Title

Data-Driven Approach for Analyzing Spatiotemporal Price Elasticities of EV Public Charging Demands Based on Conditional Random Fields

Journal

IEEE Transactions on Smart Grid, 12(5)

ISSN

1949-3053

Authors

Bao, Zhiyuan
Hu, Zechun
Kammen, Daniel M
et al.

Publication Date

2021-09-01

DOI

10.1109/tsg.2021.3080460

Peer reviewed

Data-driven Approach for Analyzing Spatiotemporal Price Elasticities of EV Public Charging Demands Based on Conditional Random Fields

Zhiyuan Bao, *Student Member, IEEE,* Zechun Hu, *Senior Member, IEEE,* Daniel M. Kammen, and Yifan Su

***Abstract*—With the increase of electric vehicle (EV) sales, the pricing strategies of public charging stations have significant impacts on their revenues and the spatiotemporal distribution of charging loads. In this paper, we quantify three kinds of price elasticity of charging demands based on the historical charging data of multiple public charging stations with different pricing schemes. The relationship between the volume-weighted average price (VWAP) and the corresponding total charging demand within a zone is studied, which does not require to change the charging prices to estimate elasticity. To evaluate the shifting of charging demands in different periods and zones, a conditional random field (CRF) model is built, which depicts the spatiotemporal correlations of charging demands. In this model, the VWAPs and the total charging demands are taken as observed variables and hidden variables, respectively. The loopy belief propagation algorithm is used to infer the loopy graph approximately, and the learning algorithm with forgetting factors is used to estimate the unknown parameters of the CRF model. The price elasticities are derived from CRF, and the elasticity matrices of charging demands are obtained. Computational results based on historical charging data verify the validity of the proposed model and method.**

***Index Terms*—Electric vehicle, charging demands, price elasticity, charging price, conditional random field.**

NOMENCLATURE

A. Acronyms

EV	Electric vehicle.
VWAP	Volume-weighted average price.
CRF	Conditional random field.
TOU	Time of use.
BP	Belief propagation.
LBP	Loopy belief propagation.
ACS	Aggregated charging station.
RTP	Real-time price.

B. Sets

$\boldsymbol{G}$	Graph of hidden variables in CRF model.
$\boldsymbol{V}$	Set of vertices of the graph (ACSs in difference periods).
$\boldsymbol{E}$	Set of edges of the graph (Paths of charging demand shifting).
$\boldsymbol{E_t}$	Paths of temporal shifting of charging demands.
$\boldsymbol{E_k}$	Paths of spatial shifting of charging demands.
$\mathcal{X}$	Set of discrete values of charging demand.
$\mathcal{D}$	Dataset of charging price-demand pairs.
$\mathcal{T}$	Set of periods.
$\mathcal{K}$	Set of zones.
$N(\cdot)$	Neighbourhoods of node $\cdot$

C. Parameters

$\boldsymbol{\omega}$	Unknown parameters of CRF model.
$\boldsymbol{\rho}$	Charging prices of ACSs in different periods.
$\boldsymbol{d}$	Charging demands of ACSs in different periods.
$P_{\boldsymbol{\omega}}(\boldsymbol{d}\|\boldsymbol{\rho})$	Joint probability for charging demands under giving charging prices and CRF parameters.
$Z(\boldsymbol{\rho})$	Partition function.
$\psi_i(d_i, \rho_i)$	Feature function for charging demand and price in the ith ACS.
$\psi_{i,j}^t(d_i^t, d_j^t)$	Feature function between charging demands in the ith and jth ACS.
$\psi_i^{t_1,t_2}(d_i^{t_1}, d_i^{t_2})$	Feature function between charging demands of periods t_1, t_2 in the ith ACS.
ω_i	Parameter for feature function $\psi_i(d_i, \rho_i)$.
$\omega_{i,j}^t$	Parameter for feature function $\psi_{i,j}^t(d_i^t, d_j^t)$.
$\omega_i^{t_1,t_2}$	Parameter for feature $\psi_i^{t_1,t_2}(d_i^{t_1}, d_i^{t_2})$.
$S_{\boldsymbol{\omega}}(\boldsymbol{d}\|\boldsymbol{\rho})$	Score of "fitness" for demands under giving charging prices and CRF parameters.
$\ln L(\boldsymbol{d}\|\boldsymbol{\rho};\boldsymbol{\omega})$	Logarithmic likelihood function.
$\|\cdot\|$	Cardinality of set $\cdot$
$E(\cdot)$	Expectation of random variable $\cdot$
$Var(\cdot)$	Variance of random variable $\cdot$
$\boldsymbol{\omega}^*$	Optimal parameters for dataset $\mathcal{D}$.
α	Forgetting factor.
γ	Learning rate.
$m_{i,j}(d_j)$	Message from node i to j for charging demand d_j.
ε	Elasticity.
$\hat{d}$	Normalized charging demand.
$\overline{d}$	Average charging demand.
$\hat{\rho}$	Normalized charging price.
$\overline{\rho}$	Average charging price.

This work was supported in part by the National Natural Science Foundation of China under Grant U1766205. (*Corresponding author: Zechun Hu.*)

Z. Bao, Z. Hu, and Y. Su are with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (baozy19@mails.tsinghua.edu.cn; zechhu@tsinghua.edu.cn; suyf19@mails.tsinghua.edu.cn).

D. M. Kammen is with Energy and Resources Group, University of California Berkeley, Berkeley, California 94720, United States. (kammen @berkeley.edu)

I. Introduction

With the technical progresses and policy incentives, the sales of electric vehicles (EVs) are increasing very fast worldwide in recent years. To recharge the EVs, charging facilities have been installed and charging service providers are expanding their public charging networks in many areas. For those EVs, e.g. electric taxis, that need convenient and/or fast charging services in urban areas, a large part of their charging demands are met in public charging stations [1]. Due to the mobility of EVs, the EV drivers can schedule to recharge at different time periods [2]–[4] and also different charging stations [5], [6] according to their convenience and the charging prices. That is, for an area with a number of charging stations, the pricing strategy of each station can potentially affect the spatiotemporal distribution of charging demands. To measure the responses of charging demands to prices, it is necessary to quantify the price elasticity of charging demand, which also plays an important role in modeling consumer behaviors [7], determining pricing strategy for charging service providers, and setting price signals for distribution system operators or load serving entities [8], [9].

The mobility of the EV charging loads can be leveraged to improve the secure and economic operation of power systems [5]. There are methods to coordinate EV charging loads: 1) Set proper charging service prices to coordinate EV charging loads [10]–[14], e.g. setting time of use (TOU) charging prices or real time pricing (RTP) in residential areas; 2) Participate electricity markets through EV aggregator, controlling charging power to minimize charging cost or maximize market revenue [15]–[21].

According to [11], a proper pricing strategy is an indirect approach to mitigate EV charging impacts. For public charging stations, transportation network is considered in some references, e.g. [12] considered EV mobility and proposed a double-layer optimization model to determine the pricing scheme for fast charging stations. An online menu-based pricing strategy was presented in [13] to reduce the violations of power distribution system security constraints, i.e. line and transformer capacity limits, voltage magnitude limit and three-phase voltage imbalance limit. Reference [10] derived the sufficient and necessary conditions of valley-filling pricing mechanisms and proposed two pricing strategies for both non-cooperative and cooperative scenarios.

EV aggregators can participate real time market and day-ahead market to gain more revenues. For day-ahead electricity market, reference [16] formulated a non-convex EV charging problem to maximize the aggregator's benefit which is solved by a distributed optimization-based heuristic method. Reference [17] proposed a decision-making framework to offer operation strategies in day-ahead market, which suits for both risk-averse and risk-seeking decision makers. A hierarchical strategy was presented in [19], where the aggregator and each EV could share limited information. An aggregative game model was proposed in [20] for the day-ahead EV charging scheduling. The interaction between the EV charging demands was considered in this game. For real time markets, the management framework for EV aggregator was developed in [18]. This framework assigned charging setting points to EVs based on evolving charging priorities. Reference [21] developed a real-time charging controller for EVs to participate ancillary services market. Some studies considered both day-ahead and real-time electricity markets to formulate their scheduling models [15].

In the case of real-time price (RTP) or time-of-use (TOU) charging price, EV drivers often adjust their charging strategies to maximize their utilities. In previous studies, EV drivers are considered to be rational [14], [22]–[24] or limited rational [25], [26]. If the EV drivers are rational, they will maximize their expected utilities [14], and the utility company or charging service provider also want to achieve their targets (profit maximization, load profile adjustment, etc.). This problem is a Stackelberg game, which can be solved under game-theoretic framework if the utility function of each EV driver is given. In [22]–[24], the authors all propose Stackelberg game models to address interests conflict between the utility company/aggregator and the customers. If the EV drivers are assumed to be limited rational, as describe in [25], their decisions are influenced by perceived subjective. The subjective decision can be captured using the prospect theory.

Few studies have been conducted to quantify the price sensitivity of EV users. Most published researches assumed the EVs were either fully controllable or responded to price in a certain pattern, e.g. linearly related to price. Till now, there are only a few published papers on the price elasticity of EV charging demands. There are three kinds of charging demand elasticity considered in the published papers: self-elasticity, temporal cross elasticity, and spatial cross elasticity. Almost all the researches considered self-elasticity [27], [28] and some of them gave extra consideration to one of temporal cross elasticity [7], [29] and spatial cross elasticity [30], [31]. However, no study has combined all the three elasticities together.

In [31], conditional random field (CRF) was applied to capture dependencies of EVs' charging decisions. In this paper, we consider the expectation of charging demands in a certain zone rather than EVs' charging decisions. Most importantly, the temporal correlation of charging demands is taken into account, which makes the CRF model a loopy graph. So, it's hard to infer the graph precisely. The relationship between unknown parameters and elasticity is derived and examined with real historical charging data. The contributions of our work are as follows:

1) A conditional random field model considering the spatiotemporal cross elasticity of charging demand is built. The CRF model is used to represent not only the correlation between charging demand and price, but also the correlation between demands, which considers the shift of charging demands in both time and space.

2) The loopy belief propagation (LBP) algorithm is used to infer the loopy graph approximately. A training algorithm with forgetting factors is proposed to estimate the parameters of the CRF model. The relationship between the parameters of the CRF model and price elasticity is derived theoretically.

3) By applying the concept of aggregated charging station (ACS), which is made up of a number of charging stations in

a same zone, the relationship between the volume-weighted average price (VWAP) for charging and the total charging demands is studied. In this way, the charging demand elasticities can be quantified even the charging prices of the stations were not changed.

II. The Idea of Price Elasticity Modeling Based on Spatiotemporal Partition

To calculate the price elasticity of electricity demand, the traditional approach is to obtain changes in the demand as a result of given changes in electricity price [32]. But for EVs, their charging demands are not static, which can be served by different charging stations at different times. This situation is similar to a market filled with many substitute goods. In this paper, we will quantify not only the temporal correlation but also spatial correlation of charging demand. The relation between price and demand is called self-elasticity of demand. The relationship between price and demand in different periods is called temporal cross elasticity of demand. The relationship between price and demand in different regions is called spatial cross elasticity of demand.

Besides, we do not need to experimentally change the prices to calculate elasticity for a certain charging station. The basic idea of this work is to collect charging data of multiple charging stations with different prices for a certain period of time. By observing the variations of charging demands in different zones, we can model the price elasticity of EV charging demands.

Because only a limited amount of historical data can be collected, random perturbations of charging demand in a single charging station cannot be neglected. And the correlation of charging demand within a single charging station is usually too weak to be precisely captured. Therefore, the research objects in this study are zones instead of single charging stations. Each zone includes multiple charging stations, thus the effect of random perturbations is reduced. In order to analyze the relationship between charging demands in different time periods, the division of time is also necessary.

An example of division for zones is shown in Fig. 1. For this urban area, there are a number of public charging stations. Their prices and pricing strategies for EV charging may be different, including different fixed prices and time of use (TOU) prices. The volume-weighted average price (VWAP) is the ratio of the total charging expenditure to total energy over a particular period, e.g. an hour, in a certain zone.

To properly divide this area into a number of zones, we should make sure each zone has multiple charging stations. A whole day is also divided into periods (could be discontinuous) such as two periods for day and night, three periods for peak, valley, and flat, etc. Then, the aggregated charging station (ACS) for a zone is formed to represent all charging stations within it. The price of an ACS is the corresponding VWAP and demands are the sum of charging demands over the same period. Besides, we consider the charging demand can transfer between ACSs in different periods.

Since we study the correlation between zones rather than single charging stations, all public charging stations in a zone

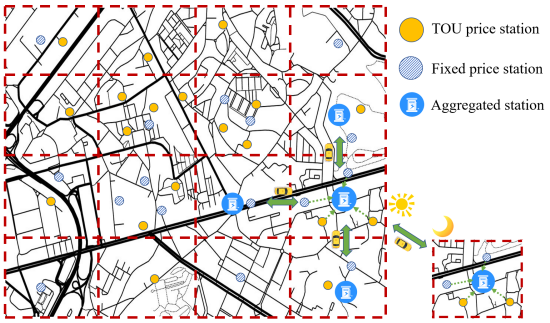


Fig. 1. Illustration of dividing an urban area into zones and aggregated charging station for a zone.

are aggregated into an ACS. Charging price of an ACS is set as the corresponding VWAP, and the corresponding demands are the total charging demands within the zone. Thus we can obtain the price-demand pair of each ACS, varying with time from the historical data. The price elasticity of charging demands can be captured from price-demand pairs based on following assumption.

Assumption: For a specific zone, the total charging demand remains unchanged if prices of all the charging stations are set equal to the VWAP in each period.

We believe that this assumption is reasonable because VWAP is the result of EV driver's decision between different types of pricing strategies of charging stations. If we set the charging price to VWAP, the total amount of charge demands will not change, given the same consumer charging demand elasticity.

III. CRF Model to Quantify Price Elasticity of Charging Demand

The elasticity of charging demand is related to a variety of factors, and the charging price is often the major one. The conditional random field model will be established in this section to represent correlations between demands to estimate self-elasticity, temporal cross elasticity, and spatial cross elasticity.

A. CRF model considering spatiotemporal cross elasticity

A conditional random field [33] has a set of observed variables and hidden random variables, which can be used to describe the correlation of random variables. We consider the prices of ACSs as observed variables and the charging demands of ACSs as hidden random variables. The observed variables are what we can directly observe or set. The hidden variables can not directly observed but can inferred (through a mathematical model) from other variables that are observed

Using the CRF model, the observed variable (charging price) can be used to infer the probability distribution of the hidden random variable (charging demand). A typical CRF model is shown in Fig. 2.

In Fig. 2, the yellow dots denote the hidden random variables, and the blue dots are the observed variables. They respectively represent the charging demands and charging prices in this paper. Each hidden random variable has a corresponding

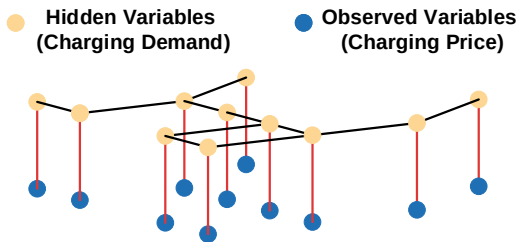


Fig. 2. Illustration of a conditional random field model.

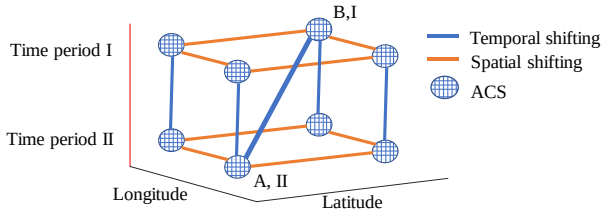


Fig. 3. Illustration of graph $\boldsymbol{G}=(\boldsymbol{V},\boldsymbol{E})$ with four zones and two periods.

observed variable. The correlations between them are denoted as red lines. Besides, there are also correlations between hidden random variables, which are denoted as black lines.

The charging demands can shift to other zones and time periods. The charging demands at different periods and zones are substitutes for each other. Thus it is not enough to study the correlations between the demand and price within a zone. We need to measure the correlations with the demands for substitutes.

A graph $\boldsymbol{G}=(\boldsymbol{V},\boldsymbol{E})$ is used to represent the correlations of charging demands. $\boldsymbol{V}$ is a set of hidden random variables, and $\boldsymbol{E}$ is a set of edges representing their correlations. Random variables that are not connected are independent. By constructing the graph $\boldsymbol{G}=(\boldsymbol{V},\boldsymbol{E})$, spatiotemporal shifting of charging demands can be described quantificationally. Edge $\boldsymbol{E}$ also represents the path of shifting charging demands.

We take charging prices as the observed variables and charging demands as the hidden variables. In the CRF model, the observed variables (charging prices here) need to be observed first, and then the probability distribution of hidden variables (charging demands here) can be inferred from them. We think it is reasonable because this inference process is similar to setting charging prices to determine charging demands.

The shifting of EV charging loads between time periods and zones is described by the correlation of corresponding charging demand. If the targeted charging station of an EV is changed from zone i to zone j, the absolute change of charging loads at zone i and zone j is equal, which means the charging demand d_i in zone i is related to the charging demand d_j in zone j. This correlation is described by the edges $\boldsymbol{E}$ in probability graph.

For an area divided into K zones and T periods, we have KT hidden random variables. Because the charging demand can shift between all periods and zones, there are C_{KT}^2 edges in this complete graph. The model will be quite complex to solve when KT is large enough. Therefore, in this paper, we choose some influential connections from all the C_{KT}^2 edges.

The graph can be made up of selected edges. The omitted edges are corresponding to the factors that can be neglected. For example, Fig. 3 illustrates four zones and the corresponding ACSs. The charging demand of each zone at a specific period can partially exchange with its adjacent zones and other periods. Besides, this graph particularly considers a edge from period II of zone A to period I of zone B because of EV users working and living in different zones. Demand shifting in time periods and zones are drawn in blue and orange lines, notated as $\boldsymbol{E_t}$ and $\boldsymbol{E_k}$, respectively.

B. Representation, learning and inference of CRF

There are three main parts to set up the CRF model. The first one is to select a proper feature function to represent price elasticity. The second one is to estimate unknown parameters $\boldsymbol{\omega}:=(\omega_i,\ldots,\omega_{i,j}^t,\ldots,\omega_i^{t_1,t_2},\cdots)$ of CRF based on historical charging data of ACS. The final part is the inference in a loopy graph with known parameters, to get the probability distribution of charging demand under observed prices.

1) Representation: We want to represent the joint probability using a proper feature function in this part. Considering the period set $\mathcal{T}$ and zone set $\mathcal{K}$, the conditional probability of charging demands under a group of charging prices in CRF can be formulated as equation (1).

$$P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})=\frac{1}{Z(\boldsymbol{\rho})}[\prod_{i\in\boldsymbol{V}}e^{\omega_i\psi_i(d_i,\rho_i)}\prod_{(i_t,j_t)\in\boldsymbol{E_k}}e^{\omega_{i,j}^t\psi_{i,j}^t\left(d_i^t,d_j^t\right)}\prod_{(i_{t_1},i_{t_2})\in\boldsymbol{E_t}}e^{\omega_i^{t_1,t_2}\psi_i^{t_1,t_2}(d_i^{t_1},d_i^{t_2})}], \tag{1}$$

Where

$$Z(\boldsymbol{\rho})=\sum_{\boldsymbol{d}\in\mathcal{X}^{|\boldsymbol{V}|}}[\prod_{i\in\boldsymbol{V}}e^{\omega_i\psi_i(d_i,\rho_i)}\prod_{(i_t,j_t)\in\boldsymbol{E_k}}e^{\omega_{i,j}^t\psi_{i,j}^t\left(d_i^t,d_j^t\right)}\prod_{(i_{t_1},i_{t_2})\in\boldsymbol{E_t}}e^{\omega_i^{t_1,t_2}\psi_i^{t_1,t_2}(d_i^{t_1},d_i^{t_2})}]. \tag{2}$$

Where $\boldsymbol{d}:=(d_i^t,\cdots),i\in\mathcal{K},t\in\mathcal{T}$ is the normalized charging demand vector, $\boldsymbol{\rho}:=(\rho_i^t,\cdots),i\in\mathcal{K},t\in\mathcal{T}$ is the normalized charging price vector, $\boldsymbol{\omega}:=(\omega_i,\ldots,\omega_{i,j}^t,\ldots,\omega_i^{t_1,t_2},\cdots)$ contains the unknown parameters. The edge set $\boldsymbol{E}=\boldsymbol{E_k}+\boldsymbol{E_t}$ includes the edges representing temporal cross elasticity and spatial cross elasticity. $\mathcal{X}$ is the discrete value set of charging demand d_i^t and $|\boldsymbol{V}|$ is the cardinality of set $\boldsymbol{V}$. Therefore, $\boldsymbol{d}\in\mathcal{X}^{|\boldsymbol{V}|}$ is to pick element in $\mathcal{X}$ for $|\boldsymbol{V}|$ times to form all possible $\boldsymbol{d}$. Partition function $Z(\boldsymbol{\rho})$ is used to normalize the probability. $\psi_i(d_i,\rho_i)$, $\psi_{i,j}^t\left(d_i^t,d_j^t\right)$ and $\psi_i^{t_1,t_2}(d_i^{t_1},d_i^{t_2})$ are feature functions for correlations, describing self-elasticity, spatial cross elasticity and temporal cross elasticity, respectively.

There is another interpretation of this joint probability equation. For each possible discrete charging demand $\boldsymbol{d}\in\mathcal{X}^{|\boldsymbol{V}|}$, we can first calculate a "fitness" score under the given feature functions.

$$S_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})=\omega_i\psi_i(d_i,\rho_i)+\omega_{i,j}^t\psi_{i,j}^t\left(d_i^t,d_j^t\right)+\omega_i^{t_1,t_2}\psi_i^{t_1,t_2}(d_i^{t_1},d_i^{t_2}) \tag{3}$$

Then, we can use softmax function to normalize these scores.

$$P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})=\frac{e^{S_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})}}{\sum_{\boldsymbol{d}}e^{S_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})}} \tag{4}$$

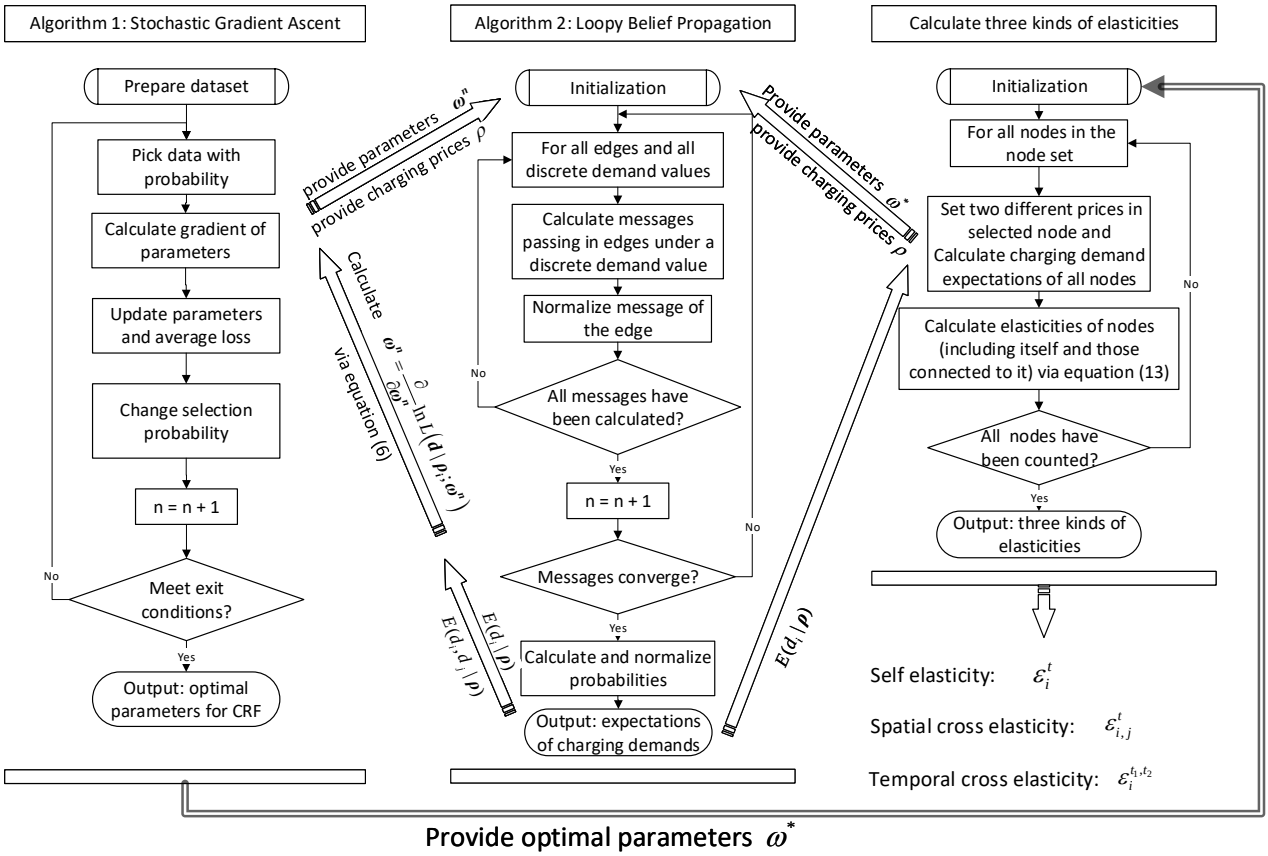


Fig. 4. Flowchart of proposed algorithm to calculate three kinds of elasticities.

The feature functions are defined as product of the corresponding variables [31].

$$\begin{aligned} \psi_i(d_i,\rho_i) &= d_i\rho_i,\ i \in \boldsymbol{V} \\ \psi_{i,j}^t(d_i^t,d_j^t) &= d_i^t d_j^t,\ \ (i_t,j_t) \in \boldsymbol{E_k}, t \in \mathcal{T} \\ \psi_i^{t_1,t_2}(d_i^{t_1},d_i^{t_2}) &= d_i^{t_1} d_i^{t_2},\ \ (i_{t_1},i_{t_2}) \in \boldsymbol{E_t} \end{aligned} \tag{5}$$

The self-elasticity $\varepsilon_k \approx \omega_k Var(\hat{d}_k|\hat{\boldsymbol{\rho}})$ can be derived from the CRF model using the above feature functions under a proper normalization of charging prices and demands. $Var(d_k|\boldsymbol{\rho})$ and $E(d_k|\boldsymbol{\rho})$ are the variance and expectation of d_k, respectively. Derivation of the price elasticities of charging demands can be found in *Appendix A*.

2) Learning: given a probability graph and corresponding feature functions, the learning problem of CRF is to find the unknown parameters $\boldsymbol{\omega} := (\omega_i,\ldots,\omega_{i,j}^t,\ldots,\omega_i^{t_1,t_2},\cdots)$ that fits the data set $\mathcal{D}$ best. Therefore, this is a parameter estimation problem aiming at finding the parameters to maximize the probability of the occurrence $P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})$ for data set $\mathcal{D}$. In this paper, the maximum likelihood estimation and the gradient descent algorithm are used to obtain the values of the parameters.

We can take logarithmic maximum likelihood function, $\ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega}) = \ln P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})$. If the feature functions are set as (5), the gradients for unknown parameters of this likelihood function can be formulated as (6).

$$\begin{aligned} \frac{\partial}{\partial \omega_i} \ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega}) &= d_i\rho_i - \rho_i E\left(d_i|\boldsymbol{\rho};\boldsymbol{\omega}\right) \\ \frac{\partial}{\partial \omega_{i,j}^t} \ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega}) &= d_i^t d_j^t - E\left(d_i^t d_j^t|\boldsymbol{\rho};\boldsymbol{\omega}\right) \\ \frac{\partial}{\partial \omega_i^{t_1,t_2}} \ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega}) &= d_i^{t_1} d_i^{t_2} - E\left(d_i^{t_1} d_i^{t_2}|\boldsymbol{\rho};\boldsymbol{\omega}\right) \end{aligned} \tag{6}$$

The derivation details in *Appendix B* [33]. In the right side of above equations, the first term is the value of the feature function under actual data, and the second term is the expected value of the feature function under parameter $\boldsymbol{\omega}$. The optimal unknown parameters is to maximize logarithmic maximum likelihood function, can be achieved by the gradient descent algorithm.

There may exist abnormal data in historical data set. The abnormal data is useful in finding the general direction of parameters in the early stage of learning, but it is harmful to search for the optimal parameters at a more precise stage later. We propose a learning algorithm considering forgetting factor, shown as Algorithm 1. This algorithm can gradually eliminate the abnormal data in the learning process.

Remark 1: When we calculate the gradients of likelihood function, the expected value of the feature function must be calculated under the given CRF parameters. But it is not easy to calculate. For example, for the first expected value $E\left(d_i\right) = \sum_{d_i} d_i P(d_i|\boldsymbol{\rho})$, the marginal probability $P(d_i|\boldsymbol{\rho})$, can be calculated directly by summing up all the joint probabilities,

Algorithm 1 Learning algorithm with forgetting factor

Input:

Forgetting factor, α;
Learning rate, γ;
Minimum and maximum number of iterations,n_{min},n_{max};
Number of periods to change selection probability, T;
Number of days of data, N;

Output:

Optimal parameters of CRF, $\boldsymbol{\omega}^*$;

Process:

Initializing $\boldsymbol{\omega^0}$, $n=0$, selection probability $p_i=\frac{1}{N}$ and loss $l_i=0$ for the ith data;
while $n<n_{max}$ **do**
  Picking the ith day's data from $\mathcal{D}$ with probability p_i;
  Calculating gradient $\Delta\boldsymbol{\omega^n}=\frac{\partial}{\partial\boldsymbol{\omega^n}}\ln L\left(\boldsymbol{d}|\boldsymbol{\rho_i};\boldsymbol{\omega^n}\right)$ using *Algorithm 2* and equation (6);
  $\boldsymbol{\omega}^{n+1}=\boldsymbol{\omega}^n+\gamma\Delta\boldsymbol{\omega}^n$;
  Updating loss $l_i+=\|\Delta\boldsymbol{\omega^n}\|_2$;
  if $n>n_{min}$ **then**
    if $mod(n,T)==0$ **then**
      Changing selection probability using average loss $p_i=\frac{e^{-\alpha\bar{l}_i}}{\sum_{k\in\mathcal{D}}e^{-\alpha\bar{l}_k}}$;
    end if
  end if
  $n=n+1$;
end while
return $\boldsymbol{\omega^n}$;

i.e. $P(d_i|\boldsymbol{\rho})=\sum_{\boldsymbol{d}\setminus d_i}\cdots\sum P\left(\boldsymbol{d}|\boldsymbol{\rho}\right)$. To get $P(\boldsymbol{d}|\boldsymbol{\rho})$, we need get scores using (3) for all $\boldsymbol{d}\in\mathcal{X}^{|\boldsymbol{V}|}$, with an exponential complexity of $O(|\mathcal{X}|^{|\boldsymbol{V}|})$. Obviously, we can not bear such an exponential complexity. Therefore, a faster inference algorithm will be introduced in the next section.

3) Inference: Given certain parameters $\boldsymbol{\omega}$ and a known charging price $\boldsymbol{\rho}$, the inference is to find the marginal distribution $P(d_i|\boldsymbol{\rho})$ and $P(d_i,d_j|\boldsymbol{\rho})$. Then it is easy to calculate expectations $E\left(d_i|\boldsymbol{\omega}\right)$, $E\left(d_i^td_j^t|\boldsymbol{\omega}\right)$, and $E\left(d_i^{t_1}d_i^{t_2}|\boldsymbol{\omega}\right)$.

The feature functions $\psi_{i,j}^t(d_i^t,d_j^t)=d_i^td_j^t$ and $\psi_i^{t_1,t_2}(d_i^{t_1},d_i^{t_2})=d_i^{t_1}d_i^{t_2}$ are equivalent for all edges. Therefore, the feature function and parameters are uniformly notated as $\psi_{i,j}(d_i,d_j)=d_id_j,(i,j)\in\boldsymbol{E}$ and $\omega_{i,j}$ in this section.

For the inference in a radial network, if we take the sum of joint probabilities directly in *Remark 1*, it will lead to an exponential complexity. In previous studies, belief propagation (BP) algorithm is usually used to infer the CRF problem. Figure 5 provides a graphical representation of a BP algorithm over a radial network. 1) First, selecting a root node for the graph (take node 1 for example); 2) Then, propagating messages from leaf nodes to root node ($4=>2;3=>2;2=>1$); 3) At last, propagating messages from leaf nodes to root node ($1=>2;2=>3;2=>4$). When all the leaf nodes receive the message, the algorithm is terminated.

The message is defined as (7):

$$m_{i,j}(d_j)=\sum_{d_i\in\mathcal{X}}e^{\omega_i\psi_i(d_i,\rho_i)}e^{\omega_{i,j}\psi_{i,j}(d_i,d_j)}\prod_{k\in N(i)\setminus\{j\}}m_{k,i}(d_i) \tag{7}$$

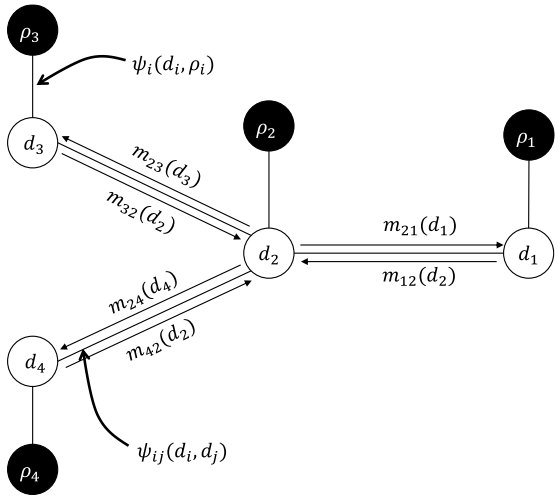


Fig. 5. Belief propagation algorithm in a radial graph.

BP algorithm can be used to accurate inference in radial graph, but in a loopy graph, the message needs to be propagated several times to get a convergent result. This iterative algorithm is called loopy belief propagation (LBP).

The LBP algorithm is widely used in probability graph inference. In [34], empirical study proves that LBP algorithm is a successful approximate inference method. Additionally, it is proved that LBP algorithm will converge to the stationary points of the Bethe approximation of the free energy in [35]. In case study, we also find that the model proposed in this paper converges very quickly with about five iterations.

Remark 2: The time complexity of LBP algorithm is linear. Specifically, if the number of iterations is K, its time complexity is $O(K|\boldsymbol{E}|)$, which is much smaller than $O(|\mathcal{X}|^{|\boldsymbol{V}|})$.

In this paper, the marginal distributions can be calculated with LBP [34], which details in *Algorithm 2*.

C. Elasticity matrices

The price elasticity of demand is usually defined as the change in demand over the change in price, as (8).

$$\varepsilon=\frac{\partial d/d}{\partial\rho/\rho} \tag{8}$$

Where d and ρ represent demand and price, respectively.

By linearizing (8), we can get a linear equation as (9). It is easy to solve the elasticity of charging demand using linear regression.

$$Ln(d)=\alpha+\varepsilon Ln(\rho) \tag{9}$$

In this paper, the spatial cross elasticity between zone i and zone j, and the temporal cross elasticity in zone i between period t_1 and t_2 are indicated as (10).

$$\varepsilon_{i,j}^t=\frac{\partial d_i^t/d_i^t}{\partial\rho_j^t/\rho_j^t},\varepsilon_i^{t_1,t_2}=\frac{\partial d_i^{t_1}/d_i^{t_1}}{\partial\rho_i^{t_2}/\rho_i^{t_2}} \tag{10}$$

Algorithm 2 Loopy belief propagation algorithm [34]

Input:

Maximum number of iterations n_{max};
Graph $\boldsymbol{G}=(\boldsymbol{E},\boldsymbol{V})$, neighbourhood of i, $N(i)$;
CRF parameters $\boldsymbol{\omega}$;
Charging prices of ACS $\boldsymbol{\rho}$;

Output:

$E\left(d_i|\boldsymbol{\omega}\right)$, $E\left(d_i^t d_j^t|\boldsymbol{\omega}\right)$, and $E\left(d_i^{t_1} d_i^{t_2}|\boldsymbol{\omega}\right)$;

Process:

Initializing messages $m_{i,j}^0(d_j)=1, \forall (i,j)\in \boldsymbol{E}, d_j\in\mathcal{X}$; $n=1$;
repeat
  for $(i,j)\in \boldsymbol{E}$ **do**
    for $d_j\in\mathcal{X}$ **do**
      Calculating message from i to j

$$m_{i,j}^n(d_j)=\sum_{d_i\in\mathcal{X}} e^{\omega_i\psi_i(d_i,\rho_i)}e^{\omega_{i,j}\psi_{i,j}(d_i,d_j)}\prod_{k\in N(i)\setminus\{j\}} m_{k,i}^{n-1}(d_i)$$

    end for
    Normalizing messages in a node:

$$m_{i,j}^n(\boldsymbol{d_j})=\frac{m_{i,j}^n(\boldsymbol{d_j})}{\sum\limits_{d_j\in\mathcal{X}} m_{i,j}^n(d_j)}$$

  end for
  $n=n+1$
until $\sum\limits_{(i,j)\in\boldsymbol{E},d_j\in\mathcal{X}} |m_{i,j}^n(d_j)-m_{i,j}^{n-1}(d_j)|<\varepsilon$ or $n>n_{max}$
Calculating belief:

$$\widetilde{P}(d_i|\boldsymbol{\rho})=e^{\omega_i\psi_i(d_i,\rho_i)}\prod_{k\in N(i)} m_{k,i}(d_i)$$

$$\widetilde{P}(d_i,d_j|\boldsymbol{\rho})=e^{\omega_i\psi_i(d_i,\rho_i)}e^{\omega_j\psi_j(d_j,\rho_j)}e^{\omega_{i,j}\psi_{i,j}(d_i,d_j)}\times \prod_{k\in N(i)\setminus\{j\}} m_{k,i}(d_i)\prod_{r\in N(j)\setminus\{i\}} m_{r,j}(d_j)$$

Normalizing belief:

$$P(d_i|\boldsymbol{\rho})=\frac{\widetilde{P}(d_i|\boldsymbol{\rho})}{\sum\limits_{d_i\in\mathcal{X}}\widetilde{P}(d_i|\boldsymbol{\rho})}$$

$$P(d_i,d_j|\boldsymbol{\rho})=\frac{\widetilde{P}(d_i,d_j|\boldsymbol{\rho})}{\sum\limits_{d_i\in\mathcal{X}}\widetilde{P}(d_i,d_j|\boldsymbol{\rho})}P(d_j|\boldsymbol{\rho})$$

Calculating expectations:

$$E(d_i|\boldsymbol{\rho})=\sum_{d_i\in\mathcal{X}} d_i P(d_i|\boldsymbol{\rho})$$

$$E(d_i,d_j|\boldsymbol{\rho})=\sum_{d_i,d_j\in\mathcal{X}} d_i d_j P(d_i,d_j|\boldsymbol{\rho})$$

If the zones and periods are both different, the elasticity $\varepsilon_{i,j}^{t_1,t_2}=0$ is zero.

Considering K zones and T periods, the elasticity matrices are formulated as (11) and (12).

$$\boldsymbol{\varepsilon}^t=\begin{bmatrix}\varepsilon_{1,1}^t & \varepsilon_{1,2}^t & \cdots & \varepsilon_{1,K}^t\\ \varepsilon_{2,1}^t & \varepsilon_{2,2}^t & \cdots & \varepsilon_{2,K}^t\\ \vdots & \vdots & \ddots & \vdots\\ \varepsilon_{K,1}^t & \varepsilon_{K,2}^t & \cdots & \varepsilon_{K,K}^t\end{bmatrix}, t\in\{1,2,\cdots,T\} \tag{11}$$

$$\boldsymbol{\varepsilon}_i^{t_1,t_2}=\begin{bmatrix}\varepsilon_i^{t_1,t_1} & \varepsilon_i^{t_1,t_2} & \cdots & \varepsilon_i^{t_1,t_T}\\ \varepsilon_i^{t_2,t_1} & \varepsilon_i^{t_2,t_2} & \cdots & \varepsilon_i^{t_2,t_T}\\ \vdots & \vdots & \ddots & \vdots\\ \varepsilon_i^{t_T,t_1} & \varepsilon_i^{t_T,t_2} & \cdots & \varepsilon_i^{t_T,t_T}\end{bmatrix}, i\in\{1,2,\cdots,K\} \tag{12}$$

Note that the diagonals of both matrices are actually self-elasticities.

To quantify elasticity matrices, arc elasticity is introduced. Arc is a section of the actual demand-price curve by giving different price ρ_1,ρ_2 and corresponding charging demand d_1,d_2. The mathematical definition is (13).

$$\varepsilon=\frac{\partial d/d}{\partial\rho/\rho}\approx\frac{\Delta d/(d_1+d_2)}{\Delta\rho/(\rho_1+\rho_2)}\approx\frac{(d_1-d_2)/(d_1+d_2)}{(\rho_1-\rho_2)/(\rho_1+\rho_2)} \tag{13}$$

D. An overview of proposed method

The flowchart of proposed method is shown in Fig. 4. *Algorithm 1* is to learn the optimal CRF parameters using the training dataset. *Algorithm 2* is to do inference under the given CRF parameters (using the charging prices to get the expected charging demands).

In *Algorithm 1*, the gradient of the logarithmic likelihood function is calculated using the expected charging demands, so *Algorithm 2* is needed. Under the optimal parameters of CRF, *Algorithm 2* is a mapping of charging prices to expected charging demands. In order to quantify the three kinds of elasticities, the optimal CRF parameters obtained by *Algorithm 1* can be used to estimate the elasticities with *Algorithm 2*.

We don't always have to compute all three kinds of elasticities. The mapping of charging prices to charging demands in *Algorithm 2* is actually the most valuable. However, the relationship between charging prices and charging demands is hidden under the CRF model. In order to get a clearer and more instructive index, we quantified three kinds of elasticities at a certain price.

Remark 3: The three kinds of elasticities can be achieved using the historical charging records. Then, the spatiotemporal distribution of charging demands can be adjusted through charging prices. There are three applications of the proposed method. 1) For a charging service provider, the proposed method can be used to get an optimal charging price strategy to maximize their revenue. 2) For utility companies who directly operate the public charging station, the proposed method can be used to set a time-of-use charging price to get a favorable charging load profile. Furthermore, using the spatial relationship of charging demands, the charging demands of heavy load zones can be shifted to other zones to reduce the reinforcement of power distribution equipment such as transformers. 3) If the utility companies can set the electricity price for charging service provider, the problem will become a Stackelberg game. Both utility companies and charging service

providers have their own optimization goals, the electricity price for charging service provider and the charging price for EV driver can be determined in an equilibrium state.

IV. Case Study Based on Simulation Data

In this section, we will apply the proposed CRF model to simulation data to verify the reliability of the model and its ability to solve a large scale problems.

A. Graphical structure and data preparation

The graphical structure used in this section is shown as Fig. 6. The graphs include three layers namely peak, flat, and valley. Each layer has 100 nodes (ACSs) connected into a grid and same ACS is connected between layers. This topology has a total of 300 nodes and 840 edges.

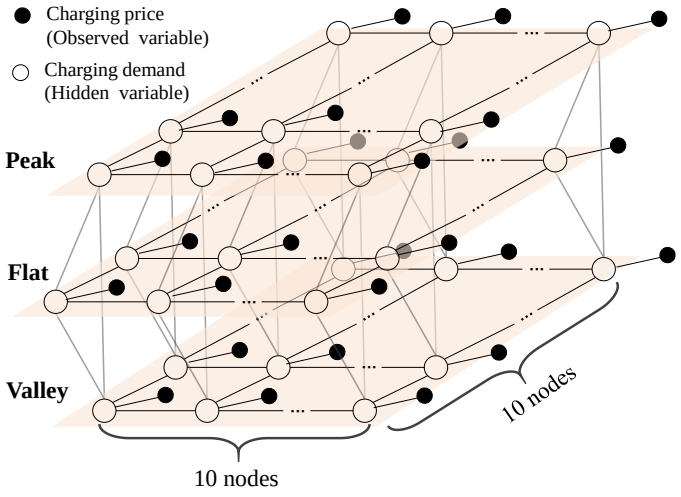


Fig. 6. Graphical structure of conditional random field in the case study for simulation data.

To generate simulation data, we set the self-elasticities to 1.6, 0.6, 0.5 for valley, flat, and peak laryer, respectively. The temporal cross elasticities are set to 0.05 (for valley-peak and valley-flat), 0.02 (for peak-flat). The spatial cross elasticities are set to 0.02 for all three layers.

A serise of demand-price pairs for all nodes are generated as training set and test set. All the data should be normalized using equation (21).

The parameters used in this case is as follow: the size of training set $N = 1000$; the learning rete $\gamma = 0.5$; the discrete set of normalized charging demand $\mathcal{X} = \{-1, -0.5, 0, 0.5, 1\}$; the error tolerance for messages converge $\varepsilon = 1e - 6$; the forgetting factor $\alpha = 5$; the inital parameters are random numbers between $[-1, 0]$; the inital messages are all set to 1; the precentage changes of prices are set to $N(0, 0.1)$.

In this case study, abnormal data will be considered. The simulation is carried out in the case of partial data error and no data error. Here we use three scenarios as follow.

1) Scenario 1: The dataset is accurate.
2) Scenario 2: The dataset is not accurate and don't use data forgetting method in training process. 30% of the training data has been added a Gaussian noise of $N(0, 0.3)$. The forgetting factor for this scenario is set to zero. (i.e. the choose probabilities of daily data won't change)
3) Scenario 3: The dataset is not accurate and use data forgetting method in training process. 30% of the training data has been added a Gaussian noise of $N(0, 0.3)$. The forgetting factor for this scenario is set to 5.

B. The result for simulation data

The test environment was MATLAB. The learning process is stopped at about 1500 iterations with Intel i9-10980XE CPU (about 30 minutes).

The parameters were tested in the test set (100 days), and the mean absolute percentage errors of the model are listed in Tab. I. Under the simulation data, CRF model has a high accuracy for all nodes.

TABLE I
Average Absolute Error on Simulation Data

	Peak	Flat	Valley	Average
Scenario 1 (high quality)	0.51%	0.51%	0.68%	**0.57%**
Scenario 2 (noise)	1.04%	1.04%	1.60%	**1.23%**
Scenario 3 (noise+forgetting)	0.51%	0.52%	0.69%	**0.57%**

If training is in a high quality dataset, error of the model is very small (see Scenario 1). When training in the dataset with partial abnormal data, the abnormal data will slow down the convergence and decreases the model accuracy (see Scenario 2). By using a forgetting factor, the abnormal data can be eliminated in the middle period of model training, and finally the accuracy can be similar to that of the data without errors (see Scenario 3).

The loss function during learning process for the high quality dataset is shown in Fig. 7:

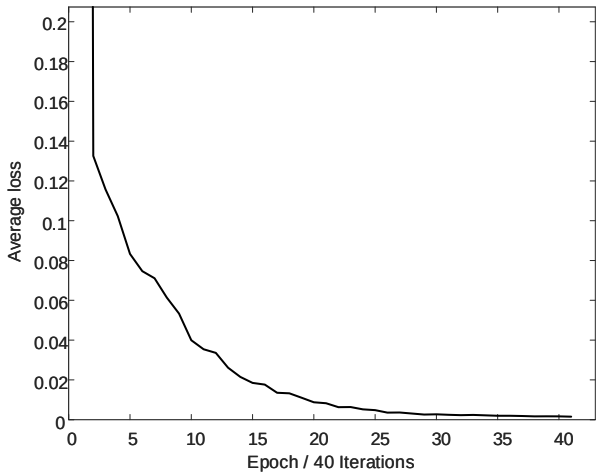


Fig. 7. Loss $l_i = \|\Delta\boldsymbol{\omega}^{\boldsymbol{n}}\|_2$ during learning process for accurate data in Scenario 1.

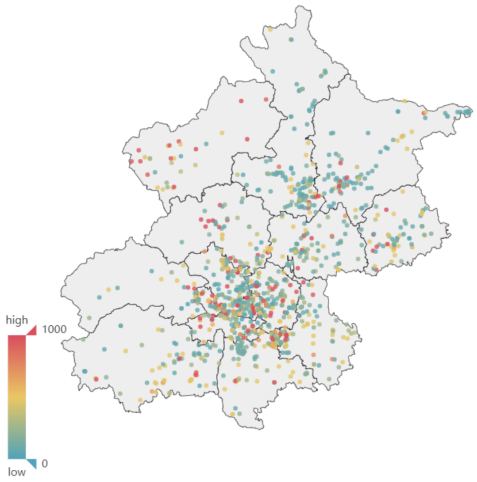
Remark 4: We use a trick in the learning process. In each iteration, the convergence of messages is time-consuming. The initial message for each iteration can be set to the converged messages in the previous iteration. With this trick,

the inference process of LBP can converge rapidly (usually one iteration).

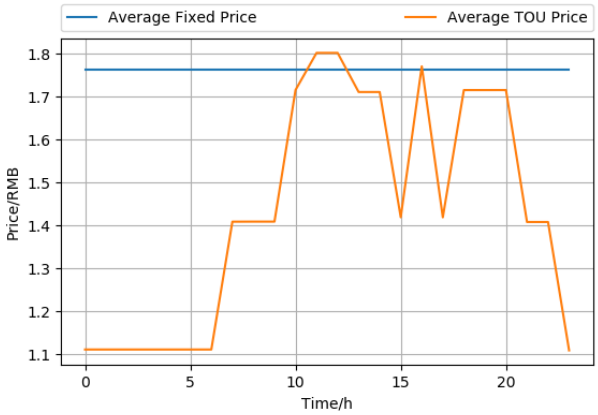
V. Case Study Based on Actual Data

A. Data preparation

We got three months of charging data from a charging service provider in Beijing. Its public charging stations cover most areas of the city. The pricing strategy can be categorized into six main groups: one is fixed price and the others are TOU prices. The locations and average prices of these charging stations are shown in Fig. 8.



(a) Locations (color bar indicates rated power)



(b) Average prices

Fig. 8. Information of public charging stations in Beijing.

The charging data is made up of charging records. We divide all charging records into different ACSs and normalize the total charging demand in every period. The charging VWAP is also calculated in every period. After that, we get the price-demand pairs for all ACSs and periods. There is only a price-demand pair for an ACS in each time period of a day.

The parameters used in this case is as follow: the size of training set $N = 90$; the learning rete $\gamma = 0.005$; the discrete set of normalized charging demand $\mathcal{X} = \{0, 0.25, 0.5, 0.75, 1\}$; the error tolerance for messages converge $\varepsilon = 1e{-}6$; the forgetting factor $\alpha = 5$; the inital parameters are random numbers between $[-1, 0]$; the inital messages are all set to 1. the minimum iterations $n_{min} = 1000$, the period of probability change $T = 200$. This simulations is implemented on a laptop computer with Intel i5-4300U CPU and 4 GB memory. All programs are coded in MATLAB.

B. Establishment of graph

Firstly, we divide the charging time into peak, flat, and valley according to the electricity price standard [36].

In this case, we have three time periods as Table II. Let $\mathcal{T} = \{H, M, L\}$ denote the set of periods, which are peak, flat, and valley. Every subgraph for time correlations has three nodes connected in pairs. The three nodes indicate charging demands of peak, flat and valley period at the same ACS.

TABLE II
Period Division and Electricity Purchase Prices

Period	Time Range	Prices (CNY/kWh)
Valley	$[23h, 0h) \cup [0h, 7h)$	0.3023
Flat	$[7h, 10h) \cup [15h, 16h) \cup [17h, 18h) \cup [21h, 23h)$	0.7697
Peak	$[10h, 15h) \cup [16h, 17h) \cup [18h, 21h)$	1.2884

Then, the subgraph for spatial correlations should be established. The city center is divided by squares, and each square represents a zone. After division, the number of EV charging stations and average EV charging power in each area are counted. The representative zones, i.e. zones with large amount of charging demands and charging stations were selected as part of the graph. In this paper, the graph shown in Fig. 9 consists of 11 zones and correlation links between adjacent zones. The charging demands can shift through the links.

(a) Number of charging stations (b) Average power of ACS/kW

Fig. 9. Selected spatial subgraph of demand within the urban area.

The final graph and its node numbers for our case study are shown in Fig. 10. This graph does not directly consider the shifting of charging demands to different periods of other zones. The graph is composed of subgraphs $(\boldsymbol{V_k}, \boldsymbol{E_k})$ about zone (in red lines) and subgraphs $(\boldsymbol{V_t}, \boldsymbol{E_t})$ about time (in blue lines). The final graph shows the topological structure of 11 zones and 3 time periods. When the price of a node (zone) changes, the charging demand of this node changes accordingly, as well as those of the nodes connected with it.

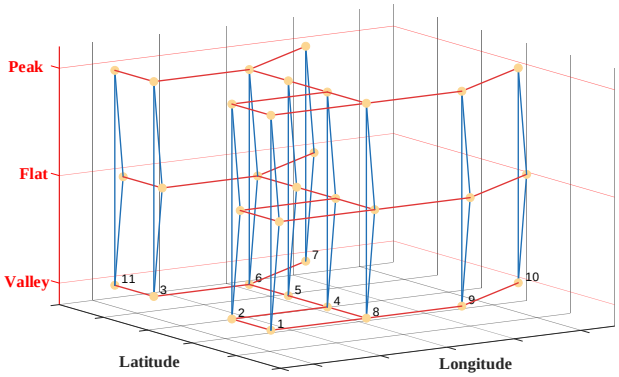


Fig. 10. Considered graph with spatiotemporal correlation of eleven zones and three periods.

C. Model evaluation

As the daily charging demand fluctuates, there are some abnormal points in the historical data. Therefore, in the process of data training, the days with large losses are gradually eliminated by reducing their selection probabilities. The algorithm is stopped at about the 8000th iteration (about 15 minutes). The average loss is shown in Fig. 11.

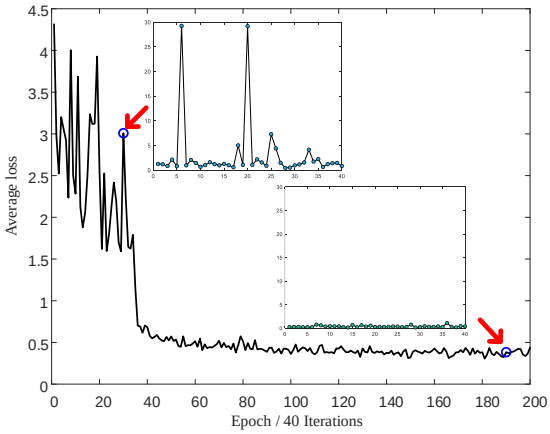


Fig. 11. Loss $l_i = \|\Delta\omega^n\|_2$ during learning process (forgetting begin at 25th epoch).

In the training process, we take every 40 iterations as an epoch. At the early stage (1-25th epoch) of the training, selection probabilities of all daily data are the same. There are some abnormal daily data in the original data, resulting in serious fluctuation of the loss function, but loss function shows a downward trend. The 40 iterations of an epoch are shown in Fig. 11 (the small window at the top left position), with two extremely daily data (6th and 20th iterations). After the 25th epoch, the selection probabilities of abnormal daily data are reduced because the average losses of them are large. Every five epochs, selection probabilities of daily data are updated according to their current average losses, so the value of the loss function is rapidly declining (20-35 epochs). After the 35th epoch, the probabilities of selecting abnormal daily data are very low, so the loss function began to decline slowly and gradually.

Using the LBP algorithm to infer the probability distribution of charging demand, we consider the expectation of charging demand as predicted value. Because the elasticity is a long-term demand prediction rather than a short-term forecasting, we will pick those days which are not eliminated in Algorithm 1 and compare the sum of their expected demand and actual demand. The accuracy is listed in Table III.

TABLE III
AVERAGE RELATIVE ERROR OF CRF MODEL

	Peak	Flat	Valley	Average
Zone1	-1.70%	-3.20%	5.20%	3.40%
Zone2	-1.50%	-1.70%	4.90%	2.70%
Zone3	-0.30%	2.60%	5.30%	2.70%
Zone4	-7.30%	-5.80%	27.80%	13.60%
Zone5	-4.00%	-2.20%	18.20%	8.10%
Zone6	-3.20%	2.80%	-2.20%	2.70%
Zone7	1.80%	4.50%	-5.00%	3.70%
Zone8	2.10%	1.10%	1.80%	1.70%
Zone9	-0.50%	-0.70%	1.80%	1.00%
Zone10	-0.50%	2.30%	-3.70%	2.20%
Zone11	-5.90%	-3.40%	5.20%	4.80%
Average	2.60%	2.70%	7.40%	**4.20%**

The average relative error is 4.2%. The prediction errors of valley charging demands are relatively large due to theirs high fluctuation. The charging demand is high in the flat and the peak periods, so the prediction errors are smaller. Generally speaking, the long-term prediction accuracy is high except for zone four, which indicates that the model is effective in predicting the long-term charging demand.

Remark 5: The impact factors over the accuracy of the model is listed as follow: a) The accuracy depends on the amount and quality of the training data. The conditional random field (CRF) is a statistic model. If the amount of training data is small and the data quality is poor, the accuracy of the CRF model will be reduced. b) The division of zones is also important. Note that there should be enough charging stations with different prices within a zone. In this way, an EV driver will have enough charging stations to choose, so as to reflect the EV driver's response to different charging prices. An extreme example is that there is only one charging station in a zone and all EV drivers can only charge at this charging station. So, the demand elasticity in this zone cannot be captured. c) The selection of hyperparameters in the model (such as the learning rate, etc.). Most of them should be adjusted during the learning process. If the learning rate is too large, the model will converge to a suboptimal solution or even divergence, leading a reduction of model accuracy. If the training step length is too small, it will take more time to converge.

D. Verification of charging elasticities of model

For the trained CRF model, we now verify the elasticities of the model. By changing the price of a specific ACS and remaining others unchanged, the demand curve can be drawn with the inference of CRF. The expected charging demand is considered as predicted value similarly.

The charging price in the zone 6 at the valley period, called *Zone6-valley* below, is set between [1.0, 2.0] CNY, and other zones are all set as 1.5 CNY. The demand curves of nodes connected with it are shown in Fig. 12.

In Fig. 12, the first and second figures show the temporal cross elasticity, the third figure shows the self-elasticity, and the fourth, fifth and sixth figures show the spatial cross elasticity. After the price of *Zone6-valley* changes, the charging demand decreases with the increase of the price in sixth zone during the valley period. However, the charging demand of nodes connected with *Zone6-valley* increases with the increase of the price of *Zone6-valley*, because these nodes are considered as a substitute for the *Zone6-valley*. And the change rate in *Zone6-valley* is much higher than that of other nodes. In addition, the relationship between logarithmic demand and logarithmic price is linear, which is consistent with the definition of price elasticity of demand in (9). The elasticity is the slope of the blue trend line. The spatial cross elasticity and the temporal cross elasticity do not change with the variation of price, but the self-elasticity does. When the charging price is higher, the self-elasticity is larger, and vice versa.

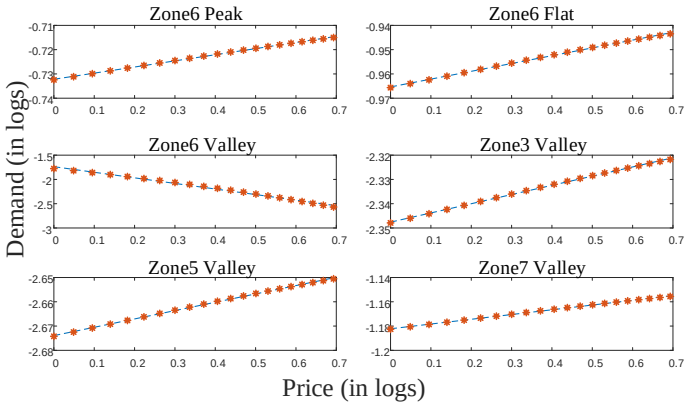


Fig. 12. Demand (in logs) changes of adjacent zones after price (in logs) changes in *Zone6-valley*.

E. Elasticity matrices

We use equation (13) to solve the temporal and spatial cross elasticity matrices under actual average charging price. Set the step of the price is 0.05 CNY and calculate the arc elasticity around average price.

An example of spatial cross elasticity matrices is shown as (14). This is a spatial cross elasticity matrix at valley. The diagonal elements, known as self-elasticity, are all negative, which means demand and price of the same zone change in the opposite direction. The nondiagonal elements, known as spatial cross elasticity, are all positive. Moreover, the absolute value of self-elasticity is much larger than that of spatial cross elasticity.

$$\varepsilon^{L} = \begin{bmatrix} & \text{Zone1} & \text{Zone2} & \text{Zone3} & \text{Zone4} & \cdots \\ \text{Zone1} & -1.391 & 0.027 & & & \\ \text{Zone2} & 0.032 & -1.742 & & 0.025 & \\ \text{Zone3} & & & -1.274 & & \\ \text{Zone4} & & 0.025 & & -1.497 & \\ \vdots & & & & & \end{bmatrix} \tag{14}$$

An example of temporal cross elasticity matrices is shown as (15). This is a temporal cross elasticity matrix at zone one. Similarly, the nondiagonal elements are positive, indicating that a charging station at different periods is the substitute for itself and charging demand can shift between different time periods. The valley-flat cross elasticity and valley-peak cross elasticity are significantly greater than the peak-flat cross elasticity. It means that it is easy to shift the charging demand to the valley period because of the bigger price difference. The prices of peak and flat periods are close to each other, so the incentive of demand shifting is not enough.

$$\varepsilon_1^{t_1,t_2} = \begin{bmatrix} & peak & flat & valley \\ peak & -0.274 & 0.0003 & 0.0180 \\ flat & 0.0004 & -0.193 & 0.0175 \\ valley & 0.0148 & 0.0111 & -1.390 \end{bmatrix} \tag{15}$$

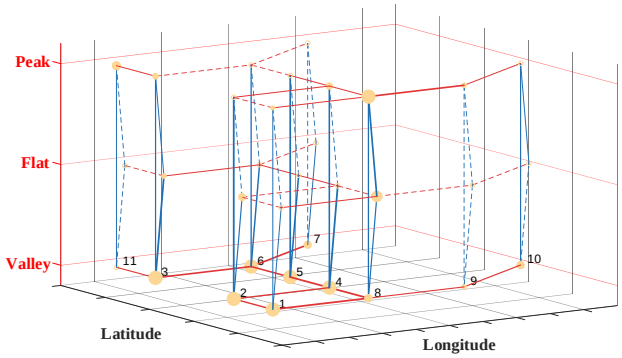


Fig. 13. Illustration of three kinds of elasticity. (Dots for self-elasticities, blue lines for temporal cross elasticities and red for spatial cross elasticities)

Fig. 13 demonstrates the self-elasticity, spatial cross elasticity and temporal cross elasticity. Note that only absolute values are considered.

The yellow node represents the self-elasticity, the blue line represents the temporal cross elasticity, and the red line represents the spatial cross elasticity. The larger the node, the thicker the line, the greater the elasticity. The dotted line indicates a small elasticity. Therefore, in a comprehensive view, the cross elasticity connected by the valley period is the greatest, and the valley time has the greatest self-elasticity.

VI. Conclusion

In this paper, we study the price elasticities of charging demands for public charging stations. At first, we divide the considered urban area into different zones. Each zone has an aggregated charging station through summing up by all public charging stations in that zone. To study the relationship between the volume-weighted average price and the total charging demand of an ACS, we bulid a conditional random

field model considering spatiotemporal shifting of charging demands. The elasticities in the CRF model are derived theoretically. The loopy belief propagation algorithm is designed to infer the loopy graph, and a learning algorithm with forgetting factors is proposed to estimate the unknown parameters of the CRF model. Three kinds of elasticity of EV charging demand, i.e. self-elasticity, temporal cross elasticity, and spatial cross elasticity, can be quantified. A case study, based on historical charging records of public charging stations, shows that the proposed method can quantify the three kinds of elasticities of charging demands effectively. These elasticities can be used to model consumer behaviors, devise pricing strategy for charging service providers, or set price signals for load serving entities.

Our next step is to solve the problem that the distribution of charging demands is not typical Gaussian distribution. Then we will focus on optimal charging strategies for public charging stations.

APPENDIX A

Since the elasticity can be formulated as (9), the self-elasticity of kth zone can be expressed as (16).

$$\varepsilon_k = \frac{\partial Ln\left[E(d_k|\boldsymbol{\rho})\right]}{\partial Ln(\rho_k)} = \frac{\partial E(d_k|\boldsymbol{\rho})/E(d_k|\boldsymbol{\rho})}{\partial \rho_k/\rho_k} \tag{16}$$

Where

$$\frac{\partial E(d_k|\boldsymbol{\rho})}{\partial \rho_k} = \int_{\boldsymbol{d}} \cdots \int d_k \frac{\partial P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})}{\partial \rho_k} \tag{17}$$

The self-elasticity can be expressed as:

$$\begin{aligned} \varepsilon_k =& \frac{\rho_k}{E(d_k|\boldsymbol{\rho})} \int_{\boldsymbol{d}} \cdots \int \left[d_k \cdot \frac{\partial P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})}{\partial \rho_k} \right] \\ =& \frac{\rho_k}{E(d_k|\boldsymbol{\rho})} \int_{\boldsymbol{d}} \cdots \int \left[d_k \cdot \frac{\partial Ln[P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})]}{\partial \rho_k} P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho}) \right] \end{aligned} \tag{18}$$

From Appendix B, we can know:

$$\frac{\partial}{\partial \rho_k} Ln\left[P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho})\right] = \omega_k \left[d_k - E(d_k|\boldsymbol{\rho})\right] \tag{19}$$

Then

$$\begin{aligned} \varepsilon_k =& \frac{\omega_k \rho_k}{E(d_k|\boldsymbol{\rho})} \int_{d_k} d_k (d_k - E(d_k|\boldsymbol{\rho})) P_{\boldsymbol{\omega}}(d_k|\boldsymbol{\rho}) \\ =& \frac{\omega_k \rho_k}{E(d_k|\boldsymbol{\rho})} [E(d_k^2|\boldsymbol{\rho}) - E^2(d_k|\boldsymbol{\rho})] \\ =& \frac{\omega_k \rho_k}{E(d_k|\boldsymbol{\rho})} Var(d_k|\boldsymbol{\rho}) \end{aligned} \tag{20}$$

If both charging price ρ and charging demand d in the CRF model are normalized to a percentage change, i.e. :

$$\hat{d}_k = \frac{d_k - \overline{d}}{\overline{d}}, \hat{\rho}_k = \frac{\rho_k - \overline{\rho}}{\overline{\rho}} \tag{21}$$

The self-elasticity will become simpler:

$$\varepsilon_k \approx \frac{E(\hat{d}_k|\boldsymbol{\rho})}{\hat{\rho}_k} = \omega_k Var(\hat{d}_k|\hat{\boldsymbol{\rho}}) \tag{22}$$

APPENDIX B

The logarithmic maximum likelihood function is

$$\begin{aligned} &\ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega}) = \ln P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho}) \\ &= [\sum_{i\in \boldsymbol{V}} \omega_i \psi_i\left(d_i, \rho_i\right) + \sum_{(i_t, j_t)\in \boldsymbol{E_k}} \omega_{i,j}^t \psi_{i,j}^t \left(d_i^t, d_j^t\right) \\ &+ \sum_{(i_{t_1}, i_{t_2})\in \boldsymbol{E_t}} \omega_i^{t_1,t_2} \psi_i^{t_1,t_2}(d_i^{t_1}, d_i^{t_2})] - Ln[Z(\boldsymbol{\rho};\boldsymbol{\omega})] \end{aligned} \tag{23}$$

Now, we consider $\frac{\partial}{\partial \omega_i} \ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega})$.

$$\frac{\partial}{\partial \omega_i} \ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega}) = \psi_i\left(d_i, \rho_i\right) - \frac{\partial}{\partial \omega_i} \ln Z(\boldsymbol{\rho};\boldsymbol{\omega}) \tag{24}$$

Because

$$\begin{aligned} &\frac{\partial}{\partial \omega_i} \ln Z(\boldsymbol{\rho};\boldsymbol{\omega}) = \frac{1}{Z(\boldsymbol{\rho};\boldsymbol{\omega})} \sum_{\boldsymbol{d}\in \mathcal{X}^n} [\psi_i\left(d_i, \rho_i\right) \prod_{i\in \boldsymbol{V}} e^{\omega_i \psi_i(d_i, \rho_i)} \\ &\prod_{(i_t, j_t)\in \boldsymbol{E_k}} e^{\omega_{i,j}^t \psi_{i,j}^t \left(d_i^t, d_j^t\right)} \prod_{(i_{t_1}, i_{t_2})\in \boldsymbol{E_t}} e^{\omega_i^{t_1,t_2} \psi_i^{t_1,t_2}(d_i^{t_1}, d_i^{t_2})}] \\ &= \sum_{\boldsymbol{d}\in \mathcal{X}^n} \psi_i\left(d_i, \rho_i\right) P_{\boldsymbol{\omega}}(\boldsymbol{d}|\boldsymbol{\rho}) = E\left(\psi_i\left(d_i, \rho_i\right)|\boldsymbol{\omega}\right) \end{aligned} \tag{25}$$

Then

$$\frac{\partial}{\partial \omega_i} \ln L(\boldsymbol{d}|\boldsymbol{\rho};\boldsymbol{\omega}) = d_i \rho_i - \rho_i E\left(d_i|\omega\right) \tag{26}$$

REFERENCES

[1] H. Cai, X. Jia, A. S. Chiu, X. Hu, and M. Xu, "Siting public electric vehicle charging stations in beijing using big-data informed travel patterns of the taxi fleet," *Transportation Research Part D: Transport and Environment*, vol. 33, pp. 39–46, 2014.

[2] M. Muratori and G. Rizzoni, "Residential demand response: Dynamic energy management and time-varying electricity pricing," *IEEE Transactions on Power systems*, vol. 31, no. 2, pp. 1108–1117, 2015.

[3] I. S. Bayram, M. Ismail, M. Abdallah, K. Qaraqe, and E. Serpedin, "A pricing-based load shifting framework for ev fast charging stations," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2014, pp. 680–685.

[4] A.-G. Paetz, T. Kaschub, P. Jochem, and W. Fichtner, "Load-shifting potentials in households including electric mobility-a comparison of user behaviour with modelling results," in *2013 10th International Conference on the European Energy Market (EEM)*. IEEE, 2013, pp. 1–7.

[5] S. Zhou, Y. Qiu, F. Zou, D. He, P. Yu, J. Du, X. Luo, C. Wang, Z. Wu, and W. Gu, "Dynamic ev charging pricing methodology for facilitating renewable energy with consideration of highway traffic flow," *IEEE Access*, vol. 8, pp. 13 161–13 178, 2019.

[6] R. Yu, W. Zhong, S. Xie, C. Yuen, S. Gjessing, and Y. Zhang, "Balancing power demand through ev mobility in vehicle-to-grid mobile energy networks," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 1, pp. 79–90, 2015.

[7] A. Kannan and V. K. Tumuluru, "Behavioral modeling of electric vehicles using price elasticities," in *2018 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*. IEEE, 2018, pp. 593–598.

[8] Q. Chen, F. Wang, B.-M. Hodge, J. Zhang, Z. Li, M. Shafie-Khah, and J. P. Catalão, "Dynamic price vector formation model-based automatic demand response strategy for pv-assisted ev charging stations," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2903–2915, 2017.

[9] R. Li, Q. Wu, and S. S. Oren, "Distribution locational marginal pricing for optimal electric vehicle charging management," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 203–211, 2013.

[10] Z. Hu, K. Zhan, H. Zhang, and Y. Song, "Pricing mechanisms design for guiding electric vehicle charging to fill load valley," *Applied Energy*, vol. 178, pp. 155–163, 2016.

[11] A. Dubey and S. Santoso, "Electric vehicle charging on residential distribution systems: Impacts and mitigations," *IEEE Access*, vol. 3, pp. 1871–1893, 2015.

[12] X. Dong, Y. Mu, X. Xu, H. Jia, J. Wu, X. Yu, and Y. Qi, "A charging pricing strategy of electric vehicle fast charging stations for the voltage control of electricity distribution networks," *Applied energy*, vol. 225, pp. 857–868, 2018.

[13] A. Ghosh and V. Aggarwal, "Control of charging of electric vehicles through menu-based pricing," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5918–5929, 2017.

[14] K. Jhala, B. Natarajan, A. Pahwa, and L. Erickson, "Real-time differential pricing scheme for active consumers with electric vehicles," *Electric Power Components and Systems*, vol. 45, no. 14, pp. 1487–1497, 2017.

[15] C. Jin, J. Tang, and P. Ghosh, "Optimizing electric vehicle charging with energy storage in the electricity market," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 311–320, 2013.

[16] J. James, J. Lin, A. Y. Lam, and V. O. Li, "Maximizing aggregator profit through energy trading by coordinated electric vehicle charging," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2016, pp. 497–502.

[17] J. Zhao, C. Wan, Z. Xu, and J. Wang, "Risk-based day-ahead scheduling of electric vehicle aggregator using information gap decision theory," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1609–1618, 2015.

[18] S. I. Vagropoulos, D. K. Kyriazidis, and A. G. Bakirtzis, "Real-time charging management framework for electric vehicle aggregators in a market environment," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 948–957, 2015.

[19] M. H. Amini, P. McNamara, P. Weng, O. Karabasoglu, and Y. Xu, "Hierarchical electric vehicle charging aggregator strategy using dantzig-wolfe decomposition," *IEEE Design & Test*, vol. 35, no. 6, pp. 25–36, 2017.

[20] Z. Liu, Q. Wu, S. Huang, L. Wang, M. Shahidehpour, and Y. Xue, "Optimal day-ahead charging scheduling of electric vehicles through an aggregative game model," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5173–5184, 2017.

[21] G. Wenzel, M. Negrete-Pincetic, D. E. Olivares, J. MacDonald, and D. S. Callaway, "Real-time charging strategies for an electric vehicle aggregator to provide ancillary services," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5141–5151, 2017.

[22] P. Shinde and K. S. Swarup, "Stackelberg game-based demand response in multiple utility environments for electric vehicle charging," *IET Electrical Systems in Transportation*, vol. 8, no. 3, pp. 167–174, 2018.

[23] M. Yu and S. H. Hong, "A real-time demand-response algorithm for smart grids: A stackelberg game approach," *IEEE Transactions on smart grid*, vol. 7, no. 2, pp. 879–888, 2015.

[24] P. salyani, M. Abapour, and K. Zare, "Stackelberg based optimal planning of dgs and electric vehicle parking lot by implementing demand response program," *Sustainable Cities and Society*, vol. 51, p. 101743, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2210670719305062

[25] K. Jhala, B. Natarajan, and A. Pahwa, "Prospect theory-based active consumer behavior under variable electricity pricing," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2809–2819, 2018.

[26] L. Hu, J. Dong, and Z. Lin, "Modeling charging behavior of battery electric vehicle drivers: A cumulative prospect theory based approach," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 474–489, 2019.

[27] N. G. Paterakis, A. Taşcıkaraoğlu, O. Erdinc, A. G. Bakirtzis, and J. P. Catalao, "Assessment of demand-response-driven load pattern elasticity using a combined approach for smart households," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 4, pp. 1529–1539, 2016.

[28] V. Gómez, M. Chertkov, S. Backhaus, and H. J. Kappen, "Learning price-elasticity of smart consumers in power distribution systems," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2012, pp. 647–652.

[29] Y. Gao, L. Mengkuo, Q. Wang, H. Liang, and J. Zhang, "Research on optimal tou price considering electric vehicles charging and discharging based on discrete attractive model," *Proceedings of the Csee*, vol. 34, no. 22, pp. 3647–3653, 2014.

[30] C. Luo, Y.-F. Huang, and V. Gupta, "Stochastic dynamic pricing for ev charging stations with renewable integration and energy storage," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1494–1505, 2017.

[31] N. Y. Soltani, S.-J. Kim, and G. B. Giannakis, "Real-time load elasticity tracking and pricing for electric vehicle charging," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1303–1313, 2014.

[32] N. G. Mankiw, *Principles of economics*. Cengage Learning, 2020.

[33] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*, ser. Adaptive computation and machine learning. Cambridge, MA: MIT Press, 2009.

[34] K. Murphy, Y. Weiss, and M. I. Jordan, "Loopy belief propagation for approximate inference: An empirical study," *arXiv preprint arXiv:1301.6725*, 2013.

[35] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Transactions on information theory*, vol. 51, no. 7, pp. 2282–2312, 2005.

[36] Development and reform commission of Beijing. (2019) Notice concerning the adjustment of general industrial and commercial sales electricity price of this city. [Online]. Available: http://fgw.beijing.gov.cn/fzggzl/zgd/zcwj/201912/t20191227_1524575.htm