Distribution grid planning and operational principles for EV mass roll-out while enabling DER integration

Deliverable (D) No: 1.4

State-of-the-Art Methods Report
A Report on Existing Tools and Methods Including Latest Innovations

Author: TRACTEBEL
Date: 20.12.2013
Version: 1.6

www.PlanGridEV.eu

Confidential (Y / N): N

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 608957.
<table>
<thead>
<tr>
<th>WP number</th>
<th>WP title</th>
<th>WP leader</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assessment and gap analysis</td>
<td>Johannes Stöckl</td>
</tr>
</tbody>
</table>

**Task title**

T1.1: Identify current grid planning rules, existing network architecture and assessment of critical load scenarios

**Main Authors**

- Pedro Rocha Almeida / Tractebel
- Stéphane Rapoport / Tractebel
- Marina González Vayá / ETH
- Thilo Krause / ETH
- Luis Barino Morales / ETH
- Stefan Ubermesser / AIT
- Johannes Stöckl / AIT
- Alexandre Dias / INESC
- Pedro Carvalho / INESC
- Luis Silvestre / EDP

**Project partners involved**

- Johannes Kathan / AIT
- Sara Ghaemi / AIT
- EDP
- RWE
- ENEL
- ENEL

**Type (Distribution level)**

- PU, Public
- PP, Restricted to other program participants (including the Commission Services)
- RE, Restricted to other a group specified by the consortium (including the Commission Services)
- CO, Confidential, only for members of the consortium (including the Commission Services)

**Status**

- ☐ In Process
- ☐ In Revision
- ■ Approved

**Further information**

www.PlanGridEV.eu
Executive Summary

This document presents a complete review of the state-of-the-art concerning the methods applied to distribution planning. It is a key element for the development of PlanGridEV as it provides a reference regarding existing methods and tools, being an important input in Work Package 4, where new and innovative methods will be developed.

On one hand, this report benefits from the experience of the Distribution System Operator (DSO) members on the PlanGridEV consortium and their knowledge about their own grids and planning rules. A complete description of DSOs’ grids and current planning rules is provided in deliverable D1.1. On the other hand, research and power system consulting partners grow upon this knowledge, exploring in depth the classical distribution planning problem and the effort made up to this date to improve and redesign distribution planning tools to meet present and future needs.

General information related to grid characteristics is provided in D1.1 for Medium Voltage (MV) and Low Voltage (LV) grid of PlanGridEV partner DSOs. Their planning activity is based on the principle of maintaining the electrical system capacity to meet future demand, while maintaining service quality levels consistent with regulatory requirements and also minimizing the environmental impact of the assets. The choice of appropriate planning criteria is important to ensure a progressive improvement of safety standards and quality of electricity energy distribution under criteria of technical and economic efficiency, along with risk analysis and environmental concerns.

Traditionally, the distribution planning problem has been solved in a stepwise process that includes some simplifications, including considering consumers as passive elements and dimensioning always for the most severe operation scenario. The main goals are to meet the highest peak load demand within the required reliability standards and for the smallest possible cost.

The first distribution planning methodologies followed a deterministic process, since the existing computational power and availability was limited. In time, parts of the process were automatized, but the main rationale remained unchanged. Recently, increasing levels of Distributed Generation (DG) plus the expected rollout of Electric Vehicles (EVs) have been introducing uncertainties. The worst case scenario that should be evaluated is no longer necessarily peak load, as off-peak conditions could cause voltage and reactive power problems in presence of DG. At the same time, there is a greater concern in developing longer term plans with the prospect of achieving better overall solutions.

Advances in smart grids and Demand Side Management (DSM) have been made as a response to these challenges. Hence, the planning assumptions must be revised to effectively integrate and consider the potential benefits of these concepts. Major advances have been made in terms of new operation scenarios including DG and/or DSM, but their integration into planning lags behind.
Additionally, the new distribution paradigm of smart grids with active participation of Distributed Energy Resources (DER) and load in network operation deeply relies on an adequate communication infrastructure. Thus, the challenges of communication must be understood and incorporated into the planning problem as alternatives to conventional reinforcements.

The analysis of all existing publications provides a good overview of existing work in the field of distribution planning. It is verified that most publications focus on the development of new optimization algorithms, while formulating the problem more or less in a similar way. Their main concern so far has been to develop robust and faster tools to address the classical planning problem. Yet, there are also a number of authors that propose the inclusion of novel features to better address the problem and its evolving needs. Some consider the reliability issue by including a cost for the non-supplied energy, but only a few consider reliability as an objective per se. Most papers consider deterministic load modelling. Nonetheless, some authors already propose ways for dealing with the uncertainties related to load and generation using stochastic or fuzzy methodologies. A few go even further and describe the usage of risk assessment methodologies to aid the planners in making the best possible decisions. Part of the works that formulated multi-objective problems proposes the development of decision-aid methodologies for selecting the best investment decisions among dominated and non-dominated solutions.

Future distribution planning tools should include better representations of the uncertainties posed by DER. DER units based on intermittent energy sources such as wind and solar require complex modelling for grid planning, where the energy availability also needs to be represented. Besides using measured data from reference DERs for grid planning, methods for generating artificial data based on statistical methods are also applicable.

There are also many models used to represent the charging flexibility of EV fleets. In the EV aggregation models to be developed in PlanGridEV, it is crucial to include network constraints at the distribution grid level. The idea is to have location-dependent aggregations to be able to formulate all network constraints explicitly. Each network node of the MV network would have its corresponding virtual storage, while EVs move from node to node will be modelled as energy transfers from one virtual storage to another. Whereas the state-of-the-art provides a good starting point, existing models will need to be adapted to the specific requirement of the planning tool to be developed in PlanGridEV.
Table of contents

Executive Summary ............................................................................................................................ 5
Table of contents ............................................................................................................................... 7
List of figures ..................................................................................................................................... 9
List of tables ..................................................................................................................................... 10
Abbreviations and Acronyms ........................................................................................................... 11
1. Introduction .................................................................................................................................. 13
   1.1. Scope of the document .......................................................................................................... 13
   1.2. Structure of the document ..................................................................................................... 13
2. Classical Distribution Planning Problem .................................................................................... 14
   2.1. Planning Process .................................................................................................................... 14
   2.2. Problem Formulation ............................................................................................................ 15
   2.3. Current DSO Tools and Methods .......................................................................................... 16
3. New Distribution System Planning Problem ............................................................................... 19
   3.1. Drivers .................................................................................................................................. 19
   3.2. Generic Problem Formulation ............................................................................................... 19
   3.3. New Approaches .................................................................................................................. 20
4. Literature Review on Distribution Planning ............................................................................... 21
   4.1. Optimization Problem and Tools .......................................................................................... 21
      4.1.1. Distribution Planning ...................................................................................................... 21
      4.1.2. Distribution System Impacts and Operation ....................................................................... 26
      4.1.3. Distribution Planning Tools ............................................................................................ 29
   4.2. Statistical Modelling of DER ............................................................................................... 31
      4.2.1. Statistical modelling of wind generation ........................................................................... 31
      4.2.2. Statistical modelling of solar generation ........................................................................... 35
      4.2.3. Statistical Modeling of EVs ............................................................................................. 41
      4.2.4. Applicability of statistical models for PlanGridEV objectives ......................................... 48
   4.3. Modelling Storage Capabilities of EV .................................................................................... 48
      4.3.1. Virtual storage-type models ............................................................................................ 49
      4.3.2. Task-type models ............................................................................................................ 51
      4.3.3. Overview and conclusions ............................................................................................... 52
   4.4. Multi-objective Optimization Approaches in Distribution Planning .................................... 53
5. Conclusions .................................................................................................................................. 56
6. References ..................................................................................................................................... 58
7. Revisions ....................................................................................................................................... 65
Annex A. Distribution Planning Tools ............................................................................................... 66
   DPLAN ........................................................................................................................................ 66
   NEPLAN/Integral ......................................................................................................................... 67
   MSPLAN ....................................................................................................................................... 67
   SynerGEE Electric ......................................................................................................................... 68
   PECO Model ................................................................................................................................. 68
   WindPRO ..................................................................................................................................... 70
   Smallworld GIS – Operation Analyser ......................................................................................... 71
   CYMDIST - Distribution System Analysis .................................................................................... 71
   PSS®SINCAL ................................................................................................................................. 72
   LAPER ......................................................................................................................................... 72
List of figures

FIGURE 1 - NEW DISTRIBUTION PLANNING PROCESS (WITH INTERACTION WITH SUB-TRANSMISSION PLANNING), ADAPTED FROM [14] ................................................................. 19
FIGURE 2 - TYPICAL WIND POWER OUTPUT WITH STEADY WIND SPEED [72] ...................... 32
FIGURE 3 - CONSTRUCTION OF A SIMULATED DATA SEQUENCE [71] ........................................ 34
FIGURE 4 - MARKOV CHAIN FOR THE REPRESENTATION OF WIND SPEED [77] ..................... 35
FIGURE 5 - COMPARISON OF MEASURED WIND SPEED DATA AND GENERATED WIND SPEED DATA [73] 35
FIGURE 6 - CLASSIFICATION OF MODEL BASED ON SPATIAL AND TEMPORAL RESOLUTION [78] ...... 36
FIGURE 8 - SIMPLIFIED DIAGRAM OF BIOLOGICAL AND ARTIFICIAL NEURON [79] .................... 38
FIGURE 9 - MLP WITH INPUTS AND HIDDEN UNITS [78] .............................................................. 39
FIGURE 10 - TOPOLOGY OF WAVELET NEURAL NETWORK [78] .................................................... 40
FIGURE 11 - MARKOV CHAIN FOR EV USAGE [63] ................................................................. 43
FIGURE 12 - DISTRIBUTION OF STATES (EXAMPLE) [89] ......................................................... 43
FIGURE 13 - MEAN ERROR OF EV ENERGY DEMAND FORECAST [89] ...................................... 44
FIGURE 14 - UNCERTAINTY IN ENERGY BOUNDS (10000 SAMPLES) [90] ............................... 44
FIGURE 15 - OPEN STREET MAP EXPORT (SCREENSHOT) ......................................................... 46
FIGURE 16 - STREET NETWORK IN MATSIM BASED ON OSM DATA ...................................... 46
FIGURE 17 - SCHEME OF CO-SIMULATION ENVIRONMENT WITH MATSIM, EVSIM AND POWERFACTORY VIA OPC CONNECTION [87] ......................................................... 47
FIGURE 18: PARETO CURVES YIELDED FOR HIGH (DASHED LINE) AND LOW (SOLID LINE) TOLERANCE IN DOMINANCE DEFINITION. ................................................................. 55
List of tables

TABLE 1 ACRONYMS ................................................................. 11
TABLE 2 – SUMMARY TABLE OF REVIEWED DISTRIBUTION PLANNING LITERATURE .......................... 22
TABLE 3 – COMPARISON OF DIFFERENT DISTRIBUTION PLANNING TOOLS .................................. 30
TABLE 4 - EXAMPLE OF THE COMPOSITION OF AN EV FLEET IN A LOW-VOLTAGE DISTRIBUTION GRID [87] ................................................................. 42
TABLE 5: MAIN CHARACTERISTICS OF REVIEWED METHODS. ........................................................ 53
### Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>AR</td>
<td>Auto Regressive</td>
</tr>
<tr>
<td>ARIMA</td>
<td>Auto Regressive Integrated Moving Average</td>
</tr>
<tr>
<td>ARIMAX</td>
<td>Auto Regressive Integrated Moving Average with Exogenous Variables</td>
</tr>
<tr>
<td>ARMA</td>
<td>Auto Regressive Moving Average</td>
</tr>
<tr>
<td>CSP</td>
<td>Charging Service Provider</td>
</tr>
<tr>
<td>DER</td>
<td>Distributed Energy Resources</td>
</tr>
<tr>
<td>DG</td>
<td>Distributed Generation</td>
</tr>
<tr>
<td>DSM</td>
<td>Demand Side Management</td>
</tr>
<tr>
<td>DSO</td>
<td>Distribution System Operator</td>
</tr>
<tr>
<td>EV</td>
<td>Electric Vehicle</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>HV</td>
<td>High Voltage</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communications Technology</td>
</tr>
<tr>
<td>LV</td>
<td>Low Voltage</td>
</tr>
<tr>
<td>MA</td>
<td>Moving Average</td>
</tr>
<tr>
<td>MLP</td>
<td>Multi-Layer Perception</td>
</tr>
<tr>
<td>MV</td>
<td>Medium Voltage</td>
</tr>
<tr>
<td>PHEV</td>
<td>Plug-in Hybrid Electric Vehicle</td>
</tr>
<tr>
<td>UCM</td>
<td>Unobserved Components Model</td>
</tr>
<tr>
<td>WNN</td>
<td>Wavelet Neural Network</td>
</tr>
</tbody>
</table>
1. Introduction

1.1. Scope of the document

This document presents a complete review of the state-of-the-art concerning the methods applied to
distribution planning. It is a key element for the development of PlanGridEV as it provides a reference
regarding existing methods and tools, being an important input in Work Package 4, where new and
innovative methods will be developed.

On one hand, this report benefits from the experience of the DSO members on the PlanGridEV
consortium and their knowledge about their own grids and planning rules. A complete description of
DSOs’ grids and current planning rules is provided in deliverable D1.1. On the other hand, research
and power system consulting partners grow upon this knowledge, exploring in depth the classical
distribution planning problem and the effort made up to this date to improve and redesign distribution
planning tools to meet present and future needs.

In particular, the following elements are studied:

- Distribution planning methodologies and algorithms
- Advanced control and operation of distribution systems that should be considered in the
  planning process
- Statistical modelling of Distributed Energy Resources (DER), namely wind generation, solar
  photovoltaic (PV) generation and Electric Vehicles (EVs)
- Modelling of storage capabilities of EV
- Multi-objective optimization approaches in distribution planning

1.2. Structure of the document

This report is composed by three main sections:

- 2. Classical Distribution Planning Problem – the classical distribution planning problem is
  formulated and discussed
- 3. New Distribution System Planning Problem – a generic review of drivers and possible
  formulation of the future distribution problem is presented
- 4. Literature Review on Distribution Planning – specific methods and algorithms are review

These sections are complemented by the work presented in deliverable D1.1, namely sections:

- 2. MV and LV Grids Overview – the MV and LV grids of the 4 PlanGridEV partner DSOs are
  described
- 3. Current Planning Rules – the DSO planning rules are exposed
2. Classical Distribution Planning Problem

Distribution planning is defined as a complex decision-making process usually formulated as a multi-objective optimization problem. The optimal solutions represent the best possible trade-off between network investment costs, energy loss costs, and reliability costs, taking into account practical, and regulatory, operational, and technical boundary conditions, [1].

2.1. Planning Process

The problem of distribution network planning

The goal of electrical distribution networks is to make electrical power available and deliver electrical energy to the consumers. Nowadays, since electrical energy is ubiquitous, the investment in distribution networks is mainly motivated by load growth and improvement of the quality of service.

The problem commonly known as distribution network planning is a real problem that DSOs face periodically. Network planning consists in establishing a decision schedule with the aim of solving problems in current networks or potential problems in the future. The appropriate definition of the decision schedule has important financial repercussions – it represents an important direct benefit to the DSO’s and an indirect benefit to consumers.

In electrified geographical areas, network energy efficiency and quality of service are important concerns for the DSOs, as power losses are a significant part of the distributed energy, and power interruptions as well as low voltage levels have important repercussions, both financially and regarding company's image towards consumers. Depending on the existing regulation quality of service is subject to an economical valuation. In Portugal the regulator sets an incentive based scheme that introduces additional revenue if interruption period goal is improved and a penalty if the goal is not achieved, [2].

Main difficulties associated with real problems resolution

The electrical distribution system has some characteristics that make the network planning problem peculiar. Some important requirements of the methods to address the problem are associated with some of these characteristics. For those requirements that are representative of important methodological difficulties, some important characteristics are here enumerated in the following:

1) **High durability of electrical facilities.** The facilities durability is significant when compared to the typical planning horizon. In the decision schedule, investment decisions taken in previous stages influence the system state in future schedule stages. Hence, the planning problem is **dynamic**.

2) **High cost of the electrical facilities.** The cost of the facilities is rather significant when compared to the operation and maintenance costs of the distribution system. For distribution networks in general, the set of economically justifiable decisions is a small part of the entire existing facilities of the system. The problem solutions usually involve a **combinatorial** set of only a few from many possible investments.

3) **High uncertainty over future system needs and expansion constraints.** It is frequent for consumer needs to be revealed in the short term, and for delays in decision making
and in obtaining authorization and funding to be uncertainty. Uncertainty over customer needs and construction delays make the planning problem **stochastic**.

4) **High diversity of observable problems and ways of acting.** It is frequent for the network to exhibit problems such as conductor rating violations, allowable bus voltage limit violations, excessive losses and deficient quality of service. The ways of acting to solve these problems can be diverse and include different kind of decisions such as conductor and transformer replacement, network reconfiguration by modifying the state of the existing switchgear equipment, network reinforcement as building of new network lines and new substations. The diversity of problems and the different possible decision types make the problem **complex**.

Thus, the role of the DSOs is to solve a diverse set of problems by defining a schedule of different type of decisions associated to significant investments and with an important impact in a system whose evolution is uncertain. With the methods currently used, finding the optimum solution to this problem is usually a hard task. The methods used by DSOs generally involve problem breakdown into smaller problems whose complexity allows targeted search by a planning engineer with the aid of analysis tools. These tools are typically computational tools that provide power flow and reliability analysis.

**2.2. Problem Formulation**

**Deterministic planning problem formulation for distribution planning optimization**

The following notation is used in the following:

- \( I \) : Investment cost function
- \( L \) : Energy loss cost function
- \( R \) : Reliability penalty function
- \( G \) : Graph of the distribution network
- \( T \) : Set of spanning tree configurations of \( G \)
- \( E \) : Set of electrically feasible configurations
- \( x \) : Graph configuration
- \( y \) : Spanning tree configuration
- \( e \) : ENS unitary price
- \( c \) : Energy unitary cost
- \( r \) : Reliability price factor \( e/c \)

Distribution networks are constituted by several feeders from different substations. Each feeder defines a radial configuration from the substation transformers to the loads. Frequently, feeder-ends are connected with each other within the same feeder and with other feeders through normally open tie-switches. Hence, the distribution network topology consists in a graph over which some open tie-switches define a radial configuration, i.e., a spanning tree. If the universe of decisions is defined to be a set of single arc decisions, the constraints are of two different classes: (i) topological constraints, such as connectivity and radiality, and (ii) power flow constraints, such as minimal node voltage and maximal branch current. The optimal planning problem consists in finding the best possible tradeoff between investment costs, energy losses costs, and reliability costs.
A graph $G=(N, A)$ is a pair of sets: one set of nodes $N$ and one set of arcs $A$. In the distribution context, the graph nodes are the load points, the busbars, and the connection points; and the graph arcs are the lines branches, cables branches, and switch devices. The problem solution must represent both a subgraph of $G$ and a spanning tree of $T$, with $G$ representing the existing network nodes, the existing network branches, and the new possible branch investments, and $T$ represents the set of all possible radial configurations for the distribution network normal operation, i.e., for power flow purposes. The optimal planning problem may be formulated as in the following:

$$\min_{x, y} I(x) + L(y) + R(x, y)$$

subject to:

$$x \in G$$

$$y \in x \cap T \cap E.$$  \hspace{1cm} (1)

Here, $L$, $R$, and $I$ stand for losses, reliability, and investment cost functions, respectively. Those functions are described next.

E. Investment Cost Function
The investment $V$ in the starting year can be computed from the summation of the graph $x$ arc investments for given cable price and installation cost $f$ and branches length $l$. The investment cost in the planning period is given by the difference between the total investment and the current value of the residual $V_r$. The current value of the residual $V_r$ represents the investment value after the planning period $N$ for a given investment amortization period $M>N$ and $\chi$.

C. Energy-Loss Cost Function
The energy losses $W_N$ in the target year $N$ can be computed from the power losses in each branch of $y$ for a specified annual load factor $\alpha$. To compute the power losses an exact AC power flow should be used. The energy losses in year $k$ of the planning period are computed from the losses of the target year $W_N$ for a given load growth $\beta$. The current value of the energy losses cost in the planning period is then computed from the current value of the energy losses in each year $k$ for a given energy cost $c$ and discount rate $\chi$.

D. Reliability Cost Function
The expected value of the ENS (or other reliability indicators) in the target year $N$ can be computed numerically. For ENS, branch faults are enumerated exhaustively (for each and every arc of $y$) and the correspondent $n-1$ configurations analysed. For each fault, the processes of isolation (with time $T_i$), reconfiguration (with time $T_s$) or repair (with time $T_r$) are analysed in $x$ and the corresponding ENS computed for a given cable failure rate $\sigma$. The expected value of the ENS in year $k$ of the planning period is computed from the ENS in the target year $E_N$ for load growth $\beta$. The current value of the reliability penalties in the planning period is then computed from the current value of the ENS in each year $k$ for a given ENS price $e$ and $\chi$ (similar process for other reliability indicators).

Network automation, like other investment hypothesis to improve reliability, changes the parametric definition of the reliability function, i.e., the values of the mean time to isolate or reconfigure faults. These hypotheses are parameter hypothesis like load growth or discount rates — their role can be accommodated by the proposed methodology.

2.3. Current DSO Tools and Methods

Classic method
The classic planning methodology consists in (i) delimiting the geographical area and the voltage levels to be analyzed, (ii) generating one or more network evolution scenarios, (iii) finding one more possible solutions for the problems found in each evolution stage and for each possible scenario, and (iv) building a schedule of decisions by adapting solutions found independently for each stage in order to obtain a chronological order of investment decisions that corresponds to the sequence of system states.

Delimiting the existing network involves defining network boundaries. A network is rarely isolated in what concerns connections to neighboring networks. However, it is frequently possible to identify influence areas of one or more distribution substations. The boundaries are typically defined by identifying a set of branches (lines or cables) whose existence does not influence significantly the performance of the network under analysis.

Generating a set of network evolution scenarios typically consists in the definition of a coherent set of possibilities for the (1) installation of new substations, (2) installation of new network branches, (3) load evolution in the existing buses and (4) new load locations.

The process of finding possible solutions for the problems found in a network, either existing or planned each of the evolution stages, typically consist of a sequence of exploration procedures and economic analysis that can be summarized as follows:

1) Analyze the network and identify the most significant problems
2) Select a set of decisions that solve the identified problems
3) Combine decisions from the set and update the existing network accordingly
4) Evaluate the tradeoff between decision costs and operational benefits and repeat steps 1 to 3 until the tradeoff is satisfactory and no significant problems are found.

Building a decision schedule from the solutions found in an independent way for each stage and for each scenario, in general, requires compromise. If one builds the schedule from the short term to the planning horizon, the long term decisions quality is subjected to the quality of the short term decisions, whereas if the schedule is built from the planning horizon to the short term, the short term decisions quality is subjected to the quality of the long term decisions.

Deficiencies of the classic planning method and the importance of optimization automation

The process of finding solutions to the distribution network planning problem when based on a targeted search lead by a planning engineer with the aid of analysis tools exhibits insuperable methodological deficiencies. Some important deficiencies are the following:

1) The universe of solutions is insufficiently explored for large-scale networks. Although the investment decision set for each schedule stage is typically not very large, the solutions set that can be obtained by combination of these decisions is typically very large. The analysis of large-scale networks is a complex and exhaustive process because it requires the observation of many parameters and because it involves simulating the best possible operation of the existing network in the search for its best possible performance.

2) The definition of the decision schedule is hierarchized. The schedules universe is not explored. The schedule is obtained a posteriori as a result of a chronological sequence of satisfactory solutions. The dynamic relationships between decisions of different stages are not conveniently explored from solutions obtained in an independent way for each stage.
3) The definition of the decision schedule does not assure robustness under uncertainty. The solutions obtained in an independent way for each scenario generally involve conflicts between investment decisions in the earlier stages.

**Optimization automation methods**

Over the last few years computer aided methodologies to approach the distribution network planning problem have been proposed with the objective of overcoming some of the inherent deficiencies of the search methods based on human oriented exploration with support on analysis tools. The generality of the recently proposed approaches are based on computational optimization methods. They are classified according to the deficiencies they intend to overcome and the employed optimization methods.

Considerable work has been carried out on optimal planning of distribution systems. The problem has been addressed by mathematical programming for linearized objectives: branch-and-bound applications can be found in [3], [4], [5]; mixed-integer programming approaches can be found in [6] and [7] together with Bender’s decomposition in [8]. Local search procedures such as branch-exchange [9], [10] and other heuristic techniques [11] have also been proposed to address the optimal planning problem. Recently, evolutionary algorithms have been taken to address more complex decision-making processes [12], [13].
3. New Distribution System Planning Problem

3.1. Drivers

The classical planning approach has to be adapted towards a new approach that takes into account the new realities in distribution systems. The increasing penetration of DG should be considered, as well as other developments such as electric vehicles, demand side management and other smart grid technologies. These new developments are essentially done at the customer level. A new activity can be added to the distribution planning process, which is customer level planning.

![New Distribution planning process](image)

*Figure 1 - New Distribution planning process (with interaction with sub-transmission planning), adapted from [14]*

**Customer-level planning** accounts for the influence of EVs, DSM or DG, which may impact overall conditions for network expansion. The inclusion of these variables may influence solution costs. The customer level planning, by incorporating EVs, DSM or DG may have a direct impact on the feeder planning and on the spatial load forecasting. Hence, new developments such as EVs, DG and DSM should be taken into account during the traditional planning activities such as feeder, substation, and sub-transmission planning.

3.2. Generic Problem Formulation

The generic problem formulation of chapter 4 is still largely valid, but has to be modified to take into account new evolutions in the distribution system. In the objective function, terms have to be added to account for the cost of installing and using demand and generation flexibility. The investment cost in this case is the cost of ICT infrastructure, while the cost of use is a financial remuneration for the grid users that offer flexibility.
The constraints have to be adapted as well. Possible consideration of radiality constraints could be relaxed, as recent developments by the Cigré Working Group 6.11, [15], indicate that planning should also consider meshed networks as a way to simplify DER integration and Active Distribution Network (ADN) implementation. The same working group indicates that new loads, DER and innovative ADN applications should be considered in distribution planning. Furthermore, the control options provided by active elements within the grid are greatly expanded due to the active participation of demand, dispersed generation and storage.

3.3. New Approaches

The conventional approach can no longer be used without being adapted because the fundamental assumptions are no longer valid. In the light of an increasing share of DG in the distribution grid, mainly based on RES and therefore not predictable, the assumption of determinism can no longer be maintained, but has to be replaced by a stochastic approach. Having a single criterion or objective function may also be no longer desirable, as distribution planning is a multi-faceted problem, where cost plays a major stake, but where there may be other objectives to pursue. Other aspects, such as reliability or environmental impact, should be taken into account. Moreover, the objective function to be considered should allow planners to incorporate criteria that could reflect different perspectives of the different involved stakeholders, in particular criteria that may be enforced by the regulator. Such flexibility would allow planners to study the distribution problem from different viewpoints, regarding external sensitivities or even those of different teams within the DSO (e.g. planning versus operation teams). Also the consumer perspective is more and more important and should be able to be captured within the new planning approaches.
4. Literature Review on Distribution Planning

This section is a survey of the state-of-the-art of current distribution planning trends. In sub-section 4.1 the work conducted on the overall planning problem is reviewed, analysing the existing methodologies to tackle the distribution planning problem. Focus is given to the formulation of the problem, including objective function and constraints, the type of load modelling, the inclusion of novel elements with observability or controllability options (such as demand response) and the chosen algorithms to tackle the problem.

The subsequent sub-sections 4.2, 4.3 and 4.4 follow a similar approach but focusing on the existing literature regarding the different modules that will feed the main planning problems with models and methods that allow dealing with Distributed Energy Resources (DER) and in particular with Electric Vehicles (EVs).

4.1. Optimization Problem and Tools

4.1.1. Distribution Planning

As it has been presented in previous sections, the distribution planning problem may be mathematically formulated as an optimization problem that may consider different objective functions, but whose traditional concern is the minimization of distribution investment costs. This problem is constrained by several technical impositions regarding the physical limits of the different elements that compose the network and the quality of service (voltage profiles, congestion in branches…) and by the maximum regulation of the controllable elements. Different algorithms have been used to tackle this problem.

An extensive review of the research developed in the context of distribution system planning is performed, framing the different research within the generic distribution planning problem presented before.

Conventional planning methods were focused on the original problem of distribution grids, which was mainly caused by the load increase that these networks had to be prepared for and power quality delivered to consumers. Therefore, deterministic scenarios would allow planning the grid for the worst operation conditions and a trade-off between quality of service and solution cost would have to be achieved. The process would be solved in a stepwise approach with limited computer aid.

There are two main drivers for distribution planning evolution. On the one hand, there is the development of the distribution system to integrate the increasing deployment of distributed energy and the smart grid concepts, which have broadened the variables and uncertainties of the planning activities. On the other hand, computational power has had great advances and motivated a lot of research in this field.

Over the years, a lot of effort has been dedicated to improve distribution planning methodologies. This section attempts to summarize the relevant advances in this area of activity. Table 2 synthesizes the surveyed works, by describing the main characteristics of the proposed methodologies. These characteristics are grouped in accordance with the elements that compose the distribution planning problem and the modelling aspects / algorithms that are considered. As the objective of cost minimization is present in all methodologies, it was not considered in the table. It is important to note that this review addressed exclusively the publications directly related to distribution planning. There are many other publications concerning the analysis of distribution networks regarding the impacts of
DG/DSM schemes or new control strategies that may be very useful to incorporate in the planning activities.

The analysis of the distribution planning existing work allows describing an overall picture of the current state-of-the-art. Naturally, the characteristic that was in all analysed works is the inclusion of investment cost and estimate of operational costs in the objective function. Also, technical constraints are typically assessed by solving power flows equations. A few works however are unclear whether this question is addressed and how it is considered. Regarding the planning stages, the most commonly addressed is feeder planning, followed by substation planning.

### Table 2 – Summary table of reviewed distribution planning literature

<table>
<thead>
<tr>
<th>Work</th>
<th>Objective Function</th>
<th>Constraints</th>
<th>Load</th>
<th>Other Variables</th>
<th>Corrective Measures</th>
<th>Algorithms</th>
<th>Planning Stage</th>
<th>Imp. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>[16], [17]</td>
<td>x x x x</td>
<td>x x</td>
<td>x x</td>
<td>x x</td>
<td>x x</td>
<td>x x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[18]</td>
<td>x x x x</td>
<td>x</td>
<td>x x</td>
<td>x x</td>
<td>x x</td>
<td>x x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[19]</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[20]</td>
<td>x x x x</td>
<td>x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[21]</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[22]</td>
<td>x x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[23]</td>
<td>x x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[24]</td>
<td>x x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[7]</td>
<td>x x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[25]</td>
<td>x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[26]</td>
<td>x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[27]</td>
<td>x x x x x x</td>
<td>x x</td>
<td>x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[28]</td>
<td>x x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[29]</td>
<td>x x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[30]</td>
<td>x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[31]</td>
<td>x x x</td>
<td>x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[32]</td>
<td>x x x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[33], [34], [35]</td>
<td>x x x</td>
<td>x x x</td>
<td>x x</td>
<td>x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[36], [37]</td>
<td>x x x x</td>
<td>x x x</td>
<td>x x x</td>
<td>x x x</td>
<td>x x x</td>
<td>x x x</td>
<td>x x x</td>
<td>x x x</td>
</tr>
<tr>
<td>[38], [39]</td>
<td>x x x</td>
<td>x x x</td>
<td>x x</td>
<td>x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[40]</td>
<td>x x x</td>
<td>x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[41]</td>
<td>x x x</td>
<td>x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>[42]</td>
<td>x x x</td>
<td>x x x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Regarding load modelling, most papers consider deterministic models. Some of these include some differentiation per consumer type in order to better distribute the load throughout the consumption nodes. There are a few authors proposing different load models, with fuzzy ([16], [17] and [18]) and
stochastic ([19] and [15]) modelling. Both techniques tackle uncertainty, but the latter considers historical values, whereas the former may also include possibility that do not rely solely in history.

DG is starting to be considered, although many times it is considered as a decision variable, [20], [21], [22] and [23]. This assumption does not suit the needs of system operators, concerning their planning needs, as the system operators do not possess direct control over DG investments.

Reliability is increasingly incorporated in the proposed methodologies. Yet, it has been done differently from case to case. The most common approach is to add a term in the objective function that monetizes the energy not supplied, [7] [20] [15] [33] [22]. It is also included as a constraint, which guarantees that the minimum requirements are met, [32], yet in this case it may be possible to find a solution with the same cost, but better reliability. In a few cases the problem is formulated as a multi-objective problem, [27] [16] [17] [18] [23], thus including a separate objective function for reliability. Thus it is possible to consider more reliability indexes beyond energy not supplied and, in this way, achieve better results. The usage of multi-objective formulation also allows adding multi-criteria decision-aid processes. This helps planners including their experience and sensitivity into the planning process.

Some of the works introduces interesting specificities, such as planning of switching elements ([7]) or performing multi-year planning with differentiated yearly investment plans.

Another survey was performed in [43], where an analysis over existing methodologies for distribution expansion planning was performed. The authors argue that mathematical programming algorithms that provide global optimum solutions are more suited to deal with the distribution expansion problem than heuristics or metaheuristics. However, they point out that there may be advantages of having risk-based analysis, such as those many times associated to some metaheuristic problems. A specific comment is made on substation siting as some work has been developed in this direction. Leaving choices open for the theoretical optimum location for the substation may some times lead to practical unfeasibility of the solution, as there may be variables that are not included in the optimization procedures, such as the existing transmission system routes. Attention was called in this work to the fact that reliability may be more than just the cost of non-delivered energy and as such it should be included in the optimization problem as a new objective and not as a part of the cost objective. Finally, the authors point out that there is some distance between the available work and actual practical implementation. This is visible, for instance, in the lack of concern of existing work with budgetary constraints of system operators. Other issues, such as upgrade possibilities or variable routing and conductor sizing should be further explored.

The analysis of all these publications provides a good overview of existing work in the field of distribution planning. It is verified that most publications focus on the development of new optimization algorithms, while formulating the problem more or less in a similar way. Their main concern so far has been to develop robust and faster tools to address the classical planning problem. Yet, there are also some of authors that propose the inclusion of novel features to better address the problem and its evolving needs. Some consider the reliability issue by including a cost for the non-supplied energy, but only a few consider reliability as an objective per se. Most papers consider deterministic load modelling. Nonetheless, some authors already propose ways for dealing with the uncertainties related to load and generation using stochastic or fuzzy methodologies. A few go even further and describe the usage of risk assessment methodologies to aid the planners in making the best possible decisions. Some of the works that formulated multi-objective problems propose the development of decision-aid methodologies for selecting the best investment decisions among dominated and non-dominated solutions.
4.1.1.1 Detailed Review of Some Publications

Currently, relevant work is being developed on the field of distribution grid planning. The conventional methods were focused on the original problem of distribution grids, which was mainly caused by the load increase that these networks had to be prepared for and power quality delivered to consumers. Therefore, deterministic scenarios would allow planning the grid for the worst operation conditions and a trade-off between quality of service and investment cost would have to be achieved.

The evolution of the distribution system into the smart grid concepts and the increasing integration of distributed energy resources have broadened the variables and uncertainties of the planning activities. Gönen and Rosado, [24], describe an optimal sizing and timing dynamic mixed-integer model for simultaneous substation and feeder expansion. A radiality constraint was added as an optional restriction, in order to meet the planners' needs according to the system operator criteria. The results emphasize that the decisions of the minimum cost investment for a specific element of expansion may not lead to a global minimum cost solution. The radiality constraints impact expansion planning increasing global cost, even though in some cases the investment costs decreased.

Tang, [7], describes a multi-stage mixed-integer nonlinear model, solved by a network-flow programming algorithm with a multi-stage interlacing strategy and a nonlinearity iterating function, to tackle the distribution planning problem. In this model reliability and cost used together in a single objective function, by considering as reliability index the outage cost. In addition to conventional considerations of investment cost in lines or cables and operational costs, the cost of investment in switching elements was also considered in the objective function.

Díaz-Dorado et al., [25], proposed the usage of evolutionary algorithms in the planning of urban distribution grids. Starting from a graph with information of all the possible line paths, their work aims at minimizing investment and losses costs, while ensuring reliability of supply. By starting with a given network configuration the evolutionary algorithm will use crossover and mutation actions on the selected configuration until it identifies the best solution. The objective function of this problem minimizes the sum of the total cost for loops, trenches and HV/MV substations. The final solution must guarantee conductor capacity and voltage drops within admissible values. Additionally, in case of the existence of loops, an open branch will be determined so as to minimize the losses. It is also considered that more than one cable can be installed on the same trench.

Gómez et al., [26], apply an ant colony system algorithm to the problem of planning primary distribution systems. The objective function was formulated as a minimization of the investment costs in lines and substations and the operational costs. The problem is restricted to energy balance at all nodes, capacity limits of new and existing circuits and substations, voltage limits and to a radial network configuration.

Miranda et al., [27], used a genetic algorithm approach to the distribution planning problem, focusing substation and feeder planning. In their work, it was assumed that new installation facilities were known beforehand and that their costs and locations had been estimated. The optimization procedure would then be used to define the best solution for the expansion of the distribution grid in analysis. However, the problem was formulated as a multi-objective problem, minimizing cost and maximizing reliability, which results in the attainment of a set of different solutions. To choose the final distribution plan, the authors suggest the application of multi-criteria decision-aid methodologies.

Ponce de Leão and Matos, [16] and [17], addressed the problem of distribution planning in a three stage process. First, fuzzy modelling is used to capture load and DG uncertainties. Due to the lack of frequency of occurrence and historical data, fuzzy was preferred over probabilistic methods. Second, a multi-objective problem was formulated aiming at the minimization of investment cost, losses and non-supplied energy. It also adds a maximization of the robustness of the solution, which deals with
the fact that some solutions might have some constraint violations, and a minimization of a severity index that is determined to assess the level of violation of constraints (how much have technical limits been surpassed). The third stage results from the need to process multiple solutions returned by the optimization process. Being non-dominated alternatives it is up to the planners to decide which the best option is. Yet, the number of solutions may still be too high and a decision-aid process, based on the planner experience, was suggested to reduce the final set of non-dominated solutions from which the planner will decide. This work only considered a choice over pre-defined reinforcements and required the usage of radial configuration networks.

Rosado and Navarro, [18], propose the usage of a Tabu search algorithm on a fuzzy optimal planning of power distribution systems. The work presented by the authors develops a multi-objective problem that simultaneously minimizes fuzzy economic cost, maximizes fuzzy reliability, minimizes the risk of the system and determines the optimal location and size of normally open feeders that maximize reliability. By introducing fuzzy modelling the uncertainties related to the load at each bus are addressed. Consequently, instead of having deterministic power flows in the grid there will be fuzzy values, which will provide a possibility of violation of the technical constraints. The Tabu search is used to generate possible solutions and avoid being trapped in a local minimum, by creating a list of visited solutions that cannot be revisited for a given number of iterations. In this way the algorithm is free to visit other search areas. The final result is a set of non-dominated solutions that allow planning engineers to choose the best option according to their experience.

Martins and Borges, [20], introduced a model for planning distribution networks where distributed generation integration is considered as an alternative to conventional grid reinforcements. A genetic algorithm was used to implement the developed optimization problem that considers the network operation to be radial, searches for the best location and relocation of normally opened lines, sites and sizes new DG units as part of the network expansion, evaluates rewiring of lines possibility and defines new load connection points. In this case DG is regarded only as a solution to distribution issues that may be installed in a controlled manner to solve local issues. The same approach was made for new load points. However, DG and load and the distribution level are subject to uncertainties and might also be included in the problem formulation as uncertainty variables that may impact grid planning positively or negatively.

Cossi and Mantovani, [28], propose an approach for the integrated planning of MV and LV grids using a Tabu search algorithm to solve the planning problem at the primary network and evolutionary algorithms for the secondary network. As usual the objective functions of each of the problems are to minimize the investment and operational costs. The most innovative part of this work deals with matching the two problems. It uses a heuristic that starts with determining the optimum planning of the primary network assuming a possible interconnection to the secondary network (taken from a list of possibilities). Then, the optimization problem is run for the secondary network. This process is repeated for all the interconnections in the list. In the end the solution with the lower global cost is chosen as the preferred solution.

Another work suggests the usage of an artificial neural network to improve the observability of the distribution network and therefore help planning activities, [21]. The neural network based on the size of the DG units returns the voltage profile on the network.

Ranjan et al., [29], propose a method that incorporates the heuristic conditions associated to distribution planning. So, an iterative process is described, in which first optimum substation location is sought, then optimal feeder path and finally the resulting solution is tested for the restrictions imposed by the heuristic rules. If there is a violation, it will be flagged and the optimization will be run again.

Sepasian et al., [30], proposed an algorithm for substation planning that includes in the objective function a minimization of 3 different costs: the cost of the direct investment in the substation, that of lines connecting the substation to the upstream grid and that of the lines connecting the substation to
the downstream grid. As the downstream grid may also be under development instead of assessing the real cost, an estimate is made with regards to an equivalent distance to the consumption zone. In order to solve this problem, a two stage approach was followed. First, a mathematical clustering algorithm was used to narrow the suitable candidate list for substation locations and, second, a genetic algorithm was applied to solve the optimization.

Asakura et al., [31], describe a long-term expansion methodology where in a first step the normal operation mode is tested for technical violations and then a contingency analysis is performed. If in any of these steps constraints violations are registered new construction plans are elaborated. This process is repeated for every year within the horizon of the planning activity. In the end the feasibility of the construction plan is tested and the optimal solution determined. It is not clear how the two last steps are conducted.

Carrano et al., [19], describe another distribution planning method based on metaheuristics. In this case an immune system inspired algorithm is presented. There is a single objective function defined as the minimization of investment cost and losses. The distinguishing point from similar work is the inclusion of uncertainty in load modelling, by defining two stochastic uncertain variables, future load on nodes and energy tax on nodes.

Varizi et al., [43] and [44], presented a two part work, where in the first place an extensive analysis over existing methodologies for distribution expansion planning was performed and on a second stage a multi-stage mathematical formulation is described to address this problem. The authors argue that mathematical programming algorithms that provide global optimum solutions are more appropriate to deal with the distribution expansion problem than heuristics or metaheuristics. However, they point out that there may be advantages of having risk-based analysis, such as those many times associated to some metaheuristic problems. Substation siting has been addressed with some of these techniques. Leaving choices open for the theoretical optimum location for the substation may sometimes lead to practical unfeasibility of the solution, as there are variables that are not included in the optimization procedures, such as the existing transmission system routes. Attention was called in this work to the fact that reliability may be more than just the cost of non-delivered energy and as such it should be included in the optimization problem as a new objective and not as a part of the cost objective. Finally, the authors point out that there is some distance between the available work and actual practical implementation. This is visible, for instance, in the general lack of concern of existing work with budgetary constraints of system operators. Other issues, such as upgrade possibilities or variable routing and conductor sizing should be further explored.

4.1.2. Distribution System Impacts and Operation

The development of new planning tools cannot be dissociated from the current trends in active management of distribution grids. A lot of effort has been put into the analyses of impacts of DER in distribution grids and the philosophies of control and management that may improve the operation conditions and contribute for the expansion of DER.

So, the most recent advances in power systems focused on the development of technologies and concepts that enable the integration of renewable energies and EVs, use demand response and contribute to DER expansion in general. These developments have been made within the broader context of the smart grids. Within this framework countless proposals were made for new innovative products and services that use intelligent monitoring, control and communications to provide better system operation, reliability and security of supply.

The activities around distribution operation can be roughly divided into the following categories:

- Control concepts: The presence of DER in the distribution network has led to the development of active control mechanisms. Besides additional grid automation, DG, EVs and demand
response are the controllable elements within distribution grids. Advanced concepts, such as those of virtual power plants, microgrids and aggregation have been developed to handle dispersed control elements. DER may contribute for balancing system operation by contributing for voltage control, congestion management or even reserve management, [45], [46], [47] and [48].

- Smart metering: Automated metering infrastructure is an enabler of DER management. Smart metering grants consumers with an active role in system operation, managing DG, EV and responsive loads.
- Data transfer and communications: Significant efforts have been dedicated to the harmonisation of the communication standards, supporting the demand for plug-and-play capability, as well as guaranteeing data exchange. Reliability and security of data are key elements for successful deployment of a proper communication infrastructure. This in turn is crucial for supporting the control concepts and enabling smart metering, [49], [50] and [51].

Active management of electricity grids has gained importance over the last decades as a crucial driver for achieving cost-effective solutions to accommodate DG integration in distribution grids at both the planning and operation stages of the distribution system, stepping away from the conventional fit-and-forget philosophy.

Within the active management paradigm demand response, including EVs, assumes a very important role as big as DG active role in system operation. Nonetheless, it also presents challenges and benefits. The main expected benefits of active load management are, [52]:

- Even though individual loads may become unavailable at any moment, aggregation of loads may guarantee the availability of demand response.
- The delay in the response of loads to the operator requests may be almost inexistent, in opposition to ramping up/down of generators.
- Because loads are distributed throughout the grid, they provide the opportunity to perform local and precise responses to contingencies.
- Using loads to provide system services could even reduce overall GHG emissions, for instance as inefficient ramping of generators is avoided.
- Spatial and temporal flexibility of loads may support an increased penetration of intermittent RES.
- Loads are already present in the power system, and effective communication platforms are becoming available. Soon, the biggest challenge for demand response will be due to the lack of the necessary demand response control strategies.

The control possibilities in the LV and MV grids may be explored with concepts such as those of microgrids, [46], or virtual power plants, [53]. Using these concepts DER may be exploited as new elements for ancillary services provision, in voltage control, congestion management or frequency regulation.

All the advanced management schemes that do not rely on local controllers require adequate communications infrastructure, reliable and capable of delivering messages to and from the elements dispersed in the grid and the central control units in due time to ensure proper control coordination.

The prospects of large EV integration reinforce the need for changing the electricity operation and planning paradigms. Hence, efforts have been made in order to determine the impacts of EV integration in the electric power system. Impact assessment studies have shown that if EVs are regarded as simple loads then there will be severe consequences for the grid technical management. For example, peak load periods may get aggravated leading to possible line congestion and voltage issues that require heavy infrastructure investments for dealing with the relatively small peak period. The initial works presented on the topic, due to the lack of quality data, had to be based in a lot of assumptions. Further studies on the issue have been trying to substitute these assumptions by
existing data or synthetic data that results from complex models that intend to represent real phenomena that influence the impact of EVs. Recent studies show increasingly more consistent results and are starting to provide the tools needed for understanding how the asserted change in the mobility paradigm will affect the electric power systems. Still, future uncertainty regarding EV, DG and DER integration in general is large and for this reason operation and planning activities are under pressure.

With the development of new business models that consider the demand response capability or even the storage capacity of EVs, this uncertainty in the electricity grids will conceivably be mitigated and grid reinforcement needs may be drastically reduced.

Being DER elements small in the power system context, in particular concerning future electricity market integration, aggregation is crucial. For instance, by aggregating several EVs, the unstable behaviour (in a market’s perspective) of a single EV (that may unexpectedly leave or re-enter the power system) is attenuated, facilitating load forecasts, and enabling bidding in markets where the results of the negotiation are binding such as the reserve market or future balancing mechanisms for distribution.

The available work on business models and the aggregating entity tend to present the role of the Aggregator in the market and the business cases and remuneration schemes in great detail. Yet, the operationalization of the management and control schemes that are inherent to the business models and the pathways for aggregators to reach the DER level and to interact with the DSO are seldom forgotten.

Furthermore, in the specific case of EVs the period when they remain connected can be used to perform the so-called smart charging schemes, which take advantage of the fact that EVs are energy dependent loads instead of power dependent loads over the period that EVs are plugged-in.

4.1.2.1 A Review of EV Impact Analyses

Confronted with prospects of large scale deployment of EVs in the electrical power system, the scientific community’s initial reaction is to identify the expected impacts on the power systems. The following issues are presently being addressed:

- Generation adequacy
- Load diagram modifications
- Electricity grids robustness
- Environmental impacts
- Economic impacts
- Regulatory issues

Heydt, [54], Collins and Mader, [55], addressed the impacts of EVs in the load diagrams at regional and national levels within the USA. Heydt considered that EVs would charge between 01h00 and 06h00. For high EV integration a significant temperature rise in distribution transformers would be expected.

Collins and Mader studied scenarios with fixed electricity and time-of-day pricing.

KeKoster et al., [56], assessed the environmental impact of the deployment of EVs, regarding several air pollutants. Schneider et al., [57], studied EV impacts on load diagrams at the distribution substation level. Both rapid charging and smart charging scenarios were tested.

Lopes et al., [58], also investigated the impacts of EVs on distribution grids, introducing a novel concern related to the grid technical restrictions. So, not only load diagrams or transformer loading
were addressed but also bus voltages and branches congestion levels. Three EV charging strategies were considered: uncoordinated charging, time-of-day tariff and smart charging. The number of EVs that could be safely integrated without reinforcement needs was quantified for a residential MV grid.

Uncoordinated charging would allow a 10% EV integration, while time-of-day tariffs and smart charging would enable 14% and 52%, respectively. Another main conclusion was that the first bottleneck will probably occur at the LV grid or at the MV/LV substation. In [59], the same authors evaluated the technical impacts of EV integration in a LV residential grid and extrapolated the results to the Portuguese context. A three-phase unbalanced load flow analysis was used as EV loads are monophasic. Richardson et al., [60], again performed another study on impact analysis of EV integration in LV grids using an unbalanced power flow method.

McCarthy and Wolfs, [61], also addressed the impacts of EVs in the distribution grids, but in this case studying the HV network.

Seeking a more realistic evaluation of the impacts on a LV grid, Clement-Nyns and Driesen, [93], and Clement et al., [62], created a methodology that distributed EVs randomly throughout the network, performing a power flow analysis for each hour of a typical day and repeated iteratively the process (assuming that batteries are fully discharged at the moment of connection to the grid). Two charging methods were evaluated: uncoordinated and coordinated.

In [63], Soares et al. developed a Monte Carlo algorithm that is intended to simulate the impacts of EVs integration in MV distribution networks. In opposition to usual simplifications, a MV grid may represent a wide geographic area that can contain residential, commercial and industrial consumers and so EVs may commute between the different types of buses. A discrete-time stochastic process was implemented to simulate EVs movement throughout a given period. Moreover, The EV battery capacity, slow charging power, energy consumption, initial battery SOC and EV travelled distance are defined according to Gaussian probability density functions. EV owner's charging behaviour is defined by a probability between several options: EV charges at the end of the day, EV charges only when it needs, EV charges whenever possible and EV charges whenever is convenient and the driver has time. The studied network would require new investments if a 50% penetration level would occur.

4.1.3. Distribution Planning Tools

There are distribution planning tools that already partially address some of the new challenges for distribution planning. Table 3 presents a comparison between some of these tools, including those used by the DSOs that participate in PlanGridEV. Each tool was classified against a set of characteristics / features that are important within the new paradigm of distribution systems planning and operation. These were grouped into the following categories:

- Main scope – identifies the type of problem that may be tackled with the tool.
- Application – distinguishes Greenfield planning from incremental planning.
- Scenarios – presents the type of load scenarios tackled by the tool.
- Optimization type – distinguishes the aims of the optimization, cost, reliability or other.
- New features – specific features that need / may need to be addressed to meet the new challenges of distribution planning.
- Planning horizon – only filled in for the tools of the PlanGridEV consortium members.

For most of the tools, the table was filled-in according to existing commercial information and so many times it is not clear what the full potential and limitations of the tool are. The table is then an exercise of normalization of information to be able to compare existing tools amongst themselves and against key characteristics for future systems. Many other characteristics / features could have been discriminated in the table, such as geographical information system handling, graphic user interface or...
other. Yet, those fall out of the scope of the project development and would divert the goal of this comparison, which is intended to be as clear and concise as possible.

In the case of the tools of consortium members some of the features may be achieved in combination with customized tools that work in cooperation with the main software. Most DSOs have also dedicated tools for performing economical analysis over their expansion plans, studying economical viability and prioritizing investments according to its technical and economical interest. Overall, all the software used by the DSOs is able to perform an actual optimization problem aiming at the incremental planning of existing grids. This is in fact the need of most DSOs across Europe since distribution systems are well mature and the driver for reinforcement is load / DG increase or aging of system elements. In the case of SynerGEE Electric, it allows only performing planning studies based on specific grid topologies and thus it does not optimize reinforcement investments. All the optimization tools work in single snapshot load conditions, either for peak load or peak generation with minimal load. All optimize cost, but DPLAN may take reliability into account by considering an associated cost within the cost minimization function. DG is considered in all the tools used by DSOs and reconfiguration modules are presented in DPLAN and SynerGEE. The planning horizon is 10 years for DPLAN and NEPLAN and 40 years for MSPLAN. NEPLAN and MSPLAN are use in combination and are complementary tools.

Table 3 – Comparison of different distribution planning tools

<table>
<thead>
<tr>
<th>Tools</th>
<th>Main Scope</th>
<th>Application</th>
<th>Scenarios</th>
<th>Optimizes</th>
<th>New Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Investment Decisions</td>
<td>Operation Simulation</td>
<td>GreenGrid</td>
<td>Incremental</td>
<td>Single Snapshot (e.g. peak)</td>
</tr>
<tr>
<td>DPLAN/INVESTE</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>NEPLAN/Integral</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>MSPLAN</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>SynerGEE</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>PECO Model</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WindPRO</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Smallworld GIS</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>CYMDIST</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>PSS®SINCAL</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DgSilent</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>LAPER</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

x – existing feature; ? – not clear but seems to be present
1Field used only for the tools of DSOs participating in PlanGridEV
2Used by DSOs participating in PlanGridEV
3Not used in the optimization process but in dedicated studies

Concerning the other available tools, these are mostly not optimizing investment plans, but aid the planning activity by allowing to perform power flow studies with or without reconfiguration capabilities or for performing reliability studies. Mostly, these tools target single snapshots of load / generation conditions. Again, all of them consider DG integration and PECO model and CYMDIST deal with reconfiguration. The PECO model that has been mainly used for consulting activities for regulators seeking tariff revision plans has had the inclusion of EV in the context of the R&D project MERGE. This is the only tool that allows both Greenfield and incremental planning. The single software that describes as a feature handling uncertainty of future expansion is PSS®SINCAL. There are some
tools with particular goals, such as the WindPRO, which is dedicated to the integration of DG and LAPER that optimizes rural electrification.

In brief, most of the available tools handle the conventional planning problem, where single loading scenarios are addressed, typically peak load, but also maximum DG may be studied. The tools seldom tackle reconfiguration of distribution networks since this introduces integer variables into the problem control variables. Regarding the optimization objectives almost only cost is computed and reliability may or may not be monetized into the objective function. However, there are no tools addressing the problem of distribution planning with the inclusion of elements such as DSM, storage or considering ICT. To address these aspects the planning tool must not study isolated snapshots of operation, but it must consider connected periods of time with inter-temporal restrictions. Storage and the activation of demand response (possible by investment on ICT) are cases where these restrictions must be present. Furthermore, there were no tools found using multi-objective problem formulation, which allows exploring in great detail the trade-offs between the different objectives, for instance cost versus reliability. Also, only one of the analysed tools included uncertainty into the formulation of the problem.

Annex A provides further information on some of the available tools for distribution planning.

4.2. Statistical Modelling of DER

Distributed energy systems (DERs) are defined as: “A device that produces electricity, and is connected to the electrical system, either "behind the meter" in the customer's premise, or on the utility's primary distribution system. A Distributed Energy Resource (DER) can utilize a variety of energy inputs including, but not limited to, liquid petroleum fuels, biofuels, natural gas, solar, wind, and geothermal. Electricity storage devices can also be classified as DERs.” [70]

For this report and the focus of PlanGridEV, two main categories of DERs are considered: Distributed generation (DG) in the means of wind and solar power generation and electric vehicles (EV). Distributed generation allows collection of energy from many sources and may give lower environmental impacts and improved security of supply. Units based on non-intermittent and storable energy sources, such as biomass or hydro, can be more easily represented in grid planning, since energy resources can be considered to be available on demand. DER units based on intermittent and not storable energy sources, such as wind and solar, require a more complex model for grid planning, where the energy availability also needs to be represented. Besides using measured data from reference DERs for grid planning, methods for generating artificial data based on statistical methods also exist. In the context of distribution planning it is important to dispose of detailed models for DER prediction and/or for generation of representative synthetic data. These are particularly important for minimizing forecast errors and for being able to derive realistic patterns for DER, yet covering also the extreme situations. The following section covers the common techniques for statistical modelling DERs and, from a future perspective also EVs.

4.2.1. Statistical modelling of wind generation

Due to its stochastic nature, the power generation from wind can vary widely and is limited to places of a high reliability of (relatively constant) wind. By this reason, the operating characteristics of single turbines or wind farms are highly dependent on the local regime of wind. This implies that the conventional power plant stochastic model is inappropriate in the case of wind power.
The most commonly used methodologies for synthetically generating wind data are based on a stochastic approach and include autoregressive moving average models based on Markov chains or wavelet analysis [71].

The output of power of a wind turbine depends on the stochastic nature of and chronological variability of the wind velocity. As shown in Figure 2, the relationship between the available wind speed and the electric power generated from the wind turbine is nonlinear.

![Figure 2 - Typical wind power output with steady wind speed [72]](image)

### 4.2.1.1 Normal and Weibull distribution

Normal and Weibull distributions are probability distribution functions and commonly used for analysing wind speed data. They are probability distribution functions and consist of independent identically distributed random numbers. The normal distribution is defined by:

$$f(x) = \frac{1}{\alpha \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

**Equation 1 - Normal distribution**

Wind data are generated by using a sequence of independent random numbers from the normal distribution. The Weibull distribution is given by:

$$f(w) = \frac{\alpha}{\beta} w^{\alpha-1} \exp\left(-\frac{w}{\beta}\right) \quad w \geq 0; \quad \alpha, \beta > 0$$

**Equation 2 - Weibull distribution**

$\alpha$ and $\beta$ are scale and shape parameters and can be determined by using a graphical method or the method of moments.
4.2.1.2 Autoregressive models

Hourly mean wind speed time series is of high dependence and requires a wind speed data generation model incorporating the dependence structure of the observation, which is not covered by normal and Weibull distributed random numbers. The autoregressive (AR) models include the correlation structure of the observation and hence generate dependent series. The first order autoregressive model covers only the effect of previous values in the series in which the observed sequence of wind speed data is used to fit a model in form of:

\[ W_i = \sum_{j=1}^{m} a_j W_{i-j} + \varepsilon_i \]

Equation 3 - First-order autoregressive model

Setting \( m=1 \), the equation above becomes:

\[ y_i = r_1 y_{i-1} + \varepsilon_i \]

Equation 4 - Markov model (1st order)

This case is also called the Markov model. The first order AR model requires only a random number of a normal distribution to be generated.

For a better preservation of the dependencies of the observations, the order of the AR model has to be increased. Especially in data sets where dependencies are very obvious, an AR model of second order has to be formalized, which can be described as:

\[ y_i = \phi_1 y_{i-1} + \phi_2 y_{i-2} + \varepsilon_i \]

Equation 5 - Second-order autoregressive model

The equation consists of the autocorrelation coefficients (\( \phi_1 \) and \( \phi_2 \)) and the random component \( r_2 \). The disadvantage of AR models lies in the need for complex techniques for determining the numerous model parameters [73].

4.2.1.3 Wavelet analysis

The wavelet approach is a real or complex-value continuous function with zero mean and finite variance and fully described in [71]. Some examples of wavelets are Morlet, Mexican hat, Shannon and Meyer. The base of the wavelet transformation is decomposing a signal (observed data) and then reconstructing it. An example (4 steps) of the Haar wavelet [74] is illustrated in Figure 3.
4.2.1.4 Markov chain

For the stochastic modelling of wind data, Markov chains from first- and second order are currently widely used [71], [73]. The Markov chain of the first order is one for which each next state depends only on immediately preceding one [75]. Markov chains of second or higher order are the processes in which the next state depends on two or more preceding ones.

The Markov chain is a method that first determines the state of the wind speed and then generates its magnitude by using a preselected distribution. It is a random process usually characterized as memory-less, which means: the next state depends only on the current state and not on the sequence of events that preceded it [76].

Figure shows a schematic diagram of the behaviour of the wind represented by a Markov chain (The wind states are represented in a growing speed order and the transition from state j−1 to state j is quantified by the transition rate λj−1;j).
In [73], the first-order Markov chain (also known as one-step) the state of wind speed in the current moment can only be defined upon the previous state. Figure 5 shows an example of results from a first-order Markov chain compared to the corresponding data set.

For statistical modelling wind turbines within a second characteristic (e.g. operational state and failure state) second-order Markov chain model can be used [77].

### 4.2.2. Statistical modelling of solar generation

Solar forecasting methods, in terms of renewable energy integration, are existing in a broad variety of techniques and approaches. Besides the statistical models, which are discussed in this report, there is also cloud imaginary and satellite based, numerical weather prediction and hybrid models existing [78]. A detailed description of all common models (including also non-statistical approaches) can be found in [79].
In [78] different methods for solar forecasting were classified within temporal and spatial resolution. Figure 6 shows the different fields of application of statistical, numeric weather prediction and satellite-based models. Further on, this report will only cover state-of-the-art methods from the statistical area.

Statistical methods of forecasting are based on historical data of solar irradiance and can be split in mainly two categories:
- Linear models or time series models
- Non-linear models

The output of power from a PV power plant is a function of the location, the time (incl. date), the area and orientation of the panels, the PV technology and the current atmospheric condition. Most of the methods shown in this report require the knowledge of clear sky conditions (characterized by the absence of clouds). Figure 7 shows a comparison of measured power output data and the expected output under clear sky conditions.
Figure 7 - Measured output (left) and power output under clear sky conditions (right) as a function of the time of the day and the day of the year [80]

As mentioned, clear sky models are a premise for a broad variety of methods. However, the focus of this report is on methods only. A comprehensive comparison of eight clear sky models against 16 independent data banks was published in [81].

### 4.2.2.1 Persistence Forecast

This method represents the simplest model. It is mostly used for validation purposes, to check whether the chosen forecast model provides better results than any trivial reference model. This method supposes that global irradiance at time \( t+1 \) is best predicted by its value at time \( t \). As shown in Figure 6, the field of application of persistence forecast is within intra-hour. For more than one hour ahead forecasting it is a very inaccurate method and therefore inoperative.

### 4.2.2.2 Regressive methods (linear Models)

**Mixed autoregressive moving average models (ARMA)**

The ARMA model is based on two elementary models: the moving average (MA) model and the autoregressive (AR) model and can be defined as:

\[
S(t) = \sum_{i=1}^{p} \alpha_i S(t-i) + \sum_{j=1}^{q} \beta_j e(t-j)
\]

Equation 6 - ARMA model

The popularity of the ARMA model is its ability to extract useful statistical properties and the adoption of the well-known Box-Jenkins methodology. ARMA models are very flexible since they can represent several different types of time series by using different order. It has been proved to be competent in prediction when there is an underlying linear correlation structure lying in the time series. One major requirement for ARMA model is that the time series must be stationary [78].

**Mixed autoregressive moving average models with exogenous variables (ARMAX)**

This is an extension of ARMA models and tries to improve the accuracy of the ARMA model by including information external to the time-series under analysis. This information could be for example: evolution of local temperature, relative humidity, cloud cover, wind speed, etc. These variables are independent from the model, but affect the output in the real world.

**Autoregressive integrated moving average model (ARIMA)**
An extension of the ARMA models, the Auto-Regressive Integrated Moving Average (ARIMA) time series models form a general class of linear models that are widely used in modelling and forecasting time series in [78]

**Mixed autoregressive moving average models with exogenous variables (ARIMAX)**

This method is similar to the ARMAX model. The previous values of exogenous time-series can also be included into the ARIMA model.

**4.2.2.3 Artificial intelligence techniques (non-linear models)**

Artificial intelligence (AI) techniques are used for several years for solar forecasting. A detailed overview of applications of AI methods for forecasting and modelling of solar irradiance is published in [79] and [82]. For this report, two of the most common AI techniques were chosen.

![Simplified diagram of biological and artificial neuron](image)

Figure 8 - Simplified diagram of biological and artificial neuron [79]

Figure 8 (a) shows a simplified diagram of a biological neuron. Antenna-like structures which extend from the cell body, or soma, allow the neuron to communicate with other cells. The structures which allow the neuron to accept input signals are called dendrites. The structure which carries signals away from the neuron is called an axon. A neuron may possess numerous dendrites but it never has more than one axon. (b) Artificial neuron with inputs $x_1$, $x_2$ and $x_3$ weighted by $w_1$, $w_2$ and $w_3$. The neuron has an embedded net function $\beta + \sum_{i=1}^{3} w_i x_i$ and transfer function $f(.)$ which are used to calculate output $z$.

**Artificial neural network (ANN)**

The most popular form of neural network is the multilayer perception (MLP) structure [78]. The MLP structure consists of an input layer, one or several hidden layers and an output layer. The input layer gathers the model's inputs vector $x$ while the output layer yields the model's output vector $y$. Figure 9 represents a one hidden layer MLP. The hidden layer is characterized by several non-linear units (or neurons). The non-linear function (also called activation function) is usually the tangent hyperbolic function $f(x)$. 

![Diagram of MLP structure](image)
\[ f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \]

Equation 7 – Tangent hyperbolic function

Neural networks with inputs (d), hidden neurons (h) and a single linear output unit defines a non-linear parameterized mapping from an input (x) to an output (y) given by the relationship:

\[
y = y(x, w) = \sum_{j=0}^{h} w_j f\left( \sum_{i=0}^{d} w_{ji} x_i \right)
\]

Equation 8 - Neural network relationship

![Figure 9 - MLP with inputs and hidden units](78)

Wavelet neural networks

Wavelet neural networks (WNN) are a combination of two techniques and inherit the advantages of neural network and wavelet transformation. The wavelet function is used as activation function (instead of the tangent hyperbolic function). As shown in the network topology in Figure 10 the WNN model can have several output layers (y) compared to the ANN model which has only one.
4.2.2.4 Comparison of ANN and classical time series models

“A comparison of ANN and classical time series models has been carried out in Reikard [83] and in Sfetsos and Coonick [84]. Both studies find that the error of a simple regression model can be reduced considerably by a factor in the range of 0.6–0.8 when using advanced models. Reikard [83] compares a regression model, the UCM Model, an ARIMA, a transfer function model, a neural network model and hybrid model.” [78].

In [78] the authors use a logarithmic scaling of the input. Results show that for the resolutions of 60, 30 and 15 min, the ARIMA model shows better results. In Sfetsos and Coonick [84], a feed-forward ANN is identified as the most appropriate. The analysis is in Reikard [83] for several stations with different climatic conditions also shows that there is a strong influence of the climatic conditions on both forecast accuracy and potential for improvement by the use of advanced models.
4.2.3. Statistical Modeling of EVs

4.2.3.1 Overview

Based on the need for aggregation of EV charging processes (e.g. areas, grid branches, fleets etc.) extensive research has been done to describe the charging profile of electric vehicles using stochastic models. The motivation, however, is manifold, ranging from forecasting the day-ahead or intraday demand to enabling smart charging. In comparison to the statistical modelling of DGs, the modelling of EVs can take the geographical movement of the vehicles also into account. Whilst DGs are always on the same geographical area, EVs are moving in stochastic patterns within a day. In general, two main approaches can be distinguished: Modelling the behaviour of vehicles (in this case EVs) based only on the temporal component of the appearance of the vehicles on a specific aggregation level, or the consideration of the whole traffic system of an area, which means keeping track of individual vehicles within one day.

4.2.3.2 EV mobility data and EV specifications

Mobility data from EVs can be available in different formats from different sources. This could be drivers logs, GPS coordinates and time-lines or, for the coverage of large numbers of vehicles in large geographical areas, origin/destination matrices (O/D matrix). O/D matrices are mainly generated from statistical surveys, are hourly based and contain information about the number of vehicles going from a specific origin in a defined hour to a known destination. Currently not all approaches are taking the geographical movement into account. Data on hourly basis are, for current power grid simulations, mostly from too low granularity. Data with granularity of 1, 15 or 30 minutes are preferred, therefore the original data set has to be adapted.

All current models use real mobility data for parameterization. Soares et al. use Portuguese mobility data. Bureaud de Castro et al. use Austrian mobility data [85] with various granularities of 15 to 60 minutes. Wu et al. and Sortomme et al. use (different) US travel survey data and Kristoffersen et al. use Danish survey data. The models are of course not based on real EV usage but on combustion engine vehicles leading to possible differences to real EV usage patterns. Sundström et al. use partly synthetic mobility data and actual vehicle data.

Besides the mobility behaviour, additional information for the representation of EVs in grid planning is also of importance.

- Driver charging behaviour:

  Soares et al. [63] define the EV customer behaviour via Monte Carlo method based on questionnaire results gathered in the FP7 project MERGE.

<table>
<thead>
<tr>
<th>EV charge at the end of the day</th>
<th>33 %</th>
</tr>
</thead>
</table>
D1.4 State-of-the-art methods report

| EV charge only when it needs | 23 % |
| EV charge whenever possible  | 20 % |
| EV charge whenever it is convenient and the driver has time | 24 % |

- **Temperature:**

  Baptista et al. [86] use winter vs. summer scenarios to describe changes in overall energy demand. Changes in EV demand are not considered there. Burnier de Castro et al. [87] use variations of household demand based on energy demand measurements in Upper Austria as well as changes of EV energy demand based on vehicle studies in climate chambers [88].

- **Fleet composition**

  Most groups define a fleet composition based on EVs available at today’s market. Exemplary for this is the fleet composition of Burnier de Castro et al. [87] is shown below. Some groups consider combustion engine cars as part of the fleet share, others consider the EV market penetration via the total amount of vehicles in the (EV) fleet.

  Table 4 - Example of the composition of an EV fleet in a low-voltage distribution grid [87]

<table>
<thead>
<tr>
<th>Vehicle class</th>
<th>Share</th>
<th>Battery capacity</th>
<th>Phases</th>
<th>Max. Aperage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premium-class</td>
<td>10 %</td>
<td>50 kWh</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>Medium-class</td>
<td>10 %</td>
<td>10 kWh</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>10 %</td>
<td>24 kWh</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Compact &amp; Minivans</td>
<td>30 %</td>
<td>26 kWh</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>30 %</td>
<td>22 kWh</td>
<td>3</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>10 %</td>
<td>16 kWh</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>100 %</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Grid connection and charging infrastructure**

  The grid can be described as single wire or 4-wire system. The advantage of 4-wire systems is the possibility to investigate the influence of asymmetric loads on the individual voltage levels. This is of special interest when considering EVs which charge on single phases (e.g. BMW i3). Basic characteristics of charging infrastructure have to be defined, as they can affect the impact of charging behaviour to the grid.

4.2.3.3 **Standard statistical models**

**Markov chains**

Typical modelling of EV charging processes is done using Markov chains. Here several discrete states of the EV (parking in public, parking in private, driving etc. – using slightly different nomenclature) are defined and connected via time depended transition probabilities. Figure 11 shows a typical setup for the Markovian process as presented by Soares et al. [63]. Kilian [89] added an additional fast charging state in order to cover an additional energy source.
Based on the vehicle data the model can be parameterized leading to probabilities for the occupation of each state. Each individual EV will then be simulated using a flow chart and Monte Carlo simulation for the transition between the states. Finally, the aggregated data will then, of course, satisfy the input data. A similar approach is also used by Vayá et al.

One drawback of stochastic modelling in this way is that it is mainly suitable for a large number of
vehicles. Kilian [89] shows that forecast accuracy is, of course, decreasing with decreasing number of vehicles and therefore not suitable for a small number of EVs in a small grid branch of an LV distribution grid.

Figure 13 - Mean error of EV energy demand forecast [89]

Similar results can be found with Mathieu et al. [90]

Figure 14 - Uncertainty in energy bounds (10000 samples) [90]
Semi-Markov Chains

Sundström et al. go further than using Markov chains but extend the model to Semi-Markov chains which seem to provide a better forecasting of a single EV owner behaviour. One further benefit for low numbers of EVs should be that the vehicle would not be able to oscillate between states (e.g. Parking and driving) as the state change probability is also a function of the time within this state which seems to be more realistic for describing EV user behaviour (minimum driving/charging time).

4.2.3.4 Agent-based Approach

Different approaches exist for agent-based electric vehicle simulation. Whilst the approach using MATSim (or similar transport simulation tools) in combination with EVSim in [87], [91], [92] allows controlling the charging process without interfering with the day plan of the agent, the method in [93] takes re-planning of the agents day plan due to the charging process into account (e.g. delay of departure to provide more time for fully charging the car). The following section about multi-agent simulations of EVs focuses on the approach of load shifting and not on rerouting or interfering with the mobility behaviour of Agents.

Multi-Agent transportation simulation (MATSim)

"MATSim provides a toolbox to implement large-scale agent-based transport simulations. The toolbox consists of several modules which can be combined or used stand-alone. Modules can be replaced by own implementations to test single aspects of your own work. Currently, MATSim offers a toolbox for demand-modeling, agent-based mobility-simulation (traffic flow simulation), re-planning, a controller to iteratively run simulations as well as methods to analyze the output generated by the modules." [94]

This is an agent based MATSim [94] approach to simulate driving patterns [95], [91]. This approach can either be used as a direct link to grid simulation via EVSim [91] or as input for parameterization of Markov Chain probabilities [95].

For simulating with MATSim, it is necessary to create, beside the mobility-agents, a street-network from the area which should be simulated. This network can be prepared with the tool OSMOSIS [96] in combination with OSM data [97]. Osmosis is a command line Java application for processing OSM data. The tool consists of a series of pluggable components that can be chained together to perform a larger operation. OSMOSIS works without graphical user interface, so after starting the package it is controlled via command line input. Currently a graphical user interface is developed in the promising project OSMembrane [98]. One method for preparing a network is to extract bounding boxes. Therefore OpenStreetMap must be used first, selecting the Export-menu and choosing the area of interest as shown in the figure below.
After selecting the network area, coordinates are found on the left hand side of the map. These coordinates are then used in the command line to create a network file:

During the process, it can be defined which level of detail and which types of roads are part of the extracted street network. Figure 16 shows a finalized street network for MATSim and contains information about street type (number of lanes, driving direction, ...) and maximum free speed.
EVSim

The tool EVSim (Electric Vehicle Simulation) was developed to allow the usage of the output of traffic simulations such as MATSim for power simulation. EVSim takes the detailed agent plans from MATSim output as input. EVSim generates for every single plan an independent Agent, which follows further on its plan (departure and arrival times, destinations …) throughout the simulation. EVSim also provides the option to add additional information to all Agents, for example: EV-Type, Battery and charger characteristics, start-SOC, etc. Locations for charging and number and specifications of charging infrastructure are also defined in EVSim. Figure 17 shows the scheme of the co-simulation environment. In a first step, the traffic simulation is done with MATSim, in the second step, the dynamic co-simulation between EVSim and the power simulation takes place. Relevant variables of EVs and the charging infrastructure are communicated via an OPC connection.

Figure 17 - Scheme of Co-simulation environment with MATSim, EVSim and PowerFactory via OPC connection [87]

This co-simulation environment is capable of simulating areas from a few single EVs up to thousands of vehicles.
4.2.4. Applicability of statistical models for PlanGridEV objectives

Due to the very different nature of solar and wind power, different approaches and methods can be found in literature. For the generation of wind speed data, traditional methods, such as, normal and Weibull distribution function, autoregressive methods and the Markov chain, are parametric methods and rely on time-series data. Depending on the availability of local data, different methods for forecasting and modelling solar power are to be favoured. In data-poor environments, where only historical data for the pre-analysis is available, endogenous stochastic methods can be used such as AR, MA, ARMA and ARIMA. In areas where more data is available, exogenous inputs can be included for stochastic learning methods. Artificial neural network modelling provides an alternative approach to physical modelling, when enough historical data is available. This technique works in data-rich and data-poor areas and is used in a wide variety of time horizons (from intra-hour to yearly averages).

Statistical methods for modelling EVs rely on mobility data for parameterization. However, these methods aren’t that dependent on local data as the agent-based approach and can therefore more easily adapted for each area of investigation. For a high level of aggregation with a high number of vehicles, the stochastic approach seems to provide results as good as the agent-based approach. For a low number of EVs in a LV grid new approaches need to be taken in order to properly assess the grid quality. One possible parameter might be a maximum probability of charging process interference for a large sample of simulations using a low number of EVs. In terms of applicability for PlanGridEV, statistical methods provide better options for implementation than agent-based approaches, which rely on different independent tools.

Currently, for EV simulations both approaches, agent-based and stochastic methods, are used. The Agent base approach provides detailed mobility behaviour with geographical information based on real mobility data. Further on it includes a whole traffic simulation which grants realistic travel times and distances, and offers the possibility of further investigations beyond the effects to the power grid. Therefore the preparation for this method is very time consuming and has to be done for each area which is investigated. This also implies the availability of the essential data (street network, mobility data …). Assumed that appropriate data is available; this method also works for very low numbers of EVs. In terms of computing capacity, the agent based approach extensive resources are needed. Based on current evaluation a Semi-Markov approach or a Markov approach with boundary conditions after state change seems to be the most promising method for low numbers of EVs.

4.3. Modelling Storage Capabilities of EV

In this section, an overview of existing models used to represent the charging flexibility of EV fleets is given. These models will serve as a starting point in the storage modelling in WP4, where appropriate modelling approaches will be developed for different scenarios. The main challenge is to adapt the existing models, which for the most part have not been envisaged for use in distribution network optimization, to the requirements of the distribution planning tool to be developed in WP4. The reviewed models of EV aggregations can be mainly classified into two types:

- Virtual storage-type models: models based on the aggregation of energy and power constraints over time
- Task-type models: models where charging is modelled as a task with a task size, arrival time and deadline
4.3.1. Virtual storage-type models

Reference [99] describes an aggregator that manages the electricity market participation of a vehicle fleet by optimizing the charging and discharging patterns. To do so, vehicles are considered to be always plugged in when they are not driving. Although the aggregator is responsible for the charging and discharging of vehicles, it cannot control the driving patterns of the vehicle fleet.

The vehicle fleet is modelled as a set of aggregate vehicles, each of them representing a number of vehicles with similar driving patterns. These driving patterns are constructed by clustering historical data. First, different variables are extracted from the historical data, namely, vehicle user, hour and day of departure, and driving distance. From this, a 24-h driving pattern for each vehicle user is constructed. Then, the vehicle fleet is partitioned into different types of vehicles (aggregate vehicles) by grouping driving patterns and following the procedure below:

1. Vehicles are first grouped according to day of departure (week-day or holiday/weekend-day) and vehicle technology
2. Driving patterns within each group in step 1 above are clustered using the k-means clustering algorithm. Each cluster is represented by a single driving pattern and weighted according to the number of vehicles within the cluster

The charging and discharging of different aggregated vehicles is decided based on their driving patterns and also based on electricity spot prices. If the total electricity load of the vehicle fleet is sufficiently small, the aggregator is considered as a price-taker, while if the aggregator has a significant market share, its impact on prices is modelled.

Reference [100] proposes a stochastic optimization approach to determine the optimal bidding strategy of an EV aggregator that participates in both the day-ahead and the regulation markets. The main sources of uncertainty affecting the bidding strategy are represented through a set of scenarios based on past observations. The uncertainties taken into account are the uncertainty in the aggregate characteristics of the EV fleet and the amount of regulation energy that will actually be requested by the system operator.

The aggregator manages a fleet of EVs and is responsible for modulating the EV charging level of each individual EV once it is plugged in. Unidirectional interaction with the grid is adopted, i.e., EVs cannot discharge energy back to the grid. However, they can deviate from their preferred operating point by changing their charging rate.

The aggregator participates in the day-ahead and regulation markets with the aim of minimizing the cost of the energy purchased and maximizing the revenue from the regulation market participation. To do so, it represents the charging flexibility of individual EVs with an aggregate model and participates in these markets as a price-taker. The details on the aggregation model are not provided in the paper. Similarly to reference [100], reference [101] proposes an optimization approach to decide on the bidding strategy of an EV aggregation agent in the electricity market. This agent participates with bids for purchasing electric energy and selling secondary reserve.

The aggregator is modelled as an electricity retailer for electric mobility that defines tariffs for its customers. The aggregator buys electric energy for EV charging and offers ancillary services in the electricity market. To do so, two different groups of clients are considered, namely, (i) clients that allow the aggregator to control the charging process during a period defined in their contract, and (ii) clients that do not allow the aggregator to control the charging process, being the aggregator just an electricity provider.

The aggregator controls the charging of type-(i) clients, which provide their state of charge needs for the next day when they arrive at home. However, since the bidding for the day-ahead market, where
the aggregator acts as a price-taker, is made before this information is available, the aggregator needs to forecast the charging requirements for the next day. Since it is not suitable to deal with each EV individually, the aggregator forecasts the charging requirements and maximum available power of the whole fleet. Particularly, the aggregator forecasts the hourly total EV energy consumption of type-(ii) clients, the energy flexibility of type-(i) clients, and the hourly spot and reserve prices. These forecasts, which are made at a given time step one day in advance, are used to define the hourly bids in both the day-ahead and the ancillary services markets. The aggregated EV flexibility is expressed as hourly constraints on the total cumulative energy purchased over time.

This paper provides also a generation mechanism for synthetic EV charging time series. To do so, EVs are initially characterized in terms of battery capacity, charging power, energy consumption, and battery state of charge. These values are defined according to truncated Gaussian probability density functions. Then, the EVs movements in a one-year period are simulated using a discrete-state, discrete-time Markov chain to define the states of all EVs at each time step of 30 min. These two steps provide the periods during which EVs are plugged in and available to charge, the network buses to which EVs are plugged in, the EV powers absorbed at each 30 min interval, the EV battery state of charge evolutions, and the EV travelled distances. Using these data, EV charging time series are generated.

Reference [102] proposes an approach for the demand side management of PHEVs following three steps: aggregation of individual PHEV charging constraints, optimization of a collective charging plan to minimize the costs for electricity supply, and real-time control to match the scheduled charging plan. First, in order to aggregate vehicles, charging constraints are classified into two types: energy and power constraints. Both power and energy constraints of individual PHEVs are aggregated to represent the collective energy and power constraints of the PHEV fleet, in a similar way as in [101].

Second, a collective charging plan for the PHEV fleet that minimizes the cost for the energy supplier is determined. To do so, aggregated energy and power constraints, as well as the PHEV charging requirements are considered.

Finally, the aggregated charging power is assigned to individual PHEVs with a heuristic defining charging urgency.

Reference [103] proposes an approach for the optimal charging of EVs including electricity grid constraints (voltage and power). The method provides an individual charging plan for each vehicle that avoids distribution grid congestion and satisfies the requirements of vehicles’ owners.

The aggregator is seen as a charging service provider (CSP) that decides the individual charging schedules of each EV. This CSP is separated from other players such as the retailer and the DSO. The CSP acts an intermediary between the EVs and the DSO. The CSP collects historical data and energy requirements of EVs, as well as information of the grid from the DSO. Then, the CSP decides the individual charging schedules for each EV so that the EV utilization is maximized, the cost of charging is maximized, and distribution grid congestion is prevented.

The batteries of EVs are simulated using a nonlinear state-dependent model. Different types of EVs are considered, namely, commuter cars, family cars, and taxis, each one with different characteristics in terms of both vehicle parameters and driving behaviour.

In order to represent the aggregated flexibility of the fleet of EVs, the usage patterns of each EV have to be considered. First, the maximum and minimum charging power consumptions of individual EVs are computed. Then, these maximum and minimum power levels are used to compute the maximum and minimum energy levels of individual EVs. Finally, the charging flexibility of the EV fleet (defined as the maximum and minimum energy and power levels over time) is determined by aggregating the flexibility of each individual EV. This is similar to the approaches in [101] and [102].
Once the optimal charging schedules are determined, these are sent to the aggregator to check their feasibility. If violations are detected, additional constraints are included into the model that limit the total power consumed by a group of vehicles. This is done iteratively until all violations are solved.

Another model based on the aggregation of power and energy constraints is presented in [95]. As in the models in [101], [102], [103], the available flexibility is defined as a set of time-varying aggregated energy and power constraints and a dynamic equation. The main difference to the previous models is that the aggregation is location-dependent: at each network node, a virtual storage is modelled. To be able to do so, the energy contributions of arriving and departing vehicles are taken into account in the dynamic equation. Moreover, in [95] the set of constraints is modelled as probabilistic constraints given EV driving pattern uncertainty. The model is included in a probabilistic optimal power flow which minimizes generation costs, given load and generator constraints, network constraints and the available (uncertain) charging flexibility.

Reference [104] proposes an algorithm for calculating load shift potentials and optimal charging of EVs.

The main requirement of customers is usually a full battery at a given time. However, there is generally a large number of load charging curves that meet this requirement as well as other technical restrictions. Under this context, this paper identifies the load shift potentials, which are defined as the range of all charging curves that meet the customers' requirements and respect all individual charging and discharging constraints over time. Although there are multiple charging curves that meet these requirements, not all these charging functions are economically optimal for consumers, and thus, the electricity supplier needs to pay for the customers' permissions to choose a different load curve.

Results reported in this paper show that the load shift potential of EVs is significant and could contribute to a better operation of the future grid. However, the willingness of individual users to participate in this will depend on the incentives they get, which requires further research.

4.3.2. Task-type models

In [105] a clustered representation of the flexible demand of a fleet of EVs as a queue of tasks is proposed. In this work, charging is controlled centrally by an aggregator, who can therefore optimize the timing of demand and offer cheaper tariffs. The rate at which vehicles charge is fixed, but charging can be initiated and interrupted by the aggregator. Charging is modelled as a number of serially executed subtasks, where tasks that are similar are clustered for an aggregated representation of the demand of the fleet. The EVs move from one cluster to the next cluster by charging, until they reach a full charge. The decision variable in this model is the number of EVs that are allowed to move from one cluster to the other, i.e., to charge. The goal of the aggregator is to buy enough power in the market to serve the EVs and, at the same time, to follow a 5-minute real time signal from the Independent System Operator.

A similar approach is found in [106], where the EV charging requests are classified into a reduced number of load types. As in the previous paper, it is assumed that the aggregator can interrupt charging, but not affect the charging rate. This approach models in more detail than the previous one the EV load, which can also include charge profiles with varying power (like the constant-current/constant-voltage charging profile), and not only constant power profiles. The non-constant power loads are approximately characterized by multiple stages of different power demands, which are dependent on the state of charge. Load types are characterized by four variables: the earliest charging start time, the latest charging completion time, the required total charging time and the power demand at each charging stage. The aggregator needs to determine, at each stage, how many requests of a certain load type should be served. This can be formulated as an integer program, and
further approximated to a linear program for large fleets. Since this problem could be computationally demanding, the authors propose a two-step procedure to solve it. In the first step, the total number of EVs of each load type to be served is determined, while in the second step the power allocated to each load type is distributed among the charging stages associated to that task. The first step is formulated as an integer linear program in the general case, while the second step is based on the earliest-stage first policy. The method is applied to a load peak minimization problem.

In [107] a model to represent aggregated demand dynamics is proposed, both for thermostatically controlled loads and for PHEVs. Individual load dynamics and aggregated dynamics are based on hybrid dynamical system load models. PHEV charging is modelled as a task, characterized by arrival time, size of the task and the deadline to be completed. The control signal envisaged would activate PHEVs with task laxity (slack time) equal or lower to the value of the signal. The hybrid system which characterizes charging has three modes: waiting, charging and complete. The aggregation of loads is based on methods from fluids dynamics, using density functions for the discrete modes of the hybrid motion. In the general case of non-homogeneous loads, clustering can be used to define sets of homogenous aggregate models. In the case of PHEVs, it is necessary to model the arrival of PHEVs as external flows entering the aggregate model. The model is proposed as an analysis tool and, potentially, for the aggregate controller design. However, the model is not used for the purpose of optimization.

The same hybrid model for individual charging described in [107] is used in [108], and it is extended to a stochastic version. A two-step approach is proposed, where the first step represents the energy planning decision (how much energy should be consumed at each time step of a given time horizon) and the second step is the real time charging control. In the second step, a power tracking policy needs to be defined, that allocates the power budget to the vehicles as precisely as possible, given vehicle constraints. To determine the optimal charging profile in the first step, the maximum and minimum consumable energy at each time step for a close-loop trajectory with a predefined policy is estimated. The energy bounds are represented as a function of the previous power schedules. The energy-planning problem can be solved with dynamic programming. The approach is tested in a setup where the charging costs given exogenous prices are minimized.

4.3.3. Overview and conclusions

Finally, Table 5 gives an overview of the main characteristics of the reviewed papers. It distinguishes between approaches where charging is only unidirectional or bidirectional (vehicle to grid), and charging can only be on/off controlled or also modulated. Moreover, some of the approaches are deterministic, while others explicitly address and model vehicle driving uncertainty (stochastic). While most of the models do not distinguish among network location (single area), [95] takes into account vehicles moving from one network node to others (multi-area). Finally, the aggregation models were mostly used in economic optimization models, such as optimal bidding problems or cost minimization given exogenous prices, while a few also considered network constraints.

In the models to be developed in PlanGridEV, it is crucial to include network constraints at the distribution grid level. With the exception of [103] and [95], none of the models explicitly addresses network constraints, and only [103] focuses on the distribution network. In [103], possible violations of distribution network constraints are iteratively detected and included into the overall optimization. In principle, the model proposed in [95] for the HV network, could also be used to model aggregations of vehicles at the MV network. In this case, the idea is to have location-dependent aggregations to be able to formulate all network constraints explicitly. Each network node of the MV network would have its corresponding virtual storage, while EV flows from node to node will be modelled as energy transfers from one virtual storage to another.
In conclusion, whereas the state-of-the-art provides a good starting point, existing models will need to be adapted to the specific requirement of the planning tool to be developed in PlanGridEV.

<table>
<thead>
<tr>
<th>Reference</th>
<th>[99]</th>
<th>[100]</th>
<th>[101]</th>
<th>[102]</th>
<th>[103]</th>
<th>[95]</th>
<th>[104]</th>
<th>[105]</th>
<th>[106]</th>
<th>[107]</th>
<th>[108]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unidirectional charging</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Bidirectional charging</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>On/off control</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Modulation</td>
<td>(x)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Stochastic</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Deterministic</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Single area</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Multi-area</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Economic optimization</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Network constraints</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 5: Main characteristics of reviewed methods.

4.4. Multi-objective Optimization Approaches in Distribution Planning

Optimization techniques should be employed in the design of electrical distribution networks, leading to a better allocation of limited financial resources. Network planning optimization should take into account both design and operational aspects of distribution networks: new networks and expansion of existing networks, installation and maintenance of facilities and conductors, energy losses and energy supply interruptions.

Most practical optimization problems in engineering address multiple objectives, but few really need to be formulated this way. Suppose that one is searching for a spanning tree solution to serve as a transportation network and that we want to minimize the sum of its arc costs (infrastructure cost), as well as to minimize the transportation effort (operation effort). Hence, there are two objectives. However, the problem can be formulated as single objective if the transportation effort can be mapped into a cost, which is possible in most cases.

In short, a multi-objective problem of the form

$$\min \{f_1, f_2, f_3, ..., f_n\}$$

can be solved as a single objective problem (weighted sum) of the form
\[ \min \sum_{i=1}^{n} k_i f_i \]

with \( f_i \) being the individual objective function and \( k_i \) the corresponding weight.

Yet, some objectives are not easy to map into a monetary base. In these cases, the problem is really multi-objective and it must be solved with an appropriate approach. The adequate analysis of an optimization problem with more than one objective should be undertaken with multi-objective optimization approaches, as these approaches can lead to meaningful results that would not be reached with single objective methods [109].

Continuing with the previous example, suppose now that the second objective of the network planning problem is the risk of transportation incidents instead of the transportation effort. Security is usually not easy to map into a monetary scale, so one will be unable to decide between «low cost and high risk» and «high cost and low risk». Given the options \((f_1=5 \text{ and } f_2=1)\), \((f_1=4 \text{ and } f_2=2)\) and \((f_1=6 \text{ and } f_2=2)\), one cannot decide between the first and second options, whereas the first option is clearly preferable to the third option, as it is better than the latter in both objectives (lower cost and lower risk). Hence, it is said that \((f_1=5 \text{ and } f_2=1)\) dominates \((f_1=6 \text{ and } f_2=2)\). It is this idea of dominance that defines the Pareto surface in multi-objective optimization. The Pareto surface separates dominated solutions from non-dominated solutions. There are several ways to determine non-dominated solutions.

An obvious way is to sample the weights \( k_i \) and solve a weighted-sum single objective problem for each sample of weights. One would come up with a set of points in the \( \mathbb{R}^n \) domain that would delimit the Pareto \( n \)-dimensional surface. Such process is simple to program although not very efficient.

Genetic Algorithms (GAs) have special features that allow more efficient Pareto optimization. In recent years, it has been recognized that GAs are particularly well suited for multi-objective optimization problems since they can simultaneously evolve an entire set of multi-objective solutions (the Pareto-optimal solutions). In this way, instead of running an optimization algorithm once a time for finding each point of such a set, the GA can find this set in a single run [110], [111].

One can build a multi-objective GA with selection operators (tournaments) that cannot decide between \((f_1=5 \text{ and } f_2=1)\) and \((f_1=4 \text{ and } f_2=2)\), by designing selection schemes by binary tournaments where the winner must dominate the loser (in the Pareto sense) to be selected.

That approach might not produce good results without the definition of a dominance tolerance. To avoid yielding strongly non-convex Pareto surfaces, dominance should be defined with a tolerance margin, \( \varepsilon \). In a minimization problem, a tolerance margin can be implemented by defining dominance as in the following:

\[
\text{Solution } x_i \text{ dominates } x_j \text{ iff:} \\
\begin{align*}
&f_j(x_i) < f_j(x_j), \text{ for some objective } j \\
&f_i(x_j) - \varepsilon \leq f_i(x_i) \leq f_i(x_j) + \varepsilon, \forall i \neq j \text{ and small } \varepsilon
\end{align*}
\]

Figure 18 shows, for two objectives \((n=2)\), a set of dominated solutions and two Pareto surfaces. In the figure, the exterior surface was yielded with high tolerance, while the strongly non-convex interior surface was yielded with small tolerance.
Figure 18: Pareto curves yielded for high (dashed line) and low (solid line) tolerance in dominance definition.

Also, note that for both low tolerance in the dominance definition and low sampling of the space (small population size), the search for the surface can be deceiving due to taking interior solutions as good objective compromises.

By formulating problems to yield Pareto surfaces, one is pulling out the best possible information for others to decide. Having obtained the Pareto set, the designer has to choose the sole solution to be implemented. At the end, in this decision-making process, someone will have to define a single objective (metric or criteria) so one can come up with a final decision. Therefore, the multi-objective problem will be turned into a pseudo multi-objective problem being solved as a single objective problem.

The problem of planning distribution networks is considered to be a difficult combinatorial optimization problem [112], [113]. It has been recognized that deterministic algorithms, such as branch-and-bound methods and dynamic programming, are unable to solve such problems beyond the scale of medium-size networks [114], [115]. This has motivated most of the recent works to employ some kind of randomized optimization algorithm [12], [26], [27], [115], [116], [117], [118], [119], [120], [121], [122]. Most of the algorithms employed in such references are GAs. GAs structured in the conventional way, however, would be inefficient for solving this class of problems [113]. In references [115] and [12], the authors have noticed the inadequacy of conventional GAs approaches to this problem, and use problem-specific GA operators combined with local search to address it. In [123] a true multi-objective optimization and problem-specific GA is employed.
5. Conclusions

This report presents a complete review of the state-of-the-art concerning the methods applied to distribution planning. It is a key element for the development of PlanGridEV as it provides a reference regarding existing methods and tools.

General information related to grid characteristics is provided for MV and LV grid of PlanGridEV partner DSOs. Their planning activity is based on the principle of maintaining the electrical system capacity to meet future demand, while maintaining service quality levels consistent with regulatory requirements and also minimizing the environmental impact of the assets. The choice of appropriate planning criteria is important to ensure a progressive improvement of safety standards and quality of electricity energy distribution under criteria of technical and economic efficiency, along with risk analysis and environmental concerns.

The first distribution planning methodologies followed a deterministic process, since the existing computational power and availability was limited. In time, parts of the process were automatized, but the main rationale remained unchanged. Recently, increasing levels of DG plus the expected rollout of EVs have been introducing uncertainties. The worst case scenario that should be evaluated is no longer necessarily peak load, as off-peak conditions could cause voltage and reactive power problems in presence of DG. At the same time, there is a greater concern in developing longer term plans with the prospect of achieving better overall solutions.

Advances in smart grids and DSM have been made as a response to these challenges. Hence, the planning assumptions must be revised to effectively integrate and consider the potential benefits of these concepts. Major advances have been made in terms of new operation scenarios including DG and/or DSM, but their integration into planning lags behind.

Additionally, the new distribution paradigm of smart grids with active participation of DER and load in network operation deeply relies on an adequate communication infrastructure. Thus, the challenges of communication must be understood and incorporated into the planning problem as alternatives to conventional reinforcements.

Today, most of the available tools handle the conventional planning problem, where single loading scenarios are addressed, typically peak load, but also maximum DG may be studied. The tools seldom tackle reconfiguration of distribution networks since this introduces integer variables into the problem control variables. Regarding the optimization objectives, almost only cost is computed and reliability may or may not be monetized into the objective function. However, tools addressing the problem of distribution planning with the inclusion of elements such as DSM, storage or considering ICT, were not identified. To address these aspects the planning tool must not study isolated snapshots of operation, but it must consider connected periods of time with inter-temporal restrictions. Storage and the activation of demand response (possible by investment on ICT) are cases where these restrictions must be present. Furthermore, no tools using multi-objective problem formulation were found. Multi-objective allows exploring in great detail the trade-offs between the different objectives, for instance cost versus reliability. Also, only one of the analysed tools included uncertainty into the formulation of the problem.

The literature review that was carried out provides a good overview of existing scientific work in the field of distribution planning. It is verified that most publications focus on the development of new optimization algorithms, while formulating the problem more or less in a similar way. Their main concern so far has been to develop robust and faster tools to address the classical planning problem. Yet, there are also a number of authors that propose the inclusion of novel features to better address
the problem and its evolving needs. Some consider the reliability issue by including a cost for the non-supplied energy, but only a few consider reliability as an objective per se. Most papers consider deterministic load modelling. Nonetheless, some authors already propose ways for dealing with the uncertainties related to load and generation using stochastic or fuzzy methodologies. A few go even further and describe the usage of risk assessment methodologies to aid the planners in making the best possible decisions. Part of the works that formulated multi-objective problems proposes the development of decision-aid methodologies for selecting the best investment decisions among dominated and non-dominated solutions.

Regarding modelling of DER, due to the very different nature of solar and wind power, different approaches and methods can be found in literature. For the generation of wind speed data, traditional methods, as there are, normal and Weibull distribution function, autoregressive methods and the Markov chain, are parametric methods and rely on time-series data. Depending on the availability of local data, different methods for forecasting and modelling solar power are to be favoured. In data-poor environments, where only historical data for the pre-analysis is available, endogenous stochastic methods can be used such as AR, MA, ARMA and ARIMA. In areas, where more data is available, exogenous inputs can be included for stochastic learning methods. Artificial neural network modelling provides an alternative approach to physical modelling, when enough historical data is available. This technique works in data-rich and data-poor areas and is used in a wide variety of time horizons (from intra-hour to yearly averages).

Statistical methods for modelling EVs rely on mobility data for parameterization. However, these methods are not that dependent on local data as the agent-based approach and can therefore more easily adapted for each area of investigation. For a high level of aggregation with a high number of vehicles, the stochastic approach seems to provide results as good as the agent-based approach. For a low number of EVs in a LV grid new approaches need to be taken in order to properly assess the grid quality. One possible parameter might be a maximum probability of charging process interference for a large sample of simulations using a low number of EVs. In terms of applicability for PlanGridEV, statistical methods provide better options for implementation than agent-based approaches, which rely on different independent tools.

In the EV aggregation models to be developed in PlanGridEV, it is crucial to include network constraints at the distribution grid level. The idea is to have location-dependent aggregations to be able to formulate all network constraints explicitly. Each network node of the MV network would have its corresponding virtual storage, while EV flows from node to node will be modelled as energy transfers from one virtual storage to another. Whereas the state-of-the art provides a good starting point, existing models will need to be adapted to the specific requirement of the planning tool to be developed in PlanGridEV.
6. References


[17] M. T. Ponce de Leão and M. A. Matos, “A Decision-Aid Method to Include Independent Fuzzy


Meeting, Minneapolis, 2010.


[90] J. L. Mathieu, M. G. Vayá and G. Andersson, Uncertainty in the Flexibility of Aggregations of


7. Revisions

7.1. Track changes

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
<th>Version</th>
<th>Changes</th>
<th>Subject of change</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedro Rocha Almeida</td>
<td>20.11.2013</td>
<td>1.0</td>
<td>First version</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Pedro Rocha Almeida</td>
<td>20.11.2013</td>
<td>1.1</td>
<td>References reviewed</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Luís Silvestre</td>
<td>27.11.2013</td>
<td>1.1</td>
<td>Document Review Process</td>
<td></td>
<td>all</td>
</tr>
<tr>
<td>Pedro Rocha Almeida</td>
<td>02.12.2013</td>
<td>1.3</td>
<td>Author revision</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Eduardo Zabala</td>
<td>13.12.2013</td>
<td>1.4</td>
<td>Review Technical</td>
<td>Review Technical Coordinator</td>
<td>-</td>
</tr>
<tr>
<td>Pedro Rocha Almeida</td>
<td>18.12.2013</td>
<td>1.5</td>
<td>Author revision</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Armin Gaul</td>
<td>20.12.2013</td>
<td>1.6</td>
<td>Final and approved</td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>
Annex A. Distribution Planning Tools

DPLAN

DPLAN is used by EDP planning teams. It allows performing expansion planning, operation planning and emergency planning. INVESTE is then used to perform a full economic evaluation on the solutions provided by DPLAN.

DPLAN includes several software modules, each focusing on a different topic:

- Investment decisions
- Topological decisions
- Voltage-current analysis
- Reliability: fault analysis and corrective switching

The optimization problem is formulated as a minimization of a value function $f(u)$, using AC power flow computation for evaluating $f(u)$, losses, reliability, among other user-defined criteria.

The discrete nature of $u$ led to usage of evolutionary algorithms instead of mathematical programming. Other features of DPLAN include the integration with DINIS (another software used by EDP) and the usage of Smallworld GIS (currently from GE energy).

Some technical and economical advantages for EDP:

- Having a detailed geographical view electrical network
- System database maintenance and updating costs reduction
- Information unification under one common model
- Flexibility and information flow between the different systems increase
- Field work support improvement
- Making internal company processes more efficient and improving the reply times
- High-resolution cartography and aerial photo maps support for technical studies and analyses
- Electrical network project co-ordination
- Network information on the Web, through SITweb

A picture from DPLAN interface is presented next.
**NEPLAN/Integral**

NEPLAN was developed by ABB and INTEGRAL by a university. It allows integrating several systems used in RWE.

It provides the strengths of the standard Software of ABB and those of the tailored product, INTEGRAL, with custom made features.

NEPLAN/INTEGRAL determines an Optimal Network configuration minimising ENS, Energy Losses and investment proposals.

It is applied to a planning horizon of 10 years and allows developing/calculating network solutions in line to the actual planning rules.

**MSPLAN**

MSPLAN was developed by a university and allows integration with different systems used in RWE. MSPLAN determines an Optimal Network configuration minimising ENS, Energy Losses and investment proposals.

It is used for a planning horizon of 40 years for checking variations of Network configuration.
SynerGEE Electric

SynerGEE Electric enables users to quickly establish whether a new customer can be connected on existing network or if a network reinforcement is required. It determines when new Overhead Lines, Underground Cables or Stations are required as a result of natural load growth.

SynerGEE Electric allows modeling the MV network for Volt-Drop and Thermal limits on Overhead lines and Underground cables. It assesses the necessity to reinforce the network to accommodate new loads and identify reinforcement.

The scope of application is:
- Network voltage: 10kV, 20kV
- Planning horizon: 1-25 years

The tool relies on a GIS model in the background, which is updated on a regular basis to reflect the dynamic nature of the network and to ensure the network being modelled reflects the ‘real-world’ situation.

SynerGEE Electric has a vast array of features that can be used for studying:
- Load Flow: Voltage and Thermal ratings
- DG and DER on the distribution network
- Reliability studies
- Phase Balancing studies
- Voltage Regulation & Transformers
- Protection

PECO Model

The PECO model was developed by IIT-UPCO, [36]. Developed by a research institution it has been used both for industrial purposes and research and development activities. As application examples, it has been used in consulting activities for regulators, with the tariff revision on Chile, Argentina and Spain, and in R&D, for the MERGE project (distribution planning with EVs).

The main potential users are utilities or regulators. The purpose of PECO is to design large distribution systems. It develops reference network models from scratch as well as it does expansion planning.

The PECO model can be used in different running modes:
- Greenfield planning, given the size and location of transmission substations and the GPS coordinates of LV, MV and HV customers
- Planning the distribution network given the location of HV/MV substations and/or the MV/LV transformers
- Given the topology of the LV or MV network, optimize conductor sizing, losses and/or reliability of supply

The software is composed by several modules (operated sequentially):
- Locating MV/LV transformers
- Locating HV/MV substations
- LV network planning
- MV network planning

PECO’s features include:
• GPS coordination of customers and transmission substations to build the whole distribution network.
• Automatic determination of the settlement's outlines and their corresponding street maps.
• Orography raster maps, forbidden ways through... given by the user.
• The optimization of both urban networks and rural networks performed simultaneously.
• Network sizing by: minimizing investment, operation & maintenance, losses subject to capacity, voltage and reliability constraints, and taking into account present demand and an estimation of its future growth.
• Network reliability assessment by simulation of the real process after a network failure has taken place.
• Reliability assessment computed using the parameter "cost of non-supplied energy", which can be different for each customer.
• Possibility of exportation of the results of the model to the Arc/View's Shape format.

PECO required input data is the following:
• Equipment
• LV, MV and HV aerial lines and underground cables
• Aerial and underground HV/MV substations
• Aerial and underground MV/LV transformers
• Protective devices for the MV network
• Capacitors and voltage regulators for the MV network
• Technical and economic parameters
• Rate of return to assess the present worth of investment, losses and operation and maintenance
• Rate of demand growth
• Number of years of linear depreciation
• The cost of energy losses
• Load and loss factors for LV, MV and HV networks
• Simultaneity factors in LV, MV and HV networks
• Demand modelling
• Definition of a settlement
• Rate of aerial/underground LV, MV and HV networks
WindPRO

WindPRO is a tool for planning single wind turbines and wind farms with modular software structure. A non-extensive list of WindPRO software modules is listed next:
• Resource: wind resource maps (GIS based)
• Meteo: meteorological data import and calculation
• eGrid: calculation of electrical grid connections
• WindPlan: planning and/or site finding based on GIS data

The eGrid module performs grid analysis for:
• Computing losses in cables and transformers
• Performing cable/transformer design
• Study steady stage voltage variations based on load conditions
• Determine short circuit power and current
• Analyse long term flicker issues

Smallworld GIS – Operation Analyser

Smallworld GIS – operation analyser provides an electricity network design process to Smallworld GIS. It manages different data formats, such as Neplan or PowerFactory and is able to simulate the connection of new consumers and also of distributed generation.

CYMDIST - Distribution System Analysis

The CYMDIST software is designed for distribution planning studies and simulating the behaviour of electrical distribution networks under different operating conditions and scenarios. The analysis functions such as load flow, short-circuit, and network optimizations, are performed on balanced or unbalanced distribution network that are built with any combination of phases and configurations.

The software is also equipped with add-on modules to perform more in-depth analyses such as reliability analysis, contingency analysis, harmonic analysis, switching (tie-points) optimization, and more.

CYMDIST may be used in “what-if” studies and to performing simulations to evaluate the impact of modifications to the system.

Analytical capabilities of CYMDIST:
• Balanced and unbalanced voltage drop and short-circuit analyses (radial, looped or meshed)
• Protective device coordination verification according to user-defined criteria for device clearance and loading
• Fault current calculations for RMS, asymmetrical and peak values for all shunt fault configurations
• Short-circuit and fault voltage analysis throughout the network taking into account pre-fault loading conditions
• Optimal capacitor placement and sizing to minimize losses and / or improve voltage profile
• Minimum fault protection analysis
• Load balancing to minimize losses
• Load allocation/estimation using customer consumption data (kWh), distribution transformer size (connected kVA), real consumptions (kVA or kW) or the REA method. The algorithm supports multiple metering units as fixed demands and large metered customers as fixed load.
• Motor starting analysis (voltage dip and maximum motor size allowable)
• Flexible load models for uniformly distributed loads and spot loads featuring independent load mix for each section of circuit
• Load growth studies for multiple years
• Feeder interconnection for load transfer simulations
• Phase merging capability
• Automatic re-conductoring and re-phasing of multiple selected sections
• Computes load equivalents and network equivalents to ease the analysis of large networks, matching exactly the power flow and short-circuit results of the non-reduced network
• Distributed generation modelling, generator impedance estimation, grid side control and protection functions

**PSS®SINCAL**

PSS®SINCAL is used for:
• Feeder planning
• Calculating routes
• Determining line data
• Determining power data
• Determining polygon data (for load density and areas)
• ISO areas handling

The integration of distributed generation on one hand allows utilities to adequately design their systems, but on the other hand provides new options for the improvement of power supply while respecting environmental protection aspects and even allows an optimized influence on the energy consumption with the help of smart meters.

Specific features:
• Specific modelling of distributed generation units
• Quasi-dynamic simulations
• Integration of smart meter data
• Impact on the protection system
• Stability analyses in unbalanced systems

**LAPER**

The purpose of LAPER is to perform rural electrification. It was developed by EDF and ADEME. The planning principle of LAPER follows 5 steps:
• Data collection: needs and means
• Initial state design
• Search for the optimal electrification pattern
• Electrification schedule
• Results