APPLYING SMART METER DATA TO LOW VOLTAGE NETWORK PLANNING

Michiel NIJHUIS
Phase to Phase – The Netherlands
michiel@phasetophase.nl

Nard VERMELTFOORT
Alliander – The Netherlands
Nard.Vermeltfoort@Alliander.com

Raoul BERNARDS
Enexis Netbeheer – The Netherlands
Raoul.Bernards@Enexis.nl

ABSTRACT

Low voltage (LV) network planning becomes more and more an issue for the distribution network operator (DNO), with the introduction of heat pumps, electric vehicles and photovoltaics. The mandatory introduction of smart meters generates detailed data of the LV-network loading. The use of these data is governed by privacy laws, which requires the DNO to have a clear and concrete vision of how the data can benefit the public interest. To this extent the applicability of smart meter data for the LV-network planning is investigated. Household load data should still be modelled to be able to use the smart meter data for long term planning studies. This also anonymises the data making its usage more acceptable. From an LV-network planning perspective, the amount of smart meter data required can be limited to a fraction of a year.

INTRODUCTION

With the introduction of electric vehicles (EV), heat pumps (HP) and photovoltaics (PV) the low voltage (LV) network needs to be reinforced in order to facilitate the transition towards a sustainable energy system [1]. If a part of the low voltage network needs to be reinforced, the network is often assessed using peak load and generation values [2]. These values used to be estimated based on the yearly energy consumption of the low voltage customers, as conventionally there are few measurement data available. With the introduction of smart meters, measurement data from users become available for the distribution network operator (DNO). This would enable the application of these measurement data to the low voltage network planning process. With the introduction of data policy regulations, the actual applicability of smart meter data might be not be straightforward. Furthermore, there are many variables which play a role when it comes to measuring data from a household. When it comes to storing and transmitting the smart meter data of millions of households, knowing how much additional data improves the application can be vital.

In order to be able to gain most information out of these data for the network planning process, the use of these data should be investigated. To be able to assess the use of smart meter data in the network planning process the characteristics of the smart meter data should be established. This will be done in section 2, followed by a discussion of the network planning process in section 3. In sections 4-6 the required characteristics of the smart meter data use will be assessed on three different aspects: time resolution, measurement duration and scenario implementation. The implementation of these effects in the network planning process is discussed in section 7.

SMART METER DATA

The data of smart meters is seen as a valuable resource to gain insight into the loading of the low voltage network and seen as an enabler for many smart grid technologies [3]. In Europe, smart meter readouts usually have a sampling time somewhere between the 5 and 30-minutes and have limited availability mainly due to privacy concern. The data produced by smart meters is protected by law. The European laws, rules and regulations on data protection and privacy for all individuals in the European union are recorded in the GDPR [4]. The smart meter data of a household is considered sensitive data, as analysis of the energy usage within a household can show personal habits of the customers, at least to some extent. Analysis or processing of smart meter data by any other party than the user itself is therefore bound to the laws, rules and regulations of the GDPR.

In accordance with the GDPR, the AVG (which is the Dutch implementation of the GDPR) states six grounds on which data processing could be allowed. Out of these six, two grounds would allow the DNO to process smart meter data: if authorised by the end-user or if there is a legitimate public interest. The process to obtain authorisation from the end-user of the smart meter is currently not (yet) implemented, making this option impossible for the time being. Therefore, only cases that represent a legitimate public interest are feasible.

Data requests are assessed on a case-by-case basis by data protection officers. The request has to be accompanied by a data protection impact assessment (DPIA). In this DPIA benefit for the public as well as the impact on the privacy of the end-user should be well-defined. Furthermore, the amount of data requests and data processing has to be reasonable and fair; meaning the least amount of (privacy sensitive) data has to be gathered to reach the desired outcome. The data protection officer has to adjudicate these DPIAs, which is not straightforward. One of the complications is that usually the “least amount of data” cannot be well defined and is thus subject to discussion. Therefore, part of this paper includes an analysis of the benefits that using smart meter data can have on planning in the low voltage grids, including the effects on the amount of data processing. Gaining insight into the minimum amount of data required, will not only lead to minimising data transactions, but can also lead to increased governmental and public support for DNO driven smart meter data application. Additionally, it helps the data protection officer to make a well-founded choice on the necessity of a data request.

LV-NETWORK PLANNING

The planning of the LV-network can be divided in three parts: the creation of new LV-networks, the expansion of LV-networks and the reinforcement of existing networks. Most of the LV-networks in Europe have been constructed in the previous century and the rate at which new LV-networks are built is relatively low. Most of the LV-network analysis is thus done on existing networks. For these networks the usage of the consumers can be obtained by reading their smart meter data. With these data the exact load values for each household can be used in the analysis. When it comes to the planning of the distribution network, two time periods are important: the 10-minute time interval (the base for the norm describing the allowable voltage variations [5]) and the hour range (which governs the thermal processes in a cable under normal loading conditions [6]). These time intervals do not coincide with the available measurement data from smart meters.

As the LV-network consists mostly of underground cables, the network will be built with a lifespan of 30 to 40 years in mind. However, the load within the LV-network can change significantly within a much shorter time period. To illustrate this point, the smart meter data from a neighbourhood in The Netherlands (de Laar, Arnhem) has been evaluated. The neighbourhood has been built in the 1980s and consist of similar houses and the socio-economic diversity is low. To analyse these smart meter data, a random selection of 20 households was taken from the neighbourhood. From this selection, the average contribution of a single household to the 10 highest peaks in the LV-network during one year, is calculated. This is done for 10000 random selections in order to generate the violin plot as shown in Figure 1.

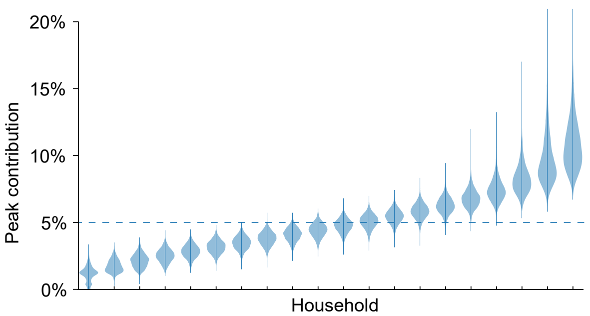

Figure 1: *The probability density functions of the peak contribution of 20 randomly selected households.*

In this figure, the probability density distributions of each of the households are shown, with a thicker line indicating a higher probability. In this figure the dotted line indicates the 5% peak contribution. This is the amount of peak contribution which can be expected if the peak contribution was equally distributed. From the figure it can be seen that about a third of the households contribute less to the peak, a third more or less contributed to the peak as expected and a third contributed more to the peak. These differences can be attributed to the differences in appliance use/behaviour of the households. Analysing a network with the current smart meter data can thus be dependent on the behaviour of a few high peak contributing households. This would not be a problem except that the energy usage within a house is not constant over decades: in 1-year time on average 3.5 out of every 100 inhabitants will have changed [7]. This can be due to people moving, moreover other changes like changes within the household composition (for instance the birth of a child) can also have a large effect on the energy usage within a household [8]. The direct use of smart meter data for network planning purposes can therefore lead to unrealistic results. Regularisation of the data through the modelling of the household load is therefore necessary. The modelling of the household load can however be greatly improved through the use of smart meter data.

TIME RESOLUTION

The time resolution of the smart meter data indicates how often a measurement value is obtained from the meter. To gain insight into how this can affect the network planning process, measurement data from 107 households with a 1-minute resolution from the CLNR project in the UK [9] is used. For these measurements the peak load within the network can be estimated for different time resolutions. This is done by averaging multiple data points for lower sampling rates.

The effect of the sampling rate on the measured peak load as well as the relationship between the various sampling times is shown in Figure 2. In Figure 2, the peak load for an increasing number of aggregated households is plotted for different sampling rates.

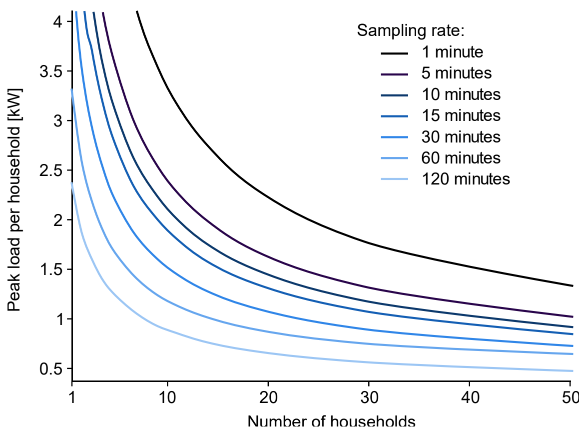

Figure 2: *The peak load per household versus the number of aggregated households for various sampling rates.*

From the figure it can be seen that with a higher sampling rate, a higher peak load is measured. This is as expected, as the shorter peaks will average out over the lower sampling rates. The figure also shows that the decrease in peak time follows the same general trend for the various sampling times. There are however differences between the curves for the different sampling times. When taking a better look at the differences between the 30-, 60- and 120-minutes sampling times at 10 aggregated households versus 50 aggregated households, it can be seen that the difference between the 30 and 60 minutes and between the 60 and 120 minutes at 10 households is almost equal while for the 50 aggregated households the 60-minute value lies much closer to the 30-minute value.

For the LV-network the time resolutions of 10-minute and 60-minute are most important. The stability of the relation between the measured time interval and the 10- and 60-minute interval is thus an indicator of how well the data can be used within the network planning process. When using data with a time resolution smaller than 10-minute the peak load at the required time resolutions can always be estimated, however if the peak at these time resolutions can also be estimated from lower sampling rates, smaller amounts of data needs to be handled. To gain more insight into the error if data with a high sampling rate is estimated based on data with a lower sampling rate, Table 1 is created. In the table the maximum and average error in the estimation of the peak load for various time resolutions is given. The estimation of the peak load is based on the following formula:

$$P_{est} = aP_{meas}^{b}$$

where, a and b are parameters, P_{meas} is the measured peak load and P_{est} is the estimated peak load.

Table 1 *The maximum and average (in parenthesis) absolute error in percentage of the estimated peak at a certain time resolution for various sampling rates.*

Sampling rate [min]	Estimated time resolution [min] → 5	10	15	30	60
5	-	-	-	-	-
10	1.1 (0.3)	-	3.4 (1.2)	-	-
15	6.1 (1.5)	4.6 (1.5)	-	-	-
30	11.7 (3.0)	10.0 (3.0)	4.4 (1.3)	-	-
60	15.0 (3.8)	13.2 (3.8)	7.1 (2.0)	2.2 (0.6)	-
120	18.5 (4.7)	16.5 (4.7)	9.8 (2.7)	4.5 (1.3)	2.0 (0.6)

From the table it can be seen that the relative average absolute error is small (<5%) in all cases, while the maximum errors can become much larger. Nowadays the 15-min values are often assumed to be a good indication of the voltage variations at a 10-min time resolution. With no adjustment this induces an average error of about 6.5% and a maximum error of about 9.7%. Using a 30-min sampling rate with a correction factor a lower error for the estimation of the 10-min time resolution can be obtained.

When looking at these figures the question remains whether the effects visualised here are still present when the energy transition takes shape. Therefore, Figure 3 has been created. This figure shows the same graph of peak load versus the number of households, only not for the basic household load, but also for households that utilize EV's and HP's.

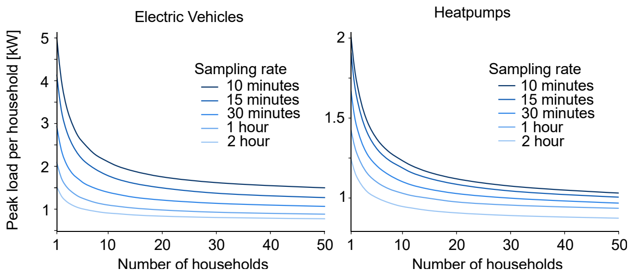

Figure 3: *The EV and HP peak load per household versus the number of aggregated households for various sampling rates.*

When looking at figure 3, the same shape can be seen as the general household load. Depending on the level of accuracy which is needed for the estimation of the peak load, a sampling rate can be selected.

MEASUREMENT DURATION

In order to get an accurate measurement of the peak loading of the network, generally one does not have to measure continuously. Only a measurement at the time of the peak loading is necessary. However, it is hard to estimate when the peak load will occur. The conventional wisdom is that the peak loading of the network will occur between 7 and 9 pm. Smart meter data has shown that this is not always true.

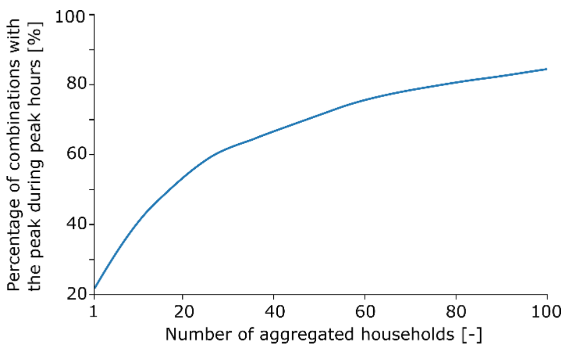

Figure 4: *The percentage of groups of households which has the peak load between 7 and 9PM for different household group sizes.*

Figure 4 shows the percentage of instances for which the peak loading occurs between 7 and 9PM for different number of aggregated households. It can be seen that for a small number of households the peak loading often does not occur at conventional peak hours, with just over 50% of the groups of 20 households having their peak between 7 and 9PM. For large groups of households these figures change considerable with almost 90% of the group of 100 households having the peak between 7 and 9PM. This shows that for medium voltage networks, one can safely assume the residential load peak will be between 7 and 9PM while this is not the case for LV-networks. Therefore, when assessing how long the measurement duration should be the measurements are assessed on a 24h bases.

An increase in measurement duration can give a better indication of the peak load of the network. As the estimated peak load of the network is dependent on the selected households from which the measurement data is used, a probabilistic approach to determining the effect of the measurement duration is taken. For different measurement durations and different number of households the peak load is calculated. To determine which part of the year should be selected, the days with the highest system peak load are selected and the measurements from these specific das are used. The system peak load is defined as the peak load of a group of 200 households. The different measurement durations are assessed with a 15-minute sampling rate.

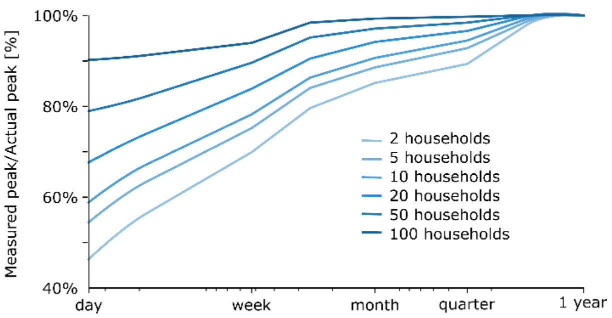

Figure 5: *The measured peak load over the actual peak load versus the measurement length for various sizes of household groups.*

In Figure 5 it is shown that the percentage error in the peak load increases substantially when the measurement duration is reduced to less than a month. The figure also shows that the estimation of the peak load of a small group of households generally requires a longer measurement duration than a large group of households. This is caused by the averaging effect of the load with larger groups of households. Next to estimating the peak load also the coincidence factor, the peak load of an individual household over the averaged peak load of a group of households, can be estimated. The results for these estimations are plotted in Figure 6.

From this figure it becomes apparent that the coincidence factor is overestimated if the measurement duration goes down, while the peak load is underestimated. Another interesting point is that the error in the coincidence factor is lowest for a small number of households and highest for larger groups of households. This can be explained from the fact that the peak load for a large group of households can still be reasonable well estimated while the estimation of the peak load of a single household will be inaccurate. While the peak load of an individual household will always be difficult to estimate, the peak loading of a small group of households will also be off as well and these errors seem to be correlated leading to a low error in the coincidence factor.

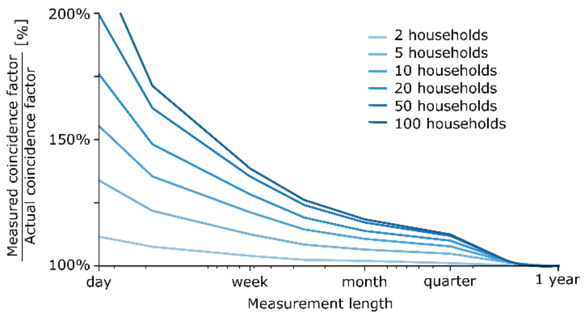

Figure 6: *The measured coincidence factor over the actual coincidence factor versus the measurement length for various sizes of household groups.*

SCENARIOS AND DISAGGREGATION

As assets such as cables and transformers have long lifetimes and are hard to replace, the electricity networks are planned and constructed for several decades (30 – 40 years) of usage. During this period the network should be able to sustain the peak loading that occurs. To account for the uncertain development in peak loads over these long timespans, scenario analysis is usually employed. In these scenarios diverse future situations are created (e.g. varying penetration levels of EV's, and HP's), and their effect on the peak loads is estimated. Measurement data from smart meters can be used to gain a better insight in the potential future peak loads for different scenarios.

Using measurement data, the current load of different types of customers can be modelled. Scenarios can then be applied by adapting the current load profiles to the future situation (e.g. by adding profiles of new technologies). To properly apply the measurement data disaggregation techniques can be used to identify individual appliances from load profiles [10]. This can be used to classify different customers, e.g. residential customers with a HP, and apply suitable load models in the scenario analysis. The accuracy of disaggregation techniques depends to a large extent on the sampling rate. Good results can be obtained using high resolution data, for lower resolution data the accuracy quickly falls and is usually limited to appliances with a high load or energy usage [11]. From a planning point of view these appliances are the most interesting, allowing for the use of disaggregation techniques.

By combining the load models of existing loads with scenario induced changes, the effects on the load profiles and peak loads can be obtained. For instance, it becomes possible to assess the effect on the load of neighbourhoods switching from a gas network to full-electric (i.e. with electric cooking and HPs), or the contribution of a certain number of EVs to the peak loads in the network.
To achieve dependable results, measurement data should be available from a diverse group of customers to provide a good representation of the population. This allows the classification of customers and applying suitable load models to different customer groups.

APPLYING SMART METER DATA TO NETWORK PLANNING

The previous section has shown that in order to incorporate smart meter data in the LV-network planning process, care must be taken with how the smart meter data is handled, in light of regulations and with respect to the actual application of the data. In the section on the measurement duration it is shown that measuring at slightly lower sampling frequencies (from 15min onwards) still allows for accurate estimation of the peak load at higher sampling frequencies. This relinquishes the DNO from the requirement of storing high frequency measurement data. In order to have an estimate of the peak load at a higher sampling rate an adjustment factor can be applied. To determine this correction factor occasional measurements with a high sampling rate are necessary. The data analysis also shows that a correction factor can be applied to gain an estimate on the loadability of a cable or for the PQ indicators time window based on a single sampling rate. This reduces the need for the use of performing load flow calculations with the different sampling rates. An average error in the peak load of 5% can be expected even with sampling only every 120 minutes.
For the measurement duration it has been shown that the peak load can be estimated with an accuracy of over 90% on a month of measurements a year. To achieve this level of accuracy the peak load should be used for larger groups of households and the coincidence factor in combination with a separately estimated individual peak load for smaller groups of households.
Measurement data disaggregation in combination with load modelling can provide a good base from which scenarios of specific appliances can be introduced. The household load model can be specified depending on household type, if measurement data from a diverse enough group of households is available.

CONCLUSIONS

Smart meter data can become an invaluable resource for the planning of LV-networks. Privacy regulations should be taken into account with the implementation of smart meter data in the LV-network planning process to address consumer privacy. A certain level of regularisation should always be applied to be able to not only correctly estimate the loading of the network at this exact moment in time, but also for the coming years. This regularisation also anonymises the data making its usage more acceptable. Different correction factors can be applied to limit the smart meter data requirements to a single sampling time. The acquisition of the data from the smart meter can be limited to one month each year without a large loss in peak load accuracy.

REFERENCES

[1] M. Nijhuis, M. Gibescu en J. F. G. Cobben, „Scenario analysis of generic feeders to assess the adequacy of residential LV-grids in the coming decades," in *IEEE PowerTech*, Eindhoven, 2015.

[2] P. M. van Oirsouw, Netten voor distributie van elektriciteit, Arnhem: Phase to Phase , 2011.

[3] B. Römerab, P. Reichhartb, J. Kranzc en A. Picotb, „The role of smart metering and decentralized electricity storage for smart grids: The importance of positive externalities," *Energy policy,* vol. 39, nr. 50, pp. 486-495, 2012.

[4] The European Parliment and Counsel, „on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," EU Regulation 2016/679, 27 April 2016.

[5] European Committee for Electrotechnical Standardization, „EN 50160: Voltage characteristics of electricity supplied by public distribution systems," 2010-07-01.

[6] International Electrotechnical Commission, „IEC 60853-1: Cyclic rating factor for cables up to and including 18/30 (36) kV," 2008.

[7] Statistics Netherlands, „Population dynamics; birth, death and migration per region," 2017. [Online]. Available: http://statline.cbs.nl/Statweb/selection/?DM=SLEN&PA=37259ENG&LA=EN&VW=T.

[8] J. Stephenson, B. Barton, G. Carrington, D. Gnoth, R. Lawson en P. Thorsnes, „Energy cultures: A framework for understanding energy behaviours," *Energy Policy,* nr. 10, pp. 6120-6129, 2010 Oct 1.

[9] Customer-Led Network Revolution, „Dataset (TC2a): Enhanced Profiling of Domestic Smart Meter Customers CLNR-L074," 2014. [Online]. Available: http://www.networkrevolution.co.uk/project-library/dataset-tc2a-enhanced-profiling-domestic-smart-meter-customers/.

[10] M. Zeifman en K. Roth, „Nonintrusive appliance load monitoring: Review and outlook," *IEEE Transactions on Consumer Electronics,* vol. 57, nr. 1, pp. 76-84, 2011.

[11] B. J. Birt, G. R. Newsham, I. Beausoleil-Morrison, M. M. Armstrong, N. Saldanha en I. H. Rowlands, „Disaggregating categories of electrical energy end-use from whole-house hourly data," *Energy and Buildings,* nr. 50, pp. 93-102, 2012.