Delft University of Technology

Efficient Assessment of Electricity Distribution Network Adequacy with the Cross-Entropy Method

Betge, J.N.; Droste, Barbera; Heres, Jacco; Tindemans, S.H.

Citation (APA)
Betge, J. N., Droste, B., Heres, J., & Tindemans, S. H. (2021). Efficient Assessment of Electricity Distribution Network Adequacy with the Cross-Entropy Method. In *2021 IEEE Madrid PowerTech, PowerTech 2021 - Conference Proceedings* Article 9494891 (2021 IEEE Madrid PowerTech, PowerTech 2021 - Conference Proceedings). https://doi.org/10.1109/PowerTech46648.2021.9494891

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Efficient Assessment of Electricity Distribution Network Adequacy with the Cross-Entropy Method

Julian N. Betge*
Dept. of Electrical Sustainable Energy
Delft University of Technology
Delft, The Netherlands
julian.betge@gmail.com

Barbera Droste
IT Data & Analytics
Alliander N.V.
Arnhem, The Netherlands
barbera.droste@alliander.com

Jacco Heres
IT Data & Analytics
Alliander N.V.
Arnhem, The Netherlands
jacco.heres@alliander.com

Simon H. Tindemans
Dept. of Electrical Sustainable Energy
Delft University of Technology
Delft, The Netherlands
s.h.tindemans@tudelft.nl

Abstract—**Identifying future congestion points in electricity distribution networks is an important challenge distribution system operators face. A proven approach for addressing this challenge is to assess distribution grid adequacy using probabilistic models of future demand. However, computational cost can become a severe challenge when evaluating large probabilistic electricity demand forecasting models with long forecasting horizons. In this paper, Monte Carlo methods are developed to increase the computational efficiency of obtaining asset overload probabilities from a bottom-up stochastic demand model. Cross-entropy optimised importance sampling is contrasted with conventional Monte Carlo sampling. Benchmark results of the proposed methods suggest that the importance sampling-based methods introduced in this work are suitable for estimating rare overload probabilities for assets with a small number of customers.**

Index Terms—**cross-entropy method, importance sampling, adequacy assessment, distribution networks, demand modelling**

I. INTRODUCTION

The energy transition has a considerable impact on electricity distribution networks, which increasingly accommodate distributed energy resources, such as renewable generation or storage. The challenge lies not only in the unpredictable nature of renewable generation, but also in the conventional design of distribution networks which facilitates unidirectional flows from the transmission level to end consumers. To enable adequate distribution network planning under these changing circumstances, new approaches and tools are needed. The review of suitable new approaches given in [1] underlines that these should be probabilistic, model risk explicitly, consider load and generation time-series, employ a multi-objective optimisation framework and incorporate a mix between network and no-network solutions. The merits of a novel planning

tool realising these requirements are demonstrated e.g. in [2], where the benefits of integrating distributed storage over traditional grid reinforcement are assessed.

A specific challenge faced by Distribution System Operators (DSOs), for which new planning approaches are needed, is to identify future congestion points in the network. This problem is a type of system adequacy assessment, where a probabilistic model of future demand is compared with the physical network capacity. Future demand may be modelled in a bottom-up manner, reflecting the uncertainty arising from various high-level technology diffusion scenarios [3].

Focusing on this challenge and in response to the need for new network planning tools, Alliander is developing the Advanced Net DEcision Support (ANDES) model for its subsidiary, the Dutch DSO Liander. In its current form, the load model makes use of measured load data from monitored customers, but the demand of unmonitored customers is represented by average category profiles. The use of averaged profiles is problematic, as these are smoother and less stochastic, leading to the underestimation of peaks and troughs. An extension to the ANDES model which randomly assigns measured smart meter profiles to unmonitored customers was proposed in [4]. This was shown to improve the prediction of peak loads, but the Monte Carlo (MC) simulation approach came at a drastically increased computational cost, preventing its adoption in the regular model.

In the context of MC assessment of system adequacy, variance reduction techniques can be employed to speed up the convergence of risk metric estimates. Among these, Importance Sampling (IS) has a particularly large variance reduction potential [5], however it is necessary to identify a suitable biasing distribution. To accomplish the latter in an automatic manner, the Kullback–Leibler divergence is used in the Cross-Entropy (CE) method. The approach has been successfully applied in power system reliability in previous studies, e.g. [6], [7].

This paper aims to contribute to the further development

of the ANDES model by showing the potential of applying the CE method in the given context. To this end, firstly an electricity demand model is specified, employing a similar MC simulation approach as in [4]. Secondly, a parameterisation of the demand model is proposed which is suitable for applying the CE method. Substantial speedups in the computation of overload probabilities can be obtained with the considered CE approach, but its effectiveness depends on the magnitude of the quantity being estimated.

II. NETWORK ASSET CONGESTION MODEL

A metric of fundamental importance in capacity planning is the network asset overload probability which indicates where congestion problems are likely to arise. A variety of assets could be considered for such analysis, but in this work we focus on (transformers in) substations. The risk metric of interest here is thus the probability of the stochastic power demand D (for a given asset) to exceed the power rating of an asset d_{cap}. For the remediation of congestion issues, it matters whether congestion occurs because the energy demanded by customers cannot be transported to them (risk of positive overload) or the energy generated by customers with PV systems cannot be transported away from them (risk of negative overload). The probabilities of both types of overloads are represented by the risk metrics r_+ and r_-:

$$r_+ = \mathbb{E}_D[\mathbb{1}_{D > d_{cap}}], \tag{1}$$

$$r_- = \mathbb{E}_D[\mathbb{1}_{D < -d_{cap}}], \tag{2}$$

where $\mathbb{1}$ is the indicator function and the expectation is taken over all realisations of the power demand D for the network asset of interest.

A. Relation to the ANDES Model

The ANDES model relies on a bottom-up approach to obtain time-resolved asset load profiles by aggregating load profiles of individual customers. For different diffusion scenarios of key low-carbon technologies, the approach allows to predict which assets might be overloaded in the future. The model covers the entire Alliander power grid, has a time horizon of up to 40 years with quarter-hourly resolution and currently comprises five future scenarios which entails a considerable computational effort and large volumes of output data [8].

An essential aspect of the modelling approach presented here is to introduce stochasticity in the electricity demand behaviour of certain customers in order to obtain a probabilistic range of possible demand behaviours. In this aspect, the modelling approach differs substantially from the ANDES model which currently makes deterministic predictions. A closer description of the ANDES model can be found in [8]. An in-depth comparison of the model presented here, the ANDES model and the model proposed by Valckx [4] is available in [9].

B. Bottom-up Demand Model

The electrical load on an asset is computed as the sum of individual demand profiles assigned to the customers connected to the asset. Only active power is considered, and losses are neglected. Each demand profile is a time series containing quarter-hourly power consumption averages, and is based on measurements from the year 2018, totalling 35,040 time steps for the entire year. The assignment of demand profiles to customers is based on the categorisation of customers used in the ANDES model, a detailed break-down of which can be found in [9]. In the following, the grouping relevant for the modelling approach adopted here is presented. Three groups of demand profile types can be distinguished therein:

1) Smart meter profiles: A set of smart meter profiles of monitored customers of the Liander grid is used as the basis for the modelling of most categories of small unmonitored customers (*Group 1*). Individually measured smart meter profiles were anonymised, but the energy-behavioural category that each profile belongs to is known. Each category contains at least 50 profiles to guarantee anonymisation. The data is owned by the DSO Liander. During model construction, filtering steps were carried out to address data quality issues, reducing the raw set of 3,773 smart meter profiles to 3,137 profiles.

Next, each energy-behavioural category was subdivided by binning the smart meter profiles and the modelled customers of each category based on their total yearly consumption γ. A target number of 50-100 smart meter profiles per bin was deemed appropriate leading to the formation of two to four bins per category. Quantile binning was used, so that the number of smart meter profiles per bin was approximately equal within each category. The customers to be modelled were likewise quantile binned according to their yearly consumption, forming the same number of bins per category as for the smart meter profiles. Customers are randomly assigned a profile from their associated category bin, scaled to the customer's measured yearly consumption.

While the energy-behavioural categories are those used in the ANDES model, the binning scheme is a new aspect introduced in this work. Overall, group 1 comprises 13 energy-behavioural categories which are split in 2-4 bins each, leading to a total of 48 bins.

2) Telemetry profiles: Group 2 comprises several thousand individual larger, commercial customers whose power consumption was measured telemetrically. These telemetric measurements are directly used to model the respective customers.

3) Average category profiles: Group 3 contains 30 different types of profile categories: two energy-behavioural categories for which no smart meter data is available (otherwise they would be placed in group 1); 20 categories based on a classification of economic activities in which large unmonitored customers fall; and eight categories based on a classification of grid connection ratings, used for the few not otherwise categorised customers. The common trait of this diverse group is that all customers in it are modelled by scaled average category profiles derived from the described categorisations. The share of this group in overall consumption is rather low, therefore the use of average category profiles is deemed acceptable.

C. Probabilistic Description of the Demand Model

We proceed to formalise the demand model described above. In the following, random variables are capitalised and vectors are set in bold. The stochastic power demand on a given network asset for all quarter-hourly time steps of a year is obtained by summing the demand profiles of all customers connected to the asset:

$$\boldsymbol{D}(\boldsymbol{\Pi}) = \sum_{i=1}^{n_s} \gamma_i \cdot \boldsymbol{s}_{b_i, \Pi_i} + \sum_{p=1}^{n_l} \boldsymbol{l}_p + \sum_{q=1}^{n_a} \gamma_q \cdot \boldsymbol{a}_{c_q}, \qquad (3)$$

where

- $\boldsymbol{D}(\boldsymbol{\Pi})$ is a resulting random annual asset demand trace or time series
$$\boldsymbol{D}(\boldsymbol{\Pi}) = \{D_t(\boldsymbol{\Pi})\} = \{D_1(\boldsymbol{\Pi}), \ldots, D_{35040}(\boldsymbol{\Pi})\},$$
- $\boldsymbol{\Pi}$ is a random vector denoting a random selection of smart meter profiles,
- $\boldsymbol{s}_{b_i, \Pi_i}$ is a normalised smart meter profile randomly drawn from bin b_i of customer i,
- $\boldsymbol{l}_p$ is the telemetry profile of customer p,
- $\boldsymbol{a}_{c_q}$ is the average category profile of category c_q of customer q,
- γ_i and γ_q are the total yearly energy consumption of customers i and q, respectively and
- n_s, n_l and n_a denote the total number of customers in groups 1-3 of the modelled asset, respectively.

The random profile selection vector

$$\boldsymbol{\Pi} = (\Pi_1, \Pi_2, \ldots, \Pi_i, \ldots, \Pi_{n_s}) \qquad (4)$$

has length n_s and indexes the smart meter profiles $\boldsymbol{s}_{b_i, \Pi_i}$, randomly drawn from bin b_i of each small unmonitored customer i connected to the asset. The elements of $\boldsymbol{\Pi}$ are the uniform discrete random variables Π_i, which have the sample space $\Omega_{\Pi_i} = \{1, \ldots, n_{b_i}\}$, with n_{b_i} being the number of smart meter profiles in bin b_i that customer i belongs to. The random selection of smart meter profiles for all group 1 customers gives rise to the stochastic properties of the demand model.

III. Monte Carlo Risk Estimation

Having defined a probabilistic demand model and the risk metrics r_+ (1) and r_- (2), this section details how various MC methods can be used to estimate these risks. In the following, only the relations for the metric r_+ for positive overloads are shown and the subscript '+' is dropped for brevity. The relations for the metric r_- follow by analogy.

A. Conventional Monte Carlo Method

The sample space of the demand model, which contains all possible demand states that can be assumed, has two dimensions: the assignment of smart meter profiles to customers and the time at which the selected profiles are evaluated. There is a natural hierarchy between these dimensions: the selection of random profiles occurs first and carries a larger overhead than the selection of time steps. Moreover, the profile selection space is vastly larger than the set of available time steps.

Together, these considerations lead to a hierarchical MC estimation scheme. The 'overload fraction' of each randomly selected profile vector $\boldsymbol{\pi}_j$ and random sample $\boldsymbol{\theta}_j$ of m time steps, $\{\theta_{j,1}, \ldots, \theta_{j,t}, \ldots, \theta_{j,m}\}$ with $\theta_{j,t} \in \{1, \ldots, 35040\}$, is quantified using the impact function $H(\cdot)$ defined as:

$$H(\boldsymbol{\pi}_j, \boldsymbol{\theta}_j) = \frac{1}{m} \sum_{t=1}^{m} \mathbb{1}_{D_{\theta_{j,t}}(\boldsymbol{\pi}_j) > d_{cap}}. \qquad (5)$$

The MC estimator for the overload probability is then the mean

$$\hat{r}_{MC} = \frac{1}{n} \sum_{j=1}^{n} H(\boldsymbol{\pi}_j, \boldsymbol{\theta}_j). \qquad (6)$$

taken over the impact values of n randomly selected profile vectors and time step samples.

As a special case, we define the *reference method*, which computes entire annual asset demand traces with $m = 35{,}040$ quarter-hourly time steps, i.e.

$$\hat{r}_{ref} = \frac{1}{n} \sum_{j=1}^{n} H(\boldsymbol{\pi}_j, \boldsymbol{\theta}^*) \qquad (7)$$

$$\boldsymbol{\theta}^* = \{1, \ldots, 35{,}040\}. \qquad (8)$$

We expect the reference method to be less efficient than the generic MC method (6), because more time is spent on analysing highly dependent samples in sequential time slots.

In both cases, the coefficient of variation or relative error of the estimator can be estimated (cf. [5]) using

$$\beta_{\hat{r}} = \frac{\hat{\sigma}_H}{\hat{r}\sqrt{n}}, \qquad (9)$$

where n denotes the number of profile selections sampled and $\hat{\sigma}_H$ is the sample standard deviation of the samples $H(\boldsymbol{\pi}_j, \boldsymbol{\theta}_j)$.

B. Importance Sampling of Profile Selections

Variance reduction techniques can lead to substantial increases in the efficiency of estimating a quantity of interest. Among these, IS is the most fundamental technique [5] and lends itself to application to the problem at hand. The core idea hereby is to sample the most spiky smart meter profiles within each bin with higher probability in order to increase the frequency of overload events.

Before being able to apply IS to the selection of profiles, the given problem needs to be parameterised appropriately. The parameterisation chosen is based on the idea to divide the smart meter profiles of each bin used in the demand model in two sets – a set of very spiky profiles and a set of profiles with more average characteristics.

The following approach was used to identify the most spiky profiles. Let $\boldsymbol{s}_b = \{s_{b,t}\}$ be a smart meter profile from a given bin b and $\tilde{\boldsymbol{s}}_b = \{\tilde{s}_{b,t}\}$ the median profile of that bin, obtained by computing the median power consumption of all profiles in the bin per time step. For brevity, the bin indices b are dropped below. The metrics Δ_+ and Δ_- have been calculated for each profile in each bin by evaluating

$$\Delta_+(\boldsymbol{s}) = \sum_{t=1}^{35{,}040} h(t), \quad h(t) = \begin{cases} (s_t - \tilde{s}_t)^2 & s_t > \tilde{s}_t \\ 0 & \text{otherwise}, \end{cases} \qquad (10)$$

and

$$\Delta_-(\boldsymbol{s}) = \sum_{t=1}^{35{,}040} h(t), \ h(t) = \begin{cases} (s_t - \tilde{s}_t)^2 & s_t < \tilde{s}_t \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

Using the threshold parameter q_{spiky}, a profile $\boldsymbol{s}$ was assigned to the set of spiky profiles for positive overloads whenever $\Delta_+(\boldsymbol{s}) \geq Q_{\Delta_+}(p = q_{spiky})$, where $Q_{\Delta_+}(p)$ is the quantile function of the median deviation metric Δ_+ which returns the estimated sample p-quantile. The two sets formed in this way for each bin b are denoted by $\mathcal{S}_{b,spiky}$ and $\mathcal{S}_{b,smooth}$. Analogously, the decision rule for negative overloads was to assign a profile $\boldsymbol{s}$ to the spiky set whenever $\Delta_-(\boldsymbol{s}) \geq Q_{\Delta_-}(p = q_{spiky})$.

The random assignment of profiles from a bin to each small, unmonitored customer (group 1) can now be parameterised in two steps. First, the customer is associated to the 'spiky' set with probability u_i and to the 'smooth' set otherwise, where the spiky set assignment probabilities u_i are determined according to

$$u_i = \frac{|\mathcal{S}_{b_i,spiky}|}{|\mathcal{S}_{b_i,spiky}| + |\mathcal{S}_{b_i,smooth}|}. \quad (12)$$

Then, a profile is randomly chosen with uniform probability from the appropriate set ($\mathcal{S}_{b_i,spiky}$ or $\mathcal{S}_{b_i,smooth}$) of the bin b_i to which the customer belongs.

The overall sampling distribution f of a modelled asset, from which random profile selections $\boldsymbol{\Pi}$ are drawn, can therefore be described as a chain of n_s Bernoulli trials

$$f(\boldsymbol{x}; \boldsymbol{u}) = \prod_{i=1}^{n_s} (u_i)^{x_i} \cdot (1 - u_i)^{1-x_i}, \quad (13)$$

parametrised with the spiky set assignment probabilities $\boldsymbol{u} = (u_1, \ldots, u_i, \ldots, u_{n_s})$ of the n_s customers modelled by a random smart meter profile (group 1). The relation determines the probability $f(\boldsymbol{x}; \boldsymbol{u})$ of encountering a given set assignment $\boldsymbol{x} = (x_1, \ldots, x_i, \ldots, x_{n_s})$ of the n_s customers, with $x_i \in \{0, 1\}$ and 1 representing the spiky set.

Parameterising the demand model in the described way enables IS using a modified vector of spiky set assignment probabilities $\boldsymbol{v}$, where the probabilities of sampling spiky profiles differ from the original probabilities $\boldsymbol{u}$. To correct the purposefully introduced bias due to sampling according to the spiky set probabilities $\boldsymbol{v}$, importance weights are used:

$$W(\boldsymbol{x}; \boldsymbol{u}, \boldsymbol{v}) = \frac{f(\boldsymbol{x}; \boldsymbol{u})}{g(\boldsymbol{x}; \boldsymbol{v})} = \frac{\prod_{i=1}^{n_s} (u_i)^{x_i} \cdot (1 - u_i)^{1-x_i}}{\prod_{i=1}^{n_s} (v_i)^{x_i} \cdot (1 - v_i)^{1-x_i}}, \quad (14)$$

where $g(\boldsymbol{x}; \boldsymbol{v})$ is the IS sampling distribution. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ be an independent and identically distributed (i.i.d.) sample of set assignment vectors from $g(\cdot; \boldsymbol{v})$ with the profile selections $\boldsymbol{\pi}_1(\boldsymbol{x}_1), \ldots, \boldsymbol{\pi}_n(\boldsymbol{x}_n)$ sampled from the appropriate sets according to these vectors. Furthermore, let $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ be a set of i.i.d. samples of m time steps, containing a sample of time steps for each of the n profile allocations. The IS estimator is then:

$$\hat{r}_{IS} = \frac{1}{n} \sum_{j=1}^{n} H(\boldsymbol{\pi}_j(\boldsymbol{x}_j), \boldsymbol{\theta}_j) \cdot W(\boldsymbol{x}_j; \boldsymbol{u}, \boldsymbol{v}). \quad (15)$$

The relative error is again estimated according to (9), but using the sample standard deviation of the product $H(\cdot)W(\cdot)$.

C. Application of the Cross-Entropy Method

Although (15) yields a valid estimator for any (non-degenerate) choice of $\boldsymbol{v}$, its performance can vary greatly. To target overload events more precisely and thus achieve better variance reduction, the probability to sample spiky profiles can be automatically increased for those customers only which are crucially involved in causing overloads using the CE method. This minimises the Kullback-Leibler divergence (a distance measure) between the IS distribution $g(\cdot; \boldsymbol{v})$ for a given asset and the theoretically optimal (but unknown) IS distribution g^*. If the distribution whose parameters are being CE-optimised belongs to an exponential family, an analytical solution to the optimisation problem can be obtained, which is the case for the Bernoulli distribution used here. The analytical solution is then estimated using a sampling-based formula – the CE method thus combines optimisation and estimation [5].

When using the CE method in the context of rare event sampling, as done here, finding the near-optimal IS distribution is often a rare event estimation problem itself and would require large sample sizes. To circumvent this difficulty, a sequence of intermediate IS distributions can be estimated which gradually approach the near-optimal IS distribution. In this section, a sequential CE algorithm for the estimation of overload probabilities of a distribution network asset is described, adapted from [6]. For a more a general discussion of the CE method, the interested reader is referred to [5], [10].

1) Set the sample size for an iteration of the CE optimisation n_{opt}, the multilevel parameter ρ and the smoothing parameter α to appropriate values. Set the maximum sample size n_{max}, the no-overload sample size $n_{max,zero}$ and the target relative error β_{target} used as stopping criteria. Choose a value for the threshold parameter q_{spiky} and assign the smart meter profiles in each bin to the 'spiky' and 'smooth' sets according to the approach described in section III-B.

2) Initialise the IS distribution parameter vector $\boldsymbol{v}$ to be estimated as $\hat{\boldsymbol{v}}_0 = \boldsymbol{u}$ to start with conventional MC sampling. Set the iteration counter of the CE optimisation process to $k = 1$. Set the initial optimisation threshold to half of the asset's rated power capacity: $d_{opt} = 0.5 \cdot d_{cap}$.

3) Generate an i.i.d. sample of set assignment vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_{opt}}$ which follow $g(\cdot; \hat{\boldsymbol{v}}_{k-1})$. Then sample profile selections $\boldsymbol{\pi}_1(\boldsymbol{x}_1), \ldots, \boldsymbol{\pi}_{n_{opt}}(\boldsymbol{x}_{n_{opt}})$ accordingly. Also, generate a set of i.i.d. samples of m time steps, $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{n_{opt}}$. Obtain the maximum asset load for each of the n_{opt} profile selection samples and store these in the vector $\boldsymbol{l}_{max}$.

4) Update d_{opt} as the $(1 - \rho)$-quantile of $\boldsymbol{l}_{max}$, thus $d_{opt} = Q_{\boldsymbol{l}_{max}}(1 - \rho)$.

5) To update the probability vector for sampling spiky

profiles, evaluate for each customer i the equation

$$\hat{v}'_{k,i} = \frac{\sum_{j=1}^{n_{opt}} \tilde{H}(\boldsymbol{\pi}_j(\boldsymbol{x}_j), \boldsymbol{\theta}_j) \cdot W(\boldsymbol{x}_j; \boldsymbol{u}, \hat{\boldsymbol{v}}_{k-1}) \cdot x_{j,i}}{\sum_{j=1}^{n_{opt}} \tilde{H}(\boldsymbol{\pi}_j(\boldsymbol{x}_j), \boldsymbol{\theta}_j) \cdot W(\boldsymbol{x}_j; \boldsymbol{u}, \hat{\boldsymbol{v}}_{k-1})}, \tag{16}$$

where $\tilde{H}(\cdot)$ uses d_{opt} instead of d_{cap} as the capacity threshold to determine overload events. Subsequently, carry out smoothed updating by evaluating $\hat{\boldsymbol{v}}_k = \alpha \cdot \hat{\boldsymbol{v}}'_k + (1-\alpha)\hat{\boldsymbol{v}}_{k-1}$ and bound all updated spiky set assignment probabilities between $1 - q_{spiky}$ and 0.9.

6) If satisfactory convergence is achieved in the optimisation stage (i.e. $\beta_{\hat{r}_{IS}} < \beta_{target}$), terminate the algorithm.

7) If $k \cdot n_{opt} > n_{max,zero}$ and d_{cap} has not been exceeded, terminate the algorithm, as the overload probability is close to zero.

8) If $d_{opt} > d_{cap}$, set $n = 0$ and proceed with step 9. Otherwise, increment k and go back to step 3.

9) Generate an i.i.d. sample of 50 set assignment vectors $\boldsymbol{x}_{n+1}, \ldots, \boldsymbol{x}_{n+50}$ according to $g(\cdot; \hat{\boldsymbol{v}}_k)$, and obtain the associated random profile selections $\boldsymbol{\pi}_{n+1}(\boldsymbol{x}_{n+1}), \ldots, \boldsymbol{\pi}_{n+50}(\boldsymbol{x}_{n+50})$. Also generate a set of i.i.d. samples of m time steps, $\boldsymbol{\theta}_{n+1}, \ldots, \boldsymbol{\theta}_{n+50}$. Set $n \leftarrow n + 50$.

10) Estimate the overload risk using (15) with $\boldsymbol{v} = \hat{\boldsymbol{v}}_k$. Estimate the relative error $\beta_{\hat{r}_{IS}}$ of the estimate using the weighted version of (9). If $\beta_{\hat{r}_{IS}} < \beta_{target}$ or if $k \cdot n_{opt} + n > n_{max}$, terminate the algorithm. Otherwise go back to step 9.

D. Generalising the Importance Distribution

If it is possible to derive a suitable general profile selection IS distribution for a wider range of assets from asset specific distributions, the optimisation stage in the CE algorithm of the previous subsection could be skipped, potentially resulting in greater computational efficiency. For finding a generalised IS distribution, here the CE-optimised IS distributions obtained by evaluating the CE algorithm on 150 MV/LV substations were used. The core idea of the generalisation approach investigated here is to replace the probabilities to pick spiky profiles of *individual customers* by probabilities for the 48 bins used in the demand model.

To obtain the latter, firstly the CE-optimised $\hat{\boldsymbol{v}}_k$ of all assets with less than 80 customers were selected. Secondly, using these, the mean of all per-customer spiky set probabilities in a bin was computed yielding average bin spiky set probabilities. Finally, a thresholding was performed such that only bins with average spiky set probabilities higher than 0.15 were assigned these higher probabilities. The initial probabilities from (12) were assigned to the remaining bins.

IV. RESULTS

The demand model and the Monte Carlo methods described in sections II and III were implemented in *R* version 3.6.1. The code was developed and run on a shared RStudio Server environment. For speed comparisons, a generic computation was carried out repeatedly to verify that computational speeds

due to different workloads of the shared RStudio Server environment varied in an acceptably small range.

The convergence target for all Monte Carlo methods was set to $\beta_{target} = 0.1$, with $n_{max} = 20{,}000$ and $n_{max,zero} = 10{,}000$ as maximum sample sizes to keep computation times in reasonable limits. To determine suitable values of the remaining method parameters, the sensitivity of computation time to the variation of individual parameters in a discrete range around an experience-based default value were explored for ten representative substations. Based on these initial experiments, the number of time steps sampled per trace was set to $m = 2{,}000$ for all methods and the additional parameters of the CE algorithm were set to $n_{opt} = 500$, $\rho = 0.05$, $\alpha = 0.6$, $q_{spiky} = 0.95$ (see [9] for a more detailed description).

To benchmark the computational efficiency of the Monte Carlo methods, the abovementioned 150 MV/LV substations were employed. For the reference method (abbr. as *Ref*), the conventional MC sampling (abbr. as *MC*) and CE-optimised profile selection IS (abbr. as *CE-IS*) nine replicates were computed for each substation. Due to limited computational resources, for the generalised bin spiky set probability IS method (abbr. as *Gen-IS*) five replicates were computed.

A. Accuracy of the Investigated IS-MC Methods

From theoretical considerations it follows that both conventional MC sampling and IS-MC yield unbiased estimates. However, contrasting the IS-MC benchmarking results with conventional MC sampling results showed that inaccurate risk estimates were often obtained for substations with a large number of customers.

The number of substation customers determines the number of parameters to be estimated and thus the dimensionality of the problem. It is a known issue that in high-dimensional settings the likelihood ratio (given in (14)) degeneracy problem leads to unstable IS distributions [11]. Concretely the issue arises e.g. when more than 15-20 customers are assigned high spiky set probabilities by the CE algorithm in a suboptimal way, which results in nearly all samples having extremely small importance weights (see (15)).

Additional experiments for a single substation (results not shown) indicate that the likelihood degeneracy problem can be ameliorated when the optimisation sample size is increased. However, for the investigated case, $n_{opt} \geq 20{,}000$ was necessary to obtain an estimate in the correct order of magnitude which strongly reduces computational efficiency.

Therefore, for the speedup quantification in the next subsection, a filtering step was included: results of both IS methods were compared to results of the reference method using Welch's t-test, and omitted if estimates were significantly different ($\alpha = 0.05$). In practice, this mostly excluded assets with more than 80 customers. Future research could investigate limiting the dimensionality of the model parameterisation or implement the improved CE method proposed in [11].

B. Efficiency of the Investigated MC Methods

The computation times t until reaching $\beta_{target} = 0.1$ were measured for the computational efficiency benchmarking

of the three considered MC sampling methods against the reference method. For those cases where the estimate did not converge to $\beta_{target} = 0.1$ before reaching $n_{max} = 20,000$, an extrapolation based on the expression for the relative error from (9) was made to approximate the total computation time needed for convergence.

The average speed-ups of the sampling methods with respect to the reference method in obtaining r_+ and r_- are shown in Fig. 1. Only nonzero estimates and cases where estimates were found to be accurate are considered here and for the remainder of this section. While all methods perform better than the reference method, it is surprising that the *CE-IS* method has a similar computational speed as the *MC* method. The *Gen-IS* method, in turn, outperforms both, as it can reap the benefits of profile selection IS without the computational cost of the CE optimisation required in the *CE-IS* method.
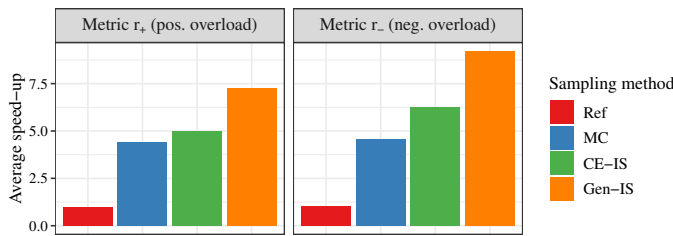

Fig. 1. Comparison of the average speed-up $\bar{t}_{ref}/t_{method}$ for all methods with respect to the reference method.

However, average speed-up comparisons across many assets with different properties hide much of the underlying detail. Generally, one of the most important determinants of the required computational effort of Monte Carlo methods is the magnitude of the quantity being estimated. Therefore, in Fig. 2, average speed-ups per order of magnitude of the estimates of metrics r_+ and r_- are shown for all methods.

Fig. 2 reveals that the profile selection IS methods were able to achieve high speed-ups between 10-50 times for estimates in the orders 10^{-5} and smaller in several cases, clearly outperforming conventional MC sampling. Furthermore, the
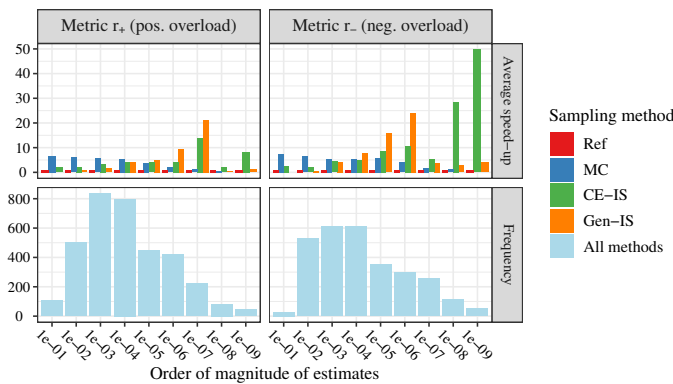

Fig. 2. Comparison of the average speed-ups $\bar{t}_{ref}/t_{method}$ (top panels) of estimates per bin for all methods with respect to the reference method. The bottom panels show histograms indicating how many estimates fall in which order of magnitude.

CE-IS method shows much higher speed-ups in the orders 10^{-8} and 10^{-9} for the r_- metric than the *Gen-IS* method. This may indicate that the CE optimisation of the IS distribution for each asset individually leads to well targeted IS distributions for very rare events – which outweighs the advantage of omitting the optimisation stage in the *Gen-IS* method. Finally, the histograms in the bottom panels of Fig. 2 show that most estimates fall in the orders 10^{-5} and larger, which are the orders of practical relevance for DSOs. This explains why the high speed-ups of the profile selection IS methods for rare events do not appear as pronounced in the average speed-ups of Fig. 1 and it suggests that conventional MC sampling is already well suited for many practically relevant cases. A refined algorithm could adapt its sampling strategy depending on the estimate order of magnitude such that the respectively most advantageous method is employed.

V. CONCLUSION

In this paper, firstly an approach to account for the stochasticity of customer behaviour in distribution network demand modelling was presented. Secondly, Monte Carlo methods for the computationally efficient estimation of network asset overload probabilities were developed. The evaluation of the methods on 150 MV/LV substations showed that cross-entropy optimised importance sampling in the current setting works well for assets with less than 50-80 customers and shows the greatest performance advantages over conventional Monte Carlo sampling for rare event probability estimates.

REFERENCES

[1] F. Pilo, G. Celli, E. Ghiani, and G. G. Soma, "New electricity distribution network planning approaches for integrating renewable," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 2, no. 2, pp. 140–157, 2013.

[2] G. Celli, F. Pilo, G. G. Soma, R. Cicoria, G. Mauri, E. Fasciolo, and G. Fogliata, "A comparison of distribution network planning solutions: Traditional reinforcement versus integration of distributed energy storage," in *2013 IEEE Grenoble Conference*. IEEE, 2013, pp. 1–6.

[3] R. Bernards, R. Moorlag, E. Rijksen, and J. Morren, "Development of a bottom-up scenario analysis network planning tool," in *2019 25th Internatioinal Conference and Exibition on Electricity Distribution*, 2019, pp. 3–6.

[4] T. Valckx, "Stochastic effects of customers on peak loads in a power distribution grid," Master's thesis, 2019.

[5] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*, 3rd ed. Hoboken, NJ, USA: Wiley, 2016.

[6] A. M. L. Da Silva, R. A. Fernandez, and C. Singh, "Generating capacity reliability evaluation based on monte carlo simulation and cross-entropy methods," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 129–137, 2010.

[7] E. Tomasson and L. Söder, "Improved importance sampling for reliability evaluation of composite power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2426–2434, 2016.

[8] P. van de Sande, M. Danes, and T. Dekker, "ANDES: grid capacity planning using a bottom-up, profile-based load forecasting approach," *CIRED-Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 2097–2100, 2017.

[9] J. N. Betge, "Monte carlo sampling techniques for the efficient estimation of risk metrics of a stochastic distribution grid power demand model," Master's thesis, 2020.

[10] G. Biondini, "An introduction to rare event simulation and importance sampling," in *Handbook of Statistics*. Elsevier, 2015, vol. 33, pp. 29–68.

[11] J. C. Chan and D. P. Kroese, "Improved cross-entropy method for estimation," *Statistics and Computing*, vol. 22, no. 5, pp. 1031–1040, 2012.