Valuation of measurement data for low voltage network expansion planning

Electric Power Systems Research 151 (2017) 59–67

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Valuation of measurement data for low voltage network expansion planning

M. Nijhuis[a,*], M. Gibescu[a], J.F.G. Cobben[a,b]

[a] Electrical Energy Systems, Eindhoven University of Technology, Den Dolech 2, Eindhoven, The Netherlands
[b] Liander N.V., Utrechtseweg 86, Arnhem, The Netherlands

ARTICLE INFO

Article history:
Received 15 August 2016
Received in revised form 5 April 2017
Accepted 13 May 2017

Keywords:
Advanced meter infrastructure
Distribution networks
Power system monitoring
Power system planning

ABSTRACT

The introduction of electric vehicles and photovoltaics is changing the residential electricity consumption. Distribution network operators (DNO) are investing in an advanced metering infrastructure (AMI) to enable cost reduction through smart grid applications. The DNO also benefits from the additional measurement data the AMI gives for the network planning process. The availability of AMI data can be limited by the cost of communication and by privacy concerns. To determine the social welfare of an AMI, the economic gains should be estimated. For the planning of the low voltage (LV) network, a method for determining the value of an AMI still needs to be developed. Therefore, a planning methodology which allows various levels of measurement data availability has been developed. By applying this approach the value of different levels of AMI from an LV-network planning perspective can be determined. To illustrated the application of this approach a case study for the LV-network of a Dutch DNO is performed. The results show that an increase in measurement data can lead to €49-254 lower LV-network reinforcement costs. A detailed analysis of the results shows that already 50% of the possible cost reduction can be achieved if only 65% of the households have AMI data available.

1. Introduction

With the electrification of heating and transportation loads and the introduction of distributed generation, the residential energy use is changing. The loading of the low voltage (LV) network can increase due to additional loads like heart pumps. Bi-directional power flows become possible through rooftop PV [1]. On the local level, the pace of and extent to which these developments take place can vary. Errors in the spatial load forecasting can have a significant effect on the required investments in the LV-network [2]. Currently, in most LV-networks, almost no measurement data is available, which generates uncertainty about the current loading of the LV-network [3]. This makes an adequate planning of the LV-network more complicated. The introduction of an advanced metering infrastructure (AMI) can, however, present the distribution network operator (DNO) with an unprecedented amount of data about the residential load.

The AMI data can be used for a multitude of applications [4], from demand response to automating the meter reading process and more accurate LV-network expansion planning [5]. AMI data offers many opportunities for real time operation of the distribution network [6–8], however for the planning of the LV-network the available opportunities are not well documented. AMI data can be used to improve the load modelling [9] and load models based on AMI data can be implemented in the network planning process [10]. The question of how much value these measurements give to the DNO remains open. Privacy concerns about the use of AMI data require anonymisation. Depending on the implemented AMI communication infrastructure a limited read-out frequency and/or measurement frequency of the AMI data may be required [11]. This limits the DNO in the amount of information which can be extracted from the AMI measurements or requires additional communication infrastructure investments [12].

To be able to make adequate choices about the extent of the implementation of an AMI, the benefits of the AMI data should be assessed. The valuation of the AMI data from a demand response perspective has already been performed [13]. The assessment of the automation of a business process like the surveying of conventional electricity meters can easily be assessed from the expenses of the DNO. The use of AMI data to accurately assess transformer loading has also been studied [14], as well as the creation of load forecasts based on AMI data [15]. The usefulness of the additional measurements to increase observability from an operational perspective,

* Corresponding author.
E-mail addresses: m.nijhuis@tue.nl (M. Nijhuis), m.gibescu@tue.nl (M. Gibescu), j.f.g.cobben@tue.nl (J.F.G. Cobben).

http://dx.doi.org/10.1016/j.epsr.2017.05.017
0378-7796/

e.g. in state estimation, is a common research topic [16], however, the data requirements from a long-term planning perspective are different. Preliminary research in using AMI data for the long-term network planning problem has been conducted [17]. The valuation of AMI data for the planning and reinforcement of LV-networks has not been covered. It is, therefore, hard to give an adequate estimation of the value of AMI data from an LV-network expansion planning perspective. This estimate is needed when determining the required characteristics of an AMI. From the point of view of generation adequacy studies into the value of additional measurement data have already been performed [18,19], however for the LV-network expansion planning the effects of the additional measurement data still need to be quantified. A uniform network planning methodology which incorporates different types of LV-measurement data is therefore needed to assess the value of AMI measurements for the LV-network planning process. Like standard network planning approaches, the methodology to provide a valuation for the AMI data from a network planning perspective should compute the LV-network cost, however, unlike the currently applied approaches, the methodology for the valuation should be able to handle load data with different levels of uncertainty in a similar way.

This paper proposes an LV-network expansion planning approach capable of handling these different levels of load uncertainty. How this approach subsequently can be applied to determine the value of AMI data from a network planning perspective is also shown in this paper. In Section 2 the effect of measurement data on the planning of the LV-network is discussed. In Section 3, the different levels of LV-measurement data availability are discussed, in combination with the methodology for incorporating this data in the LV-network expansion planning process. In Section 4, the application of this methodology is presented through a case study for the largest DNO in the Netherlands. Conclusions on the value of AMI data from a network expansion planning perspective are shown in Section 5.

2. Use of measurement data in LV-network expansion planning

The use of additional measurement data in LV-network planning reduces the uncertainty about the current loading of the LV-network. This reduction in uncertainty can lead to better investment decisions and thus to lower LV-network cost. If more measurement data is available the number of possible loading states of the LV-network reduces. This results in a tighter probability distribution of the load, which reduces the uncertainty and allows for more efficient investment decisions. Assume the loading of the network is normally distributed with mean μ and standard deviation σ. The network can be planned for the 98th percentile value of the peak load, or roughly the value of $\mu+2\sigma$. This is illustrated in Fig. 1a where a normal distribution of the loading and the ideal planning level are shown. In this illustration, the network loading can be seen as the level of the peak load in the network. Where the distribution of this loading comes from the uncertainty about the actual level of peak load as there are no measurements available. A reduction of uncertainty would result in a lower variance, so $\sigma_{new}<\sigma_{old}$. The mean of the new probability density function can also change. Through the reduction of uncertainty, the mean of the probability density function moves closer to the actual value of the loading. Two situations are likely to occur $\mu_{new}+2\sigma_{new}<\mu_{old}+2\sigma_{old}$ or $\mu_{new}+2\sigma_{new}>\mu_{old}+2\sigma_{old}$. Both situations are shown in Fig. 1b and c, respectively.

$$\mu_{new} + 2\sigma_{new} < \mu_{old} + 2\sigma_{old}$$

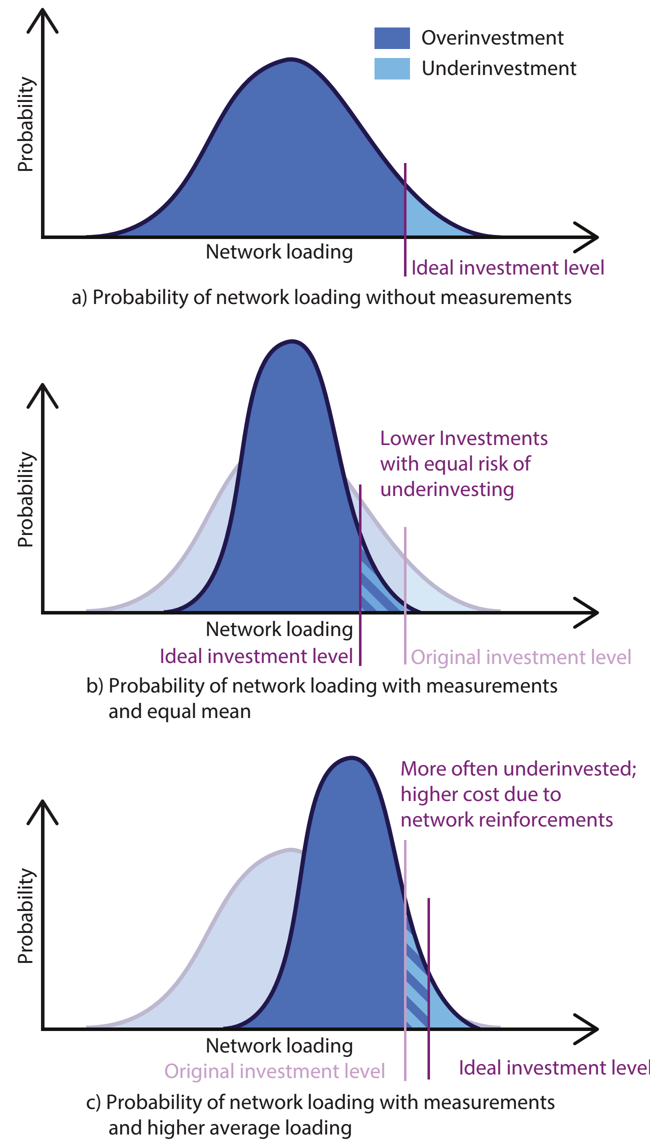

Fig. 1. Value of measurement data in the LV-planning process.

If the network is still planned with the same level of risk, the 98th percentile value of the peak load, the network investments will decrease. The value of the 98th percentile of the load is lower thus a less strong and cheaper network can be constructed with the same expected chance of overloading. This difference is the cost saving for the DNO from a network planning perspective.

$$\mu_{new} + 2\sigma_{new} > \mu_{old} + 2\sigma_{old}$$

If the network is still planned with the same 98th percentile value of the load, the DNO will construct a stronger and more expensive network. However, as the old network was planned with a percentile value for the risk which turned out to be lower than the 98*th* percentile, more cases in which the network has to be reinforced at a later stage will occur (more than the planned 2% of the cases). In the long run, this reduction in the need for future reinforcements of the network generates a lower total investment cost for the DNO, as unforeseen reinforcements can assumed to be suboptimal.

3. Methodology

To assess the value of data in the LV-network from a planning perspective and to consequently calculate the break-even point at which investing in additional measurements becomes economically infeasible, an LV-network expansion planning approach needs to be defined. Based on this planning approach the expected cost with different levels of data availability can be calculated. As the current approach to network planning is based on a simple peak load calculation, this approach first needs to be adjusted to allow for the inclusion of additional measurement data. Hereafter, a generic LV-network expansion planning approach capable of handling different levels of measurement data availability is defined. The implantation of different levels of measurement data availability is discussed hereafter. Finally, the metrics for the calculation of the added value of the additional measurements are determined.

The LV-network planning process is traditionally based on the assessment of the network over a time horizon of 40 years. From scenarios about future changes in the residential load, a single load growth factor is distilled. This yearly load growth factor is combined with an adjusted peak load per household. By taking into account the coincidence factor of the loads, the peak load in the whole network can be determined. The network is designed for this future peak load, by estimating the voltage levels and component loadings during the assumed peak load conditions [20]. In most LV-networks, only a single measurement at the MV/LV transformer is available. In addition, only the maximum loading of the transformer is recorded. This measurement data often needs to be collected manually. Therefore most of the MV/LV transformers only have one measurement point per year. For most residential customers only data on the yearly energy consumption exists, as it is needed for billing purposes.

3.1. LV-network planning with AMI data

In the current network planning approach, the emphasis is on creating a network that is strong enough to handle most load scenarios for the coming 40 years. This generally creates a more robust network than strictly necessary. As data about the current loading of the LV-network is usually unavailable, reinforcing the LV-network incrementally depending on the load growth is not the most cost effective option. With the introduction of an AMI or additional measurements in the LV-network, a more incremental approach for LV-network planning becomes feasible. The state of the LV-network can be better monitored, leading to a better planning and evaluation of possible network reinforcements. In this section, an approach is defined with respect to the planning of the LV-network for different levels of measurement data availability. An overview of the proposed planning approach can be seen in Fig. 2.

The step I in the adjusted planning process the future scenarios of the household load for the coming 40 years are generated based on the scenarios discussed in Appendix A. In step II these scenarios are used in a bottom-up household load modelling approach [21] to develop a set of load curves for the household load. These individual profiles need to be aggregated in order to be utilised in the LV-network planning process. Aggregated load profiles can be created from the individual load curves by clustering them to a limited number of load profiles per scenario and per time step [22,23], this is done in step III. Through a Monte Carlo approach random combinations of load curves are assigned to the households. In step IV the load curves are assigned to the households in a random manner. This assigning of the load curves is repeated multiple times to be able to assess different loading situations. In step V a single combination of load curves is transformed into different load levels based on the measurement data availability as discussed in Section 3.2.

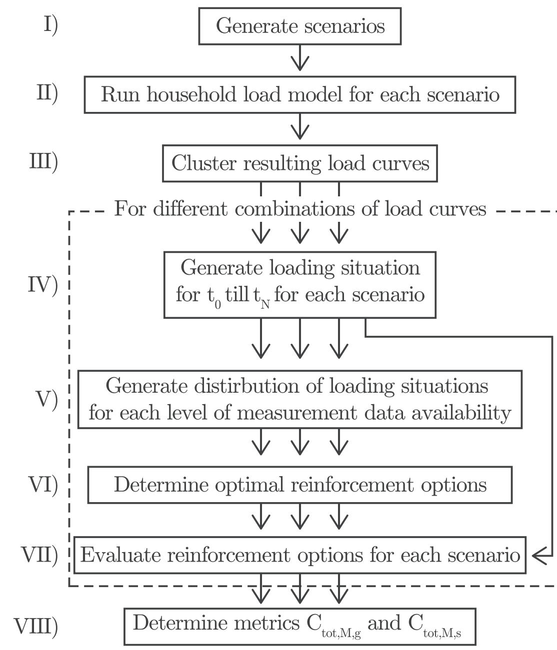

Fig. 2. Adjusted planning process.

Based on these loading situations the network planning optimisation is performed to determine the best reinforcement options in step VI. In step VII these reinforcement options are evaluated based on the original load curves as determined in step IV. By evaluating them with the original load curves a difference in cost is found with respect to the optimal reinforcements as determined in step VI. This difference in cost is the basis for the metrics with respect to the value of additional measurement data determined in step VIII.

The network loading is evaluated once every half year (to account for seasonal variation). Depending on the observability of the network, the loading can be determined to a certain accuracy level at each evaluation interval. Based on this information a new decision can be made on whether or not to invest in network reinforcements. If reinforcements are required, but not anticipated, the DNO has to make an unplanned network reinforcement, resulting in a higher cost. This cost is implemented in the method by using a penalty for any unanticipated reinforcements.

3.2. Measurement data integration in LV-network planning

Depending on the LV-network measurement approach, data with different levels of detail may be available. Based on the most logical locations of measurement devices, a number of different levels of data availability can be distinguished. These different levels are depicted in Fig. 3.

Fig. 3 shows the possible locations of the measurement devices in the LV-network. The purple indicates the situation now with just maximum current measurements at the MV/LV transformer. The blue indicates the next step in implementing measurements, with the maximum current also measured at each of the outgoing LV-feeders. In addition to just measuring the maximum current at the LV-feeder, time-series measurements at the beginning of the feeder can also be employed, this is indicated by the yellow in the figure. The DNO can also choose to collect AMI data every two months for planning purposes. This option is evaluated at two different participation rates for consumers. The red in the figure indicates the situation if 50% of the consumers are willing to let the DNO use their energy usage data for improved network planning, and the green indicates the situation if 100% of the consumers give their energy usage data to the DNO. The way these different levels of availability

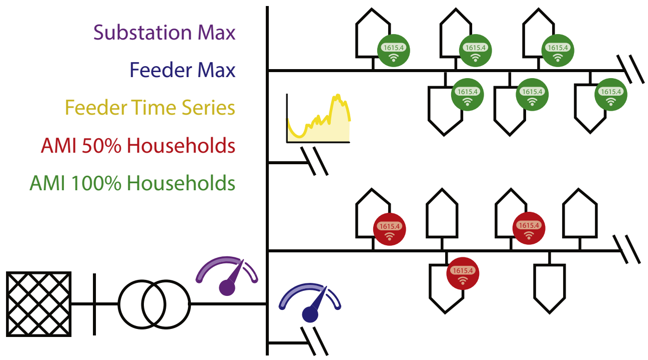

Fig. 3. Different levels of data availability in the low voltage network.

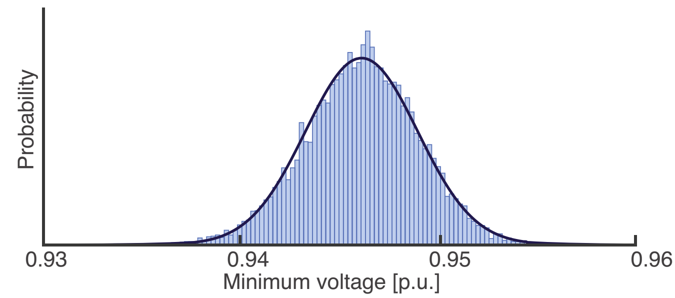

Fig. 4. Distribution of the minimum voltage for a LV-feeder within a range of ±0.1% variation in feeder peak loading.

affect the planning of the LV-network is discussed for each of the five options below.

3.2.1. MV/LV transformer maximum

If only the maximum MV/LV transformer load is known, the loading of a feeder needs to be estimated. To be able to assess the loading of a single LV-feeder, a two-step approach is taken. First of all the variation of the maximum feeder load with respect to the maximum MV/LV transformer loading is estimated. Based on measurements of the maximum feeder current and the maximum MV/LV transformer current, the possible range of LV-feeder currents can be estimated. This is done by using the following formula:

$$P_{fdr} = (1 \pm 0.15)\frac{P_{trf} \cdot N_{HH,fdr}}{N_{HH,trf}} \quad (1)$$

with P_{fdr} the feeder loading, P_{trf} the transformer loading, $N_{HH,fdr}$ the number of customers connected to the feeder and $N_{HH,trf}$ the number of customers connected to the entire MV/LV substation. For a measured maximum MV/LV transformer power there is a 30% range (±0.15) for the maximum feeder loading. As there are no detailed measurements available the actual state of the feeder is considered to be unknown. For each of the maximum feeder loadings within this 30% range, a number of loading situations exist. These loading situations are assessed in the same way as the situation with only a maximum LV-feeder measurement available. This assessment is discussed in the next section. Only the feeder loading is estimated through (1), instead of measured.

3.2.2. LV-feeder maximum

Measurements of the peak current at each of the outgoing feeders can be performed. The maximum current at the beginning of each of the feeders is known in this case. By measuring the current at each feeder, a distinction can be made between feeders with a high and low load. As only the maximum loading of the feeder is recorded, variation in the load profile of the feeder is still possible. The households contributing to the peak load can be, for instance, all in the beginning of the feeder, uniformly distributed over the feeder or at the end of the feeder. The actual loading situation is taken from an assumption on the distribution of possible loading situations. This is illustrated in Fig. 4.

Fig. 4 shows the possible minimum voltage which occurs in the feeder for different combinations of household loads. The maximum current at the beginning of the feeder only varies by ±0.1%, while the minimum voltage lies in the range of 0.935–0.955 (p.u.). To construct the figure, load data from 120 smart meters (15-min data for 3 months) from a single neighbourhood are randomly distributed over a feeder with 67 connected consumers. This distribution of the voltage is generated for a range of feeder loadings. The 98*th* percentile of the voltage distribution is subsequently chosen as acceptable risk level. The load current I_m at bus m is assumed to be equal to the load current at the other buses $I_0 = I_1, \ldots, = I_m, \ldots, = I_{N_b}$ with N_b the number of buses. The load current I_m which results in a voltage deviation equal to the 98*th* percentile value is chosen as the current that is applied in the optimisation procedure as explained in Section 3.3.

3.2.3. LV-feeder profile

In addition to measuring the peak loading, the entire load profile at the beginning of the feeder can also be measured. It is assumed that the profile is measured with a 15-min time step and the data is sent every half year. The main difference from a practical point of view is not in the measurements of the LV-feeder but in the amount of data which needs to be transmitted from the MV/LV substation to the servers of the DNO. In this situation, the investments are assessed for each single set of load curves. This set of load curves $I_{L,0}, \ldots, I_{L,n}$ is used to generate a probability density function of the possible voltage deviations by changing which load curve is assigned to which household (e.g. $I_0 = I_{L,1}, I_1 = I_{L,0}$ or $I_0 = I_{L,0}, I_1 = I_{L,1}$). The assignment of load curves to the households that represent the 98*th* percentile value from the probability density function of the voltage deviations is taken and used in the optimisation procedure.

3.2.4. AMI 50% participation

If an AMI is introduced at the household level, it remains possible that consumers will opt-out of AMI data retrieval by the DNO. In this case, it is assumed that a randomly chosen 50% of the households chose to opt out of the data-sharing. This situation is assessed by giving 50% of the households the same load curve over multiple iterations, while for the other 50% of the households the load is chosen based on the 98*th* percentile value of the peak load distribution of the smart meter measurements. For the households without smart meter measurements, the loads are assumed to be equal $I_0 = I_1, \ldots, = I_{\frac{n}{2}}$ while for the households which are measured, individual load curves are assigned $I_0 = I_{L,0}, \ldots, I_{\frac{n}{2}} = I_{L,\frac{n}{2}}$. The investment options are subsequently assessed for different sets of households with the same load curve over multiple iterations.

3.2.5. AMI 100% participation

If all the households allow the DNO to retrieve their load profile data, the complete loading situation is observable for the DNO. This allows the DNO to accurately assess the required network investments. Though the future network loading is still uncertain, the DNO can determine the chance of voltage or current violations with a high accuracy. For this situation, each load current $I_0, \ldots, I_n$ is assigned an individual value from the available set of load curves $I_{L,0}, \ldots, I_{L,n}$ (e.g. $I_0 = I_{L,0}, \ldots, I_n = I_{L,n}$).

3.3. Planning implementation

The goal of the network planning process is to find the lowest expected cost at which the network can function adequately.

The network planning is often limited to find the lowest investment and operational cost (energy losses) under the long term load uncertainty [24]. The availability is often also included in the LV-network expansion planning, however the majority of the LV-networks within Europe are radial networks constructed with underground cables. This lead to low failure rates and little network reconfiguration options, which in turn generates only small differences in availability of the different network expansion options. For these reasons availability is not taken into account in this work. The approach applied in this paper consist of taking binary reinforcement decisions d from the set of possible options $\mathbb{D}$ at time t, evaluated over the set of discrete times $\mathbb{T}$, to obtain the lowest possible cost C_{tot}. To account for the time value of money the total cost $C_{tot}(t)$ are discounted by the deprecation rate r. The solution needs the voltages at all the nodes n, U_n to be within the limits U_{min} and U_{max} and the branch currents I_{br} in the network to be within the current limits (1 (p.u.)) at all times. This can be expressed as:

$$\min \sum_{t \in \mathbb{T}} \frac{C_{tot}(t)}{(1+r)^t}$$

$$\begin{aligned} \text{s.t.} \quad & d \in \{0,1\} && \forall d \in \mathbb{D} \\ & I_{br,n}(t) \leq 1\,(\text{p.u.}) && \forall n = 0, \ldots, N_b \\ & U_{\min} \leq U_n(t) \leq U_{\max} && \forall n = 0, \ldots, N_b \end{aligned} \tag{2}$$

In reality, it can happen that the voltage and maximum current constraints are violated. These constraints are therefore modelled as a penalty function instead of a hard constraint. The value of the penalty function represents the cost for the DNO when either a violation of the voltage or the maximum current occurs. The penalty function is assumed to be a constant C_{pen}, independent of the severity of the violation. The penalty cost C_{pen} consists of a part of direct damages, for instance, due to a violation of the maximum current limit, the fuse at the beginning of the LV-feeder will trip, resulting in an outage and the need for a replacement fuse. The other part of the penalty function is the additional cost associated with an urgent network reinforcement instead of a planned network reinforcement. Therefore the penalty cost C_{pen} is a DNO specific value.

The total cost C_{tot} consists of multiple parts: The fixed investment cost C_{fix}, the penalty cost C_{pen}, the cost to reinforce the branches of the feeders $d_{cbl}p_{cbl} \cdot l_{branch}$, the cost of adding new feeders $d_{cbl}p_{fdr} + p_{cbl}l_{0\rightarrow mid}$, the cost of an additional MV/LV substation $d_{MV/LV}p_{MV/LV} + p_{cbl,MV}l_{0\rightarrow end}$ and the cost associated with energy losses $p_e\sum_{m=0}^{N_b} L_m$.

$$\begin{aligned} C(t) = {} & C_{fix} + C_{pen} + d_{cbl}p_{cbl}l_{br,n} + d_{fdr}p_{fdr} + d_{fdr}p_{cbl}l_{0\rightarrow mid} \\ & + d_{MV/LV}p_{MV/LV} + d_{MV/LV}p_{cbl,MV}l_{0\rightarrow end} + p_e\sum_{m=0}^{N_b} L_m \end{aligned} \tag{3}$$

with p_{cbl}, p_{fdr}, $p_{MV/LV}$, $p_{cbl,MV}$ and p_e being the prices of one meter LV-cable, an additional feeder at the LV busbar, an additional MV/LV-substation, one meter of MV cable and one kWh of electricity respectively, $l_{br,n}$ the length of branch n, $l_{0\rightarrow mid}$ and $l_{0\rightarrow mid}$ the distance between the MV/LV-substation and the middle and end of the feeder, d_{cbl}, d_{fdr} and $d_{MV/LV}$ the vectors of binary decision variables corresponding to whether or not to reinforce a branch, add a new feeder or add an additional MV/LV substation respectively.

The branch currents I_{br}, bus voltages U_n and branch losses L_m are computed from a backwards forwards sweep load flow calculation:

$$U_i = U_{i,0} - \sum_{n=i}^{N_b} \left(Z_n I_{br,n}\right) \tag{4}$$

where U_i is the voltage at bus i, bus indexes i, n and m running from the busbar (index 0) to the end of the feeder (index N_b), $U_{i,0}$ the base voltage (400 V L-L), Z_n the impedance of the section between node n and $n-1$, $I_{br,n}$ the branch current as calculated by:

$$I_{br,n} = \sum_{m=n}^{N_b} \left(\frac{U_{m,0}I_m + L_m}{U_m}\right) \tag{5}$$

with I_m the current of the load at bus m. The loss L_m between bus m and bus $m+1$ is calculated by:

$$L_m = \alpha Z_n I_{br,n}^2 \tag{6}$$

with α a factor to account for the fact that the loss calculation is now only based on a single peak load moment instead of the entire load curve. The factor α is calculated at the beginning of the optimisation by calculating the loss by applying the entire load curve and dividing it by the loss calculated during peak load times.

The voltage limits that are used during the assessment of the LV-feeders are chosen as ±0.05 (p.u.). This more stringent limit is applied instead of the generally applied limits of ±0.1 (p.u.), as the assessment is based only on the LV-feeders. The voltage fluctuations in the MV-network are not taken into account. The residential load currents I_m are constructed based on the amount of available measurement data as discussed in the previous section. In order to have an idea of the time evolution of the load currents, the load current needs to be constructed for all the scenarios s in the scenario space $\mathbb{S}$.

For the assessment of the possible network reinforcement alternatives, genetic algorithms [25] and particle swarm optimisation [26] are often used. Computation wise there is little difference in the performance of these methods [27,28]. A genetic algorithm is applied in this paper to determine the best reinforcement options. It is assumed that the best reinforcement options are equal for loading situations where $\sum_{n=1}^{N_B} Z_{poc,n} \times I_n$ differs less than 1% of the maximum value, with N_B the number of buses, $Z_{poc,n}$ the impedance at the points of connection and I_n the current injected at bus n. This allows for a lower number of computations while determining the best reinforcements option for each scenario and for each level of measurement data availability.

3.4. Metrics to valuate network planning data

To valuate the available measurement data for the network planning process the difference in total cost with different levels of data availability is calculated by (2). Not all networks might have the same need for measurement data from a network planning perspective. Networks for which no problems are expected in the coming years, even under extreme load growth scenarios, obviously do not benefit from having measurement data available. The question remains for which type of networks the additional measurement data holds more value. As the LV-network consists of many different types of feeders, whether or not the value of additional measurement data differs for these feeders should be calculated. This value can be calculated by adjusting the minimisation of the cost as given in (2) with an objective function which assumes an equal probability of each scenario s in the scenario space $\mathbb{S}$.

$$C_{tot,M,g} = \frac{1}{|\mathbb{S}|}\min_{d \in \mathbb{D}} \sum_{s \in \mathbb{S}} C_{tot,M,s,g} \tag{7}$$

where $C_{tot,M,s,g}$ is the total cost of LV-feeder g under scenario s and with the level of measurement data availability M. By comparing these total cost $C_{tot,M,g}$ for a specific feeder g, but different levels of data availability M, the value of the measurement data for a specific feeder g can be determined. The optimisation is done based on the evaluation of the cost under all scenarios, as it is assumed that the scenarios are indistinguishable at the time that the investments are made. In this case, the scenarios are assumed to have an equal probability. If more information on the individual probabilities of specific scenarios is present the sum in Eq. (7) can be changed into a weighted sum.

The expected load growth also should have a significant effect on the value of the additional measurement data, therefore a metric to quantify the effects of using different load growth scenarios is presented. If the loading is expected to remain more or less constant in the coming years, additional measurement data will not add any value, while scenarios with a fluctuating load growth should give more value to the measurement data. To assess the effects of the scenarios on the value of measurement data, the objective function as given in (2) is changed to generated the expected cost for each scenario s, while taking into account all the LV-feeders:

$$C_{tot,M,s} = \frac{1}{|\mathbb{G}|}\sum_{g \in \mathbb{G}} \min_{d \in \mathbb{D}} C_{tot,M,s,g} \quad (8)$$

where $\mathbb{G}$ is the set of generic LV-feeders g. The objective function (8) is very similar to the objective function in (7), where the objective in (7) is averaged based on the scenarios s, (8) is averaged based on the feeders g. The main difference is that in (7) the minimisation is performed based on the performance under all the scenarios, while in (8) the minimisation is performed for each scenario individually.

Another aspect which can be determined by the use of this methodology is the incremental value of an increase in the penetration of AMI. If more and more consumers connected to the same feeder allow for the use of their AMI data for network planning purposes the expected total investment cost will decrease. Also with an increasing AMI usage, there will be a point when the additional measurements at the feeder will no longer generate a lower expected cost. It is interesting to determine at which point it is no longer useful for the DNO to increase the measurement set from a network planning perspective. This is calculated by determining the expected cost at different levels of AMI penetration in the same manner as done for the 50% AMI penetration.

4. Case study

To illustrate the application of the proposed approach for the valuation of measurement data for the planning of an LV-network, a case study on the network of Liander (Dutch DNO) is performed. For this case study, first the network which is considered is discussed. This is followed by an elaboration on the creation of the household load for the coming years. Next to this, the reinforcement options and costs are given. Based on this information the analysis of the value of measurement data from a network planning perspective can be performed. The results are given in Section 4.1.

To gain an idea of the value of measurement data for the DNO, a representative subset of the network of the DNO needs to be selected for the case study as the complete network is too large for the analysis (>80,000 LV-feeders). The selection of the representative LV-feeders is done by the application of a clustering approach to the entire LV-network of Liander. By applying a clustering approach to the entire network generic feeders can be extracted from the network. These generic feeders are representative for the whole network. A fuzzy k-medians clustering approach is used to generate a set of generic LV-feeders based on the network data. The following network parameters are used in the clustering approach: impedances, cable length, number of branches, branch depth and the number and type of connected customers. Next to these network parameters, the graph theory concepts of degree distribution, sequence and the centrality of the power, impedance and length are also used for the clustering. By using these parameters feeders can be rebuilt from the cluster centres which would represent the structure and loading of the original feeder. For a more detailed description of the applied clustering approach [29] can be consulted. By applying this clustering approach a subset of 96 feeders is extracted from the entire LV-network. These feeders are subsequently analysed with a peak load of 125% of the current peak load, to determine which feeders would require reinforcements within the next 5 years. This creates a set of 26 generic feeders that are used to valuate the measurement data from an LV-network planning perspective. The other feeders are not taken into account as reinforcements are not required independent of the amount of available measurement data.

Table 1
Cost of the components used in the case study.

Symbol	Component	Cost (€)
p_{cbl}	LV-cable (\km)	45,000
$p_{cbl,MV}$	MV-cable (\km)	70,000
p_{msr}	Additional MV/LV substation	25,000
p_{fdr}	LV-busbar expansion for an additional feeder	7000
p_e	Electricity price (\MWh)	40
C_{pen}	Penalty cost	8500
C_{fix}	Fixed investment cost	1300

For the creation of the scenarios which are used to assess the adequacy of the LV-feeders, a scenario space has been constructed based on all the combinations of future scenarios for electric vehicles (EV), photovoltaics (PV), heat pumps, economic growth and appliance efficiency. Appendix A gives an overview of these scenarios. The household load is generated by using a bottom-up Monte Carlo Markov Chain approach [21]. In this method, the behaviour of each household member is modelled using Markov Chains, based on data from time use surveys. The appliances are modelled individually and switch on or off based on the modelled behaviour of the users. This generates a load curve for each individual appliance within a household. The appliances are subsequently aggregated to generate the load curve of the household. For each of the individual households within a network, this approach generates a unique load curve. In order to create the load curve of a household for the coming years, scenarios on how the appliances within the household will develop, need to be implemented. These scenarios are implemented by updating the appliances in the households based on the penetration rate of new appliances (EV, PV, etc.) for a certain scenario. In the same way, trends with respect to the changes in appliance ratings (e.g. efficiency and size of appliances) are also taken into account. With this model, the load curve for 2000 individual households for one day each half year for the coming 40 years has been created for the scenarios as described in Appendix A. To make the analysis of possible loading situations computationally feasible the household load curves have been clustered by using a k-means clustering approach 10 representative load curves for each half year of each scenario are generated [23]. More advanced approaches exist to reduce the dimensionality of load profiles [30,31], however the accuracy of k-means approach was found to give an acceptably small error (a Wasserstein distance of 0.984 was found when comparing the voltages calculated with the original load profiles to the voltages created with clustered load profiles).

The reinforcement options are assessed based on their expected value. For the calculation of the expected value, the data in Table 1 is used. The depreciation rate r is chosen as 3% per year and for simplicity, a single cable type is used for replacements: a 4 × 150AL

Table 2
Expected discounted cost for the coming forty years for LV-network reinforcements with different levels of data availability per LV-feeder.

	Cost (€)	Percentage
100% AMI	520	67.2%
50% AMI	680	87.8%
Feeder time series	674	87.1%
Feeder maximum	725	93.7%
Substation maximum	774	100%

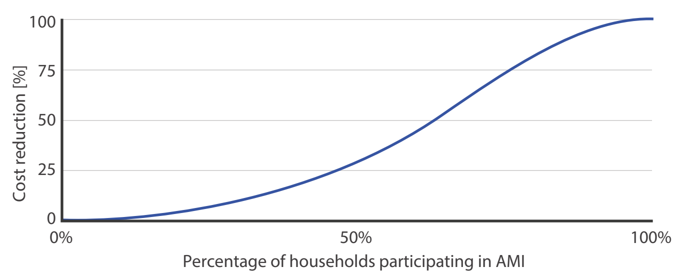

Fig. 5. Mean expected cost with decreasing AMI participation.

Table 3
Characteristics of the feeders for which the results are shown in Fig. 6.

	A	B	C	D	E
Length (m)	900	450	500	890	820
Connections	9	36	45	15	95
$\bar{Z}_{PoC}$ (mΩ)	133	110	105	162	140

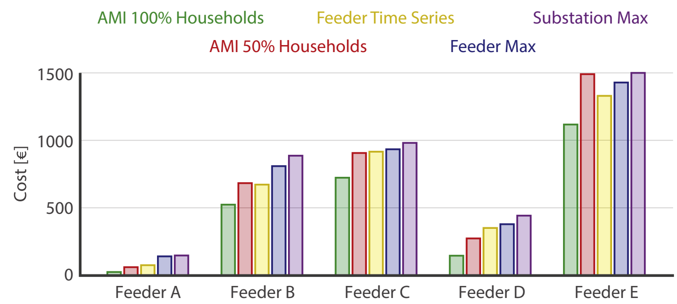

Fig. 6. Expected cost network reinforcements for five feeders with different levels of data availability.

LV-cable. With this data, the expected cost of the reinforcements can be computed for the different levels of measurement data availability.

4.1. Results

The approach described in the previous section has been applied to valuate the measurement data for the LV-network of Liander. The results are shown in this section. First, the cost of the different levels of data availability is given. Secondly, the results are analysed on the different LV-feeders characteristics, followed by an analysis based on the chosen scenarios.

4.1.1. Valuation of measurement data

The cost of the network per LV-feeder are calculated for different levels of data availability and the results are shown in Table 2.

As expected, the table shows an overall reduction in expected cost as the availability of the measurement data increases. The use of feeder data instead of substation measurements can reduce the expected cost for feeder reinforcement by about 10%, compared to the current situation. This translates to a cost saving of €49 over 40 years per LV-feeder. For these savings to be achieved, additional measurement devices will have to be installed, maintained and the measurement data has to be collected. Most likely this results in an uneconomical business case if only the benefits from the network expansion planning perspective are taken into account. The possibilities of having time series feeder measurements and acquiring AMI data from 50% of the households generate a more or less equal expected cost that is slightly lower than the expected cost using a measured feeder maximum. Acquiring all the data from the AMI infrastructure can lead to an additional 20% lower expected cost. The majority of the analysed feeders hardly need any reinforcements, even in the most extreme scenario. The added value of measurement data for these feeders is zero from a network planning perspective.

The expected cost between the 100%, 50% and 0% AMI data availability cases differ significantly. To further illustrate this the expected costs for intermediate levels of AMI availability have been computed.

In Fig. 5 the expected cost for increasing levels of AMI data availability is shown. It can be seen that the expected cost as a function of the level of AMI data availability follows an S-curve. The inclusion or exclusion of just a couple of households from the AMI data does not have a large effect on the expected cost. If the level of available AMI data can be increased it pays off most if a level of data availability of at least 40% is already present. Increasing the level of data availability past the 90% no longer generates any significant additional returns. For a level of AMI data availability of 65% of all the households, a 50% cost reduction is achieved.

4.1.2. Feeder characteristics

Next to the results for the combined LV-feeders, also the results for an individual feeder can be given. From all the feeders used in the calculations, five feeders have been selected. The characteristics of these feeders are shown in Table 3.

The table shows the length, number of connections and average impedance at the point of connection (PoC) for each of the feeders. Feeders A and D are rural feeders with a large length and a small number of customers. Feeders B and C are suburban feeders with an intermediate length and amount of customers and feeder E is an older urban feeder with a long length, high number of customers and high average impedance at the PoC. For these five feeders, the results for the calculation of the network cost with different levels of data availability are plotted in Fig. 6.

The trend of a decrease in expected cost with an increase in available measurement data is visible in the figure of the individual feeders as well. There are however large differences in the cost savings which can be achieved for the individual feeders. For feeder A the availability of all the AMI data for the LV-network planning process can lead to a cost saving of about 80%, while for feeder E the cost saving is only 20%. This difference can be explained by the fact that feeder A is, for a large percentage of the analysed scenarios, strong enough to not require any reinforcements, while Feeder E needs reinforcements reasonably early. The use of additional measurement data can in the case of feeder A avoid any unnecessary reinforcements. For feeder E the large differences between the loading of the individual households occur early on, as the penetration rates are small in the beginning. If one knows the sum of all the loads on the feeder, the small number of households which use a certain technology is taken into account, while if a random 50% of the household have a known loading, the small number of households might be excluded. This leads to relatively high expected cost for feeder E. A similar picture to feeder A is present at feeder D. For feeder E the use of AMI data for half of the households does not assist much in the grid planning process, while for the other feeders it gives an additional benefit. For the suburban feeders B and C, the reduction in expected cost is between the reduction of

Table 4
Characteristics of the scenarios for which the results are shown in Fig. 7.

	Low	Low PV, EV	Medium	High PV, EV	High
PV	Low	Low	Medium	High	High
EV	Low	Low	Medium	High	High
Heat pump	Low	Medium	Medium	Medium	High
GDP	Low	Medium	Medium	Medium	High
Efficiency	High	Medium	Medium	Medium	Low

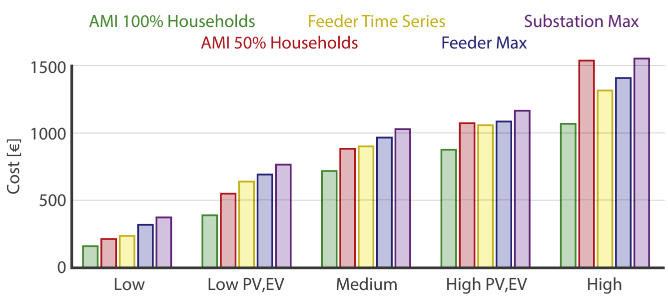

Fig. 7. Expected cost of network reinforcements for the five scenarios from Table 4 with different levels of data availability.

feeders A and E. The additional measurement data has more influence on feeder B in comparison to feeder C. This can be explained by the slightly higher number of customers connected to feeder C. The error induced by the use of average values is, therefore, lower and the benefits of the additional measurement data as well. Overall feeders with a lower need for reinforcements can have more benefits from AMI data in comparison to the feeder reinforcement cost, however the absolute benefits are higher for the feeders that require more reinforcements.

4.1.3. Scenario influence

Next to the feeder characteristics, the uncertainty of the future loading of the network is also expected to have a large influence when it comes to the value of the measurement data. To gain more insight in the effect of the scenarios on the value of measurement data, the results have also been analysed for different scenarios. From the scenarios which have been used in the computation, the subset depicted in Table 4 is used for a more in-depth analysis. The five scenarios which have been chosen consist of the two most extreme scenarios, an intermediate scenario and two scenarios in which just the level of EV and PV penetration is altered from the intermediate scenario. The calculated expected cost for these five scenarios and the five cases of data availability is shown in Fig. 7.

The expected cost for the different scenarios confirms the idea that the relative value of the measurement data is larger as the chance of having to reinforce is smaller. This can be seen as the relative difference between the 100% AMI case and the substation maximum case decreases as the amount of load growth in the scenarios increases. As the amount of load growth increases, the benefits of the usage of only half the AMI data decreases, as it is expected that the difference in loading between households becomes larger. For the scenario with the least load growth the expected costs are close to the case with 100% AMI data available, while for the highest load growth, the 50% AMI case generates almost no profit compared to the substation maximum case. The value of having measurement data available on a feeder level instead of just at the substation level is significantly higher for the extreme scenarios in comparison to the intermediate scenarios. For the extreme scenarios, the difference between the penetration rates at different time steps for different feeders can be larger, giving rise to an increased value of knowing the loading of the feeder.

5. Conclusion

An approach for the valuation of the use of measurement data for the LV-network planning process is presented. The approach is based on using a single LV-network planning approach that allows for the use of different levels of measurement data, through altering the applied load values. This approach allows for the determination of the value of measurements in the LV-network, from a network planning perspective. With this approach, the cost-benefit analysis for installing an AMI can be improved. Based on the proposed method the value of the possible deferral of the LV-network reinforcements can also be included. The approach has been applied to the LV-network of a Dutch DNO. In general, an increase in available measurement data leads to lower reinforcement cost as the network-wide results show. The cost reduction that can be achieved in practice does, however, differ significantly based on the scenario and the type of feeder which is used in the analysis. Lower load growth scenarios and more robust feeders, which still require some level of finely-tuned reinforcements are the main beneficiaries of the uncertainty reduction through additional measurement data. The analysis also shows that on average a level of 65% of AMI data availability can already achieve a cost reduction of 50%.

This work allows the DNO to decide whether or not to invest in additional measurement devices in the LV-network, by not only including the operational benefits for an AMI but also the benefits from the deferral of LV-network reinforcements. As the use of additional measurement data is required for the implementation of many smart grid solutions such as demand side management (DSM), the approach can be extended to include DSM alternatives in the planning process. As future work, a more accurate assessment of the benefits of applying these smart grid technologies can be performed by their inclusion in the LV-network planning methodology.

Appendix A. Load scenarios

For the evaluation of the network expansion planning, the residential load for the coming decades needs to be determined, as LV-cables tend to have a lifetime easily eclipsing the 30 years. The main technologies which are expected to increase in penetration rate and have a large influence on the LV-network are PV, EV and heat pumps. The penetration rate is defined as the number of households that use a certain technology. Next to the increasing penetration of these technologies, the changes in electricity usage should also be incorporated in the scenarios. Scenarios on the economic growth and the change in appliance efficiency are also used. The scenarios are modelled through the use of the following equation:

$$pt(yr) = \frac{pt_{\max}}{1 + e^{-a(yr)}} \quad (9)$$

with pt the penetration level at year yr, pt_{max} the maximum penetration level and a the penetration rate parameter. This approach is applied for the all the scenario drivers, expect for the economic

Table A.5
Overview of the different parts of a scenario and how they are implemented.

	High		Medium		Low	
	pt_{max}	a	pt_{max}	a	pt_{max}	a
PV	82	0.16	32	0.18	7.2	0.19
EV	63	0.09	28	0.11	4.5	0.11
Heat pump	46	0.06	5.2	0.06	0.1	0.05
Efficiency	0.52	0.13	0.59	0.12	0.96	0.17
Growth	1.25%		0.5%		0%	

growth where an exponential growth function is used. In Table A.5 an overview of the employed scenario parameters is given.

For the EV and PV the following assumptions have been made with respect to the modelling of these technologies. The rated power of the PV systems are assumed to uniformly distributed between 2 and 5.5 kW per household. For the EV charging, slow charging at a rate between 3 and 8 kW is assumed. The distribution of departure and arrival times of the EVs as well as the distance driven are modelled based on data from mobility studies according to the method described in [32]. The EVs start charging as soon as they arrive back at the household.

References

[1] M. Nijhuis, M. Gibescu, J. Cobben, Assessment of the impacts of the renewable energy and ICT driven energy transition on distribution networks, Renew. Sustain. Energy Rev. 52 (August) (2015) 1003–1014, Available: http://linkinghub.elsevier.com/retrieve/pii/S1364032115007716.

[2] D.A.G. Vieira, M.M.A. Cabral, T.V. Menezes, B.E. Silva, A.C. Lisboa, Measuring the spatial error in load forecasting for electrical distribution planning as a problem of transporting the surplus to the in-deficit locations, Proceedings – 2012 11th International Conference on Machine Learning and Applications, ICMLA 2012, vol. 2 (December 2012) 475–480.

[3] M. Neaimeh, R. Wardle, A.M. Jenkins, J. Yi, G. Hill, P.F. Lyons, Y. Hübner, P.T. Blythe, P.C. Taylor, A probabilistic approach to combining smart meter and electric vehicle charging data to investigate distribution network impacts, Appl. Energy 157 (January) (2015) 688–698.

[4] E. Liu, M.L. Chan, C.W. Huang, N.C. Wang, C.N. Lu, Electricity grid operation and planning related benefits of advanced metering infra-structure, 2010 5th International Conference on Critical Infrastructure, CRIS 2010 – Proceedings (September 2010) 1–5.

[5] J. Foosns, E. Tnne, T. Pynten, Power system planning in distribution networks today and in the future with smart grids, in: 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013), No. 1426, June 2013, http://dx.doi.org/10.1049/cp.2013.1233.

[6] D. Bernardon, G. Iop, M. Gundel, J. Fernandes, A real-time operation analysis system for distribution networks, Electr. Power Syst. Res. 78 (March (3)) (2008) 346–352, Available: http://linkinghub.elsevier.com/retrieve/pii/S037877960700048X.

[7] A. Alam, J. Zhu, V. Frey, R. Zhao, T. Hoang, S. Larson, S. Mundade, S. Ruelas, A. Herasimava, Improving operational planning study models with historical generation and load data, in: 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), vol. 2016, July, IEEE, May 2016, pp. 1–5, Available: http://ieeexplore.ieee.org/document/7520038/.

[8] A.A. Mohamed, O.A. Mohammed, Real-time load emulator for implementation of smart meter data for operational planning, in: 2012 IEEE Power and Energy Society General Meeting, IEEE, July 2012, pp. 1–6, Available: http://ieeexplore.ieee.org/document/6345619/.

[9] X. Zhang, S. Grijalva, M.J. Reno, A time-variant load model based on smart meter data mining, in: 2014 IEEE PES General Meeting – Conference & Exposition, vol. 2014, October, IEEE, July 2014, pp. 1–5, Available: http://ieeexplore.ieee.org/document/6939365/.

[10] R. Brakken, T. Ait-Laoussine, S. Steffel, D. Sikes, Integrating customer load and GIS data for improved distribution planning & operations (updated), in: 2005/2006 PES TD, IEEE, 2006, pp. 1056–1063, Available: http://ieeexplore.ieee.org/document/1668648/.

[11] W. Luan, D. Sharp, S. Lancashire, Smart grid communication network capacity planning for power utilities, Transmission and Distribution . . . (April 2010) 1–4, Available: 10.1109/TDC.2010.5484223%5Cnhttp://ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=5484223.

[12] W. Luan, D. Sharp, S. LaRoy, Data traffic analysis of utility smart metering network, in: 2013 IEEE Power & Energy Society General Meeting, IEEE, 2013, pp. 1–4, Available: http://ieeexplore.ieee.org/document/6672750/.

[13] P. Siano, Demand response and smart grids – a survey, Renew. Sustain. Energy Rev. 30 (November) (2014) 461–478.

[14] Y.L. Lo, S.C. Huang, C.N. Lu, Transformational benefits of AMI data in transformer load modeling and management, IEEE Trans. Power Deliv. 29 (April (2)) (2013) 742–750.

[15] B.P. Hayes, M. Prodanovic, State forecasting and operational planning for distribution network energy management systems, IEEE Trans. Smart Grid 7 (March (2)) (2016) 1002–1011.

[16] R. Singh, B.C. Pal, R.B. Vinter, Measurement placement in distribution system state estimation, IEEE Trans. Power Syst. 24 (May (2)) (2009) 668–675.

[17] G. Roupioz, X. Robe, F. Gorgette, First use of smart grid data in distribution network planning, 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013) (June 2013) 1–4, Available: http://www.ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=6683438.

[18] J. Pan, S. Rahman, Multiattribute utility analysis with imprecise information: an enhanced decision support technique for the evaluation of electric generation expansion strategies, Electr. Power Syst. Res. 46 (2) (1998) 101–109, Available: http://www.sciencedirect.com/science/article/pii/S0378779698000224.

[19] A.H. van der Weijde, B.F. Hobbs, The economics of planning electricity transmission to accommodate renewables: using two-stage optimisation to evaluate exibility and the cost of disregarding uncertainty, Energy Econ. 34 (January (6)) (2012) 2089–2101, http://dx.doi.org/10.1016/j.eneco.2012.02.015.

[20] H. Willis, Power Distribution Planning Reference Book, vol. 2, CRC Press, March 2004.

[21] M. Nijhuis, M. Gibescu, J. Cobben, Bottom-up Markov Chain Monte Carlo approach for scenario based residential load modelling with publicly available data, Energy Build. 112 (January) (2016) 121–129, Available: http://linkinghub.elsevier.com/retrieve/pii/S0378778815304436.

[22] F.L. Quilumba, W.J. Lee, H. Huang, D.Y. Wang, R.L. Szabados, Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities, IEEE Trans. Smart Grid 6 (March (2)) (2015) 911–918.

[23] M. Nijhuis, R. Bernards, M. Gibescu, J. Cobben, Stochastic household load modelling from a smart grid planning perspective, in: 2016 IEEE International Energy Conference (ENERGYCON), IEEE, April 2016, pp. 1–6, Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7513954.

[24] W.H. Kersting, Distribution System Modeling and Analysis, 2002.

[25] V. Miranda, J.V. Ranito, L.M. Proena, Genetic algorithms in optimal multistage distribution network planning, IEEE Trans. Power Syst. 9 (January (4)) (1994) 1927–1933, Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=331452.

[26] I. Ziari, G. Ledwich, A. Ghosh, Optimal integrated planning of MVLV distribution systems using DPSO, Electr. Power Syst. Res. 81 (October (10)) (2011) 1905–1914, Available: http://linkinghub.elsevier.com/retrieve/pii/S0378779611001271.

[27] C. Gamarra, J.M. Guerrero, Computational optimization techniques applied to microgrids planning: a review, Renew. Sustain. Energy Rev. 48 (2015) 413–424.

[28] R. Shi, C. Cui, K. Su, Z. Zain, Comparison study of two meta-heuristic algorithms with their applications to distributed generation planning, Energy Proc. 12 (2011) 245–252.

[29] M. Nijhuis, M. Gibescu, J. Cobben, Clustering of low voltage feeders from a network planning perspective, Cired 2015, Lyon, France (June 2015).

[30] R. Li, F. Li, N.D. Smith, Multi-resolution load profile clustering for smart metering data, IEEE Trans. Power Syst. 31 (November (6)) (2016) 4473–4482, Available: http://ieeexplore.ieee.org/document/7428953/.

[31] R. Al-Otaibi, N. Jin, T. Wilcox, P. Flach, Feature construction and calibration for clustering daily load curves from smart-meter data, IEEE Trans. Ind. Inf. 12 (April (2)) (2016) 645–654, Available: http://ieeexplore.ieee.org/document/7404272/.

[32] R.A. Verzijlbergh, The Power of Electric Vehicles – Exploring the value of flexible electricity demand in a multi-actor context (Ph.D. dissertation), October 2013, Delft.