder}[\dot{P}_{\mathrm{EVA}}, \dot{u}_{\mathrm{EVA}}] = f[&P_{\mathrm{EVA}}(t-1), P_{\mathrm{EVA}}(t-2), ..., P_{\mathrm{EVA}}(t-N); \\ &u_{\mathrm{EVA}}(t), u_{\mathrm{EVA}}(t-1), ..., u_{\mathrm{EVA}}(t-N)] \end{aligned} \tag{20}$$

$$\dot{u}_{\mathrm{EVA}} = [u_{\mathrm{EVA}}(t), u_{\mathrm{EVA}}(t-1), ..., u_{\mathrm{EVA}}(t-N)] \tag{21}$$

$$\dot{P}_{\mathrm{EVA}} = [P_{\mathrm{EVA}}(t-1), P_{\mathrm{EVA}}(t-2), ..., P_{\mathrm{EVA}}(t-N)] \tag{22}$$

We have the one time-step-ahead prediction as shown in Eq. (23).

$$\hat{P}_{\mathrm{EVA}}(t) = F\{\mathrm{DS}(t), e_{\mathrm{EVA}}(t), \mathrm{Encoder}[\dot{P}_{\mathrm{EVA}}, \dot{u}_{\mathrm{EVA}}]\} \tag{23}$$

Considering that we need to predict multi-step-ahead $\hat{P}_{\mathrm{EVA}}(t+k)$, we could repeatedly use one-step-ahead prediction as the next input, which takes the form of a multi-step prediction problem as in Refs. [26,27]. Therefore, we have Eq. (24).

$$\begin{aligned} \hat{P}_{\mathrm{EVA}}(t+k) = F\{&\hat{P}_{\mathrm{EVA}}(t+k-1), ..., \hat{P}_{\mathrm{EVA}}(t); \mathrm{DS}(t+k), ..., \mathrm{DS}(t); \\ &e_{\mathrm{EVA}}(t+k), ..., e_{\mathrm{EVA}}(t); \mathrm{Encoder}[\dot{P}_{\mathrm{EVA}}, \dot{u}_{\mathrm{EVA}}]\} \end{aligned} \tag{24}$$

Denoting $F(\cdot)$ as Decoder$[\cdot]$, we finally have Eq. (25).

$$\begin{aligned} \hat{P}_{\mathrm{EVA}}(t+k) = \mathrm{Decoder}\{&\hat{P}_{\mathrm{EVA}}(t+k-1), ..., \hat{P}_{\mathrm{EVA}}(t); \mathrm{DS}(t+k), ..., \mathrm{DS}(t); \\ &e_{\mathrm{EVA}}(t+k), ..., e_{\mathrm{EVA}}(t); \mathrm{Encoder}[\dot{P}_{\mathrm{EVA}}, \dot{u}_{\mathrm{EVA}}]\} \end{aligned} \tag{25}$$

The same procedures could be applied to $\hat{P}_{\mathrm{DHWSA}}(t)$, thus we have Eq. (26)

$$\begin{aligned} \hat{P}_{\mathrm{DHWSA}}(t+k) = \mathrm{Decoder}\{&\hat{P}_{\mathrm{DHWSA}}(t+k-1), ..., \hat{P}_{\mathrm{DHWSA}}(t); \\ &\mathrm{DS}(t+k), ..., \mathrm{DS}(t); \ e_{\mathrm{DHWSA}}(t+k), ..., e_{\mathrm{DHWSA}}(t); \\ &\mathrm{Encoder}[\dot{P}_{\mathrm{DHWSA}}, \dot{u}_{\mathrm{DHWSA}}(t)]\} \end{aligned} \tag{26}$$

This form is similar to the form of Seq2Seq model in the field of natural language process (NLP) [28]; thus, we could use the most advanced model, called "Transformer" in NLP to achieve such prediction.

3.2. Transformer

"Transformer" was first introduced in 2017 [29]. Prior to that, to achieve a Seq2Seq model, various types of RNNs and convolutional neural networks (CNNs) were the major choices. However, these RNN/CNN-based models are associated with many problems, particularly RNN-based models. One of the most severe problems in RNN-based models is the long-time dependency. When proceeding with long-sequence input, RNN-based models could suffer gradient exploration and gradient vanish [23,24], which results in poor performance. Although some updated RNN structures such as long short-term memory (LSTM) [30] and gate recurrent unit [31] have been proposed, these updated RNNs also fail as the sequence becomes longer.

The transformer shows its great ability in proceeding with long-time dependency due to its multi-head attention and self-attention mechanisms, which are the core of the model. These mechanisms enable the model to determine the inputs that are more important and those that are less dynamic. It is easy for the transformer to learn the time-coupling relationship through the training dataset. In our study, one day was divided into 96 time steps, which made the input sequence longer; thus, there is no doubt that the transformer would be a suitable model to solve this problem. In addition, the multi-head attention and self-attention mechanisms also make the transformer applicable to more general cases. As we need to predict the flexibility of two DR resources, we used this stronger tool.

The structure of the transformer is relatively complex; therefore, we will not explain the detailed mechanism. More details about the transformer model can be found in Ref. [29].

3.3. TCN-based input embedding module

In ordinary transformer, the input needs to pass through an "input embedding layer" before it is sent to the internal structure. In NLP, the function of the input embedding layer is to use a much lower-dimensional vector to represent the relatively high-dimensional one-hot encoding of words. More generally, this can be regarded as a type of feature extraction method. Based on this, we found a way to achieve feature extraction of the input in our study.

CNN may be a good choice for our study, as the convolution operation is a strong tool for extracting information from the original input. Considering that the inputs in this study are time-series, it is natural that a structure that can capture sequence information may work. Among different kinds of CNN-based models, the model called "TCN" proposed in 2018 [32] is one of the most advanced CNN-based models to handle temporal sequence. In the later section of this paper, we show how to use a type of TCN that could achieve such feature extractions in the original transformer model.

The core structures of TCN are shown in Fig. 4. The left figure, which is called "dilated causal conv," represents the integration of causal convolution and dilated convolution, which enables a shorter sequence to represent a longer sequence. The number of blue lines connected to a square is "kernel size" (in this figure, the kernel size = 3). "d" refers to dilation factors, representing the distance between two adjacent blue lines connected to a square. As we can see, when the kernel size is 3, a network of depth 4 (the number of layers) makes the final output (the blue square on the top layer) contains all the information of the input. The right figure, which is called "residual block in TCN," represents a designed structure to make the network deeper without suffering degradation [33]. This structure is called "residual connection," which is proposed in Ref. [33]. Weight norm [34] scales the layer parameters, and thus improves the model's performance.

Finally, by connecting several residual blocks in series, we got the overall model of TCN, namely "TCN-based input embedding module."

It should be noted that in the TCN model, the lengths of the input and output are identical; however, we could only choose the outputs of the last 96 time steps on the top of the TCN structure as the input of the transformer. There are two reasons for this. On the one hand, the last 96 outputs contained all the information of

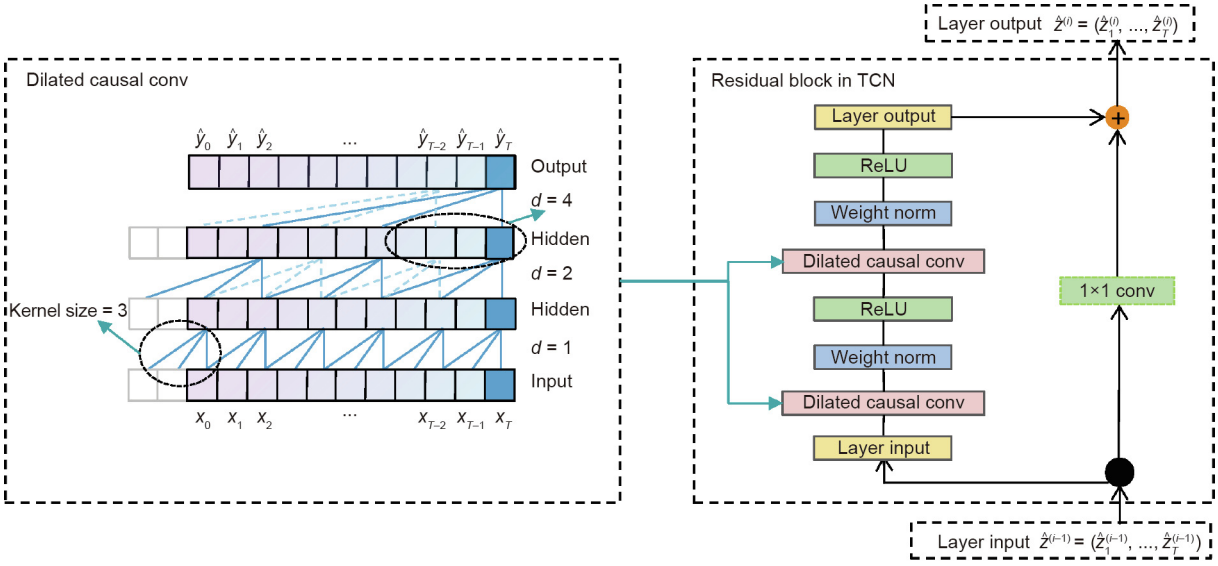

Fig. 4. The core structures of TCN. ReLU: rectified linear unit.

the input sequence by reasonable design; on the other hand, a shorter sequence could decrease the computational burden of the transformer. We call the overall structure the TCN-combined transformer model, which is plotted in Fig. 5. It can be seen that this model structure exactly achieves the form of Eqs. (25) and (26).

3.4. Training details

3.4.1. Model inputs and outputs

In Eqs. (25) and (26), we set up a general form for prediction. In this section, we make the form in Eqs. (25) and (26) more detailed. As mentioned earlier, the outputs of the model are $\hat{P}_{\mathrm{EVA}}(t)$ and $\hat{P}_{\mathrm{DHWSA}}(t)$ under different DS values. As for the input, in addition to $\hat{P}_{\mathrm{EVA}}(t+k-1),\ldots,\hat{P}_{\mathrm{EVA}}(t)$, $\mathrm{DS}(t+k),\ldots,\mathrm{DS}(t)$, and $\dot{P}_{\mathrm{EVA}}$, we need to specify variables including N, $\dot{u}_{\mathrm{EVA}}$, $\dot{u}_{\mathrm{DHWSA}}$, e_{EVA}, and e_{DHWSA} based on the characteristics of the problems. First, we set $N=192$ for all scenarios. As one day was divided into 96 time steps in this study, $N=192$, which indicates that the inputs over the past two days were considered. Other variables depend on different cases, including EVA/DHWSA and strategies 1 and 2. They are introduced individually.

3.4.2. Model inputs and outputs for the EVA

For strategy 1, we set $\dot{u}_{\mathrm{EVA}}=\dot{0}$ and $[e_{\mathrm{EVA}}(t+k),\ldots,e_{\mathrm{EVA}}(t)]=\dot{0}$, which means there were no external factors in the encoder and decoder. In this scenario, Eq. (25) can be rewritten as Eqs. (27) and (28).

$$\hat{P}_{\mathrm{EVA}}(t+k)=\mathrm{Decoder}\{\hat{P}_{\mathrm{EVA}}(t+k-1),\ldots,\hat{P}_{\mathrm{EVA}}(t);\ \mathrm{DS}(t+k),\ldots,\mathrm{DS}(t);\ \mathrm{Encoder}[\dot{P}_{\mathrm{EVA}}]\} \tag{27}$$

$$\dot{P}_{\mathrm{EVA}}=[P_{\mathrm{EVA}}(t-1),P_{\mathrm{EVA}}(t-2),\ldots,P_{\mathrm{EVA}}(t-192)]. \tag{28}$$

For strategy 2, considering that the base power is set according to the ToU tariff, we took ToU Tariff as the external factor, both in $\dot{u}_{\mathrm{EVA}}$ and e_{EVA}. We have Eqs. (29)–(31):

$$\hat{P}_{\mathrm{EVA}}(t+k)=\mathrm{Decoder}\{\hat{P}_{\mathrm{EVA}}(t+k-1),\ldots,\hat{P}_{\mathrm{EVA}}(t);\ \mathrm{DS}(t+k),\ldots,\mathrm{DS}(t);\ e_{\mathrm{EVA}}(t+k),\ldots,e_{\mathrm{EVA}}(t);\ \mathrm{Encoder}[\dot{P}_{\mathrm{EVA}},\dot{u}_{\mathrm{EVA}}]\} \tag{29}$$

$$\begin{aligned}\dot{u}_{\mathrm{EVA}}&=[u_{\mathrm{EVA}}(t),u_{\mathrm{EVA}}(t-1),\ldots,u_{\mathrm{EVA}}(t-192)]\\&=[c_{\mathrm{ToU}}(t),\ldots,c_{\mathrm{ToU}}(t-192)]\end{aligned} \tag{30}$$

$$[e_{\mathrm{EVA}}(t+k),\ldots,e_{\mathrm{EVA}}(t)]=[c_{\mathrm{ToU}}(t+k),\ldots,c_{\mathrm{ToU}}(t)]. \tag{31}$$

3.4.3. Model inputs and outputs for DHWSA

The characteristics of DHWSA were similar to those of EVA; thus, for strategy 1, the input and output were the same as those of EVA, as shown in Eqs. (32) and (33).

$$\hat{P}_{\mathrm{DHWSA}}(t+k)=\mathrm{Decoder}\{\hat{P}_{\mathrm{DHWSA}}(t+k-1),\ldots,\hat{P}_{\mathrm{DHWSA}}(t);\ \mathrm{DS}(t+k),\ldots,\mathrm{DS}(t);\ \mathrm{Encoder}[\dot{P}_{\mathrm{DHWSA}}]\} \tag{32}$$

$$\dot{P}_{\mathrm{DHWSA}}=[P_{\mathrm{DHWSA}}(t-1),P_{\mathrm{DHWSA}}(t-2),\ldots,P_{\mathrm{DHWSA}}(t-192)] \tag{33}$$

For strategy 2, similarly we have Eqs. (34)–(36)

$$\hat{P}_{\mathrm{DHWSA}}(t+k)=\mathrm{Decoder}\{\hat{P}_{\mathrm{DHWSA}}(t+k-1),\ldots,\hat{P}_{\mathrm{DHWSA}}(t);\ \mathrm{DS}(t+k),\ldots,\mathrm{DS}(t);\ e_{\mathrm{DHWSA}}(t+k),\ldots,e_{\mathrm{DHWSA}}(t);\ \mathrm{Encoder}[\dot{P}_{\mathrm{DHWSA}},\dot{u}_{\mathrm{DHWSA}}]\} \tag{34}$$

$$\begin{aligned}\dot{u}_{\mathrm{DHWSA}}&=[u_{\mathrm{DHWSA}}(t),u_{\mathrm{DHWSA}}(t-1),\ldots,u_{\mathrm{DHWSA}}(t-192)]\\&=[c_{\mathrm{ToU}}(t),\ldots,c_{\mathrm{ToU}}(t-192)]\end{aligned} \tag{35}$$

$$[e_{\mathrm{DHWSA}}(t+k),\ldots,e_{\mathrm{DHWSA}}(t)]=[c_{\mathrm{ToU}}(t+k),\ldots,c_{\mathrm{ToU}}(t)] \tag{36}$$

3.4.4. Hyperparameter settings

To train the model, the hyperparameters must be specified. The hyperparameters used in this study are listed in Table 3. The Adam algorithm performs well in network training, and can be easily applied to many machine learning tasks [35]. We chose the Adam optimizer to train the proposed network.

3.4.5. Training dataset

We needed to specify the training dataset. We created the training dataset according to the required variables, namely, the inputs and outputs in Eqs. (27)–(36). The training dataset for the flexibility prediction of EVA and DHWSA is introduced as follows. Please

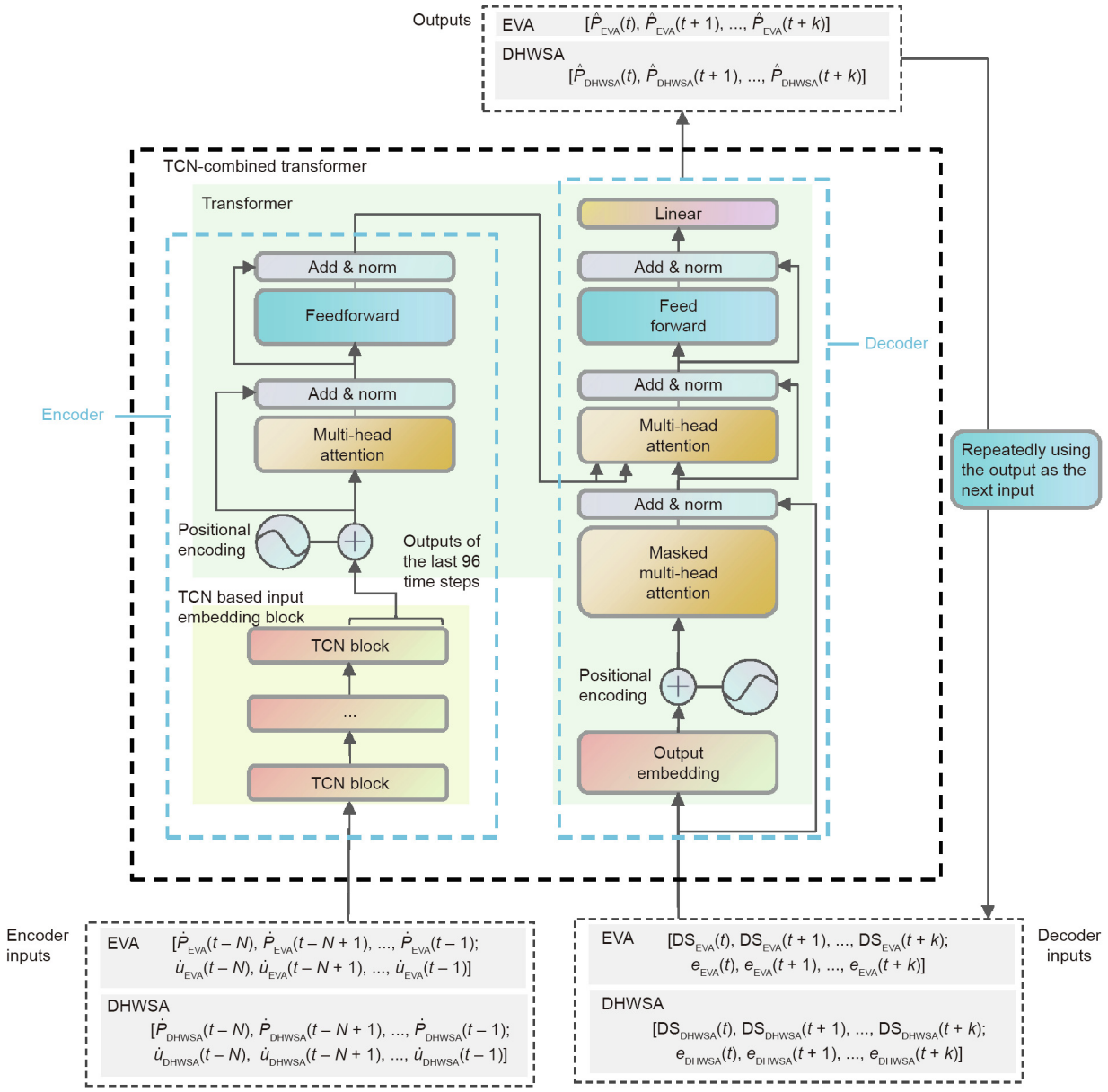

Fig. 5. TCN-combined transformer.

Table 3
Hyperparameter for our TCN-combined transformer model.

Model structure	Hyperparameter	Value
TCN-based input embedding block	Number of residual blocks	4
	Number of channels	64
	Kernel size	4
	Dropout rate	0.1
Transformer	Encoder layers	3
	Encoder heads	8
	Encoder feedforward layer dims	256
	Decoder layers	3
	Decoder heads	8
	Decoder feedforward layer dims	256
	Dropout rate	0.1

note that we needed to train several different models as the characteristics are different for different scenarios (EVA or DHWSA under different strategies).

3.4.6. Training dataset for EVA

To train the TCN-combined transformer model, we assumed that the historical information of all the distributed EVs, including the arrival time, departure time, battery capacity, and state of charge, can be collected by EVA. Then, based on Eqs. (2)–(8), we generated the base electricity consumption power under strategy 1. Based on Eqs. (2)–(8) and (15), we obtained the base electricity consumption power under strategy 2. Thus, we obtained the base schedule power P_{EVA} with no DS in strategies 1 and 2. Based on the base power, according to Table 1, EVA can derive the EV charging/discharging power $\hat{P}_{EVA}$ under DS = 1 and −1 on each day.

3.4.7. Training dataset for DHWSA

The basic information of DHWS, including physical parameters, water temperature, and switching information, was needed for the training of the DHWSA network. After collecting the historical data of the DHWS, based on Eqs. (9)–(14), we produced the base

electricity consumption power under strategy 1. Based on Eqs. (9)–(15), we produced the base electricity consumption power under strategy 2. Thus, we can obtain the base schedule power P_{DHWSA} with no DS in strategies 1 and 2. Based on the base power, according to Table 1, EVA can derive the EV charging/discharging power $\hat{P}_{DHWSA}$ under DS = 1 and −1 on each day.

Please note that for strategy 2, the ToU tariff on each day needed to be collected for both EVA and DHWSA.

3.5. Application

After training, our TCN-combined transformer model could provide accurate predictions. In real time, once we want to predict the flexibility multi-step-ahead, a new set of EV/DHWS data is used as input to the training process of EVA/DHWSA under different electricity consumption strategies. The corresponding power of EVA/DHWSA activated by different values of DS (including DS = 0 and DS ≠ 0) can be generated for one-step-ahead. For example, if we need to predict the up-regulation flexibility at time step $t + k$, we can make DS($t + k$) = −1 as the decoder inputs, and obtain the network output of the prediction power at $t + k$ under DS($t + k$) = −1. Next, we make DS($t + k$) = 0 as the decoder inputs, and obtain the network output of the prediction power at $t + k$ under DS($t + k$) = 0. Using Eq. (16), we obtain the up-regulation flexibility using the prediction power under DS($t + k$) = −1 and by subtracting the prediction power under DS($t + k$) = 0. Then, using the output as the next input, the corresponding power multi-step-ahead can be derived. The multi-step-ahead flexibility can be easily obtained using Eqs. (16) and (17).

3.6. Summary of proposed algorithm

In this section, we introduce Algorithm 1 as a summary of the overall procedure of the proposed TCN-combined transformer-based flexibility prediction. Note that in application, the EVA/DHWSA only needs to repeat step 6 after the models are trained.

Algorithm 1. Algorithm for the TCN-combined transformer-based DR flexibility prediction.

1. Prepare 4 TCN-combined transformer models to predict the flexibility of EVA and DHWS under strategies 1 and 2, respectively.
2. Get historical dataset of EVA and DHWSA considering their characteristics for strategy 2; ToU tariff should also be collected.
3. Specify the inputs and outputs of the model for EVA and DHWSA under different strategies based on Eqs. (27)–(36).
4. Set the model hyperparameters for EVA and DHWSA under different strategies.
5. Train the 4 models.
6. Input the real-time data of EV/DHWS to the network and achieve a one-step-ahead prediction. Then, repeatedly use the output as the next input to achieve the multi-step-ahead prediction; thus, the multi-step-ahead flexibility can be predicted using Eqs. (16) and (17).

4. Case study

To demonstrate the effectiveness of the proposed TCN-combined transformer-based method in the flexibility prediction of EVs and DHWs, simulations of case studies were carried out, and results are presented in this section. We used a computer with an Intel® Core™ i7-7500U CPU @ 2.70–2.90 GHz, 8 GB of RAM, and RTX2060, as well as a cloud server with 6 × Intel® *Xeon*® CPU E5-2678 v3 @ 2.50 GHz, 11 GB of RAM, and RTX2080 Ti to run all the simulations.

4.1. Basic data

The simulation was based on a residential area with 2000 DHWSs and 1000 EVs. The data and parameters of the parking behavior of the EVs were obtained from Ref. [36]. The data and parameters of the usage behavior of the DHWSs were obtained from Ref. [37]. The key parameters mentioned above are listed in Table 4. The two strategies data came from 256 weekdays per year. We selected the first 196 days to form the training dataset of the transformer network. To avoid over-fitting of the network, the remaining 197–226 days were selected as the validation data set. Finally, the 30 days from 227th to 256th days were randomly selected to test the performance of the proposed algorithm. The ToU tariffs for EV charging stations and residential areas are shown in Fig. 6.

Owing to the differences between EVs and DHWSs in power consumption time, the studied horizon for DHWSs in one day was from 00:00 to 24:00, while the studied horizon for EVs was from 12:00 to 12:00 on the next day. In this study, one time step was 15 min; thus, one day can be divided into 96 time steps.

4.2. Convergence performance of the transformer network

In this study, the mean square error between the target values and the network outputs in the training dataset was used as the loss function. Fig. 7 illustrates the network learning performances under the two strategies of the two loads. As shown in Fig. 7, all loss functions experienced a steep drop over the previous period of epochs, and as epochs increased, they decreased marginally and steadily at approximately 10^{-5}. Despite the slight fluctuations, the validation loss functions maintained a downward trend without any abnormal surge, which indicates that the learning processes avoided overfitting. Overall, considering the total training time of 4 h (600 epochs) and 6 h (1000 epochs) as well as a high-grade convergence performance, it was proven that the proposed network structure achieved a balance between complexity and performance.

4.3. Flexibility prediction results analysis

To evaluate the prediction performance of EVA and DHWSA flexibility under different maintenance times, we compared the results under strategy 1 (without considering optimization power consumption) and strategy 2 (considering optimization power consumption) in six scenarios: ① DS = −1 maintained for 15 min; ② DS = −1 maintained for 30 min; ③ DS = −1 maintained for 60 min; ④ DS = 1 maintained for 15 min; ⑤ DS = 1 maintained for 30 min; ⑥ DS = 1 maintained for 60 min. Moreover, to validate if our proposed multi-step-ahead flexibility prediction method can predict a long-range slot, we selected a 1-day predicted slot, which means we needed to test the accuracy of the 96-time-step-ahead prediction; hence, k = 96. Thus, our multi-step-ahead prediction was achieved by repeatedly making a one-step-ahead prediction.

Table 4
Some Parameters of EVs and DHWSs.

Parameter	Value	Parameter	Value
P_{fast}	9 kW	E_{max}	0.9 × 30 kW·h
P_{slow}	3.3 kW	E_{exp}	0.85 × 30 kW·h
P_{dis}	−3.3 kW	P_b	3 kW
E_{ms}	0.5 × 30 kW·h	V_{tank}	300 L

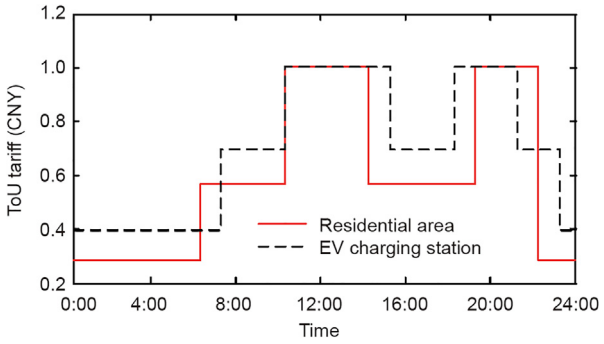

Fig. 6. ToU tariff for EV charging station and residential area.

In real time, it can be used to predict any-step-ahead flexibility at any time, depending on the situation.

To illustrate the superiority of the proposed TCN-combined transformer-based algorithm, we also used the LSTM method (a type of updated RNN structure, as introduced in Section 3.2) to predict the flexibility described in this section. The trained LSTM contained one input layer, two hidden layers, and 128 nodes in each hidden layer.

4.3.1. Flexibility prediction results in one day

Figs. 8–11 demonstrate the flexibility prediction results of a randomly selected day in the test dataset. It can be seen from the figures that the peak values of flexibility are slightly more difficult to predict than with other values by both the TCN-combined transformer and LSTM. However, in general, compared with the real value, the predictions of the TCN-combined transformer were more accurate than those of the LSTM. Note that the real value was derived by sending DS to EVs and DHWSs in real time, as discussed in Sections 2.4 and 2.5.

As shown in Figs. 8–11, different maintenance times had different flexibility values. Flexibility size decreased with an increase in maintenance time. As the DR resources are decentralized, and with large uncertainty, it is difficult for them to provide aggregated stable flexibility over a long period of time. Thus, the maintenance time of DR resources in the DR program should be set at a smaller value, which is conducive to activate more potential DR flexibility.

In addition, the flexibilities using strategies 1 and 2 were different between periods. For EVA, the flexibility mainly appeared between 16:00 and 8:00. The positive flexibility (absorbing power from the grid) under strategy 2 was greater than that under strategy 1, while the negative flexibility (releasing power to the grid) under strategy 2 was less than that under strategy 1. For DHWSA, the flexibility appeared all day. The period between 8:00 and 24:00 for strategy 2 provided more positive flexibility and less negative flexibility, while the period between 0:00 and 8:00 for strategy 2 provided less positive flexibility and more negative flexibility. Furthermore, the lower the maintenance time, the more significant is the difference.

Moreover, under the same strategy, DR resources exhibited different characteristics in different scenarios. For EVA, an EV has the potential to discharge once it is charged higher than the minimum SOC. However, when it is close to the departure time, discharging would hardly exist to meet the EV energy demand. Thus, when DS = −1, there was a negative flexibility between 16:00 and 8:00, with a maximum of −4.5 MW/15 min. When DS = 1, the EV SOC increased rapidly in the fast-charging mode; however, the EV did not continue to provide flexibility owing to the limitation of the maximum battery capacity. Therefore, the EVA provided a maximum flexibility of approximately 3 MW/15 min from 16:00 to 24:00 when DS = 1. For DHWSA, the positive and negative flexibilities were distributed relatively uniformly during the predicted slot. Due to the changes in the use of the DHWS (e.g., bathing in the evening, showering in the

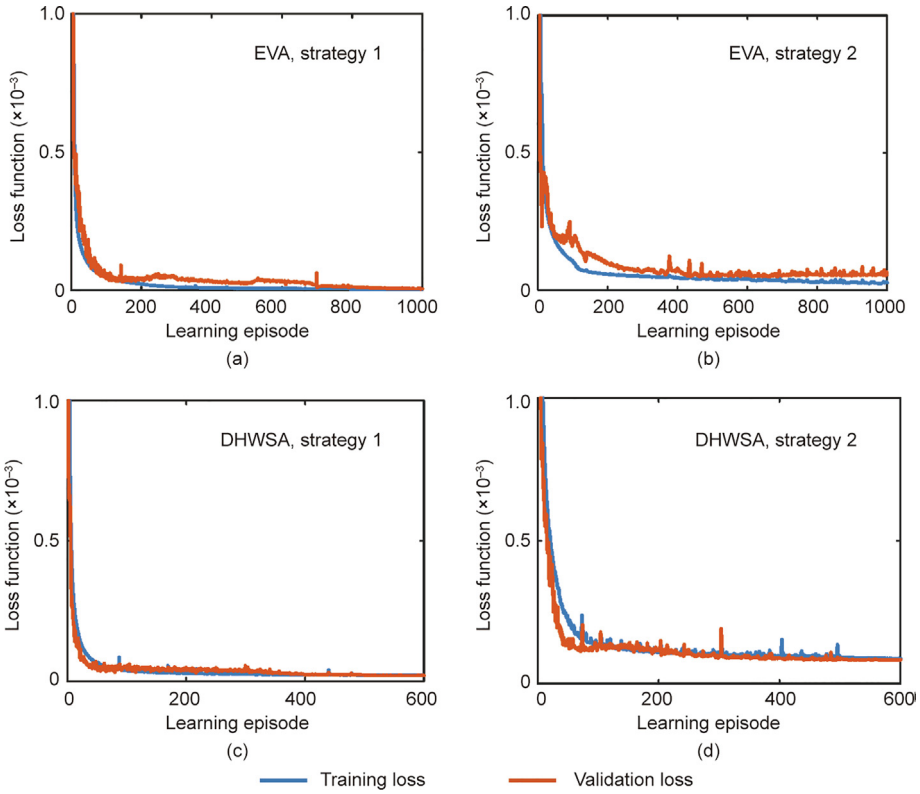

Fig. 7. Loss function of EVA and DHWSA under two scenarios.

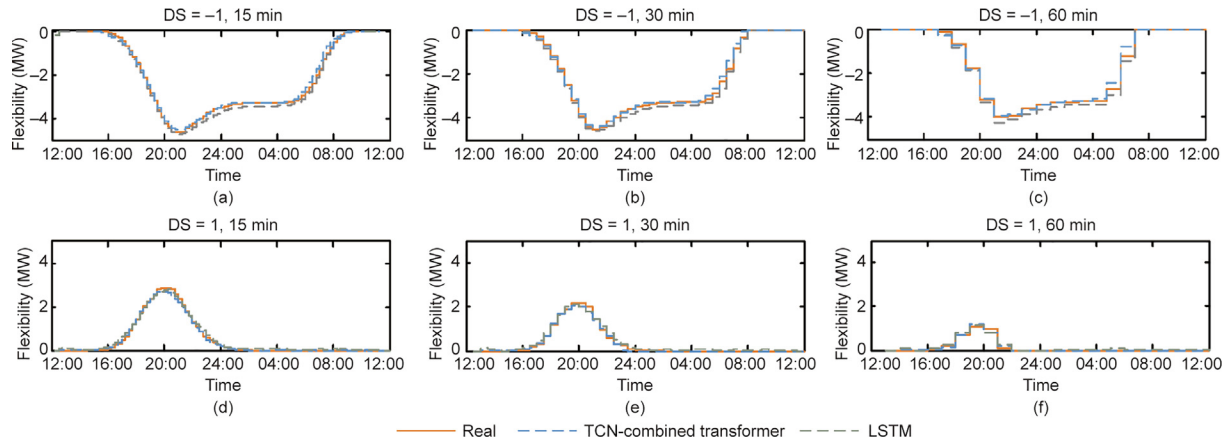

Fig. 8. The flexibility prediction results of EVA under strategy 1.

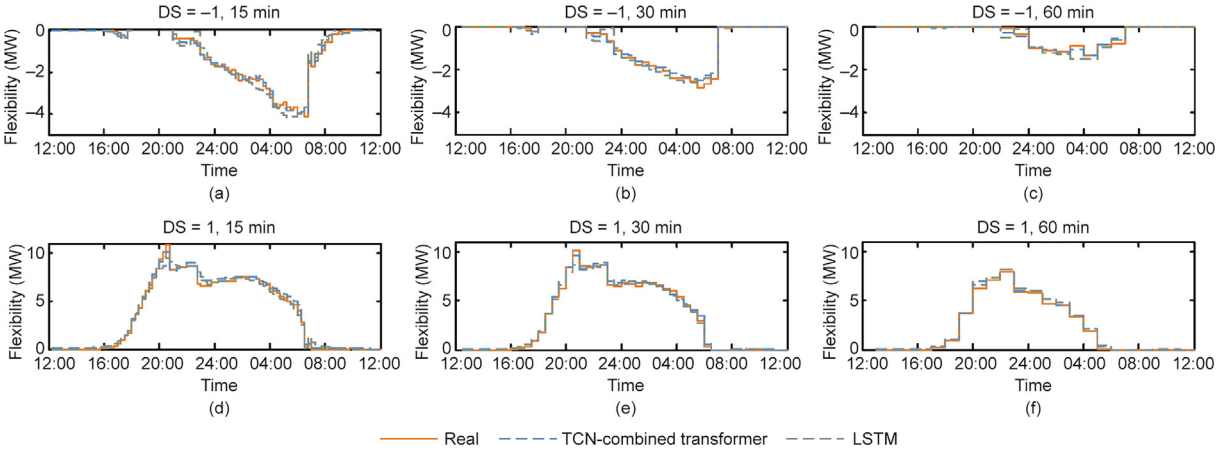

Fig. 9. The flexibility prediction results of EVA under strategy 2.

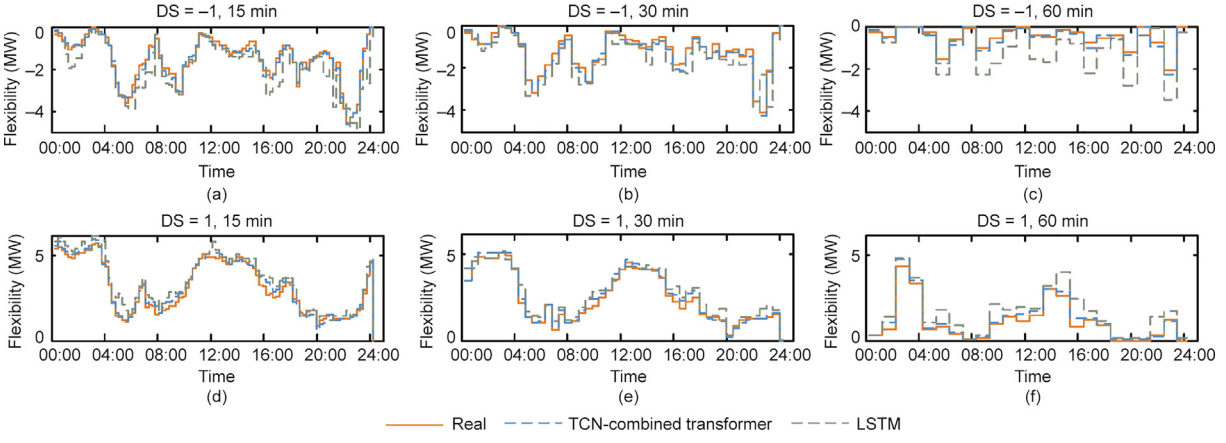

Fig. 10. The flexibility prediction results of DHWSA under strategy 1.

morning, barely no usage between 13:00 and 16:00, and delay of temperature after usage), which were described in Ref. [37], the negative flexibility reached low points of approximately −4.5 to −4.8 MW/15 min at around 17:00, while the positive flexibility reached a peak point of approximately −5 MW/15 min between 00:00 and 8:00.

4.3.2. Prediction accuracy for 30 consecutive days

Tables 5–8 provide information on the mean absolute error (MAE, the average absolute value of the error between the predicted and real values) of the 30-day prediction under different strategies and scenarios. In Figs. 8–11 and Tables 5–8, it can be seen that the predicted flexibility of the proposed TCN-combined

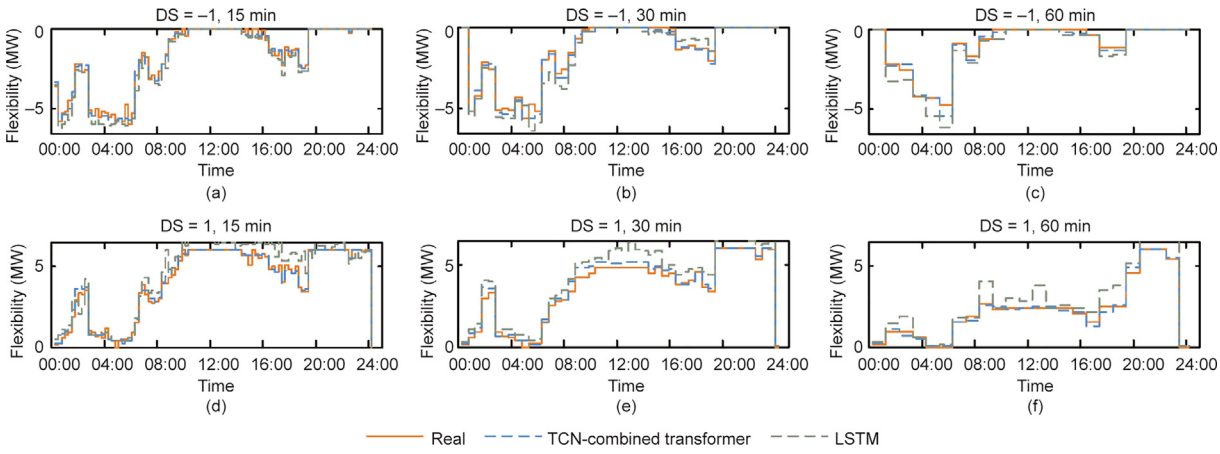

Fig. 11. The flexibility prediction results of DHWSA under strategy 2.

transformer was closer to the real value than that of the LSTM. In the simpler scenario, that is, EVA prediction under strategy 1, the TCN-combined transformer performed slightly better than the LSTM. As the scenario became more complex, such as EVA prediction under strategy 2 and DHWSA prediction under both strategies 1 and 2 (the DHWSA experienced greater complexity owing to the large number of DHWS users), the prediction accuracy of the TCN-combined transformer was relatively better than that of the LSTM. As described in Section 3.2, LSTM is an updated RNN structure, which can significantly improve the long-time dependency problem of the RNN. With the increase in influencing factors, such as the ToU tariff, number of users, and randomness of user behavior, the LSTM still experienced difficulty of fully capturing the information in the input sequence. In general cases, with the multi-head attention and self-attention mechanisms, the TCN-combined transformer can easily learn time-series information through the input sequence. It was found that the predicted flexibility of the TCN-combined transformer was close to the real value with high accuracy, and that the TCN-combined transformer is more suitable for flexibility prediction of DR resources in different electricity consumption scenarios.

4.4. Future application prospect for the proposed flexibility prediction method

As discussed in Section 2.1, the negative flexibility could provide up reserve, and the positive flexibility could provide down reserve. From the prediction, the flexibility of aggregated DR resources is obtained through multi-step-ahead, and the schedulable flexibility of the aggregator's owned resources can be predicted at the day-ahead and real-time levels. The obtained prediction results can provide a reference for the decision-making of aggregator or superior scheduling organizations, thus achieving further cooperation between the aggregator and operators of different levels of electric power systems, such as transmission system operators (TSOs) and distribution system operators (DSOs).

At the day-ahead level, by inputting different DS signals, the amount and maintenance time of the flexibility can be obtained and used to grasp the overall distribution of flexibility and make a general assessment of resource flexibility. The aggregator can trade the flexibility of DR resources in the electrical energy and ancillary service markets based on the day-ahead forecast results. For DR users, electricity cost savings can be achieved. For the power system operator, the power consumption of DR resources and the reserve capacity of the power system can be ensured in advance, which is conducive for reducing the impact of random electricity consumption behavior of DR resources and uncertain renewable energy production on the power system, thereby promoting load balance and ensuring part of the emergency reserve. For example, according to Fig. 11, under strategy 2, there was adequate downward reserve flexibility of EVs in the period of 20:00 and 24:00. The upward reserve flexibility was adequate between 4:00 and 8:00 in short maintenance time (15–30 min), distributing uniformly and decreasing between 24:00 and 8:00 when the maintenance time reached 60 min. Thus, in the day ahead, based on this projected result, aggregators are inclined to offer downward reserve between 20:00 and 24:00 and upward reserve between 4:00 and 8:00 in the reserve market, as a more robust decision. Thus, the reserve capacity can be used for system operation, such as frequency regulation at the TSO level or congestion alleviation at the DSO level.

At the real-time level, the flexibility and maintenance time of DR resources defined in the day ahead can be updated by continuously rolling the prediction. As real-time prediction is an ultra-short-term prediction, it can achieve a more accurate prediction than that at the day-ahead stage. The updated flexibility is reported to the system operator accordingly; thus, the system operator is aware of the available reserve in a timely manner. In addition, the aggregator may need to participate in intra-day market, balancing market, and real-time operation; thus, the updated flexibility is necessary in the aggregator optimization problem, which defines the upper and lower boundaries of power of distributed energy resources.

Table 5
MAE for 30 d of the EVA prediction results under strategy 1.

DS	Prediction method	15 min (MW)	30 min (MW)	60 min (MW)
1 (down)	LSTM	0.055	0.056	0.042
	TCN-combined transformer	0.028	0.027	0.015
−1 (up)	LSTM	0.105	0.102	0.106
	TCN-combined transformer	0.085	0.083	0.088

Table 6
MAE for 30 d of the EVA prediction results under strategy 2.

DS	Prediction method	15 min (MW)	30 min (MW)	60 min (MW)
1 (down)	LSTM	0.230	0.180	0.158
	TCN-combined transformer	0.177	0.155	0.148
−1 (up)	LSTM	0.147	0.116	0.092
	TCN-combined transformer	0.094	0.073	0.052

Table 7
MAE for 30 d of DHWSA prediction results under strategy 1.

DS	Prediction method	15 min (MW)	30 min (MW)	60 min (MW)
1 (down)	LSTM	0.273	0.2519	0.274
	TCN-combined transformer	0.147	0.1480	0.121
−1 (up)	LSTM	0.282	0.1987	0.178
	TCN-combined transformer	0.185	0.1930	0.175

Table 8
MAE for 30 d of DHWSA prediction results under strategy 2.

DS	Prediction method	15 min (MW)	30 min (MW)	60 min (MW)
1 (down)	LSTM	0.234	0.191	0.169
	TCN-combined transformer	0.139	0.141	0.138
−1 (up)	LSTM	0.221	0.217	0.179
	TCN-combined transformer	0.140	0.163	0.143

5. Conclusions

A TCN-combined transformer-based algorithm was developed in this study to predict both the size and maintenance time of EV and DHWS flexibility in the DR program. The prediction is based on the network training of historical power consumption behaviors of EVs and DHWS, as well as the DSs. The accuracy of the flexibility prediction was verified through the case study of a group of residential EV and DHWS end-users under different power consumption strategies. We can infer that the flexibility size would decrease with an increase in maintenance time. In general, the proposed TCN-combined transformer-based flexibility prediction method can support DR scheduling in the daily grid operation.

The main challenge in the application of deep learning technology in the field of power grids is the lack of sufficient data as there is usually not enough data in engineering for such a learning-based method. Most of the measured data were of low quality. To further apply our proposed flexibility method, we should take advantage of the rapid development of smart meters and actively consider and carry out beneficial attempts to obtain real-world data. In addition, a further development in data generation technology is highly important to supplement a large amount of data when obtaining accurate data is difficult. As some real-world data can be lost in the transmission process, corresponding data repair technology can also be helpful in improving the quality of data.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (51877078 and 52061635102) and the Beijing Nova Program (Z201100006820106).

Compliance with ethics guidelines

Junjie Hu, Huayanran Zhou, Yihong Zhou, Haijing Zhang, Lars Nordströmd, and Guangya Yang declare that they have no conflicts of interest or financial conflicts to disclose.

References

[1] Hu J, Zhou H, Li Y, Hou P, Yang G. Multi-time scale energy management strategy of aggregator characterized by photovoltaic generation and electric vehicles. J Mod Power Syst Clean Energy 2020;8(4):727–36.
[2] Kumar A, Sekhar C. Demand response based congestion management in a mix of pool and bilateral electricity market model. Front Energy 2012;6(2):164–78.
[3] Hu J, Yang G, Ziras C, Kok K. Aggregator operation in the balancing market through network-constrained transactive energy. IEEE Trans Power Syst 2019;34(5):4071–80.
[4] Xue Y, Yu X. Beyond smart grid—cyber–physical–social system in energy future. Proc IEEE 2017;105(12):2290–2.
[5] International energy agency. Harnessing variable renewables: a guide to the balancing challenge. Paris: OECD Publishing; 2011.
[6] Chen Z, Sun Y, Ai X, Malik SM, Yang L. Integrated demand response characteristics of industrial park: a review. J Mod Power Syst Clean Energy 2020;8(1):15–26.
[7] Fattahi J, Samadi M, Erol-Kantarci M, Schriemer H. Transactive demand response operation at the grid edge using the IEEE 2030.5 standard. Engineering 2020;6(7):801–11.
[8] Li Z, Guo Q, Sun H, Wang J. Storage-like devices in load leveling: complementarity constraints and a new and exact relaxation method. Appl Energy 2015;151:13–22.
[9] Sanandaji BM, Vincent TL, Poolla K. Ramping rate flexibility of residential HVAC loads. IEEE Trans Sustain Energy 2016;7(2):865–74.
[10] Hu K, Li W, Wang L, Cao S, Zhu F, Shou Z. Energy management for multi-microgrid system based on model predictive control. Front Inf Technol Electron Eng 2018;19(11):1340–51.
[11] Vivekananthan C, Mishra Y, Ledwich G, Li F. Demand response for residential appliances via customer reward scheme. IEEE Trans Smart Grid 2014;5(2):809–20.
[12] Wu J, Xue Y, Xie D, Yue D, Wen F, Zhao J. Evaluation and simulation analysis of reserve capability for electric vehicles. Autom Electr Power Syst 2018;42(13):101–7. Chinese.
[13] Wang J, Jia Y, Mi Z, Chen H, Fang H. Reserve service strategy of electric vehicles based on double-incentive mechanism. Autom Electr Power Syst 2020;44(10):68–76. Chinese.
[14] Yao W, Zhao J, Wen F, Xue Y, Chen F, Li L. Frequency regulation strategy for electric vehicles with centralized charging. Autom Electr Power Syst 2014;38(9):69–76. Chinese.
[15] Zhang B, Xu G. Rolling horizon optimization for grid-connected electric vehicles considering demand difference. Autom Electr Power Syst 2020;44(13):106–14. Chinese.
[16] Han B, Lu S, Xue F, Jiang L. Day-ahead electric vehicle aggregator bidding strategy using stochastic programming in an uncertain reserve market. IET Gener Transm Distrib 2019;13(12):2517–25.
[17] Wang Yi, Hug G, Liu Z, Zhang N. Modeling load forecast uncertainty using generative adversarial networks. Electr Power Syst Res 2020;189:106732.

[18] Chen K, Chen K, Wang Q, He Z, Hu J, He J. Short-term load forecasting with deep residual network. IEEE Trans Smart Grid 2019;10(4):3943–52.
[19] Hong T, Pinson P, Wang Yi, Weron R, Yang D, Zareipour H. Energy forecasting: a review and outlook. IEEE Open Access J Power and Energy 2020;7:376–88.
[20] Divshali PH, Evens C. Behaviour analysis of electrical vehicle flexibility based on large-scale charging data. In: 2019 IEEE Milan PowerTech; 2019 Jun 23–27; Milan, Italy; 2019.
[21] Sajjad IA, Chicco G, Napoli R. Definitions of demand flexibility for aggregate residential loads. IEEE Trans Smart Grid 2016;7(6):2633–43.
[22] Paridari K, Nordström L. Flexibility prediction, scheduling and control of aggregated TCLs. Electr Power Syst Res 2020;178:106004.
[23] Hochreiter S, Bengio Y, Frasconi P, Schmidhuber J. Gradient flow in recurrent nets: the diffifficulty of learning long-term dependencies. In: John FK, Stefan CK, editors. A field guide to dynamical recurrent networks. New York: IEEE Press; 2001. p. 237–44.
[24] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Netw 1994;5(2):157–66.
[25] Lu Z, Li H, Qiao Y. Power system flexibility planning and challenges considering high proportion of renewable energy. Autom Electr Power Syst 2016;40 (13):147–57. Chinese.
[26] Parlos AG, Rais OT, Atiya AF. Multi-step-ahead prediction using dynamic recurrent neural networks. Neural Netw 2000;13(7):765–86.
[27] Bao Y, Xiong T, Hu Z. Multi-step-ahead time series prediction using multiple-output support vector regression. Neurocomputing 2014;129:482–93.
[28] Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. 2014. arXiv:1409.3215v3.
[29] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. 2017. arXiv:1706.03762v5.
[30] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9 (8):1735–80.
[31] Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014. arXiv:1412.3555v1.
[32] Bai S, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. 2018. arXiv:1803.01271.
[33] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27–30; Las Vegas, NV, USA; 2016.
[34] Salimans T, Kingma DP. Weight normalization: a simple reparameterization to accelerate training of deep neural networks. 2016. arXiv:1602.07868.
[35] Kingma DP, Ba JL. Adam: a method for stochastic optimization. 2017. arXiv:1412.6980v9.
[36] Luo Z, Hu Z, Song Y, Xu Z, Lu H. Optimal coordination of plug-in electric vehicles in power grids with cost-benefit analysis—part II: a case study in china. IEEE Trans Power Syst 2013;28(4):3556–65.
[37] Sandels C, Widén J, Nordström L. Forecasting household consumer electricity load profiles with a combined physical and behavioral approach. Appl Energy 2014;131:267–78.