A Comparison of Scenario Generation Methods for the Participation of Electric Vehicles in Electricity Markets

Ida Græsted Jensen[a,*], Niels Framroze Møller[a], Giovanni Pantuso[b], Nina Juul[a]

[a]*Technical University of Denmark, DTU Management Engineering, Produktionstorvet, Building 426, DK – 2800 Kongens Lyngby, Denmark*
[b]*Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK - 2100 Copenhagen, Denmark*

Abstract

We consider the procurement of electricity for a large fleet of electric vehicles operating in electricity markets. Due to uncertain regulating prices, this problem has typically been modelled as a stochastic program. In this study, we address the issue of generating scenario trees, i.e. simplified representations of the uncertainty necessary to solve the corresponding stochastic programs. A trade-off between accurate descriptions of the uncertainty and tractability of the stochastic program is sought. Based on data describing electric mobility and electricity prices in Denmark, general-purpose scenario generation strategies are tested and compared. Such strategies include state-of-the-art property matching methods and time-series analysis. The results show that the co-dependence between the regulating prices at different hours of the day plays a crucial role when generating scenario trees for these problems, making copulas an important property to consider. This information can help decision makers to achieve better (cheaper) electricity procurement by accurately preprocessing the uncertainty in the regulating prices.

Keywords: Electric vehicles, Regulating Market, Stochastic Programming, Scenario Generation

1. Introduction

The use of electric vehicles (EVs) is becoming increasingly popular due to factors such as governmental incentives towards low-carbon transportation (see Coffman et al.[1]) and more economically feasible technologies (see Nykvist and Nilsson[2] which set a positive outlook for the future usage of EVs). Consequently, managing their charging activities becomes necessary to avoid voltage and congestion problems in the grid and to reduce the costs of charging for the end users. In fact, an uncoordinated simultaneous charging of a large number of EVs is likely to cause unexpected peak loads to the grid. Mediators between large fleets of EVs and the electricity markets are thus needed. Such a mediator is called an *EVs aggregator* (EVA). The role of the EVA is essentially that of pooling and coordinating the participation of large fleets of EVs in electricity markets to procure the required power, see e.g., Brooks et al.[3], Kristoffersen et al.[4], Vagropoulos and Bakirtzis[5], Bessa and Matos[6],

---

*Corresponding author: idje@dtu.dk

Juul et al.[7], Tavakoli et al.[8], Vagropoulos et al.[9]. By participating in the electricity markets, an EVA can lower the electricity bill by charging at low prices and being paid for providing ancillary services (i.e. up- or down-regulation). However, this requires EVAs to consider market dynamics when planning the charging of the EVs and possibly adapt their usage. In general, the owners of the EVs agree to a charging management scheme and benefit from lower electricity prices. However, in many cases, the owner of the EVs and the EVA can be the same subject as, for instance, in car-sharing services and in large companies using EVs for deliveries.

Among the papers studying the participation of EVs or EVAs in electricity markets, several use stochastic programming as a modelling tool. For example, Al-Awami and Sortomme[10] study the problem of mitigating trading risk by coordinating the stochastic supply (from e.g. wind) scheduling with responsive demand given by unidirectional vehicle-to-grid services. The authors used a mixed-integer linear stochastic program. Khalid et al.[11] proposed an optimal bidding algorithm for a unidirectional vehicle-to-grid which considers the uncertainty in regulating prices. The problem is formulated using a stochastic program with the objective of maximizing the total profit of the EVA participating in the regulating market. The authors demonstrate the advantage of the stochastic program over the deterministic counterpart using simulation. Vagropoulos and Bakirtzis[5] consider an EVA participating in both the day-ahead energy and regulation markets and use a stochastic program to determine the optimal bidding strategy. The authors test the model using an EVA representing 1000 EVs. Finally, Juul et al.[7] also consider the possibility of acting both in the day-ahead and regulating markets. The authors compare several charging strategies for a fleet of EVs, ranging from simply charging whenever the vehicle is parked to coordinating with the participation in electricity markets. Among the strategies explored, the authors propose a stochastic program which considers uncertainty in the regulating prices. Similar to Khalid et al.[11], the authors show the advantages of using a stochastic program rather than a deterministic program.

However, solving stochastic programs in practice creates additional challenges compared to classical deterministic mathematical programs. Among these, there is the need of a preprocessing phase known as *scenario generation* which has the purpose of providing suitable input data for the uncertain parameters of the problem. In fact, except for trivial cases, solving a stochastic program requires a (possibly small) discrete version of the (typically continuous and multivariate) probability distribution describing the uncertainty. Such discrete version is referred to as a *scenario tree*. Several techniques have been proposed for generating scenario trees (see e.g. Dupacová et al.[12], Pflug[13], Høyland and Wallace[14], Høyland et al.[15], Kaut[16]). Eventually, as stressed in[17], when generating scenario trees, the aim is not necessarily that of obtaining the best approximation of the original distribution in a statistical sense. Rather, the aim is obtaining a discrete distribution which allows decision makers to make as good a decision as they would with the original distribution, or as close as possible to that. The difference between the expected result of using the solution to the true problem (i.e. the often intractable problem based on the original distribution) and the expected result of using the solution obtained by means of a scenario tree is referred to as the *bias* (see e.g. Kaut and Wallace[18] and King and Wallace[17]) and is the key measure used throughout this paper.

In this paper we focus on an EVA that participates in electricity markets in order to charge

a fleet of EVs and on the problem of providing suitable input parameters (i.e. a scenario tree) for the corresponding stochastic program. The scope of this study is to compare general-purpose scenario generation strategies with the aim of identifying those that can provide a descriptive but small representation of the uncertainty. That is, a scenario tree is considered descriptive if it leads to solutions whose expected result (e.g. cost) is sufficiently close to the expected result of the solution obtained using the true distribution (the latter of which is achievable only if the original problem can be solved, as in this paper). That is, descriptiveness is measured in terms of bias (the difference between the results of using the true versus an approximated random process, as explained in Section 3). Furthermore, in this context, a scenario tree is *small* if the number of scenarios it contains allows the stochastic program to be solved in an acceptable time limit.

Within the field of optimisation in energy, ad-hoc scenario generation methods have been proposed in order to model the stochasticity generated by renewable energy sources. For instance, Pinson et al.[19] propose a method that permits the generation of statistical scenarios of short-term wind generation. The method is based on the transformation of a series of prediction errors into a multivariate Gaussian random variable. The method, which requires the computation of a covariance matrix, can also be adopted to treat general random variables. Wang et al.[20] capture the dependency between renewable sources, choosing the best fitting copula for scenario generation. Wang et al.[21] generate a probabilistic net load forecast using point forecasts and a copula-based sampling method to construct the conditional error distribution. Toubeau et al.[22] seek to obtain multivariate scenarios by sampling the predicted multivariate distribution using a copula-based strategy such that temporal information and cross-variable dependencies are considered. Lee and Baldick[23] use the generalised dynamic factor model to generate load and wind power scenarios in such a way that the correlation structure between load and wind is preserved. Chen et al.[24] and Vagropoulos et al.[25] use neural networks in order to be able to capture both the linear and non-linear the dependencies in data. Chen et al.[24] focus on generative models, where scenarios are generated based on unsupervised (machine) learning from historical data.

In this paper we focus only on general purpose scenario-generation method as they are adaptable to any source of uncertainty, including renewable production sources. Several general-purpose scenario generation methods among the most widely used are evaluated. These include the discretisation of a distribution estimated through time-series analysis, property matching (and particularly moment matching, see Høyland and Wallace[14] and Høyland et al.[15]) and copula matching (see Kaut[16]).

For testing and comparing the scenario generation methods, the stochastic programming model introduced by Juul et al.[7] is applied to a case study based on Denmark. The reason for choosing the model in Juul et al.[7] is twofold. First, it is sufficiently representative of the underlying decision process, as it includes both the day-ahead and regulating markets and, second, it is simple enough to allow us to solve the problem using the original description of uncertainty. In particular, the second point is necessary to calculate the bias. The problem considered has a 24-hours planning horizon.

The remainder of this paper is organised as follows. Section 2 briefly summarises the problem and introduces the corresponding mathematical model. Section 3 clarifies how scenario generation methods are assessed and introduces a number of scenario generation methods that were tested. Then Section 4 reports the results of a computational study based

on mobility and electricity market data from Denmark. The final conclusions are presented in Section 5.

2. Problem Description and Mathematical Model

An EVA must procure electricity to charge a fleet of EVs to guarantee given driving tasks for each vehicle. The driving tasks are assumed to be deterministic; therefore the EVA knows for each hour how much each vehicle drives. This is the case, for example, of a fleet of EVs engaged in delivery operations for which the routing has been predetermined or the case of EVs booked in a car-pooling service. The hourly load for the fleet is thus known. The EVs in the fleet can be charged at any time during the planning horizon except for the hours when they are driving, and the *charging strategy* can be defined as the charging amount for each EV and for each hour in the planning horizon.

The EVA has the option to buy electricity at the spot price. In addition, the EVA participates in the regulating market. This allows the EVA to emend the original spot-market purchase if profitable. In particular, at a given hour of the day, the EVA can offer up- or down-regulation to the grid. If the grid needs up-regulation, i.e., injection of electricity, the EVA can consider selling to the grid part of the electricity bought in the spot market. In this case, the EVA would emend the original charging strategy by stopping or curtailing the charge scheduled for some EVs. In addition, if the grid needs down-regulation, i.e., in case of excess of production, the EVA can consider buying additional electricity, thus charging the EVs batteries more than originally planned (if there is enough battery capacity left in the fleet). Therefore, the EVA can lower the electricity bill by being remunerated for ancillary services.

However, when the charging strategy is decided, only the hourly spot prices are known, while the status of the grid is uncertain, as well as the regulating price in each hour in the planning horizon. That is, we assume that when the charging strategies are decided, the spot market has been cleared and the spot prices realised. Therefore, the charging strategy and the consequent amount of electricity to buy in the spot market must be decided under uncertainty. The problem was modelled by Juul et al.[7] as a two-stage stochastic program where first-stage decisions establish a charging amount for each hour to be bought in the spot market, while second-stage decisions amend the original plan by curtailing or increasing the planned charging if profitable.

To introduce the mathematical model for the problem, let $\mathcal{J}$ be the set of vehicle types, where each vehicle type has a different driving pattern, and $\mathcal{T}$ represent the set of time periods. In addition, the following parameters are known to the decision maker. Let C^{MAX} be the maximum charging which can be done in a time period, η be the charging efficiency, D_{jt} be the charge consumed by the driving of vehicles of type j at time period t, L_j^{MIN} and L_j^{MAX} be the minimum and maximum storage capacity of vehicle type j, respectively, and P_t^S be the spot price for time period t. However, the regulating prices are uncertain and depend on the status of the grid at a specific time period t. Let ω be a random event representing the status of the grid in each hour of the planning horizon, and let $P_t^R(\omega)$ be a random variable representing the regulating price at time period t and $\xi = (P_1^R(\omega), \ldots, P_{|\mathcal{T}|}^R(\omega))$ the collection of regulating prices. The probability distribution of ξ is assumed to be known. Let x_{jt} be a first-stage decision variable representing the charge purchased at the spot price for vehicle

4

type j at time period t, with $x = (x_{11}, \ldots, x_{|\mathcal{J}|,|\mathcal{T}|})$. Let $\lambda_{jt}^{UP}(\omega)$ and $\lambda_{jt}^{DOWN}(\omega)$ be second-stage decision variables representing the amount of up- and down-regulation, respectively, and $l_{jt}(\omega)$ be a second-stage decision variable representing the storage level on-board vehicles of type j in time period t. Notice that second-stage decision variables are adaptable to the status of the grid, i.e. are made once the regulating prices become known. The problem can thus be formulated as the following two-stage stochastic program.

$$\min_{x} f(x, \xi) = \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} P_t^S x_{jt} + E_\xi \left[\sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} P_t^R(\omega) \left(\lambda_{jt}^{DOWN}(\omega) - \lambda_{jt}^{UP}(\omega) \right) \right] \tag{1a}$$

$$\text{s.t. } l_{jt}(\omega) = l_{j,t-1}(\omega) + \eta(x_{jt} + \lambda_{jt}^{DOWN}(\omega) - \lambda_{jt}^{UP}(\omega)) - D_{jt}, \qquad j \in \mathcal{J}, t \in \mathcal{T}, \tag{1b}$$

$$L_j^{MIN} \leq l_{jt}(\omega) \leq L_j^{MAX}, \qquad j \in \mathcal{J}, t \in \mathcal{T}, \tag{1c}$$

$$x_{jt} + \lambda_{jt}^{DOWN}(\omega) \leq C^{MAX}, \qquad j \in \mathcal{J}, t \in \mathcal{T}, \tag{1d}$$

$$x_{jt} - \lambda_{jt}^{UP}(\omega) \geq 0, \qquad j \in \mathcal{J}, t \in \mathcal{T}, \tag{1e}$$

$$D_{jt}(x_{jt} - \lambda_{jt}^{UP}(\omega)) = 0, \qquad j \in \mathcal{J}, t \in \mathcal{T}, \tag{1f}$$

$$D_{jt}\lambda_{jts}^{DOWN}(\omega) = 0, \qquad j \in \mathcal{J}, t \in \mathcal{T}, \tag{1g}$$

$$x_{jt}, l_{jt}(\omega), \lambda_{jt}^{UP}(\omega), \lambda_{jt}^{DOWN}(\omega) \geq 0, \qquad j \in \mathcal{J}, t \in \mathcal{T}. \tag{1h}$$

Objective function (1a) consists of the cost of the electricity bought at the spot price plus the expected cost or revenue for offering down or up regulation, and constraints (1b) monitor the power inflow and outflow in the batteries of the vehicles between two consecutive time periods. The storage level at a specific hour is equal to the storage level at the previous hour plus the electricity bought from the market. Constraints (1c) state that the battery level must be within a given upper- and a lower-bound. Constraints (1d) ensure that the maximum charging capacity for every time period is respected. Constraints (1e) state that up-regulation can be provided by selling part of the electricity bought at the spot price. Up-regulation is provided without actually discharging the battery but rather by giving up part of the electricity bought at the spot price. Constraints (1f)-(1g) state that cars are not charged while driving, i.e. when $D_{js} > 0$. Finally, constraints (1h) set the domain for the decision variables.

3. Analysis Framework

This section first discusses the general issue of generating and evaluating scenario trees for problem (1) in Section 3.1, and then introduces suitable general-purpose state-of-the-art scenario generation methods in Section 3.2.

3.1. Scenario Trees and Bias

In most practical cases, the vector of regulating prices ξ is a continuous random variable fully described by a (multivariate) probability density function or by a discrete distribution with a very large sample space. This, in turn, renders problem (1) intractable. Consequently, a discrete approximation must be used, $\hat{\xi}$ of ξ which contains a suitably small number of possible outcomes and giving up something in terms of quality of the solution. That is, the solution to a different problem is obtained, namely:

$$\hat{x} = \operatorname*{argmin}_{x} f(x, \hat{\xi}) \tag{2}$$

The expected loss incurred in this transition from ξ to $\hat{\xi}$ is referred to as the *bias* (see, e.g., Kaut and Wallace [18]) and is calculated as follows:

$$b_{\hat{\xi}} = f(\hat{x}, \xi) - z = f(\hat{x}, \xi) - \min_x f(x, \xi)$$

Therefore, the following two objectives are the focus when generating scenario trees $\hat{\xi}$: 1) having a bias as small as possible, and 2) having a number of outcomes small enough to be able to solve problem (2). Especially in sampling methods, these objectives are conflicting, as increasing the cardinality of $\hat{\xi}$ to infinity eventually produces ξ, thus typically reducing the bias.

To calculate the bias, the original problem (1) must be solvable. However, obviously, if (1) were solvable, there would not be the need to find $\hat{\xi}$ in the first place. Thus, in most real-life applications, (1) cannot be solved (e.g. when ξ is continuous) and must thus revert to bounds on the bias (see e.g. Mak et al. [26] and Pflug [13]), or (1) can be solved in an amount of time which is not suitable for the decision maker. In the latter case, the bias can be calculated exactly by solving (1) only once in a possibly very long amount of time. This provides an idea of how much the solutions to $\hat{x}$ could be improved. This paper works with a case falling in the second category, as the goal is to estimate the bias exactly. The procedure to calculate the bias can be summarised as follows:

1. solve the stochastic program (1) built on ξ and store its optimal objective value z
2. solve the stochastic program (2) built on a scenario tree $\hat{\xi}$, which approximates ξ, and store its optimal first-stage solution $\hat{x}$
3. solve the stochastic program (1) built on ξ, but such that the first-stage variables are fixed to $\hat{x}$, and store its optimal objective value $\hat{z}$
4. calculate the bias as $b_{\hat{\xi}} = |\hat{z} - z|$

3.2. Scenario Generation Methods

In the following sections, three general-purpose state-of-the-art scenario generation methods are summarised for problem (1).

We do not consider scenario reduction techniques, such as Dupacová et al. [27], Heitsch and Römisch [28], and Morales et al. [29] in order to isolate the effects of the scenario generation methods considered. Scenario reduction techniques have the scope of reducing the number of scenarios while preserving the probabilistic information they contain. Such techniques might be particularly beneficial in cases where a mere application of a scenario generation technique results in a scenario tree too large for solving the resulting problem. We refer the reader to the above-mentioned literature for further details.

3.2.1. Scenario generation based on time-series regressions

The idea of this approach is to use the historical time-series data to estimate regression equations for the regulating prices and then use these equations to construct the scenario trees. First, the conditional distribution of the regulating price for hour 1, day τ, conditional on the available information up to but not including day τ was computed. This information thus contains knowledge of all spot prices from day τ (as these are known in advance) and all past realised spot and regulating prices (e.g. from days, $\tau - 1, \tau - 2$). To be specific, this

conditional distribution is estimated in the form of a histogram constructed from adding the estimated conditional mean and residuals from a regression equation for the regulating price for hour 1. The latter equation can thus be written as a regression model with $AR(k)$ errors, i.e.

$$p_{1,\tau}^r = \lambda^T x_\tau + u_{1,\tau},$$

where $p_{1,\tau}^r$ is the logarithm of the regulating price in hour 1, day τ, x_τ is a regressor vector including a constant, the 24 spot prices (also in logarithms) from day τ, indicators for the day of the week and a sine-cosine term for the annual seasonality. The corresponding coefficient vector is λ with superscript T denoting the transpose. The error term, $u_{1,\tau}$ follows an auto-regressive equation, i.e. $u_{1,\tau} = \rho_1 u_{1,\tau-1} + ... + \rho_k u_{1,\tau-k} + e_{1,\tau}, e_{1,\tau} \sim i.i.d(0,\sigma^2)$. Note that in this regression, redundant terms such as lagged daily indicators were removed. Note also that, although in the initial estimation spot prices for *all* hours of the current day were included in the regression, only the spot price in hour 1 turned out to be significant. Thus, the spot prices for the remaining hours were excluded from the regression. In the proposed computations, the outcomes for hour 1 are thus taken as the outcomes associated with the histogram, i.e. the midpoints. Second, for any hour t of the remaining hours, $t = 2, .., 24$, the outcome of the regulating price is computed as the estimated conditional mean, and is still conditional on the same information as above and the outcome of the regulating price for hour 1 (the above-mentioned midpoints). The conditional mean for the regulating price of hours $t = 2, .., 24$ are estimated from similar regressions as that for hour 1, but they are augmented with the regulating price from hour $t - 1$. Note that, in order to obtain statistically well-specified regression models for all hours, $t = 1, 2, .., 24$, the lag length, k, in the auto-regressive error was allowed to differ across the different hours, taking a value in $\{1, 2, .., 8\}$. In the same vein as for hour 1, the residuals from the regressions for hours $t = 2, .., 24$ could be used to construct similar estimated conditional distributions (histograms). However, this would quickly result in an enormous number of scenarios/outcomes, making the computations excessively cumbersome. For example, by allowing for as little as two outcomes for each hour, there would be 2 to the power of 24 (16,777,216) scenarios. Hence, for tractability, only one value for hours $t = 2, .., 24$ was used, namely the estimated conditional mean. In the following, the different scenario trees, i.e. with different numbers of scenarios, simply correspond to different numbers of outcomes for the regulating price in hour 1 (number of bars in the histogram).

3.2.2. Property Matching

The general idea behind property matching (see Høyland and Wallace[14]) is rather simple: scenario trees $\hat{\xi}$ are created such that only selected statistical properties of the original ξ are matched. Any statistical property can be matched, including extreme values and inter-temporal properties such as regression to the mean or autocorrelations. For example, if ξ has a mean of 5 and a variance of 2.5, a moment matching method could be used to match the first two moments of the distribution to obtain $\hat{\xi}$ with a mean of 5 and a variance of 2.5. However, all other statistical properties are generally different. A particular advantage of Property Matching methods is that in practice they do not require the original probability distribution to be fully specified. In fact, they can handle cases, as the one presented in our computational study, in which a decision maker is only able to estimate statistical properties

7

from historical data, but for which there is no reason to believe that the data belongs to a particular probability distribution.

A common approach, used in our computational study, was proposed by Høyland et al.[15] and consists of matching the first four moments of the marginal distributions and the correlations, where the first moment is the mean, the second is the variance, and the following equation specifies the third ($n = 3$) and fourth ($n = 4$) moment:

$$\frac{1}{\sigma^n} E_\xi [\xi - \mu]^n$$

where σ is the standard deviation and μ is the mean. The method begins by generating discrete univariate random variables matching the desired first and second moment. Then, it transforms them so that the resulting random vector is consistent with the given correlations. Since the transformation distorts the third and fourth marginal moments, the initial third and fourth moment are chosen so the target ones are obtained.

3.2.3. Copula-Based Scenario Generation

The method of Kaut[16] is in the spirit of Høyland and Wallace[14] and is based on the idea of using copulas instead of correlations to describe the co-dependence between the random variables. Copula is the joint cumulative distribution function of any multivariate random vector with uniform margins. However, following Sklar's Theorem (see Sklar[30]), a general multivariate cumulative distribution function is fully determined by the marginal cumulative distributions and by the copula. Thus, while correlations describe only the linear dependence between the random variables, the copula provides a full description of the dependence. In addition, copulas and margins are independent.

The method of Kaut[16], similarly to the method employed by Høyland et al.[15], takes as input the target margins, but it uses copulas instead of correlations. It includes two steps. In the first step, a heuristic generates copula samples. However, since a copula is a multivariate distribution with uniform margins, the scenarios are samples from $[0, 1]$. Therefore, in the second step, the margins are transformed to the target margins using the inverse cumulative distribution functions. Thus, the scenarios obtained have both the correct copula and marginal distributions.

4. Computational study

This section first introduces a Danish case study and then presents the results obtained by testing different scenario generation methods for this case study.

4.1. The Danish Case study

The uncertain parameters in our problem are the regulating prices. For the Danish case, historical realisations of hourly regulating prices for the period January 1, hour 1, 2006, through December 31, hour 24, 2014, were obtained from the Danish TSO, Energinet.dk (see their market data). Since strong evidences that the data belongs to a particular probability distribution were not obtained, agreeing with the discussion in King and Wallace[17], Chapter 4, instead of fitting an arbitrary distribution, the available historical data was considered as our true (large) discrete sample space, with every realisation being equally probable, and

this is referred to as the *reference tree*. In particular, separate seasons were distinguished, as well as weekdays and weekends. Therefore, problem (1) was solved for a specific day, using a tree comprising all the available regulating price realisations for weekdays or weekends of the season where the day belongs. For example, if the problem for Tuesday, August 15, was solved, the reference scenario tree would contain all the available regulating prices for summer weekdays.

[Table 1 about here.]

Table 1 reports the number of scenarios in the reference trees used. For our case, problem (1) was solved for all the reference scenario trees as this is necessary to estimate the bias. However, this is a consequence of using a mathematical model (1) which is not computationally difficult. In more general cases, using scenario trees of the sizes reported in Table 1 might not be possible and scenario generation would be the only way out.

As in Juul et al.[7], a Nissan Leaf model with storage capacity L_j^{MAX} of 24 kWh was assumed as the reference car, and the minimum battery level admissible (L_j^{MIN}) was assumed as 0 kWh, i.e. the battery can be fully utilised. This study applies the same assumptions used in Juul et al.[7] that the maximum charging capacity in each hour is $C^{MAX} = 6.9$ kWh and the charging efficiency is $\eta = 0.9$. Driving demand was obtained from Kristoffersen et al.[4]. The authors analysed private transport activity data from Western Denmark for the period January 2006 to December 2007 and provided a number of driving patterns representing different possible utilisations of EVs. In this study, all 20 patterns were included with each driving pattern representing a different vehicle (set $\mathcal{J}$). Based on the distance travelled each hour and the characteristics of the car used (Nissan Leaf), the corresponding driving loads D_{jt} were computed. An hourly resolution and a planning horizon of 24 hours were used.

Problem (1) was solved for 8 days of the year representing the categories season and weekday/weekend. To select representative days for each season, the spot prices for each of the eight categories described were analysed, and a representative day was defined as a day where the spot price for all hours of this day and the preceding six days were not outliers using the outlier definition by Tukey[31]. Here, an outlier is defined when being outside the range $[Q1 - 1.5 \cdot IQR, Q3 + 1.5 \cdot IQR]$, where $Q1$ and $Q3$ are the first and third quartile and IQR is the interquartile range, i.e. $Q3 - Q1$. The outlier detection were made for each of the eight categories. The six preceding days are the actual preceding days even though this means that they are not in the same category. After this procedure was performed, one representative day for each category was chosen.

4.2. Results

For each target day, scenario trees with 10, 20, 30, 40, 50, and 100 scenarios were generated, using each scenario generation method discussed in Section 3.2. These scenario tree sizes are considerably smaller than the reference tree (see Table 1) and are therefore preferable in most real-life situations from a computational perspective. In figure 1, the average computational time for each size of the scenario trees is shown together with the average computational time for the reference trees for weekend days and weekdays.

[Figure 1 about here.]

9

For each scenario tree generated, the bias was calculated as explained in Section 3. Table 2 reports the bias for the different scenario trees and for the different days.

[Table 2 about here.]

Table 2 shows that the property matching and the copula matching method outperform the discretisation of the empirical distribution obtained via time-series analysis. The property matching method manages to nullify the bias in six of the eight cases when generating 100 scenarios. The copula matching method nullifies the bias for a number of cases already with 40 scenarios, and, except for the winter days, it provides a rather small gap already with 30 scenarios.

To understand the reason for the performance gap between the copula matching and the property matching methods and between these and the discretisation method, it is necessary to look again at how these methods generate scenario trees. The property matching and the copula matching method focus on the first four marginal moments, and the results show that these capture the shape of the distribution rather well. Instead, the regression model captures most of the uncertainty through a linear relationship between the random variables, and the discretisation is applied to the residuals. An important element of difference is in the relationship between the random variables. The regression model describes a linear relationship only between the regulating price at two consecutive hours, while the property matching method considers, through correlation, a more general linear co-dependence between all the random variables. Finally, the copula matching method generalises this relationship even further. In fact, it does not limit the co-dependence between the random variables to a linear relationship, but it captures a more general relationship through copulas. Our results show that, for our case, the co-dependence between the random variables (i.e., the regulating prices) is a crucial element when generating scenario trees.

In addition, the computation based on the regression model always provides the same bias. The reason for this is that the different numbers of outcomes generated are merely based on different resolutions of a histogram based on the same numbers. The only difference is therefore in terms of the resolution fineness of the histogram (a finer resolution corresponds to more outcomes). As long as the outcome number is not very low, e.g. going from 2 to 3 outcomes, the distribution of regulating prices relative to the spot price, which is what matters for the optimisation problem, is the same and has the same bias results.

In order to compare our results with random sampling, we have solved the optimisation problem using random sampling for each of the combinations of days and seasons as in Table 2. In Figure 2, the bias generated by using random sampling is compared to that generated using copula matching. We have tested with 40 scenarios as this is where copula matching starts nullifying the bias. Copula matching shows to perform better than the random sampling for all the combinations of days and seasons.

[Figure 2 about here.]

In fig. 3, the average policies applied for the best and worst performing day for the copula-matching is shown together with the spot price, the forecast of the regulating price, and the realised regulating price.

10

[Figure 3 about here.]

Thus, it is most lucrative for the EVA to plug in the vehicles during morning (7:00 to 10:00) and evening (19:00 to 01:00) during weekdays or afternoon and early evening (13:00-19:00) during weekends. During these time slots, the EVA can provide up-regulation to the grid when it is most needed (and to receive the corresponding compensation). In addition, both charging and up-regulation are scheduled in hours where the forecasted regulating price is higher than the spot price, such that the aggregator earns money on up-regulation in all these hours. Instead, down-regulation is spread out over time periods where the forecasted regulating price is lower than the spot price. This in turn allows the EVA to purchase electricity when the grid needs down regulation, and the corresponding electricity price is low. This particularly benefits a system that heavily relies on (uncontrollable) wind production, such as the Danish system. By means of smart mediators, EVs can thus be seen as a component of the electricity system which helps control stochastic production and not as a threat to the stability of the system.

5. Conclusion

EVs constitute an important element of current mobility systems and will certainly be even more important in the future. However, a coordination of their charging activities is necessary in order not to jeopardise the security of electricity systems. Furthermore, the storage capacity of EVs allows them to provide ancillary services to electricity markets. However, in the configuration of most current electricity markets, the participation of EV owners requires an intermediate player, the EVA, which operates on behalf of a large number of EV owners.

Several studies have investigated how the decisions of an EVA can be facilitated. A considerable number of these use stochastic programming to manage the uncertainty which is inherent, since the charging strategy has to be decided before regulating prices are known. In this paper, we have looked at the pre-processing phase for stochastic programs known as scenario generation. The aim of scenario generation is that of providing a description of the uncertainty which is simple enough to keep the stochastic program tractable while still adequately representing the true probability distribution of the characterising uncertainty. Three methods were compared to model the uncertainty: 1) estimation of the scenario tree, which is based on time-series analysis of historical data; 2) property matching, which generates scenarios matching only certain given statistical properties of the true distribution (e.g. marginal moments and correlations); and 3) copula matching, which, in addition to margins, describes the co-dependence between the random variables using the copula function.

Our results show that the co-dependence between the (uncertain) regulating prices at different hours is an important property of the uncertainty to capture in scenario trees. In particular, the results suggest that it is preferable to use copulas to model this co-dependence, while the linear dependence modelled by correlation coefficients or by the linear terms of a regression model appear insufficient. These findings might help practitioners selecting and fine-tuning their scenario generation methods.

In turn, better scenario trees allow the decision maker to make better charging decisions. It emerges in fact that the best policies consist of charging during night when the system

11

needs down-regulation and to provide up-regulation early in the morning and in the evening. This is particularly beneficial for systems reliant on uncontrollable sources, as it helps to mitigate peaks in demand and in production. However, it remains to assess whether these strategies are conflicting with best-practices in terms of battery utilisation. Taking these issue into account might require additional constraints on the optimisation problem and possibly amend the pattern highlighted in hour results.

Acknowledgements

References

[1] Makena Coffman, Paul Bernstein, and Sherilyn Wee. Electric vehicles revisited : a review of factors that affect adoption. *Transport Reviews*, 37(1):1–15, jan 2016. ISSN 0144-1647. doi: 10.1080/01441647.2016.1217282. URL https://www.tandfonline.com/doi/full/10.1080/01441647.2016.1217282.

[2] Björn Nykvist and Måns Nilsson. Rapidly falling costs of battery packs for electric vehicles. *Nature Climate Change*, 5(4):329–332, mar 2015. ISSN 1758-678X. doi: 10.1038/nclimate2564. URL http://www.nature.com/doifinder/10.1038/nclimate2564.

[3] Alec Brooks, Ed Lu, Dan Reicher, Charles Spirakis, and Bill Weihl. Demand dispatch. *IEEE Power and Energy Magazine*, 8(3):20–29, 2010.

[4] Trine Krogh Kristoffersen, Karsten Capion, and Peter Meibom. Optimal charging of electric drive vehicles in a market environment. *Applied Energy*, 88(5):1940–1948, may 2011. ISSN 03062619. doi: 10.1016/j.apenergy.2010.12.015. URL http://linkinghub.elsevier.com/retrieve/pii/S0306261910005222.

[5] Stylianos I. Vagropoulos and Anastasios G. Bakirtzis. Optimal Bidding Strategy for Electric Vehicle Aggregators in Electricity Markets. *IEEE Transactions on Power Systems*, 28(4):4031–4041, nov 2013. ISSN 0885-8950. doi: 10.1109/TPWRS.2013.2274673. URL http://ieeexplore.ieee.org/document/6582687/.

[6] R.J. Bessa and M.A. Matos. Optimization models for an EV aggregator selling secondary reserve in the electricity market. *Electric Power Systems Research*, 106:36–50, jan 2014. ISSN 03787796. doi: 10.1016/j.epsr.2013.08.006. URL http://linkinghub.elsevier.com/retrieve/pii/S0378779613002162.

[7] N. Juul, G. Pantuso, J. E. Iversen, and T. Boomsma. Strategies for Charging Electric Vehicles in the Electricity Market. *International Journal of Sustainable Energy Planning and Management*, 7:67–74, 2016.

[8] Ahmad Tavakoli, Michael Negnevitsky, Duy Thanh Nguyen, and Kashem M. Muttaqi. Energy exchange between electric vehicle load and wind generating utilities. *IEEE Transactions on Power Systems*, 31(2):1248–1258, 2016.

[9] Stylianos I. Vagropoulos, Georgios A. Balaskas, and Anastasios G. Bakirtzis. An investigation of plug-in electric vehicle charging impact on power systems scheduling and energy costs. *IEEE Transactions on Power Systems*, 32(3):1902–1912, 2017.

[10] Ali T. Al-Awami and Eric Sortomme. Coordinating Vehicle-to-Grid Services With Energy Trading. *IEEE Transactions on Smart Grid*, 3(1):453–462, mar 2012. ISSN 1949-3053. doi: 10.1109/TSG.2011.2167992. URL http://ieeexplore.ieee.org/document/6075307/.

[11] M. W. Khalid, Ali T. Al -Awami, and Eric Sortomme. Stochastic-programming-based bidding strategy for V2G services. In *IEEE PES ISGT Europe 2013*, pages 1–5. IEEE, oct 2013. ISBN 978-1-4799-2984-9. doi: 10.1109/ISGTEurope.2013.6695468. URL http://ieeexplore.ieee.org/document/6695468/.

[12] Jitka Dupacová, Giorgio Consigli, and Stein W. Wallace. Scenarios for Multistage Stochastic Programs. *Annals of Operations Research*, 100(1-4):25–53, 2000.

[13] G. Ch. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming*, 89(2):251–271, 2001.

[14] K. Høyland and S. W. Wallace. Generating Scenario Trees for Multistage Decision Problems. *Management Science*, 47(2):295–307, 2001.

[15] K. Høyland, M. Kaut, and S. W. Wallace. A Heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24:169–185, 2003.

[16] Michal Kaut. A copula-based heuristic for scenario generation. *Computational Management Science*, 11(4):503–516, oct 2014. ISSN 1619-697X. doi: 10.1007/s10287-013-0184-4. URL http://link.springer.com/10.1007/s10287-013-0184-4.

[17] A. J. King and S. W. Wallace. *Modeling with Stochastic Programming*. Springer New York, New York, 2012.

[18] M. Kaut and S. W. Wallace. Evaluation of Scenario-Generation Methods for Stochastic Programming. *Pacific Journal of Optimization*, 3:257–271, 2007.

[19] Pierre Pinson, Henrik Madsen, Henrik Aa Nielsen, George Papaefthymiou, and Bernd Klöckl. From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy*, 12(1):51–62, 2009. ISSN 10954244. doi: 10.1002/we.284.

[20] Tao Wang, Hsiao-dong Chiang, and Ryuya Tanabe. Toward a Flexible Scenario Generation Tool for Stochastic Renewable Energy Analysis. *Power Systems Computation Conference (PSCC) 2016*, 2016.

[21] Yi Wang, Ning Zhang, Qixin Chen, Daniel S. Kirschen, Pan Li, and Qing Xia. Data-Driven Probabilistic Net Load Forecasting With High Penetration of Behind-the-Meter PV. *IEEE Transactions on Power Systems*, 33(3):3255–3264, 2018. doi: 10.1109/TPWRS.2017.2762599.

[22] Jean-François Toubeau, Jérémie Bottieau, François Vallée, and Zacharie De Grève. Deep Learning - based Multivariate Probabilistic Forecasting for Short - Term Scheduling in Power Markets. *IEEE Transactions on Power Systems*, pages 1–12, 2018. doi: 10.1109/TPWRS.2018.2870041.

[23] Duehee Lee and Ross Baldick. Load and Wind Power Scenario Generation Through the Generalized Dynamic Factor Model. *IEEE Transactions on Power Systems*, 32(1): 400–410, 2017. ISSN 08858950. doi: 10.1109/TPWRS.2016.2562718.

[24] Yize Chen, Yishen Wang, Daniel Kirschen, and Baosen Zhang. Model-Free Renewable Scenario Generation Using Generative Adversarial Networks. *IEEE Transactions on Power Systems*, 3:3265–3275, 2018. ISSN 0885-8950. doi: 10.1109/TPWRS.2018.2794541. URL http://arxiv.org/abs/1707.09676.

[25] Stylianos I. Vagropoulos, Evaggelos G. Kardakos, Christos K. Simoglou, Anastasios G. Bakirtzis, and João P S Catalão. ANN-based scenario generation methodology for stochastic variables of electric power systems. *Electric Power Systems Research*, 134: 9–18, 2016. ISSN 03787796. doi: 10.1016/j.epsr.2015.12.020. URL http://dx.doi.org/10.1016/j.epsr.2015.12.020.

[26] Wai Kei Mak, David P. Morton, and R. Kevin Wood. Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, 24(1):47–56, 1999. ISSN 01676377. doi: 10.1016/S0167-6377(98)00054-6. URL http://www.sciencedirect.com/science/article/pii/S0167637798000546.

[27] J. Dupacová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming. *Mathematical Programming*, 95(3):493–511, 2003.

[28] Holger Heitsch and Werner Römisch. Scenario reduction algorithms in stochastic programming. *Computational optimization and applications*, 24(2-3):187–206, 2003.

[29] Juan M. Morales, Salvador Pineda, Antonio J. Conejo, and Miguel Carrion. Scenario reduction for futures market trading in electricity markets. *IEEE Transactions on Power Systems*, 24(2):878–888, 2009.

[30] Abe Sklar. Random variables, distribution functions, and copulas: a personal look backward and forward. *Lecture Notes–Monograph Series*, 28(May):1–14, 1996. doi: 10.1214/lnms/1215452606. URL http://www.jstor.org/stable/4355880.

[31] J. W. Tukey. *Exploratory Data Analysis*. Addison-Wesley series in behavioral science. Addison-Wesley Publishing Company, 1977. ISBN 9780201076165. URL https://books.google.dk/books?id=UT9dAAAAIAAJ.
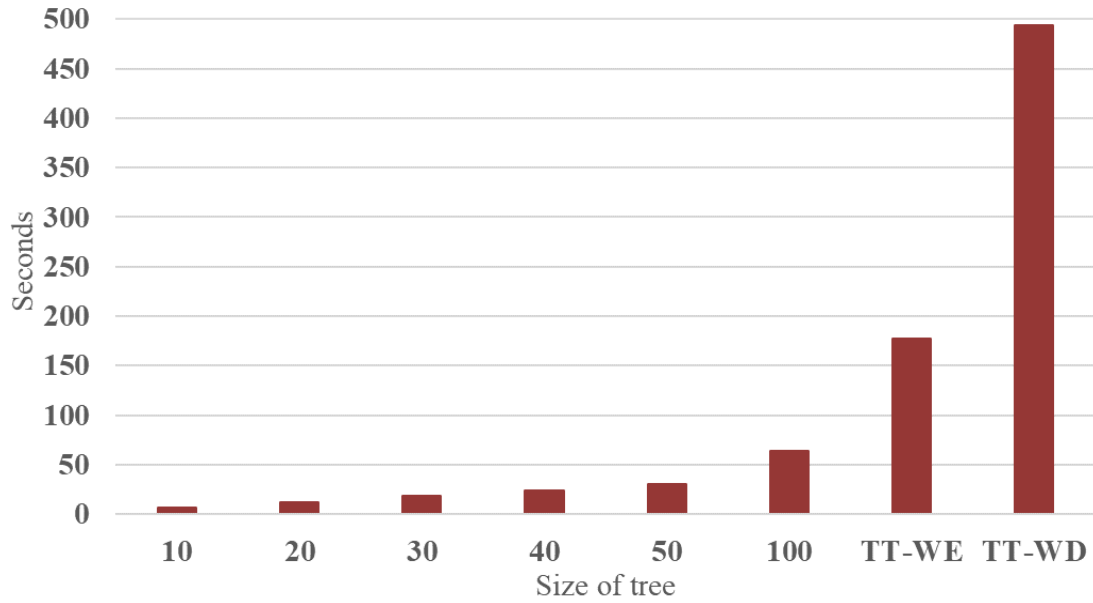
Figures

Figure 1: Average computational time of scenario trees and reference trees. The average size of the reference trees is 235 for weekend days (TT-WE) and 586 for weekdays (TT-WD).
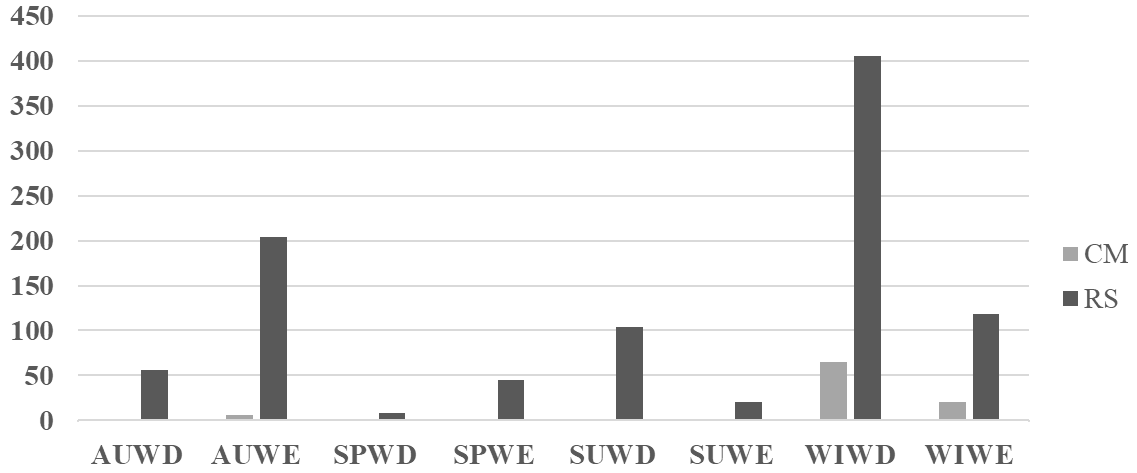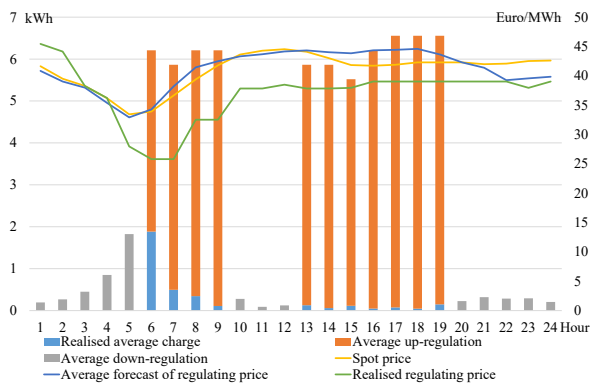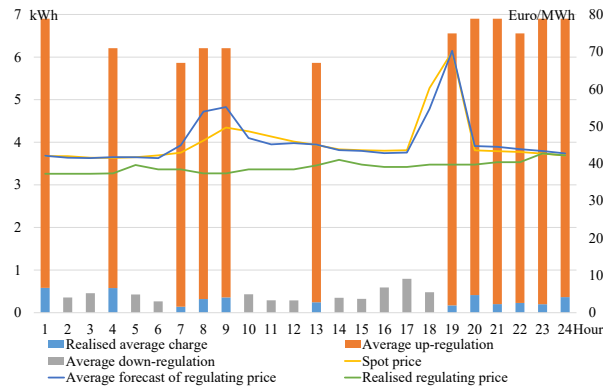
Figure 2: The bias of random sampling (RS) compared to copula matching (CM) for all days tested using the scenario trees with 40 scenarios. Abbreviations: AU = Autumn, SP = Spring, SU = Summer, WI = Winter, WE = Weekend, WD = Weekday.

(a) Summer weekend-day, SUWE

(b) Winter weekday, WIWD

Figure 3: Policies for charging of EVs for the best and worst day with copula-matching

Tables

Season	Day Type	Reference Tree Size
Winter	Weekday	581
Winter	Weekend	231
Spring	Weekday	591
Spring	Weekend	237
Summer	Weekday	591
Summer	Weekend	237
Autumn	Weekday	585
Autumn	Weekend	234

Table 1: Number of scenarios in the reference trees

Method	Scenarios	AUWD	AUWE	SPWD	SPWE	SUWD	SUWE	WIWD	WIWE	Average
CM	10	99.89	53.90	4.14	9.77	31.07	3.02	201.01	643.74	130.82
	20	70.72	19.47	1.44	9.77	11.09	5.79	162.56	433.10	89.24
	30	7.69	6.40	1.69	9.77	5.54	1.36	101.84	118.40	31.59
	40	0.00	6.40	0.00	0.00	0.00	0.00	65.45	20.11	11.50
	50	0.00	0.00	0.00	13.54	0.00	0.00	100.95	2.30	14.60
	100	0.00	0.00	0.00	0.00	0.00	0.00	64.54	2.20	8.34
PM	10	19.75	168.66	31.23	61.10	71.80	35.27	124.51	460.52	121.61
	20	107.57	56.35	31.58	60.02	58.13	4.90	23.74	187.33	66.20
	30	223.61	0.00	23.61	10.30	49.55	2.43	6.25	0.00	39.47
	40	9.01	13.90	29.01	24.61	19.47	0.00	7.96	158.33	32.79
	50	15.93	0.00	8.11	43.18	15.70	5.92	6.25	77.97	21.63
	100	0.00	0.00	0.00	37.34	0.00	0.00	6.25	0.00	5.45
TS	10	324.06	450.04	66.57	316.85	202.01	87.58	192.66	1534.68	396.81
	20	324.06	450.04	66.57	316.85	202.01	87.58	192.66	1534.68	396.81
	30	324.06	450.04	66.57	316.85	202.01	87.58	192.66	1534.68	396.81
	40	324.06	450.04	66.57	316.85	202.01	87.58	192.66	1534.68	396.81
	50	324.06	450.04	66.57	316.85	202.01	87.58	192.66	1534.68	396.81
	100	324.06	450.04	66.57	316.85	202.01	87.58	192.66	1534.68	396.81

Table 2: Bias results for different days and scenario generation methods and number of scenarios. The bias is expressed as the % increase from the optimal value obtained using the reference tree. Scenario generation methods are indicated as CM (copula matching), PM (property matching) and TS (time-series analysis). Abbreviations: AU = Autumn, SP = Spring, SU = Summer, WI = Winter, WE = Weekend, WD = Weekday.