Electrical Power and Energy Systems 71 (2015) 15–25

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Stochastic modeling for scheduling the charging demand of EV in distribution systems using copulas

V. Tavakoli Bina [a], D. Ahmadi [b,*]

[a] Faculty of Manegment, University of Tehran, Tehran, Iran
[b] Electrical Engineering Faculty, Pooyesh Higher Education Institute, Qom, Iran

ARTICLE INFO

Article history:
Received 22 August 2014
Received in revised form 22 January 2015
Accepted 1 February 2015

Keywords:
Copula
Demand response
Electric vehicles
GAMS
Stochastic modeling

ABSTRACT

Charging demand of electric vehicles (EV) has potentially a significant influence on the power grid. If this charging demand coincides mainly with the peak demand of the power grid, then additional active power has to be supplied to fulfill load management purposes. Thus, it is necessary to estimate and schedule the charging demand of the EV in order to lower the peak demand. Various estimation techniques are available such as Gaussian mixture model and copula. This paper uses copula for data estimation because copula imposes no restriction on the marginal distributions of the available data. Meanwhile, uncertain estimated data requires error elimination. Clayton copula is selected for flexile part of load profile, and Gaussian copula for non-controllable part of the load profile based on the two consecutive days (TCD) classification. Hence, the created scenarios were applied to an optimization problem that flattens the load profile as much as possible using general algebraic modeling system (GAMS). Then, this research concentrates on two new semi-automatically proposals concerned with the day-ahead charging demand response (DR) strategies. These strategies contribute to removing the estimation errors created due to the uncertainties. In order to examine the efficiency of the proposed strategies, the day-ahead pricing (DAP) with inclining block rates (IBR) model is assumed for retail electricity markets.

Introduction

Technology of plug in EV is potentially an effective way to solve the fossil fuel energy crisis, green house emission and air pollution [1]. However, increasing the EV penetration in the distribution systems can overload the grid during the peak times. Therefore, this additional demand will strongly influence on the generation, transmission and distribution systems. However, the peak demand can be reduced by redistributing it to the less congested periods such as early mornings or late at nights. In other words, when load-shifting strategies are applied to control flexible loads, future investment on creation of new power plants is lowered [2,3].

Moreover, unlike the usual existing loads, the EV locates within the class of controllable loads because they are being equipped with batteries that enable them to store electrical energy. Therefore, it would always be interesting to schedule the charging time of the EV in order to fill the valley hours by reducing the peak demand in smart grids [3,4]. Recently, many lectures have investigated the influence of high penetration of the EV in the network [5–7]. In [7], the Gaussian copula (GC) function was proposed to estimate the charging impact of the EV on load curve stochastically for 2020–2025.

It is assumed in [7] that all EV can be charged everywhere across the grid. In other words, the charging profile of a huge number of EV was modeled from the *perspective of power network* using the GC. In practice, the majority of studies focus on online scheduling and management of the charging strategies of the EV [1,3,8–11]. Clearly, the cost of the online DR, which is known as fully automated DR, is high and complicated in implementation; this is due to the need for installing numerous hardware such as smart meters, control instruments and communication equipment. In recent years, aggregators join two or more customers to manage their electricity in smart grids [12]. Aggregators suffer lack of historical demand data (HDD) for estimating short-term future data [13–15]. The HDD are collected from the meters installed for the customers.

This research concentrates on estimating the next day hourly charging demand of a limited number of EV from the *perspective of an aggregator or a distribution system operator* (DSO) by employing both Archimedean and Elliptical copulas. Copula was chosen for estimation because copula is capable of describing nonlinear dependence among multivariate data independent of their marginal probability distributions mathematically. Five different types

* Corresponding author.
E-mail address: ahmadi.danial@gmail.com (D. Ahmadi).

http://dx.doi.org/10.1016/j.ijepes.2015.02.001
0142-0615/

of copula are examined in order to model the charging demand of 960 EV as flexible part of the load profile; analytical assessment of the five copulas results in Clayton copula as the best copula among them. Then, it is necessary to predict non-controllable part of the next day load profile (960 residential units). Hence, one year collected data are classified into five classes of two consecutive days (TCD). Again, five different copula functions were examined (see Table 1), resulting in the GC as the best copula for predicting non-controllable loads. Then, a stochastic modeling is proposed for estimating the next day hourly demand, applying all predicted scenarios to an optimization problem that is solved by the GAMS under a stochastic process. In addition, this paper introduces two *semi-automatic* day-ahead DR strategies that concentrate on charge scheduling of the EV. The outcomes will help aggregators to reduce both their peak demand and electricity payment. Moreover, two case studies are considered to verify the proposed strategies on charging schedule; an aggregator supplies 960 residential units located in phase 1, 2, and 3 of Ekbatan complex in western Tehran. Then, the proposed DR strategies were applied to the prepared residential case. Simulations on the carried out case studies show the effectiveness of the two proposed algorithms. The outline of the proposed research work performed in this paper is shown in Fig. 1.

Modeling principles

Fig. 1 demonstrates the modeling principles of this research, including exact data gathering, data generation, producing day-ahead DR scenarios as well as arranging an optimization problem to find out the best-predicted load profile. The following subsections describe the chain of modeling elements that form the entire work.

Exact data gathering

As the *first step*, samples of a limited number of EV are taken from the owners of the EV, including their traveled distances and times to plug in. In practice, the EV owners are asked by the aggregator to *predict* their possible traveled distances and times to plug in for *the next day*. Thus, two points are crucial; first, the number of nominated EV has to be limited due to various economical, practical and social restrictions. Hence, only 50 EV owners contribute to the procedure of exact data gathering. Second, these predicted data by the EV owners are *uncertain data* for the next day, i.e. may differ slightly from the next day exact data.

Data generation

It is necessary to estimate data for the remaining residential EV owners (910 out of 960) based on those of 50 taken samples. Copula was chosen in this research for generating required data because of the discussions raised in the following subsections.

Why copula?

In brief, the most interesting advantage of copula functions is their capability in estimating marginal and rank correlation of samples of random variables [16]. The modeling principles of copulas allow easy modeling and estimation of multivariate distribution function. In other words, compared to other estimating methods, copula inter-relates multivariate data nonlinearly. This research uses copula to estimate the bivariate charging demand of the EV. Since these variables (traveled distances and times to plug in) are statistically dependent, a joint distribution function is built using copula. Subsequently, predicted charging demand of the EV is determined according to the obtained distribution functions.

Moreover, sometimes the number of monitored exact data has to be much smaller than that of the original dataset in practice. For example, installing GPS on all EV to collect the required data (traveled distances and times to plug in) is costly, imposing further possible social restrictions on the customers. Nevertheless, the aggregator needs to estimate the required demand of recharging *all* EV for the next day in order to pre-schedule flexible loads optimally. Hence, the total charging demand of all EV has to be predicted using a limited number monitored exact data.

Suitability of copula for creation of a large dataset

Here it is demonstrated a simulation in which a small dataset, taken from a limited number of EV, has to be expanded in order to prepare estimate of the load profile for all the available EV as a large dataset. For example, assume the small dataset includes 50 exact data; once copula estimates 20 datasets of size 50 EV from the exact dataset, while once again copula is used to create a large dataset of size 1000 EV directly from the exact dataset. Simulations are shown in Fig. 2(a)–(c), in which 1000 EV data are compared correspondingly for the two ways of data estimation. Both the estimated data (Fig. 2(a)) and their probability distribution functions (Fig. 2(b)–(c)) confirm the two ways of estimating with copula are quite the same. In fact, copula works on correlation of the exact dataset, achieving much more acceptable data in estimation of large datasets compared to other solutions. Thus, according the introduced simulations, a copula function can work on predicting a large dataset directly; e.g. predicting distances and times to plug in for 960 houses (owning the EV) in the considered case study straightforward. However, the estimated data by copula are uncertain as well.

Mathematical description of copula

Definition of copula

Copulas are multivariate distribution functions (C) of m variables whose one-dimensional marginal distributions are uniformed within the interval $[0,1]^m$ with the following properties [16]:

(a) $C(1, \ldots, 1, u_l, 1, \ldots, 1) = u_l$ for all $1 \leqslant l \leqslant m$;
(b) $C(u_1, u_2, \ldots, u_m)$ is increasing in each component u_l, $l \in \{1, 2, \ldots, m\}$;
(c) The range of C is the unit interval $[0,1]$;
(d) For $u_l \leqslant v_l$, $1 \leqslant l \leqslant m$, C satisfies the rectangle in equality [17];

$$\sum_{l_1=1}^{2} \cdots \sum_{l_m=1}^{2} (-1)^{l_1+\ldots+l_m} C(u_{1,l_1}, \ldots, u_{m,l_m}) \geqslant 0 \tag{1}$$

where m is the number of dependent outcomes that should be modeled with all marginal distributions of the random vectors (u_1, u_2, …, u_m). It can be showed from the definition that copulas are capable of describing nonlinear dependence among multivariate data in isolation from their marginal probability distributions [16]. Copulas can also serve as a powerful tool for both modeling and simulating nonlinearly-interrelated multivariate data, and uniform continuity and existence of all partial derivatives [18]. Consider the joint probability distribution of m-random variables X_q ($q = 1, \ldots, m$), $H(x_1, \ldots, x_m)$;

$$H(x_1, x_2, \ldots, x_m) = Pr(X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_m \leqslant x_m) \tag{2}$$

where continuous marginal probability distributions are denoted by $F_q(x_q) = Pr(X_q \leqslant x_q) = u_q$. According to [16], Sklar's theorem dictates that the relationship between $H(x_1, \ldots, x_m)$ and $F_q(x_q)$, $(q = 1, \ldots, m)$;

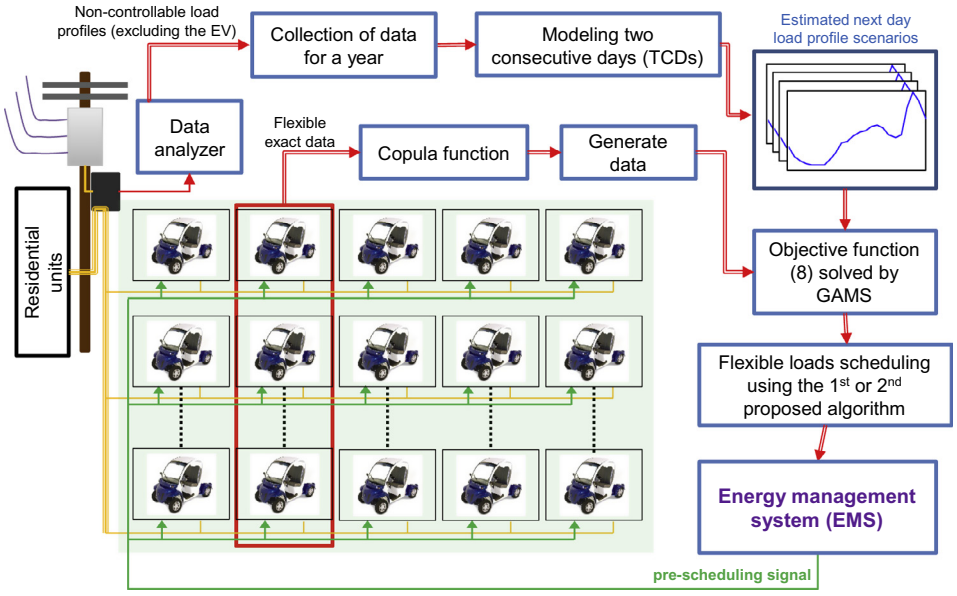

Fig. 1. The outline of the research work performed in this paper, including the EV, estimation and generation of next day hourly scenarios for both flexible and non-controllable load profiles, finding out the optimized scenario and scheduling the charging demand for the EV using two suggested DR strategies.

Table 1
Characteristics of archimedean and elliptical copulas and their measures of dependence.

Copula	Dependence structure characteristics	Archimedean generation function	Relation between the copula parameter and the Kendall's τ coefficient
Gaussian	Symmetric about center point, weak tail dependencies, left and right tail dependencies go to zero at extremes	–	$\mathbf{Rho} = \sin(\pi\tau/2)$ $\tau \in [-1,1]$
t	Symmetric about center point, weak tail dependences, left and right tail dependencies go to zero at extremes	–	$\mathbf{Rho} = \sin(\pi\tau/2)$ $\tau \in [-1,1]$
Clayton	Symmetric about center point, strong left tail dependence and weak right tail dependence, right tail dependence goes to zero at right extreme	$\frac{1}{\alpha}(t^{-\alpha}-1)$	$\alpha = 2\tau/(1-\tau)$ $\tau \in (0,1)$
Gumbel	Symmetric about center point, strong right dependence, weak left tail dependence, left tail dependence goes to zero at left extreme	$(-\ln t)^{\alpha}$	$\alpha = 1/(1-\tau)$ $\tau \in [0,1)$
Frank	Symmetric about center point, very weak tail dependencies, left and right tail dependencies go to zero at extremes	$-\ln\frac{\exp(-\alpha t)-1}{\exp(-\alpha)-1}$	$\tau = 1-\frac{4}{\alpha}+\frac{4}{\alpha^2}\int_0^{\alpha}\frac{t}{\exp(t)-1}dt$ $\tau \in [-1,1]$

which can be established by using the copula function $C^m(u_1, \ldots, u_m)$ as follows:

$$H(x_1,x_2,\ldots,x_m) = C^m[F_1(x_1),F_2(x_2),\ldots,F_m(x_m)] = C^m(u_1,u_2,\ldots,u_m) \quad (3)$$

In mathematical terms, a copula function $C^m(u_1, \ldots, u_m)$ is the m-dimensional probability distribution on a unit hypercube $[0,1]^m$ with uniform marginal probability distributions on [0,1], and is defined as follows [18]:

$$C^m(u_1,u_2,\cdots,u_m) = Pr(U_1 \leqslant u_1, U_2 \leqslant u_2, \cdots, U_m \leqslant u_m) \quad (4)$$

where u_m represents a sample of a standard uniform random variables U_q ($q = 1, \ldots, m$). Some samples of t, Gaussian, Gumbel, Frank, and Clayton copulas which are known as the Archimedean and Elliptical copulas are illustrated in Fig. 3 in bivariate form. Moreover, the Kendall correlation τ is one of the most accurate methods in estimating the copula parameters. Hence, the parameter **Rho** of the Gaussian (Normal) copula can be estimated as **Rho** which is equal to $\sin(\pi.\tau/2)$, and the parameter α of the Gumbel copula as $\alpha = 1/(1-\tau)$. Other methods and more details for estimation of the Archimedean copulas parameters were summarized in Table 1 [16–18].

The GC is known as Elliptical copula. It is the most familiar among all copulas and is distributed over the unit cube $[0,1]^m$. The m-dimensional GC is defined as bellows:

$$C^m(u_1,u_2,\ldots,u_m;\mathbf{Rho}) = \varphi_m(\varphi^{-1}(u_1),\varphi^{-1}(u_2),\ldots,\varphi^{-1}(u_m);\mathbf{Rho}) \quad (5)$$

where $\varphi^{-1}(.)$ is the inverse cumulative distribution function of a standard normal distribution function $\varphi(.)$; and $\varphi_m(.;\mathbf{Rho})$ is the m-dimensional standard multivariate normal distribution function with mean vector zero and covariance matrix equal to the correlation matrix, **Rho**. To simulate dependent multivariate or bivariate data using a copula must specify each of the following:

- The copula family and any shape parameters.
- The rank correlations among variables and.
- Marginal distribution for each variable.

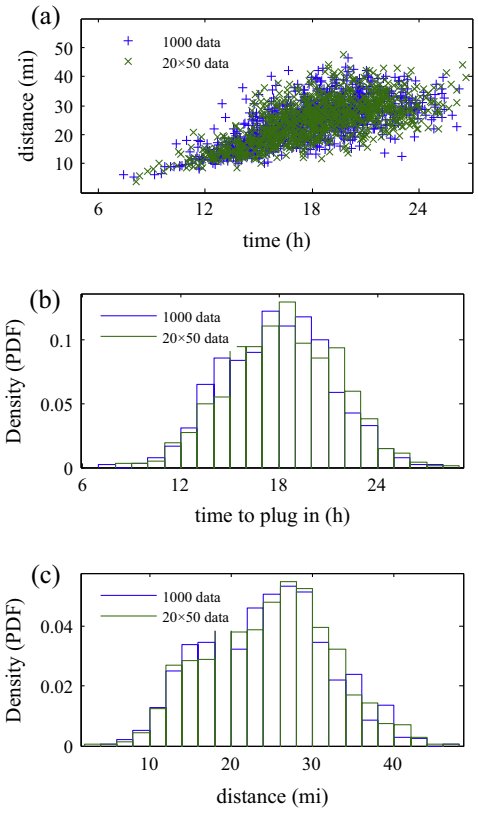

Fig. 2. Copula estimates a large dataset of size 1000 data from 50 exact data; once 20 sets of size 50 data (in green) and another time one set of size 1000 data (in blue), (a) estimated data, (b) the PDF of estimated traveled distances, and (c) the PDF of estimated times to plug in. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Seeking a proper copula for data creation

Earlier studies in [7,10] showed that the approximate plug in times for recharging the EV have been during the afternoon when people arrived home from work. These charging times usually coincide with the time of the peak demand on the load profile [7]. Fig. 4 shows the predicted times to plug in and distances traveled at the night before through the internet by about 5% of 960 EV (50 EV owners). Then, the parameters of copula functions can be worked out by means of an approximation to Kendall's rank correlation (see relationships in Table 1). For example, the parameter **Rho** of the GC is calculated for Fig. 4 as the dependencies between the times to plug in and distances traveled as follows:

$$\mathbf{Rho} = \begin{bmatrix} 1 & 0.73 \\ 0.73 & 1 \end{bmatrix} \tag{6}$$

There are two copula families, Archimedean and Elliptical, including various functions with their own characteristics. Table 1 lists five different types of copula functions that are examined by this research for seeking a proper one. This examination is taken place as follows:

- The exact dataset of size 50 EV were applied to five different copula functions listed in Table 1, resulting in five different large datasets of size 960 EV estimates as shown in Fig. 5(b)–(f) for the GC, *t* copula, Gumbel copula, Frank copula, and Clayton copula.
- Questioners were distributed among 960 EV owners to fill out their detailed exact times to plug in and distances traveled for all 960 EV after the occurrence in the predicted day. Fig. 5(a) shows the realized collected dataset in order to compare it with 960 EV estimates in Fig. 5(b)–(f) by various copula functions. Fig. 5(b)–(f) show that the produced estimates by different copulas for charging 960 EV are very similar to those of the exact dataset shown in Fig. 5(a).
- The predicted datasets are used to pre-schedule required demand of the EV in the next day using the battery charge information listed in Table 2 regarding four typical EV in the USA market [11]. In fact, the listed data in Table 2 are used to

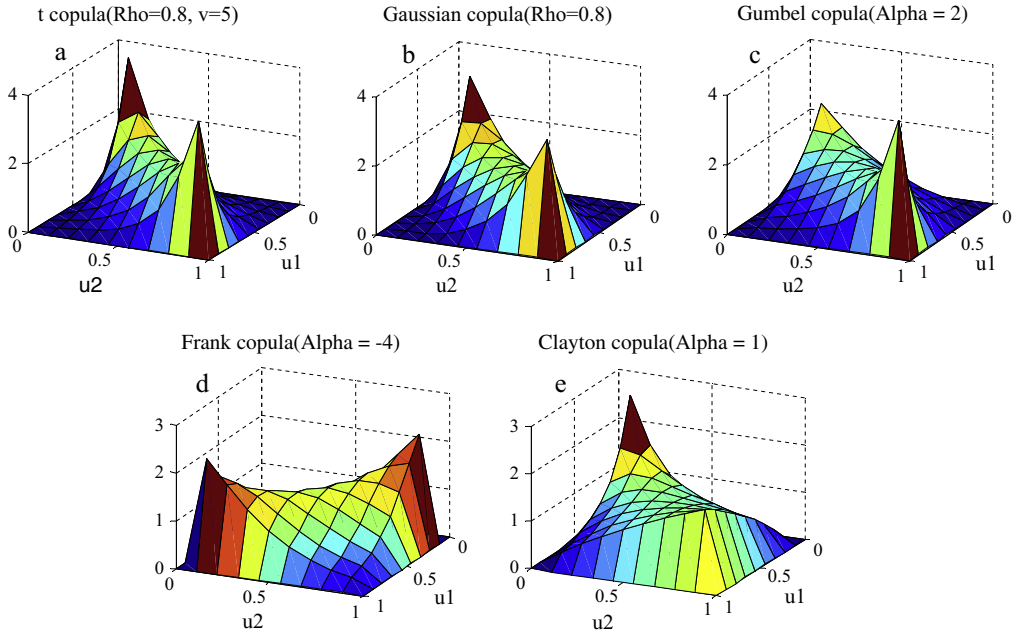

Fig. 3. Sample of different bivariate copula functions ((a): *t* copula (**Rho** = 0.8, ν = 5); (b): Gaussian copula (**Rho** = 0.8); (c): Gumbel copula (α = 2); (d): Frank copula (α = −4); (e): Clayton copula (α = 1)).

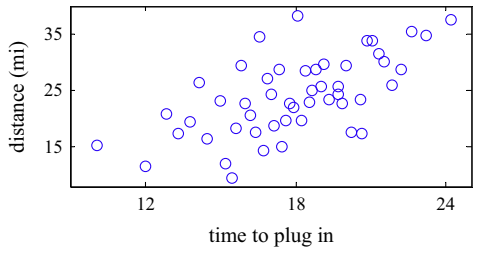

Fig. 4. Predicted times to plug in and traveled distances that is estimated by 50 EV owners.

transform the traveled distances into active power. Therefore, the converted load profile are demonstrated in Fig. 6 for 960 EV; both the actual and predicted data (with five copula functions). Black solid line in Fig. 6 introduces the exact collected load profile, while other five load profiles are the estimates by the GC, *t* copula, Clayton, Gumbel and Frank.

- Performances of five copulas are tested using the well-known mean absolute percentage errors (*MAPE*) as below:

$$MAPE = \left[\frac{1}{24}\sum_{i=1}^{24}\left|\frac{d_i - D_i}{d_i}\right|\right] \times 100 \quad (7)$$

where d_i and D_i are the real and predicted demand at the ith hour. Table 3 shows the calculated *MAPE* for all five copulas with respect to the exact data, introducing Clayton copula as the best copula among the five studied copulas to model the load profile for 960 EV. Additionally, the fitted distributions using copula are ranked in Table 4 according to the Akaike both information criterion (AIC) and Hannan-Quinn information criterion (HQIC) [19]. The higher the absolute values listed in Table 4, the better the fitting. Therefore, the Clayton copula is singled out as the best copula to model the charging demand according to both the AIC and the HQIC.

Hence, data generation takes place using Clayton copula for 960 EV.

The TCD seeking algorithm for the next day non-controllable load profile

One major problem for the aggregators is the lack of the HDD to schedule flexible loads optimally in smart grids [20]. The HDD is needed to estimate the hourly demand in different scaled terms [13–15]. Non-controllable loads are considerable portion of the load profile in which the collected HDD is classified according to the TCD. This classification scheme includes five classes for the TCD as shown in Fig. 7(a) (Friday is considered as a holiday). Aggregator collects the HDD over a full year for non-controllable loads for the case study (960 units).

The HDD as the exact dataset

Fig. 7(b) shows the general form of hourly demand for the TCD, where (d_1, d_2, …, d_{24}) and (d_{25}, d_{26}, …, d_{48}) illustrate the hourly demand for today and the next day in each class, respectively (totally 48 h). Thus, 52 collected data per year (or 1 per week) are available for every class of the TCD. In fact, the exact dataset for the HDD includes 48 vectors (hours) of size 52 weeks, i.e. the HDD forms a matrix of 52×48 for non-controllable loads. Actually, every hour during a year could be treated as a random variable; totally 48 variables that are correlated with each other at different times.

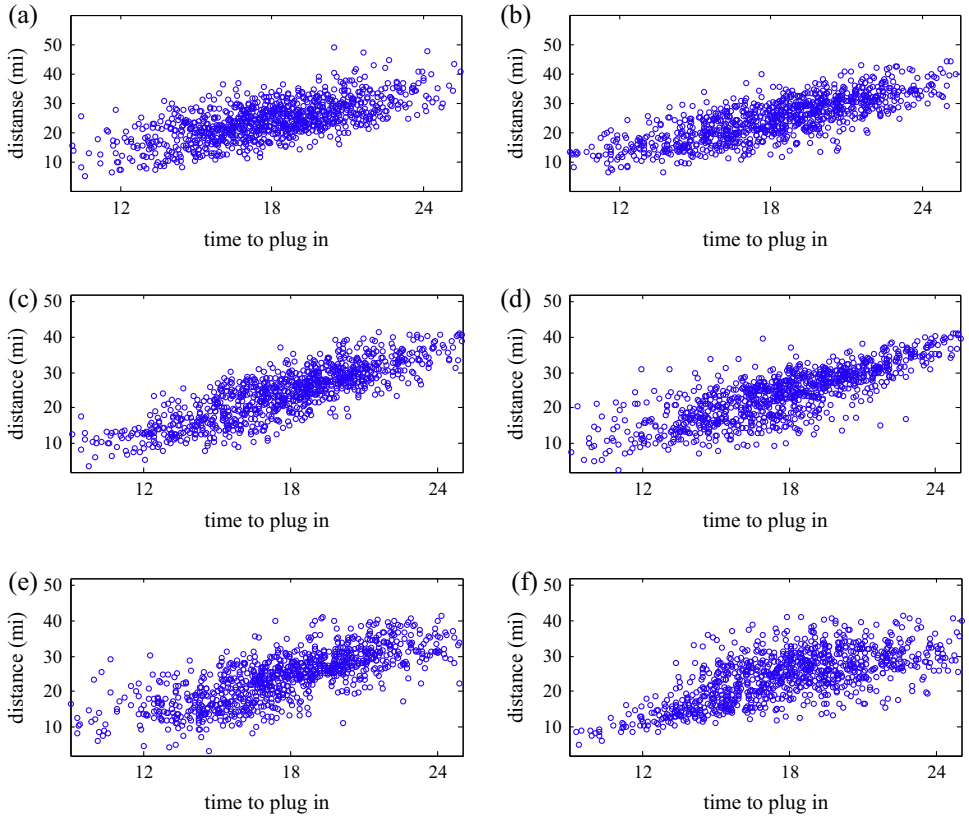

Fig. 5. Scattered plot of the times to plug in and the distances traveled for 960 users based on the actual data and using five types of copula function to model EV ((a): actual data; (b): GC (**Rho** = 0.73); (c): t copula (**Rho** = 0.85, ν = 1.4e7); (d): Gumbel copula (α = 2.64); (e): Frank copula (α = 7.8); (f): Clayton copula (α = 1.78)).

Table 2
Usual EV in the USA market [15].

Model	Battery capacity	Energy available	EV-Range	Max. Charge power rates
GM-Ch. Volt	16 kW h	8 kW h	40 mi	1.9 kW 3.3 kW*
Nissan-LEAF	24kW h	19.2kW h	100 mi	1.8 kW 3.3 kW* 49 kW
Volvo C30	24kW h	22.7kW h	93 mi	3.6 kW
Tesla Roadster	53kW h	37.1kW h	244 mi	1.8 kW 9.6 kW* 16.8 kW

* Max. charge power rates is considered mi: Mile.

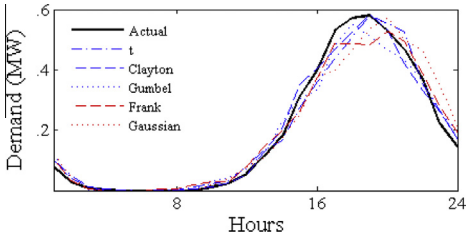

Fig. 6. The charging demand profile of 960 EV based on the actual and simulated data.

Table 3
MAPE of estimation the charging demand of EV using Archimedean copulas.

Copula	*t*	Clayton	Gumbel	Frank	Gaussian
MAPE	11.74%	8.94%	12.65%	13.20%	18.18%

Table 4
AIC and HQIC information criteria for modeling the charging demand of EV.

Familiar of copula	AIC	HQIC
Clayton	−4461.05	−4456.94
Gumbel	−3595.62	−3593.56
Frank	−3294.67	−3290.75
t	−2773.45	−2769.35
Gaussian	−2000.71	−1996.61

Selecting copula for scenario generation

These exact 48 datasets could be the variables of a multivariate copula function. Since the number of collected data is small for every hour (52), copula is capable of creating large datasets for all 48 variables (see Section 'Suitability of copula for creation of a large dataset'). Once again, both the AIC and HQIC information criteria selects the best performance for the five different copulas

Table 5
AIC and HQIC information criteria for modeling hourly demand of TCD.

Familiar of copula	AIC	HQIC
Gaussian	−20284.55	−16121.43
t	−18927.87	−14767.97
Gumbel	−3405.62	−3404.34
Frank	−3354.52	−3353.24
Clayton	−2562.54	−2561.26

under a typical load profile. Table 5 lists the worked out AIC and HQIC, showing the GC is the best of the five copulas to model the hourly energy consumption for the TCD.

Scenario generation by the GC

The parameters of the multivariate GC consist of 48 random variables in the TCD. Table 1 introduces a relationship that works out the dependencies between all these 48 variables; Table 6 shows the calculated correlation matrix. Having calculated the parameter **RHO**, the exact dataset (52×48) can be applied to the GC to generate as many scenarios as needed in practice. Here 1000 scenarios were generated in order to expand the exact dataset to a large 1000×48 dataset. Fig. 8 illustrates the estimated 1000 scenarios for a specific TCD (Tuesday–Wednesday).

Seeking for the TCD analogy

The first 24 columns in the large dataset shows today predicted scenarios, while the next 24 columns introduce the next day scenarios. It should be added that every row of the first 24 columns (of 1000×48 dataset) could be compared with that of the available today load profile (of 52×48 dataset). The most similar rows to the available today profile are selected according to the *MAPE* definition in (7). This paper assumes the maximum allowable *MAPE* to be equal to 8% ($MAPE \leqslant 8\%$) as the entrance condition for the best scenarios in choosing the next day load profile. All selected TCD-based scenarios are forwarded to the next step in Section 'Scenario selection and the DR proposals' to select the next day load profile (the last 24 columns) stochastically.

Scenario selection and the DR proposals

Since 187 TCD-based scenarios along with the predicted EV data are available, a stochastic process is required to work on the available data. Thus, an optimization problem can be formed in order to minimize the peak demand as well as the cost of electricity.

Optimization problem

Assume e_i and D_i^j are the charging demand of the EV (flexible load) and predicted demand of non-controllable loads (for the jth scenario) at the ith hour in the next day, respectively ($e_i + D_i^j$ = total demand at the ith hour). Also, P_j is the required

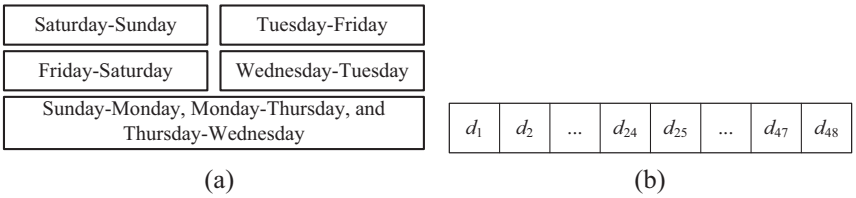

Fig. 7. The TCD structure, (a) five classes of two consecutive days, and (b) the general form of hourly demand in the TCD.

Table 6
Coefficient matrix, **Rho**, of the 48 variables GC ($m = 48$).

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}	u_{12}	u_{13}	u_{14}	u_{15}	u_{16}	u_{17}	u_{18}	u_{19}	u_{20}	u_{21}	...	u_{47}	u_{48}
u_1	1	0.82	0.84	0.83	0.85	0.85	0.74	0.85	0.73	0.72	0.76	0.8	0.75	0.67	0.59	0.66	0.7	0.61	0.61	0.51	0.41	...	0.46	0.49
u_2	0.82	1	0.9	0.81	0.89	0.89	0.77	0.75	0.67	0.76	0.8	0.82	0.78	0.76	0.7	0.72	0.72	0.67	0.7	0.64	0.55	...	0.54	0.57
u_3	0.84	0.9	1	0.95	0.96	0.97	0.87	0.91	0.89	0.91	0.9	0.89	0.89	0.83	0.79	0.82	0.83	0.83	0.83	0.66	0.56	...	0.59	0.58
u_4	0.83	0.81	0.95	1	0.97	0.94	0.9	0.95	0.89	0.88	0.84	0.86	0.88	0.79	0.76	0.83	0.87	0.87	0.78	0.66	0.54	...	0.59	0.51
u_5	0.85	0.89	0.96	0.97	1	0.97	0.9	0.92	0.87	0.86	0.83	0.86	0.86	0.79	0.78	0.85	0.85	0.81	0.76	0.7	0.61	...	0.55	0.53
u_6	0.85	0.89	0.97	0.94	0.97	1	0.9	0.91	0.89	0.92	0.89	0.89	0.89	0.82	0.78	0.84	0.87	0.85	0.83	0.68	0.59	...	0.65	0.64
u_7	0.74	0.77	0.87	0.9	0.9	0.9	1	0.88	0.8	0.79	0.84	0.79	0.85	0.84	0.83	0.88	0.93	0.87	0.81	0.87	0.8	...	0.56	0.49
u_8	0.85	0.75	0.91	0.95	0.92	0.91	0.88	1	0.89	0.86	0.86	0.91	0.88	0.79	0.71	0.78	0.86	0.85	0.79	0.68	0.58	...	0.6	0.51
u_9	0.73	0.67	0.89	0.89	0.87	0.89	0.8	0.89	1	0.94	0.85	0.86	0.9	0.77	0.69	0.78	0.81	0.87	0.87	0.57	0.45	...	0.6	0.52
u_{10}	0.72	0.76	0.91	0.88	0.86	0.92	0.79	0.86	0.94	1	0.88	0.88	0.9	0.77	0.66	0.74	0.79	0.86	0.88	0.55	0.42	...	0.68	0.56
u_{11}	0.76	0.8	0.9	0.84	0.83	0.89	0.84	0.86	0.85	0.88	1	0.94	0.92	0.92	0.76	0.75	0.83	0.87	0.91	0.64	0.54	...	0.7	0.66
u_{12}	0.8	0.82	0.89	0.86	0.86	0.89	0.79	0.91	0.86	0.88	0.94	1	0.92	0.85	0.67	0.73	0.81	0.84	0.85	0.58	0.46	...	0.65	0.64
u_{13}	0.75	0.78	0.89	0.88	0.86	0.89	0.85	0.88	0.9	0.9	0.92	0.92	1	0.87	0.67	0.75	0.82	0.89	0.92	0.69	0.52	...	0.76	0.63
u_{14}	0.67	0.76	0.83	0.79	0.79	0.82	0.84	0.79	0.77	0.77	0.92	0.85	0.87	1	0.87	0.83	0.85	0.83	0.82	0.7	0.61	...	0.68	0.65
u_{15}	0.59	0.7	0.79	0.76	0.78	0.78	0.83	0.71	0.69	0.66	0.76	0.67	0.67	0.87	1	0.93	0.88	0.8	0.67	0.67	0.66	...	0.53	0.6
u_{16}	0.66	0.72	0.82	0.83	0.85	0.84	0.88	0.78	0.78	0.74	0.75	0.73	0.75	0.83	0.93	1	0.96	0.88	0.73	0.7	0.64	...	0.6	0.62
u_{17}	0.7	0.72	0.83	0.87	0.85	0.87	0.93	0.86	0.81	0.79	0.83	0.81	0.82	0.85	0.88	0.96	1	0.95	0.81	0.74	0.65	...	0.66	0.64
u_{18}	0.61	0.67	0.83	0.87	0.81	0.85	0.87	0.85	0.87	0.86	0.87	0.84	0.89	0.83	0.8	0.88	0.95	1	0.91	0.68	0.54	...	0.73	0.65
u_{19}	0.61	0.7	0.83	0.78	0.76	0.83	0.81	0.79	0.87	0.88	0.91	0.85	0.92	0.82	0.67	0.73	0.81	0.91	1	0.71	0.57	...	0.7	0.59
u_{20}	0.51	0.64	0.66	0.66	0.7	0.68	0.87	0.68	0.57	0.55	0.64	0.58	0.69	0.7	0.67	0.7	0.74	0.68	0.71	1	0.95	...	0.37	0.26
u_{21}	0.41	0.55	0.56	0.54	0.61	0.59	0.8	0.58	0.45	0.42	0.54	0.46	0.52	0.61	0.66	0.64	0.65	0.54	0.57	0.95	1	...	0.23	0.21
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...
u_{47}	0.46	0.54	0.59	0.59	0.55	0.65	0.56	0.6	0.6	0.68	0.7	0.65	0.76	0.68	0.53	0.6	0.66	0.73	0.7	0.37	0.23	...	1	0.8
u_{48}	0.49	0.57	0.58	0.51	0.53	0.64	0.49	0.51	0.52	0.56	0.66	0.64	0.63	0.65	0.6	0.62	0.64	0.65	0.59	0.26	0.21	...	0.8	1

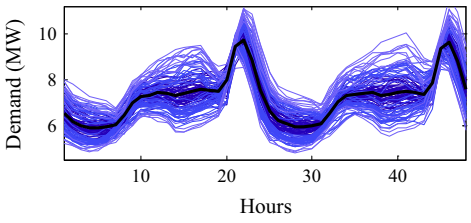

Fig. 8. Predicted 1000 scenarios for a typical TCD (created by the GC).

electrical energy to recharge the EV during the jth hour based on the predicted times to plug in and traveled distances using Clayton copula. If P_{total} is the total electricity needed to completely charge all EV based on the predicted data, and then the resultant charging schedule minimizes not only the peak demand but also the cost of electricity. The optimization problem is proposed as follows:

$$\min f(z) = \sum_{s}\sum_{i=1}^{23}\left[(D_{i+1}^{s} + e_{i+1}) - (D_i^s + e_i)\right]^2 \tag{8}$$

$$s.t.\begin{cases} e_i \leqslant \sum_{j=1}^{i} P_j \\ \sum_{i=1}^{24} e_i = \sum_{j=1}^{24} P_j = P_{\text{total}} \end{cases}$$

The proposed objective function in (8) minimizes the difference between total demands of every two consecutive hours, flattening the total load profile as much as possible. The first constraint in (8) restricts the selection of the flexible load e_i at the ith hour to smaller than or equal to the sum of predicted EV demand by Clayton copula. The second constraint forces the sum of all selected flexible loads by the optimization problem for 24 h is identical to all predicted demands for 24 h. The GAMS solves (8) as a stochastic process to optimally schedule the charging times of the EV. These resultant outcomes introduce e_i for 24 h that exclude any DR strategies. At the same time, bivariate and multivariate copula functions predict both the charging demands of 960 EV and non-controllable loads. Thus, one should expect estimation errors that affect the desired scheduling considerably. Hence, it is necessary to focus on proposing the DR strategies for the day-ahead load profile.

The DR proposals

This section introduces two new DR strategies in order to flatten the load profile as much as possible. Before describing the DR proposals, the following pre-assumptions obtained either by the optimization problem (8) or the real demands:

- e_i: The *next day* pre-scheduled demands for charging the EV during the ith hour based on the selected scenario using the GAMS (see optimization problem described by (8)).
- P_{iP}: *Today* real demand, compared to outcomes of the GAMS (e_i), required for charging the EV during the ith hour as realized available data.
- P_i: Modified today scheduled demand for the EV during the ith hour for removing the estimating errors.

The first proposed DR algorithm

Here a semi-automatically day-ahead DR strategy is proposed that requires bidirectional communication infrastructure. Main steps of the first proposal are as follows (see flowchart of Fig. 9):

- Step 1: If $P_{iP} \leqslant e_i$, then the charging demand of the EV is set to P_{iP} at the ith hour for the next day.
- Step 2: If $P_{iP} > e_i$, then the additional electrical energy required to charge the EV during the ith hour, $P_{iP} - e_i$, will be shifted to the remaining hours ($j \in \{i+1, i+2, \ldots, 24\}$) proportional to $\frac{e_j}{\sum_{n=i+1}^{24} e_n}(P_{iP} - e_i)$ as shown in Fig. 9.

The second proposed DR algorithm

The second proposed strategy requires unidirectional communication infrastructure to perform semi-automatically DR as shown in Fig. 10. First, the optimization problem in (8) was solved for 23 hourly demands ($e_i, i \in \{1, 2, \ldots, 23\}$). The main steps are as follows:

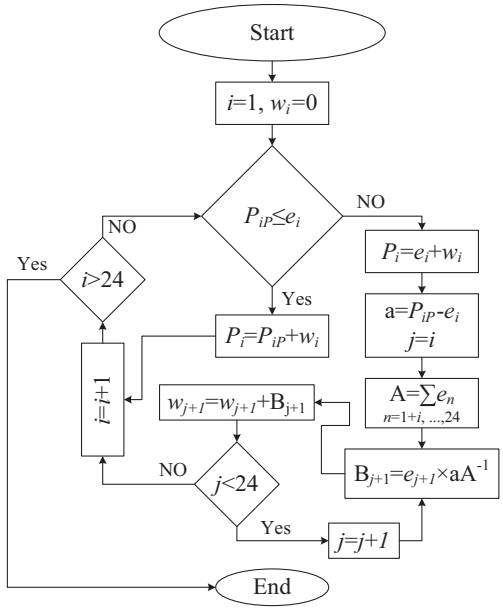

Fig. 9. The first proposed algorithm to schedule the charging demand of the EV.

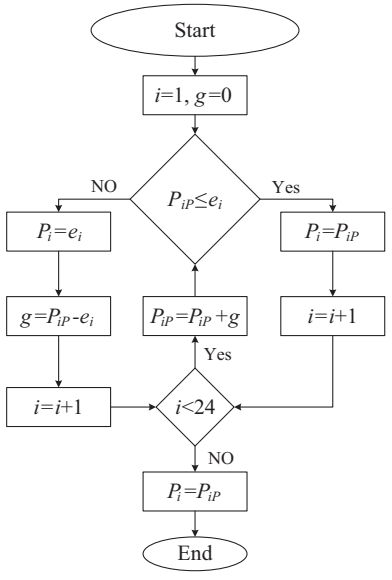

Fig. 10. The second proposed algorithm to schedule the charging of EV.

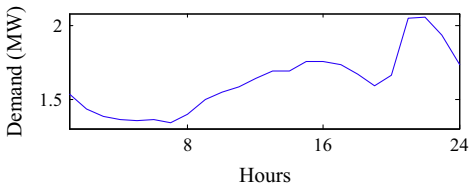

Fig. 11. Hourly load profile of the 960 residential consumers (non-controllable part) excluding their EV in 11/29/2011.

- Step 1: If $P_{iP} \leqslant e_i$, then the exact consumed energy would be equal to P_{iP}.
- Step 2: If $P_{iP} > e_i$, then the additional demand will be shifted to the next hour.
- Step 3: The rest of the additional charging demand in the current day is shifted to the Mth hour which was not pre-scheduled in order to guarantee the full charging of all EV.

The next section examines all proposed algorithms along with estimations and optimization in order to verify the suggestions.

Case studies and discussions

Assume the case study investigates the load profile of 960 households that participated in the day-ahead DR programming. This investigation tracks assessment of data prediction for flexible (EV) and non-controllable loads, seeking for the TCD-based analogy and the two DR strategies. The performed simulations demonstrate the efficiency of the proposed strategies. Let us assume all 960 EV are charged up to the maximum charge power rates (MCPR) according to the listed data in Table 2. Two circumstances are studied in simulations as below:

- The group of EV includes 960 vehicles of the same brand, namely GM-Chevy Volt. Table 2 shows that the MCPR for GM-Chevy Volt is 3.3 kW h.
- The group of EV includes 960 vehicles of the same brand, namely Tesla Roadster. Table 2 shows that the MCPR for Tesla Roadster is 9.6 kW h.

Data generation

Two copula functions are responsible for creating the required data; first, Clayton copula that predicts distances travelled as well as times to plug in by 960 EV from an exact small dataset (see Sections 'Exact data gathering' and 'Data generation'). The small dataset gathered from 50 EV owners on November 29, 2011 that gave their estimates on the two named variables for November 30, 2011. Then, Clayton copula generates the large dataset of size 960 EV for November 30, 2011. Second, the GC is used to create non-controllable load profile. Available historical data are the hourly non-controlled demand for one year up to November 29, 2011. Then, these data were applied to the GC in order to correlate random variables and generate new data. Table 6 shows the calculated parameters (52×48) of the multivariate GC (**Rho**) for the target TCD (November 29 and 30, 2011). Eventually, the GC creates a dataset of size 1000×48, where 187 out of 1000 scenarios were chosen ($MAPE \leqslant 8\%$) based on the hourly demand in November 29, 2011. Fig. 11 illustrates the non-controlled demand on November 29, 2011excluding the EV charging demand. Fig. 12 shows the created 1000×48 datasets in blue lines (by the GC) as well as 187×48 selected datasets in green lines ($MAPE \leqslant 8\%$). Black line in Fig. 12 depicts the actual non-controlled demand on November 29 and 30, 2011.

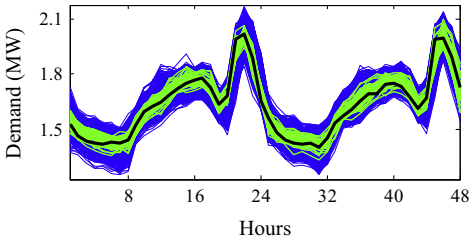

Fig. 12. 1000 estimated hourly demand of TCD for 960 residential units excluding their EV.

Case study 1: 960 EV of brand GM Chevy-Volt

The first case study simulates prediction of the day-ahead load profile (960 residential units) including non-controllable loads plus the flexible EV charging powers. First, flexible part (the EV) of the load profile excludes any DR strategies as shown by black dashed lines in Fig. 13(a) (see details of both flexible and non-controllable loads in Fig. 13(b)). Then, the flexible part goes through the two DR proposals; blue dotted lines (the DR strategy in Fig. 9) and red dotted lines (the DR strategy in Fig. 10) in Fig. 13(a) demonstrate the outcomes (see details of both flexible and non-controllable loads in Fig. 13(c)–(d)). Meanwhile, it is useful to obtain the effect of actual data (after the occurrence) when optimized by (8) and undergone by the two DR strategies. (Note that actual data is practically unavailable before the occurrence.) Blue[1] line in Fig. 13(a) provides the realized load profile (see details of both flexible and non-controllable loads in Fig. 13(e)). Hence, the closeness of the predicted data can be examined with those of the actual data in Fig. 13(a).

According to the simulations (see Fig. 13), the predicted demand is acceptably close to the actual scheduling. To put it on a firm basis, the peak demands in Fig. 13(a) are lowered by 22.435% (0.46 MW) when applying the proposed DR strategies. Additionally, a typical 12-level tariff rate structure introduced in [21] for the DAP with IBR as listed in Table 7. This could be used as an assessing means of the proposed day-ahead DR algorithms. Comparing different load profiles in Fig. 13(a) reveals that the energy cost of the aggregator is $26,631 per day when excluding the DR. Energy cost reduces to $21,736 per day for the DR strategy in Fig. 9 and $21,721 per day for the DR strategy in Fig. 10. Interestingly, the actual energy cost is $21,643 per day for the actual data after the occurrence (very close to those of the predicted data). Therefore, the cost of energy by scheduling the charging of the EV is reduced by 18.53%, 18.44% and 18.88% for the first DR strategy, the second DR strategy and the actual data, respectively.

Case study 2: 960 EV of brand Tesla-Roadster

The second case study was also simulated like that of the first case. Fig. 14(a) describes exactly the same situation as that of Fig. 13(a), but for the second case study in which brand of the EV has changed. In addition, Fig. 14(b) introduces one non-controllable load profile along with four flexible load profiles for the EV (excluding the DR, including two proposed DR strategies and actual data plus DR). Comparing the load profiles 'excluding the DR' with those of 'including the DR' in Fig. 14(a) shows that the peak demand decreases by 17.84% (0.446 MW). Moreover,

[1] For interpretation of color in Fig. 13, the reader is referred to the web version of this article.

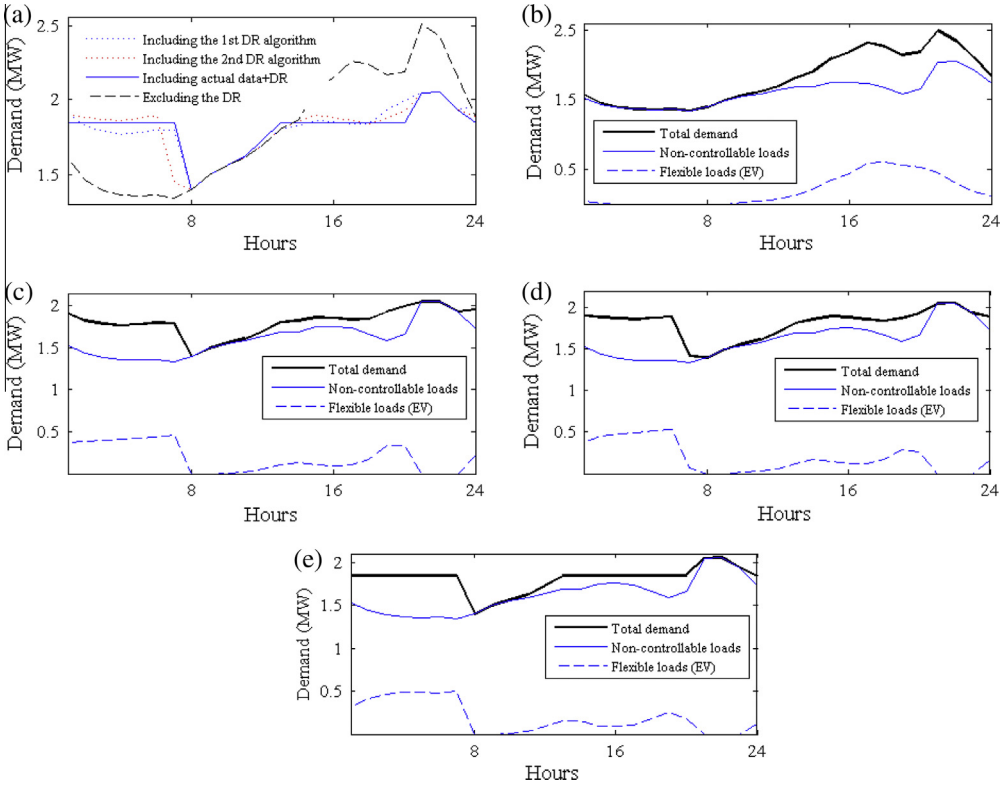

Fig. 13. The load profile of 960 end-users under four different types of charging management in the first case study, (a) total load profiles, (b) load profile decomposition for the case "excluding the DR", (c) load profile decomposition for the case "The first DR algorithm", (d) load profile decomposition for the case "The second DR algorithm", and (e) load profile decomposition for the case "Actual data plus the DR".

Table 7
Assumed DAP with IBR model for retail market.

Demand (MW)	Price ($/MW)	Demand (MW)	Price ($/MW)
0.0–0.5	110.00	1.7–1.8	445.01
0.5–0.8	126.50	1.8–1.9	511.76
0.8–1	145.47	1.9–2	588.52
1–1.1	167.29	2–2.1	676.80
1.1–1.2	192.39	2.1–2.2	778.32
1.2–1.3	221.24	2.2–2.3	895.07
1.3–1.4	254.43	2.3–2.4	1029.33
1.4–1.5	292.60	2.4–2.5	1183.73
1.5–1.6	336.49	2.5–2.6	1361.30
1.6–1.7	386.96	2.6–2.8	1565.49

the mentioned tariff rate structure in Table 7 applied to the load profile to obtain the energy cost. Comparing different load profiles in Fig. 14(a) demonstrates that the energy cost of the aggregator is $26,542 per day when excluding the DR. Energy cost reduces to $21,732 per day for the DR strategy in Fig. 9 and $21,771 per day for the DR strategy in Fig. 10. Interestingly, the actual energy cost is $21,721 per day for the actual data after the occurrence (very close to those of the predicted data). Therefore, the cost of energy by scheduling the charging of the EV is reduced by 18.12%, 17.97% and 18.16% for the first DR strategy, the second DR strategy and the actual data, respectively. Table 8 illustrates comparing load profiles in Fig. 13 with those of Fig. 14 as well as the aggregator payments for the two case studies. It can be concluded from Table 8 that the predicted data with the proposed DR algorithms produce very close outcomes to those of the actual data plus the DR program.

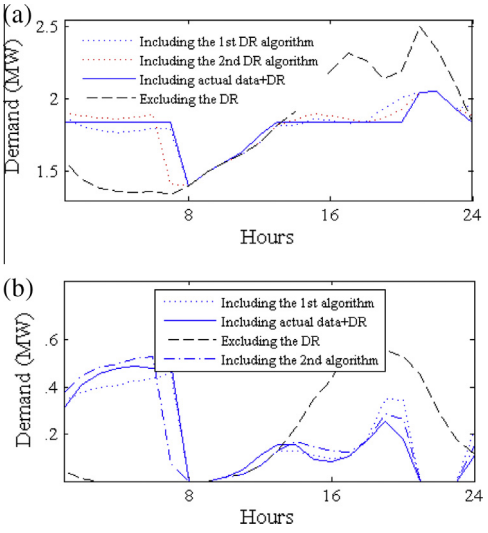

Fig. 14. Four different types of flexible load management (the EV) in the second case study, (a) total load profiles, and (b) flexible parts of the load profiles.

Table 8
Energy cost ($) of an aggregator in a day with and without applying different DR strategies.

Scenario	Excluding DR	1st proposed DR algorithm	2nd proposed DR algorithm	Optimal DR
1	26,631	21,736	21,721	21,643
2	26,542	21,732	21,771	21,721

Conclusion

This paper uses Clayton copula to predict distances travelled as well as times to plug in for a large dataset of size 960 EV (flexible part of the next day load profile) from an exact small dataset. In addition, the GC creates large datasets of size 1000×48 from available small datasets of size 52×48 for each TCD-based combination (non-controllable part of the next day load profile for 960 residential units). Copula functions are selected based on three tests, namely *MAPE*, AIC and HQIC. Then, 1000 scenarios of 48 h load profile is narrowed down to 187 scenarios by applying the condition $MAPE \leqslant 8\%$. Further, an optimization problem is arranged in which flexible load (one scenario) plus non-controllable load (187 scenarios) is flatten for two consecutive hours as much as possible. Resultant non-controllable scenario is fixed for the rest of study, while the flexible load (the EV) goes through two suggested DR strategies. The main idea is to flatten the whole day-ahead load profile in order to lower the peak demand, reducing the energy payment by the aggregator. Hence, two case studies are arranged in which a 12-level tariff rate structure is considered for the DAP with IBR that investigate the proposed DR strategies. Moreover, actual data after occurrence are collected, where passes through a DR program. This helps evaluation of the DR strategies to be compared analytically according to the DAP with the IBR. Simulations and analytical comparisons confirm the closeness of the predicted load profile along with the suggested DR strategies to the actual data plus the DR program.

References

[1] Cao Yijia. An optimized EV charging model considering TOU price and SOC curve. IEEE Trans Smart Grid 2012;3:388–93.
[2] Ortega-Vazquez MA. Electric vehicle aggregator/system operator coordination for charging scheduling and services procurement. IEEE Trans Power Syst 2013;2:1806–15.
[3] Ahn C. Optimal decentralized charging control algorithm for electrified vehicles connected to smart grid. J Power Sources 2011;196:10369–79.
[4] Sortomme E. Optimal charging strategies for unidirectional vehicle-to-grid. IEEE Trans Smart Grid 2011;2:131–8.
[5] Li Gan. Modeling of plug-in hybrid electric vehicle charging demand in probabilistic power flow calculations. IEEE Trans Smart Grid 2012;3:492–9.
[6] Verzijlbergh Remco A. Network impacts and cost savings of controlled EV charging. IEEE Trans Smart Grid 2012;3:1203–12.
[7] Lojowska A, Kurowicka D, Papaefthymiou G. Stochastic modeling of power demand due to EV using copula. IEEE Trans Power Syst 2012;27:1960–8.
[8] Shao Shengnan. Demand response as a load shaping tool in an intelligent grid with electric vehicles. IEEE Trans Smart Grid 2011;2:624–31.
[9] He Yifeng. Optimal Scheduling for charging and discharging of electric vehicles. IEEE Trans Smart Grid 2012;3:1095–105.
[10] Wu Di. Load scheduling and dispatch for aggregators of plug-in electric vehicles. IEEE Trans Smart Grid 2012;3:368–76.
[11] Shao Shengnan. Grid Integration of electric vehicles and demand response with customer choice. IEEE Trans Smart Grid 2012;3:543–50.
[12] http://en.wikipedia.org/wiki/Deregulation_of_the_Texas_electricity_market. http://www.wikipedia.org/.
[13] Taylor JW. Short-term load forecasting with exponentially weighted methods. IEEE Trans Power Syst 2012;1:458–64.
[14] Al-Hamadi HM, Soliman SA. Fuzzy short-term electric load forecasting using Kalman filter. IEE Proc-Gener Transm Distrib 2006;2:217–27.
[15] Filho K, Lotufo ADP, Minussi CR. Short-term multimodal load forecasting using a modified general regression neural network. IEEE Trans Power Delivery 2011;4:2862–9.
[16] Nelsen RB. An introduction to copulas. 2nd ed. New York: Springer; 2006.
[17] Schmidt T. Coping with copulas. In: Rank J, editor. Copulas: from theory to applications in finance. Risk Books; 2007.
[18] Goda K. Statistical modeling of joint probability distribution using copula: application to peak and permanent displacement seismic demands. Struct Saf 2010;32:112–23.
[19] Help File for Model Risk Version 4, © Vose Software (2007).
[20] Azadeh. Forecasting electrical consumption by integration of neural network, time series and ANOVA. Appl Math Comput 2007;186:1753–61.
[21] Ahmadi D, Tavakoli Bina M. Modeling and estimating the energy consumption of household electrical equipment having stochastic time of use using Gaussian copula. In: 27th international power system conference (PSC'12).