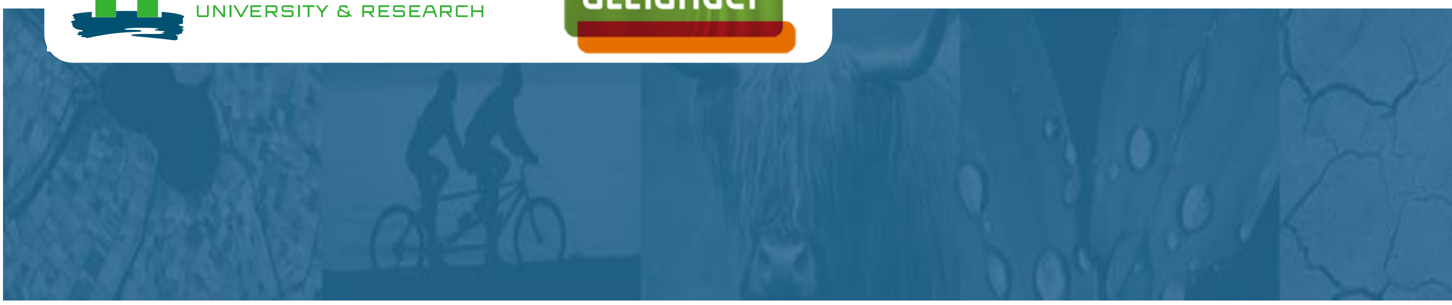
Geo-information Science and Remote Sensing

Thesis Report GRS-2018-44

Modeling the spatial economic efficiency of power distribution networks

Bas Wiltenburg
August 27th, 2018

Confidential information and results were removed in this report.



Modeling the spatial economic efficiency of power distribution networks

A research contributing to the deployment of SmartGrid and MicroGrid applications, with the aim of lowering the costs of the overall energy supply system.

Bas Wiltenburg

Registration number 940428960120

Supervisors:

dr.ir. S (Sytze) de Bruin
ir. M.H.G.I. (Matthijs) Danes

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

August 27th, 2018
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2018-GRS-2018-44
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Preface

This research was performed at the request of and in collaboration with Alliander. Alliander is one of the three main utility companies in the Netherlands and responsible for about one third of the Dutch regional power distributions networks. Because this thesis report contains of confidential and business-sensitive information, dissemination of this report is only possible with approval of Matthijs Danes, the person concerned on behalf of Alliander.

Acknowledgement

I would like to thank Alliander for giving me the opportunity for doing my thesis in and for this organisation. It was a good time in which I have learned many new things about a domain which was unknown for me at the moment I started my thesis. I would like to thank my supervisor Matthijs Danes in particular for his enthusiasms, guidance and involvement for this thesis project. Your enthusiasm kept me motivated during the entire research period. Besides, I also want to thank Dennis Briene, Robert Bennink, Martin Stronks and Aart van den Dolder for their contribution to this thesis. Moreover, I would like to acknowledge my supervisor Sytze de Bruin from the Wageningen University for his guidance and rapid and good feedback. Additionally, a special thanks to my partner Liliana, for her great support in this period.

Abstract

Today’s power system of the Netherlands is characterized by the growing introduction of electric vehicles (EV), heat pumps and distributed generation sources (DG). For grid operators, this causes the need for investment increases, which will result in higher costs of the overall energy supply system and consequently higher tariffs for the customers. Nonetheless, distributed generation sources also introduced new and innovative concepts such as MicroGrid and SmartGrid. In these concepts, customers no longer need to be connected to the national transmission grid and using electric energy generated at the consumption side of the power supply chain. By using MicroGrid and SmartGrid applications, the total costs of the overall energy supply system can be reduced by decreasing network utilization, reducing network losses and prevent new investment in feeder lines, transformers and substations not need to be upgraded or replaced. In order to keep the costs of the grid network as low as possible, this thesis contributes to the problem where SmartGrid and MicroGrid concepts may decrease the total costs of the existing energy supply system.

In this research, a method is proposed which calculates the spatial economic efficiency of the existing grid network in order to determine which parts of the system decrease its total economic efficiency and consequently may be potential for MicroGrid applications. This method involves a network analysis for each customer in the grid network and allocates asset costs to asset dependent customers. The network analysis uses a network graph which includes each individual distribution asset of the grid network. By allocating the grid network costs to asset dependent customers, a ‘heat map’ of relative expensive and less expensive areas of the grid network is derived. By using an iterative greedy algorithm, the economically inefficient areas of the original energy supply system are identified, which at the same time can be applied as MicroGrid system as well.

The costs of the grid network are allocated to the customers based on two different cost allocation methods. The first allocation method distributes the costs of a specific asset linearly over its dependent customers while the second allocation method allocates the costs based on the extent-of-use of the asset dependent customers. A scenario analysis is performed in order to assess the impact of applying the computationally less expensive linear cost allocation method. In addition, the scenario analysis also assesses the impact of using not specific asset costs, as administrated by the financial department of Alliander. The method was implemented on the power distribution network supplied by the high-voltage substation ‘Hoorn Holenweg’, located in the municipality of Hoorn, the Netherlands. Results of the method showed that economically inefficient areas were only found in rural and sparsely populated areas. Moreover, the use of not specific asset costs caused a substantial variability in the method’s output. The cost allocation method, however, did not have a significant impact on the calculated economically inefficient areas.

Keywords

Power distribution networks, Cost allocation, Cost distribution, Economic efficiency, MicroGrid, Graph database

Content

List of figures

List of tables

List op appendixes

.

Glossary

Circuit breaker: a circuit breaker is an asset in the grid network that is used to interrupt the current flow.
Civil costs: costs incurred by a contractor for labour, material and equipment.
Distributed generation: refers to the generation of electric energy by smaller-scale facilities connected to the medium-voltage and low-voltage distribution networks.
Distribution board: divides an electrical power feed into multiple circuits.
EAN: is an abbreviation of European Article Number, what refers to an unique identification number for each electricity- and gas connection in Europe.
EV: is an abbreviation of electric vehicle
Feeder line: a power section transferring power from a high-voltage substation to a transformer station.
Fishbone structure: an outgoing distribution network that does not interconnect with any other network.
Grid operator: company which supplies electricity and operates and maintains the power distribution system.
High-voltage substation: transforms high voltage (50kV) of the transmission system to medium voltage (10kV) of the power distribution system.
HV: abbreviation of high voltage.
Joint: is where cables interconnect.
Linear cost allocation method: distributes the asset's cost linearly over its dependent customers.
Load-based cost allocation method: distributes the asset's cost based on the relative load-contribution of its dependent customers.
Loop structure: power distribution networks are mutual interconnected in loops.
Low-voltage cable: cable in the low-voltage distribution network (230V)
LV: abbreviation of low voltage.
Main power section: a set of aggregated cables which together constitute to a specific route in the network.
Medium-voltage cable: cable in the medium-voltage distribution network (10kV)
MicroGrid: is an electrical distribution network comprising various distributed power generators, storage devices and controllable loads that can operate either in an interconnected way or isolated from the utility service provider as a controlled entity (De Leone, 2017).
MV: abbreviation of medium voltage.
Neo4j: open source graph database.
SmartGrid: is a model of energy management in which the users are engaged in producing energy as well as consuming it while having information systems fully aware of the energy demand-response of the network and of dynamically varying prices (Pagani & Aiello, 2016).
Transformer station: transforms medium voltage (10kV) to low voltage (230V) in a power distribution network.

1 Introduction

1.1 Context and background

The current grid network of the Netherlands is designed based on an energy system where the generation of energy takes place at locations where it is more economical and efficient to generate (Provoost, 2009). Over the years, this traditional system of centralized electric energy generation has resulted in large-scale centralized power plants and increased distances between generation and consumption. Because of this, the electricity network of the Netherlands is made up of a high-voltage transmission network and regional distribution networks respectively connecting the large-scale power plants to substations and the substations to the customers. In this way, a one-directional power supply chain has formed in which electric energy is transported from central generation points to customers (Provoost, 2009).

However, energy technologies are changing fast. Due to rapid technological developments concerning storage and distributed generation, the connection to the national transmission grid becomes no longer a requirement in order to provide a customer with electricity. Distributed generation refers to the generation of electricity by smaller-scale facilities connected to the medium-voltage and low-voltage power distribution networks (Jasemi, et al., 2016). In the Netherlands, this integration mainly concerns of private wind- and solar energy (Bosman, 2012), which is encouraged to adapt the existing fossil-based electricity system to a new one, which is more environmentally friendly and economically sustainable. However, these changes have a significant impact on the utilization of the existing grid network. After all, the power flow in the system becomes multi-directional since energy is generated on the consumption side of the power supply chain as well.

Thence, the integration of distributed generation has introduced two new and innovative conceptual applications, namely: SmartGrid and MicroGrid. The concept of SmartGrid is to control power flow, support electrification from existing equipment (such as electrical vehicles and mobile appliances) and monitor and modify energy spending patterns of actual consumption in real time in homes and buildings (Kyriakopoulos & Arabatzis, 2016). In a more visionary acceptation, the SmartGrid is a model of energy management in which the users are engaged in producing energy as well as consuming it while having information systems fully aware of the energy demand-response of the network and of dynamically varying prices (Pagani & Aiello, 2016). The MicroGrid concept slightly differs from the SmartGrid concept. A MicroGrid is an electrical distribution network comprising various distributed power generators, storage devices and controllable loads that can operate either in an interconnected way or isolated from the utility service provider as a controlled entity (De Leone, et al., 2017). In other words, a MicroGrid is a smaller low-voltage, or in some cases medium-voltage, grid network with resources, storage and controllable loads with a total installed capacity in the range of a few kWs to a couple of MWs (Anastasiadis, et al., 2018). According to Raju, et al., (2017), these new conceptual applications are positioned to hold a critical part in bringing into practice the large-scale implementation of distributed energy resources in all working conditions, for both grid-connected as well as off-grid modes.

1.2 Problem definition

Parallel to the upcoming integration of distributed generation units, today's power systems are characterized by the growing introduction of electric vehicles (EV) and heat pumps connected to the distribution networks as well (Soares, et al., 2015). Globally, the last years have seen a rapid increase in the numbers of electric vehicles, rising from a few thousand in 2009 to some 740.000 by the end of 2014 (Galvin, 2017). The establishment of charging stations imposes an additional burden on the power grid, as the high charging loads of fast charging stations will degrade the operating parameters of the distribution network (Deb, et al., 2018). Many researchers demonstrate the adverse impact of EV charging loads on different parameters of the distribution network like voltage profile, harmonics and peak load (Deb, et al., 2018). This comes in addition to the effects caused by distributed generation units. After all, the increasing amount of distributed generation installed in the grid may also cause the voltage to rise to unacceptable levels during periods of high generation and low consumption (Weckx, et al., 2015). For keeping the grid voltage within limits, both the demand increase as well as the distributed generation might cause significant problems. Research performed by Reza (2006) indicates that from the transmission system stability point of view, if higher energy consumption and generation levels are coming up, sufficient voltage support must be installed. For the grid operators this would imply that, when current developments continue, required capacity expansions will cause significant economic impacts on the costs of the existing power distribution networks. In order to keep the costs of the grid network as low as possible, this thesis contributes to the problem where SmartGrid and MicroGrid concepts may decrease the total costs of the overall energy supply system.

1.2.1 Missing information for SmartGrid applications

One of the concepts of SmartGrid is the application of a real-time locational marginal price system to control the grid voltage and avoid damage to the grid network. With locational marginal prices, the grid operator can steer the reactive power consumption and active power curtailment of distributed generation units to guarantee a safe grid operation (Weckx et al., 2015). In such a system, generators or loads that locate in a manner that reduces line loading or uses fewer assets, are allocated lower costs (Sotkiewicz & Vignolo, 2007). This means that customers are charged with prices that can vary over short time intervals, in which the price depends on the contribution of the customer to the (total) grid network's expenditures at the specific time step. According to De Oliveira-De Jesus & Antunes, (2018), marginal pricing in power systems could lead to a reduction of energy use, lower energy losses and alleviate congestion in transmission lines. This implies that customers would be exposed to lower energy prices (De Oliveira-De Jesus & Antunes, 2018). However, in order to determine the user's contribution to the total grid network expenditures at each moment in time, costs need to be specified on a level of individual assets. Besides, the costs arising from each asset should be allocated to the customers based on the extent the customers actually use the specific assets. By doing this, a 'heat map' of costs for each part of the distribution grid can be derived, showing which spots are costing the most money (St. John, 2014). This information can then be used in SmartGrid applications where dynamic electricity prices help grid operators to control the grid voltage.

Moreover, the derived information of assigned costs to customers can also be used in order to determine which parts of the grid network have most potential for MicroGrid applications. After all, areas which are very expensive in the conventional energy supply system may become less expensive when the energy supply system functions as an own entity, isolated from the national transmission grid. This research is specifically focusing on modeling the spatial economic efficiency of assets and customers in order to determine the areas where MicroGrid concepts may lower the total costs of the overall energy supply system. Hence, the method proposed in this research localizes the areas in the original grid network which decrease its total economic efficiency and at the same time can be applied as MicroGrid as well. However, in order to conclude whether the customers in uneconomic areas should actually be transferred to a MicroGrid system, the method has to include the costs of MicroGrids as well. Besides, you have to calculate how much MicroGrids save costs in power plants not built or feeder lines, transformers and substations not upgraded or replaced in order to balance out which combination of technologies and strategies actually result in the most optimal system (St. John, 2014). However, in this research, the MicroGrid costs as well as the impact on feeder lines, transformers and substations are not assessed. This means that the modeled MicroGrid areas are based on an economic optimization of the existing energy supply system only. The problem whether economically inefficient areas should actually be transferred to a MicroGrid system is not solved in this research.

1.2.2 Scope of this research

Due to the shared nature of the grids, the calculation of the customers grid network costs can be rather complex. After all, the cost for providing a service to one user depends on the services being provided to other users, as well as on how users are utilizing the system (Parsons & Sakhrani, 2010). Moreover, costs need to be determined on a level of individual assets. In order to calculate the customers costs at each moment in time, each individual customer should be allocated to moment-related costs arising from the capital, depreciation, operating and energy loss expenditures of all its dependent assets. However, due to limited time and absent real-time data, this research only focuses on the allocation of depreciation costs to customers on an annual basis. The determination of capital, operating and energy loss expenditures arising from each specific asset at each moment in time, are outside the scope of this research. Therefore, the developed method should be seen as a part of a much larger entity in the future, which also has to include the capital, operating and energy loss expenditures as well as real-time cost allocation and a solid optimization strategy calculating the most optimal energy supply system.

In this research, the depreciation costs of assets are based on data received from the financial department of Alliander. This data consists of average costs of assets for each valuation year. This means that the costs of assets as administrated by Alliander are far from specific and does not reflect the costs which actually has been incurred in order to buy and install the specific distribution asset. According to Europacable, (2014), it generally can be said that about 1/3 of the investment costs derive from the cost of the cable and up to 2/3 will derive from the cost of installations, notably civil works. Therefore, the use of average costs exclude the spatial variation of civil costs, which in reality is very unlikely. In order to assess the variability of the method's output to the average asset costs as administrated by Alliander, a sensitivity analysis is performed in which the costs of assets are differentiated and corrected for its topographical location. When the outcome appears to be very sensitive for the spatial corrected asset costs, a conclusion is that the use of asset costs as administrated by the financial department of Alliander has a large impact on the robustness of the method and the reliability of the results. Furthermore, the sensitivity analysis also includes the cost allocation method. After all, various methods have been proposed in literature for allocating grid network costs to the customers (Picciariello, et al., 2015). Because costs of the grid network can be allocated to the customers in many different ways, two different cost allocation methods are compared in order to assess the variability of the method's output to the cost allocation method applied. When the locations of economically inefficient MicroGrid areas appear to be very sensitive to the cost allocation method applied, a conclusion is that a precise and well-considered calculation of the user's contribution to the total utility's costs is very important.

1.2.3 Potential applications supported by the method

The method proposed in this research also supports other potential applications related to the deployment of SmartGrid and MicroGrid concepts. The method developed in this research is able to present the spatial economic efficiency of the grid network on any cross-section and is able to assess the economic impact of a conventional capacity expansion in the grid network. Although the method does not include real time cost allocation, it is capable of adding distributed generation sources to the grid network. This is a relevant application for assessing whether the addition of a distributed generator can prevent a capacity expansions in upstream feeder lines or transformers. In this research, potential applications of the method are exposed by calculating the economic impact of a capital investment and an added distributed generation source.

1.3 Research objective and research questions

The objective of this thesis is to develop a method for identifying both economically efficient and economically inefficient areas in power distribution networks, where the latter can be applied as MicroGrid as well.

To reach this objective, four research questions are answered:

1. What are suitable cost allocation methods for distributing costs of an electric power distribution network from an economic perspective?
2. How can available data and the cost allocation method be integrated into a model determining the economic (in)efficient locations in the service area of Alliander?
3. Where are the economic efficient spots in the distribution networks of Hoorn Holenweg and where are potential areas for MicroGrids?
4. What is the sensitivity of the results to variations in costs and cost allocation?

2 Literature review

2.1 Cost allocation methods

Modeling the economic efficiency of a power distribution network requires allocating the grid network costs to the customers. According to Bonbright, et al., (1961), cost allocation encompasses several principles including cost recovery, transparency, simplicity, stability, equity and cost causality. The first five of these principles are related to consumer protection principles while the last criterion is referring to an economic efficiency principle. The cost-causality based allocation principle implies that the allocation of costs should accurately reflect each network user's contribution to the total network costs (Picciariello, et al., 2015). Hence, in order to determine the economic efficiency of a specific customer, this principle has to be applied.

Causality-based allocation methods are well studied in literature. Various studies have been done focusing on the allocation of shared network costs to users. Vignolo, (2007) developed the 'Amp-mile method', which calculates an 'extent-of-use' for each distribution asset in the grid network. The 'extent-of-use' can be defined as the load's or generator's impact on a distribution asset relative to the total flow on that specific asset for each moment in time. By allocating costs depending on the extent-of-use, cost allocation is largely determined by the location of a specific injection or withdrawal in the grid network. Generators or loads that locate in a manner that reduces line loading or uses fewer assets, are allocated to lower costs (Sotkiewicz & Vignolo, 2007). Because assets consist of relative contribution-factors of all users at all times, a locational charge for each individual customer can be obtained.

Another method, developed by Li & Tolley, (2007), uses the unused capacity of an existing network to reflect the cost of advancing or deferring future investments as a consequence of adding generation or load at each node of a distribution network (Picciariello, et al., 2014). For network assets that are affected by the injection or load there will be a cost associated with accelerating the investment or a benefit associated with its deferral (Wang , et al., 2018). Based on a fixed growth rate of demand, the method calculates the remaining years until an asset no longer satisfies the required capacity, calculates the present value of the future investment needed and compares this with the outcome of the same calculation including incremental additions of power injections at specific nodes in the grid network.

A third method is the use of so-called 'Reference Network Models'. This method has similarities with the method developed by Li & Tolley, (2007) but determines an optimal network capacity (reference network) through an optimization process where required capital costs and annual network operating costs are traded off (Strbac & Mutale, 2005). In this method, the obtained optimal reference network (modeled based on a prediction of future variations of load and generation) is used in order to compare the optimum capacities of the reference network with the capacities of the existing grid network. By comparing the reference network with the existing system capacities it is possible to identify areas of over- and underinvestment (Strbac & Mutale, 2005). The allocation of costs depends on the forward-looking investment costs in which each customer is classified as either demand-dominated or generator-dominated. When the maximum demand-minimum generation is being the critical condition, generated dominated customers clearly reduces the demand for distribution network capacity and therefore are rewarded with lower cost allocation (Strbac & Mutale, 2005). Demand dominated customers, on the other hand, accelerate the investment and therefore will be assigned higher costs.

2.1.1 Applicability of theoretical methods

The use of Reference Network Models is not considered to be suitable for this research. First of all due to the complexity involved in determining the optimum reference network, which is outside the scope of this research. Secondly, because it considers a most critical condition: when maximum demand-minimum generation is being the critical condition, a generator-dominated customer is allocated lower costs. However, when the other condition is almost as critical as the most critical one (but at a different time), the generator-dominated customer affect the distribution network almost as much but is not allocated with the same costs. In today's distribution networks of the Netherlands, where both conditions can be critical during the same day, this cost allocation method hence does not reflect the utility's capital and operational expenditures very precisely.

The second method of Li & Tolley, (2007) allocated costs based on the impact of a customer on the total unused capacity of a specific distribution asset. Besides, it takes into account the additional increment of generation and

load, which means that the method allocated costs based on the capacity a customer may need in the future, but currently not uses. Hence, this method is effective when modeling the economic efficiency of power distribution networks in which the adequacy of the distribution network is relevant as well.
In relation to this research, main problem of this method as well as the 'Amp-mile method' developed by Vignolo, (2007), is the fact that their application to medium- and low voltage grids is rather complex since facility by facility calculations are needed (Picciariello, et al., 2014). Moreover, to be able to determine the user's relative impact on a distribution asset at all times, energy generation and consumption data of all users at all times need to be available. However, the problem is that at this moment customers are not obligated to share their real time energy data for analyzing purposes. This makes the implementation of reliable real time load-flow calculations difficult. Due to this problem and because real-time facility-by-facility calculations cause extreme computation times, it is considered to be too time-consuming to fit into the scope of this research.
The cost allocation method applied in this research therefore calculates a static 'extent-of-use' of each customer for each asset and hence is based on a static energy supply route between customer and substation. However, this assumes that customers are supplied by the high-voltage substation only, which means that in this research, distributed generation units do not affect the customer's energy supply route.

2.2 Calculation of energy supply routes

In order to calculate the customer's costs, it must be known which assets are actually used by each specific customer connected to the grid network. However, the contribution of a customer to the total cost of an asset is different for all assets. After all, each asset provides electricity to a different number of customers. Hence, the cost for providing electricity to one customer depends on the supply of electric energy the asset provides to other customers (Parsons & Sakhrani, 2010). This means that, in order to obtain this information, a network analysis must be performed for each individual customer, calculating the route the current flow travels to provide the customer with electrical energy. By performing a network analysis for each individual customer, customers can be linked to the assets which are part of the customer's energy supply route and assets can be linked to the customers whom actually use the specific asset. In the context of this research, this means that the ability to efficiently perform network analyzes over a large number of assets is very important.
The execution of network analyzes require a network that includes all levels of the system, from the high-voltage substation to the low-voltage customers. The energy supply route of each customer can then be computed by querying the mutual relations of all network assets interconnected with the specific customer. Due to the fact that the network consists of many assets as well as many customers, an efficient and fast querying of mutual relations is required.
At this moment, the grid network topology of Alliander is stored in relational databases only. However, according to Medhi & Baruah, (2017), querying the interconnectivity of objects can be rather complex when using a relational database. To relate one information to another, a foreign key is necessary to join an object in a specific table with another object. To infer relationships in a relational database, multiple joins are required (Yoon, et al., 2017). However, in a relational database, this will become soon too computationally expensive (Medhi & Baruah, (2017). In case of a network with more than 60.000 assets, a relational database may not be optimal to deal with.
In contrast, various studies show that a graph database seems to be more suitable for highly connected data. A graph database uses nodes and edges (relationships) to represent and store data. Each node represents an entity and each edge represents a specific relationship between two nodes (Yoon, et al., 2017). This structure allows it to store particular objects with relations between them (Medhi & Baruah, (2017), which ensures that no expensive joins are needed. According to Miller (2013), the relational database model is optimized for aggregated data while the graph database is optimized for highly connected data. This is confirmed in Medhi's research, in which the performance of querying connected data in a graph database is compared with the performance of the same query in a MySQL database. Results showed that, for retrieving connected data, the graph database gives much better results. Hence, chosen was to build the required network, which has to include all assets between the substation and the customers, in a graph database structure. In order to create the network in a graph database structure, the open source database software Neo4j was used.

3 Study area and data

3.1 Study area

This research was performed for all power distribution networks supplied by the high-voltage substation of Hoorn Holenweg. This substation supplies electric energy to most of the inhabitants of the municipality of Hoorn, the Netherlands. The grid network chosen in this research includes urban, industrial as well as rural areas. An overview of the grid network of Hoorn Holenweg is presented in Figure 1.

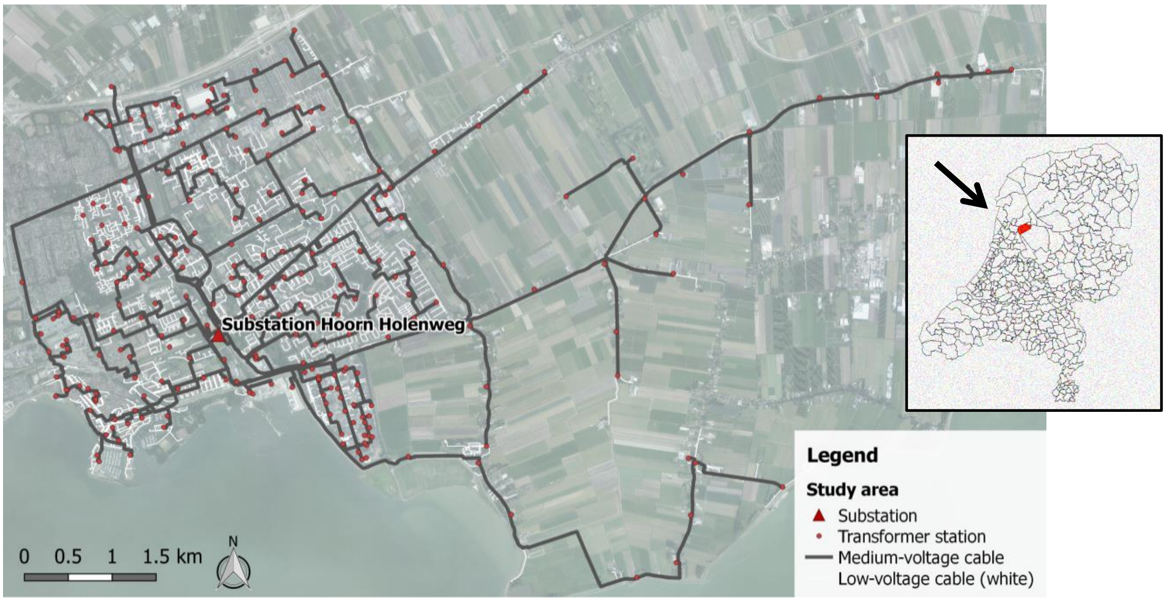


***Figure 1**: Study area.*

Besides the fact that the grid network of Hoorn Holenweg covers various topographical areas, it also includes the two most common forms of power distribution network design: i.e., the *fishbone structure* and *loop structure*. Figure 2 presents both design principles. In the loop structure, distribution networks are mutual interconnected in loops (Europacable, 2014). This means that distribution networks are connected with other distribution networks by closed circuit breakers. This may apply to both medium-voltage networks as well as low-voltage networks. In case of a breakdown on one branch, low-voltage customers can be supplied through another branch from another transformer with very little interruption of the service. In a fishbone structure, a distribution network does not interconnect with any other network. Hence, this structure is characterized by an overall shorter cable infrastructure. Most of the distribution networks are based on this latter principle.

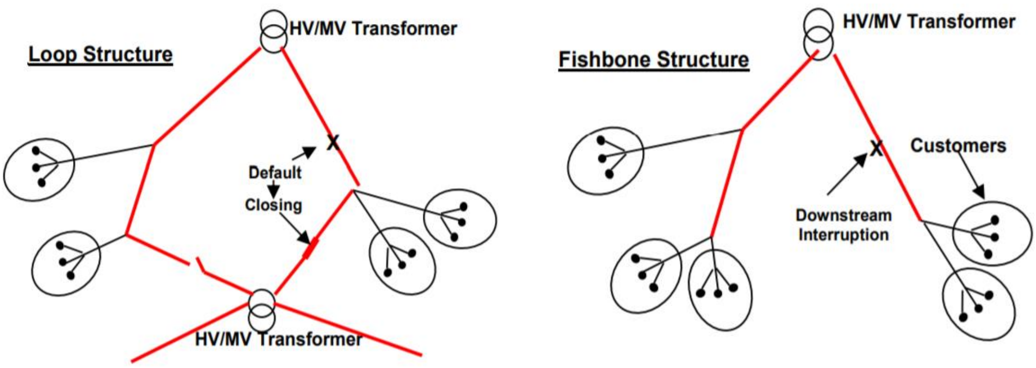


***Figure 2:** Loop structure (left) and fishbone structure (right) design principles of power distribution networks (source: Europacable, 2014)*

3.2 Data

3.2.1 Network data

In total, 22108 customers (both wholesale customers and normal households) are connected to the grid network of Hoorn Holenweg. In the context of this research, a digital network graph was built of the grid network of Hoorn Holenweg which included the substation and all its downstream medium-voltage cables, station constructions, installations, transformers, low-voltage cables, distribution boards and customers. In order to build the digital network, various databases were used that store the different components of the physical grid network. Besides the actual assets which are subject to depreciation costs, the network graph had to include the configuration of the system as well. This configuration models whether the current flow between two transformer stations can pass, or is interrupted. The normal mode of circuit breakers determines the configuration of the grid network. Hence, the location and normal mode of circuit breakers were used to enrich the network with its normal configuration. An explanation of how the network was built, and what conversions and integrations were carried out, is included in Appendix I.

3.2.2 Asset cost data

In order to determine the annual depreciation costs of each asset in the grid network of Hoorn Holenweg, it must be known how the grid network's assets are activated on the financial statement of Alliander. In the financial administration of Alliander, seven different types of assets are distinguished: medium-voltage cables, low-voltage cables, medium-voltage installations, low-voltage installations, transformers, station constructions and distribution boards. Nonetheless, the grid network consists of numerous other assets as well, such as joints, circuit breakers, terminators, security components and ICT-operating systems. In the financial administration of Alliander, the cost of most of these assets are integrated into the cost of cables.
The cost of cables are administrated as a cost per meter for each valuation year and include the average civil costs which had been incurred by the utility company as well. Because costs are administrated as an average cost per meter, each cable of the same length from the same year has the same cost, independent of the capacity and material of the cable and the civil costs which actually has been incurred. The only distinction in cost is related to the fact whether a cable is a medium-voltage or low-voltage cable of a single-core or multi-core type. Single-core cables are valued at 1/3 of the average multi-core cost per meter.
The cost of a transformer station is based on the average costs of the typical components of a transformer station. Hence, the average cost of the building, a medium-voltage installation, a low-voltage installation and a transformer constitute to the total cost of a transformer station. These average costs are administrated for all valuation years. The cost of a specific transformer station hence is based on the valuation years of the specific components only, in which no distinction is made between, for example, the actual size of the installations.
However, not all assets are part of the asset population subject to depreciation. Assets not paid by Alliander are not activated on the financial statement. Power grids on private land (customer's cables) usually do not have depreciation costs, and neither do power grids laid out for public lightning and transformer stations installed for wholesale customers. The financial department of Alliander maintains a database with all assets and contains several columns describing the assets financial status. This database was used in order to determine whether an asset was subject to depreciation costs or not. Another database was used to retrieve the relevant asset's properties, such as the length, valuation year and the number of cores. Guidelines in Alliander's accounting manuals were used for calculating the actual annual depreciation of each asset, which was dependent on the lifespan and residual value of the asset.

3.2.3 Customers data

The tariffs paid by the customers are regulated and hence were determined by using public tariff sheets available on the website of Alliander[1]. The tariff paid by a customer consists of a tariff for the connection, a tariff for the transport of energy and a tariff for the measuring services. For normal customers, the tariff depends on the nominal capacity of the grid connection only. Hence, the nominal capacity of the customer's grid connection was used in order to determine its annual revenue.

For wholesale customers, however, the transportation tariff is based on their actual energy consumption. This means that the determination of revenues requires the monthly energy consumption of the customer during high tariff hours and low tariff hours, its peak consumption and its contracted transport capacity. For all wholesale customers connected to the grid network of Hoorn Holenweg, these data were provided by the customer's administration department of Alliander. Based on these data and by using the regulated tariff provisions for wholesale customers, the annual tariff of each wholesale customer was computed.

Besides the tariffs paid by the customers, the method proposed in this research used the annual energy consumption of customers to allocate the depreciation costs of the grid network. The annual energy consumption of each customer and the nominal capacities of the customer’s grid connections were extracted from a specific customers database.

[1] https://www.liander.nl/uwtarieven

4 Methods

4.1 Allocatiing grid network costs

The cost-causality based allocation method implemented in this research defined a static extent-of-use of each customer for each asset. To examine the robustness of the method with regard to the cost allocation method applied, two different cost allocation methods were compared: a linear cost allocation method and a load-based cost allocation method was applied.

4.1.1 Linear cost allocation method

The linear cost allocation method distributed the asset's depreciation cost linearly over the customers whom actually used the specific asset. Formula 1 calculates the customer's cost arising from one specific distribution asset. To calculate the total annual cost of a customer, this formula was applied to all assets which were part of the customer's calculated energy supply route. Hence, in order to calculate the annual total costs of all customers, for each unique asset in the grid network was calculated which and how many customers were dependent on the asset (see also section 4.2: network analysis). Figure 3 shows an example of the calculation of the total cost allocated to a specific customer. In this figure, the costs besides the nodes represent fictive annual depreciation costs of assets and the numbers inside the nodes represent the number of customers whom using the specific asset. The arrows represent the direction of the current flow.

$$C = \sum_{i=1}^{n} \left(\frac{d_i}{a_i}\right) \qquad (1)$$

where

C is the total allocated cost to a customer [€]
n are all the assets used by the specific customer
d_i is the annual depreciation cost of asset i [€]
a_i is the total number of customers whom using asset i

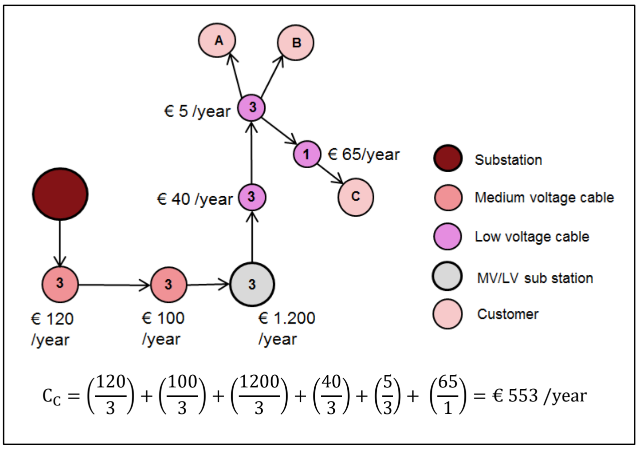


***Figure 3:** Sample calculation of the customer's allocated cost based on the linear cost allocation method.*

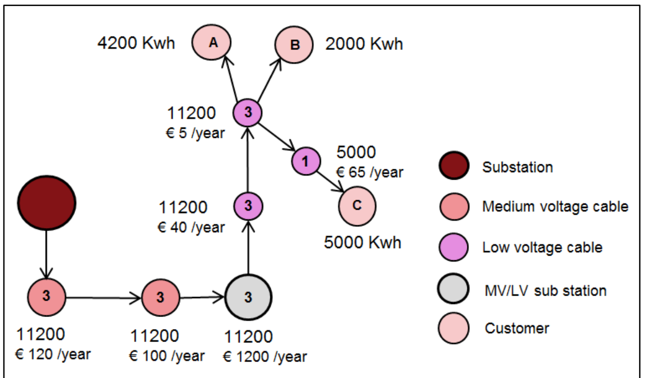
4.1.2 Load-based cost allocation method

In the load-based cost allocation method, the assets costs were distributed to the customers based on the relative load-contributions of the customers to the total annual energy load on the asset. The total annual energy load of an asset was calculated by summing the annual energy consumption of all the asset dependent customers. An extent-of-use of each dependent customer was calculated by dividing the annual energy consumption of the customer by the total annual load on the asset. Subsequently, the customer's cost arising from the asset was calculated by multiplying the relative load-contribution of the customer by the depreciation cost of the asset. To calculate the total annual cost of a customer, Formula 2 was applied to all assets which were part of the customer's calculated energy supply route. Figure 4 shows an example of the calculation of the customer's cost based on the load-based cost allocation method. In comparison with Figure 3, the additional numbers beside the nodes represent the total annual energy loads on the assets.

$$C = \sum_{i=1}^{n} \left(\frac{p}{P_i}\right) * d_i \qquad (2)$$

where

C is the total allocated cost to a customer [€]
n are all the assets used by the specific customer
p is the annual energy consumption of the customer [Kwh]
P_i is the total annual energy load on asset i [Kwh]
di is the annual depreciation cost of asset i [€]



$$C_C = \left(\left(\frac{5000}{11200}\right) * 120\right) + \left(\left(\frac{5000}{11200}\right) * 100\right) + \left(\left(\frac{5000}{11200}\right) * 1200\right) + \left(\left(\frac{5000}{11200}\right) * 40\right) + \left(\left(\frac{5000}{11200}\right) * 5\right) + \left(\left(\frac{5000}{5000}\right) * 65\right) = € \, 719 \text{ /year}$$

Figure 4: *Sample calculation of the customer's allocated cost based on the load-based cost allocation method.*

4.2 Network analysis

4.2.1 Building network graph

A network analysis was performed for each individual customer connected to the grid network of Hoorn Holenweg. The network on which the network analyzes were performed, was built in a graph database and contained representations of all assets between the substation and the customers that were subject to depreciation costs. This paragraph describes the most essential elements of the network graph and its data infrastructure. An explanation of the construction of the network graph is provided in Appendix I.

Each node in the network graph represented an asset, customer or property. Different types of nodes were distinguished by its label. The labeled property graph built for this research consisted of thirteen different labels, of which nine represented different asset types. The other labels represented main power sections, properties and customers. A main power section is a set of aggregated cables which together constitute to a specific route between two transformer stations. This means that a main power section is not an actual asset itself. However, main power sections were added to the graph to perform specific queries more efficiently. Nodes labelled as 'property' consisted of one generic property (attribute) of either an asset or a customer. Each unique property-node only occurred once in the graph and had many relationships to all the nodes the property applied to. Relevant property-nodes in the graph were, among others, the different nominal capacities of the customer's grid connections and the different types of cables (singe-core or three-core cables). Figure 5 presents an overview of the data infrastructure of the network graph with respect to the medium-voltage power sections, medium-voltage cables and property-nodes. An overview of all the different assets the network graph contained, is presented in Table 1.

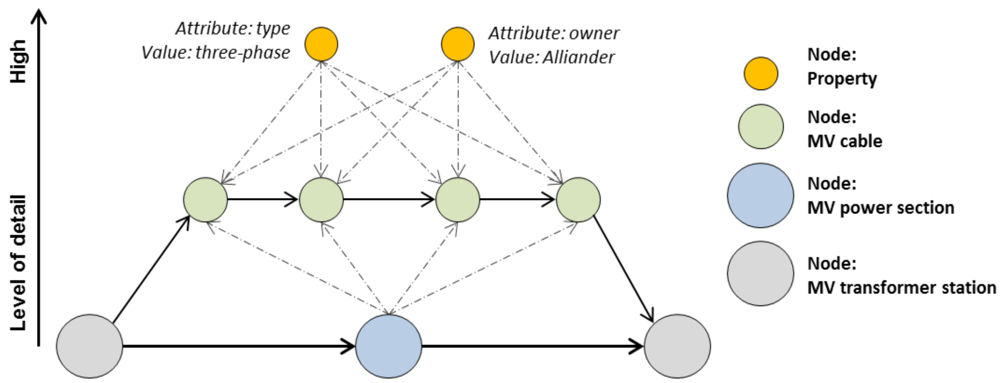


Figure 5: *Data infrastructure of network graph regarding to medium-voltage main powerlines,*

In the network graph, customers were added as nodes with label EAN. EAN is an abbreviation of European Article Number, what refers to an unique identification number for each electricity- and gas connection in Europe. Each EAN node had one relationship with an asset-node (the cable to which the customer was connected) and multiple relationships with property-nodes, for instance with the nominal capacity of its grid connection. In addition, EAN-nodes were enriched with an attribute 'annual energy-consumption', which was used in order to calculate the customer's load-contribution to each specific distribution asset in its energy supply route.

Besides main power sections, properties and customers, most of the labels represented the different sorts of assets which actually made up the network. As described in section 3.2.1, the network graph only contained assets which may have a depreciation cost, which were cables, station constructions, installations, transformers, distribution boards and the substation. To determine the annual depreciation cost of each asset, each asset-node was enriched with its valuation year, length (if applicable) and an attribute telling whether the asset was included in the population of assets subject to deprecation costs or not.

Table 1: *Different nodes of the network graph*

Label of node	Description	Asset	Important attributes
HV substation	High-voltage substation	Yes	Id, valuation year
Transformer station	MV/LV transformer station	Yes	Id, valuation year
MV installation	Medium-voltage installation in a transformer station	Yes	Id, valuation year
LV installation	Low-voltage installation in a transformer station	Yes	Id, valuation year,
Transformer	Transformer	Yes	Id, valuation year
MV cable	Medium-voltage cable	Yes	Id, valuation year, length
LV cable	Low-voltage cable	Yes	Id, valuation year, length
LV distribution board	Distribution board in low-voltage distribution network	Yes	Id, valuation year
PL distribution board	Distribution board for public lighting	Yes	Id, valuation year
EAN	Customer	No	EAN, annual energy consumption
MV power section	A route between two stations (section)	No	Id
LV power section	An outgoing low-voltage distribution network (section)	No	Id
Property	A generic property of an asset or customer. Each unique property-node has one attribute.	No	Important attributes for customers: - Capacity of grid connection - Commercial or household Important attributes for cables: - Number of cores

The interconnectivity and configuration of the grid network of Hoorn Holenweg was added to the network graph by adding relationships between physical interconnected assets. Because the network graph represented the grid network according to its standard configuration, the relationships between asset-nodes were enriched with a property telling whether an open circuit breaker interrupted the distribution network or not. Hence, each relationship in the graph had a property which can be 'open' or 'closed'. The shortest Euclidean route between the high-voltage substation and the customer was not always the route the current flow could actually travel to provide the specific customer with electricity. Figure 6 shows the interconnectivity of asset-nodes in the network graph. In this figure, main power sections and properties are not presented. In summary, the network graph built contained of over 22000 customers, 63000 low-voltage cables, 1300 medium-voltage cables, 250 distribution boards and 320 stations with transformers and installations.

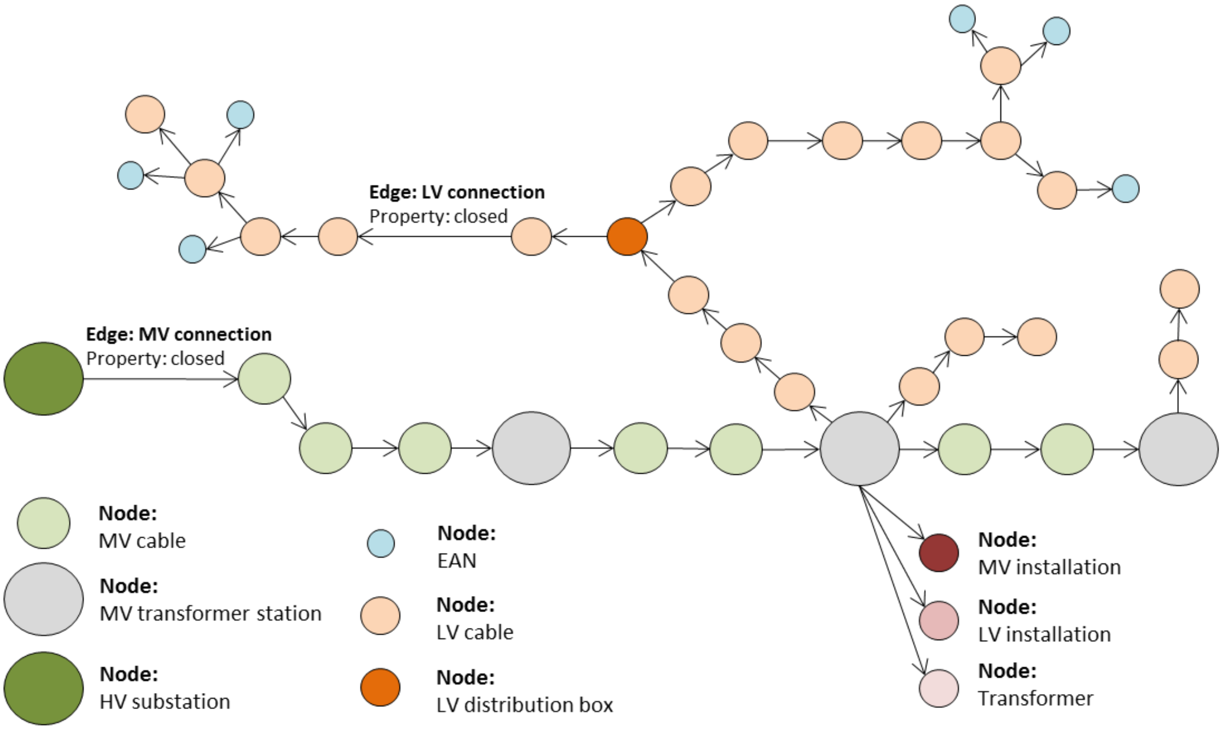


Figure 6: *Data infrastructure of network graph regarding to all assets subject to depreciation costs*

Much attention was paid to the construction and validation of a reliable network graph. However, because of topological errors in the network data used, the network graph contained of several errors and unlikely configurations. Appendix VI includes a description of the main errors in the constructed digital network.

4.2.2 Shortest weighted path query

Shortest path queries were used to calculate the route the current flow travels to provide each customer with electricity. A shortest path query is to find a path from a start node to a target node in a graph, satisfying the distance or the weight on this path is smallest (Chen, et al., 2018). The BFS (Breadth-First-Search) algorithm is the fastest and simplest algorithm of graph queries, which finds the shortest path in an unweighted graph (Chen, et al., 2018). Breadth-first algorithms conduct searches by exploring the graph one layer at a time. They begin with nodes one level deep away from the start node, followed by nodes at depth two, then depth three, and so on until the entire graph has been traversed (Chao, 2016). The Dijkstra algorithm conducts the BFS-algorithm in a weighted graph (Chen, et al., 2018). This algorithm knows a higher level of analysis and takes into account a weight parameter in order to find the shortest path between two different nodes. With respect to this research, this algorithm was used in order to find the shortest weighted path between each customer and the substation. The lengths of cable-nodes were used to obtain the total resistance of the path (see Figure 7). The length of each cable was stored as attribute of each node and hence could easily be used as weight parameter. Despite that the length of a path is not the only factor determining the total resistance, in this research, the total length was considered to be a sufficiently accurate enough approach to reality.

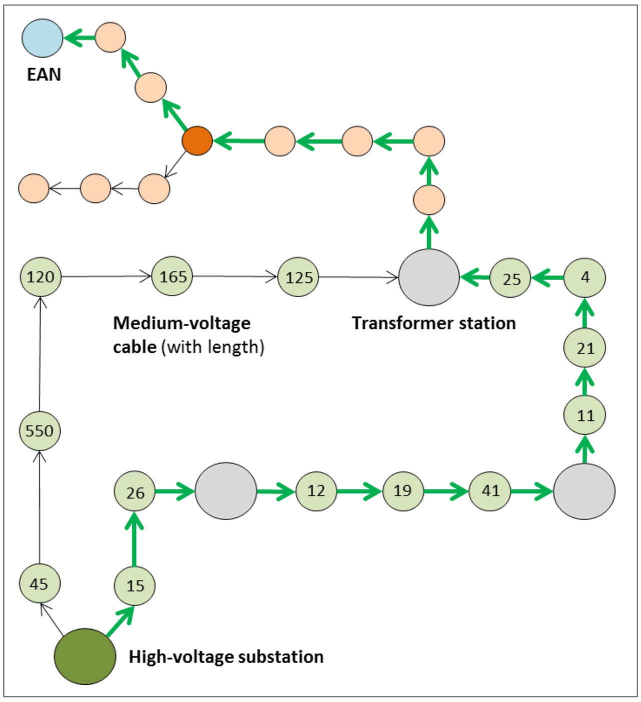


Figure 7: *Sample result of Dijkstra algorithm*

In addition to the lowest resistance, a second criterion in the network analysis was that the path could not contain any open circuit breaker interrupting the current flow. For the network analysis, this caused that an extra condition was set that only approved the path when it consisted of ‘closed’ relationships only. Each relationship connecting two different asset-nodes was enriched with the property ‘open’ or ‘closed’, telling whether the current flow could actually pass or was interrupted. Regarding to the implementation of this condition, the Cypher function *‘shortest path’* was used. Cypher was the graph-query language used in this research. The function ‘*shortest path*’ calculates the single shortest path from a start node to a target node, in which weight parameters and additional conditions can be set as well. Hence, by using the *shortest path* function of Cypher, a shortest weighted path query was built which also included the extra conditional statement checking whether the current flow could actually pass each individual cable-asset in the path. Query 1 and Query 2 show the cypher queries which were applied in order to compute the route the current flow travels to provide one specific customer with electricity.

Query 1: *Computes the shortest weighted path from a customer to its closest transformer station*

```
1 MATCH p = shortestPath((ean:EAN)-[:LVCABLE_CONNECTION*]-(t:TRANSFORMER))
2 WHERE ALL (r IN relationships(p) WHERE r.position = 'closed') AND ean.EAN  = '%s'
3 WITH p, t reduce(length = 0, lvcable IN nodes(p) | length + COALESCE(lvcable.length, 0))
4 AS dist ORDER BY dist LIMIT 1
5 WITH p, t UNWIND nodes(p) AS lv_nodes
6 RETURN lv_nodes.length, lv_nodes.id, lv_nodes.valuation_date, lv_nodes.type,
7 lv_nodes.distribution_board_id, t.id
```

In Query 1, the shortest weighted path(s) were calculated between a known customer (single start-node) and its interconnected unknown transformer stations (multiple target-nodes possible) that could physically supply the customer with electric energy. Hence, this query returned the energy supply route over the low-voltage distribution network of the system. When the customer was connected to multiple transformer stations, the transformer which had the shortest total path length was returned as being the customer's energy supply transformer station *(t.id)*.
In the shortest weighted path query, the brown colors represent labelled nodes and relationships and the green colors represent defined variables, in which variable *p* is the path variable that returned all the nodes and relationships of the path queried. In line one, the query included a not specified directional relationship unlimited in depth. The current flow may possibly come from both ways. In the network graph, the relationship which connects two mutual connected low-voltage assets was labelled as 'LVCABLE_CONNECTION' and consisted of the property 'position', which could had the value 'open' or 'closed'. Hence, in line two, a where statement required all relationships in the path to be 'closed'. In line three, the unweighted shortest path(s) that were stored in variable *p* were enriched with the weight property 'length' of all the nodes in the path. By sorting and limiting all possible routes by the total length of the paths, the energy supply route between the customer and its closest transformer station was obtained. The unwind statement in line five unlisted all the nodes in the computed shortest weighted path and stored the nodes in a new variable *lv_nodes*. The return statement returned all the attributes requested from these nodes. Attributes that were requested of each node were: its length and valuation date (to determine its depreciation cost), its type (cable or distribution board) and its id. When the path contained a distribution board, the id of this asset was returned by *lv_nodes.distribution_board_id*.
Query 2 is similar to Query 1, but instead computes the shortest weighted path between the transformer station returned by Query 1, and the known high-voltage substation of Hoorn Holenweg. Hence, this query returned the energy supply route over the medium-voltage distribution network of the system. The return statement of Query 2 returned the id's, length's (when applicable), valuation date's and types of all the asset-nodes part of the customer's energy supply route over the medium-voltage distribution network of the system. Because the medium-voltage distribution networks contained of cables, transformers and installations, *mv_nodes.type* returned whether the id and valuation date were related to a cable, transformer or installation.

Query 2: *Computes the shortest weighted path from a known transformer station to a known substation*

```
MATCH (t:TRANSFORMER) WHERE t.id = '%s' WITH t
OPTIONAL MATCH (os:SUBSTATION) WHERE os.id = '%s' WITH t, os
p = shortestPath((os)-[:MVCABLE_CONNECTION*]-(t))
WHERE ALL (r IN relationships(p) WHERE r.position = 'closed')
WITH p, reduce(length = 0, mvcable IN nodes(p) | length + COALESCE(mvcable.length, 0))
AS dist ORDER BY dist LIMIT 1
WITH p UNWIND nodes(p) AS mv_nodes
RETURN mv_nodes.length, mv_nodes.id, mv_nodes.valuation_date, mv_nodes.type
```

4.2.3 Create dataset of energy supply routes

The energy supply routes of customers were stored in a new created dataset. This dataset contained the total length of each customer's energy supply route, in which a distinction was made between the length over the low-voltage distribution network and the length over the medium-voltage distribution network. Moreover, a distinction was made between the length over valuated assets and not valuated assets (depreciated or not subject to depreciation costs at all). Table 2 shows an example of the resulting customers energy supply routes dataset.

Table 2: *Example of resulting dataset of customer's energy supply routes*

	Medium-voltage network route			**Low-voltage network route**			**Total lenght [m]**	
Ean-code	Total length [m	Valuated length [m	Relative [%	Total length [m	Valuated length [m	Relative [%	Valuated	Total
87168592000140XXXX	3457.63	3457.63	100.00	394.39	363.09	92.06	3820.71	3852.01
87168590000205XXXX	3457.63	3457.63	100.00	672.33	576.14	85.69	4033.76	4129.96
87168590000205XXXX	3457.63	3457.63	100.00	676.32	577.24	85.35	4034.86	4133.95
87168590000207XXXX	3457.63	3457.63	100.00	673.42	587.57	87.25	4045.20	4131.05
87168590000207XXXX	3457.63	3457.63	100.00	681.14	594.70	87.31	4052.32	4138.77

4.2.4 Create datasets of asset dependent customers

The network analysis returned the id's of all assets which were part of the energy supply route of a customer. Hence, by performing the network analysis for each individual customer in the grid network, all assets could be linked to its dependent customers. In order to allocate the cost of each individual asset, five different datasets were created which stored the dependent customers of five different sorts of assets. These datasets stored the dependent customers of distribution boards (1), public lighting boards (2), low-voltage cables (3), medium-voltage cables (4) and transformer stations (5), respectively. The actual transformers and installations did not have an own dataset since these assets were part of the transformer stations and therefore had the same set of dependent customers. The datasets were created by adding the customer, after the computation of its energy supply route, to the EANS column of the assets which were actually used by the specific customer. Hence, after the computation of all energy supply routes, each asset was linked to all the customers whom actually made use of the asset. The static tables were stored as text-files for the implementation of the cost allocation methods as described in section 4.1. Table 3 shows an example of a resulting asset-customer dataset.

Table 3: *Example of a resulting dataset of asset dependent customers*

Asset id {type: integer}	**EAN-code** {type: string}
422478836	"87168590000611xxxx, 87168590000611xxxx, 8716859000061xxxx, 87168590000611xxxx, 87168590000611xxxx, 87168590000611xxxx, 87168590000656xxxx, 8716859000065xxxx"
443549504	"87168590000611xxxx, 87168590000611xxxx, 8716859000061xxxx, 87168590000611xxxx, 87168590000611xxxx, 87168590000611xxxx, 87168590000656xxxx"

4.2.5 Calculate the cost and economic efficiency of customers

The tables generated in section 4.2.4 contained all the assets in the grid network with is dependent customers. Based on these results, the cost allocation methods of section 4.1 were applied which calculated the customer's cost arising from each specific asset in the grid network. This calculation was based on the number of dependent customers and their annual energy consumption. By iterating over these datasets and applying the cost allocation formulas for each asset, a new dataset was created with the total allocated annual depreciation cost of each customer. However, when assets were not used by any customer, these assets were not included in the datasets generated in section 4.2.4. Nonetheless, also assets which were not used by any customer are subject to depreciation costs. The costs of these assets, as well as the annual depreciation costs of the high-voltage substation, were linearly distributed over all the customers for both the linear cost allocation method as well as the load-based cost allocation method.

To calculate the actual economic efficiency of a customer, the customer's tariff (as determined in section 3.1.3) was subtracted by the customer's allocated annual cost. In this calculation, costs arising from operations, energy losses and overhead were added to the allocated depreciation costs of the customer. However, these costs were estimated based on the ratio of depreciation costs to the total costs of Alliander as stated in its annual report of 2017. This report showed that the total costs of Alliander in 2017 were € 1.535.000.000, of which € 396.000.000 was the result of depreciation[2]. This amounts to approximately 25%. Accordingly it was assumed that the customer's allocated depreciation costs represented 25% of the total cost of the customer.
For visualization purposes, an Inverse Distance Weighted interpolation was performed of the assigned costs to customers as well as the economic efficiencies of customers.

4.2.6 Calculate the spatial economic efficiency of the power distribution networks

The energy supply routes dataset generated in section 4.2.3 were used for calculating the economic returns of all individual assets in the grid network. To compute the economic efficiency of each individual asset, the revenue of each customer was allocated to the assets in the same way as the costs of assets were allocated to the customers. Based on the calculated length of the customer's energy supply route, a revenue per meter per customer was calculated by dividing the customer's revenue by the total length of its energy supply route. Accordingly, a total annual revenue of each asset was calculated by multiplying the revenue per meter of all the asset dependent customers, with the length of the cable. The economic return of each cable was then computed by subtracting the total allocated revenue by its annual deprecation cost.

The economic efficiency of an asset was also computed using a different approach. By summing the costs and revenues of customers downstream of the asset, the economic efficiency of the asset was presented in a way that showed whether the asset provided electric energy to relative expensive areas or relative economic efficient areas. This implies that, for each possible section in the grid network, a cross-section can be made of revenues and costs gained and incurred by the utility in order to provide electricity to the customers located downstream of the specific asset. Via this approach, not the economic return of a specific asset was calculated, but the total value of revenues minus allocated costs of all customers for whom the asset was laid out. Based on the datasets obtained in section 4.2.4 (assets with its dependent customers) and the results of section 4.2.5 (annual costs and revenues of the customers), this economic insight for each asset in the grid network was obtained.

4.3 Localize potential MicroGrid areas

Final objective of the method was to localize economically inefficient areas which at the same time are applicable for MicroGrid concepts as well. In this method, the modeled MicroGrid areas are based on an optimization of the existing grid network only and hence does not constitute to the economic most optimal energy supply system. The optimization of the existing grid network was performed by applying an iterated greedy algorithm. An iterated greedy algorithm is a simple and effective metaheuristic, which iteratively applies a constructive heuristic to an incumbent solution and uses an acceptance criterion to decide whether the newly constructed solution should replace the incumbent solution or not (Pan, et al., 2008). In order to localize the economically inefficient MicroGrid areas, this approach was used to iteratively isolate economically inefficient distribution networks from the original grid network. In the iterated greedy algorithm, an economic efficiency criterion was used to decide whether a distribution network decreased the total economic efficiency of the system or not. In this algorithm, the economic efficiency criterion referred to the economic efficiency of a low-voltage power section, that was calculated by summing up the costs and revenues of all customers whom were dependent on the power section's cables. Therefore, a low-voltage power section and its dependent customers were isolated from the existing grid network when their summed allocated costs where higher than their summed revenues. However, when transferring isolated customers to another energy supply system (MicroGrid), the costs allocated to those customers had to be reallocated to the customers whom were still part of the original grid network. Hence, after each iteration, the greedy algorithm reallocated the costs of affected assets to a lower number of customers, what required the algorithm to reevaluate the acceptance criterion after each iteration. Based on this approach, low-voltage power

[2] https://2017.jaarverslag.alliander.com/verslagen/jaarverslag-2017

sections and its customers were isolated from the original system until all power sections in the grid network fulfilled the economic efficiency criterion.
The reallocation of costs was performed using the asset dependent customers datasets generated in section 4.2.4. The disconnected customers were removed from the datasets, after which the cost of each affected asset was reallocated to its new, reduced group of dependent customers. This means that the remaining asset dependent customers were allocated higher costs. The customer’s cost difference arising from each affected asset, was calculated by comparing the customer’s cost in the new situation with the customer’s cost in the original situation. This difference then was used to calculate the increase in cost for each customer due to each affected asset. The total cost increase of each customer determined the new economic efficiencies of the remaining low-voltage power sections.

4.3.1 Defining potential MicroGrid areas

As a result of iteratively isolate low-voltage power sections from the existing grid network, it may be possible that all outgoing low-voltage power sections of a specific transformer station were isolated from the grid network. In such a scenario, the transformer station was no longer used by any customer whom was still part of the original grid network. An area was defined as a potential MicroGrid area when the entire service area of the transformer station was isolated from the original grid network. If the isolated transformer station was located at the end of a medium-voltage network, its supplying medium-voltage power section was no longer used either and hence was part of the designated MicroGrid area as well.

4.4 Impact of a capital investment and a distributed generation source

As described in section 1.2.3, this research exposed potential applications with respect to the deployment of other MicroGrid and SmartGrid concepts. Hence, the method in this research was used to calculate the economic impact of a conventional capacity expansion as well as the alternative of adding a distributed generation source to the grid network. The economic impact of a conventional capacity expansion was calculated by adding extra costs to the power section upgraded, after which the additional costs were distributed over its dependent customers by using the linear cost allocation method. To calculate the economic impact of a distributed generation source added to the grid network, new energy supply routes were calculated based on the method described in section 4.2. Thereafter, the depreciation costs of the network were reallocated to the customers. The impact of adding a distributed generation source was obtained by comparing the results of the conventional energy supply system with the results including the added distributed generation source.

4.5 Sensitivity analysis

The customer's allocated cost as well as the economically inefficient MicroGrid areas are sensitive to the used average asset costs as well as the choices made regarding the cost allocation method applied. To assess the robustness of the method's outputs, a sensitivity analysis was performed. The information gained by a sensitivity analysis is to map the variability of the output according to the variability of the input (Laoun et al., 2016) and identify which input factors are most important (Saltelli, 2002).

4.5.1 Scenario analysis

The sensitivity analysis was performed by varying two essential input factors in eight different scenario calculations. In these scenario calculations, the first differentiated input factor was the cost allocation method applied. Moreover, the method used an average cost per meter cable which did not include any spatial differentiation of costs. This means that a cable located in an urban area was deemed just as expensive as a cable in a rural area. However, assets located in complex urban areas should essentially be more costly than assets located in non-complex (rural) areas. Urban areas generally involve more road reparation work, are characterized by a more complex underground cable infrastructure and require more additional traffic measures. Hence, the used average cost per meter was designated as second variable input factor.
Based on the actual costs incurred by Alliander in recent projects, it was found that there exists a strong varying differentiation between the costs of projects in rural, urban and industrial areas and that drilled cables also had a significant impact on the actual asset costs. Moreover, it was found that the average cost per meter in Hoorn, calculated based on all recent projects between 2012 and 2018, was 17% lower than the used cost per meter as provided by the financial department of Alliander. This means that the used cost per meter may have caused an overestimation of the total grid network's costs. However, because the actual costs incurred by the grid operator even showed high variations inside one specific land use type, the variability of the method's output to the used average cost per meter was assessed by calculating four different cost-scenarios. In these scenarios, the variability to the used average cost per meter was calculated based on an average, best-case and worst-case spatial correction of asset costs. This means that, for each specific cost allocation method, four different cost-scenarios were distinguished:

1. The average cost per meter approved and provided by the financial department of Alliander;
2. The costs of [1] plus an average spatial differentiation factor and additional drilling costs;
3. The costs of [1] plus a best-case spatial differentiation factor and additional drilling costs;
4. The costs of [1] plus a worst-case spatial differentiation factor and additional drilling costs.

Table 4: *Different scenarios distinguished*

	Cost 1	Cost 2	Cost 3	Cost 4
Linear allocation	*Scenario 1*	*Scenario 3*	*Scenario 5*	*Scenario 7*
Load-based allocation	*Scenario 2*	*Scenario 4*	*Scenario 6*	*Scenario 8*

Table 4 shows all different calculated scenarios. In scenario 1 and 2, the average costs per meter provided by the financial department of Alliander were used. These scenarios therefore did not include any spatial correction of costs. In all other scenarios, the used cost per meter was reduced by 17% and was corrected for the asset's topographical location (urban, rural or industrial).

4.5.2 Spatial classification of cables

To calculate the costs of a cable in the scenarios 3 to 8, the cables needed to be classified as being located in an urban, industrial or rural environment. Hence, a spatial overlay of the grid network's topology with a geographical dataset was performed. For this overlay, the Dutch topographical data Top10NL was used. The Top10NL topographical map distinguishes various land use classifications which were, in the context of this research, aggregated to urban, industrial or rural. In Appendix II, the result of the topographical land use classification is provided. Besides the general classification whether an asset was being located in an urban, industrial or rural environment, the scenarios 3 to 8 also took into account whether a specific cable might has been drilled. This means that for each cable-asset in the grid network was determined whether the cable intersected with water.

When a cable asset was assigned as being drilled, additional costs were added to this cable depending on the year the cable was commissioned. In 2017, the average additional drilling costs were €83,49 per meter. To determine whether a cable was drilled, a spatial overlay was performed with the Top10NL waterbodies layer. However, not all cables that intersected with water were assigned as being drilled. With respect to the determination of drilled cables, the following criteria were used:

- Waterways needed to be at least 6 meters wide;
- Waterways needed to be part of the main drainage system;
- Cables which intersected with harbors were always drilled;
- The cable needed to intersect with at least 10% of its total length;
- If more than 25% of the length of the entire power section intersected with water, all cables of the section were assigned as being drilled;

If the above conditions applied, the intersecting length was computed and compared with the total length of the cable. In addition, in order to verify whether the entire power section should be assigned as being drilled, the intersecting length of the cable was added on the total intersecting length of cables which belong to the same power section. Based on these criteria, a total length of 2643 meter medium-voltage cable was assigned as being drilled. An overview of these cables is provided in Appendix II.

4.5.3 Calculation of spatial differentiation factors

The scenarios 3 to 8 differentiated from each other due to an average, best-case and worst-case cost-scenario. This means that, in these scenarios, the depreciation cost of each cable were multiplied by 0,83 (17% lower average costs in Hoorn) and a spatial differentiation factor depending on the asset's location and the specific cost-scenario that applied. In the average cost-scenario (Cost 2), the spatial differentiation factors were calculated by averaging the actual expenditures made in four comparable projects in Hoorn that were situated in a specific topographical area (urban, rural or industrial), divided by the cost per meter provided by the financial department of Alliander and reduced by 17%. This assumed that the grid network's relative asset population in the specific topographical area, also applied to the relative asset population on which the costs of the financial department were based. In the best-case cost-scenarios (Cost 3), the spatial differentiation factors were calculated by taking, instead of the average incurred cost per meter, the lowest possible cost per meter of a specific topographical area. This resulted in lower spatial differentiation factors for each area, causing that the total assets depreciation costs became lower. Regarding to the worst-case cost-scenarios, higher spatial differentiation factors were calculated by taking the highest possible cost per meter incurred in the specific topographical area. Formulas 3, 4 and 5 present how the spatial differentiation factors were calculated. Table 5 presents, for the different topographical areas, the actual costs per meter incurred in recent projects in Hoorn. The actual calculated spatial differentiation factors are presented in Table 6.

$$f_{avg} = \frac{I_{avg}}{C * 0{,}83} \quad (3) \qquad f_{best-case} = \frac{I_{min}}{C * 0{,}83} \quad (4) \qquad f_{worst-case} = \frac{I_{max}}{C * 0{,}83} \quad (5)$$

where

f_{avg} is an average spatial differentiation factor for a specific topographical area;
$f_{best-case}$ is a best-case spatial differentiation factor for a specific topographical area;
$f_{worst-case}$ is a worst-case spatial differentiation factor for a specific topographical area;
C is the average cost per meter provided by the financial department of Alliander;
I_{min} is the minimum cost per meter in Hoorn for a specific topographical area;
I_{max} is the maximum cost per meter in Hoorn for a specific topographical area;
I_{avg} is an average cost per meter in Hoorn for a specific topographical area;

Table 5: *Actual costs incurred in different topographical areas in Hoorn*

Cost-scenario	Rural area [€/m]	Industrial area [€/m]	Urban area [€/m]
Average	60,40	123,20	118,50
Best-case	44,40	88,10	108,30
Worst-case	93,50	140,80	168,60
Cost finance department (2015)		Average costs incurred in Hoorn (2015)	
€ 136 /meter		€ 113 /meter	

Table 6: *Calculated spatial differentiation factors*

Cost-scenario	Rural area	Industrial area	Urban area
Average	0,53	1,09	1,05
Best-case	0,39	0,78	0,96
Worst-case	0,83	1,25	1,49

4.5.4 Quantifying sensitivity of customers costs

In order to assess the variability of the customers costs to each of the two input factors, the range of the customer's allocated cost was calculated. Range is a measure of spread and was used to measure the variability of the assigned costs to customers to each of the variable input factors. The range due to spatial corrected asset costs was calculated by subtracting the customer’s maximum cost in scenario 1, 3, 5 and 7 (linear cost allocation) by the customer’s minimum cost in these scenarios. The variability to the cost allocation method applied was assessed by calculating the difference between the customer’s cost in scenario 1 (linear cost allocation) and scenario 2 (load-based cost allocation).

4.5.5 Quantifying sensitivity of the modeled MicroGrid areas

In order to assess the variability of the modeled MicroGrid areas, set theory was used. Set theory is a mathematical concept that studies sets in a universe of objects. In this mathematical concept, a universe specifies a very general collection of elements, such as all cars or all people. Sets can then be defined within such a universe for elements that fulfil a certain condition, such as the set of people older than 75 years old (Molenaar, 1998). In order to assess the impact of the method’s variable input factors to the calculated MicroGrid locations, this mathematical concept was used. In this context, the universe specified all potential MicroGrid locations obtained from all different scenario results. Hence, a set can be defined as the resulting MicroGrid locations of one specific scenario. The intersection of two scenarios enabled a descriptive analysis regarding to whether the scenarios were very identical or whether they were very different. In these intersections, all members in set 2 should also be member of set 1 and vice versa. This means that when set 2 had no members occurring in set 1, the intersection resulted in an empty set. The union of two sets (two scenario results) is the set that contains all elements that were members of one of these two sets or both sets (Molenaar, 1998). Figure 8 shows the concept of an intersection set and an union set. When the difference between the intersection set and the union set is relatively large, the different scenarios resulted in considerably different results. Hence, in order to assess the consistency of the method’s calculated MicroGrid locations, the intersection sets and union sets of all scenario combinations were calculated. If the differences between those sets were very large, the method’s calculated MicroGrid areas would be considered inconsistent. However, if the union and intersection sets were identical, the method’s output would be considered consistent. This (in) consistency was used as a measure of sensitivity to the two different variable input factors.

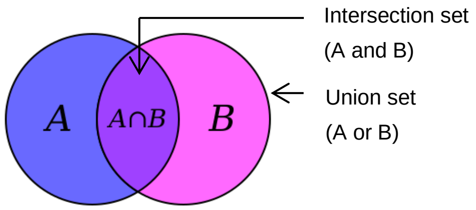


Figure 8: *Principle of set theory*

5 Results case study

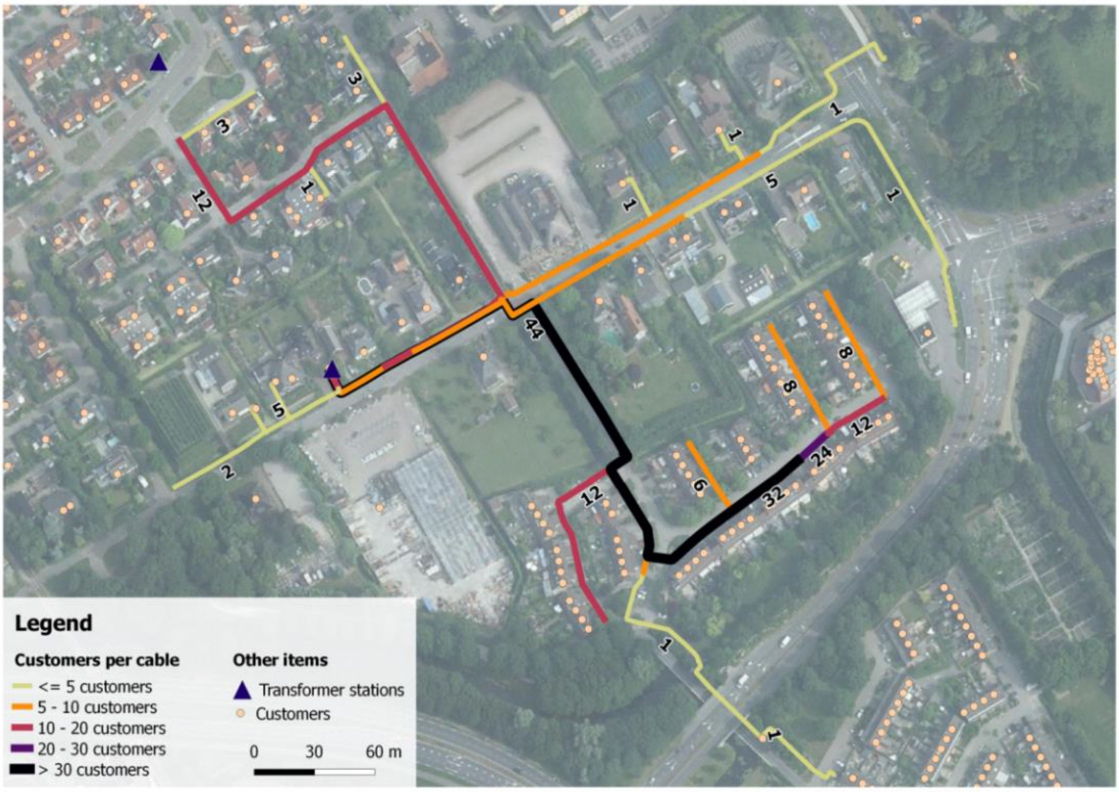
5.1 Network analysis of customer's energy supply routes

A network analysis computed the energy supply route of each individual customer in the grid network. The total calculated length of the energy supply route of each customer is shown in Figure 9.

Figure 9: *Length of each customer's energy supply route*

As a result of the network analysis, each distribution asset was linked to its dependent customers. In Figure 10, the medium-voltage cables in the grid network are visualized based on their number of dependent customers while in Figure 11 the number of dependent customers for a small selection of low-voltage cables is presented. Similar results were obtained for the transformer stations, distribution boards and all other cable assets.

Figure 10: *Number of asset dependent customers of each medium-voltage cable*



Source: Publieke dienstverlening op de kaart (PDOK), Nederland op de kaart

Figure 11: *Number of asset dependent customers of a small selection of low-voltage cables.*

5.2 Customers allocated costs

Figure 12:

Figure 13:

5.3 Customers economic efficiencies

Figure 14

Figure 15

5.4 Spatial economic efficiency of the grid network

Figure 16

Figure 17

Figure 18

5.5 MicroGrid areas

Main aim of this research was to determine which parts of the grid network decrease the total economic efficiency of the network and at the same time can be applied as MicroGrid as well. Potential MicroGrid areas were calculated by using the method as described in section 4.3.1 and consisted of areas which decreased the total economic efficiency of conventional energy supply system. Based on respectively the linear cost allocation method and the load-based cost allocation method, Figures 19 and 20 show the results regarding to which areas were computed as economically inefficient.

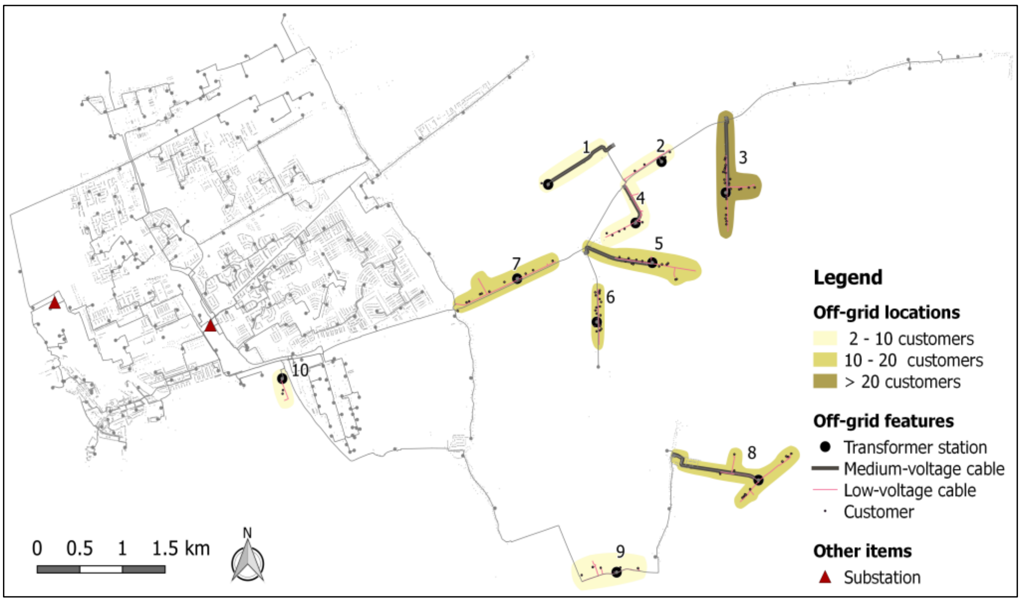


***Figure 19**: Computed economic inefficient areas based on the linear cost allocation method .*

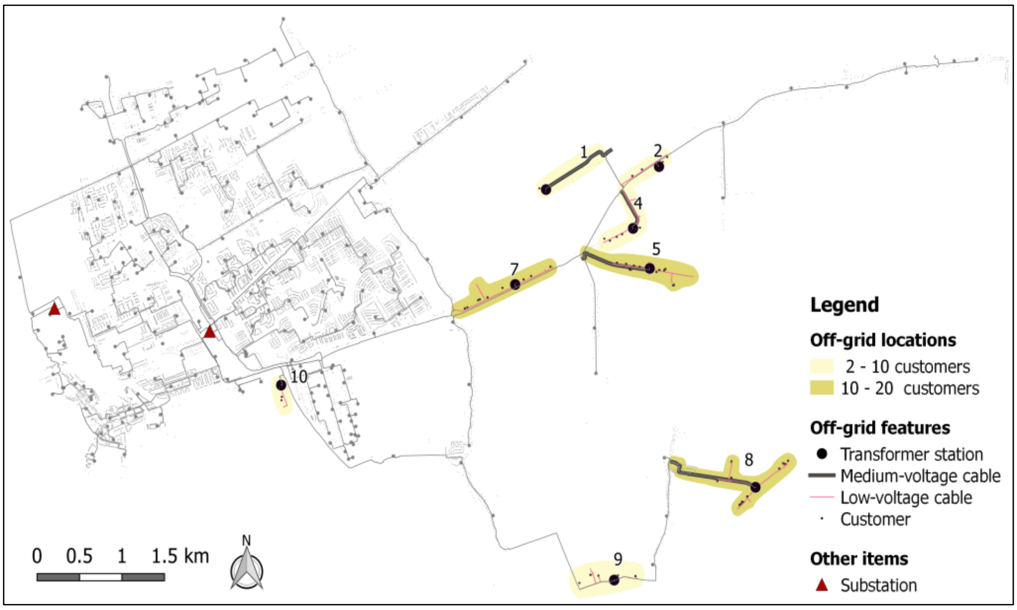


***Figure 20:** Computed economic inefficient areas based on the load-based cost allocation method.*

Figure 21

5.6 Sensitivity analysis

5.6.1 Sensitivity of the customer's allocated cost

Figure 22:

Figure 23:

5.6.2 Sensitivity of MicroGrid areas

Set theory was used to assess the sensitivity of the calculated MicroGrid areas to the variability of input factors. For each combination of scenarios, the union set and intersection set were calculated. Table 7 presents the result of the calculated union sets while Table 8 presents the result of the intersection sets. Both tables need to be considered jointly and we are particularly interested in the ratio between intersection and union. Scenario 1 shows that the linear cost allocation method resulted in 10 economically inefficient areas, of which 8 areas also occurred in the result based on the load-based cost allocation method (scenario 2). Moreover, the result shows that only scenario 7 resulted in MicroGrid areas which deviated from scenario 1. Table 8 shows that 6 of the 10 areas in scenario 1 occurred in all other scenarios. In the combinations of scenarios 4, 5 and 6, the union sets were equal to the intersection sets. This means that the results obtained in these scenarios were identical.

Table 7: *Unions sets derived from the scenario analysis results*

Union sets

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8
Scenario 1	10	10	10	10	10	10	12	10
Scenario 2		8	9	8	8	8	11	8
Scenario 3			7	7	7	7	10	8
Scenario 4				6	6	6	10	7
Scenario 5					6	6	10	7
Scenario 6						6	10	7
Scenario 7							10	10
Scenario 8								7

Table 8: *Intersection sets derived from the scenario analysis results*

Intersection sets

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8
Scenario 1	10	8	7	6	6	6	8	7
Scenario 2		8	6	6	6	6	7	7
Scenario 3			7	6	6	6	7	6
Scenario 4				6	6	6	6	6
Scenario 5					6	6	6	6
Scenario 6						6	6	6
Scenario 7							10	7
Scenario 8								7

6 Discussion

MicroGrid areas

Table 8 shows that the modeled economically inefficient areas were quite consistent regarding to the cost allocation method applied. Scenario 1 shows that the linear cost allocation method resulted in 10 economically inefficient areas, of which 8 areas also occurred in the result based on the load-based cost allocation method (scenario 2). Moreover, the union of scenario 1 and 2 shows that the total number of areas was 10 as well. This means that the load-based cost allocation method did not result in any new areas not obtained in scenario 1. Therefore, the cost allocation method did not affect the locations of economically inefficient areas very significantly. This implies that, although the cost allocation method had a large impact on the customer's individual allocated cost, the impact was much smaller on an aggregated level. Customers with a high annual energy consumption balanced out the lower costs of low-consuming customers quite well. However, the intersection set of scenario 1 and 2 still shows a difference of two areas. Figure 19 represents the modeled MicroGrid areas obtained in scenario 1 and Figure 20 represents the modeled MicroGrid areas in scenario 2. The difference between those figures shows that locations 3 and 6 did not intersect with scenario 1. The fact that locations 3 and 6 were economically efficient in scenario 2, can possibly be explained by the fact that both areas contained more low-voltage customers (19 and 28 respectively). Because most of the high consuming and thus more expensive customers were directly connected to the medium-voltage distribution network, lower costs were allocated to the customers connected to the low-voltage distribution network. This means that, when an area contained more low-voltage customers, it was affected more severe when high consuming customers in other areas were utilizing the same assets. Moreover, the relative load contributions of high consuming customers were more significant in a sparsely populated areas. This caused that low-voltage customers were allocated significant lower costs. The low population density in combination with the higher number of low-voltage customers, might explain why locations 3 and 6 were economically inefficient in scenario 1 and economic efficient in scenario 2.

Additionally, Table 8 shows that based on spatial corrected asset costs, 6 out of the 10 areas in scenario 1 (no spatial correction of costs) were economically inefficient at all times. This implies that, regarding the grid network of Hoorn Holenweg, the modeled economically inefficient areas had more sensitivity to spatial corrected asset costs. Table 7 shows that the spatial corrected asset costs only changed the locations of MicroGrid areas in scenario 7 (worst-case cost scenario). The union of scenario 1 and 7 resulted in 12 areas. This means that the worst-case cost scenario resulted in two new economically inefficient areas only. This implies that, regarding the grid network of Hoorn Holenweg, the modeled economically inefficient areas still were quite consistent.

Discussion of methods

The main aim of this research was to determine which parts of the grid network decrease the total economic efficiency of the network and at the same time can be applied as MicroGrid. However, the iterated greedy algorithm as presented in section 4.3 isolated customers from the original grid network but did not include calculations on the MicroGrids itself. This means that the iterated greedy algorithm localized the economically inefficient areas but did not perform a network optimization in which the costs of MicroGrids were traded-off with the costs of the existing energy supply system. This means that the iterated greedy algorithm cannot be used to conclude whether the economically inefficient areas should actually be transferred to a MicroGrid system or not. In addition, modeling MicroGrids also require the calculation of costs saved by power plants not built and feeder lines, transformers and substations not upgraded or replaced (St. John, 2014). This would imply that the iterated greedy algorithm is not be suitable for designing the economic most optimal energy supply system, including MicroGrid and SmartGrid applications.

In order to calculate the spatial economic efficiency of the grid network, section 4.2.5 explained that the method involved an estimation of costs arising from operations, energy losses and overhead. The customer's allocated depreciation cost was assumed to be 25% of the total annual cost of the customer. This means that when a customer used more assets and thus was allocated higher depreciation costs, it also was allocated higher costs arising from operations, energy losses and overhead. However, this introduced incorrect proportions: if a customer used many depreciated assets, it was allocated lower depreciation costs and hence was allocated lower costs arising from operations, energy losses and overhead as well. Moreover, the estimation of energy loss and operating expenditures arising from each asset did not take into account relevant predictors, such as the location, age and quality (material) of the asset. This implies that the economic efficiency of customers as presented in Figure 14 and 15, was affected by inaccurate estimations of operating and energy loss expenditures.

Section 4.1 explained that cost allocation was based on the annual energy consumption of the customers. This means that the customer's load as a result of energy generation was not taken into account. According to Weckx et al., (2015), distributed energy generation may also cause the voltage to rise to unacceptable levels. This implies that energy generating customers may also cause the investment needs to increase. However, the cost allocation methods in this research did not take into account the extra costs of the utility company caused by electric energy added to the grid network. This means that energy generating customers in no case were allocated higher costs.

Moreover, section 4.1 explained that the customers costs were calculated by using a static extent-of-use of each customer to each asset in the grid network. This extent-of-use was based on a non-dynamic energy supply route of the customer, in which was assumed that its electric energy came from the nearest substation at all times. However, because today's power system of the Netherlands is subject to increasing numbers of small-scale distributed energy sources (Jokic, 2007), this assumption becomes less plausible each day. In the long term, this implies that the distributed generation units should be integrated in this method as well. However, when this is the case, the cable's length which was used as weight parameter in the Dijkstra shortest path algorithm, would also need to be replaced. Distributed generation sources involve dynamic energy generation profiles, which means that real time load-flow calculations are needed for calculating the energy supply route of each customer at each moment in time.

In section 2.1, Li & Tolley, (2007) proposed a cost allocation method which takes into account the unused capacity of the existing system. By integrating the unused capacity in the calculation of the customers grid network costs, forward looking investment costs will cause areas of underinvestment to become more expensive. This implies that, based on the unused capacity of the system, it can be modeled whether the addition of a distributed generation unit postpone or accelerates the forward looking investment needed. However, this method did not include the unused capacity of the existing system, which means that cost allocation was not very committed to the forward looking investment costs in the energy supply system.

Despite that the method included limitations, fundamental aspect of the method was the use of a network graph. Due to the data structure of the network graph, the method performed the network analysis efficiently, which makes it a suitable tool for real time network analyzes as well. Moreover, changes in the physical system of the grid network can easily be added to the digital network graph too. This implies that a network graph can be very useful for calculating how a combination of distributed generation sources can help solve the most expensive areas first. Appendix V represents two potential applications of using a network graph. In this result, the network graph was used to model the economic impact of a conventional grid expansion as well as the addition of a distributed generation source to the grid network. At this moment, however, this results do not yet represent a realistic economic impact. After all, the method was not yet able to calculate real time energy supply routes and did not include the unused capacity of the existing system. Regarding the calculated economic impact in Appendix V, this means that the method assumed that the added distributed energy generator did not had limitations regarding to its energy generation profile, which in reality is not a plausible assumption. This means that customers to which the added generator was closest, per definition were subject to shorter routes and hence were allocated lower costs. However, in Sharma et al., (2011) and Wu et al., (2000), load flow analysis methods are proposed using graph theory principles. This implies that the graph database can also be used for other applications solving the present limitations of this method. This means that, in the long term, the network graph can be used to calculate how distributed generation sources can decrease network utilization, reduce energy losses and prevent conventional grid investments.

7 Conclusions and recommendations

Regarding research question 1, a suitable cost allocation method has to calculate the contribution of each customer to the total energy load on each asset for different moments in time. This impact should be calculated based on the customer's energy consumption as well as the customer's energy generation. Besides, the unused capacity of the network should be integrated in the cost allocation method as well.

Regarding research question 2, the network graph computed energy supply routes efficiently. Moreover, this research showed that the implementation of a graph database is suitable for modeling the economic impact of system's changes, such as grid expansions and added distributed generation sources.

Regarding research question 3, economically inefficient areas were only found in rural and sparsely populated areas, for both the linear cost allocation method as well as the load-based cost allocation method.

Regarding research question 4, cost allocation caused a significant variability in the customer's individual allocated cost. Nonetheless, the consistency of the modeled MicroGrid areas was quite high. The locations of economically inefficient areas were not very sensitive to the cost allocation method applied. The spatial corrected asset costs caused a lower consistency of modeled MicroGrid areas, especially in sparsely populated areas. In addition, the calculated cost correction factors are not one-to-one applicable to other study areas. Hence, it cannot be concluded whether the impact of the unspecific asset costs is the same for other grid networks.

Recommendations for cost allocation

The method proposed in this research implemented static cost allocation. This involved several limitations. In order to enable the method to assess whether the addition of a distributed generation source can actually prevent a required investment, dynamic and time-dependent energy supply routes need to be computed. Future work can contribute to this problem by implementing a variable extent-of-use cost allocation method, based on load flow calculations taking into account the time-dependent energy generation and demand. Moreover, by integrating the unused capacity of the existing system as well, the method can be used for directing the market to locations where large scale distributed generation units, such as wind- and solar farms, reduce the costs of the overall energy supply system. Based on a spatial prediction of growth rate of demand and generation, the unused capacity can be integrated by calculating the present value of future investment needed.

Recommendations for improvement

This method did not perform a network optimization in which the costs of MicroGrids were traded-off with the costs of the existing energy supply system. This means that this method cannot be used to conclude whether the economically inefficient areas should actually be transferred to a MicroGrid system. Further research should focus on the integration of costs of MicroGrid applications as well and should propose a new optimization algorithm which balance out which combination of technologies result in the most optimal energy supply system.

The specification of operating and energy loss expenditures arising from each asset in the grid network was outside the scope of this research. Hence, the total customer's cost was estimated based on its allocated depreciation cost. However, this caused incorrect proportions. A better approach for estimating the customer's energy loss and operating expenditures, is to calculate the proportion of the customer's energy supply route to the total length of all routes. The customer's proportion can then be multiplied with the total estimated costs of the grid network arising from overhead, operation and energy losses. Besides, MicroGrid and SmartGrid applications have potential to decrease network utilization and thus reduce network losses. Future work can contribute to this problem by integrating specific costs arising from operations and energy losses as well.

In this method, the customer's load as a result of energy generation was not considered. To prevent energy generating customers to be underestimated in costs, a possibility is to allocate the grid network's costs based on the nominal capacity of the customer's grid connection. After all, the nominal capacity of the customer's grid connection is a good measure of the customer's potential energy consumption, as well as its potential energy generation. This means that customers with high levels of generation, will also be allocated higher costs.

Fundamental aspect of the method was the use of a network graph. However, despite that many attention was paid to the construction of the network graph, topological errors still affected the results. Further validations and improvements of the network graph are recommended to obtain more reliable results.

References

Anastasiadis, A., Konstantinopoulos, S., Kondylis, G., Vokas, G., & Salame, M. (2018). Carbon tax, system marginal price and environmental policies on Smart Microgrid operation. *Management of Environmental Quality: An International Journal*, 29(1), 76-88.

Blom, M., Bles, M., Leguijt, C., & Rooijers, F. (2012). *Maatschappelijke kosten en baten van Intelligente Netten.* Delft: CE Delft.

Bonbright, J., Danielsen, A., & Kamerschen, D. (1961). *Principles of public utility rates.* New York: Columbia University Press.

Bosman, M. (2012). *Planning in smart grids.* University of Twente: Enschede, Netherlands.

Chao, J. (2016, August 1). *Neo4j*. Retrieved from Graph Databases for Beginners: Graph Search Algorithm Basics: https://neo4j.com/blog/graph-search-algorithm-basics/

Chen, W., Chen, Z., Liu, J., & Yang, Q. (2018). A novel shortest path query algorithm. *Cluster Computing*, 1-12.

De Leone, R., Giovannelli, A., & Pietrini, M. (2017). Optimization of power production and costs in microgrids. *Optimization Letters*, 11(3), 497-520.

De Oliveira-De Jesus, P., & Antunes, C. (2018). Economic valuation of smart grid investments on electricity markets. *Sustainable Energy, Grids and Networks*.

Deb, S., Tammi, K., Kalita, K., & Mahanta, P. (2018). Impact of Electric Vehicle Charging Station Load on Distribution Network. *Energies*, 11(1), 178.

Europacable. (2014). *An Introduction to Medium and Low Voltage Cables in Distribution Networks as support of Smart Grids.* Brussels: Europacable.

Galvin, R. (2017). Energy consumption effects of speed and acceleration in electric vehicles: Laboratory case studies and implications for drivers and policymakers. *Transportation Research Part D: Transport and Environment*, 53, 234-248.

Jasemi, M., Adabi, F., Mozafari, B., & Salahi, S. (2016). Optimal Operation of Micro-grids Considering the Uncertainties of Demand and Renewable Energy Resources Generation. *International Journal of Renewable Energy Development*, 5(3), 233-248.

Jokic, A. (2007). Price-based optimal control of electrical power systems. *Dissertation Abstracts International*, 68(04).

Kyriakopoulos, G., & Arabatzis, G. (2016). Electrical energy storage systems in electricity generation: energy policies, innovative technologies, and regulatory regimes. *Renewable and Sustainable Energy Reviews*, 56, 1044-1067.

Laoun, B., Naceur, M., Khellaf, A., & Kannan, A. (2016). Global sensitivity analysis of proton exchange membrane fuel cell model. *International Journal of Hydrogen Energy*, 41(22), 9521-9528.

Li, F., & Tolley, D. (2007). Long-run incremental cost pricing based on unused capacity. *IEEE Transactions on Power Systems*, 22(4), 1683-1689.

Medhi, S., & Baruah, H. (2017). Relational database and graph database: A comparative analysis. *Journal of Process Management. New Technologies*, 5(2), 1-9.

Mendez, V., Rivier, J., De La Fuente, J., Gomez, T., Arceluz, J., Marin, J., & Madurga, A. (2006). Impact of distributed generation on distribution investment deferral. *International Journal of Electrical Power & Energy Systems*, 28(4), 244-252.

Miller, J. (2013). Graph database applications and concepts with Neo4j. *Proceedings of the Southern Association for Information Systems Conference* (pp. Vol. 2324, p. 36). Atlanta, GA, USA: Georgia Southern University.

Molenaar, M. (1998). *An introduction to the theory of spatial object modelling for GIS.* Enschede: Taylor & Francis.

Pagani, G., & Aiello, M. (2016). From the grid to the smart grid, topologically. *Physica A: Statistical Mechanics and its Applications*, 449, 160-175.

Pan, Q., Wang, L., & Zhao, B. (2008). An improved iterated greedy algorithm for the no-wait flow shop scheduling problem with makespan criterion. *The International Journal of Advanced Manufacturing Technology*, 38(7-8), 778-786.

Parsons, J., & Sakhrani, V. (2010). Electricity Network Tariff Architectures A Comparison of Four OECD Countries. *Center for Energy and Environmental Policy Research*, No. 10-008. Available at SSRN: https://ssrn.com/abstract=1711198 or http://dx.doi.org/10.2139/ssrn.1711198.

Picciariello, A., Ramirez, C., Guillén, J., Marín, P., & Söder, L. (2014). Distribution Network Tariffs and Distributed Generation: Need for an Innovative Methodology to Face New Challenges. Application to a Case Study. *Energy & the Economy, 37th IAEE International Conference.* International Association for Energy Economics.

Picciariello, A., Reneses, J., Frias, P., & Söder, L. (2015). Distributed generation and distribution pricing: Why do we need new tariff design methodologies? *Electric power systems research*, 119, 370-376.

Provoost, F. (2009). *Intelligent distribution network design.* Eindhoven: Eindhoven University of Technology.

Raju, L., Morais, A., Rathnakumar, R., Ponnivalavan, S., & Thavam, L. (2017). Micro-grid Grid Outage Management Using Multi-agent Systems. *Recent Trends and Challenges in Computational Models (ICRTCCM), 2017 Second International Conference* (pp. 363-368). IEEE.

Reza, M. (2006). *Stability analysis of transmission system with high penetration of distributed generation.* Delft: TU Delft.

Saltelli, A. (2002). Sensitivity analysis for importance assessment. *Risk analysis*, 22(3), 579-590.

Shahriari, S., Varjani, A., & Haghifam, M. (2012). Cost reduction of distribution network protection in presence of distributed generation using optimized fault current limiter allocation. *International Journal of Electrical Power & Energy Systems*, 43(1), 1453-1459.

Sharma, D., Chaturvedi, A., Purohit, G., & Shivarudraswamy, R. (2011). Distributed load flow analysis using graph theory. *World Academy of Science, Engineering and Technology*, 50, 219-222.

Soares, T., Pereira, F., Morais, H., & Vale, Z. (2015). Cost allocation model for distribution networks considering high penetration of distributed energy resources. *Electric Power Systems Research*, 124, 120-132.

Sotkiewicz, P., & Vignolo, J. (2007). Towards a cost causation-based tariff for distribution networks with DG. *IEEE Transactions on Power Systems*, 22(3), 1051-1060.

St. John, J. (2014, August 21). *Greentech Media* . Retrieved from Distributed Marginal Price: The New Metric for the Grid Edge? Pinpointing the value of the distributed, intelligent grid with Integral Analytics: https://www.greentechmedia.com/articles/read/distributed-marginal-price-dmp-the-new-metric-for-the-grid-edge#gs.VcOWepg

Strbac, G., & Mutale, J. (2005). *Framework and methodology for pricing of distribution networks with distributed generation.* Centre for Distributed Generation and Sustainable Electrical Energy.

Vignolo, M. (2007). *Cost-causality based tariffs for distribution networks with distributed generation.* Montevideo: Universidad de la Republica (Uruguay). Facultad de Ingenieria, 2007.

Wang , Z., Gu, C., & Li, F. (2018). Flexible operation of shared energy storage at households to facilitate PV penetration. *Renewable Energy*, 116, 438-446.

Weckx, S., D'hulst, R., & Driesen, J. (2015). Locational pricing to mitigate voltage problems caused by high PV penetration. *Energies*, 8(5), 4607-4628.

Wu, F., Ni, Y., & Wei, P. (2000). Power transfer allocation for open access using graph theory-fundamentals and applications in systems without loopflow. *IEEE transactions on power systems*, 15(3), 923-929.

Yoon, B., Kim, S., & Kim, S. (2017). Use of graph database for the integration of heterogeneous biological data. *Genomics & informatics*, 15(1), 19-27.

Zhou, K., Fu, C., & Yang, S. (2016). Big data driven smart energy management: From big data to big insights. *Renewable and Sustainable Energy Reviews*, 56, 215-225.

Appendixes

Appendix I Construction of the network graph

The in relational databases stored network data were converted to a graph database structure in order to efficiently perform the network analysis of each individual customer in the grid network. In this appendix is explained how the data originally was stored in the relational databases and how the conversion to the graph database was performed. This explanation is divided in two sections, in which the first section describes the conversion of the medium-voltage network topology and the second section the conversion of the low-voltage network topology.

1. Construction of the medium-voltage network topology

With respect to the medium-voltage distribution networks, two main type of assets were distinguished, namely: the medium-voltage cables and medium-voltage transformer stations. The relational tables which were used in order to convert the medium-voltage network topology to the graph, are described in the following paragraph.

Relational tables storing the medium-voltage grid network topology

- Table 'MV_NRG_MS_VELD'

This relational table consisted of all the medium-voltage power sections interconnecting all different transformer stations. Important attributes of this table were the power section's object ID, the ID's of the connecting transformer stations, the ID of its substation and the direction of the current flow. The substation ID referred to the power section's supplying high-voltage substation. Since the study area of this research was defined by the distribution networks supplied by the substation of Hoorn Holenweg, this substation ID was used in order to obtain all the power sections supplied by the substation of Hoorn Holenweg.

- JSON files

JSON files stored all the individual medium-voltage cables belonging to a particular medium-voltage power section. This means that each individual medium-voltage power section had an own JSON file storing all its particular cable assets. Besides the cables, this table also included the interconnectivity of the cables (which cables were connected to each other) via a cable's subsequence number. Important attributes of this table therefore were: the station ID's of the cable's departure- and destination station, the ID of the cable's medium-voltage power section, the cable's subsequence number and the cable's own asset ID.

- Table 'MV_NRG_STATION'

This table consisted of all the high-voltage substations, medium-voltage transformer stations, control stations and low-voltage distribution boards of Alliander. Important attributes of this table were: station ID, type of station, station number and the date of valuation. Via the station number, several of other data, stored in other tables, were obtained as well, such as the installations and transformers the stations actually contained.

- Table 'MV_NRG_MS_SCHAKELAARS'

This table consisted of all the circuit breakers in the medium-voltage grid network of Alliander. A circuit breaker is an asset in the grid network that can be used to interrupt the current flow. If a circuit breaker is closed, the current flow can pass, otherwise the current flow is interrupted. Circuit breakers were important assets since they enable different configurations and prevent many users from being without power for a long time during a breakdown in the grid network. However, because these assets were not assigned to any costs, circuit breakers were included as individual assets (nodes) in the network graph neither. Nonetheless, the location and position of a circuit breaker was still relevant since it defined which parts of the grid network could actually be used by a specific customer. With respect to the construction of the graph database, this means that the positions of circuit breakers were integrated in the graph database by means of a relationship property telling whether the standard mode of a circuit breaker was closed or open. This implies that at the location of an open circuit breaker, the relationship between the two assets was enriched with a property-value 'open', telling that the current flow is interrupted. In order to acquire the circuit breakers locations, important attributes of this table were the station number near which the circuit breaker is active, its medium-voltage power section ID (on which side of the station is the circuit breaker located), the direction of the current flow (how does the circuit breaker affect the current flow) and the position of the circuit breaker in normal mode. Figure 1 represents how the grid-configuration data of circuit breakers was stored in the network graph.

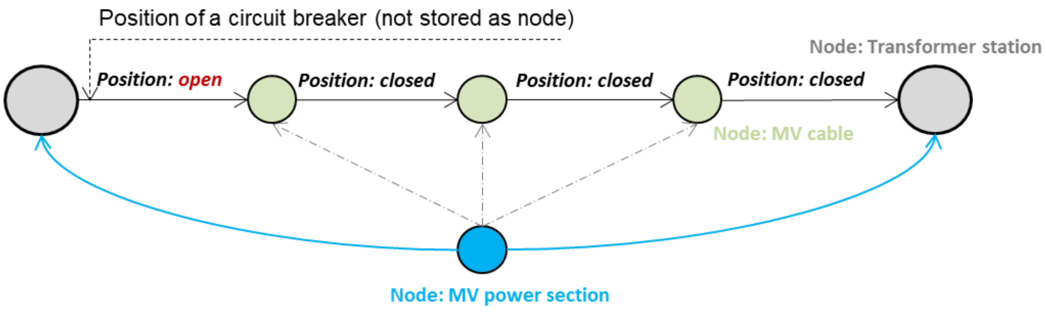


Figure 1: Enrich network with its configuration

- Table 'MV_NRG_GVB_AANSLUITING'

This table consisted of all the (commercial) high energy consumption customers directly connected to the medium-voltage grid network. This group of customers were connected to an own medium-voltage transformer station and therefore did not contribute to the load on the low-voltage distribution networks. In order to relate the customers with its belonging medium-voltage transformer station, important attributes of this table were the EAN code of the customer and the station ID to which it was connected.

- Table 'MV_NRG_MS_KABELS'

This table contained the properties and specifications of all unique medium-voltage cables in the medium-voltage distribution networks of Alliander. By using the ID's of cables, relevant asset data was derived. Attributes relevant for this study were the year of construction, valuation date, cable length, location of the cable, type and owner of the cable.

- Table 'MV_NRG_VERMOGENSTRANSFORMATOR'

This table contained the properties and specifications of all transformers. In order to link a transformer to its belonging medium-voltage transformer station, the foreign attribute 'station number' was used.

- Tables 'MV_NRG_LS_INSTALLATIE' & 'MV_NRG_MS_INSTALLATIE'

This tables contained the properties and specifications of all installations which were part of a medium-voltage transformer station. The foreign attribute 'station number' was used to link the asset with its belonging medium-voltage transformer station.

Write medium-voltage network topology to the graph database

Firstly, the high-consuming customers and medium-voltage assets were stored in the network graph. Medium-voltage cables and stations were derived from the JSON files, that stored the medium-voltage grid network topology. To write a specific medium-voltage cable to the graph database, the following query was used:

Query 1: add node to network graph

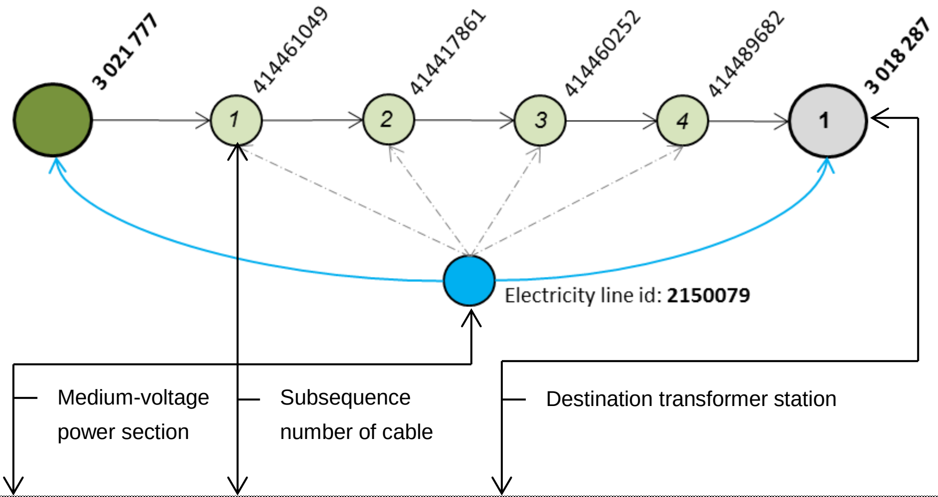
```
MERGE (msk:MSKABEL {asset_id: 414460252 })
```

Adds an unique node to the graph database. If a node with exactly the same attributes already exists, no node is added.

The node has a label *MSKABEL* and is assigned to a variable named *msk*.

The node properties with their corresponding attribute values.

The actual tracing information of cables (how the medium-voltage cables were actually connected to each other) was derived from the JSON files as well. Figure 2 shows a subset of the JSON data and the results of the conversion of this data to the network graph. To determine the interconnectivity of a cable and its position in the entire grid network, the attributes ‘station_to’, ‘station_from’ and ‘cable order’ were used. In case of a subsequence number of 1 or a subsequence number equal to the total number of cables belonging to the specific medium-voltage power section, the specific cable was connected to its departure station or destination station respectively. The medium-voltage transformer stations were added to the graph database by using a query similar to query 1, however, the label was different and properties were obtained from table ‘MV_NRG_STATION’.



asset_id	electricity_line_id	cable_order	station_number_to	station_number_from	HV_substation	station_order
414460252	2150079	1	3 018 287	3 021 777	3 021 777	1
414461049	2150079	2	3 018 287	3 021 777	3 021 777	1
414417861	2150079	3	3 018 287	3 021 777	3 021 777	1
414489682	2150079	4	3 018 287	3 021 777	3 021 777	1

Figure 2. Relational database translated to graph database

To interconnect the medium-voltage cables in the network graph, a relationship labeled as ‘MV cable connection’ was added. This relationship consisted of the property ‘position’, which determined if the current flow can pass or was interrupted (based on the location of an open circuit breaker). For adding a specific relationship between two nodes, the following cypher query was used:

Query 2: Adding a relationship to the graph database

```
MATCH (msk1:MSKABEL) WHERE msk1.asset_id = 414460252,
MATCH (msk2:MSKABEL) WHERE msk2.asset_id = 414461049
CREATE UNIQUE (msk1)-[r:MSKABEL_VERBINDING {position: 'closed'}]->(msk2)
```

Creates a unique relationship between the nodes assigned to variable msk1 and msk2

msk1 is the node selected in the first match-statement.

Specify the relationship with its label, direction and properties. The label name is 'MV cable connection' and the relationship has a property 'position'.

To complete the medium-voltage network topology, the actual distribution assets were enriched with their particular asset data, such as the material, activation date and length of the cable. This information was stored in the table "MV_NRG_MS_KABELS" and could be related to the topology via its asset id. By adding particular asset data to the graph, a distinction was made between unique asset data and data applicable for a large number of assets (general characteristics). Information which was unique for each individual asset was stored as a cable node property while general characteristics were stored as nodes with a relationship to the asset-nodes it applied to. A This structure reduced the size of the database since general characteristics were only added to the database ones.

2. Construction of the low-voltage network topology

Relational data tables storing the low-voltage grid network topology

- Table 'LV_NRG_TRACE'

To construct the low-voltage distribution network, a tracing table (node-link) was used in which all different types of assets occurred. Hence, the LV_NRG_TRACE table consisted of joins, security components, customer connection's and end closures as well. When a specific distribution asset was connected to multiple other assets, the asset was stored as an object as often as it had connections to other assets. The attributes *asset_id* and *object_1_id* represented the asset while the attribute *object_2_id* represented the asset connected to *object_1_id* (see Figure 3). Important attributes of this table were: station number, power section id, id object 1, id object 2, type object 1, type object 2, asset id, cable length and depth.

- Table 'MV_NRG_LS_KABELS'

This table consisted of the properties of all specific low-voltage cables in the grid network of Alliander. By using the asset ID as foreign key, important asset information of relevant low-voltage cables was derived. Attributes relevant for this study were the valuation date, location, type and number of cores of the cable.

- Table 'MV_NRG_LS_AANSLUITING'

This table consisted of all the customers connected to the low-voltage distribution networks of Alliander. Important attributes of this table were the asset id, EAN code, standard annual energy consumption, nominal capacity and type of the customer. By using the asset id as foreign key, a customer was related to the specific low-voltage cable to which it is connected.

Write low-voltage network topology to the graph database

The LV_NRG_TRACE table stored all the individual low-voltage distribution assets of Alliander. Hence, this table consisted of tens of thousands of objects. In order to obtain the relevant low-voltage assets only, a subset was made based on the medium-voltage transformer stations that occurred in the already constructed medium-voltage network topology. Each medium-voltage transformer station consisted of several outgoing low-voltage distribution networks of which each outgoing network was characterized by an unique low-voltage power section id. Based on this id, cables were added to the network graph in which the first cable of a specific low-voltage power line (depth is 0) was connected to the transformer station. However, in contrast to medium-voltage power sections, the last cable object of a low-voltage power section is usually not connected to another transformer station but ends with an end closure object (fishbone structure). However, in some cases the last cable was connected to a low-voltage distribution boards which again splitted up in several outgoing low-voltage distribution networks. When these outgoing low-voltage networks connected with another transformer station, the low-voltage distribution network was designed according to the loop-structure design principle.

Because cables are interconnected by joints, in the relational database, cable objects were always followed by a joint object (node-link structure). In the context of this study, however, joints were not relevant and therefore were not added to the network graph. Because the LV_NRG_TRACE table consisted of all the different types of distribution assets, this table could not directly be implemented in the graph database without adaption. In the construction of the graph database, joints and security components were removed and a relationship labeled as 'LV cable connection' connected the cable objects which were physically interconnected. Figure 3 shows how assets and the interconnectivity of assets were stored in the relational database and how they were implemented in the network graph. In this figure, black arrows represent the interconnectivity of assets as it is in practice while the green dotted arrows represent the interconnectivity of assets as implemented in the graph database.

Object_1_type	Object_1_Id	Object_2_Id	Object_2_type	Asset id
Cable	454278062	454278068	Connection socket	454278062
Connection socket	467874482	454278062	*Cable*	467874482
Cable	806033878	467874482	Connection socket	806033878
Customer connection	454125017	806033878	*Cable*	454125017
Cable	806033875	454125017	Customer connection	806033875

Connected to (graph)

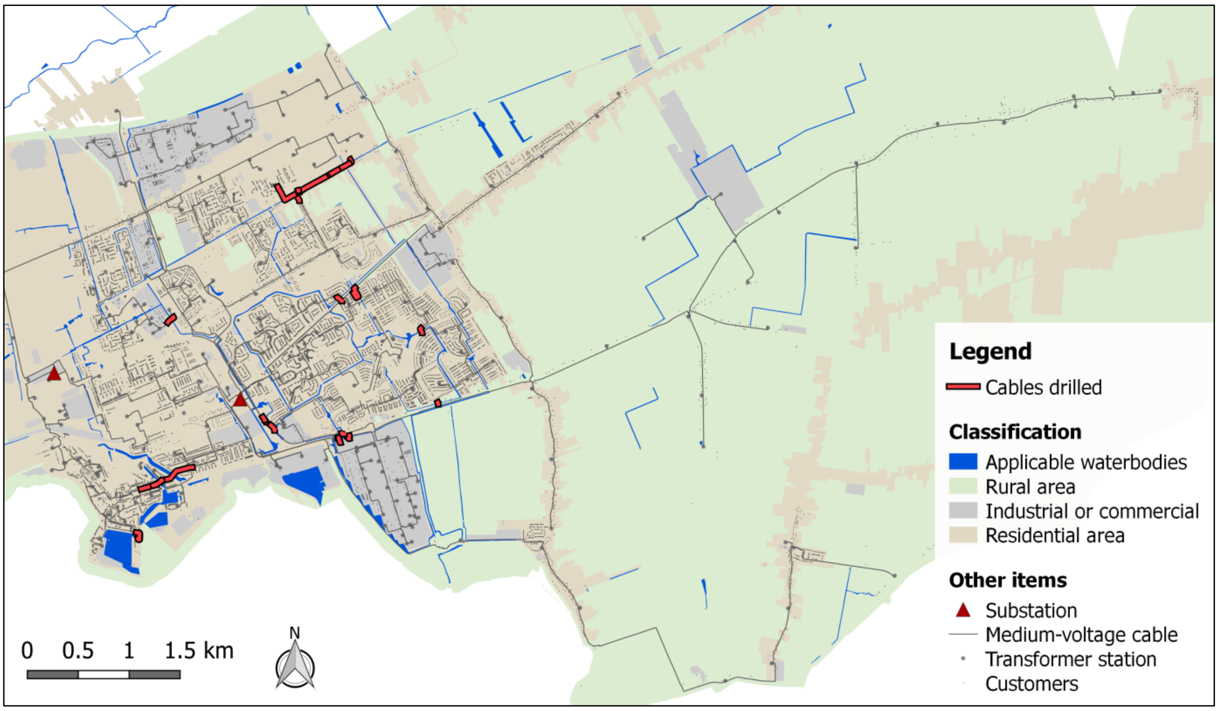
Connected to (oracle)

LV- cable objects

Socket object, customer connection object or security object

Figure 3: Low-voltage network data as stored in the relational database and graph

Appendix II Land use classification result



Appendix III Charts of customers allocated costs

Appendix IV Charts of the economic efficiency of assets

Appendix V Potential applications supported by the method

In addition to the generation of economic insights of assets and customers, another objective of this study was to calculate the economic impact of a traditional grid expansion versus the impact of adding a distributed generation source to the grid network. Figure 1 shows the economic impact on the customers allocated network costs as a result of a traditional capital investment in the grid network (grid expansion) as well as the integration of a distributed generation source.

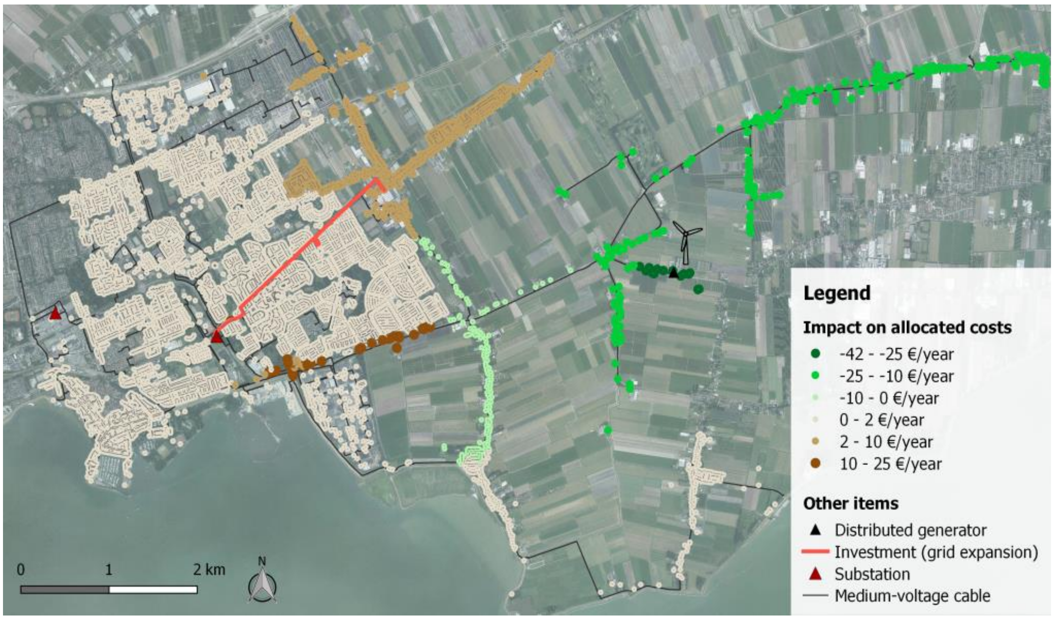


Figure 1: *The economic impact of a capital investment and a power unlimited distributed generator*

Figure 2 shows the locations of not used medium-voltage power sections, showing that the customers located downstream of the unused cables no longer made use of the substation's energy supply. As a result, the depreciation costs of the substation were not distributed over these customers, causing that all other customers were allocated slightly higher depreciation costs.

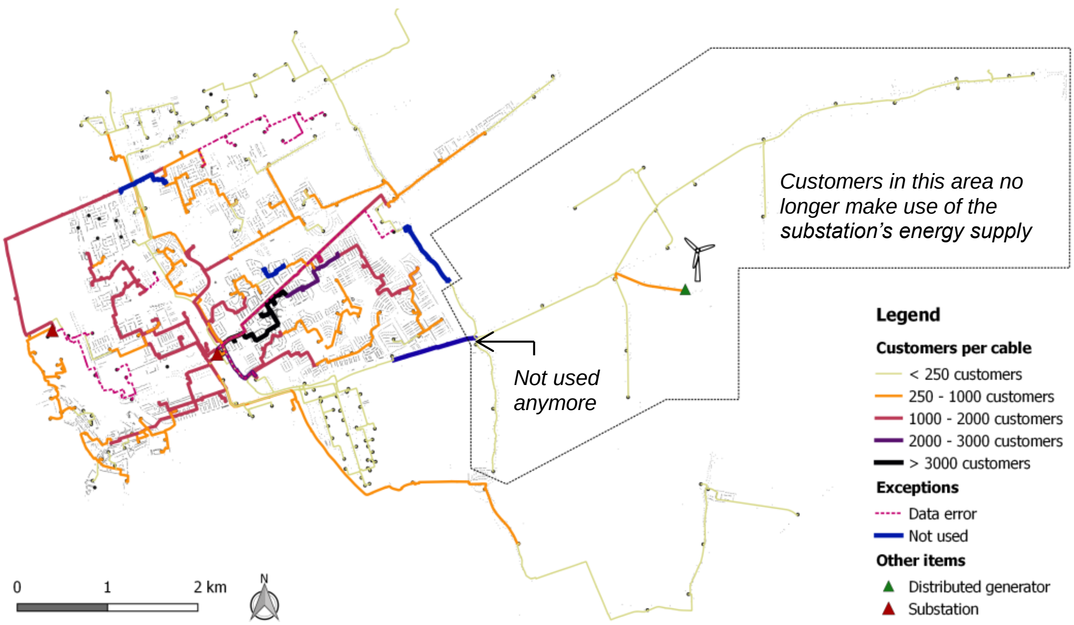


Figure 2: *Number of asset dependent customers of all medium-voltage cables after adding a power unlimited distributed generator*

Appendix VI Errors in digital network graph

Table 1: Explanation of relevant errors in the network graph

No.	Name	Description
1	Loop structure in low-voltage distribution networks	Based on the network built for this research, it was found that a notable number of low-voltage distribution networks were connected according to the loop structure design principle. However, regarding to the operation of low-voltage networks, it is unlikely that these networks will actually be found in this structure. This implies that that the network data of especially the low-voltage distribution networks may contained of topological errors. However, it cannot be validated whether the loop structured low-voltage distribution networks were actually erroneous or not: it is a suspicion which must be checked manually on location. This means that, in the context of this research, loop-structured low-voltage distribution networks were assumed as being correct.
2	Erroneous medium-voltage distribution networks	The study area of this research was defined by the high-voltage substation of Hoorn Holenweg. However, two medium-voltage networks were digitally connected to the substation of Hoorn Holenweg, while physically supplied with electrical energy by another substation (see Figure 1). This means that these networks in reality do not belong to the substation of Hoorn Holenweg. During the construction of the network graph, this caused that these medium-voltage distribution networks were not enriched with their downstream low-voltage network topology. Therefore, the two erroneous medium-voltage networks were not interconnected with any low-voltage customer. The customers directly connected to the erroneous medium-voltage networks (wholesale customers) were, however, connected. The allocated costs to these three customers therefore were highly affected by the incomplete network topology of their dependent assets and thus needed to be removed in the method's output.
3	Uncertainty in valuation dates (1)	A substantial population of assets had a valuation date of April 30th. When an asset was subject to a valuation date of April 30th, the actual valuation year of the asset was predicted based on data of other (adjacent) assets. This implies that the valuation year of the specific asset was not 100% certain.
4	Uncertainty in valuation dates (2)	A small population of assets consisted of predicted valuation years with very high uncertainty. Assets which had a valuation date of December 25th were enriched with a completely unknown valuation year. The valuation years of these assets could not be predicted based on adjacent assets. Hence, the entire asset population with no traceable information was systematically enriched with valuation years based on the relative populations of assets in the existing valuation years. Assets characterized by a valuation date of December 25th hence consisted of highly uncertain data.

Figure 1: *Erroneous medium-voltage distribution network.*
The black ellipse shows that two medium-voltage networks were not physically connected to the high-voltage substation of Hoorn Holenweg, but to another substation. Furthermore, it can be seen that the outgoing distribution networks (visualized in red) were not enriched with their downstream low-voltage distribution networks. Three wholesale customers were directly connected to these network.