D 2020

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

TECHNOLOGY DIFFUSION MODELS IN POWER SYSTEM PLANNING AND POLICY DESIGN

FABIAN HEYMANN

Thesis submitted to the Faculty of Engineering of University of Porto
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Sustainable Energy Systems

Fabian Heymann

Technology diffusion models in power system planning and policy design

Universidade do Porto

Faculdade de Engenharia

FEUP

ACADEMIC DISSERTATION IN *SUSTAINABLE ENERGY SYSTEMS*

Dissertation submitted to the Faculty of Engineering of University of Porto in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Faculty of Engineering
University of Porto

Porto, May 2020

Supervisor:	Vladimiro Miranda, Ph.D., Full professor, Faculty of Engineering and Center of Power and Energy Systems, University of Porto and INESC TEC Porto, Portugal
Co-supervisor:	Filipe Joel Soares, Ph.D., Senior researcher, Center of Power and Energy Systems, INESC TEC Porto, Portugal
Co-supervisor:	Pablo Duenas, Ph.D., Research scientist, MIT Energy Initiative Massachusetts Institute of Technology Cambridge (MA), United States of America

The author gratefully acknowledges funding under MIT Portugal Scholarship PD/BD/114262/2016 and project SmartGP/0002/2015 under the framework of ERA-Net Smart Grids Plus, financed/co-financed by FCT. In addition, the author of this work is grateful for the strong institutional support provided by INESC TEC and MITEI – the hosting instutions of the research that culminated into this thesis.

Deixe-me ir
Preciso andar
Vou por aí a procurar
Sorrir pra não chorar
Quero assistir ao sol nascer
Ver as águas dos rios correr
Ouvir os pássaros cantar
Eu quero nascer
Quero viver

Cartola - Preciso Me Encontrar

Abstract

This thesis is dedicated to the study of technology adoption patterns and their prediction in space and time.

For the case study of Continental Portugal, a combined spatial and non-spatial framework to characterize technology adopters is presented. The framework is using non-parametric inference methods rooted in Information Theory and explorative spatial data mining techniques.

On this ground, a rigorous analysis of various models to represent technology adoption dynamics is conducted.

Providing a simulation-based spatiotemporal technology adoption model, studies assess the uncertainties introduced by poor representations of technology adoption dynamics in electricity network planning.

In addition, the simulation-based model was used to investigate the effects of energy policy changes on technology adoption patterns. Such studies allow to define optimized incentive designs derived from steering technology adoption towards reduced systemic impacs.

Resumo

Esta tese é dedicada à análise de padrões temporais e espaciais de adoção de novas tecnologias.

Foram analisadas metodologias baseadas em métodos espaciais e não-espaciais combinados para caracterizar os utilizadores mais propensos à adoção de novas tecnologias, utilizando, como caso de estudo, Portugal Continental. Estas metodologias utilizam métodos de inferência não paramétricos provenientes do campo da Teoria da Informação e técnicas exploratórias de *data mining*.

Diversas metodologias para caracterizar as dinâmicas de adoção de novas tecnologias foram analisadas de forma rigorosa, utilizando o caso de estudo selecionado.

Posteriormente, o modelo espácio-temporal desenvolvido foi utilizado para avaliar as incertezas introduzidas por representações mais simplistas e menos rigorosas da dinâmica de adoção de novas tecnologias no planeamento de redes de elétricas.

Adicionalmente, o modelo de simulação foi utilizado para investigar os efeitos que diversas políticas energéticas terão, potencialmente, nas dinâmicas de adoção de novas tecnologias. Este modelo de simulação permite também otimizar as políticas de inventivo à adoção de novas tecnologias, de forma a alcançar os objetivos traçados pelas entidades decisoras.

Acknowledgements

One of the first PhD thesis I read contained a strong metaphor, comparing the research period with a sailing trip.

Looking back on my time in research, I cannot agree with this image. I rather like to look at this time as a joyful but energy-sapping kajak tour. Instead of relying on the probabilistic nature of winds, the kayaker persistently uses his force, circumventing rocks and choosing river branches either ending in waterfalls or eventually leading him to the safe boathouse. Although the kajaker is mostly alone out on the waters, his trip requires strong support from his team which remains onshore.

The time has come to thank my team:

Thank you, Vladimiro Miranda, who in a sense, initially provided me with the kayak to start the trip together with the full trust in my abilities. You represented an unshakable compass through the intellectual adventures to come. You propelled the voyage, setting new, exiting horizons and making me outgrowing myself.

Thank you both Filipe Joel Soares and Pablo Duenas, for being the strong currents of my trip. You kept me going every meter and made me elegantly passing any rocks or icebergs that crossed my way in the turbulent waters. Apart from being the perfect company for the past years, you would also let me look over your shoulders and learn about the daily life of a researcher.

I thank all my dear Portuguese friends, colleagues and professors from FEUP and INESC TEC, my paddles, who helped me to translate spikes in neural activity into force, and, by times, also sufficiently distracted me from my project. You taught me unconditional hospitality, curiosity and patience.

Let me also thank you, Célia Couto, Sandra Oliveira Nunes and Paula Castro. If the paddles were already taken, you were the reliant, live-saving coating of the kajak. You made sure that the trip was never ended, holding back all the threatening waves pressing against my boat.

I am also grateful to the Portuguese Foundation of Science and Technology (FCT) that financed my research. It opened me the door into a new world and allowed me to mingle with all the brilliant minds in the MIT Portugal Program. Especially being non-Portuguese, I will not forget the hospitality and support I received from FCT and many other Portuguese institutions.

Luckily, I already had the chance to express my gratitude to representatives of the Portuguese Government, explaining the Portuguese Minister of Higher Eduction on a round table that a part of the attractiveness of Portugal as a science hub also roots in the beauty of its people.

I would also like to thank my most inner circle of emotional support, all my family members and my lover that strengthend the trust in my abilities and helped me going further than I ever expected. A special thanks goes to the members of *The Three Noses*, that provided me with strength and confidence since the first day in research until the last minutes before thesis defense.

I am also very grateful for the collaboration and idea exchange with inspiring thought leaders and personalities in the electricity sector such as Prof. Ignacio J. Pérez-Arriaga (MIT), H. Lee Willis (Quanta Technology), Tom Osterhus (Integral Analytics) or Richard Tabors (Tabors Caramanis Rudkevich), who I met during my stay at the MIT Energy Initiative. In addition, I benefited at large from the excellent collaboration with Joel David Melo Trujillo (Federal University of ABC, Brazil), whose work and advice left a strong mark on my thesis."

Obviously, I feel deeply bond to my old kajak club, my *professoras* Armanda and Vera and all my teammates that made me spend so many saturday mornings on the Mondego river (in my blue Ferrari). Human-machine interaction and human existance becomes so easy out on the water.

I realized that this thesis also tells a lot about myself, reflecting my strong passion for multidisciplinary research. Hopefully, it will be a rich source for new transversal concepts to the interested reader.

Finally, readers are empowered with a new, surprising perspective: Spatial science and power system planning can be good dancing partners.

Duisburg, 13th May 2020

Fabian Heymann

Contents

Figures

Tables

Acronyms and symbols

Acronyms

at	DER allocation technique
ACC	Acccuracy
ANN	Artificial Neural Networks
DER	Distributed Energy Resources
DSO	Distribution System Operators
EV	Electric Vehicles
FITs	Feed-in tariffs
GIS	Geographic Information Systems
HP	High Performance Households
HVAC	Heating, Ventilation and Air Conditioning Devices
IDC	Incentive Design Combinations
IT	Information Theory
LISAs	Local Indicators of Spatial Association
LMI	Low-or-Medium Income Households
MLR	Multi-Linear Regression models
PV	Photovoltaics
RN	Randomized Allocation
RMSE	Root Mean Square Error
T/D	Transmission/Distribution
TSO	Transmission System Operators
UA	Uncertainty Analysis

Symbols

α	load growth rate
a_i	mean centered values
af	adoption-influence
afc	adoption favoring criteria
AP_t	accumulated products adopted until period t.
b_j	means for all neighbor values of polygon i.
bs	battery size
c	socio-demographic criteria
cp_h	normalized hourly EV charging profile
CT	change in Theil's T
D	distance metric
dp	generic distance parameter
ED	Euclidean distance
$E(I_i)$	expected value of I_i
FN	false negatives
FP	false positives
GF_{EV}	global, annual EV forecasts
GF_{PV}	global, annual PV forecasts
gp_h	normalized hourly PV generation profile
h	hour
H	Information Entropy
H(X,Y)	joint entropy of X and Y
HV/MV	high voltage/ medium voltage
I	Moran's I
IC	investment costs

IG	Information Gain
IGR	Information Gain Ratio
IP	investment plan
IS	innovativeness scores
i	polygon index
L	natural load
lp_h	normalized hourly load profile
M	total market potential
MAE	Mean Average Error
MD	Manhattan distance
MI	Mutual Information
MV	medium voltage
NL	net-load
N_{EV}	number of EV
n_t	number of first-time purchases
N_{PV}	number of PV
NO	number of observations
occ	other census criteria
p	coefficient of innovation
PA	peak load added
pccs	per-capita car-share ratio
pcder	aggregated per-capita share of DER per census cell
pcpv	per-capita PV potential
PL	Peak load
PNL	Peak-net-load
pval	p-value

pr	PV module performance ratio
q	coefficient of imitation
r	spatial census tract
r.area	total roof-top area
RF	reverse flow hours
sc	scenario
sa	HV/MV substation service area
smi	simulated Moran's I values above determined
SMI	total number of simulated Moran's I values
ST	cellular adoption states/stages
TC	HV/MV transformer upgrade costs per unit
TCAP	transformer capacities
TDER	maximum DER within spatial census cells
TEV	maximum EV potential
TN	true negatives
TP	true positives
TPP	total aggregated population in all census cells
TPV	maximum PV potential
TT	Theil's T
TTC	total aggregated HV/MV transformer upgrade costs
u.fract	usable roof fraction
$Var(I_i)$	variance of I_i under randomization hypothesis
w_{ij}	spatial weight matrix
$\hat{y}$	lagged mean
y_i	polygon value
ys	forecasting horizon in years

1 Introduction

The adoption of new technologies such as distributed energy resources (DER) requires holistic frameworks to analyse, compare and predict their diffusion patterns in space and time. This thesis is the first contribution of a sophisticated framework that allows for a more accurate and reliable representation of technology diffusion processes in electricity network planning and policy decision processes.

1.1 BACKGROUND

Worldwide, residential consumers have been adopting new distributed energy resources (DER). This concept is nowadays taken in a broad sense, including photovoltaics (PV), electric vehicles (EV), distributed storage, together with electrified heating, ventilation and air conditioning devices (HVAC) [1], [2]. Such appliances impact on the individual electricity demand profiles and thus redesign the geography of consumers and generation sites [3], [4]. Hence, the substantial change of electricity network planning techniques and energy policy tools under the large-scale adoption of such appliances is deemed necessary [3]–[5].

Electricity networks link power generation to electricity demand. Thus, the geography of consumer patterns, that is, their location in space, consumption times and magnitude, eventually determine the layout of electricity networks [6]. DER are changing the distributed electricity consumption+generation morphology and therefore impact on network planning [7], [8]. However, traditional network planning routines rarely consider the structure and propensity of consumers to adopt new technologies. The reason for this is that usually planning has its roots in forecasting, and a common forecasting assumption is that the future is structurally and in human behaviour similar to the past and present.

However, new technologies may be disruptive of a sustained pattern and make such assumption invalid. Consequently, usual oversimplified representations of DER adoption dynamics are unable to capture the large-scale technology diffusion [9] and this will result in suboptimal network investment decisions [3], [10], [11]. Recent studies suggest that the prediction of future spatial distributions of DER such as PV, EV, and HVAC may bring in high economic value to energy utilities. In a first case study, absent or existing, accurate DER adoption forecasts could decrease or increase network companies revenues by several millions of U.S. dollars per TWh consumed [12].

On the other hand, the installation and operation of DER at the consumer's site, or their generalized use (such a with EVs) does offer potential benefits, including self-consumption, arbitrage trade, shifted consumption and flexibility provision [13]. Energy policies that stimulate the use of DER in a synergetic way can provide substantial value additions – both for the individual consumer level as well as for the electricity system as a whole [14]–[18]. Examples of such benefits are reduced electricity bills through increased self-consumption, price arbitrage, and revenues from providing system services like demand response. Relevant system benefits are reduced CO_2 emissions, reduced peak loads, and less congestion and losses in distribution grids. At the same time, new business models are emerging that are also possible cause for other types of disruptions; e.g. net metering introduced potential perspectives of serious unbalance in the soundness of the traditional financial operation of distribution utilities and storage is pressing business models to be supported more on CAPEX than OPEX.

The usability of synergetic potentials requires complete understanding of the mechanisms that drive technology adoption and models capable of predicting their occurrence within electricity networks. In order to exploit such potential benefits and mitigate problems derived from DER, locational and temporal signals for DER utilization must become well-concerted [17].

The implementation of incentive schemes to induce technology adoptions or behaviour changes is a common tool in public policies. Incentive designs affect the nature and grade of consumers' decision to adopt a certain technology. Therefore, they represent the first and necessary condition for determining the spatial distribution of DER within a social system. Electricity tariffs, on the other hand, contribute to steer the temporal use of a technology at a location. They therefore represent the, second, sufficient condition to exploit DER synergies.

This thesis is dedicated to analysing the first, necessary condition – spatiotemporal patterns of DER adoption, under a diversity of stimuli. The impacts of new technologies in the power system structure and operation are modelled as a diffusion process, and the spatiotemporal progressive impacts are observed with appropriate tools and tested in study cases, in a globally novel approach.

The thesis presents a sophisticated framework that allows for a more accurate and reliable representation of technology diffusion processes in electricity network planning and policy decision processes. This study is of the utmost importance, in a time of change of paradigms.

1.2 Research questions

The thesis presents a framework to characterize technology adopters as well as a rigorous analysis of various models to represent technology adoption dynamics in electricity network planning and energy policy studies. The presented tools enable a detailed characterization of technology diffusion processes in space and time and allow for both insight and foresight, having a high applicability both to policy analysis and electricity network planning.

To understand the lines of investigation pursued, a set of research questions was identified and organized. The first block of research questions aims at surveying and further-developing approaches to categorize DER adopters. Furthermore, DER adoption drivers and descriptive analysis of resulting patterns shall be investigated on, showcasing methodological advances on a case study (Continental Portugal). Eventually, produced outcomes are further used to develop a spatiotemporal model for DER adoption forecasting.

Previously mentioned research gaps are addressed in Chapter 2 and Chapter 3. The corresponding research questions these chapters answer are:

- How can one describe and compare technology adoption patterns?
- How to predict technology adoption patterns in space and time?
- Which components do spatiotemporal technology adoption models typically consist of?
- Along which criteria can technology adoption models be categorized?

Instead, Chapter 4 and Chapter 5 are directed towards research questions that originate in current challenges in electricity network planning and policy analysis and design. Those are:

- What are the effects of different technology representations on electricity network planning?
- How do different policy designs affect system expansion costs and distributional effects?
- Can orchestrated incentive designs reduce such costs/effects?

Finally, a global, guiding research question that has been followed throughout this thesis is:

- How can one enhance the representation and modelling of technology diffusion dynamics in electricity network planning and policy design studies?

1.3 THESIS CONTRIBUTIONS

The thesis presents a set of conceptual and mathematical models applied to case studies. Such models allow to characterize, compare and predict technology diffusion dynamics. The resulting framework is used to explore a variety of current challenges present in electricity network planning and energy policy design throughout this thesis' chapters.

The thesis introduces a spatiotemporal DER adoption forecasting model that can mimic large-scale technology diffusion dynamics in space and time. Apart of a thorough validation process, it has been consecutively applied to explore interlinked topics in electricity network planning and energy policy design.

One of the originalities of this work lies in its multidisciplinary embeddedness and comprises contributions to a wide array of academic fields, including diffusion research, electricity network planning and energy policy design and markets.

To allow a reader to understand the lines of development of the research reported in this thesis, the main contributions of the thesis are immediately referred to below, and will be further commented in the Conclusions chapter. They can be differentiated along conceptual, methodological and case study-level:

Main conceptual advances:

- Measuring DER adoption pattern dispersion using spatial data mining.
- Establishment of a framework to compare different technology adoption forecasting techniques, including a thorough analysis of the potentials and limitations of technology diffusion models.

Main methodological advances:

- Extension of technology adopter characterizations to non-parametric approaches rooted in Information Theory.
- Development of a flexible, census data-based spatiotemporal model to forecast DER adoption.
- New metrics for measuring the uncertainty in flows between the transmission-distribution boundary.
- Modelling the spatial DER adoption patterns under different policy designs.
- Methodology to estimate the trade-offs of DER system integration cost versus adoption asymmetries.

Main case studies:

- Holistic description of Portuguese DER technology adopters and adoption drivers.
- Network expansion cost estimates under different representations of the DER adoption process in Porto Municipality and Continental Portugal.
- Study of the effect of various policy designs on electricity network expansion and technology distribution asymmetries in Continental Portugal.

1.4 THESIS STRUCTURE

The thesis comprises six chapters. While this chapter provides a brief description of the background and scope of this thesis, Chapter 2 systematically reviews past research that has been directed to the characterization of energy technology adopters. Differentiating between spatial and non-spatial approaches, Chapter 2 further presents a combined non-spatial/spatial analysis of EV, PV and HVAC adopter in Continental Portugal.

Chapter 3 is dedicated to the introduction of models that are currently available to model technology diffusion in space and time. It introduces the main theoretical frameworks on technology adoption and diffusion. After highlighting current state-of-the-art models on energy technology diffusion, this section further presents the spatiotemporal discrete-state technology diffusion model that has been introduced in the frame of this thesis. Furthermore, this chapter contains a comparison of the proposed model to current state-of-the-art models used to forecast technology adoption patterns in space and time.

Chapter 4 presents the application of the developed spatiotemporal technology diffusion model to electricity network planning problem. Both distribution and transmission network case studies have been conducted. This chapter shows the models relevance under a spatial perspective of electricity network planning.

Building on the outcomes of Chapter 3 and Chapter 4, Chapter 5 analyses the effects of changes in policy designs. It introduces and discusses different energy policy instruments currently used to stimulate the diffusion of energy technologies. It further investigates how changes in energy policy instruments affect the spatial distribution and thus, the impact of energy technologies. Additionally, the effect of different incentive designs in adoption asymmetry and the allocation of technologies' benefits are assessed.

Finally, Chapter 6 presents the main conclusions drawn from the studies contained in this thesis. This chapter not only presents a thorough discussion on model development and simulation outcomes, but further provides a detailed outlook to future applications and open research questions.

The wide range of contributions presented in this thesis is reflected in its intrinsic structure: each of the chapters is presented as a self-contained section. Therefore, a dedicated conceptual review and the presentation of previous works is placed at the beginning of each of such chapters. Such introduction contextualizes the models and numerical analysis presented hereafter. The thesis structure and research questions addressed are shown in Figure 1.1.

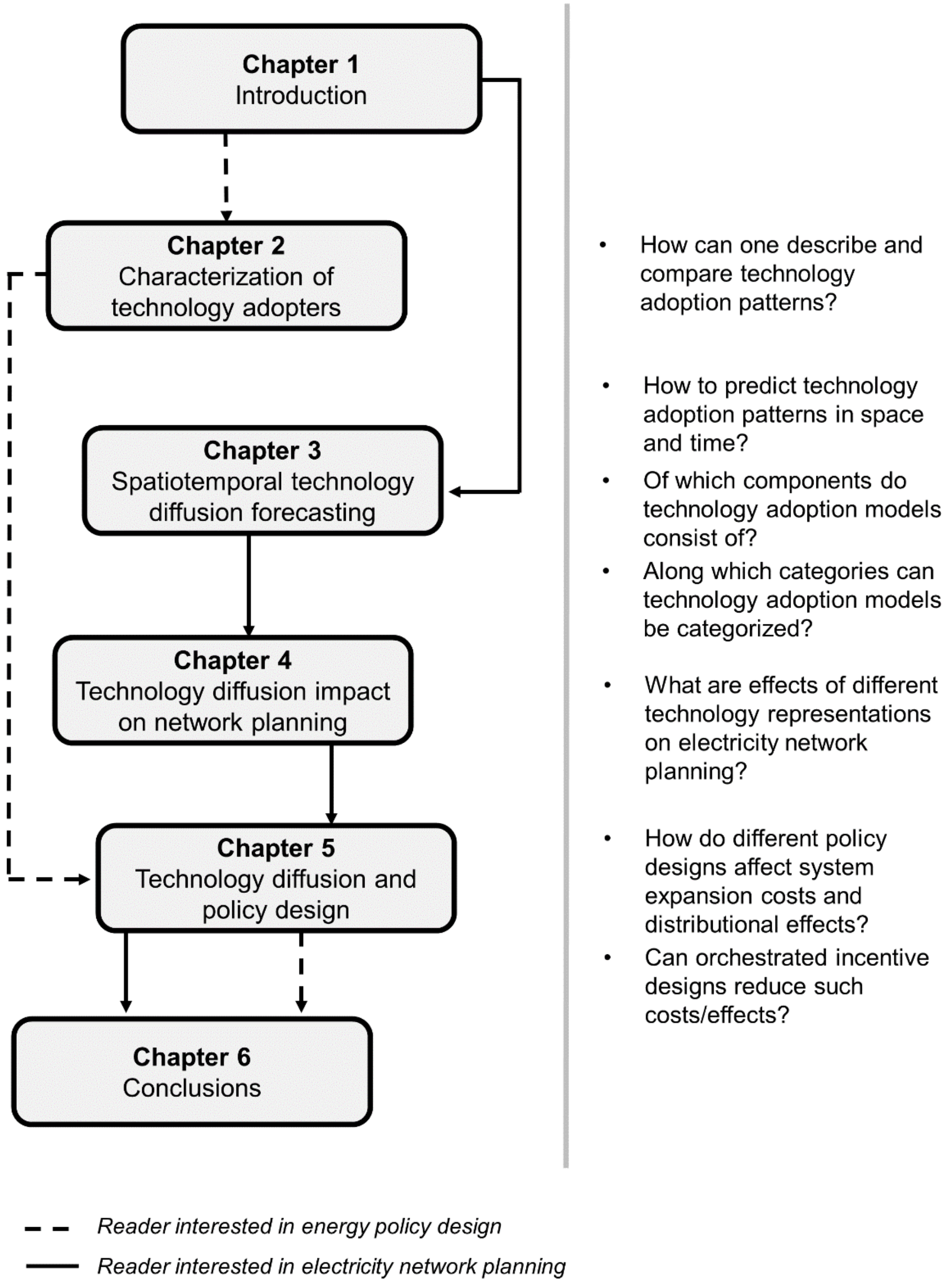


Figure 1.1 Thesis structure and questions addressed

1.5 A NOTE ON CASE STUDIES AND DATA SOURCES

The proposed models and idea-sets of this thesis have been widely tested and applied to realistic case studies. Using openly available census data-set from the Portuguese Institute of Statistics [19], this thesis embraces studies from municipal distribution system level to analysis that spans over both major parts of the national distribution system and distribution-transmission interface of Continental Portugal.

The advancements presented especially in Chapter 2 and 3 have been made possible through the provision of detailed information on energy technology adopters. The complete set of residential EV adopters by the end of year 2016 could be retrieved through the Portuguese charging infrastructure platform operator. Further adopter data-sets including a subset of residential PV and HVAC adopters have been provided through the Portuguese energy agency (ADENE). The joint availability of adopter information comprising various technologies allowed for unique insights that have been impossible to retrieve within many other studies.

Finally, information on the Portuguese electricity distribution system, notably the geographical location, installed capacity and peak-load values of HV/MV transformers have been publicly available under [20], [21]. Based on those, HV/MV transformer service areas, that is, spatial zones that are supplied by a given HV/MV transformer, could be computed using Voronoi diagrams. This process will be detailed in Chapter 4.

All models presented in this thesis are flexible in the sense that presented approaches can be transferred to any other distribution or transmission case study in case respective input data are provided. It is assumed that for industrial implementation, distribution system operators can reconstruct and refine population information unleashing extensive customer information stored in their client data-bases.

Transmission grid planners and operators, on the other hand, may exploit already existing data-sets such as the TIGER products [22] that cover the United States of America perimeter or the European census hub [23]. Such data-bases come with sufficient granularity that is required for transmission planning and can be crossed with information on the installation of energy technologies in each census polygon.

Although the thesis addresses aspects of power system planning, the presented spatiotemporal technology diffusion model represents a convenient tool to build spatially resolved scenarios for many future technologies or other network industries such as water, ICT or transportation.

1.6 SCIENTIFIC PUBLICATIONS

The scientific advancements presented in this thesis were published in five international journals and seven international conference contributions. Currently, two additional publications are prepared. All international publications are shown below:

International journal papers

I. F. Heymann, J. Silva, V. Miranda, J. Melo, F. J. Soares, and A. Padilha-Feltrin, "Distribution network planning considering technology diffusion dynamics and spatial net-load behavior," *Int. J. Electr. Power Energy Syst.*, vol. 106, pp. 254–265, 2019.

II. F. Heymann, P. Duenas, F. J. Soares, V. Miranda, I. Perez-Arriaga, R. Prata, "Orchestrating incentive designs to reduce adverse system-level effects of large-scale EV/PV adoption – The case of Portugal", *Applied Energy*, vol. 256, 113931, 2019.

III. L. Rodrigues, H. M. Bolognesi, J. D. Melo, F. Heymann, and F. J. Soares, "Spatiotemporal model for estimating electric vehicles adopters," *Energy*, vol. 183, 2019.

IV. F. Heymann, F. J. Soares, P. Duenas, and V. Miranda, "Explorative Spatial Data Mining for Energy Technology Adoption and Policy Design Analysis," *Springer Lect. Notes Artif. Intell.*, vol. 11804, pp. 1–11, 2019.

V. F. Heymann, M. Lopes, F. vom Scheidt, J.M. Silva, P. Duenas, F. J. Soares and V. Miranda, "DER Adopters analysis with Spatial Autocorrelation and Information Gain Ratio under different Census-data Aggregation Levels," *IET Renew. Power Gener.*, vol. 14 , issue 1, 2019.

International conferences

VI. F. Heymann, P. Vilaca, J.V. Silva, P. Duenas, J. Melo, F. J. Soares and V. Miranda, "Vertical Load Uncertainty at the T/D Boundary under different spatial DER allocation techniques," *IEEE Sustainable Energy Systems and Technologies Conference*, Porto, 2019.

VII. F. Silvestro, F. Pilo, F. Heymann, M. C. Alvarez-Herault, M. Braun, J. C. Araneda, J. Taylor, "Review of transmission and distribution investment decision making processes under increasing energy scenario uncertainty" *25th International Conference on Electricity Distribution (CIRED)*, Madrid, 2019.

VIII. F. Heymann, V. Miranda, M. Lopes, F. J. Soares, and A. Dias, "Synergies between Electric Vehicles and Distributed Renewable Generation?", *11th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2018)*, Dubrovnik, 2018.

IX. F. Heymann, J. Melo, P. Duenas, F. Soares, and V. Miranda, "On the Emerging Role of Spatial Load Forecasting in Transmission / Distribution Grid Planning," in *11th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2018)*, Dubrovnik ,2018.

X. F. Heymann, P. Duenas, F. Soares, and V. Miranda, "Explorative ex-ante consumer cluster delineation for electrification planning using image processing tools", *11th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2018)*, Dubrovnik, 2018.

XI. F. Heymann, N. Neyestani, F. J. Soares, V. Miranda, "Mapping the impact of daytime and overnight electric vehicle charging on distribution grids", *IEEE Vehicle Power and Propulsion Conference (VPPC)*, Belfort, 2017.

XII. F. Heymann, C. Pereira, V. Miranda, F. J. Soares, "Spatial load forecasting of electric vehicle charging using GIS and diffusion theory", *2017 IEEE PES International Smart Grid Technologies Conference Europe (ISGT)*, Turin, 2017.

<u>Under review/ development</u>

XIII. F. Heymann, F. vom Scheidt, J. Melo, P. Duenas, F. J. Soares, V. Miranda, "Forecasting DER adoption in space and time: Model development, calibration and parameter choice", *under development*, 2019.

Chapter references

[1] IRENA - International Renewable Energy Agency, "Renewable energy highlights," Abu Dhabi, United Arab Emirates, 2019.

[2] REN21, "Renewables 2019 global status report 2019," Paris, 2019.

[3] EPRI, "Developing a Framework for Integrated Energy Network Planning (IEN-P)," Palo Alto, California, 2018.

[4] K. Bell, "Methods and Tools for Planning the Future Power System: Issues and Priorities," *Model. Requir. GB Power Syst. Resil. Dur. Transit. to Low Carbon Energy*, no. Paper 5, pp. 1–35, 2015.

[5] I. J. Perez-Arriaga, J. D. Jenkins, and C. Batlle, "A regulatory framework for an evolving electricity sector: Highlights of the MIT utility of the future study," *Econ. Energy Environ. Policy*, vol. 6, no. 1, pp. 71–92, 2017

[6] A. Zvoleff, A. S. Kocaman, W. T. Huh, and V. Modi, "The impact of geography on energy infrastructure costs," *Energy Policy*, vol. 37, no. 10, pp. 4066–4078, 2009.

[7] M. De Gennaro, E. Paffumi, H. Scholz, and G. Martini, "GIS-driven analysis of e-mobility in urban areas: An evaluation of the impact on the electric energy grid," *Appl. Energy*, vol. 124, pp. 94–116, 2014.

[8] R. Bernards, J. Morren, and H. Slootweg, "Development and Implementation of Statistical Models for Estimating Diversified Adoption of Energy Transition Technologies," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1540–1554, 2018.

[9] Federal Energy Regulatory Commission, "Distributed Energy Resources - Technical Considerations for the Bulk Power System," 2018.

[10] G. Feismann, "Handbook / White Paper of Novel Planning Rules for EV in Smart Grids," 2016.

[11] F. Heymann, J. Silva, V. Miranda, J. Melo, F. J. Soares, and A. Padilha-Feltrin, "Distribution network planning considering technology diffusion dynamics and spatial net-load behavior," *Int. J. Electr. Power Energy Syst.*, vol. 106, pp. 254–265, 2019.

[12] P. Gagnon, B. Stoll, A. Ehlen, G. Barbose, and A. Mills, "Estimating the Value of Improved Distributed Photovoltaic Adoption Forecasts for Utility Resource," May, 2018.

[13] European Commission, "Clean Energy For All Europeans," Brussels, 2016.

[14] A. Chaouachi, E. Bompard, G. Fulli, M. Masera, M. De Gennaro, and E. Paffumi, "Assessment framework for EV and PV synergies in emerging distribution systems," *Renew. Sustain. Energy Rev.*, vol. 55, pp. 719–728, 2016.

[15] D. B. Richardson, "Electric vehicles and the electric grid: A review of modeling approaches, Impacts, and renewable energy integration," *Renew. Sustain. Energy Rev.*, vol. 19, pp. 247–254, 2013.

[16] M. H. Moradi, M. Abedini, S. M. R. Tousi, and S. M. Hosseinian, "Optimal siting and sizing of renewable energy sources and charging stations simultaneously based on Differential Evolution algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 1015–1024, 2015.

[17] F. Heymann, J. M. Silva, V. Miranda, and F. J. Soares, "Synergies between Electric Vehicles and Distributed Renewable Generation?," in *11th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2018)*, 2018.

[18] P. Nunes, T. Farias, and M. C. Brito, "Day charging electric vehicles with excess solar electricity for a sustainable energy system," *Energy*, vol. 80, pp. 263–274, 2015.

[19] INE, "Import of georeferenzed Census 2011 data (BGRI)," 2011. [Online]. Available: http://mapas.ine.pt/download/index2011.phtml. [Accessed: 01-Jan-2017].

[20] EDP Distribuição, "National distribution grid - High Voltage and substations," 2016. [Online]. Available: http://edp-distribuicao-rede.wntech.com/Nut/4. [Accessed: 05-Sep-2017].

[21] EDP Distribuição, "Development and investment plan for the national distribution grid 2015-2019," Lisbon, 2015.

[22] United States Census Bureau, "TIGER Products," 2018. [Online]. Available: https://www.census.gov/geo/maps-data/data/tiger.html.

[23] EUROSTAT, "Population: demography, population projections, census, asylum & migration – Overview," 2019. [Online]. Available: https://ec.europa.eu/eurostat/web/population/overview.

2 Technology adopter analysis

The recent introduction of policy support schemes, legal standards and decreasing cost structures of distributed energy resources (DER) have coincided with rising market penetration levels of these resources. While a growing body of literature mainly specializing on one technology has been established, few efforts have been directed to more holistic, multi-technology analyses and their interdependencies. In this chapter, a joint analysis of EV, PV and HVAC technologies in Portugal is provided. The analysis combined non-spatial and spatial methods to characterize current adopter groups. Outcomes are compared to previous literature findings and provide fundamental understanding of adoption drivers. These drivers are direct input to spatiotemporal adoption forecasting models introduced subsequently.

2.1 From technology adoption to technology diffusion

To understand technology diffusion processes, one must first understand adopters. This thesis addresses both adoption and diffusion phenomena. While adoption has been defined as the decision to start to use or discontinue to use an innovation by an individual or an organization, diffusion is the percolation of an innovation through the fabric of a given (social) system. Thus, adoption on individual or organization basis eventually results in system-wide diffusion of a technology, idea or behaviour [1].

The predictive modeling of technology diffusion processes requires either a detailed understanding of the forces that drive technology adoption on the individual level or a reliable high-level framework that can analyse and forecast macroscopic diffusion patterns.

While in-depth understanding of adopter processes can be used to model large-scale technology diffusion processes *bottom-up*, other macro-type diffusion forecasting models recognize and extrapolate (spatial) patterns from *top-down*. The variety of studies on energy technology adopters can therefore be divided into two types:

- Micro-level non-spatial analysis (typically bottom-up)
- Macro-level spatial analysis (typically top-down).

While the former type of analysis does rely on georeferenced adopter information (e.g. household location using Longitude/Latitude coordinates) or empirical analysis such as surveys, latter relies on the incremental changes of patterns from a macroscopic "birdseye" perspective only. This chapter presents both perspectives, confronting consolidated literature findings with outcomes retrieved for the Portuguese case study.

2.2 Non-spatial adopter characterization

Literature findings

With the rising uptake of distributed energy resources (DER), research has been directed to the characterization of adopters. In this chapter, we will present recent outcomes on the description on a subset of DER, in particular, residential EV, PV and HVAC adopters. In this work, HVAC, as a potential demand response resource, is considered a DER technology in line with [2].

Studies on the diffusion of electric vehicles have been conducted considering different geographical perimeters. While several studies have covered US [3]–[5], UK [6], [7], and Germany [8], [9], other world regions are clearly under-represented. For example, there is a general lack of studies on major Asian car manufacturing nations such as China, Japan and South Korea, at least available in English. Some studies explicitly focus on electric vehicle type (e.g. plug-in hybrid or battery EV) [4], [10], while other consider the broader range of existing technologies [7], [9], [11].

Most studies agree that early adopters of EV have above average income, are male and belong to a certain age group, mostly close to working or active driving age. Interestingly, subsets of these criteria are also repeatedly named for the adoption of domestic PV modules [12]–[15] and HVAC [16], [17]. Apart of the similarities among EV, PV and HVAC, studies about PV and HVAC technologies discovered no link to the adopter's gender.

Part of the studies contributes to a mixed picture, suggesting that socio-demographic factors might be less important [7] or that there would be no evidence on the effect of household income to EV adoption [3]. While PV and HVAC adoption is sometimes associated with house ownership and houses being free standing or semi-detached [12], [16], [17], EVs are additionally associated with family households with children or home ownership [6], [10]. Additionally, it was suggested that EV vehicle adoption of smaller vehicle sizes would correlate with higher urban built-up density [18].

Some studies link early adopters of EVs to environments characterized by high employment and population density [11], while others suggest EV would most probably be found in rural areas given that car-based trip distances are more adequate [8]. Other residential electrical appliances considered show a similar picture. It was found that early adopters of HVAC systems are expected to have similar socio-economic characteristics, given the above average material wealth expressed by ground values and their concentration in areas possessing lower population densities. In addition to that, the results suggest a positive correlation between HVAC adoption to home ownership [17]. On the other hand, there is no information available if HVAC adoption would correlate similarly to education levels or household composition (family, singles).

There is a recent study of global scope that relied on 12,000 respondents in Asia, Europe, Latin America, Middle East and North America, that, in contrast to previous findings, covered various energy- and resource-efficient technologies, such as energy and water consumption and the use of efficient technologies, food, waste and recycling preferences as well as transport choice [16]. The sample was stratified along age and other socio-demographic characteristics. Another particularity of the dataset is that it further contains beliefs and attitudes of respondents, such as the participation in non-governmental organizations or scepticism towards technological progress. However, EVs are not included.

The study confirms that above-average income, higher education households that live in their property are more likely to become early adopters of renewable energy technologies or energy efficiency programs than low-income, renting individuals.

What looks as a coherent pattern across the literature and globally in this study, however, can be further differentiated locally. Given different income-to-energy technology acquisition ratios for all countries analyzed, income was only a decisive factor in case low-medium income households have had constrained access to credits or given a generally high purchasing power in the respective country [16].

Interestingly, the same study further detects the importance of population concentration (urban, rural) for technology adoption. In line with the findings of [19], outcomes suggest that detached houses within household ownership are stronger connected to technology adoption than rented flats (potentially urban). These authors argue that technology adoption might, for many appliances, require dedicated space. This adoption discrepancy for rural and urban citizens has been detected for a range of technologies analysed in [16].

Table 2.1 synthesizes the literature findings of residential DER adopters in past studies.

Table 2.1 Literature-based description of residential EV, PV and HVAC adopters

Criteria	EV	PV	HVAC
Gender	• Male [6], [7]	-	-
Age	• Below 55 [4] • Young middle aged [3] • Modus 41-50years [9] • Aged 18 – 64 years [10]	• 80% of between 30-60 years [12] • Between 35-55 years [13]; over 55 [20]	• Negative correlation with respect to older survey respondents [16]
Household	• Households with children [10]	• Smaller households [21] • Families [20]	-
Education	• Bachelor degree or higher [3], [10]	• Above average education [15],[14], [21] [20]	-
Income	• Above average income [6], [7], [10] • >35,000$/year/HH [11]	• Above average median household income [13] [21]	• Positive correlation to higher income groups [16]
Location	• Towns with population less than 100,000 [8] Potential adopters living in urban properties [15] • Adopters living in urban or suburban areas with access to a garage, owning at least two cars [5] • Smaller vehicle sizes in higher density areas [18]	• Suburban areas negatively correlated with PV uptake, higher diffusion in city centre [13] • Polluted areas [21]	• Rural areas, as appliances might require dedicated space not available in dense-urban residencies [16], [19] • Low population density [17]
Home ownership	• Positive correlation to house ownership [6]	• Positive correlation to house ownership [12]	• Positive correlation to house ownership or negative correlation to renting [12], [38]
Environ-ment	-	• Solar irradiation [22]	-
Other	• High resident worker density [11] Full time jobs [8] • Positive environmental attitude [4] • Household with at least 2 cars [7], [9] • High average floor area [7] • Supportive social environment [10]	• High electricity consumption [21] • Negative correlation to unemployment [23] • Negative correlation to new buildings [23] • Local organizations promoting PV [24]	• Above average floor area [17] • Areas with higher home values [17] • Households with participation in non-governmental organizations or environmental initiatives [16] • Low pop. density areas [17]

Analysis of Portuguese adopters

This section contains thesis outcomes of the analysis of Portuguese EV, PV and HVAC adopters that have been presented in [25]. Here, geographical adopter locations have been spatially interfaced with census data on neighbourhood level. Given that each neighbourhood is described through more than 120 socio-demographic census variables, a detailed picture of residential EV, PV and HVAC adopters could be provided.

The above-mentioned work constructed census variable rankings which have been retrieved using the non-parametric information gain ratio (IGR) method, that roots in Information Theory. It is computed by dividing the information gained ($\boldsymbol{IG}$, in terms of Entropy reduction) through a certain variable ($\boldsymbol{X}$) in relation to its output and the intrinsic information of a split ($\boldsymbol{ISP}$) after an observation ($\boldsymbol{Y}$) [26]:

$$IGR(X,Y) = \frac{IG(X,Y)}{ISP(X,Y)} \tag{2.1}$$

A detailed description of IGR and its comparison to other approaches is presented in Chapter 3. The table below (Table 2.2) lists the Top-15 ranked socio-demographic census criteria that have been associated with the occurrence of EV, PV or HVAC technology in each census cell.

The IGR-based analysis suggests that EV adoption can be linked to the population share of educated, elderly citizen groups. In addition, the variable ranking suggests that the share of newer buildings (2006-2011) across census cells relates to EV adoption.

Besides, results suggest a strong relation of family groups with younger children to PV adoption. Furthermore, PV adoption seems to relate to low-rise buildings (“Buildings with 1 or 2 floors”). On the other hand, results suggest that HVAC adopter households relate to a certain building construction period (1981-1990). Furthermore, former adopters relate to residencies with large apartment sizes (“Class. family acc. w/ 100-200 m2“ and “Class. family acc. w/ 200 m2 +“) and parking availability, which might be explained by the general need for further space for such appliances as suggested by [16], [19].

It is important to highlight that, differently from analyses that use correlation coefficients, IGR provides exclusively insights in the strength of association (contribution to response variable) and does not allow to distinguish the positive or negative nature (sign) of such relation. Therefore, the presence of seemingly counterintuitive variables such as “Female res. aged 65 and +” and “Male residents aged 25-64” suggests that both variables relate to the presence of EV, with the former variable probably relating negatively to EV adoption while, in line with other outcomes, one might expect the latter positively relate to EV adoption.

Table 2.2 Top-15 socio-demographic criteria associated with adoption of each DER (neighborhood level)

Rank	EV	PV	HVAC
1	Buildings mainly non-residential	Res. w/ 3rd cycle of elem. educ.	Classical residences
2	Residents attending high school	Children aged 10-13	Female res. aged 65 and +
3	Class. families w. people aged 65+	Residents finished high school	Class. resid. w/ park. for 1 veh.
4	Female res. aged 65 and +	Female children aged 0-4	Class. resid. w/ park. for 3 veh.
5	Male residents aged 25-64	Male residents over 64 years old	Classic buildings w/ 1 or 2 div.
6	Buildings built 2006-2011	Classic residence for acc.	Classic buildings w/ 3 or + div.
7	Children aged 0-4	Female children aged 5-9	Buildings built 1981-1990
8	Male residents aged 20-64	Children aged 0-4	Class. family acc. with 3/4 rooms
9	Residents with university degree	Res. w/ 1st cycle of elem. educ.	Female residents
10	Residents without econ. activity	Res. working in the muni. of resid.	Classic isolated buildings
11	Residents aged 25-64	Male children aged 0-4	Class. resid. w/ park. for 2 veh.
12	Male res. aged 65 and +	Male children aged 15-19	Class. family acc. w/ 100-200 m2
13	Employed residents	Groups with 2 children, not married	Buildings with 1 or 2 floors
14	Residents aged 20-64	Groups with 1 child, not married	Male residents aged 20-24
15	Residents over 64 years old	Residents without econ. activity	Class. family acc. w/ 200 m2 +

2.3 EFFECTS OF DATA AGGREGATION

Previous studies on data aggregation effects

As many previously cited works (e.g. [6], [17], [11], [27]–[28]) rely on census data that aggregated individual observations, effects of spatial scale should be considered.

Aggregation of information to spatial units such as neighbourhood-, borough-, municipal- or even district-level has been extensively discussed in spatial science context [29]–[32]. All these works highlight that conclusions drawn from statistical analysis applied to aggregated data need to be treated carefully. The work of [31] analyses statistical dependencies using regression analysis on a georeferenced census dataset that covers the Buffalo Metropolitan area. Results show that the standard errors of parameter estimates increase as data are spatially aggregated.

Likewise, it has been observed that the goodness of fit increases with data aggregation. As mentioned by [32], a common drawback of aggregated census data might lie in the smoothing of observations, that, as a result, cannot provide sufficient variance to draw meaningful conclusions. Furthermore, multi-collinearity might rise with aggregation levels.

Such effects have been reported in [30], [32]. For example, [31] finds highly unreliable results given the strong deviations under each aggregation level. The authors find that an increase of the share of elderly in each household by 10% can result in a family income decrease estimate ranging from $308–$2,654. Likewise, results suggest accuracy of the multi-linear regression can simply be increased by aggregating data.

In [28], three statistical phenomena are discussed and synthesized that may negatively impact the reliability of drawn conclusions using spatially aggregated data.

- The ecological fallacy describes the attempt to draw conclusions from aggregated macro-level data to individual observations;
- The individualistic fallacy, on the other hand, is linked to efforts that try to infer macro-level coherences from individual observations;
- Cross-level fallacy implies drawing conclusions from one subgroup of a population to another subgroup.

However, researchers might need to rely on aggregated data. Authors of [32] present a discussion of the advantages of both aggregated and disaggregated observations for population analysis.

Aggregated information is typically easier to retrieve and requires less computational resources for analysis. Disaggregated data on the other hand might better reflect individual behaviour and should be the default option especially when policy efficacy is investigated. However, its access is limited and costly (e.g. using surveys).

As the work of [32] on the correlation of census variables to vehicle ownership suggests, using the lowest aggregation level might lead to the best model fit with the highest correlation values. That is in line with the suggestions of [31] on coping with spatial aggregated data. Herein, the authors suggest the following:

- Avoiding the use of aggregation whenever possible;
- Analyzing various aggregation levels and report results and respective deviations.

Curiously, none of the reviewed works (Table 1) that analyse EV, PV or HVAC adopters with aggregated census data explicitly mention or investigate such phenomena (e.g. [9], [12], [17], [24]).

In contrast, this work compares the retrieved results (rankings of socio-demographic census variables) across three different spatial aggregation levels – municipal-level, borough-level, neighborhood-level – to gain insight into the stability of retrieved rankings in case different census aggregation levels would be used.

Hence, above mentioned results (Top-15 census rankings) have been differentiated for all three considered spatial data aggregation levels (municipalities, boroughs (blocks of houses), neighbourhoods). Results will provide insights in the stability of such rankings across census aggregation levels.

Variable rankings considering aggregated census data

In this section, the Top-15 census variable rankings using an IGR approach are compared. While the IGR approach is presented in more detail in Chapter 3, the changes in rankings across aggregation levels itself receive undivided attention in this subchapter.

The tables below (Table 2.3 and Table 2.4) show the variable rankings retrieved using census data aggregated to borough and municipal census units. At the borough level, the presence of PV adopters is now strongly associated to census variables related to age. The presence of various similar census variables such as “Male children aged 10-13”, “Female children aged 14-19“, “Female children aged 5-9” or “Female children aged 0-4” suggests a (likely positive) link of PV adoption to families with children and adults.

Table 2.3 Top-15 socio-demographic criteria associated with adoption of each DER (borough level)

Rank	EV	PV	HVAC
1	Buildings mainly non-residential	Female residents aged 25-64	Male residents aged 20-24
2	Residents aged 65 and +	Male children aged 10-13	Class. family acc. occ. by landlord
3	Pensioners and retired residents	Female children aged 14-19	Residents aged 20-24
4	Children aged 0-4	Female children aged 5-9	Class. family acc. rented
5	Residents aged 25-64	Male residents aged 20-24	Children aged 15-19
6	Groups with children under 15	Residents aged 20-64	Employed residents
7	Employed residents	Classic families w/ 1 or 2 people	Residents attending high school
8	Resident attending university	Class. families w/ people aged 65+	Male children aged 0-4
9	Resid. w/ 2nd cycle of prim. educ.	Groups with children under 15	Residents aged 65 and +
10	Residents wo/ econ. activity	Residents unable to read or write	Residents aged 20-64
11	Residents aged 20-64	Resident employed	Groups w/ child below 6 years
12	Class. families w/ people aged 65+	Class. family acc. rented	Residents wo/ econ. activity
13	Residents w/ university degree	Residents aged 25-64	Classic buildings w/ 1 or 2 div.
14	Male children aged 5-9	Children aged 15-19	Pensioners and retired residents
15	Male res. aged 65 and +	Female children aged 0-4	Residents unable to read or write

The Top-15 variable subset associated to EV adoption shows a more mixed presence of census variables related to basic and higher education (“Resid. w/ 2nd cycle of prim. Educ”, “Residents with university degree” or

"Resident attending university"). Furthermore, the presence of children and older people (latter such as "Class. families w/ people aged 65+", "Residents aged 65 and +" or "Pensioners and retired residents") show a strong association to EV adoption.

Table 2.4 Top-15 socio-demographic criteria associated with adoption of each DER (municipal level)

Rank	EV	PV	HVAC
1	Male residents aged 25-64	Class. family acc. w/ 3/4 rooms	Class. family acc. with 3/4 rooms
2	Buildings with 5 or more floors	Buildings built between 1961-1970	Employed residents
3	Residents unable to read or write	Female residents	Residents w/ econ. activity
4	Residents w/ university degree	Male residents aged 25-64	Res. working in the muni. of resid.
5	Residents wo/ econ. activity	Male residents aged 20-64	Classic buildings built for 3+ acc.
6	Resident employed	Residents without econ. activity	Buildings w/ 5 or more floors
7	Residents aged 25-64	Class. family acc. occ. by landlord	Res. studying in the muni. of resid.
8	Female residents aged 20-64	Res. working in the tert. sector	Residents w/ university degree
9	Class. families w/ people aged 65+	Resident employed	Res. w/ post high school educ.
10	Res. working in the tert. sector	Residents aged 25-64	Pensioners and retired residents
11	Residents aged 65 and +	Female children aged 5-9	Female res. aged 65 and +
12	Pensioners and retired residents	Residents finished high school	Female residents aged 0-4
13	Male res. aged 65 and +	Class. family acc. rented	Residents unable to read or write
14	Residents aged 20-64	Class. family acc. with 1/2 rooms	Class. families w/ people aged 65+
15	Class. family acc. w/ 3/4 rooms	Residents aged 20-64	Female residents

The top-ranked variable "Buildings mainly non-residential" seems counterintuitive as residential EV adoption is concerned. However, given the absence of association sign of the IGR method, the association may be of negative nature.

Census variables related to HVAC adoption show very diverse variable ranking, mixing adults and family-with-children type variables (“Male residents aged 20-24”, “Residents aged 20-24” or “Children aged 15-19“) with building ownership variables (“Class. family acc. occ. by landlord“ or “Class. family acc. Rented”). Such findings are very in line with the literature that suggest home ownership being strongly associated with HVAC adoption.

Comparing such findings to the coarsest data aggregation level (municipal), the Top-15 subsets of census variables strongly reorder and substitute again. Eventually, one finds the Top-15 variable rankings for each technology and aggregation level rarely similar and highly unstable. Such effects will be quantified in the following section.

Ranking stability under census data aggregation

Table 2.5. shows the rank similarity or overlap of two data aggregation levels. Interestingly, highest overlap is achieved for EV, with 10 census variables listed in both Top-15 level for boroughs and neighbourhoods. Likewise, nine census variables are contained in the Top-15 rankings for municipal and borough aggregation levels. The Top-15 ranked census variables associated to PV and HVAC adoption, on the other hand, share only two variables across borough and neighbourhood aggregation level and five respectively four variables across municipal and borough census data aggregation. However, the former analysis is complementary to the high rank instability and rank substitution observed across all tables concerned (Table 2.4, 2.5 and 2.6). For all three technologies, variables show strong shuffling comparing census aggregation to municipality, borough and neighbourhood level.

Table 2.5 Top-15 rank similarity across aggregation levels

	EV	PV	HVAC
Municipalities - Boroughs	0.60	0.33	0.27
Boroughs - Neighbourhoods	0.67	0.13	0.13

In general, rank similarities seam to vary less across data aggregation levels if more observations are available. For example, for EV, there where by far more observations where available then for PV or HVAC. Therefore, EV Top-15 ranks have a higher similarity across data aggregation levels.

The are several reasons why the similarities of Top-15 census variables associated to EV, PV and HVAC adoption are of high interest:

- In case technologies are adopted by the same adopter groups (and thus, in the same locations), synergies behind the meter or on neighbourhood level might be exploited. Such parallel adoption behaviour is, for example, assumed under a microgrid paradigm.

- If the subset of technologies that may be used synergistically (e.g. EV-PV [32]) occurs in within the same location or in geographic proximity, such synergies might be exploited without crossing higher electricity network hierarchies.

- The adoption of various electrical appliances such as EV, PV and HVAC in the same or close geographical location (mixed adoption clusters) might adversely affect electricity network planning and operation.

- Under the presence of support schemes that foster technology adoption, a highly congruent characterization of EV, PV and HVAC adopters might ground on very concentrated and unequal participation in such subsidy schemes.

A comparison of the top-ranked census criteria for each technology suggests small overlap of EV, PV and HVAC adopters (Table 2.6).

Table 2.6 Top-15 rank similarity across technologies

	EV-PV	EV-HVAC	PV-HVAC
Municipalities	0.40	0.53	0.27
Boroughs	0.20	0.27	0.13
Neighbourhoods	0.20	0.07	0.13

It is noteworthy that household preferences for EV-PV are more frequent that PV-HVAC or EV-HVAC, while the former is generally assumed beneficial for synergetic use [33]. However, analysis under higher census data aggregation levels (Boroughs and Municipalities) surprisingly suggests that EV-HVAC are more common than EV-PV.

Such results, again, display the high instability aggregated census data introduces in such kind of analysis.

2.4 SPATIAL ADOPTER CHARACTERIZATION

Geostatistical analysis tools

This subchapter is dedicated to the question, if DER adoption would follow a spatially homogeneous pattern. Hence, the aim at this step is to determine the degree or absence of autocorrelation that DER adopter patterns expose, given that strong positive autocorrelation (e.g. DER adopter clusters) might potentially impact the planning of electricity networks. One of the most common metrics for spatial autocorrelation is Moran's I [34].

Moran's I reveals the tendency of polygons having similar (linear correlated) values when compared to their neighbors. It is a global autocorrelation test similar to Geary's C or the global Getis-Ord G [35]. It is a dimensionless, appealing metric, because like linear (non-spatial) regression, it produces outputs within [-1,1], where a value of 1 represents absolute spatial autocorrelation, 0 spatial randomness with no distinct pattern and -1 complete dissimilarity similar to a checkerboard pattern [36]. Latter occurs, if all spatial objects are neighbored by the most dissimilar values of the population. As major input serves a weight matrix $\boldsymbol{w_{ij}}$ that represents the distance structure of the polygon under analysis and its neighboring structure.

In the simplest case, the neighbourhood structure incorporates the degree of adjacency, taking values of "0" (is not neighbour) or "1" (is neighbour). The formula sums the differences between each polygon ($\boldsymbol{i}$) value $\boldsymbol{y_i}$ and its neighbourhood polygons' ($\boldsymbol{j}$) values $\boldsymbol{y_j}$ with respect to the global mean $\boldsymbol{\hat{y}}$ (the so-called lagged mean or spatial lag). This is then divided by the variance of each value y_i with respect to the global average $\boldsymbol{\hat{y}}$ and consecutively multiplied with the number of observations $\boldsymbol{NO}$ by the spatial weight matrix $\boldsymbol{w_{ij}}$.

Although Moran's I can also be retrieved analytically (it represents the value of the Ordinary-Least-Squared fitted slope plotting polygon values against their lagged correspondents), it is typically computed as stated below:

$$I = \frac{\mathrm{NO}}{\sum_{i=1}^{NO}\sum_{j=1}^{NO} w_{ij}} \frac{\sum_{i=1}^{NO}\sum_{j=1}^{NO} w_{ij}(y_i-\hat{y})(y_j-\hat{y})}{\sum_{i=1}^{NO}(y_i-\hat{y})^2} \tag{2.2}$$

While the output gives a first indication of the spatial autocorrelation structures, respective significance levels (p-values) can be either obtained by comparing the variances to predefined distributions or through simulation approaches, with latter having been widely advocated in [34]. The previously introduced analysis of spatial autocorrelation (Moran's I) provides insight in the global dispersion/concentration of spatial patterns.

In addition, attempts have been made to break geographical variation down to study local situations.

In this light, Luc Anselin suggested a new type of model, namely the so-called "local indicators of spatial association (LISAs)" [37]. Such should comply with two requirements:

- The LISA value of each observation should provide insights to the spatial clustering around that value
- The sum of all LISA observations should be proportional to a global metric of spatial autocorrelation (e.g. all LISA values should sum to a global autocorrelation value).

The latter requirement can be met using index decomposition techniques. In the same work [37], Anselin suggested a LISA based on the decomposition of Moran's I, to retrieve a Local Moran I. Here, the autocorrelation value associated to each observation is $\boldsymbol{I_i}$, whereas $\boldsymbol{a_i}$ are the mean-centred values and $\boldsymbol{b_j}$ are the means for all neighbour values of polygon $\boldsymbol{i}$. Thus, $\boldsymbol{I_i}$ can be retrieved following:

$$I_i = a_i \sum_j w_{ij} b_j \tag{2.3}$$

Using a permutation Monte-Carlo sampling approach as in the test-statistic approach of Eq.2.3, a significance test may be conducted using [38]:

$$z(I_i) = \frac{I_i - E[I_i]}{\sqrt{Var[I_i]}} \tag{2.4}$$

Here, values of $I_i > 0$ indicate that a cluster of similar values (higher or lower than average) is present. Likewise, values of $\boldsymbol{I_i} < 0$ indicate a combination of dissimilar values (e.g. high values surrounded by low values). In R programing language, LISAs can be computed using the "localmoran" command of the "spdep" package. This command returns the local Moran's I statistic for each polygon, the expected value $\boldsymbol{E(I_i)}$ and variance $\boldsymbol{Var(I_i)}$ under the randomization hypothesis, the test statistic (Eq. 2.4) as well as the p-value of the above statistic assuming approximate normal distribution [34].

Results for Portuguese adopters

The EV, PV and HVAC adopter coordinates have been plotted and interfaced with the Portuguese 2011 census dataset provided by the Portuguese National Institute of Statistics (INE). Using total EV, PV, HVAC quantities and total resident numbers in each municipality, adoption shares for all 279 municipalities have been obtained (Figure 2.1). Given the availability of adopter datasets, only municipalities on Continental Portugal have been considered.

A first visual inspection (Figure 2.1) suggests EV adoption clusters mainly along the shorelines and specifically around Lisbon and Porto metropolitan areas. The municipality groups with higher shares of PV and HVAC adopters on the other hand locate particularly around the southern border and towards the country's interior parts.

After the retrieval of Moran's I values using the equation provided above (Eq. 2.2), p-values are computed. In this work, the p-value (***pval***) is approximated using a simulation-based approach firstly presented in [37]. Here, the probability of obtaining Moran's I values above the observed one is calculated using the following formula:

$$pval = \frac{smi+1}{SMI+1} \tag{2.5}$$

Here, ***smi*** is the number of simulated Moran's I values above the determined one, while ***SMI*** represents the total number of simulations. While Fig. 2.1 shows the distribution of EV, PV and HVAC in Portugal, the figure below (Fig. 2.2) shows the one-sided exceedance probability distributions for obtaining values larger than the retrieved Moran's I values determined. Thus, the simulation draws a predefined number of Moran's I values that would occur if observed polygon values under spatial randomization (null hypothesis).

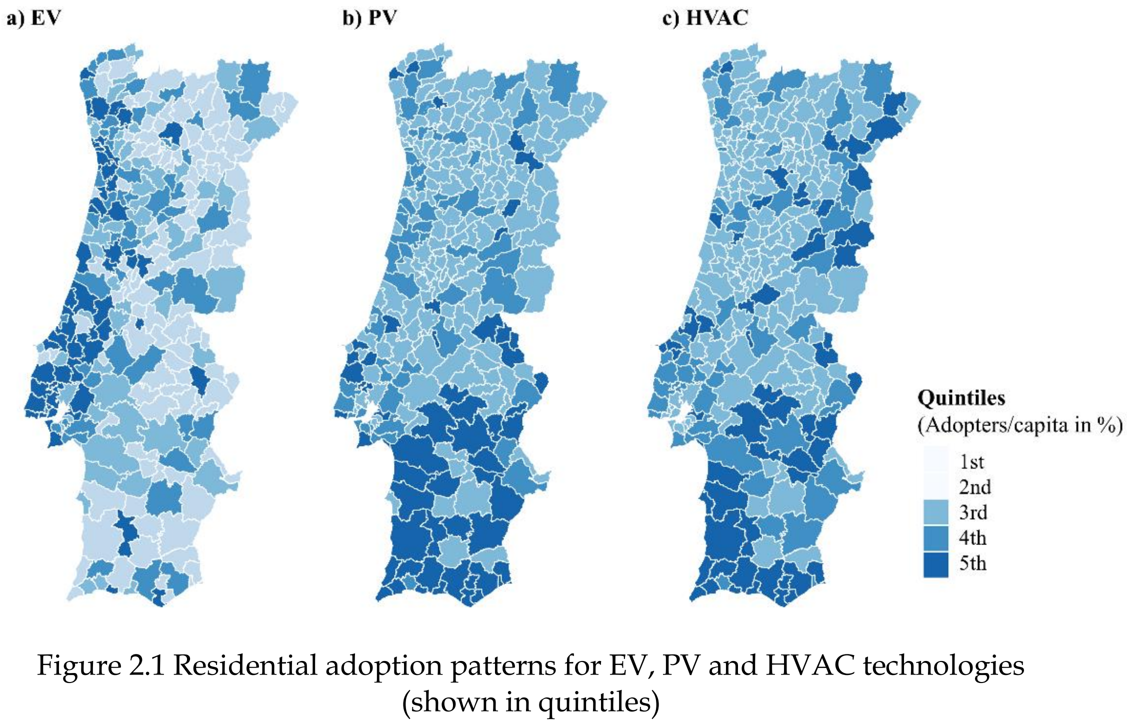


Figure 2.1 Residential adoption patterns for EV, PV and HVAC technologies (shown in quintiles)

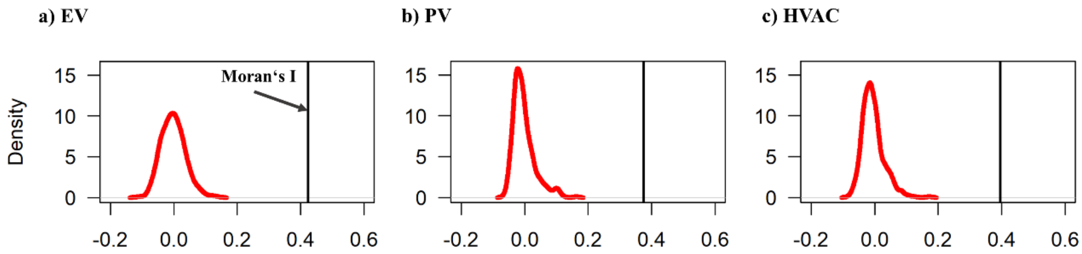


Figure 2.2 Exceedance probability density of estimated Moran's I values of EV, PV and HVAC

The distribution of the exceedance probability was generated while randomly alternating observed values among polygons during 600 permutations under randomization with equal probability and no repetition. It is important that the number of permutations is smaller than the possibilities to rearrange the polygon values in order to avoid double counting effects that may skew results eventually.

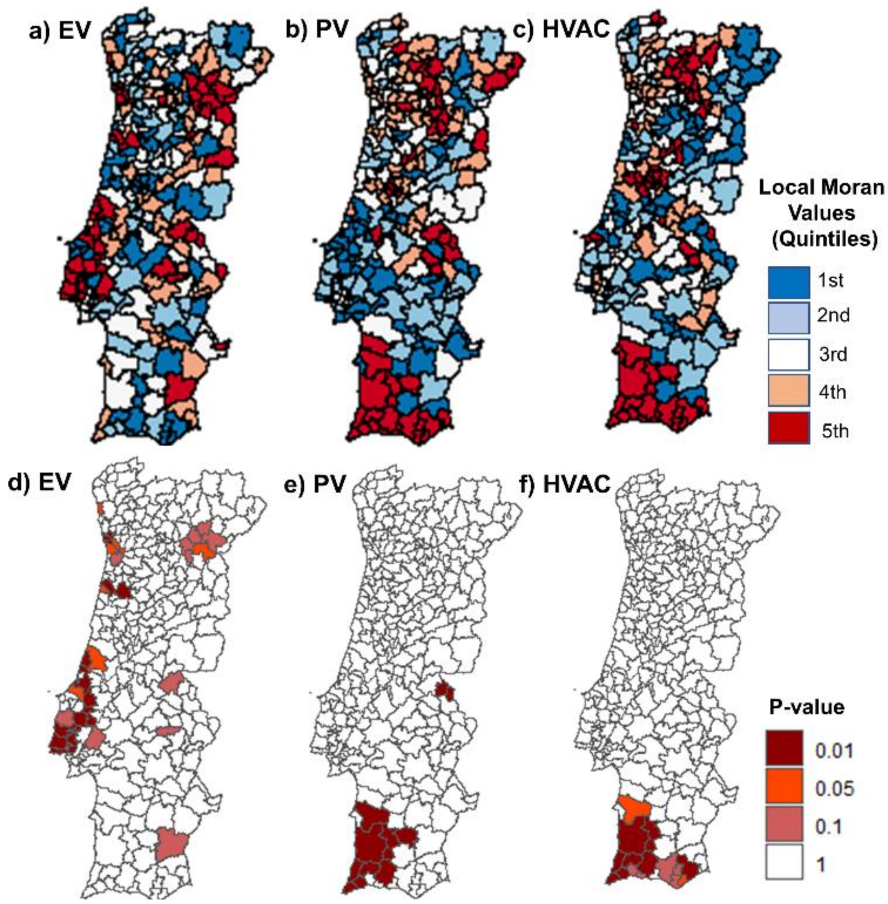


Figure 2.3 Spatial distribution of Local Moran I (a, b, c) and respective p-values (d, e, f).

Table 2.7 Moran's I results for each technology

Technology/ Value	EV	PV	HVAC
Moran's I	0.42346	0.37526	0.39532
p-value	<0.01	<0.01	<0.01

Results suggest strong evidence for the rejection of the Null hypothesis of spatial randomization with p-values smaller than 0.1% in across all technologies. The vertical bar in Fig. 2.2 indicates the realized Moran's value and the estimated Moran's I value for each technology on Municipal aggregation level.

It should be noted that Moran's I's outcome dependency on the predefined neighbourhood structure and situation of boundary polygons without a complete neighbourhood matrix has been criticized in [34]. However, authors of the same work admitted that no optimal treatment of these cases has been found so far.

Regarding local patterns of spatial association, outcomes show spatial hotspots along the Southern cost (PV, HVAC) and Western costs (EV) as well as in some isolated areas in Northern-central Portugal (EV, PV, HVAC). Furthermore, all technologies adopter distributions suggest cold spots in the Northern or Central areas of Continental Portugal (EV, PV, HVAC). Taking the test statistics analysis into account, the hotspots along the urban centres at Portugal's Western coastline (EV) and the Southern costal hotspots (PV; HVAC) suggest being significant at levels <1%.

Several techniques to extend the local autocorrelation analysis, considering common critiques on the necessary normality assumption (of $\boldsymbol{I_i}$s) and multiple hypothesis tests have been proposed. The interested reader might find an extensive overview of such extensions together with case study applications in [35].

Chapter summary and conclusions

The chapter presented mixed non-spatial and spatial analysis of residential EV, PV and HVAC adopter groups in Portugal. While outcomes provide insights in the spatial patterns current DER support schemes cause, on the one hand side, a deep understanding of adoption patterns and its drivers are fundamental to forecast future dynamics on the other hand.

As major outcome, findings suggest such adopters represent different population subgroups. This has major implications on network planning and political attempts to exploit synergies of joint technology adoption (e.g. EV-PV, PV-HVAC) behind the meter. The main conclusions of this chapter can be summarized in the following way:

- A fine delineation of consumer preferences is required in order to model technology diffusion while considering the spatial and socio-demographic structure of a study area.

- Previous studies analysed DER georeferenced adopters often with census data. Mostly, regression models were fitted to detect the "driving forces" behind technology adoption.

- Although Portuguese EV, PV and HVAC adopter characterizations show strong discrepancies, domestic adoption of appliances may coincide at the geographical location. Under current adoption, especially EV-PV may be found co-located, which may be due to the strong similarity of early adopter groups in Portugal.

- The spatial study of DER adoption patterns, e.g. using geostatistical tools (Moran's I), has so far been neglected. This work came first to compare spatial autocorrelation of DER.

- Analysing Portuguese EV, PV and HVAC adopters using spatial autocorrelation, the outcomes reveal a strong spatial clustering of the mentioned technologies.

- The comparison of three census data aggregation levels revealed a strong instability once census variables where ranked according to their association strength to DER adoption.

Chapter references

[1] K. F. Kee, “Adoption and Diffusion,” in *The International Encyclopedia of Organizational Communication*, 2017, pp. 1–15.

[2] Federal Commission Energy Regulatory, “Distributed Energy Resources - Technical Considerations for the Bulk Power System,” 2018.

[3] M. K. Hidrue, G. R. Parsons, W. Kempton, and M. P. Gardner, “Willingness to pay for electric vehicles and their attributes,” *Resour. Energy Econ.*, vol. 33, no. 3, pp. 686–705, 2011.

[4] R. Curtin, Y. Shrago, and J. Mikkelsen, “Plug in Hybrid Electric Vehicles,” University of Michigan, p. 45, 2009.

[5] Deloitte, “Gaining Traction: A customer view of electric vehicle mass adoption in the US automotive market,” *Deloitte Consult. LLP*, 2010.

[6] A. R. Campbell, T. Ryley, and R. Thring, “Identifying the early adopters of alternative fuel vehicles: A case study of Birmingham, United Kingdom,” *Transp. Res. Part A Policy Pract.*, vol. 46, no. 8, pp. 1318–1327, 2012.

[7] J. Anable, S. Skippon, G. Schuitema, and N. Kinnear, “Who will adopt electric vehicles? A sementation approach of UK consumers. Proceedings to ECEEE 2011: Summary Study,” *ECEEE 2011 SUMMER STUDY*, pp. 1015–1026, 2011.

[8] D. Biere, D. Dallinger, and M. Wietschel, “Ökonomische Analyse der Erstnutzer von Elektrofahrzeugen,” *Zeitschrift für Energiewirtschaft*, vol. 2009, no. 2, pp. 173–181, 2009.

[9] P. Plötz, U. Schneider, J. Globisch, and E. Dütschke, “Who will buy electric vehicles? Identifying early adopters in Germany,” *Transp. Res. Part A Policy Pract.*, vol. 67, pp. 96–109, 2014.

[10] J. Saarenpää, M. Kolehmainen, and H. Niska, “Geodemographic analysis and estimation of early plug-in hybrid electric vehicle adoption,” *Appl. Energy*, vol. 107, pp. 456–464, 2013.

[11] T. D. Chen, “Where are the electric vehicles? A spatial model for vehicle-choice count data,” *J. Transp. Geogr.*, vol. 43, pp. 181–188, 2015.

[12] V. Vasseur and R. Kemp, “The adoption of PV in the Netherlands: A statistical analysis of adoption factors,” *Renew. Sustain. Energy Rev.*, vol. 41, pp. 483–494, 2015.

[13] C. Lee, "Influence of local environmental , social , economic and political variables on the spatial distribution of residential solar PV arrays across the United States," *Energy Policy*, vol. 47, pp. 332–344, 2012.

[14] D. G. Labay and T. C. Kinnear, "Exploring the Consumer Decision Process in the Adoption of Solar Energy Systems," *J. Consum. Res.*, vol. 8, no. 3, pp. 271–278, 1981.

[15] T. Islam and N. Meade, "The impact of attribute preferences on adoption timing: The case of photo-voltaic (PV) solar cells for household electricity generation," *Energy Policy*, vol. 55, pp. 521–530, 2013.

[16] N. Ameli and N. Brandt, "Determinants of Households' Investment in Energy Efficiency and Renewables," *Environ. Res. Lett.* , vol.10, 044015, 2014.

[17] D. S. Noonan, L.-H. C. Hsieh, and D. Matisoff, "Spatial Effects in Energy-Efficient Residential HVAC Technology Adoption," *Environ. Behav.*, vol. 45, no. 4, pp. 476–503, 2013.

[18] C. R. Bhat, S. Sen, and N. Eluru, "The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use," *Transp. Res. Part B Methodol.*, vol. 43, no. 1, pp. 1–18, 2009.

[19] C. C. Michelsen and R. Madlener, "Homeowners ' preferences for adopting innovative residential heating systems : A discrete choice analysis for Germany," *Energy Econ.*, vol. 34, no. 5, pp. 1271–1283, 2012.

[20] J. Sommerfeld, L. Buys, K. Mengersen, and D. Vine, "Influence of demographic variables on uptake of domestic solar photovoltaic technology," *Renew. Sustain. Energy Rev.*, vol. 67, pp. 315–323, 2017.

[21] N. Balta-Ozkan, J. Yildirim, and P. M. Connor, "Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach," *Energy Econ.*, vol. 51, pp. 417–429, 2015.

[22] V. Rai and S. A. Robinson, "Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental factors," *Environ. Model. Softw.*, vol. 70, pp. 163–177, 2015.

[23] S. Dharshing, "Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany," *Energy Res. Soc. Sci.*, vol. 23, pp. 113–124, 2017.

[24] A. Palm, "Local factors driving the diffusion of solar photovoltaics in Sweden: A case study of five municipalities in an early market," *Energy Res. Soc. Sci.*, vol. 14, pp. 1–12, 2016.

[25] F. Heymann *et al.*, "DER Adopters analysis with Spatial Autocorrelation and Information Gain Ratio under different Census-data Aggregation Levels," *IET Renew. Power Gener.*, vol. 14 (1), 2019.

[26] A. G. Karegowda, A. S. Manjunath, G. Ratio, and C. F. Evaluation, "Comparative study of attribute selection using Gain Ratio and correlation based Feature Selection," *Int. J. Inf. Technol. Knowl. Manag.*, vol. 2, no. 2, pp. 271–277, 2010.

[27] C. Davidson, E. Drury, A. Lopez, R. Elmore, and R. Margolis, "Modeling photovoltaic diffusion: an analysis of geospatial datasets," *Environ. Res. Lett.*, vol. 9, no. 7, p. 074009, 2014.

[28] S. Dharshing *et al.*, "Regional distribution of photovoltaic deployment in the UK and its determinants : A spatial econometric approach," *Chem. Phys. Lett.*, vol. 67, no. 2015, pp. 113–124, 2017.

[29] W. A. V Clark and K. L. Avery, "The Effects of Data Aggregation in Statistical Analysis," *Geogr. Anal.*, vol. 8, (4), pp. 428–438, 1976.

[30] S. Openshaw, "Ecological fallacies and the analysis of areal census data," *Environ. Plan. A*, vol. 16, pp. 17–31, 1984.

[31] A. S. Fotheringham and D. W. S. Wong, "The modifiable areal unit problem in multivariate statistical analysis," *Environ. Plan. A*, vol. 23, no. 7, 1991.

[32] M. K. Adjemian and J. Williams, "Using census aggregates to proxy for household characteristics : An application to vehicle ownership University of California," *Transportation (Amst).*, vol. 36, no. 2, pp. 223–241, 2009.

[33] P. Nunes, T. Farias, and M. C. Brito, "Day charging electric vehicles with excess solar electricity for a sustainable energy system," *Energy*, vol. 80, pp. 263–274, 2015.

[34] E. Gómez-Rubio, Virgilio; Bivand, Roger; Pebesma, *Applied Spatial Data Analysis with R*, Second Edi., 2013.

[35] L. Comber and C. Brunsdon, *R for Spatial Analysis and Mapping*. SAGE Publications Ltd, 2015.

[36] M. Gimond, "Intro to GIS and Spatial Analysis [ebook]," 2019. [Online]. Available: https://mgimond.github.io/Spatial/index.html.

[37] L. Anselin, "Local indicators of spatial association - LISA," *Geogr. Anal.*, vol. 27, pp. 93–115, 1995.

[38] A. C. A. Hope, "A Simplified Monte Carlo Significance Test Procedure," *J. R. Stat. Soc. Ser. B*, vol. 30, no. 3, pp. 582–598, 1968.

3 Models for technology adoption forecasting

With the ongoing adoption of new technologies, there has been a growing interest in forecasting such dynamics in space and time. This interest has been strengthened by the need of network planners and policy makers to assess the effects of large-scale adoption of energy technologies. This chapter presents spatiotemporal DER adoption forecasting models, their components and limitations. Then, focusing on one deterministic simulation model variant developed within this thesis, a calibration considering most sensitive model parameters is performed. Finally, the presented model is compared to other models that can be found in the emerging literature on DER adoption models.

3.1 TECHNOLOGY DIFFUSION MODELING

Why we need technology diffusion modeling

The transformation towards low-carbon, decentralized power systems coincide with a strong uptake of new distributed energy resources (DER). While it has been observed that, during the initial stage, small-scale renewable energy technologies and EV adoption patterns differed both in space and time [1], [2], the previous chapter strongly suggested that DER adoption patterns are indeed autocorrelated.

Such evidence has important implications to power system planning. Given their accountability for the system security, both electricity network planners and operators as well as policy makers have a high interest in DER diffusion modelling [3]. Strong deviations of the realized DER diffusion process can lead to system disturbances and additionally increase the utility (or distribution and transmission system operators – DSO/TSO) capital and operation costs.

A recent report of the U.S. National Renewable Energy Laboratory (NREL) has been the first to provide a systematic quantification of the DER adoption forecast errors [4]. Results suggest that for an electric utility (or DSO) with 1 TWh delivered to its client-base, DER adoption forecasts can decrease the companies' capital and operational costs by up to 7 million US dollars. Such figures are assuming a 15-year planning horizon. Hence, besides the necessity of network planners to design a cost-efficient and reliable system, the abovementioned research outcomes suggest the value that improved DER adoption forecasts may provide.

In a recent study [5] on electricity network planning requirements, a list of 10 major pressing challenges to power system planning have been formulated. Such challenges included a finer spatial and temporal granularity of DER adoption patterns, DER adoption location forecasting and an improved accountancy of uncertainty sources in the planning process.

Likewise, the need for improved representations of DER adoption processes in transmission and distribution network planning has been expressed in [5]–[8].

First studies showed that the use of different DER allocation techniques in a municipal distribution network can lead to strong expansion cost over-/underestimation. In the respective case study, forecast errors could account to 4 million Euros over a 15-year horizon for a typical municipal distribution network [9]. On transmission planning level, a comparison of the DER allocation techniques resulted in DER installation forecast variations of roughly about 100 MW installations for some service areas [10].

Therefore, this chapter thoroughly addresses diffusion models that are applied to predict several small-scale energy technologies that are likely to change the spatial morphology of residential peak demand [6].

As one core innovation, this work analyses the effect of a variety of different DER allocation techniques currently used in the literature and industry. As shown below, most of these allocation techniques are rather coarse and do not include modifiable temporal or spatial differentiations. Given the previously stated need for such increase in data resolution/finer granularity, we will shortly discuss how a perfect DER diffusion model could look like and then show how a flexible spatial DER diffusion model can be built.

The ideal technology diffusion model

This subchapter provides some reflection on a potential *ideal* shape of a DER diffusion forecasting model to be developed, stretching both on spatial and temporal dimension as well as desirable input/output formats.

- **Input data**: Spatiotemporal diffusion models rely on spatial and temporal data. Spatiotemporal data consists of so-called spatial objects. Such objects can be uniquely determined (and thus, separated) by the following five dimensions: Space, time, scale, attributes and relationships [11].

  DER diffusion modelling under ideal conditions would require the complete description and full availability of the collection of spatial objects that represent the study area. Recognizing that diffusion occurs in social systems, an accurate and detailed description of the respective social system under analysis grounded on individual or household level would be required.

  Furthermore, under ideal conditions, a flexible model would adjust to likely situations of missing data as well as dynamically incorporate changing system variables (e.g. population structure, large-scale trends in technology preference or cost structures).

- **Spatial scale**: Considering DER diffusion forecasting for the residential sector, the targeted spatial resolution is the household or consumer level. Under an optimal modelling framework, this requires locational information of households/individuals that could turn DER adopters. Alongside, the availability of information that allows to infer their adoption preference is fundamental. Eventually, spatial predictions on the future distributions of adopters and adopter preferences under demographic changes or other societal trends would be required.

- **Temporal scale**: Depending on the application domain (network planning, policy analysis, economic analysis), DER diffusion forecasting time steps of quarters or years ahead are common practice [12]–[14]. Under ideal conditions, one could envision time horizons of months or weeks. However, under current conditions, there lies little use in such improvements.

- **Output and validation**: Results of an ideal DER forecasting model would be easily interpretable and ready to interface with network planning or other, e.g. market modelling, tools. Taking multi-temporal DER adoption observations, such model would be calibrated accounting for spatiotemporal covariance structures, even considering non-stationary conditions as outlined in [15].

This sketched, ideal model has not been developed within this thesis. From the current standpoint of model developments [3], it is questionable if such model could be implemented within the next decade due to computational burden and data privacy.

Moreover, it is not clear if an ideal model would be desirable, and, more importantly, how to prove a given model being ideal. As stated in [16], models cannot be proven to be true. This is because models are closed systems (real systems are open), data inputs are typically already contaminated with inference and assumptions and many different models can produce the same results (non-uniqueness/under-determination).

Today, DER diffusion models are far from the Ideal, mainly due to the following reasons:

1) DER adoption observations with multiple time steps have not been widely accessible. That limits a potential, step-wise validation of the technology diffusion model using observations stretching over a longer time period.

2) Official population information and census data on individual level are not publicly available due to consumer data protection and privacy concerns. That renders a multi-year observation of structural population changes intractable.

3) Due to the described lack of multi-temporal adopter observations, it is extremely difficult to model adoption-interdependency structures among neighbouring observations or considering several technology diffusion processes.

The availability of adopter observations for multiple time-steps represents a fundamental requirement for assuring model quality. Especially in the starting phase of adoption processes, adopter numbers are minimal and observations scarce. Paradoxically, large-scale technology diffusion eventually brings substantial adopter observations with different time steps. Likewise, rising availability of open source data allows for data-driven model building and allows to eventually improve the representation of technology diffusion processes.

However, with increasing maturity of the diffusion process, the knowledge on adopters is likely to have consolidated and interest from policy maker or network planner's perspective would fade away.

Still, the reader should be reminded with the words of Box: "Models, of course, are never true, but fortunately it is only necessary that they are useful" [17]. Therefore, the use of DER diffusion models can be found in their application as "heuristic tools for learning about the world and generating (and testing) interesting hypotheses" [18]. While certain abstractions both on spatial and temporal dimensions had to be made during the model development process, the resulting model still represents a strong improvement in granularity to current DER representations used in industry and academia practices [9].

3.2 Diffusion model components and key parameters

Technology diffusion theory introduces a set of basic assumptions that have been formulated in the major work on diffusion research by Everett Rogers, first published in 1962. Such assumptions have been extracted from Roger's latest version of the Diffusion of Innovations [19] and are listed below:

- Diffusion processes are non-repetitive and therefore unique for each innovation and social system considered;

- Diffusion processes are influenced both by the time and social network structure that accommodate the diffusion process;

- Only complete diffusion processes are considered. Incomplete or backward diffusion processes are therefore neglected. In fact, backward diffusion processes could be modelled as diffusion of the negative practice.

Therefore, any technology diffusion model should consider the previously stated assumptions, translating such into model type, parameters choice, that is, the overall model architecture. For example, the technology diffusion model to be developed should be flexible and transferable to other social systems, building on datasets that are typically largely available for various social systems. Hence, dataset types that can be exploited to characterize the social system under analysis are required.

Furthermore, given the different diffusion speeds of technology adoption (e.g. such as described for EV in [20]), the diffusion model to be developed requires an adjustable temporal domain.

The presented model is a spatial simulation model, meaning that "models that represent the change in spatial patterns through time" [18]. It is a simulation model since the underlying relationships are, through equations and iterations, embedded into a computational framework.

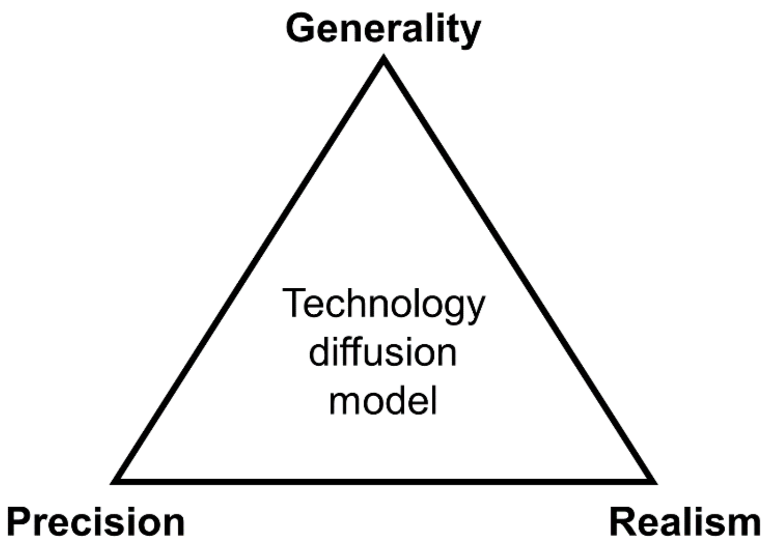


Figure 3.1 Trade-off of diffusion simulation model building (inspired by [18])

Simulation models are used to mimic changes in space through time. A major difficulty is to integrate different process rates in spatial simulation models. While limitations under asynchronous updating rules might be light, major difficulties arise for simultaneous updates under different process rates. Another way to differentiate updating processes is with discrete time steps or following event-driven timelines [18].

The proposed model developed within the frame of this thesis consists of three large major modules that work independently but constantly update and exchange information. All modules are embedded in one programming language (R-Studio). The functioning of the three modules are briefly described below, while their construction is explained in dedicated chapters hereafter:

- **Global DER forecast module**: This module computes global DER diffusion forecast curves for each technology under analysis. Model calibration grounds on historical DER growth or stock rates;

- **Cellular adopter module**: Once DER adopter quantities have left the global DER forecast module, they enter the cellular adopter module. Here, through a predefined definition of cell states (or adoption stages), and, assuming an adoption sequence along adopter's innovativeness scores, the globally forecasted DER quantities are translated into spatial patterns;

- **Adoption pattern mapping**: Eventually, adoption patterns are mapped for each time step using Geographic Information Systems (GIS), that serve as a visual interface and provide ready-available map stacks for decision makers and network planners.

A well-described drawback on the way to link pattern and process is the fact that the same process might produce different patterns or different processes might result in the same pattern [18]. Furthermore, the importance of the chosen scale of space and time needs to be highlighted, as this choice has a strong impact on the pattern and process analysis [18], [21], [22].

However, it is understood that the building process of models does require a constant trade-off between generality, precision and realism [18] (compare Fig. 3.1). While the effects of spatial data aggregation have been presented in the former chapter, a latter chapter presents an investigation on the effects of temporal discretization to model outputs.

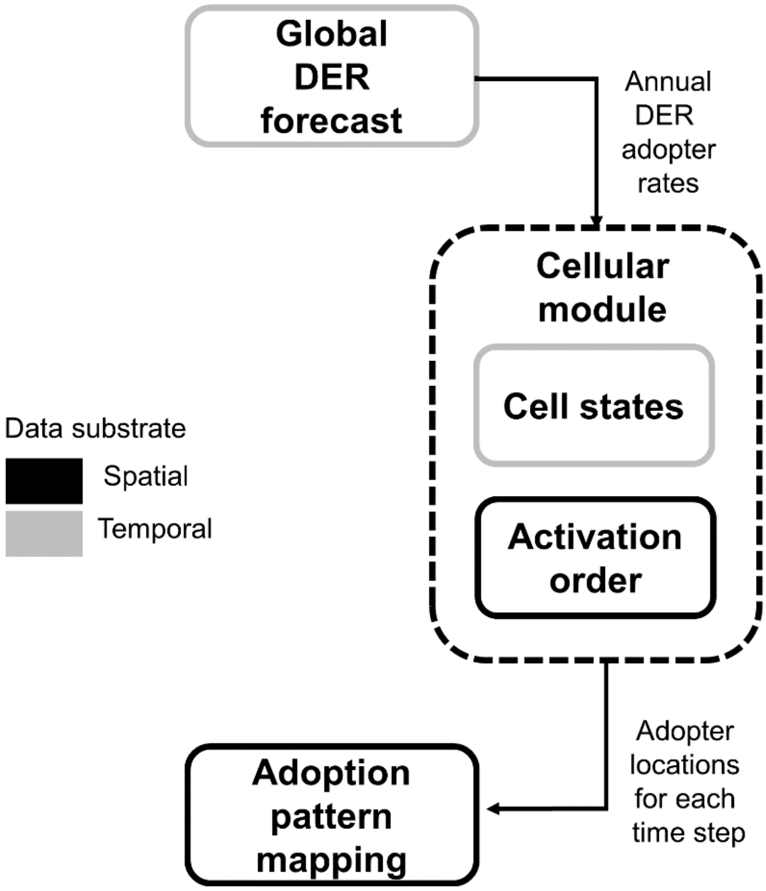


Figure 3.2 Components of the proposed spatiotemporal DER diffusion model.

3.3 TEMPORAL: BUILDING GLOBAL TECHNOLOGY DIFFUSION FORECASTS

Temporal adoption forecasts can be derived with the well-established Bass model. The Bass model was proposed in 1969 and represents a mathematical diffusion model similar to an S-curve model used for load growth studies [23]. The model originates in the field of marketing [24] and has been applied across a variety of disciplines, such in the fields of medicine and energy engineering [25]–[27].

It is a widely used, appealing model due to:

- A relatively low model complexity;
- The provision of product/appliance adoption curves without costly and time-consuming empirical surveys;
- Its descriptive results that are comprehensive and intuitive rather than relying on a deep analysis of the underlying processes [20].

The model relies on a few parameters that are determined beforehand. As input, it uses an estimation of the total market potential (***M***) of the analyzed technology. It further requires two coefficients ($\boldsymbol{p}$, $\boldsymbol{q}$) that are commonly described as innovation and imitation coefficients [20].

The number of first time purchases $\boldsymbol{n_t}$ can be expressed by the following formula [20]:

$$n_t = \frac{dAP_t}{dt} = p(M - AP_t) + q\frac{1}{M}AP_t(M - AP_t) \qquad (3.1)$$

Here, $\boldsymbol{AP_t}$ represents the accumulation of products (e.g. DER technologies) adopted until period $\boldsymbol{t}$. Given the technologies' respective $\boldsymbol{p}$, $\boldsymbol{q}$ values together with the estimated final adopter market $\boldsymbol{M}$, the total cumulative fraction of the analysed product that is adopted at time $\boldsymbol{t}$ can be estimated using the integration of the density function of Eq. 3.1 [20]:

$$N_t = F(t) \times M = \left(\frac{1 - e^{-(p+q)t}}{1 + \frac{p}{q}e^{-(p+q)t}}\right) \times M \qquad (3.2)$$

Here, the two coefficients ($\boldsymbol{p}$, $\boldsymbol{q}$) and the estimated market potential (***M***) eventually determine the evolution of the diffusion rate, and thus, the shape of the aggregated adoption curve. The current literature provides a wide array of adoption scenarios with diverse sets of $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{M}$ value tuples across markets and technologies (including EV and PV that will be analysed in subsequent thesis chapters) [27][20].

The work of [20] has demonstrated the sensitivity of model parameter choice ($\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{M}$) to achieved results, showing that different model parameter sets might result in strong forecast deviations. For example, [20] recorded that variations of the $\boldsymbol{p}$, $\boldsymbol{q}$ parameters could result in 10-100 fold deviation of technology quantities forecasting under a given time horizon.

Therefore, model parameters should be chosen carefully and the impact of potential estimate errors (e.g. over/under-estimated market potential ***M***) thoroughly assessed [20].

3.4 SPATIOTEMPORAL: DEVELOPMENT OF THE CELLULAR MODEL

The proposed model consists of a spatial module. It runs a discrete-state, deterministic simulation, where technology adopters from a forecasted global stock are annually allocated into census cells (consumers). That global stock is direct output of the previously introduced Bass DER adoption forecast for each year under analysis.

As a discrete, spatial model, the modelled diffusion process is broken down into spatial census cells with a determined cell size and cell state at a given time step. Cell states are built discrete, considering 4-year intervals (or one-year, two-year and 20-year periods in the model calibration study later in this chapter). That way, each cell passes through a pre-defined development pathway, that is expressed through the share of residents adopting a given technology for each time step. Given the previously introduced assumptions of diffusion processes hold, only forward adoption behaviour to a maximum adoption state (100% of residents/ potential adopters) is considered.

Currently, the representation of the temporal adoption behavior is constrained for adoption states for the sake of a reduced computational burden. In future simulation embedding an increased availability of computational power, the presented workflow can easily be broken down into even smaller time periods (years, quarter years or months). The census cells consecutively fill up (consumers adopt DER) until reaching a saturation point - the maximum potential of DER that can be adopted in each census cell. As for many other diffusion processes, such behavior can be modelled using S-curves – an approach also used for load growth studies [23].

Neighbour or peer effects are important factors that can accelerate the uptake of technologies. Studies for PV have shown that residencies in the direct vicinity (e.g. 100m) of PV installations have a positive correlation to adopt such technologies next [13], [28]. On the other hand, absence or even negative influence of energy technology adopters outside this 100m radius has been documented [13]. Therefore, the states presented as part of diffusion model can be also understood as an implicit form to consider the positive effect of neighbour influence. Through the discretization of adoption behaviour, each census cell increases its adoption share from state to state and, therefore, automatically produces adopter clusters within each cell.

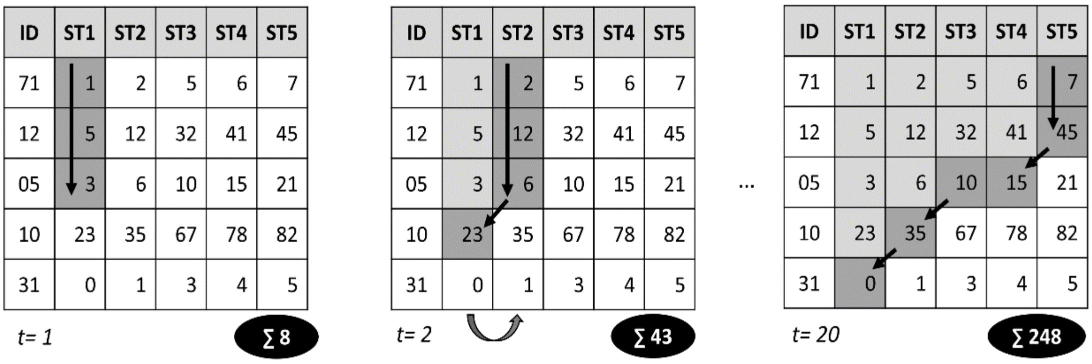


Figure 3.3 Spatial cellular adoption module for 5 adoption states

The adoption module relies on the following inputs: It first uses the innovativeness scores and global, annual DER forecasts (***GF**$_{EV}$*, ***GF**$_{PV}$*) of the Bass model detailed in the previous subchapter. It passes the following three steps:

- As first step, all census cells are ranked according to their innovativeness scores (***IS***) under each policy design considered. As shown in Figure 3.3., each of those predefined policies see census cells in a predefined order that is established through the innovativeness scores. The number in each cell correspond to the resource quantity adopted at each state. Throughout the allocation of resources within a chosen energy policy design, the order remains unchanged.

- Secondly, given an annual, global adoption forecast, cells are filled always starting from the highest to lowest ranked cell. That is, cell states (States 1-5) are actualized (+1) and respective adopters per each technology determined.

    - If global forecasts allow to change all census cells to proceed one stage, the allocation jumps to the top ranked cell and re-starts until depleting the EV/PV stock of a year.

    - The DER adopters are allocated each year along the 20-year time horizon. Once a census cell reached its final stage (State 5), it is excluded from further allocation.

- Eventually, DER adopter numbers are aggregated to the electricity network service area considered using spatial aggregation tools.

A scheme of the change in adoption states of each cell along the years (20) is shown in Figure 3.3. The 20-year timespan of this work was chosen to be in line with a typical distribution planning horizon [23].

However, at this development stage of the model, enhanced neighbourhood interaction beyond the above-mentioned mechanism (e.g. between the census cells) is not considered. This is due to the fact that there are currently mixed results on peer influence over larger distances (>100m) in [13], [29], and the scope of this study, which is the analysis of policy frameworks and not optimizing DER adoption forecasts.

Furthermore, there is increased difficulty in modeling peer-interaction through real social networks. As observed by [30], contagious spread through social networks in socio-demographic context does not necessarily imply that diffusion does occur in the agent's direct neighbourhood. However, modelling realistic social networks is an arduous task that requires extensive data collection for each study area as well as additional theoretical advancements that lie outside the scope of this thesis.

We expect such extension might be valuable once diffusion models are applied using data of finer granularity just like disaggregated consumer locations and their connection to distribution feeders. Once such level of granularity and multi-temporal adoption observations are available, peer-to-peer interaction modelling will become implementable. That way, works can further validate the results of [13], [28], [31] that quantified the influence on households adoption behaviour of neighbours, however, on much more granular scale than pursued in this work.

3.5 SPATIAL: ESTABLISHING ADOPTION PREFERENCE MAPS

The idea that technologies are differently embedded in societies as a result of intrinsic differences in their social structure and the spatial distribution thereof has drawn increasing attention on technology diffusion research [19].

In this context, the theory of *Diffusion of Innovations* by Everett Rogers ranks among the most prominent works that provide a rigorous and extensive, theoretical backbone to the analysis of technology diffusion processes. Rogers defines diffusion as "the process in which an innovation is communicated through certain channels over time among the members of a social system" [19].

As innovation, he describes an item such as an idea, project or practice, that is perceived as new by a potential adopter such as an individual or household. According to this definition, it becomes clear that while electric vehicles is generally assumed as an innovative technology, although the technology itself has been used for more than 100 years [32], if the social system perceives and reacts towards a given technology as something new, innovation theory can be applied.

The social structure is expected to necessarily affect the diffusion process as the diffusion of innovations takes place in the social system. Rogers characterized adopters into five groups.

1) Innovators;
2) Early adopters;
3) Early majority;
4) Late majority;
5) Laggards.

Innovativeness is used for categorization in diffusion research. Empirical results suggest that each of the following adopter categories are characterized by a strong homophily inside each group. Homophily is the degree to which individuals share certain social, demographic or economic characteristics [19]. In Roger's work, adopter categories are delineated by their "innovativeness", a continuous variable that aims to reflect the individual's characteristic to adopt an innovation earlier than its peers inside a social system [19].

As shown in Figure 3.4.a, Rogers defined each adopter group as a distinct area under the adoption curve. As mentioned above, adopters have been differentiated relying on the innovativeness concept that allows to grouped individuals/households into groups given a predefined innovativeness score interval.

In Figure 3.4.b, the translation of such division into the adoption process is shown. Here, the position of adopter groups along the cummulative adoption curve suggests that Innovators and Early adopters are the first adopting a given innovation. Then, once a critical mass has reached, the uptake accelerates during the diffusion through Early and Late majority groups.

Another pillar of Rogers theory states that adoption is a dynamic process, thus its rates possess temporal variability. Incomplete or non-adoption is generally not included in Diffusion theory [19]. Although Rogers does not provide a mathematical framework to predict the diffusion speed and adoption rates of innovations which would be applicable to the modelling of energy technologies, his framework of ranking consumer groups along innovativeness can serve as an allocation formula for a given innovation that spreads inside a population structure [19].

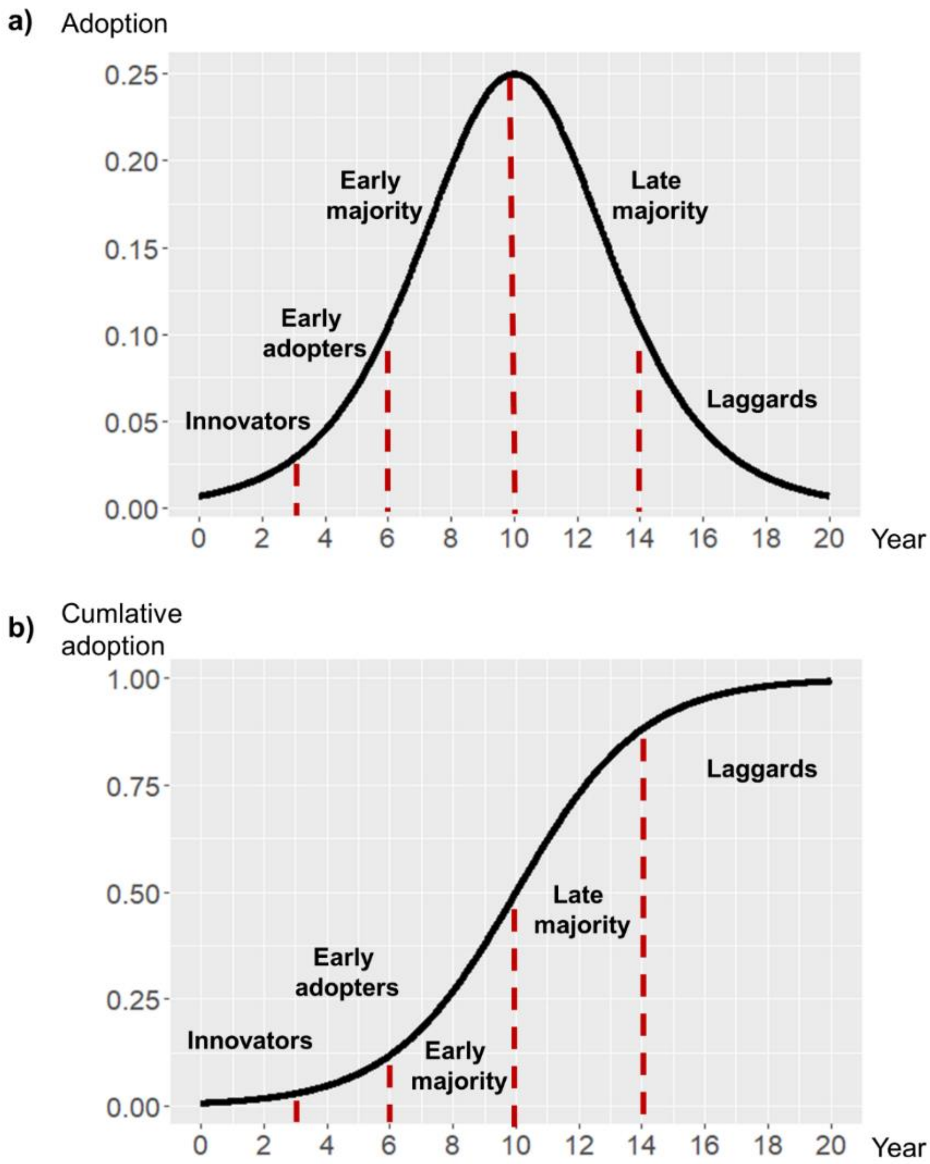


Figure 3.4 Adoption dynamics according to Rogers.

Therefore, in the context of DER diffusion forecasting, Roger's theory itself is not directly applicable. It rather can serve to discriminate neighbourhoods or census tracts along a virtual innovativeness score that is used to establish the adoption order of all census cells.

That given, a global forecast can be allocated over a district/city or quarter while identifying regions of preference (early adopters) and inertia (late adopters, laggards). Here, innovativeness is a key concept that originates in the field of marketing. The basic idea is that individuals or groups of individuals can be ranked according to their preference to adopt an innovation earlier than their peers inside a social system [19].

The construction of cell-wise innovativeness scores is therefore the third step in development of the proposed spatiotemporal DER diffusion forecast model. To build innovativeness scores for each spatial census unit, we used a subset of the socio-demographic census variables that reportedly have an established causal link to DER adoption.

It is important to highlight two characteristics that such cell-wise adoption modelling brings. On the one hand, it should be noted that cell states changes can be updated sequentially or simultaneously, with different results in case of cell state interdependence [18]. The presented model implemented a sequential state actualization. However, the thorough testing of both forms and the impact on their results lie outside the scope of this thesis.

Another important issue is the boundary determination of the technology that is analysed. This point is related to the definition where one innovation/technology starts and ends. For example, in PV technology, typical multi-crystalline modules have lately been joined by products that exploit thin film technology [33]. Although the same technological PV backbone is concerned, the two products might attract different adopter groups and it could be argued that they should be studied separately.

Rogers notes that "a technology cluster consists of one or more distinguishable elements of a technology that are perceived as being closely interrelated" [19]. Thus, the study of each innovation independently of others may be misleading.

We model the positive response of a certain adopter group to a determinate energy policy design. That is, we relate a predefined, numeric adoption factor to each socio-demographic variable that was associated to an adoption of DER under each energy policy design. Modelling the innovativeness scores assumes that a certain population subgroup, based on its socio-demographic characteristics, has a specific, predefined, likelihood to adopt a certain technology.

Thus, sequential modelling of the adoption process follows a black-box approach, where it is assumed that certain socio-demographic census variables would trigger DER adoption stronger than the remaining variables. Given the census dataset with about 120 socio-demographic criteria ($\boldsymbol{c}$) and over 17,000 spatial census tracts ($\boldsymbol{r}$=17,337), a vector with innovativeness scores (***IS***) for each technology and incentive design was retrieved following:

$$IS = \begin{pmatrix} x_{1,1} & \cdots & x_{1,c} \\ \vdots & \ddots & \vdots \\ x_{r,1} & \cdots & x_{r,c} \end{pmatrix} \times \begin{pmatrix} af_1 \\ af_2 \\ \cdots \\ af_c \end{pmatrix} \tag{3.3}$$

Here, the census dataset, consisting of a matrix of ***c*** columns which represent a number of census criteria and ***r*** rows, that correspond to spatial cells/neigborhoods, is multiplied with a predefined vector of adoption-influence (***af***). An earlier implementation of innovativeness scores in a similar use case based on census data can be found in [34]. The next subchapter will shed some light on several ways to construct such adoption-influence vectors, applying mathematical parametric and non-parametric techniques to a set of DER adopter observations.

3.5. Calibration of temporal and spatial model parameters

Global technology adoption forecast scenarios

For the development of global DER forecasts, we use technology adoption forecasts contained in [35]. The report presents three storylines: "Sustainable Transition", "Distributed Generation" and "Global Climate Action", that are considered within this thesis. The three storylines resprent three distinct potential futures of European power systems which have a high internal consistency. It is an appealing set of scenarios to use due to the strong involvement of policy makers, industry experts and external stakeholders during the scenario development process.

Table 3.1 Storylines for Portugal towards 2035 used within this thesis (modified after [35], [36]).

Storyline/ Technology	Sustainable Transition	Distributed Generation	Global Climate Action
EV (adopters)	999,917	114,324	688,270
PV (in MW)	10,452	3,060	9,824

The DER forecasts within these reports are direct input to the DER diffusion model, that relies on two DER estimations: the global adoption forecasts for every year and the estimation of the potential amount of DER that could be adopted by all spatial census cells (the theoretical upper adoption boundary of a given DER technology in a system). In general, such theoretical upper boundary can be computed using:

$$TDER = TPP \times \ pcder \tag{3.4}$$

Here, ***TDER*** represents the maximum potential of DER to be accomodated within all spatial census cells, whereas ***TPP*** and ***pcder*** are the total aggregated population contained in all census cells and the allowable per-capita share of DER respecively.

Given the various technologies that have been considered as DER as well as exponentially increasing interactions, a pre-selection has been made. In [37], DER have been defined as distributed, small-scale resources located within the electricity distribution system. Such definition embraced a variety of technologies, such as distributed generation (PV, CHP), energy storage (e.g. batteries) or energy efficiency/ controllable loads (e.g. EV, HVAC).

As the following part of this thesis involves increasing combinatorial analysis of DER adoption patterns, we will focus on two DER technologies that have received most attention. Such technologies under analysis have been EV and PV modules. This is in line with other studies that have been dedicated to analyse strong potential synergies of EV and PV technologies, especially in residential environments [38]–[41]. However, it should be noted that all modules presented in this thesis are flexible in the sense that every given technology can be analysed, given that spatial adopter locations and census data is provided.

Looking now at the case of the selected two technologies, the maximum EV and PV potential (***TEV***, ***TPV***) can be calculated as following:

$$TEV = TPP \times pccs \tag{3.5}$$

$$TPV = TPP \times \ pcpv \tag{3.6}$$

Here, ***TPP*** is the total population of continental Portugal where ***pccs*** represents the current car-share ratio (approximately 0.45) that has been derived using the ratio of light private passenger cars to the overall Portuguese population [42], [43] (assuming perfect substitution). Likewise, the total PV potential is derived using the per-capita PV share ***pcpv***. As this report aggregates both large-scale and dispersed PV installations in residencies, the forecast was corrected using the current ratio of dispersed PV (0.5) to overall PV installations in Portugal as stated in [44]. The calculation of the total EV and PV potentials at all residencies were based on [42] and [9].

For the calculation of the roof-top-based PV potential at consumers residencies, the total roof-top area (***r.area***) multiplies the average per-capita rooftop area with the total resident number ***TPP***. Multi-cristalline PV cell technology with 200 W per panel and a 1.5 m² cell size is assumed. Together with an estimated usable roof fraction (***u.fract***) of 0.3 similar to [45], an 80% performance ratio (***pr***) and a per capita rooftop area of 13 m² [46], we retrieve a potential per capita PV capacity of 0.4 kW. Other conversion technologies (mono-christalline, thin-layer) or installations (ground-based, building envolope integrated) have been neglected. However, given the structure of the established methodology, such technologies may be integrated in future extensions.

The respective ***pcpv*** and ***pccs*** values have been established as the following:

$$pccs = 0.45 \tag{3.7}$$

$$pcpv = \frac{r.area \times u.fract}{cell.size \times 5} \times pr \tag{3.8}$$

Here, the per-capita PV potential included the previously estimated total rooftop area (***r.area***) at all Portuguese residencies multiplied with the usable fraction (***u.fract***). Their product has been divided by a typical cell size of 1.5 m² and the conversion factor from a 200 W cell to 1 kW (5). The resulting theoretical EV and PV adoption potentials on Continental Portugal are shown below (Table 3.2.).

Table 3.2 Total theoretical adoption potential for EV and PV on Continental Portugal

Theoretical potential/ Quantity	TPP (in residents)	TEV (in adopters)	TPV (in kW)
Value	10,047,621	4,522,000	5,023,811

Obviously, technical parameters had been fixed before for model simulation runs, while sensitivities were eventually assessed for alterations of the DER technical parameters.

Likewise, technical characteristics such as technology sizes, efficiencies and cost structures will evolve through time. Therefore, the presented model integrates an interface to conveniently replace such values for future model extensions or actualizations.

Choosing the spatial input data substrate

The spatiotemporal diffusion model relies on accessible spatial data. Spatial data are data that have a stated relation to space. Shi suggested that over 70% of real-world phenomena can be related to and described with spatial data [47].

However, compared to data that lack spatial reference, spatial data are inherently different from any other data given that the typical independence assumption for observations does not hold as well as patterns of spatial heterogeneity [48].

Spatial data consist of spatial objects. Any of such spatial objects can be uniquely determined (and thus, separated) by the following five dimensions: 1) Space, 2) Time, 3) Scale, 4) Attribute and 5) Relationships [11]. The figure below shows the division between spatial and non-spatial data together with typical data attributes used in spatial science [49].

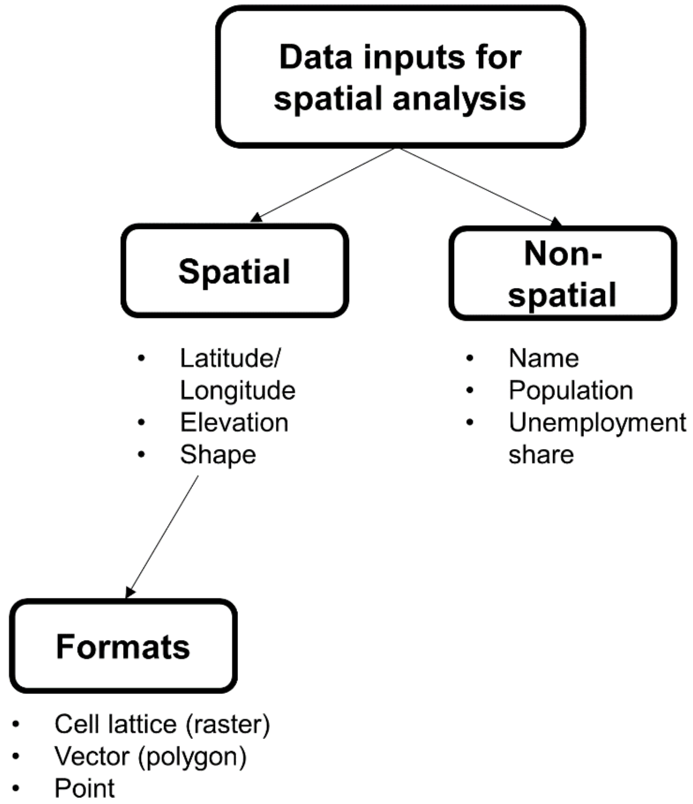


Figure 3.5 Spatial data and spatial analysis (inspired by [49]).

While common spatial data formats are spatial points, vectors (lines or polygons) and raster (cell lattices), latter two are most commonly used in spatial simulation models (e.g. in [9], [12], [13], [50]). Spatial point data are commonly neglected due to the computational burden of distance pattern analysis such data type implies. The number of computations increase with n(n-1)/2 for a given number of objects (n). For example, equal distance calculations between 100 spatial point objects require roughly 5,000 calculations, whereas 1,000 objects imply approximately 500,000 calculations [18]. The following table sums up the major characteristics of spatial cell lattice and vector formats, building on findings described in [18].

Table 3.3 Advantages and drawbacks of spatial vector and raster data formats

	Vector format (spatial polygons)	Cell lattice format (raster cells)
Advantages	• Polygon delineations typically match with census or other political, administrative or socio-demographic entities. • Mostly, lossless data handling as no data transformation required. • Realistic neighbourhood structure.	• Fast processing of equal-sized cells. • Wide arrange of applications existing (e.g. Cellular Automata). • Once transformed into a global lattice model, it is straightforward to fusion data from various sources.
Drawbacks	• Slower processing if compared to cell lattice. • Difficulties to match or merge spatial data of different polygon extents. • Difficulties to incorporate distance-relations that possess isotropic character due to typically heterogeneous polygon sizes.	• Loss of information due to conversion of inhomogeneous spatial input data. • Difficulties to incorporate distance-relations that possess anisotropic character, that is directed, non-homogeneous effects in space. • Mostly deterministic applications. Probabilistic extensions produce highly variable patches.

Discretization of the adoption processes

The adoption potential of each census cell is discretized into adoption stages, in order to allow for the dynamic modeling of technology adoption. In our work, we define 4 variations of discrete state configurations.

Such discretizations are achieved through multiplying the theoretical adopter potential with respective adoption share values per cell. Discretization states have been generated extracting the adoption shares per year given using an S-curve model. Adopter numbers per cell are rounded to integer. The four modifications considered four different state intervals (20-year interval, 4-year interval, 2-year interval and 1-year interval). Such, with respect to adoption shares, are stated below:

$$ST(20) = \{100\%\} \tag{3.9}$$

$$ST(4) = \{5\%,\ 27\%, 73\%, 95\%, 100\%\} \tag{3.10}$$

$$ST(2) = \{\ 2\%, 5\%, 12\%, 27\%, 50\%, 73\%, 88\%, 95\%, 98\%,\ 100\%\} \tag{3.11}$$

$$ST(1) = \begin{Bmatrix} 1\%, 2\%, 3\%, 5\%, 8\%, 12\%, 18\%, 27\%, 38\%, 50\%, 62\%, \\ 73\%, 82\%, 88\%, 92\%, 95\%, 97\%, 98\%, 99\%, 100\% \end{Bmatrix} \tag{3.12}$$

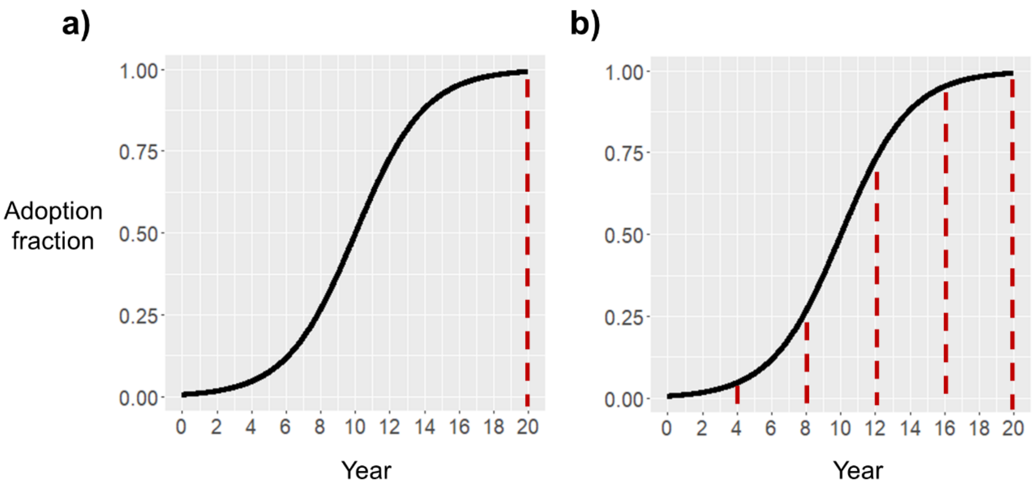


Figure 3.6 Schematic discretization of DER adoption behavior in the cellular model to 20-year (a) and 5-year time steps (b).

Fig. 3.6. shows the diffusion process cut in different adoption states (red lines) considering two of four exemplary temporal discretization forms that have been considered in this thesis. While (a) shows a discretization equivalent to a binary adoption behaviour model (not adopt/ full adopt), (b) displays discretization with finer temporal granularity using 4-year intervals.

Thus, a) and b) result in one and five adoption stages that each census cell can achieve, respectively. Temporal resolution has consecutively become finer to 2-year and 1-year time-steps. The discretization to 20 one-year and 10 two-year steps is shown below (Figure 3.7).

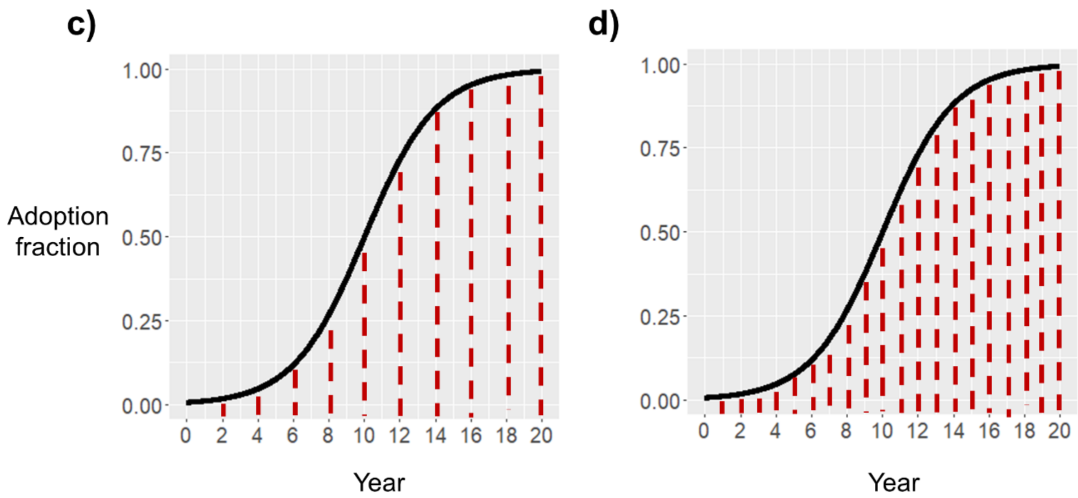


Figure 3.7 Schematic discretization of DER adoption behavior in the cellular model to 2-year (c) and 1-year time steps (d).

Establishing the adoption order

The proposed spatiotemporal model relies on cell rankings that are built using innovativeness scores. Such scores result from variable rankings, that is, vectors that weight the census variables of each census cell. In this work, such weight vectors are retrieved using outputs of variable ranking methodologies.

Various methodologies have been proposed to do variable ranking and variables selection, especially in machine learning [51]–[53]. Although the authors of [51] highlight the difference between variables (original) and feature (constructed from variables), most works use variables and features interchangeably (e.g. [54], [55]). Aware of this difference, we adopted the term *feature* only in the case synthetically manufactured variables are concerned.

The main goal of these routines is to define the amount of information provided by each variable and discard irrelevant or redundant variables to eventually minimize computational efforts [53], [55]. The variable selection algorithms are divided into two main categories: (a) subset selection, and (b) ranking of features [56].

Both parametric and non-parametric methodologies that can establish variable rankings have been used. In the following, a subset of commonly used methodologies is listed below.

<u>Simple variable ranking with literature values</u>

A simple implementation of such variable rankings can be achieved manually. Herewith, previously defined (e.g. literature-based) weights are allocated to each census variable. Given the sum of a weighting vector with all variable values, a cell ranking can be achieved.

For example, if literature results suggest education, income and age play a major role in the adoption of a certain technology, such variables ("adoption favouring criteria" – ***afc***) could receive a higher associated weight than other census criteria (***occ***) or subsets that have been negatively linked with technology adoption. In the latter case, negative weights can be given. The example below shows the case, where a 20-fold higher importance to variable subset has been given (***afc***)

$$af_n = \begin{Bmatrix} 1.00 & for\ n \in \{afc\} \\ 0.05 & for\ n \in \{occ\} \end{Bmatrix} \tag{3.13}$$

Such weight vector should numerically discriminate the importance (expressed as numeric weight) of adoption favouring criteria (***afc***) and all other census criteria (***occ***). Eventually, af_n can be substituted by values derived from enhanced inference tools (such as in [13], [57]). Such tools, along multiple linear regression, will be explained in the following.

<u>Variable ranking with Linear Regression models</u>

This section addresses multi-linear regression models (MLR). These are linear regression models with multiple input parameters [58] that are widely used. Commonly named advantages include simple implementation, good interpretability and small computational effort, compared to other techniques. On the downside, MLR can only model linear relationships within the data [58].

In this work, we are especially interested in the variable weights that are established once a model is calibrated. Therefore, the data set is split into training set (75%) and test set (25%). The model is fitted in *R version 3.5.*, using *glm.fit*, which minimizes iteratively reweighted least squares. Note that in our data set, 28 input variables are not linearly independent. Following [58], such variables have been dropped. This led to a matrix with all variables being linearly independent. As a result, model inputs do not mutually influence parameter estimates anymore [58].

<u>Variable ranking with Artificial Neural Networks and the Olden approach</u>

Artificial neural networks (ANN) are universal function approximators. They resemble the shape and wiring of neural connections of the human brain and typically consist of multiple layers of neurons. ANN can outperform certain other methods such as linear regressions due to their ability to capture non-linear relationships [58]. One key drawback of ANN is that they have no direct explanatory power, an issue addressed by Olden [59]. While ANN weights have been transferred to rule-based reasoning in [60], latter approach is not suitable to quantify the relative weighting of input variables required for variable ranking.

Building on a former attempt to illuminate the relation of neural weights to input variables presented in [59], a variable weighting methodology that outperforms all former implemented attempts has been introduced in [61]. Latter methodology consists of a consecutive multiplication of neuron weights of each predictor throught the “input - hidden layer - output” typology chosen. Predictor variables’ aggregated weights are added up and ranked eventually. Given that this approach provides positive and negative importance values, such approach can be adapted to achieve variable ranking (e.g. from highest to lowest importance value).

The composition of such ANN weights vectors can be used to achieve census cell ranking required in the technology diffusion model presented in this work.

The ANN typology employed in this work consists of one input layer, one output layer and one hidden layer. The input layer consists of neurons for each of the 122 sociodemographic attributes. The consecutive layer, also known as hidden layer, instead contains 12 neurons and has been determined after a testing period. Furthermore, different activation functions for the hidden neurons are tested. The non-linear, hyperbolic tangent (tanh) function is selected eventually. The cost function used is the sum of the squared errors (SSE). Resilient backpropagation with weight backtracking (*rprop*) is set as adaptive optimization algorithm to minimize

SSE. *Rprop* is a supervised learning algorithm that is faster than training with backpropagation and does not require additional free parameter values [62].

Before training the ANN, the input data is transformed using min-max transformation. Furthermore, the data set is balanced by retaining approximately the same number of negative observations than positive observations. Balancing data sets have shown to improve predictive performance of ANN [61]. It has been performed for all three methodologies (MLR, ANN, IGR).

Variable ranking with Information-Theoretic Criteria

In this work, variable ranking based on mutual information (MI) was tested as well. This approach measures the MI of a variable associated to a class label (dependent variable). MI measures the mutual dependency between the two variables and is intrinsically linked to entropy [63]. Entropy measures the level of impurity in a group of examples and thus represents a measure of system's unpredictability. The use of Information Theoretic (IT) criteria does not rely on normality assumptions and dependency discoveries can go beyond linear relationships, unlike most of the other models currently used to identify drivers of DER adoption. In direct comparison to linear regression models, the IT approach displays the strength of association of a given variable to the outcome variable in absolute terms. That is, IT-based analysis does not provide information on the sign (positive/negative character) of such association.

Shannon introduced the notation of Information Entropy in 1948 in his work on Information Theory (IT) (see [63]). He defined the entropy $\boldsymbol{H}$ (in bits) of a discrete random variable X as:

$$H(X) = -\sum_{i=1}^{N} P(x_i)\, log_2\, P(x_i) \tag{3.14}$$

Here, x_i is a possible value of X. If any observation about the given data is made, new information can then be recomputed. The difference between the two information values is called Information Gain ($\boldsymbol{IG}$), a measure that has been introduced in the context of decision trees [64]. Here, IG is used to detect the variable for the decision tree root node.

The entropy change represents the information that is gained by the observation and can be used to identify the branch node that provides most information for a decision tree. IG can be written as:

$$IG(X,Y) = H(X) + H(Y) - H(X,Y) \tag{3.15}$$

Here, ***H(X,Y)*** is the joint entropy of ***X*** and ***Y***. As observed by [52], IG is biased towards choosing attributes with large number of values. This can cause overfitting, resulting in selecting variables as root-nodes that are non-optimal. Information gain ratio (***IGR***) attempts to correct the information gain calculation by introducing a split information value (***ISP***), in order to reduce its bias. After calculating the split information as described in [52], IGR can be computed.

The IGR represents the relation between IG and the intrinsic information of a split after an observation (Y):

$$IGR(X,Y) = \frac{IG(X,Y)}{ISP(X,Y)} \tag{3.16}$$

An overview of other IT-based variable selection procedures and an approach that compares rankings based on their compressibility have been discussed or presented in [54] and [65].

3.6 RESULTS OF THE SPATIAL DIFFUSION MODEL CALIBRATION

Accuracy metrics

The predictive power of the proposed diffusion model has been tested considering two aspects:

- Quantification of DER adoption per HV/MV service areas;
- Identification of the HV/MV service areas that receive DER adopters.

As described above, model configurations include four temporal discretizations as well as different cell rankings are compared to the real occurrence of DER using the Mean Absolute Error (MAE) [66] and contingency tables in order to retrieve accuracy measures (ACC) [67].

Furthermore, the average deviation of EV adopters per HV/MV substation service area is calculated. As the adoption of DER is at early stage and our model uses discrete adoption steps (e.g. States 1-5), we chose the Mean Absolute Error (MAE) metric over Mean Square Error, as latter is overly penalizing outliers that are likely as we use discretized stages. MAE is retrieved using following formula:

$$MAE = \frac{1}{NO} \sum_{i=1}^{NO} |y_i - \hat{y}_i| \tag{3.17}$$

This metric evaluates the average deviation of EV adopters per each HV/MV substation service area, dividing the absolute, summed deviation of the estimated number of EV/PV ($\hat{\mathbf{y}}$) from the real observed quantities ($\mathbf{y}$) by the number of observations ($\boldsymbol{NO}$). Similarly, the Root Mean Square Error (RMSE) is calculated. Different to the MAE, the RMSE represents an aggregated, squared summation of the differences of all paired observations. Eventually, the summation is subject to root extraction as shown below:

$$RMSE = \left[\frac{1}{NO} \sum_{i=1}^{NO} (y_i - \hat{y}_i)^2 \right]^{1/2} \tag{3.18}$$

On the other hand, the accuracy metric provides insights in the performance of the spatiotemporal DER forecast globally (e.g. compared to a perfect accuracy of 1.0) and in comparison to current standard approaches in use. Hence, during the first evaluation, we transform the evaluation of the adoption forecast into a classification problem, as available adopter datasets used in this work have only one time-stamp, and given their quantity being very small (EV <3,000) with regard to the total population (>10 million inhabitants). However, such value should be carefully analysed, given that a wrongly forecasted census cell can already turn a HV/MV service area from a non-adopter into an adopter.

The accuracy can be retreived counting all correctly predicted positives (true positives – TP), all wrongly predicted positives (FP) and all truly identified negatives (TN), all wrongly predicted negatives (FN) while applying the following formula [68]:

$$ACC = \frac{TP+TN}{TP+FP+TN+FN} \tag{3.19}$$

In the presented case, positives are HV/HMV substations classified as adopter substations and negatives are non-adopter HV/MV substations. All three metrics will provide insights that allow for the optimization of the spatiotemporal model parameters.

Outcomes

As the presented model consists of several, adaptable modules, the performance of different model configurations is of interest. Outcomes will indicate:

i) Which model configurations outperform others given a pre-established set of performance indicators (absolute, global performance);

ii) Which model parameters are likely to bring further, potential improvements (relative, localized performance).

Thus, the technology diffusion model was tested comparing outcomes of 16 model configurations, resulting from a combination of four cell ranking techniques as well as four temporal discretization intervals (Eq. 3.09. – Eq. 3.12.). The 16 model configurations are visualized below (Figure 3.8).

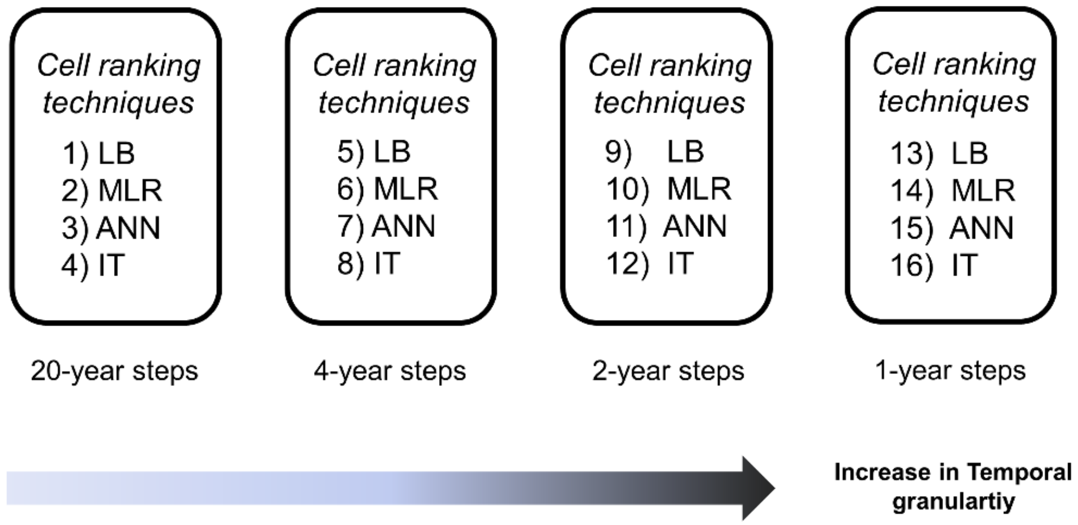


Figure 3.8 Technology diffusion model configurations

Results show that neither temporal full discretization or the finest granularity (1-year intervals) achieve the highest performance in both MAE and RMSE. While such discretization results in MAE values well below 10 adopters per HV/MV service area, configurations with 2-year and 4-year intervals produce deviations equivalent to 6-8 adopters per HV/MV service area. Curiously, error levels for 20-year and 1-year intervals are similar as well as 2-year and 4-year intervals (Figure 3.9).

Interestingly, a high similarity of outcomes gained by using the first three cell ranking techniques can be detected. This is noteworthy, as such techniques (simple literature-based weight allocation, regression weights and artificial intelligence-based analysis) differ in complexity. One explanation may lie in the similarity of our study cases' variable weight vectors to previous literature findings. The interested reader may find an extensive analysis in [29].

On the other hand, IT variable ranking resulted in the highest RMSE and MAE errors. This is not surprising as ANN and MLR indicated several criteria being negatively associated to EV adoption. However, the IT method used (Gain Ratio), as all IT approaches, cannot discriminate between positive and negative association. Furthermore, the implemented approach suffers in case multiple correlated input variables are present. Such variables can skew the resulted ranking.

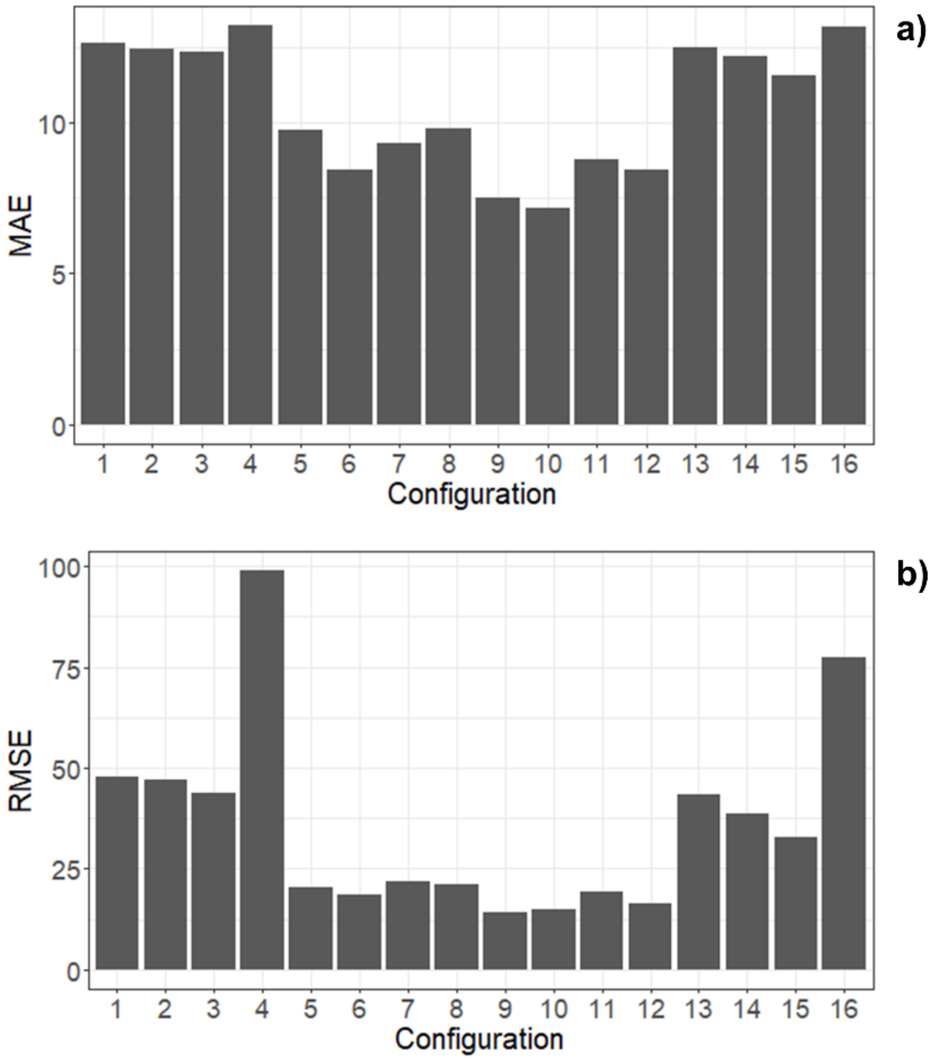


Figure 3.9 Results of error analysis of all 16 configurations.

The classification accuracies of all 16 model modifications strengthens drawn conclusions from the error analysis. Please note that Figure 3.9. shows both MAE and RMSE respectively (a, b).

Again, a strong effect of temporal model configurations to model accuracies are discovered. Looking across all performance terms, MLR is the outperforming method for construction variable weight vectors. As the literature-based (LB) approach is only slightly behind MLR accuracy and error values, it provides a robust way to merge judgmental knowledge of the planning outputs with mathematical frameworks from similar case studies.

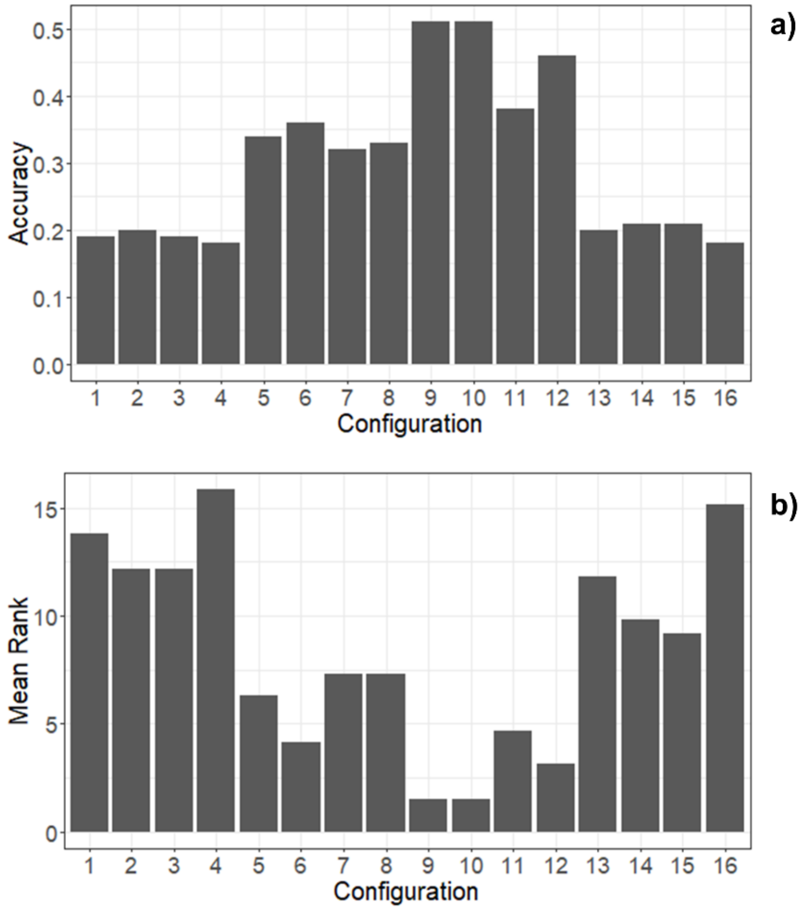


Figure 3.10 Results of error analysis of all 16 configurations.

Discussion of the calibration process

All 16 configurations have been implemented on a 64-bit Windows setup with 2.0 GHz, using R version 3.5.1. Before model validation, the input data set has been transformed using min-max transformation. Although numerous normalization methods do exist (z-transform, min-max, etc.), the effect of different techniques lies outside the scope of this work.

Another important aspect is the ratio of positive to negative observations. As DER adoption processes are in an early phase with low overall penetration levels, adopter to non-adopter observations tend to be unbalanced. In this work, the data set has been balanced by retaining approximately the same number of negative observations than positive observations.

Balancing data sets has shown to improve predictive performance of ANN [62]. It has been performed for all three methodologies (MLR, ANN, IGR).

As discussed in [69]–[71], balancing can have strong effects on modelling outcomes. While several strategies to cope with unbalanced data have been presented (e.g. in [69]), it is still a strongly debated topic whose deeper analysis represents a promising research avenue of future work. Therefore, it is not further discussed in this work.

The performances of different model configurations have been presented above. Outcomes suggest that trade-off between granularity and model complexity can be made. Especially, it has been observed that the construction techniques of census cells' weight vectors add little to model performance if compared to the simplest case (manual literature-based, weight allocation).

On the other hand, model performance increases with finer temporal granularity. However, using a granularity equivalent to less than 2 years seems to decrease model performance both in terms of errors and accuracy. Given little real-world observations and the unknown adoption time that has been aggregated to the base year, an explanation of such behaviour remains fully speculative.

In general, 2-year or 4-year adoption intervals seem most suitable, with the latter demanding significantly less computational resources (roughly half than for 2-year intervals).

3.7 UNCERTAINTY IN SPATIAL DIFFUSION MODELS

Uncertainty in spatial data

All spatiotemporal technology diffusion models rely on spatial data. Given that spatial data-sets typically arise from various sources (e.g. census institutes, utilities, urban planners) with varying capturing methods (air-borne laser, satellite, field surveys), uncertainty is introduced. In order to fully understand the predictive accuracy of spatiotemporal forecasting methods, such uncertainty needs to be addressed. While in technology diffusion models, uncertainty in both spatial and temporal data exists, latter is mainly concerned with the uptake rates of a given technology. As such rate is (as in e.g. [13], [14] and this work as well) calibrated separately against real observations, further analysis on temporal inaccuracies lies outside the scope of this work.

According to [18], uncertainty in spatial data can occur because of:

- Classification errors;
- Locational errors;
- Errors in nature/magnitude of value/relation.

As research on spatial data quality is dynamically emerging, there is no widely accepted definition on spatial data uncertainty. It has been characterized in different ways, considering either the effects of uncertainty on modelling processes [18] or the origin of spatial data uncertainty [11]. Former work divides uncertainty into effects on:

- Location (the position of spatial points, lines, polygons, raster cells and values);
- Level (the occurrence of certain attributes in a specific region);
- Nature (the relationship among spatial attributes and spatial data typologies).

Apart from a lack of agreement on the definition of spatial data uncertainty, in [18] it is stated that modellers rather struggle to correctly represent uncertainty then removing it. Obviously, given the simplifying character of models in general, latter is never possible. It should be remembered that uncertainty is handled differently, if deterministic or stochastic models are used. Deterministic models typically do not integrate uncertainty. On the other hand, within specified boundary conditions, stochastic models include a random component.

Former produce the same outcomes for the same input. Deterministic models thus assume that given a sufficiently granular description of states (current, past), perfect forecasts on future states would be possible. In a way, realistic models buy "tractability, but at the cost of realism" [18].

Stochastic models on the other hand, recognize that many natural processes (and measurement processes as well), possess a variability in process rates or measurement systems among others. Stochasticity can be introduced artificially by adding a random variable, that, e.g. could follow a predefined distribution (Gaussian, Binomial, Poisson, etc.). Using probabilistic approaches (Monte Carlo) further implies a change in verbal expression of the results. Typically, results are expressed in the probability to exceed a certain value.

As stated in [11], one of the fundamental reasons for the existence of spatial data inaccuracies is the difference between a complex, dynamically evolving world that can only be described through continuous numerical representations and the discrete, simplified computer environment.

It originates from processes such as data conversion, interpolation methods, digitization of cartographic material, field surveys, airborne and terrestrial laser scanning as well as remote sensing methods among others [47]. An overview of spatial data capture methods and their reported accuracies is listed below. While ground station surveys are the most accurate methods until today, spatial data production using laser scanner or remote sensing technologies have been becoming increasingly accurate during the past decades [11].

Table 3.4 Spatial data capture methodologies and typical accuracies (based on [47])

Spatial data capture method	Accuracy
Map digitization	NA
Aerial Photogrammetry	NA
Ground station survey	1mm – 1ppm
Global Positioning System (GPS)	10 – 20 m
Laser scanner	mm-cm (terrestrial)
	0.1 m (aerial)
Remote Sensing (Satellite images)	60 – 0.5m
	(strong, recent improvements)

In spatial models, uncertainties can be modeled in different ways. In [47], Shi presents three approaches that relate to different modules within a given spatial analysis task. According to this division, uncertainty can be assessed in the spatial data itself, in the spatial model that is built on the input data and links input to output, eventually, in the spatial analysis (compare Fig. 3.11). Latter includes all spatial operations conducted during modeling.

The three categories with exemplary subcategories are listed below:

1) Modelling Uncertainties in Spatial Data (Input to model)

    a. Modelling Positional Uncertainty in Spatial Data;
    b. Modelling Attribute Uncertainty;
    c. Modelling Integrated Positional and Attribute Uncertainty.

2) Modeling Uncertainties in Spatial Model (relation of inputs to outputs)

    a. Modelling Uncertain Topological relationships;
    b. Modelling Uncertainty in DEM.

3) Modelling Uncertainties in Spatial Analyses (model building blocks – routines)

    a. Modelling Uncertainty in Overlay Analysis;
    b. Modelling Uncertainty in Buffer Analysis;
    c. Modelling Uncertainty in Line Simplification.

This work considers uncertainties in the spatial model (topological relationships) only (Figure 3.11). However, under the presented layout of the spatiotemporal technology diffusion model, such topologies (the allocation of spatial census cells to HV/MV transformer locations) are directly linked to aspects of positional uncertainty in spatial data, and, attribute uncertainty (spatial cell is supplied by transformer A or B). Likewise, uncertainties in the spatial model incorporate spatial analysis uncertainties as the DER adopters that have been spatially superimposed with the spatial census data-set.

Uncertainty can also be introduced due to fusion of data from different sources or spatial data models. However, the analysis of combinatorial uncertainty generation due to data fusion lies outside the scope of this work and is therefore not considered. Furthermore, uncertainties in census data, adopter locations and spatial operations are not considered.

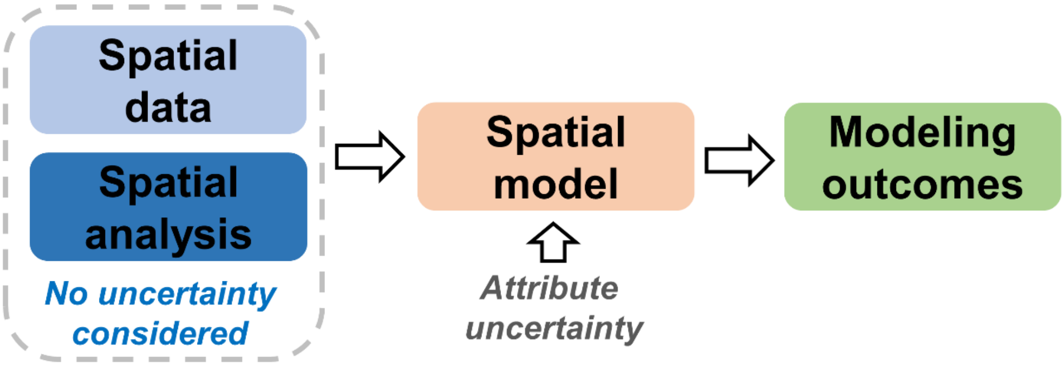


Figure 3.11. Uncertainties assessed in the proposed methodology within this thesis

However, from an electricity network planner or energy policy officer perspective, it is the understanding of the effects of such uncertainties on model outcomes that is fundamental. This is the domain of error analysis methods that can assess model-based uncertainty. According to [18], commonly used error analysis methods are the error analysis, sensitivity analysis, uncertainty analysis and robustness analysis that are described below (Table 3.5).

The approach pursued during this thesis is twofold: First, uncertainties in the development of the spatiotemporal technology diffusion model are assessed. Here, the spatial simulation model is calibrated considering topological spatial data uncertainty while shuffling various methodological configurations.

Secondly, uncertainty of external, global conditions is translated into sets of discrete scenarios that are used to hedge the risk inaccurate results can pose to decision makers.

Table 3.5 Error analysis methodologies (based on [18])

Error analysis method	Description
Error analysis	Considers error propagation and checks if added uncertainty of used data sources amplify or compensate on aggregated model basis (error amplification or error compensation)
Sensitivity analysis	Is often used as a local analysis where one parameter is varied within a predefined interval (e.g. +/- 10%) while holding all other parameter values constant. Typically, the sensitivity analysis relies on the proportional change of the analysed variable; If the ratio of change in output given the input is exceeding 1.0, we call this "sensitive", while changes below 1.0 are associated with a "robust" parameter.
Uncertainty analysis	The basic goal of uncertainty analysis (UA) is to isolate the most sensitive variables and rank them according to their sensitivity for the model output. Compared to sensitivity analysis, UA is a more general approach that considers the interaction of multiple variables. A common approach is to identify the most sensitive variables and then associate a probability distribution to each variable (the choice of the probability distribution is crucial). As the UA as a multivariate analysis includes parameter interaction, parameter space sampling becomes crucial to reduce computational costs.
Robustness analysis	The structural uncertainty of models (or model uncertainty) is defined as the way the models' structure affects its outcomes. It was proposed by Jansen (1998) to coarsen spatial/temporal resolutions or reducing model complexity to assess the effects of model uncertainty. As extreme, the model structure could be completely replaced by another representation (robustness analysis). It should be noted that the repetitive checking and running of model simulations during its development and refinement stages is an informal way of conducting a robustness analysis.

The spatiotemporal technology diffusion model proposed in this thesis is a deterministic model. However, for the analysis of model outcomes, it is very helpful to assess the uncertainty due to measurement methodologies and uncertainties that are produced and propagated through spatial analysis processes [11].

As mentioned above, uncertainty of topological relationships has been considered within this work. This is achieved through the analysis of positional uncertainty in spatial input data (the relation of census cells to HV/MV transformers). In other words, the uncertainty of HV/MV transformer locations with respect to the position of spatial census cells is modelled. Such relation contains both positional aspects (the distance of census cells to HV/MV transformers) and attribute aspects (each census cell is assigned to one transformer area).

To assess the impact that erroneous positional information has on model outcomes, projected HV/MV substation coordinates (in meters) are synthetically altered. Using a random distribution with equal probability and replacement, HV/MV substation transformer positions have been permutated considering positional 10 – 100m errors along both latitude and longitude axis.

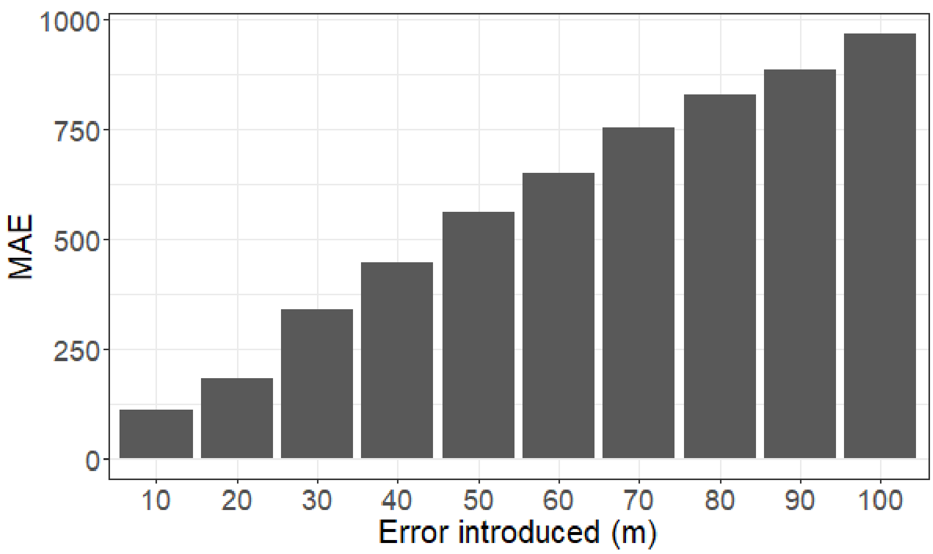


Figure 3.12 Simulated evolution of Mean Average Error (MAE) along growing positional error

The figure above (Fig. 3.12) shows the simulated evolution of the MAE of population counts that are contained in each HV/MV service area after and before permutation. As seen, the MAE grows with the error introduced.

Given an average population of roughly 26,000 residents within one HV/MV service area and considering the smallest service area with approximately 3,200 residents, it becomes clear that errors until 20m will have even under worst case conditions only minor impact (<6%) on the model outcomes. This is especially noteworthy, recalling 20m being the maximum error that may be introduced under global positioning systems (GPS) which also ranges among the largest uncertainty sources (Table 3.4.).

Reducing uncertainty in spatial data with data quality management

Given the previously described ways that can impact on the accuracy of spatial data and the reliability of spatial models, spatial data quality management is a crucial process for spatial models. A list of spatial data quality elements has been presented by [72]. Such quality elements include:

- Lineage;
- Positional accuracy;
- Completeness;
- Logical consistency;
- Semantic accuracy;
- Temporal information;

Recently, the growing use of spatial data and rising concerns to handle inaccuracies have led to standardization efforts. As such, ISO 19113 2005, an industry norm on Geographic Information – Quality principles has been developed and recently actualized (**ISO 19157 2013).** ISO 19157:2013 established a set of definition and standard processes to assess the quality of geographic data in a standardized way. It includes [73]:

- A definition of data quality components;
- Specifications of data quality measures;
- Detailed description of the general procedures for evaluating geographic data quality;
- Development of principles for reporting data quality.

3.8 Comparison to other model implementations

The scientific method requires validation. While one way to achieve validation is the confrontation of modelling outcomes with real observations, another way to accomplish validation is through comparison with other methodologies. In the following, first, other attempts to model DER diffusion in space and time will be presented. Eventually, such models are compared to the presented approach.

Former developed models include popular agent-based models (ABM) [74]–[77], spatial regression models [13], [78], [79] and simulation-based approaches other than ABM [9], [14]. It is noteworthy, that all works are very recent and have been simultaneously developed within the last 5-10 years.

A recent report differentiates existing DER diffusion models into eight categories: 1) Time series, 2) Regression, 3) Machine Learning, 4) Bass Diffusion Models, 5) Customer Behaviour, 6) ABM, 7) Combined Market Penetration and 8) Macroeconomic models. Although the importance of sufficient spatial granularity is mentioned, a model delineation along spatial dimension is not presented [4]. Furthermore, research on spatial regression or simulation has not been included.

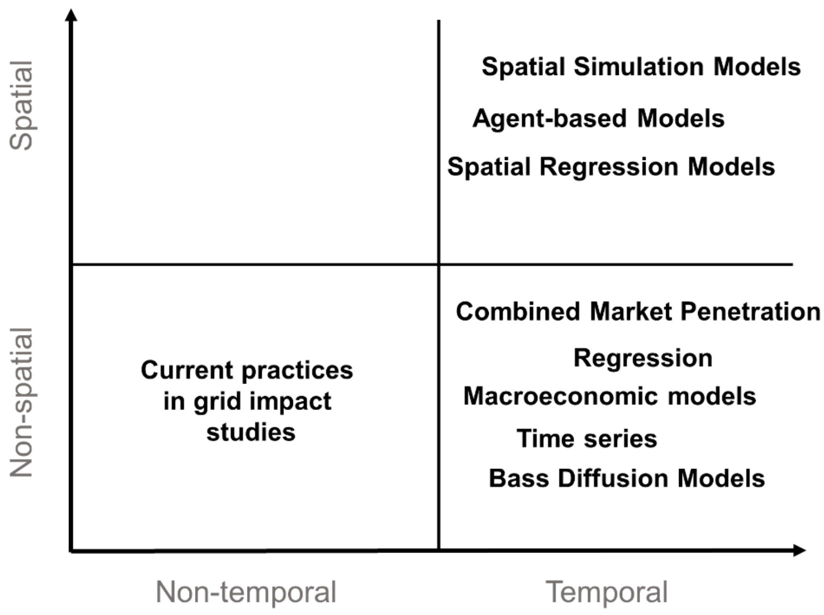


Figure 3.13 An overview of technology diffusion models.

Figure 3.13 displays all diffusion model types presented in [4], current practices in grid integration studies [9] as well as spatial regression and spatial simulation model families. It can be observed, that most models consider the temporal aspect of technology adoption, while spatial aspects are commonly neglected. Only spatial simulation studies (other than ABM), agent-based models and spatial regression models can be used to forecast the spatial patterns of technology adoption. In this work, ABM have been separately analyzed from spatial simulation models as former have the particularity to model single individuals or agents which is typically not the case for spatial simulation models.

As shown in [9], current studies that assess the grid impact of electrical appliances to electricity networks commonly neglect the temporal and spatial dimension of technology adoption, rather relying on fixed capacity extrapolations or random allocations.

The model family of the top-right corner will be discussed in more detail in the following.

As one of the most cited works on technology diffusion using agent-based modelling, the work of [80] includes theoretical considerations within an empirically driven, agent-based model to forecast technology adoption (PV) on household level. Model validation was delivered though predictive accuracy, RMSE of marginal adoption rates (temporal dimension with changing price and subsidy patterns), spatial accuracy and demographic accuracy. Spatial prediction errors were analysed with fuzzy numerical similarity and wavelet verification. However, the authors state that model results heavily relied on expensive, household-level survey data. In addition, social relationships had to be modelled relying on the small world algorithm that randomly assigns influence links throughout the population under analysis.

A similar ABM approach to model EV diffusion was developed in [77]. Using the models uniquely for modeling large-scale interactions of EV with the distribution grid infrastructure, the authors integrated agent-based traffic simulation software MATSim (Multi Agent Transportation Simulation) together and a Vehicle Technology Assessment Model), and eventually, power system simulation software. The latter includes as well routines to manage the grid interactions with plug-in electric vehicles [77]. While the work provides insights in the joint modelling of transport and electricity network flows, it does not consider the structure of the social system in which these interactions are embedded.

The study by [76] models EV diffusion with ABM approaches, where the utility to adopt for each agent rises by the increasing popularity of the technology as well as the rising numbers of peers which adopt EV. As basic assumption, income was inversely correlated with risk-averse behaviour. Low income groups are therefore especially constrained to adoption. The model outcomes could not exploit real observations to support model validation. Due to lack of data, the authors have been exploring different patterns produced rather than matching outcomes with real EV adoption data. Results showed that dense clusters could be generated even if mild per-effects are at work. The authors stretch on the difficulty to model social interactions and the need of further calibration, e.g. to validate the synthetically established social network typologies employed [76].

A study on PV diffusion using agent-based modelling was conducted by [81]. The study focused on the Italian residential sector, simulating PV diffusion with ABM for yearly adoption increments. Sub-modules represented the Italian social system and the agents' individual investment decision (household level) using census data. The study displayed another drawback of agent-based models. A major limitation noted during the implementation of the model was the need of computational resources. Looking at country-level, the study included about 10 million Italian households of a total of 23.9 million households, living in one- or two-family houses. The computation was speeded up by aggregating households together. Calculation efforts without aggregation would have required 8 bytes of hard-drive memory for the 20 attribute values of each agent and a total of about 1.5 GB for one simulation step. Additionally, the whole simulation would have lasted about 12 hours and required 30 GB hard-drive storage capacity [81].

Spatial regression models represent another popular model type that is applied to forecast technology diffusion processes in time and space. The most recent work has been developed in [13]. Here, the authors developed a two-stage process to estimate the adoption of residential EV, PV and HVAC appliances. After starting with an empty logistic regression model, census variables from a data-set partly crossed with real observations have been consecutively selected. Then, for each technology, a final model has been established given predefined performance criteria. Eventually, such models where used to forecast technology adoption over a 10-year horizon. Using a non-spatial regression model, the developed model neglects spatial interaction. A further drawback mentioned by the authors is the reliance on expensive survey data and adopter observations that are often not available (e.g. the authors in this work model HVAC adoption using literature values).

On the other hand, authors of work [78], [79] present geographically weighted local regression models (GWR) that are applied to predict spatial adoption patterns of load growth and household appliances with high power demand (electric stoves). Both studies show both advantages of spatial regression techniques (ease to implement) and limitations (relatively high error of 50-60%).

Further typical limitations of spatial regression models lie in the reliance of mostly simplified (e.g. binary) neighbourhood structures [82], difficulties to predict hotspots with values far above local averages [83] and multicollinearity of predictor variables. In addition, such models are bounded to detect linear dependencies among spatial data-sets.

The work in [79] presents a hierarchical Bayesian model that has been developed to generate spatiotemporal PV adoption patterns in a Brazilian city. The paper includes a forecast for residential uptake of PV panels, considering different energy policy developments. However, in the presented way, such model relied heavily on sets of parameters, synthetic distributions of model variables and globally set hyperparameters that are chosen beforehand and fixed. Thus, they cannot track dynamic changes of model parameters.

As previously shown, the model presented within this work [9] uses a cellular module build on adoption stages that is sequenced through an innovativeness score-based ranking. Such simulation models use discretized uncertainty, that is, predetermined scenarios that are translated into spatial adoption patterns. Thus, the spatial simulation model takes advantage of a deterministic model typology that is most suitable in case discrete scenario sets are used.

However, the deterministic model can be seen as a building block of a stochastic model. Combining a distribution of scenarios with Monte-Carlo methods, stochastic processes can be simulated. Furthermore, it is obvious as the census-based nature of the presented simulation model requires reliable census data that is openly accessible. In fact, there is a rising availability of census data (e.g. census data is freely available covering all European countries [84], and many other, such as the United States [85], and Brazil [86] among others). Thus, census data represents a convenient source to build such models.

An alternative, spatial simulation model that relies on little population information only, has been introduced by [14]. The presented approach includes a threefold process. Building on a spatial model that produces global PV adoption time series, the authors use a support-vector based model to estimate the adoption probability in each cell for a given time horizon. Eventually, the global PV adoption forecast is allocated among the cells that are classified as growth cells. While the model provides a framework that does not rely on census data and census polygons, the authors arbitrarily chose a cell lattice with 800-meter resolution that was enriched with information of six predictor variables. Such arbitrary aggregation, as shown in Chapter 2, may negatively impact on the studies explanatory power.

Table 3.6 Advantages and limitations of spatiotemporal technology diffusion models

	Spatial regression models	Simulation-based models	Agent-based models
Advantages	• Can fully capture social system of rich census data • Simple to implement	• Can fully capture social system of rich census data • Simple to implement • Flexible adoption of new features	• Enhanced modeling with detailed representation of decision processes • Neighbourhood interaction
Limitations	• Relies on census data availability • Neighbourhood interaction assumed static	• Relies on census data availability • Computationally demanding • Neighbourhood interaction requires cell lattices or simplified neighborhood structures (if in vector format)	• Relies on census data availability • Requires costly empirical survey data • Computationally demanding • Neighbourhood interaction requires cell lattices or simplified neighbourhood structures (if in vector format)

Previously presented diffusion model typologies require complex error calculations that account for the spatial and temporal interdependency of adoption behaviour. Such stage-wise analysis can only ground on multi-year adoption observations that are still rare to find (one exception with detailed, quarterly adoption time series is presented in [14]). An implementation of all diffusion model types presented for a joint case study could provide further insights in the relative performance of technology diffusion models. Until today, such comparison has not been conducted and the realization of such experiment remains future work, not covered within this thesis.

As discussed in [9], current techniques in industry and academia tend to avoid the use of technology diffusion models, relying mostly on randomized or very simplified extrapolation techniques. Such approaches take a few network parameters into account (e.g. installed transformer capacities) and do not rely on a model that possesses both spatial and temporal dimension.

A comparison of the developed model to current industry standards is therefore presented in a dedicated chapter on the interface of electricity network planning and technology diffusion (Chapter 4).

Finally, for the frictionless implementation of a spatiotemporal diffusion model in utilities' departments and for energy policy makers, Occam's razor rule can serve as a guiding principle: The concept conveys the idea that in case of multiple, competing models that are used for inference, the simplest/least complex one should be used [18]. The spatial simulation model that has been presented in this chapter satisfies this requirement, providing a very flexible forecasting tool that relies on inexpensive minimum input data and little computational resources.

Chapter summary and conclusions

The chapter introduced spatiotemporal DER adoption models. Such models typically consist of a global DER forecast, a cellular module, consisting of cell states and an activation order, as well as a module for adoption pattern mapping. In the model presented within this thesis, the cellular module builds on granular census data. Using elements of Roger's theory on the *Diffusion of Innovations*, an activation order is built on innovativeness scores, that allow to rank census cells along the adoption process. Comparing different model modifications (e.g. time discretization or ranking estimations), the trade-off between data granularity, model complexity and model accuracy can be identified. Outstanding results can be summarized as the following:

- Highly accurate spatiotemporal technology diffusion models are currently impossible given the, by definition, reduced amount of adopter observations.

- Likewise, adoption forecasts under mature technology diffusion are eventually of little use given that market penetration is already close to saturation.

- Currently, few approaches that integrate spatial to temporal diffusion forecasts are available. Developed models can be grouped into 1) Spatial simulation models, 2) Spatial regression models, and 3) Agent-based models.

- Compared to other model families, simulation models possess both the highest flexibility and simplicity, which make them attractive for implementation in network planning or energy policy design.

- As in the proposed model, all spatiotemporal technology diffusion model families do possess a temporal model, a spatial model and a spatiotemporal linkage that determines the activation order of spatial cells.

- Model calibrations need to consider both spatial (resolution of the cellular model and adoption sequencing) and temporal (discretization of adoption process and forecasting intervals) variations.

- Uncertainty is present in all spatial data and analysis processes. Therefore, acknowledging and documenting inaccuracies is fundamental to assess the impact of uncertainty to model outcomes.

Chapter references

[1] IEA, *World Energy Outlook 2016*. Paris, 2016.

[2] European Commission, *EU Reference Scenario 2016: Energy, Transport and GHG emissions trends to 2050*. 2016.

[3] S. Coley *et al.*, "Guidance on Solar PV Adoption Forecast Methods for Distribution Planning,", 20 pp, Palo Alto, USA, 2018.

[4] P. Gagnon, B. Stoll, A. Ehlen, G. Barbose, and A. Mills, "Estimating the Value of Improved Distributed Photovoltaic Adoption Forecasts for Utility Resource Planning," Denver (CO), 2018.

[5] EPRI, "Developing a Framework for Integrated Energy Network Planning (IEN-P)," Palo Alto, California, 2018.

[6] K. Bell, "Methods and Tools for Planning the Future Power System: Issues and Priorities," *Model. Requir. GB Power Syst. Resil. Dur. Transit. to Low Carbon Energy*, no. Paper 5, pp. 1–35, 2015.

[7] ICF, "Integrated Distribution Planning,", USA, 2016.

[8] F. Pilo, S. Jupe, , F. Silvestro, C. Abbey et al., "Planning and optimization methods for active distribution systems," *CIGRE Brochure number 591*, Paris, France, Aug. 2014.

[9] F. Heymann, J. Silva, V. Miranda, J. Melo, F. J. Soares, and A. Padilha-Feltrin, "Distribution network planning considering technology diffusion dynamics and spatial net-load behavior," *Int. J. Electr. Power Energy Syst.*, vol. 106, pp. 254–265, 2019.

[10] F. Heymann *et al.*, "Vertical Load Uncertainty at the T/D Boundary under different spatial DER allocation techniques," in *IEEE Sustainable Energy Systems and Technologies Conference*, 2019, Porto, pp. 1–6.

[11] W. Shi, *Principles of Modeling Uncertainties in Spatial Data and Spatial Analyses*. CRC Press, 2009.

[12] S. A. Robinson and V. Rai, "Determinants of spatio-temporal patterns of energy technology adoption: An agent-based modeling approach," *Appl. Energy*, vol. 151, pp. 273–284, 2015.

[13] R. Bernards, J. Morren, and H. Slootweg, "Development and Implementation of Statistical Models for Estimating Diversified Adoption of Energy Transition Technologies," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1540–1554, 2018.

[14] T. Zhao, Z. Zhou, Y. Zhang, P. Ling, and Y. Tian, “Spatio-Temporal Analysis and Forecasting of Distributed PV Systems Diffusion: A Case Study of Shanghai Using a Data-Driven Approach,” *IEEE Access*, vol. 5, pp. 5135–5148, 2017.

[15] B. Finkenstädt, L. Held, and V. Isham, *Statistical Methods for Spatio-Temporal Systems*, vol. 45, no. 1972. New York: Taylor & Fancis Group, LLC, 2007.

[16] N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, validation, and confirmation of numerical models in the earth sciences,” *Science (80-.).*, vol. 263, no. 5147, pp. 641–646, 1994.

[17] G. E. P. Box, “Some problems of statistics and everyday life,” *J. Am. Stat. Assoc.*, vol. 74, no. 365, pp. 1–4, 1979.

[18] D. O’Sullivan and G. L. W. Perry, *Spatial Simulation: Exploring Pattern and Process*, 1st ed. Wiley-Blackwell, 2013.

[19] E. M. Rogers, *Diffusion of innovations*, 5th Edition, New York Free Press, USA,2003.

[20] J. Massiani and A. Gohs, “The choice of Bass model coefficients to forecast diffusion for innovative products: An empirical investigation for new automotive technologies,” *Res. Transp. Econ.*, vol. 50, pp. 17–28, 2015.

[21] W. A. V Clark and K. L. Avery, “The Effects of Data Aggregation in Statistical Analysis,” *Geogr. Anal.*, vol. 8 (4), no. October, 1976.

[22] A. S. Fotheringham and D. W. S. Wong, “The modifiable areal unit problem in multivariate statistical analysis,” *Environ. Plan. A*, vol. 23, 1991.

[23] H. L. Willis, *Power distribution planning reference book*, Second Edi. New York, 2004.

[24] F. M. Bass, “A New Product Growth for Model Consumer Durables,” *Manage. Sci.*, vol. 5, no. 16, pp. 215–227, 1969.

[25] L. M. M. Benvenutti, A. B. Ribeiro, F. A. Forcellini, and M. U. Maldonado, “The Effectiveness of Tax Incentive Policies in the Diffusion of Electric and Hybrid Cars in Brazil,” *41st Congr. Latinoam. Din. Sist. São Paulo*, no. November, 2016.

[26] T. W. Valente, S. R. Dyal, K. H. Chu, H. Wipfli, and K. Fujimoto, "Diffusion of innovations theory applied to global tobacco control treaty ratification," *Soc. Sci. Med.*, vol. 145, pp. 89–97, 2015.

[27] H. Duan, G. Zhang, S. Wang, and Y. Fan, "Peer interaction and learning: Cross-country diffusion of solar photovoltaic technology," *J. Bus. Res.*, vol. 89, no. August 2017, pp. 57–66, 2018.

[28] B. Bollinger and K. Gillingham, "Peer Effects in the Diffusion of Solar Photovoltaic Panels," *Mark. Sci.*, vol. 31, no. 6, pp. 900–912, 2012.

[29] F. Heymann *et al.*, "DER Adopters analysis with Spatial Autocorrelation and Information Gain Ratio under different Census-data Aggregation Levels," *IET Renew. Power Gener.*, vol 14 (1), 2019.

[30] L. Bian, "A conceptual framework for an individual-based spatially explicit epidemiological model," *Environ. Plan. B Plan. Des.*, vol. 31, no. 3, pp. 381–395, 2004.

[31] M. Graziano and K. Gillingham, "Spatial patterns of solar photovoltaic system adoption: The influence of neighbors and the built environments," *J. Econ. Geogr.*, vol. 15, no. 4, pp. 815–839, 2015.

[32] VW, "Basics of Electric Vehicles Design and Function," *Report of Self Study Program 820233*, p. 62, Herndon (VA), 2013.

[33] KEMA, "Report - Study on the Impact of Distributed Generation on the National Electricity System,"Arnhem, 2011.

[34] F. Heymann, C. Pereira, V. Miranda, and F. J. Soares, "Spatial load forecasting of electric vehicle charging using GIS and diffusion theory," *2017 IEEE PES Innov. Smart Grid Technol. Conf. Eur.*, Turin, pp. 1–6, 2017.

[35] ENTSO-E, "TYNDP 2018: Scenario Report," Brussels, 2018.

[36] ENTSO-E, "ANNEX II : Methodology of Scenario report," Brussels, 2018.

[37] H. Jiayi, J. Chuanwen, and X. Rong, "A review on distributed energy resources and MicroGrid," *Renew. Sustain. Energy Rev.*, vol. 12, no. 9, pp. 2465–2476, 2008.

[38] A. Chaouachi, E. Bompard, G. Fulli, M. Masera, M. De Gennaro, and E. Paffumi, "Assessment framework for EV and PV synergies in emerging distribution systems," *Renew. Sustain. Energy Rev.*, vol. 55, pp. 719–728, 2016.

[39] M. H. Moradi, M. Abedini, S. M. R. Tousi, and S. M. Hosseinian, "Optimal siting and sizing of renewable energy sources and charging stations simultaneously based on Differential Evolution algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 1015–1024, 2015.

[40] D. B. Richardson, "Electric vehicles and the electric grid: A review of modeling approaches, Impacts, and renewable energy integration," *Renew. Sustain. Energy Rev.*, vol. 19, pp. 247–254, 2013.

[41] P. Nunes, T. Farias, and M. C. Brito, "Day charging electric vehicles with excess solar electricity for a sustainable energy system," *Energy*, vol. 80, pp. 263–274, 2015.

[42] ACAP, "Estatísticas do Sector Automóvel," *Assoc. Automóvel Port.(Portuguese Automotive Association)*, Lisbon, 2013.

[43] INE, "População residente em cidades (N.º) por Local de residência," Lisbon, 2016.

[44] Portuguese General Direction of Energy and Geology (DGEG), "Rapid Statistics - Renewables No.168," Lisbon, 2018.

[45] S. Izquierdo, M. Rodrigues, and N. Fueyo, "A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations," *Sol. Energy*, vol. 82, no. 10, pp. 929–939, 2008.

[46] J. Melius, R. Margolis, and S. Ong, "Estimating Rooftop Suitability for PV: A Review of Methods, Patents, and Validation Techniques," Denver (CO), 2013.

[47] W. Shi, *Principles of Modeling Uncertainties in Spatial Data and Spatial Analyses*, 1st Editio. CRC Press, 2009.

[48] P. Ardilly et al., *Handbook of Spatial Analysis - Theory and Application with R*, no. October. Montrouge Cedex, 2018.

[49] A. Bhardwaj, "Spatial Data Mining," in *Data Mining Techniques and Tools for Knowledge Discovery in Agricultural Datasets*, IASRI, 2012, pp. 153–166.

[50] J. Melo, E. Carreno, A. Padilha-Feltrin, and C. Minussi, "Grid-based simulation method for spatial electric load forecasting using power-law distribution with fractal exponent," *Int. Trans. Electr. ENERGY Syst.*, no. 26, pp. 1339–1357, 2016.

[51] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *J. Mach. Learn. Res.*, vol. 3, no. 3, pp. 1157–1182, 2003.

[52] A. G. Karegowda, A. S. Manjunath, G. Ratio, and C. F. Evaluation, "Comparative study of attribute selection using Gain Ratio and correlation based Feature Selection," *Int. J. Inf. Technol. Knowl. Manag.*, vol. 2, no. 2, pp. 271–277, 2010.

[53] G. Doquire, M. Verleysen, "Feature Selection for Multi-label Classification Problems," *IWANN 2011: Advances in Computational Intelligence pp 9-16,* 2011.

[54] G. Brown, "A New Perspective for Information Theoretic Feature Selection," in *12th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2009, no. 1, pp. 49–56.

[55] M. Bennasar, Y. Hicks, and R. Setchi, "Feature selection using Joint Mutual Information Maximisation," *Expert Syst. Appl.*, vol. 42, no. 22, pp. 8520–8532, 2015.

[56] J. Pohjalainen, O. Räsänen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits," *Comput. Speech Lang.*, vol. 29, no. 1, pp. 145–171, 2015.

[57] T. Islam and N. Meade, "The impact of attribute preferences on adoption timing: The case of photo-voltaic (PV) solar cells for household electricity generation," *Energy Policy*, vol. 55, pp. 521–530, 2013.

[58] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, 2nd Edition, USA2009.

[59] J. D. Y. J. Olden D. A., "Illuminating the 'black box': a ramdomization approach for understanding variable contributions in artificial neuronal networks.," *Ecol. Modell.*, vol. 154, pp. 135–150, 2002.

[60] A. R. G. Castro and V. Miranda, "An interpretation of neural networks as inference engines with application to transformer failure diagnosis," *Int. J. Electr. Power Energy Syst.*, vol. 27, no. 9–10, pp. 620–626, 2005.

[61] J. D. Olden, M. K. Joy, and R. G. Death, "An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data," *Ecol. Modell.*, vol. 178, no. 3–4, pp. 389–397, 2004.

[62] M. Riedmiller, Heinrich Braun, "A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm", *Proc. of the IEEE Intl. Conf. on Neural Networks*, pp. 586 - 591, 1993.

[63] J. C. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*, 1st ed. Springer Publishing Company, Incorporated, 2010.

[64] J. R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol. 1, pp. 81–106, 2007.

[65] J. H. Collier and A. H. Konagurthu, "An information measure for comparing top k lists," *arXiv*, vol. October, 2013.

[66] R. M. Forte, *Mastering predicitive analysis with R.* , 414 pp, Packt Publishing, 2015.

[67] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, 1991.

[68] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, pp. 861–874, 2006.

[69] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 42, no. 4, pp. 463–484, 2012.

[70] Y. Sun, A. K. C. Wong, and M. S. Kamel, "Classification of Imbalanced Data: a Review," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 23, no. 04, pp. 687–719, 2009.

[71] Q. Wei and R. L. Dunbrack, "The Role of Balanced Training and Testing Data Sets for Binary Classifiers in Bioinformatics," *PLoS One*, vol. 8, no. 7, 2013.

[72] J. L. Morrison, "Spatial Data Quality," in *Elements of spatial data quality*, S. Guptill and J. Morrison, Eds. 1995, pp. 1–13.

[73] International Organization for Standardization, "ISO 19157:2013 - Geographic information -- Data quality," 2013. [Online]. Available: https://www.iso.org/standard/32575.html.

[74] S. A. Robinson, M. Stringer, V. Rai, and A. Tondon, "GIS-Integrated Agent-Based Model of Residential Solar PV Diffusion," *32nd USAEE/IAEE North Am. Conf. July 28-31, 2013*, pp. 1–19, 2013.

[75] M. D. Galus, M. Zima, and G. Andersson, "On integration of plug-in hybrid electric vehicles into existing power system structures," *Energy Policy*, vol. 38, no. 11, pp. 6736–6745, 2010.

[76] D. McCoy and S. Lyons, "Consumer preferences and the influence of networks in electric vehicle diffusion: An agent-based microsimulation in Ireland," *Energy Res. Soc. Sci.*, vol. 3, no. C, pp. 89–101, 2014.

[77] M. D. Galus, "Agent-based modeling and simulation of large scale electric mobility in power systems," Doctoral thesis ETH Zurich, no. 20288, p. 314, 2012.

[78] J. D. Melo, "Spatial Pattern Recognition of Urban Sprawl Using a Geographically Weighted Regression for Spatial Electric Load Forecasting," *18th International Conference on Intelligent System Application to Power Systems (ISAP),* Porto , 2015.

[79] J. Villavicencio Gastelu, J. D. Melo Trujillo, and A. Padilha-Feltrin, "Hierarchical Bayesian Model for Estimating Spatial-Temporal Photovoltaic Potential in Residential Areas," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 971–979, 2018.

[80] V. Rai and S. A. Robinson, "Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental factors," *Environ. Model. Softw.*, vol. 70, pp. 163–177, 2015.

[81] J. Palmer, G. Sorda, and R. Madlener, "Modeling the Diffusion of Residential Photovoltaic Systems in Italy : An Agent-based Simulation Johannes Palmer , Giovanni Sorda and Reinhard Madlener May 2013 Institute for Future Energy Consumer Needs and Behavior (FCN)," no. 9, pp. 1–48, 2013.

[82] E. Gómez-Rubio, Virgilio; Bivand, Roger; Pebesma, *Applied Spatial Data Analysis with R*, Springer, 2nd Edition, USA, 2013.

[83] G. Hoek *et al.*, "A review of land-use regression models to assess spatial variation of outdoor air pollution," *Atmos. Environ.*, vol. 42, no. 33, pp. 7561–7578, 2008.

[84] EUROSTAT, "2011 Census data-base," 2011. [Online]. Available: https://ec.europa.eu/eurostat/web/population-and-housing-census/census-data/2011-census.

[85] United States Census Bureau, "Census data," 2019. [Online]. Available: https://www.census.gov/programs-surveys/geography/data.html.

[86] Brazilian Institute for Geography and Statistics, "Census micro data," 2018. [Online]. Available: https://ww2.ibge.gov.br/home/estatistica/populacao/censo2010/resultados_gerais_amostra/resultados_gerais_amostra_tab_uf_microdados.shtm. [Accessed: 08-Nov-2019].

4 Electricity network planning under technology diffusion

The large-scale adoption of DER impacts on the load profiles of residential consumers. Therefore, electricity network planning methodologies need to be updated, incorporating models that can realistically represent DER adoption dynamics within connected consumer groups. After introducing common, current simplifications to model DER adoption in electricity network impact studies, this chapter presents the application of a spatiotemporal DER adoption model to distribution and transmission planning study cases. Results show the range of uncertainties network planners face from applying different DER allocation techniques.

4.1 A BRIEF INTRODUCTION INTO ELECTRICITY NETWORK PLANNING

Electricity networks are typically large-scale, capital-intensive infrastructures. Therefore, independent system operators (ISOs), distribution or transmission system operators conduct short-, medium- and long-term planning exercises that aim to reduce the risks associated to investment decisions. Such risks may arise from changes in influencing factors that are endogenous to electricity planning (component costs, demand and generation patterns) or exogenous (capital costs, policy changes).

This section introduces general electricity network planning aspects while showing how newly introduced electrical appliances such as DER might impact on the planning of large electricity infrastructures. After presenting traditional approaches that are currently used to reduce the spatial and temporal uncertainty in network expansion planning, two case studies are presented.

The case studies compare DER representations currently used to model technology diffusion to a diffusion model with higher temporal and spatial resolution. This spatiotemporal model has been discussed in Chapter 3. Through the application to real world problems, the chapter allows to quantify uncertainties that arise if different technology uptake representations are used.

Central aspects of distribution system planning

Distribution networks are recognised as the lower, mostly radial networks that connect most of consumers, commercial or residential but also industrial (when not of large scale) to the power grid. In Europe, distribution system operation, maintenance and planning are typically carried out by distribution system operators (DSO). The design and exploitation of a reliable and safe network at a worthy profit and at reasonable cost for all consumers connected is achieved through distribution network planning (DNP).

The planning exercise is subject to many constraints but also must be undertaken considering a diversity of alternative functions, not in the sense of multiple criteria, but derived from distinct market designs, which condition the corporate purpose of the builder and owner of the system. For instance, while in some countries the owner of assets is also the commercializer of energy, in other countries these functions are unbundled. Also, the growing presence of new renewable sources at consumer level is questioning the traditional planning paradigms, by changing the nature of the network, from provider of energy to provider of backup support service.

Different types of distribution network planning can be identified [1]:

- Expansion planning;
- Operation planning;
- Greenfield planning.

The latter represents the design of electrical distribution networks from scratch, e.g. for new developing areas. Greenfield planning possesses the highest complexity among planning types due to the multitude of network components, interdependent cost structures and feasible designs [1].

Distribution network planners rely on spatial and temporal load and generation forecasts that serve as input to various planning tools. Such tools may include trade-off analysis of a set of reinforcement or non-wires alternatives and eventually inform investment decisions [2], [3].

Historically, distribution network expansion has been driven by demand growth [4]. Nowadays, such growth is expected to be driven by new load types or generator instalments [5]. Therefore, it has been argued that commonly used distribution grid planning tools lag behind the current requirements of networks which see high shares of DER [5]–[7]. Particularly, EV charging, distributed PV and electrified heating systems are foreseen to change future demand patterns and thus affect distribution network planning [5]. Furthermore, distributed storage, especially at consumer level, threatens to be a sort of game changer, if the drop in costs that is currently being witnessed opens a window of economic feasibility for such option.

Power flow analysis, fault analysis, dynamic analysis, and power quality analysis and advanced optimization are typical tools that interface with the load and generation forecasts. The interested reader may find a comprehensive overview over distribution network planning aspects and evolving planning requirements in restructured power industries in [3], [6], [8], [9].

Central aspects of transmission system planning

A principal aim of transmission planning is the addition of transmission lines and substations at the lowest possible cost without compromising security of supply [10], [11]. Furthermore, transmission system operators (TSO) are typically legally obliged to provide non-discriminatory network access. Thus, they enable the efficient functioning of wholesale markets [12].

If compared to distribution network assets, transmission network assets possess a higher capital intensity while asset lifetime may extend to 40 years and beyond [11].

Likewise, transmission network planning is typically a decade-long process, where upfront negotiation and licensing periods of 10 years are usual [11]. Transmission network planners forecast the evolution of demand, framing uncertainties even with scenarios that consider different, potential future market frameworks [5].

It should be highlighted that investment decision processes on transmission assets can strongly differ given the diverse jurisdictions those processes are embedded in. For example, in regulated markets, the transmission expansion problem (TEP) may consist of minimizing investment costs while assuring to meet demand and reliability target levels defined by a regulator or other state authority [10].

In a deregulated, unbundled market context, transmission companies may be incentivized to reduce investment costs while considering non-wires alternatives. Later market environments complicate transmission network planning as planners need to anticipate load evolution and generation expansion under limited information exchange between formerly integrated activities. In addition, non-overlapping planning horizons under market integration and sector coupling increase the complexity of respective optimization approaches [5], [11].

Other relevant aspects of transmission planning are listed below (as in [11]):

- Most countries use interactive transmission expansion methodologies. In the most common (e.g. European) case, a TSO submits a plan that is subsequently evaluated and approved by a regulatory authority.

- The construction/reconfiguration of lines or substations might be open to a competitive bidding process. For example, in Europe, project developers and investors can build merchant lines that connect regions. Such projects require authorization through the European Energy Regulatory Agency (ACER).

- From a modelling perspective, transmission expansion planning (TEP) is a multi-stage process with decision being taken at several time steps.

Although many studies use deterministic forecasts for the transmission expansion planning, generation expansion and generation cost uncertainties are usually incorporated in non-deterministic models.

Generally, TEP is a problem that can be addressed either through static, sequential static and dynamic planning models. Although dynamic planning represents the ideal approach for a long-term horizon, sequential static models are implementable at lower computational cost, data input and complexity.

With deregulation and the separation from power generation and delivery, further challenges to transmission companies appear. Market integration efforts (e.g. the European Energy Union) make interdependency assessments necessary and thus increase the system variables under optimization [11].

In addition, the consideration of DER in transmission planning becomes increasingly important. This is due to the impact of such resource on transmission network flows that can heavily influence the techno-economic optimality of TEP (e.g. through the temporal load reduction of HV/MV substations or changes in transmission lines losses, among others) [10].

Currently, recent literature surveys revealed that the majority of transmission studies do not include DER-related uncertainties [10], [13]. Likewise, transmission planning remains often constrained to the consideration of peak conditions and a single scenario for generation expansion [5].

Comparison of transmission and distribution planning aspects

Table 4.1. showcases differences between distribution and transmission planning. Those can be found in the cost of individual projects, the typical amount of assets to be examined as well as project selection practices and commissioning processes.

While distribution planners are concerned with assets that tend to serve relatively few customers, a major uncertainty lies in the prediction of spatial load patterns for a given time horizon. For transmission planners, there is a lower concern for spatial uncertainty, given that the longer project lead-times cause higher temporal uncertainty if compared to distribution planning. This is reflected in the coarser spatial resolution typically considered in transmission planning [4], [13].

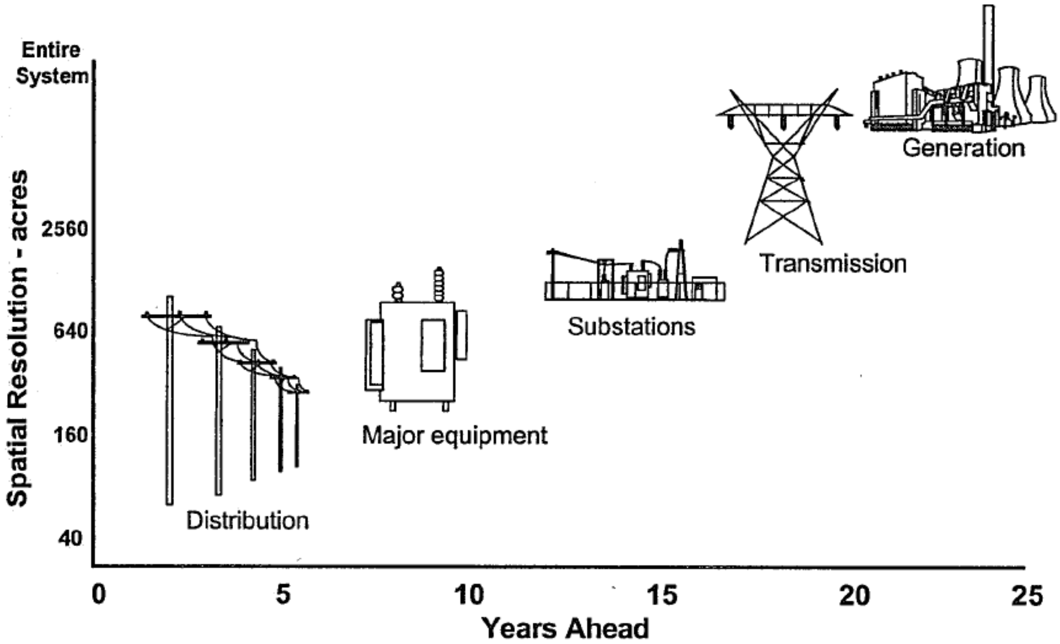


Figure 4.1 Planning horizon and spatial granularity in T/D planning (extracted from [4]).

The difference in transmission and distribution project lead-times and spatial resolutions considered for both planning processes are shown in Figure 4.1. Interestingly, the cited source suggests generation planning relying on the coarsest spatial resolution with the lowest lead-time (up to 25 years).

It should be expected that recent uptake dynamics of renewable power generators with relatively smaller generation capacities and unit costs (if compared to large fossil-fuel based power plants) [14] decreases average project lead-times and spatial planning resolutions.

The growing research dedicated to spatial aspects of generation planning might provide additional evidence to this hypothesis (e.g. [15]–[17]).

Table 4.1. Comparison of major transmission and distribution system characteristics (from [13])

	Transmission grid	Distribution grid
Primary drivers of system expansion	• Non-discriminatory access provision • Reliability • Reduce congestion • System stability constraints	• Non-discriminatory access • Reliability • Thermal and voltage constraints
Forecast horizon	15-20 years	1-10 years
Spatial forecast granularity	5 – 10 km^2	0.25 – 2.5 km^2
Typical number of customers served by primary assets	Billions to hundreds of thousands	Tens of thousands to single loads
Load flow	Balanced three-phase	Unbalanced three-phase
Investment cost per component	Relatively high	Relatively low
Commissioning process	Single projects with public hearings, env. impact assesses. and regulatory approval	Aggregated asset evaluation or indirect rate case evaluation by regulatory authority

Current challenges in power system planning

As the power sector is evolving, new challenges require the adjustment of traditional network planning and operation principles. In [5], the key challenges for power system planning separated along systemic boundaries have been summarized.

Such main findings for distribution and transmission planning are listed in the following:

- For distribution networks, rising levels of DER might trigger reverse flows (e.g. through PV during midday) or branch overloading (e.g. through EV charging), which require remedial actions [7], [18], [19].

- Both transmission and distribution system operators require new tools to hedge uncertainty in generation and load patterns. Former may arise under the variability of renewable generation under limited controllability [5], [7], [20].

- For both transmission and distribution system planners, the fusion of new data from various sources and formats and their feed-in into powerful analytical tools are required [7], [21].

- Additional cooperation to improve net-load forecasts at transmission-distribution interfaces is required [7], [22], [23].

In addition to the above-mentioned, the dynamic evolution of new generation and demand technologies on the consumer side add additional uncertainty to electricity network system planners [7].

In another study [7], the Electric Power Research Institute (EPRI) has listed 10 urgent challenges to electric power system planning. Such includes the incorporation of operational detail; increasing modelling resolution/ finer granularity; integrating generation, transmission, and distribution planning; addressing uncertainty and managing risk; improving forecasting methodologies; and improving modelling of customer behaviour and interaction, among others. Interestingly, most of the above-mentioned aspects are linked to the representation and modelling of uncertainty under DER adoption and its impact on investment decisions and risk.

The following sections of this chapter address such DER adoption uncertainty in electricity network planning.

4.2 UNCERTAINTY THROUGH DER ADOPTION REPRESENTATIONS

Sources of uncertainty in electricity network planning

In electricity network planning, uncertainty may be defined as the (unknown) divergence between a system state and a modeller's representation of that state. Uncertainty potentially grows, if future and past system states are to be represented - that is, if a temporal evolution to the system representation is added.

As the main function of electricity networks is to link electricity generators to consumers, the evolution of the generation as well as the demand side are of fundamental interest. While foresight of the technical and economic evolution of electricity network components themselves is relevant as well, latter aspects lies outside the scope of this work.

In general, power system-related uncertainty can be rooted to different aspects. Such are [24]:

- Fossil fuel price uncertainty;
- Environmental regulation and energy policy uncertainty;
- Load forecast uncertainty;
- Renewable power generation forecast uncertainty;
- Cost evolution of power system components;
- Distributed generation and storage business model evolution uncertainty.

Apart of such developments, unbundling and liberalization triggered a shift from fuel price to market price risks [24]. Deregulated markets tend to bring an increasing number of stakeholders that complicate system planning as conflicting interests and limited information flows are ubiquitous.

The above-mentioned restructuring of power systems saw dynamic increase in renewable energy installations across the world [14]. These resources heterogeneously impact on load profiles in time and space and such variable patterns tend to decrease certainty for system planners [5].

Eventually, new legislations (e.g. incentive programs) often trigger behavioural changes both at supply and demand side. With regard to the adoption of DER and other technologies, legislative changes represent another form of uncertainty as patterns of resource utilization may change. The position of uncertainty sources within electricity network processes is shown below (Figure 4.2).

In electricity network studies, it is sometimes argued that uncertainty is typically addressed in two ways [9]:

- Through scenarios;
- Through sensitivities.

Scenarios represent discretized load and generation forecasts and may summarize different political measures (e.g. low-carbon emission-oriented policy). In this case, it is usual to associate each scenario with a weight sometimes denoted as “subjective” probability, estimated from the judgment of experts.

Scenarios may also derive from a discretized representation of continuous probability density functions. This option is becoming more relevant with the need to represent new renewable generation such as wind or photovoltaic, but also the emerging phenomenon of electric mobility. In this case, actual scenario probabilities derive from the discretization process and true stochastic programming models may be built.

Sensitivities are analysed observing changes in model outputs after changing key constraints or data. These variables are determined beforehand and their effect on final outcomes is estimated keeping the other model parameters constant. Sensitivities represent constrained partial derivatives and may be associated, in mathematical models with derivatives, to dual variables of Lagrange multipliers. If no derivatives are available or one is trying to assess the effect of major changes, then sensitivities may be associated to discrete step changes.

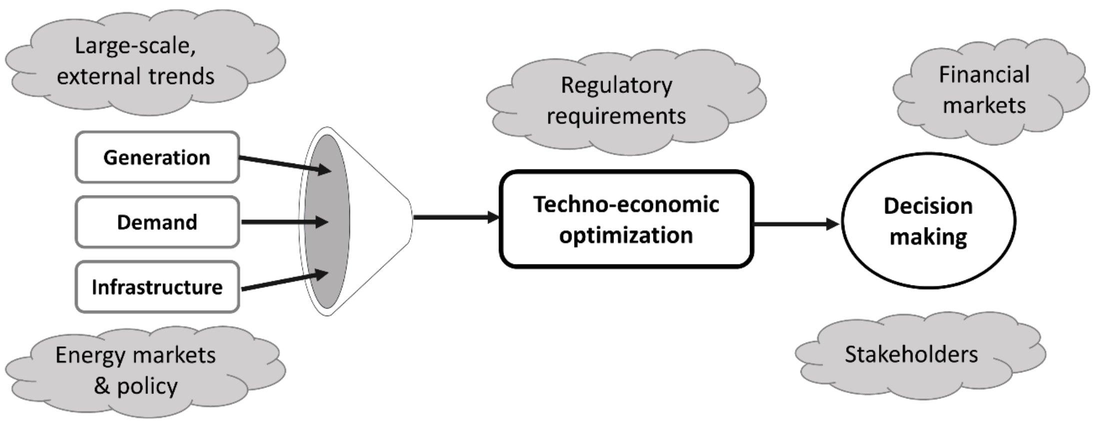


Figure 4.2 Uncertainty sources in T/D planning and investment decisions (based on [13]).

Alongside the adoption of DER, new uncertainties occur. They relate to the evolution of load and generation patterns in space and time. Until recently, electricity network planners had few tools to describe such patterns. Even more, attempts to model the impact of DER on electricity networks have widely neglected data-sets from other areas (e.g. economics , demography, human geography) relevant to DER adoption.

While the development of spatiotemporal DER adoption models have been presented earlier within this thesis, the following section summarizes current approaches employed in research and industry. It will be shown that such standards cannot capture the planning uncertainties introduced by DER adoption dynamics.

From the perspective of an electricity network planner, location and adoption time of DER are important inputs to estimate the grid impact of such technologies. As the peak-load estimate is a fundamental planning criterion [6], the planner further needs to assess the likely impact of adopted resources on such peak value [25]. Consequently, a rising amount of research has been dedicated to studying the potential impact of DER. Especially, distribution networks have been assessed [26]–[31], as these have been more vulnerable to adverse effects that DER may bring [32][9].

Until today, very simplified techniques to represent DER uptake dynamics across test networks have been in use. Such approaches relied either on random allocation or extrapolation techniques (Table 4.2.).

For example, the contributions of [29], [30] assessed the impact of the utilization of EV and PV on test distribution systems. While the effects on distribution networks largely depend on the quantity and position of such resources along distribution feeders [18] resources have been allocated equally across the case studies' network. Likewise, there have been presented some studies that assessed the grid impact of distributed PV generation using completely randomized allocations of PV systems across the test network under study [31].

Similar approaches have been developed, simplifying DER adoption dynamics through linear extrapolations using installed capacities or peak loads at busbar or transformer locations. Such studies have been conducted for both EV and PV technologies [26]–[28].

The mentioned studies show that, currently, DER are modelled using randomization or very simplified methods that rely on extrapolations, equal shares or synthetic probability distributions. Consequently, such grid impact studies cannot properly represent non-linear adoption behaviour across time and space. This is especially noteworthy as first studies suggest DER adoption clusters during an early uptake phase [33].

Therefore, due to the static character of current approaches to model DER adoption, heterogeneous adoption patterns that relate to underlying socio-demographic structures cannot be represented. Hence, the use of more granular geodata (e.g. georeferenced census data-set), allows to enrich network data with very granular consumer information that could complement circuit-based power flow analysis with spatial forecasts. Consequently, DER adoption scenarios can be enhanced while uncertainty decreases.

Table 4.2 DER allocation methods in distribution networks

Methodology	Technology	Reference
Deterministic:		
- Extrapolation using busbar capacities	EV	[27]
- Extrapolation using peak demand	EV/PV	[28],[26]
Randomized:		
- Single-step random allocation (e.g. equal assignment)	EV/PV	[29], [30]
- Multi-step iterative allocation (e.g. using Monte Carlo)	PV	[31]

The hierarchical dimension of DER representations is shown in Figure 4.3. In general, the input data requirements increase towards consumer-level DER adoption forecasts. For example, relevant consumer data may consist of census information (social and demographic structure), house value or income information (purchasing power) and electricity consumption data (contracted power, annual consumption, peak demand). Given the multitude of data sources to be combined and model alignment to such specific source, model transferability is expected to decrease.

Until now, representations that consider the network's underlying population structure or even consumer structure have been mostly neglected. Noteworthy exceptions are the spatiotemporal models presented in Chapter 3 (e.g. [34]–[37]).

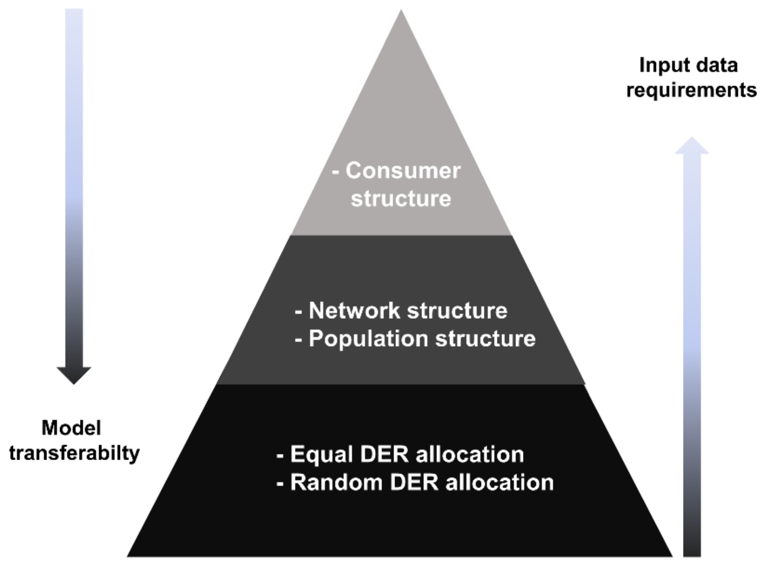


Figure 4.3 Granularity levels of DER adoption representations in T/D planning.

4.3 DISTRIBUTION SYSTEM PLANNING UNDER TECHNOLOGY DIFFUSION

Introduction

The strong uptake of DER connected to distribution networks calls for an actualization of the planning tools currently in use [6]. Therefore, research has been dedicated to enhanced distribution system expansion models in the past decades. The aim of most of these tools is to address the growing need for planning tools that can cope with the uncertainties introduced by DER adoption.

Proposals to adopt multi-stage distribution network expansion planning under multiple criteria and resorting to meta-heuristics may be identified as early as 1993, with the seminal paper [39]. The actual industrial implementation of such techniques waited for advancement in geographical representations and more computing power. An excellent example is referred to in [40], for a model adopted by the Portuguese distribution utility EDP Distribuição. Recent advancements in distribution network expansion included planning models stretching over multi-stage horizons, incorporating load growth and energy prices uncertainties [25], [38], using probabilistic approaches or search algorithms (e.g. Genetic Algorithms) [41], [42]. In addition, new models that integrate optimized DER placement in distribution networks have been proposed [43]–[45].

However, missing knowledge about technology diffusion processes and partially, limited data availability, have hindered the development of realistic DER diffusion forecast models.

As shown earlier, electricity network planners still rely on very randomized or equal-share allocations of DER in distribution grids. It should be expected that the use of these approaches result in very misleading conclusions as DER are adopted by well-studied, distinct population groups [46]–[49] – a fact that explains why research found clustered, heterogeneous spatial DER adoption patterns [50] [33].

Hence, electricity distribution planners should use planning models that can provide spatiotemporal DER adoption forecasts. Results of these models should take into account potential clustering of early DER adopters while providing insights in the net-load effects of new appliances in residential environments.

Such tools would allow grid planners to screen all HV/MV substation service areas of a given distribution network service area for potential overloading, before detailed power-flow analysis would be required.

The methodology presented in this chapter compares the developed spatiotemporal DER diffusion forecast model to traditional approaches that have been used to allocate such resources in distribution networks. Exploiting a very granular census data, a set of alternative peak-net-load scenarios has been developed.

it is important to assess the distribution network planner's risk to over- or underestimate the grid impact triggered by the adoption of distributed energy resources. The proposed methodology consists of five subroutines, identified in Figure 4.4..

In order to understand the essence and the application of the methodology, it will be described through its application to a case study: the distribution network in the city of Porto, Portugal. This will be done without loss in the generality of the methodology proposed.

The coarse steps of the methodology are summarized below:

i. A spatial routine computes HV/MV substation service areas (eight, in the case of Porto Municipality), using transformer locations (XY coordinates). HV/MV transformer service areas are retrieved using Voronoi diagrams.

ii. A geolocation routine relates census polygons with varying spatial extent and population characteristics to HV/MV service areas.

iii. The spatiotemporal DER diffusion forecast model that generates EV and PV adoption patterns using census data as input.

iv. A net-load analysis that considers hourly load, EV charging and PV generation time series.

v. A risk analysis routine compares the outputs for different DER adoption forecast techniques, analysing hourly net-load time (***NL***) series (8760 h) for each HV/MV transformer.

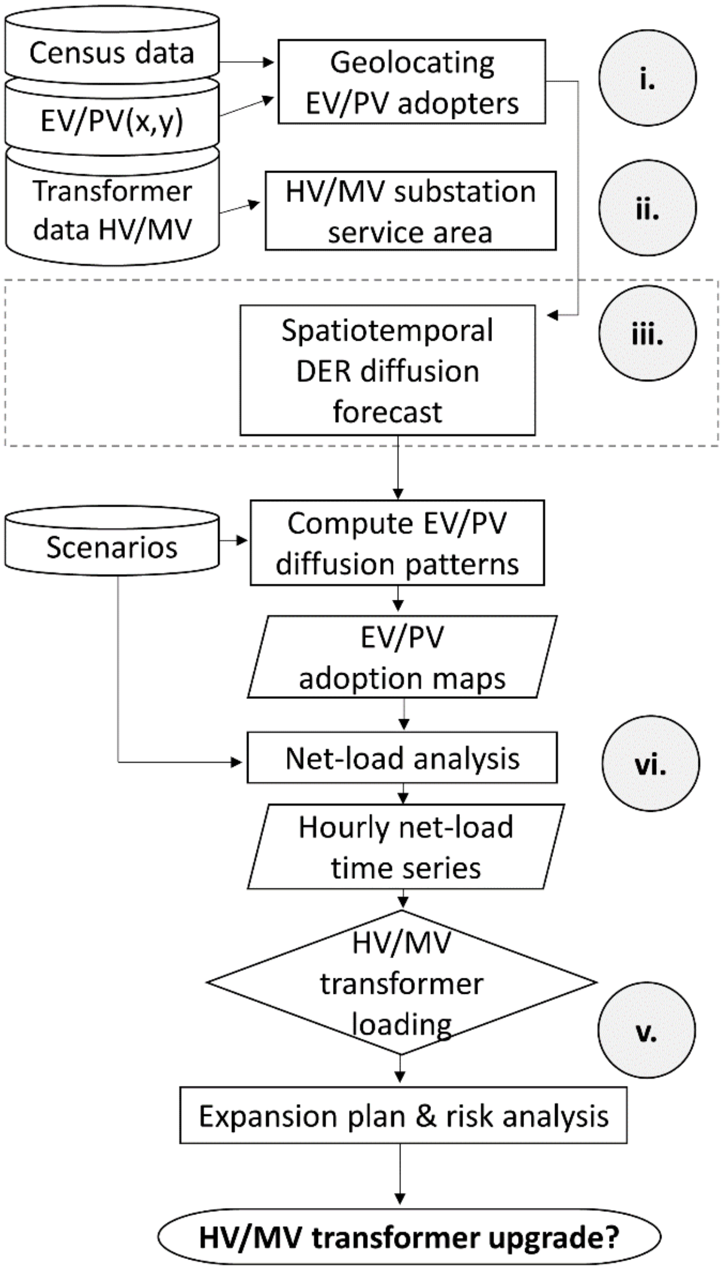


Figure 4.4 Distribution grid expansion planning using spatial net-load forecasting

It is expected that the proposed methodology can contribute to a better understanding of the way DER adoption will affect distribution grid expansion decisions. Given that the model consists of flexible sub-routines, a high transferability to any other appliance or modelling time horizon of interest is guaranteed. Regarding the spatiotemporal DER adoption model introduced earlier (Chapter 3), the analysis presented in this chapter relies on additional data sources. Such are:

- **Demand growth and DER adoption scenarios**. Former load growth values ($\boldsymbol{\alpha}$) have been extracted from the national distribution system operators' investment plan [51]. The plan foresees load to grow around 1-2% annually. As some service areas in Porto Municipality receive new Metro connections in the coming years, load growth values at the upper boundary of that range have been allocated (**).

  Likewise, global DER forecasts have been provided in [52]. This report contains consistent storylines that have a European scope and have been defined considering multiple stakeholders from the electricity industry, politics and other organizations.

  The three scenarios used in this work (*Baseline*, *High growth* and *Very High growth*) resemble the scenarios *Sustainable Transition*, *Global Climate Action* and *Distributed Generation* of the mentioned report. Values have been downgraded to Porto municipality given the population share calculated from [53]. It is assumed that population count and structure will remain stable over the analysis horizon ($\boldsymbol{ys}$). The resulting scenarios for Porto municipality are shown below (Table 4.3).

Table 4.3 Scenarios towards 2035

	Baseline	High Growth	Very High Growth
Annual peak load growth**	0% /1%	1.5% /2%	1.5% /2%
Added EV [in 1000]	5 (6%)	31 (34%)	45 (50%)
Added PV [MW]	26 (20%)	41 (33%)	78 (70%)

** Higher values applied to service areas that expect additional Metro connection (Campo Alegre, Vitoria). As has been shown in [4], [54], access to mobility infrastructure tends to increasing load growth.

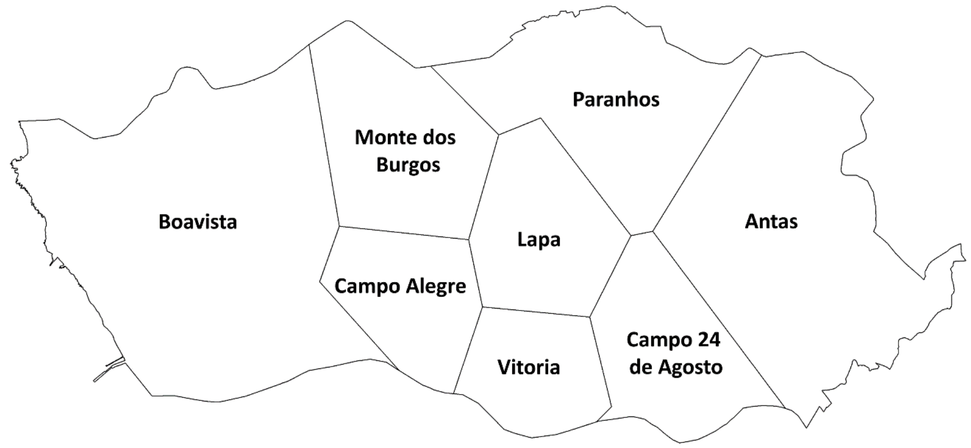


Figure 4.5 Approximated HV/MV transformer service areas in Porto

- **HV/MV substation service areas.** The spatial extent of high voltage/ medium voltage (HV/MV) substation service areas have been approximated through a Voronoi diagram. That way, service areas are built around each service area centre (HV/MV substation). The underlying algorithm maximizes the spatial extent of each service area until interfering with the neighbouring substations' service area. For Porto Municipality, the resulting diagram for its eight HV/MV substation is shown in Figure 4.5.

  A more detailed explanation of the underlying mathematical model as well as an earlier application to HV/MV substation planning is available in [55], [56]. All substation characteristics (coordinates, winter peak loads' and installed transformer capacities (***TCAP***)) were retrieved online [57]. Confronting retrieved service areas to their real extent [51] suggests a high congruence of the developed approach, which shows its adequacy.

- **Time series and other model parameters.** Hourly load profiles in a typical Portuguese MV network and PV generation profiles have been retrieved from [58] and normalized using max-min value ranges. An EV charging profile has been obtained by a Portuguese mobility solution developer. MV load, EV and PV time series are stretched to the forecasted peak load and residential EV charger and PV module capacities in each service area (in MW). For the EV model, the analysis considers two virtual charging power rates: First, for overnight charging events, an average of 5.89 kW per EV connected is assumed. This value aggregates 30% of EV adopters that would charge with 12kW and a vast majority (70%) with 3.7kW. On the other hand, midday charging is assumed with 9.82 kW per EV connected (assuming 10% charge with 40kW, 30% with 12kW and 60% with 3.7kW).

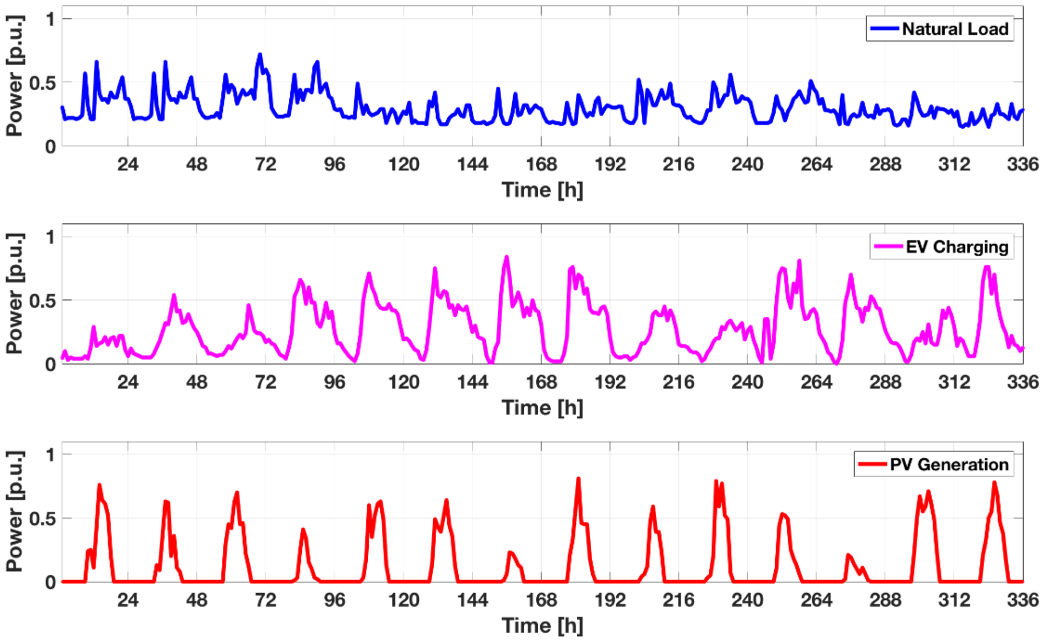


Figure 4.6. Hourly MV-load, EV charging and PV generation time series covering 14 days.

Forecasted DER adoption patterns

At this this step of the analysis, the outcome of the spatiotemporal DER adoption forecast model is employed. It should be recalled that the result of this model is a set of maps which represent spatial distributions of newly added loads/generators in a given population. Figure 4.7. displays added EV charging power (in kW) and PV module capacities (in installed kW) per census cell. EV charging patterns differentiate midday charging (a), considering charging at the workplace as well as incoming commuters [50] and overnight charging at the adopter's residences (b). Outcomes show results for Porto Municipality using the proposed diffusion forecast methodology (*Baseline scenario*) by 2035 and suggest divergent spatial patterns for both EV and PV adoption (c) and EV midday (a) and overnight charging events (c).

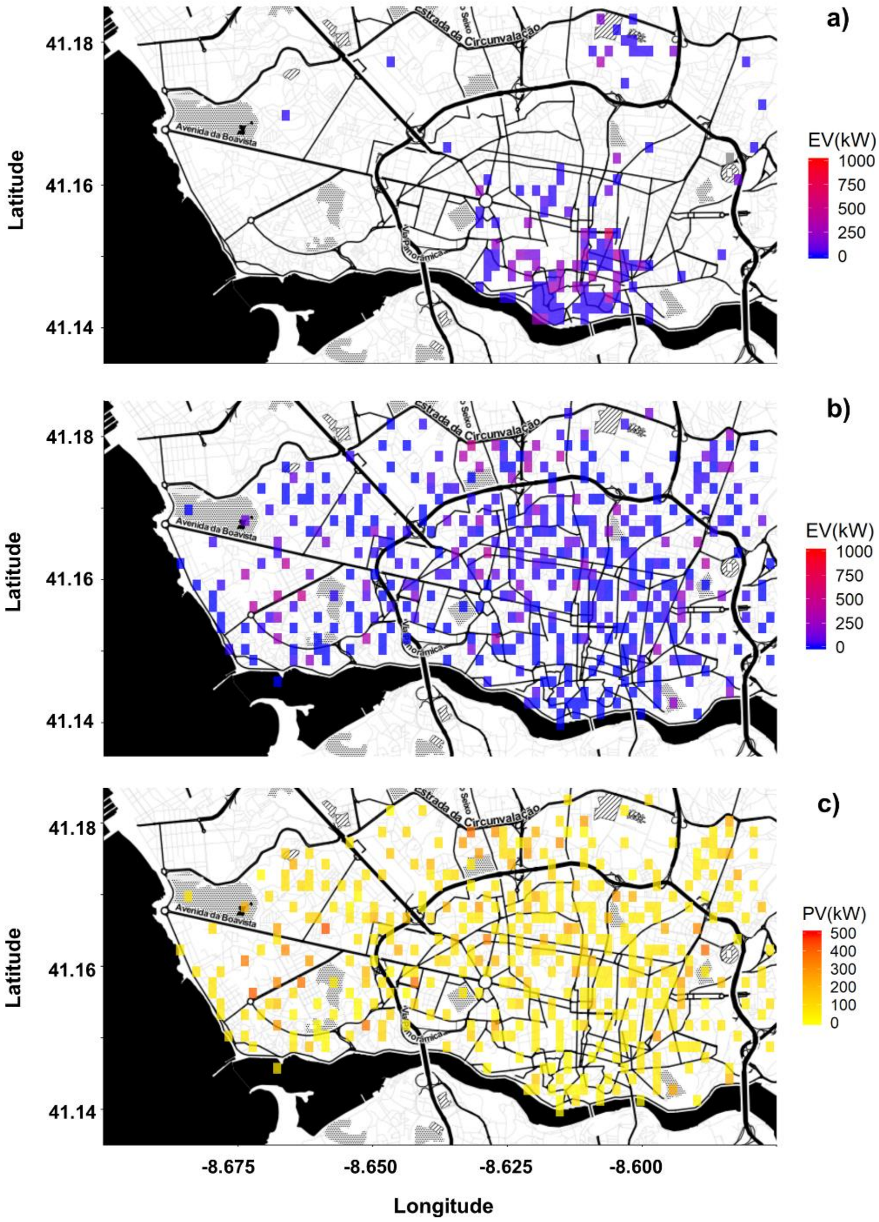


Figure 4.7. Expected midday EV charging peak (a), overnight EV charging peak (b) and roof-integrated PV peak generation (c).

As a general tendency, residential PV and overnight residential EV charging show a more dispersed pattern. In contrast, midday charging events concentrate around commercial zones in the inner-city centre. Because of the strong spatial difference in overnight and midday EV charging patterns, increases in HV/MV substation peak loads will occur in a very heterogeneous way. PV adoption patterns demonstrate a more homogeneous distribution similar to EV overnight charging events instead. It is noteworthy that PV and EV adoption patterns do not fully overlay, suggesting asymmetric effects on the net-load of each service area (Figure 4.7.). Given the additional imbalance of EV/PV capacities forecasted, very different absolute spatial EV/PV concentrations (in kW) are achieved.

Net-load calculation

As in [19], this work assesses netting effects of DER. In our study, we consider EV charging and PV generation as addition and subtraction to the natural load respectively. We define the natural load as the utilized load pattern of a medium voltage (MV) network curve unaffected by future DER adoption.

Aggregating forecasted DER capacities and natural load profiles as a single curve per substation service area, we retrieve the net-load combining EV charging (***EV***), PV generation (***PV***) and natural load (***L***) (Eq. 7) for each time step. Therefore, the net-load (***NL***) for each hour (***h***) in a given service area (***sa***) is:

$$NL_{sa,h} = L_{sa,h} + EV_{sa,h} - PV_{sa,h} \tag{4.1}$$

Ideal load flow conditions are assumed, given that the purpose of this work is to analyse the impact of large-scale technology adoption. Therefore, LV or MV network losses or other constraints are neglected. Typical hourly load profiles of a representative Portuguese MV network (lp_h) and PV generation (gp_h) were normalized and stretched to installed PV capacities, EV charger capacities and peak demand in each service area (Figure 4.5).

EV (cp_h) time series was multiplied with the aggregated EV charging power (in MW) of all census cells in a given service area. Different charging behaviour for midday (at work) and overnight (at the residence) have been assumed. While a truncated charging curve for the period between 6am and 8pm was used for midday charging, the remaining hours were truncated for overnight charging time series. Eventually, hourly natural load ($L_{sa,h}$), EV charging ($EV_{sa,h}$) and PV generation ($PV_{sa,h}$) time series (***t***), all in MW, were derived for each service area (***sa***):

$$L_{sa,h} = lp_h \times PL \times (1 + \alpha)^{ys} \tag{4.2}$$

$$EV_{sa,h} = cp_h \times (N_{EV} \times s \times cr)/1{,}000 \tag{4.3}$$

$$PV_{sa,h} = gp_h \times (N_{PV} \times pr) \tag{4.4}$$

Natural peak loads have been updated from the studied reference year (2015) using load growth factor (α) over the forecasted time horizon of 20 planning years (ys). Natural load time series originated in the multiplication of a normalized hourly load profile (lp_h) with peak load values (PL) per HV/MV transformer. Added EV charging load per service area ($EV_{sa,h}$) was retrieved through the multiplication of EV adopters (N_{EV}) with a charging simultaneity factor of (s), overnight (5.89 kW) or midday (9.82 kW) charging rate (cr) and the normalized EV charging profile (cp_h).

On the other hand, PV generation times series of each service area ($\boldsymbol{PV_{sa,h}}$) were constructed by multiplying panel capacities ($\boldsymbol{N_{PV}}$) with a typical performance ratio of 0.8 similar to [59] ($\boldsymbol{pr}$) and a normalized generation profile.

As this study considered the netting effects of DER on each service area's natural load curves, this work introduced the peak-net-load ($\boldsymbol{PNL}$) as planning criteria. For each service area ($\boldsymbol{sa}$) and scenario ($\boldsymbol{sc}$), PNL is calculated such as:

$$PNL_{sa,sc} = \max NL_{sa,sc,h} \tag{4.5}$$

Illustrative results of the net-load analysis are shown in Figure 4.8. It displays load duration curves for all scenarios (a) and comparing four DER allocation techniques earlier discussed (b). Figure 4.8.a shows how increasing DER adoption towards the *Very High Growth* (VHG) scenario shift the net-load duration curve upwards. Furthermore, outcomes suggest overloading for the VHG scenario. On the other hand, comparing the evolution of net-load behaviour among all DER allocation techniques (diffusion model, extrapolation of DER quantities using relative ratios of HV/MV transformer capacities and peak-loads or equal shares) unveils a strong difference of the spatiotemporal diffusion model to the remaining.

While one expects more than 500 hours of HV/MV transformer overloading by 2035 using a spatiotemporal DER diffusion model, DER allocation based on extrapolation would suggest no capacity constraints. This is significant risk to network planners that rely on simplified extrapolation-based DER allocation, as such techniques might level-out adverse effects that concentrated DER adoption might cause.

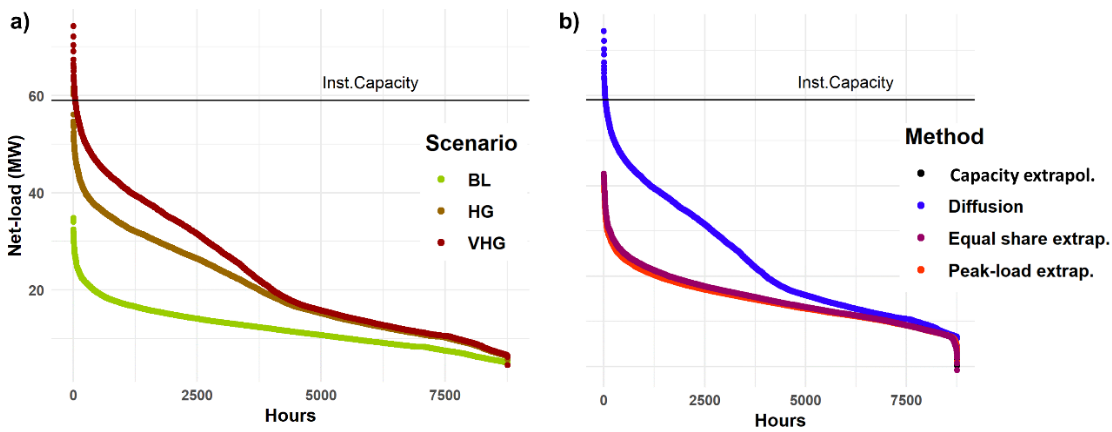


Figure 4.8. Net-load duration curves for the Vitoria area in Porto by 2035, using diffusion forecasts (a) and the discussed methods for Very High Growth conditions (b).

Furthermore, we are interested in assessing the effect of different DER allocation techniques on expansion cost estimates. In order to assess the network expansion cost, we calculate the cost of HV/MV transformers that would require capacity expansion/replacement under each scenario. A standard threshold of the Portuguese distribution system operator to upgrade HV/MV substations has been considered. If the ratio of peak-net-load (***PNL***) to installed transformer capacity (***TCAP***) exceeds 0.9, network upgrades are triggered. For the Portuguese case study, a typical upgrade cost of roughly 2 million € per HV/MV transformer is used [51]. Thus, the cost required for network expansion (in the following called investment costs (***IC***)) for each HV/MV substation service area (***sa***) within all scenarios can be approximated with the following formulation:

$$IC_{sc,sa} \begin{cases} 2{,}000{,}000€ \;\; for \max\limits_{sa,sc} PNL \; \geq 0.9 * TCAP_{sa} \\ 0€ \; for \; \max\limits_{sa,sc} PNL \; < 0.9 * TCAP_{sa} \end{cases} \tag{4.6}$$

Given the sum of all HV/MV service area's investment costs for a given scenario (***sc***), one can now calculate the value of the investment plan (***IP***) that would address each scenario's investment needs. Thus, transformer upgrade cost is seen equivalent to overall network expansion cost, although the former typically represent only a share of investment needs [60].

$$IP_{sc} = \sum_{sa=1}^{8} IC_{sc} \tag{4.7}$$

In the presented analysis, a radial network configuration has been assumed while effects of line losses and detailed power flow calculations are excluded from the scope of analysis. It should be further noted that the presented investment analysis does not intend to fully replace a detailed cost analysis, but rather provides indication on the effects of various DER adoption dynamics. Comparing the expansion costs under a base case (***s*** = 0.5) and using a spatiotemporal DER diffusion model, one would plan for a 4 and 6 million € ***IP*** for High growth and Very High growth scenarios, respectively. In case the system would have evolved according to this forecast, network planners using extrapolation-based DER allocation would have underspent 2 and 4 million €, respectively (Figure 4.9.b).

If investment costs are compared for higher simultaneity factors, which is equivalent to higher EV charging rates or higher adopter numbers per HV/MV transformer, equal-share or extrapolation-based DER allocation overestimate network expansion costs. A reasonable explanation is that spatiotemporal DER models predict adoption patterns based on population structure.

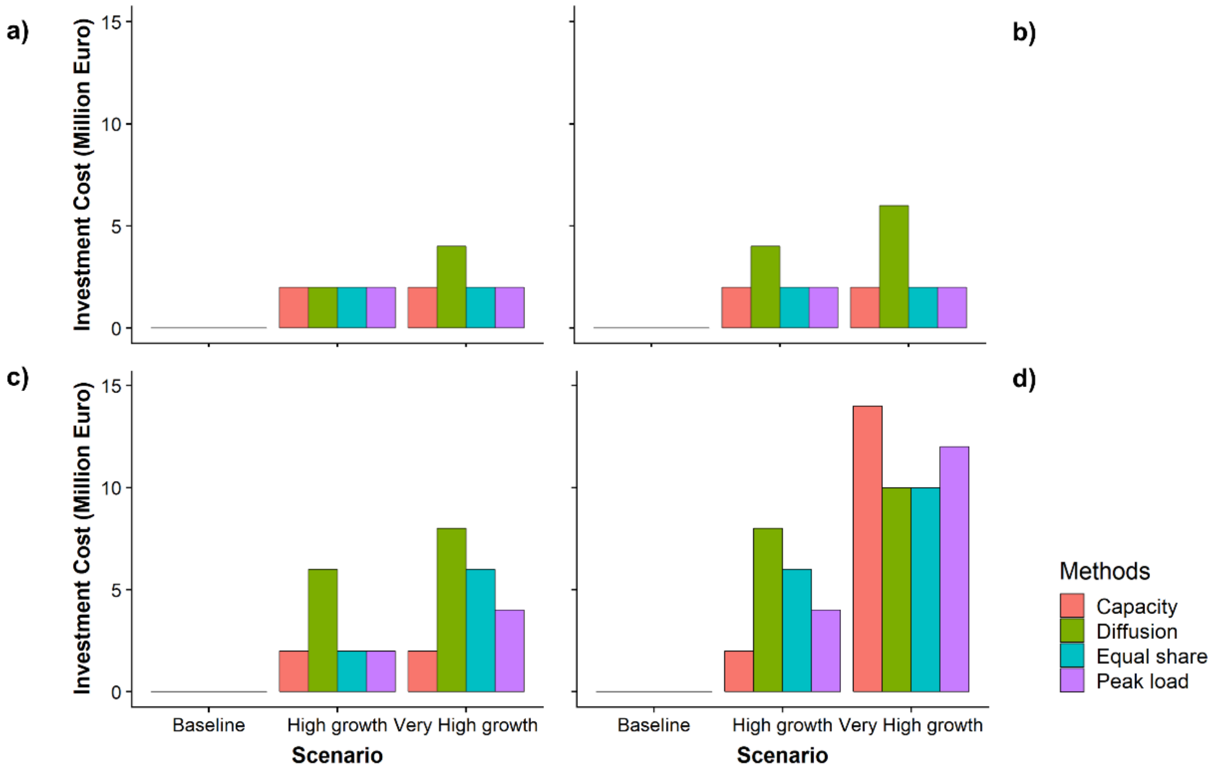


Figure 4.9. Investment costs using different DER allocation methods and charging simultaneity rates, with a) s=0.25, b) s=0.5, c) s=0.75, d) s=1.0.

Typically, urban, population-dense HV/MV service areas (such as in Porto municipality) possess higher capacity overheads, which allow for higher DER penetration before upgrades are necessary. On the contrary, equal-share or extrapolation-based DER allocation fills service areas with DER disregarding the internal demand structure, and potentially, lower overcapacities installed.

As a general outcome, results show that the spatial variability of the technology adoption processes leads to underestimations of the grid impact of DER. For peak-net-load situations, current state-of-the-art uptake forecasts of EV chargers and roof-integrated PV panels lead to too conservative net-load estimates if compared to the proposed diffusion forecast (Figure 4.8). In fact, due to clustering of early EV and PV adopters, installed transformer capacities might be surpassed in certain neighbourhoods where current state-of -the-art DER deployment forecasts based on extrapolation would still indicate available excess capacity.

Finally, the peak-net-load of all service areas, considering the three scenarios, have been analysed. Results suggest that conventional approaches are incapable to predict adoption clusters (e.g. Vitoria or Campo 24 de Agosto) that are foreseen if a spatiotemporal DER adoption

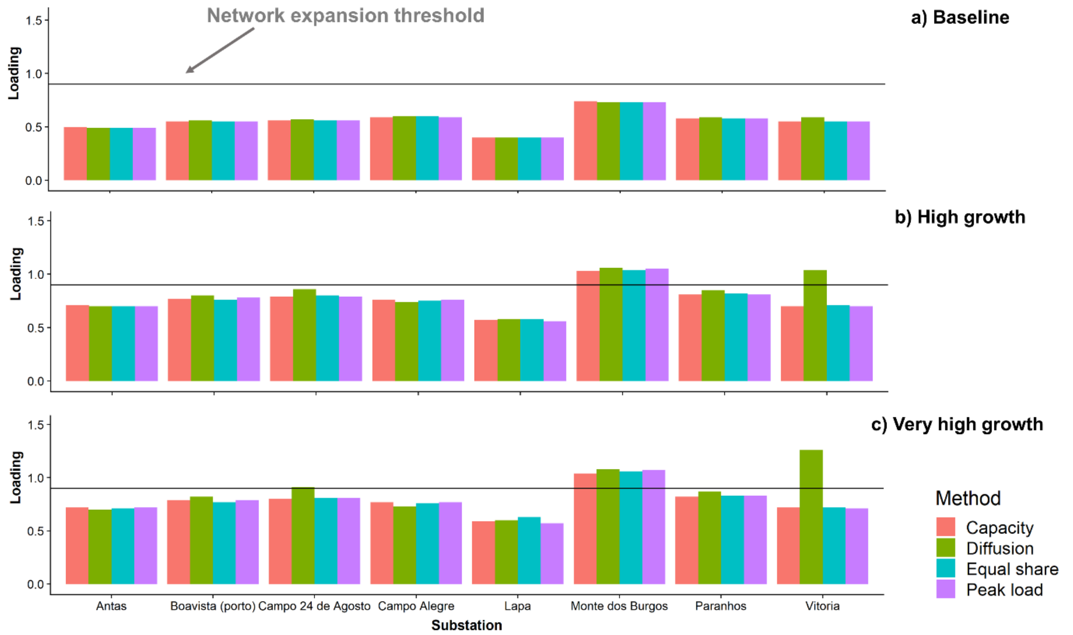


Figure 4.10. Comparison of HV/MV transformer loading using the presented approaches and across all substations, methodologies and scenarios analyzed (with s=0.5).

model is used. If latter models are more accurate (as suggested in Chapter 3), conventional extrapolation-based or equal-share DER allocation models tend to underestimate network expansion costs (Figure 4.10). Instead, if stronger EV uptakes (or higher charging rates/simultaneity factors) are considered, such approaches might overestimate expansion needs (Figure 4.9.).

Uncertainty assessment

As the presented approach consists of the forecasting of spatiotemporal technology adoption patterns under limited historical observations, traditional forecasting validation is not convenient. In Chapter 3, different approaches to assess locational uncertainty in the spatial module and temporal uncertainty (time-step discretization) have been presented.

While the focus of this chapter lies on the interaction of DER adoption and network planning, EV and PV generation uncertainties have been additionally assessed. By altering parameters that impact their time series a sensitivity analysis can be provided.

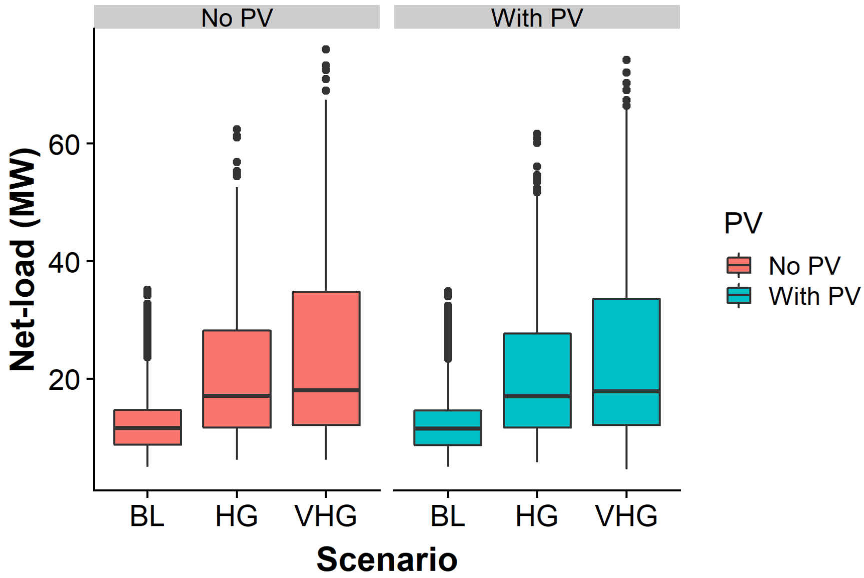


Figure 4.11. Net-load variability for all three scenarios (diffusion forecast)

To assess the sensitivity of ***PNL*** **(*PNL*** is the peak-net-load) to EV module parameters (e.g. simultaneity rate, charging rate), these have been altered and outcomes compared. Changes in the simultaneity rate (0.5) to 0.25 and 0.75 result in a change of approximately 30% in net-peak-load. Furthermore, the effects of PV module inclination changes to the evolution of peak net-loads have been assessed. Here, simulations suggest very light impact on ***PNL*** estimates, showing that panel orientation changes (full North or full South) may result in roughly 30% changes in the PV peak generation upwards and downwards, respectively. A detailed comparison of different module technologies and the likely evolution of efficiency lay outside the scope of this work.

Finally, the net-load variability for Vitoria service area has been assessed. Figure 4.11 provides evidence that load variability will rise during the 20-year planning horizon, which strong increases under High growth and Very high growth scenarios. Considering the 95% confidence interval, load variability rises from less than 10 to over 30 MW respectively.

According to the analysis, variability is expected to reach roughly 25 MW under the "Very High Growth" scenario. Comparing peak-net-load evolution along scenarios with or without PV (in case of a temporary unavailability due to a weather event) provide insights in the origin of the variability. Here, outcomes suggest only a minimal increase in net-loads of approximately 5% in case of PV unavailability (Figure 4.11.). This could be explained by the non-coincidence of PV generation time series and EV charging time series and the charging behaviour considered (dumb charging). Therefore, increased EV charging and load growth remain most likely influencing factors for the in net-load variability observed.

4.4 TRANSMISSION SYSTEM PLANNING UNDER TECHNOLOGY DIFFUSION

Vertical load flows at the transmission/distribution interface

One important challenge to transmission and distribution system operators under strong DER uptake is the development of new tools that can hedge uncertainty in generation and load patterns [5]. As DER tend to be installed at distribution grid level, an improved cooperation to improve net-load forecasts at transmission-distribution interfaces is required.

As centralized generation usually connected to higher voltage levels (transmission) is progressively complemented by low-/medium voltage-level DER installations, a strong DER uptake is assumed to impact on the vertical load exchange at the distribution-transmission interface [11].

As argued in [61], [62], such uptake of DER that embeds in distribution networks may require increasing coordination of distribution and transmission expansion planning. Fostering data exchange between TSOs and DSOs may eventually reduce planning uncertainties [61]. An outlook how concerted transmission and distribution planning could be conducted is shown in Figure 4.2 (based on [61]). It is noteworthy that under unbundling, such integration of transmission and distribution planning is facing additional, institutional and legal complications that may be absent in vertically integrated electricity businesses as present in the US.

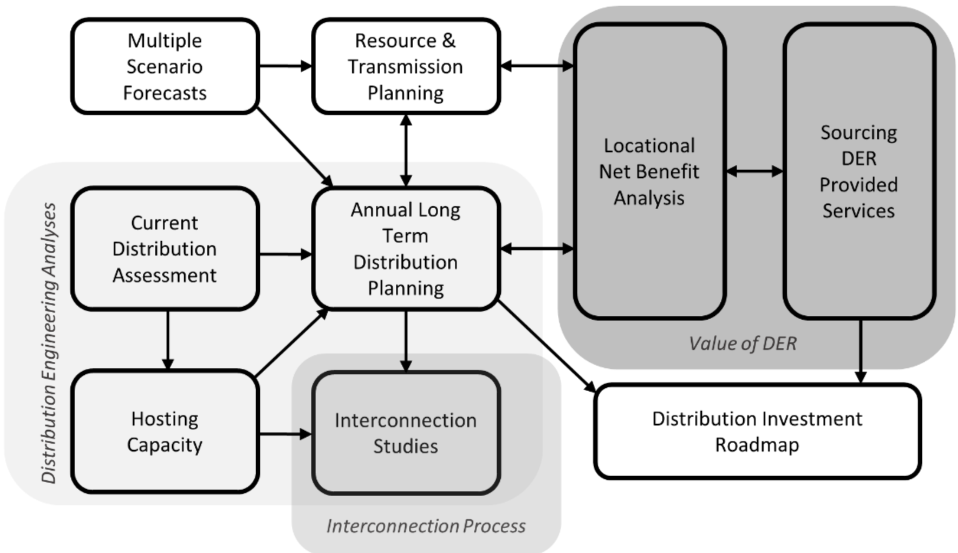


Figure 4.12 Integrating distribution and transmission system planning (inspired by [61]).

Recently implemented unbundling policies led to a separation of vertical integrated utilities. Therefore, transmission and distribution system planning has been separated. Consequently, transmission system planners do not have full oversight of distribution network configurations and development plans, resulting in additional uncertainty in transmission grid planning [5].

In this chapter, the proposed spatiotemporal DER diffusion model is applied to a use case on transmission network expansion planning under DER diffusion. The reader may recall that the proposed spatiotemporal DER diffusion model grounds on a high-resolution census data-set with over 17,000 census cells for Continental Portugal. The spatiotemporal model produces forecasted EV and PV adoption on a 20-year horizon. Model outcomes are net-load curves for each of the distribution network's HV/MV transformer service areas. In this chapter, these HV/MV transformer service areas are aggregated into transmission service areas, which are the areas served by one distinct transmission entry point. These outcomes allow transmission system planners to analyse the vertical load diagram between the distribution and transmission interfaces under various DER adoption scenarios.

In order to understand how uncertainties of the representation of DER adoption patterns propagate from HV/MV substation service areas towards transmission service areas, we analyse retrieved DER forecasts across four typical DER allocation techniques commonly applied for grid impact studies. These allocation techniques have been previously introduced (Chapter 4.1.2).

Finally, the chapter presents two new planning criteria that can provide transmission system planners additional indication, providing insights into the system's net-load behaviour at the transmission/distribution interface (T/D).

A spatial model of the transmission/distribution interface

One innovation provided in this chapter is the geometric approximation of transmission service areas. This is achieved through combining the following two data-sets: A georeferenced HV/MV transformer positions (Figure 4.13.a) that is accessible through the major Portuguese distribution system operator [57] and tabular information of the linkage of the HV/MV substation transformers to transmission entry points (retrieved from [60]).

The HV/MV transformer service areas have been calculated passing the following steps:

- As first step, HV/MV substation service areas are approximated through a Voronoi diagram as in [56] (Figure 4.13.b));
- Retrieved spatial service area polygons are merged into transmission service areas, based on the information of their connection to each transmission entry point, available in [60];
- Resulting polygons are related to the Portuguese census data-set through spatial intersection (Figure 4.13.c).

This way, each transmission entry service area is linked to detailed information of the population subgroup, assumed fed by each specific transmission entry node.

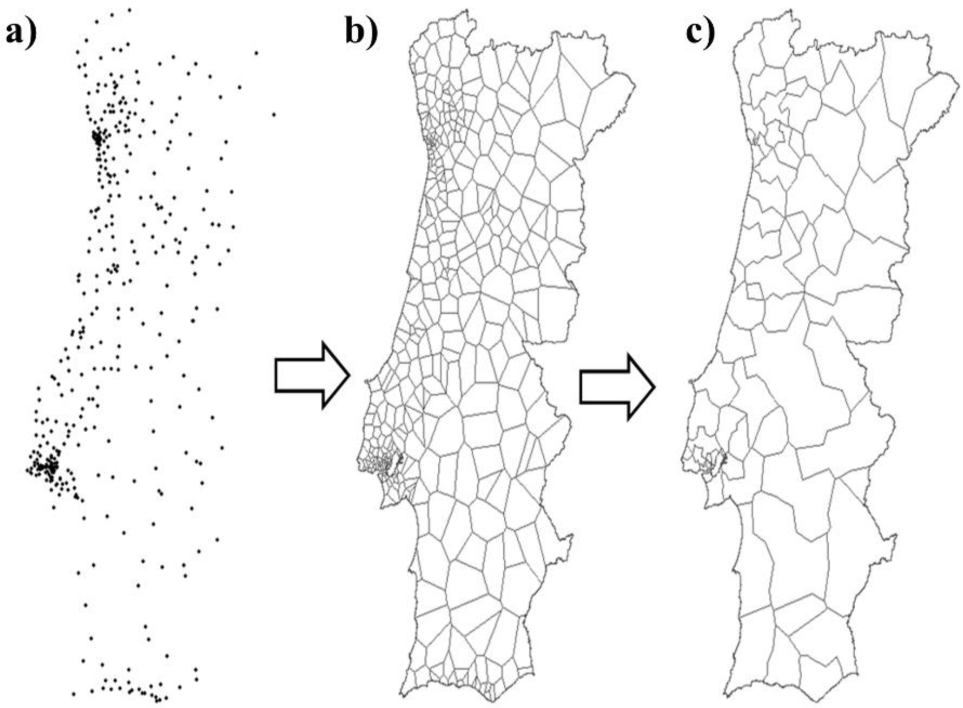


Figure 4.13. Retrieval of spatial T/D network service areas

In the continental Portuguese Distribution network system, there are 391 HV/MV substations that can be linked to 63 transmission entry points [57], [60]. Former data-set provides detailed information that allows for a characterization of HV/MV substations that is sufficient to support long-term network expansion planning (e.g. the location, peak load, installed capacity are provided).

The developed methodology is again illustrated with the Portuguese case and consists of five steps. Outcomes are vertical load diagrams that are computed and analysed for each of the 63 entry points to the transmission network. The first four steps are similar to the methodology presented in Chapter 4.3.

i. First, a geolocation routine relates census polygons with varying spatial extent and population characteristics to HV/MV service areas.
ii. Second, a spatial routine computes 391 HV/MV substation service areas using transformer locations (XY coordinates). HV/MV transformer service areas are aggregated to each T/D connection point.
iii. The spatiotemporal DER diffusion forecast model generates EV and PV adoption patterns using census data as input.
iv. A net-load analysis that considers hourly load, EV charging and PV generation time series.
v. Vertical load flow analysis routine that uses the hourly net-load time (***NL***) series (8760 h).

A flowchart of the process is provided in Figure 4.14.

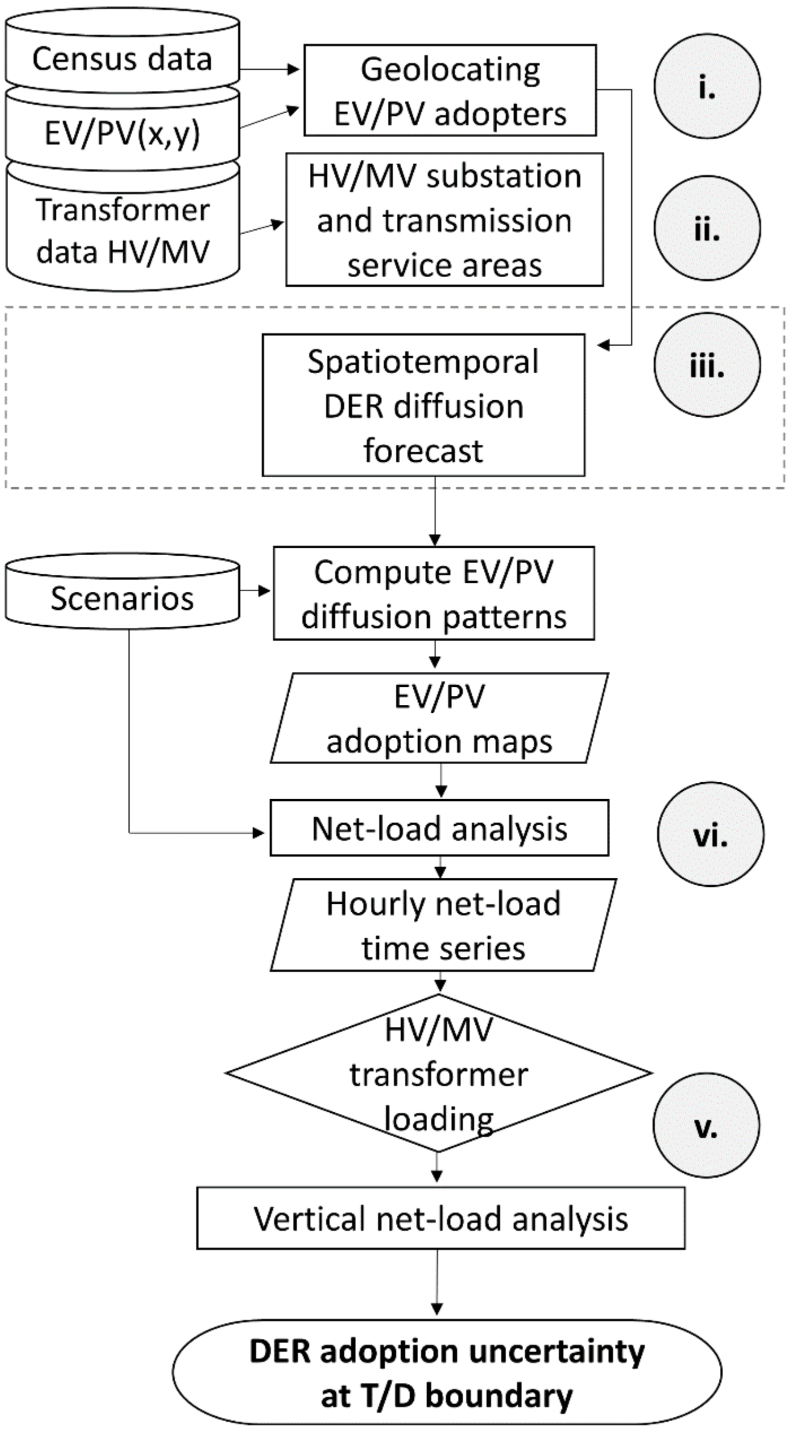


Figure 4.14. Transmission grid expansion planning using spatial net-load forecasting.

Inspired by previous work presented in [35], the netload (***NL***) is calculated using the following expression:

$$NL_{sa,h} = L_{sa,h} + EV_{sa,h} - PV_{sa,h} \tag{4.8.}$$

Here, the hourly ($\boldsymbol{h}$) netload for each HV/MV transformer service area ($\boldsymbol{sa}$) results from subtracting a scaled PV curve (***PV***) from the addition of transformer load (***L***) and aggregated EV charging (***EV***). The net load of each transmission service area is retrieved through the addition of individual HV/MV transformer ***NL*** values. It should be noted that peak-load coincidence behaviour is neglected. Therefore, resulting values represent boundary upper estimates.

The assessment of DER uptake on transmission system planning is conducted using time-series analysis. In particular, changes of transmission grid entry point load diagrams, including the change in flow directions have been analysed. Likewise, the peak-load behaviour at each T/D connection point has been analysed.

A principal component of this chapter is the comparison of four large-scale DER allocation techniques ($\boldsymbol{at}$) and the effect of their use on vertical load flow retrieved estimates. Eventually, the presented approach quantifies such effects for the resulting load flow estimates at the T/D boundary of the continental Portuguese power system. The following metrics are analysed:

The reverse flow hours (***RF***) is the sum of hours at a given transmission system entry point that sees reverse flows. That way, ***RF*** provides insights in the hours each transmission system service area would see PV generation surpassing consumption (including EV charging additions). ***RF*** are calculated for each transmission service area as well as the four DER allocation techniques considered:

$$RF_{at} = \sum_{h=1}^{8760} \begin{Bmatrix} 1 \; if \; NL_h < 0 \\ 0 \; if \; NL_h \geq 0 \end{Bmatrix} \tag{4.9.}$$

Here, $\boldsymbol{at}$ is the set of DER allocation techniques typically used in grid impact studies (equal-share, random assignment, extrapolated with peak demand, extrapolated with installed transformer capacity). It should be noted that the metric sums all hours with a net load below zero, disregarding the magnitude of the deviation.

As second indicator used in this work, the peak load added (***PA***) is estimated under each DER allocation technique and across all 63 transmission system service areas.

The principal goal of this metric is to provide a rough indication of estimated peak load addition under each of the four DER allocation techniques. A $\boldsymbol{PA}$ of zero serves as the reference value (peak load in year zero of the natural load). The formula considers the peak load difference before and at the end of 20 years of EV/PV adoption. Again, $\boldsymbol{PA}$ is retrieved for each transmission service area ($\boldsymbol{sa}$) and across all DER allocation techniques ($\boldsymbol{at}$).

$$PA_{al} = \sum_{sa=1}^{63}(PL_{Year\ 20} - PL_{Year\ 0}) \tag{4.10.}$$

Input data sources to the vertical load analysis

Apart from the previously introduced spatiotemporal DER adoption model and the spatial model of the T/D interface, this analysis relies on two additional data sources:

1) Global values of EV/PV adoption scenarios for 2035, corresponding to a 20-year planning horizon pursued in this work. For the sake of this study, the *Distributed Generation* scenario has been chosen (compare Table 4.4). Electricity consumption levels and EV/PV installation rates from 2015 served as base year [52].

   The global EV/PV adoption time series was established with the Bass model [63]. The Bass model and its formulas have been introduced in Chapter 3. The model's coefficients $\boldsymbol{p}$ and $\boldsymbol{q}$ have been retrieved, calibrating the model with historical uptake values [64], [65]. Final coefficients are shown in Table 4.4.

   The distributed PV capacity adoption forecast in residencies (in kW) was adjusted taking into account the current ratio of dispersed PV to overall PV installations in Portugal reported in [65]. Furthermore, total EV and PV potentials (***M***) at Portuguese residencies have been estimated using inputs from [66] and [35].

2) Typical values for load, EV charging and PV generation subroutines. As in subchapter 4.2., natural load is assumed to include installed EV/PV at residencies. Again, normalized MV, EV charging and PV generation time series have been used. A per capita PV potential of 0.4 kWpeak/capita was estimated. Furthermore, an EV adopter charging rate of 5.9 kW with a 0.5 simultaneity rate has been chosen.

Table 4.4. Bass model parameters of the proposed methodology.

Technology	p	q	M
Electric vehicles	0.000618	0.873600	999,917
Photovoltaics	0.000618	0.873600	3,867,000

Outcomes of the vertical load flow analysis

The results of the analysis are shown in Figure 4.15 and Figure 4.16. Outcomes suggest that large differences in the estimation of EV charger or PV installations in each transmission service area might occur, depending on the DER allocation technique employed. In other words: The choice of which technique is used to allocate DER across the 63 transmission system service areas may result in very different net-load estimations.

With regard to installed capacities allocated into each service area, maximum deviations of up to 48 MW for EV charging and 97 MW of residential PV installations per service area can be observed (Figure 4.15, Figure 4.16). Likewise, large differences in minimum, maximum and installed capacity ranges of both EV charging and PV installation forecasts are visible. Their regional differences are displayed in Figure 4.15 and Figure 4.16 respectively.

A comparative analysis of the effect of DER allocation techniques to the previously introduced metrics ***PA*** and ***RF*** is shown in Figure 4.17. Showing the aggregated reverse flow hours (from distribution to transmission) (a) and peak-load increments (b) over the Portuguese transmission system, it becomes clear that the four DER allocation techniques might result in very different ***RF*** values.

While all allocation techniques predict similar peak load additions, a strong variation of reverse flow hour estimates can be observed. While former range between 1,200 and 1,400 MW peak load addition only, ***RF*** hours vary in between 5,000 and 30,000 hours.

The differences in ***PA*** are summing added capacities over all transmission system service areas. As such capacity expansion is exclusively due to EV charging added to the natural load profiles (demand growth has been neglected in this study), similar peak load addition estimates under all four DER allocation techniques come unsurprising. Assuming that the diffusion forecast provides most accurate DER diffusion forecasts (as indicated by outcomes of Chapter 3), estimates suggest that equal share DER allocation might lead to overestimate ***RF*** hours.

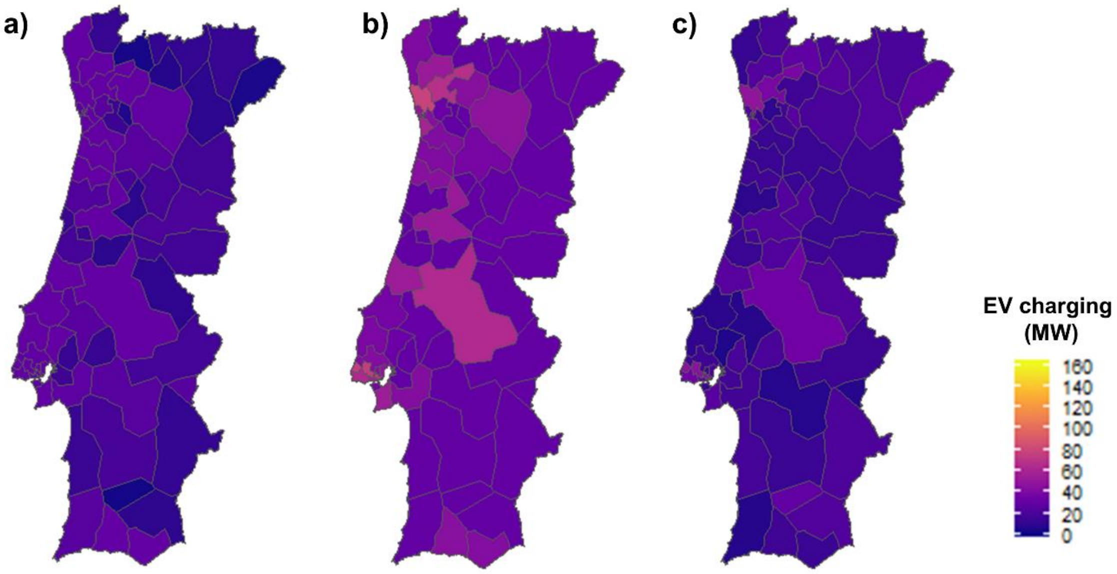


Figure 4.15. EV diffusion for different EV allocation techniques, where a) max, b) min and c) maximum difference.

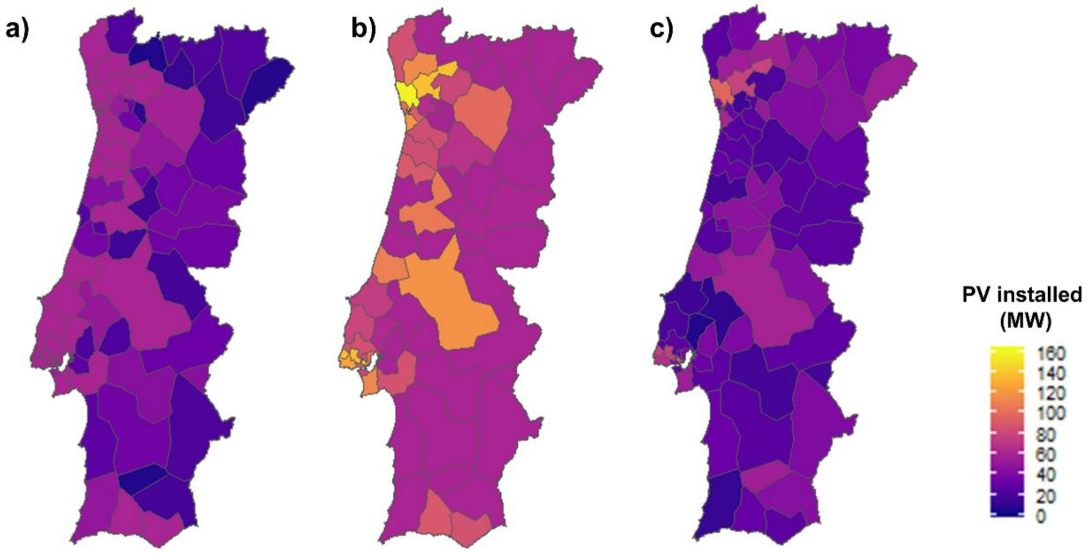


Figure 4.16. PV diffusion for different PV allocation techniques, where a) max, b) min and c) maximum difference.

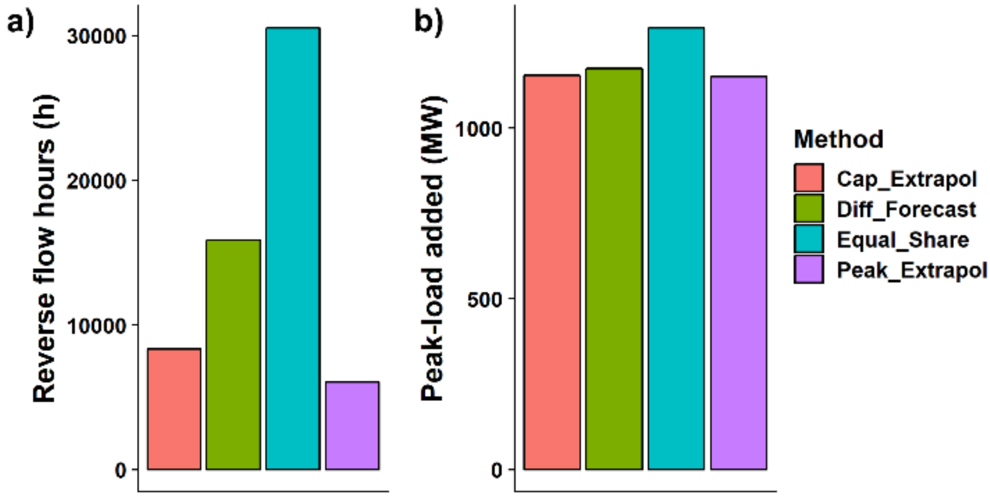


Figure 4.17. Estimated reverse flow hours (a) and peak-load increments (b) over the transmission system under different DER allocation techniques.

In our simulation, results show a roughly doubled RA hours estimate under equal share extrapolation if compared to the diffusion forecast. On the other hand, peak-load or installed transformer capacity extrapolations tend to strongly underestimate ***RF*** hours.

In the following, such effects have been showcased for two specific transmission system service areas. In Figure 4.18., load curves for the two transmission service areas that possess either the maximum amount of reverse flow hours (a) or the maximum addition to existing peak load (b) are shown. Situation **a)** could be observed under an equal-share of DER allocation in a service area within Northern Portugal. Here, the application of the equal DER allocation technique leads to a strong amplification of the initial load curve (Figure 4.18.a). For the considered time horizon, approximately 20 MW of EV charging and almost 60 MW of PV installations have been added to the respective transmission service area. The low initial demand level (less than 5 MW peak), imbalanced EV/PV adoption and low temporal charging/generation complementarity are resulting in frequent reverse flows (Figure 4.18.a).

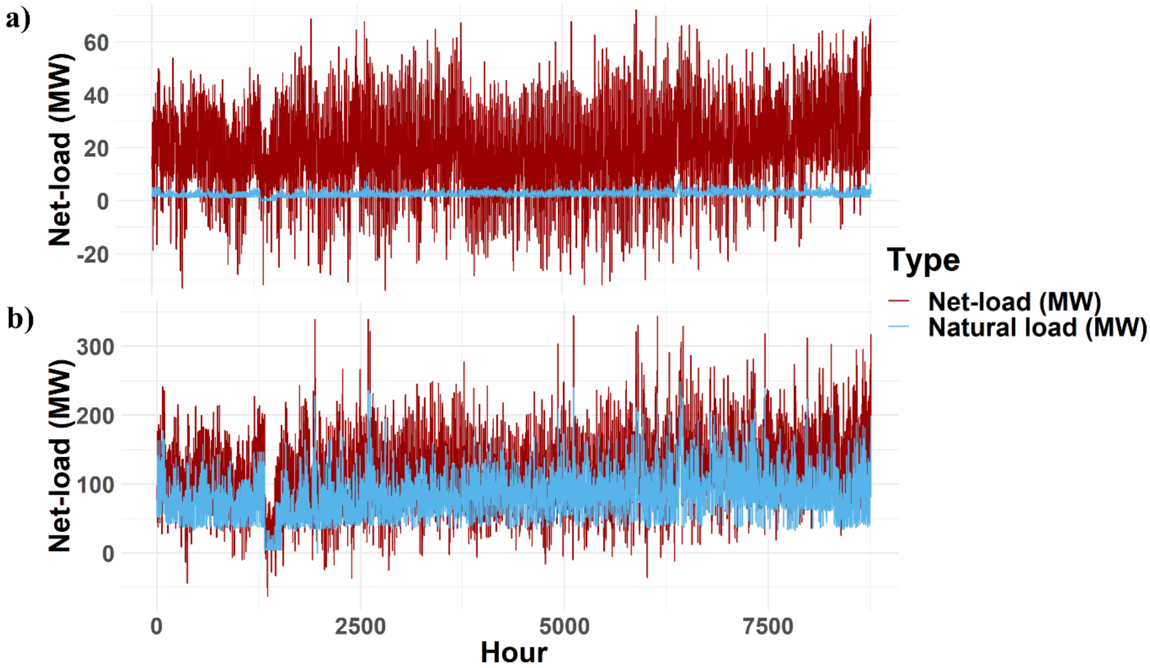


Figure 4.18. Load curves for transmission service area with most frequent reverse flows (a) and maximum peak increment (b).

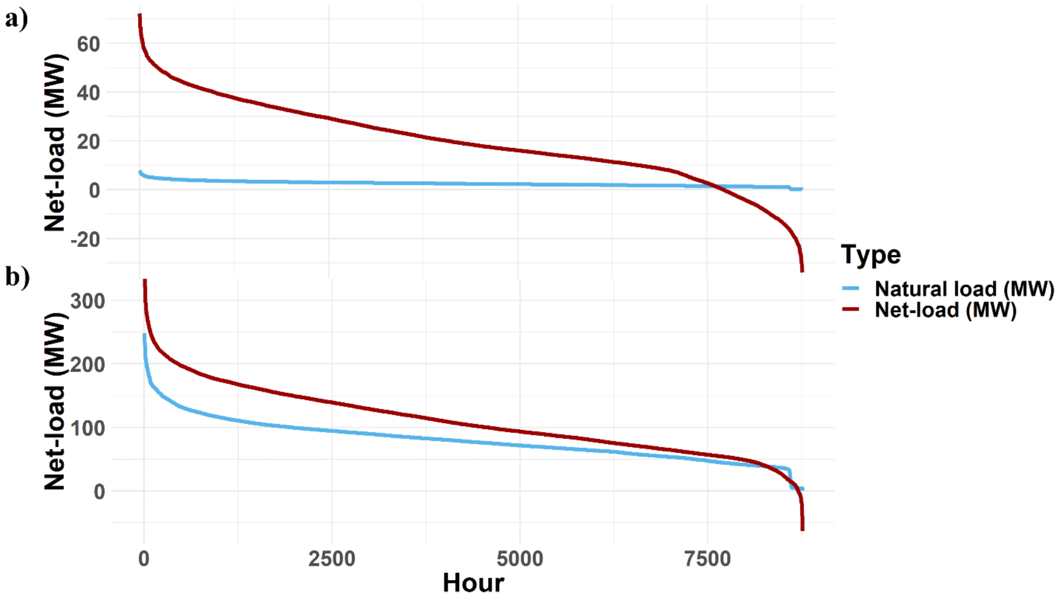


Figure 4.19. Load duration curves for transmission service area with most frequent reverse flows (a) and maximum peak increment (b).

In contrast, situation **b)** represents a transmission service area that locates in a southern, densely populated zone inside Lisbon Metropolitan area. This service area experiences the maximum peak load addition under all DER allocation techniques and service areas considered. In this specific case, the highest PA occurs under the DER forecast, which is sensitive to population counts within each transmission service area. Results of the time series analysis further validate the intuition, that reverse flows are less likely under high load level regimes (Figure 4.19).

The outcomes highlight the importance of carefully selecting a DER diffusion model that allows to understand the uncertainty range different technology uptake representations might add to a modelling process. In addition, assessing the sensitivity of DER model choice on transmission network expansion decisions is a prerequisite for economically efficient network planning.

Chapter summary and conclusions

In the profound transition of power systems, electricity network planners require tools to forecast the evolution of demand patterns affected by DER adoption. Currently, equal-share or extrapolation-based DER allocation are commonly used approaches to model DER uptake in grid impact studies.

Instead, this chapter presents a comparative analysis, relying on the spatiotemporal DER adoption model developed in Chapter 3. The presented methodologies allow for distribution planners to screen HV/MV substation service areas with a robust, transferrable impact analysis tool that is capable to assess the impact of any new grid-connected appliance.

For transmission planners, the methodology presented introduced two new planning criteria: The reverse flow hours (***RA***) and peak-load added under each DER allocation technique (***PA***). Both criteria can be calculated for each transmission service area. The resulting approach allows to show the uncertainty of vertical load flows, that is, the flows between transmission and distribution networks, under a given DER allocation technique.

The highlights of this chapter can be summarized in the following:

- The adoption of new technologies such as DER alters load patterns in residential environments and shifts the focus from peak-load to peak-net-load as a planning criterion.

- The currently used DER allocation models tend to underestimate distribution network expansion costs during light DER uptake and overestimate such costs under very high penetration of DER.

- The use of simplified DER allocation models further propagate uncertainty into vertical load diagrams at the transmission-distribution interface. For example, considering the analysed scenario, EV and PV adoption forecast per transmission service area can vary up to 60 MW or 100 MW, respectively.

- Although the presented tool has been applied to distribution expansion planning, the presented approach is applicable to similar analysis in any other network industries such as water, ICT or transportation.

Chapter references

[1] G. A. Jimenez-Estevez, L. S. Vargas, and R. Palma-Behnke, “An evolutionary approach for the greenfield planning problem in distribution networks,” *IEEE Int. Conf. Neural Networks - Conf. Proc.*, pp. 1744–1749, 2007.

[2] M. Coddington and K. Schneider, “Utility Distribution Planning 101,” in *Distribution Systems and Planning Training for Midwest Public Utility Commissions*, Berkeley, 2018.

[3] H. L. Willis, *Power distribution planning reference book*, Second Edi. New York, 2004.

[4] H. L. Willis, *Spatial Electric Load Forecasting*, Second Edi. New York: CRC Press, 2002.

[5] K. Bell, “Methods and Tools for Planning the Future Power System: Issues and Priorities,” *Model. Requir. GB Power Syst. Resil. Dur. Transit. to Low Carbon Energy*, no. Paper 5, pp. 1–35, 2015.

[6] G. Feismann, “Handbook / White Paper of Novel Planning Rules for EV in Smart Grids,” 2016.

[7] EPRI, “Developing a Framework for Integrated Energy Network Planning (IEN-P),” Palo Alto, California, 2018.

[8] elvolveDSO, “Development of methodologies and tools for new and evolving DSO roles for efficient DRES integration in distribution networks,” 2014.

[9] I. J. Pérez-Arriaga, S. Ruester, S. Schwenen, C. Battle, and J.-M. Glachant, *From distribution networks to smart distribution systems: rethinking the regulation of european electricity DSOs.*, Florence, 2013.

[10] R. Hemmati, R. A. Hooshmand, and A. Khodabakhshian, “State-of-the-art of transmission expansion planning: Comprehensive review,” *Renew. Sustain. Energy Rev.*, vol. 23, pp. 312–319, 2013.

[11] S. Lumbreras and A. Ramos, “The new challenges to transmission expansion planning. Survey of recent practice and literature review,” *Electr. Power Syst. Res.*, vol. 134, pp. 19–29, 2016.

[12] ABB Consulting, “Transmission Planning - An Introduction,” 2003.

[13] F. Silvestro *et al.*, "Review of transmission and distribution investment decision making process under increasing energy scenario uncertainty," in *25th International Conference on Electricity Distribution (CIRED)*, 2019, Madrid, pp. 3–6.

[14] IEA, *World Energy Outlook 2016*. Paris, 2016.

[15] T. Höfer, Y. Sunak, H. Siddique, and R. Madlener, "Wind farm siting using a spatial Analytic Hierarchy Process approach: A case study of the Städteregion Aachen," *Appl. Energy*, vol. 163, pp. 222–243, 2016.

[16] S. Izquierdo, M. Rodrigues, and N. Fueyo, "A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations," *Sol. Energy*, vol. 82, no. 10, pp. 929–939, 2008.

[17] B. Resch *et al.*, "GIS-Based Planning and Modeling for Renewable Energy: Challenges and Future Research Avenues," *ISPRS Int. J. Geo-Information*, vol. 3, no. 2, pp. 662–692, 2014.

[18] M. Rylander, J. Smith, and W. Sunderman, "Streamlined method for determining distribution system hosting capacity," *IEEE Trans. Ind. Appl.*, vol. 52, no. 1, pp. 105–111, 2016.

[19] A. Kaur, L. Nonnenmacher, and C. F. M. Coimbra, "Net load forecasting for high renewable energy penetration grids," *Energy*, vol. 114, pp. 1073–1084, 2016.

[20] J. R. Aguero and H. L. Willis, "Spatial Electric Load Forecasting Methods for Electric Utilities," Raleigh, 2007.

[21] THEMA Consulting Group, "Data Exchange in Electric Power Systems: European State of Play and Perspectives," Oslo, 2017.

[22] North American Electric Reliability Corporation, "Distributed Energy Resources: Connection Modeling and Reliability Considerations," no. February, p. 50, 2017.

[23] D. Alvarado, A. Moreira, R. Moreno, and G. Strbac, "Transmission Network Investment with Distributed Energy Resources and Distributionally Robust Security," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 5157 – 5168, 2018.

[24] P. Linares, "Multiple Criteria Decision-Making and Risk Analysis as Risk Management Tools for Power Systems Planning," *IEEE Trans. Power Syst.*, vol. 22, no. 7, p. 54, 2002.

[25] A. Bagheri, H. Monsef, and H. Lesani, "Integrated distribution network expansion planning incorporating distributed generation considering uncertainties, reliability, and operational conditions," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 56–70, 2015.

[26] M. Emmanuel and R. Rayudu, "The impact of single-phase grid-connected distributed photovoltaic systems on the distribution network using P-Q and P-V models," *Int. J. Electr. Power Energy Syst.*, vol. 91, pp. 20–33, 2017.

[27] J. F. Soares, P. Almeida, M. Galus, P. Barbeiro, and J. Pecs Lopes, "Active Management of Electric Vehicles Acting as Distributed Storage," in *Smart Grid Handbook*, C.-C. Liu, S. McArthur, and S.-J. Lee, Eds. John WIley & Sons Ltd., 2016.

[28] A. Rodriguez-Calvo, R. Cossent, and P. Frías, "Integration of PV and EVs in unbalanced residential LV networks and implications for the smart grid and advanced metering infrastructure deployment," *Int. J. Electr. Power Energy Syst.*, vol. 91, pp. 121–134, 2017.

[29] A. Zecchino and M. Marinelli, "Analytical assessment of voltage support via reactive power from new electric vehicles supply equipment in radial distribution grids with voltage-dependent loads," *Int. J. Electr. Power Energy Syst.*, vol. 97, no. April 2017, pp. 17–27, 2018.

[30] F. J. Ruiz-Rodriguez, J. C. Hernández, and F. Jurado, "Voltage unbalance assessment in secondary radial distribution networks with single-phase photovoltaic systems," *Int. J. Electr. Power Energy Syst.*, vol. 64, pp. 646–654, 2015.

[31] F. Shahnia, R. Majumder, A. Ghosh, G. Ledwich, and F. Zare, "Sensitivity analysis of voltage imbalance in distribution networks with rooftop PVs," *IEEE PES Gen. Meet. PES 2010*, 2010.

[32] H. Schuster, J. Kellermann, and T. Bongers, "Development of methodologies and tools for new and evolving DSO roles for efficient DRES integration in distribution networks," 2014.

[33] M. S. Elnozahy and M. M. A. Salama, "A comprehensive study of the impacts of PHEVs on residential distribution networks," *IEEE Trans. Sustain. Energy*, vol. 5, no. 1, pp. 332–342, 2014.

[34] R. Bernards, J. Morren, and H. Slootweg, "Development and Implementation of Statistical Models for Estimating Diversified Adoption of Energy Transition Technologies," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1540–1554, 2018.

[35] F. Heymann, J. Silva, V. Miranda, J. Melo, F. J. Soares, and A. Padilha-Feltrin, "Distribution network planning considering technology diffusion dynamics and spatial net-load behavior," *Int. J. Electr. Power Energy Syst.*, vol. 106, pp. 254–265, 2019.

[36] T. Zhao, Z. Zhou, Y. Zhang, P. Ling, and Y. Tian, "Spatio-Temporal Analysis and Forecasting of Distributed PV Systems Diffusion: A Case Study of Shanghai Using a Data-Driven Approach," *IEEE Access*, vol. 5, pp. 5135–5148, 2017.

[37] S. A. Robinson and V. Rai, "Determinants of spatio-temporal patterns of energy technology adoption: An agent-based modeling approach," *Appl. Energy*, vol. 151, pp. 273–284, 2015.

[38] C. L. T. Borges and V. F. Martins, "Multistage expansion planning for active distribution networks under demand and Distributed Generation uncertainties," *Int. J. Electr. Power Energy Syst.*, vol. 36, no. 1, pp. 107–116, 2012.

[39] V. Miranda, J. V. Ranito and L. M. Proença, "Genetic Algorithms in Optimal Multistage Distribution Network Planning," *IEEE Transactions on Power Systems*, Vol. 9, No. 4, pp.1927-1933, November 1994

[40] P. M. S. Carvalho e L. A. F. M. Ferreira, "Large-Scale Network Optimization with Evolutionary Hybrid Algorithms: Ten Years' Experience with the Electric Power Distribution Industry," *Computational Intelligence in Expensive Optimization Problems*, vol. 2, pp. 325-343, Springer Berlin Heidelberg, 2010.

[41] Y. Wang and D. Infield, "Markov Chain Monte Carlo simulation of electric vehicle use for network integration studies," *Int. J. Electr. Power Energy Syst.*, vol. 99, no. January, pp. 85–94, 2018.

[42] H. Falaghi, C. Singh, M. R. Haghifam, and M. Ramezani, "DG integrated multistage distribution system expansion planning," *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 8, pp. 1489–1497, 2011.

[43] R. Hemmati, R. A. Hooshmand, and N. Taheri, "Distribution network expansion planning and DG placement in the presence of uncertainties," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 665–673, 2015.

[44] Z. Hu and Y. Song, "Distribution network expansion planning with optimal siting and sizing of electric vehicle charging stations," *Proc. Univ. Power Eng. Conf.*, 2012.

[45] M. H. Moradi, M. Abedini, S. M. R. Tousi, and S. M. Hosseinian, "Optimal siting and sizing of renewable energy sources and charging stations simultaneously based on Differential Evolution algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 1015–1024, 2015.

[46] T. D. Chen, "Where are the electric vehicles? A spatial model for vehicle-choice count data," *J. Transp. Geogr.*, vol. 43, pp. 181–188, 2015.

[47] V. Vasseur and R. Kemp, "The adoption of PV in the Netherlands: A statistical analysis of adoption factors," *Renew. Sustain. Energy Rev.*, vol. 41, pp. 483–494, 2015.

[48] C. Lee, "Influence of local environmental , social , economic and political variables on the spatial distribution of residential solar PV arrays across the United States," *Energy Policy*, vol. 47, pp. 332–344, 2012.

[49] J. Saarenpää, M. Kolehmainen, and H. Niska, "Geodemographic analysis and estimation of early plug-in hybrid electric vehicle adoption," *Appl. Energy*, vol. 107, pp. 456–464, 2013.

[50] F. Heymann, N. Neyestani, F. J. Soares, and V. Miranda, "Mapping the Impact of Daytime and Overnight Electric Vehicle Charging on Distribution Grids," in *2017 IEEE Vehicle Power and Propulsion Conference (VPPC)*, Belfort, 2017.

[51] EDP Distribuição, "Development and investment plan for the national distribution grid 2015-2019," Lisbon, 2015.

[52] ENTSO-E, "TYNDP 2018: Scenario Report," 2018.

[53] INE, "Import of georeferenzed Census 2011 data (BGRI)," 2011. [Online]. Available: http://mapas.ine.pt/download/index2011.phtml. [Accessed: 01-Jan-2017].

[54] V. Miranda and C. Monteiro, "Fuzzy inference applied to spatial load forecasting," *Proc. IEEE PowerTech Budapest*, 1999.

[55] M. De Berg, O. Cheong, M. Van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, vol. 17. 2008.

[56] A. Navarro and H. Rudnick, "Large Scale Distribution Planning – Part II : Macro-optimization with Voronoi ' s Diagram and Tabu Search," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 752–758, 2009.

[57] EDP Distribuição, "National distribution grid - High Voltage and substations," 2016. [Online]. Available: http://edp-distribuicao-rede.wntech.com/Nut/4. [Accessed: 05-Sep-2017].

[58] ERSE, “Electricity profiles,” 2017. [Online]. Available: http://www.erse.pt/pt/electricidade/Paginas/default.aspx. [Accessed: 01-Jan-2017].

[59] B. P. Rand, J. Genoe, P. Heremans, and J. Poortmans, “Solar Cells Utilizing Small Molecular Weight Organic Semiconductors,” *Prog. Photovolt Res. Appl.*, vol. 15, no. February 2013, pp. 659–676, 2007.

[60] EDP Distribuição, “Distribution Network Development and Investment Plan 2017-2021 (in Portuguese),” 2016.

[61] ICF, “Integrated Distribution Planning,”USA, 2016.

[62] H. Gerard, E. Rivero, and D. Six, “Basic schemes for TSO-DSO coordination and ancillary services provision – SmartNet Project Deliverable 1.3,”, Belgium, 2016.

[63] F. M. Bass, “A New Product Growth for Model Consumer Durables,” *Manage. Sci.*, vol. 5, no. 16, pp. 215–227, 1969.

[64] European Alternative Fuels Observatory, “Electric vehicle statistics,” 2017. [Online]. Available: http://www.eafo.eu/vehicle-statistics/m1. [Accessed: 01-Jan-2017].

[65] Portuguese General Direction of Energy and Geology (DGEG), “Rapid Statistics - Renewables No.168,” 2018.

[66] ACAP, “Estatísticas do Sector Automóvel,” *Assoc. Automóvel Port. (Portuguese Automotive Association)*, Lisbon, 2013.

5 Optimal policy design using spatiotemporal DER adoption forecasts

Technology adoption is often accelerated through what has been called support schemes or incentive designs. It has been found that programs directed to residential users tend to favour high-income, above-average educated adopter groups, who receive twofold benefits. On the one hand, the adoption of an innovation allows reducing operational costs. On the other hand, the upfront price may be decreased by direct financial incentives born by all remaining taxpayers. In order to avoid widening socio-economic gaps as a result of DER adoption, DER incentive designs are currently under review. This chapter investigates the system-level effects of different incentive-design combinations for EV and PV. Exploiting the capabilities of the previously introduced spatiotemporal technology adoption forecasting model, incentive design effects on distribution network expansion costs and spatial adoption asymmetries are evaluated. Outcomes provide indication to policy makers on how potential synergies under orchestrated EV and PV incentive designs may be exploited. For example, outcomes suggest that, at system level, distribution network expansion costs can be reduced while minimizing DER adoption asymmetries, if specific incentive designs are combined.

5.1 INCENTIVE SCHEMES AND THEIR DESIGN

Accelerating technology diffusion through incentive schemes

The perceived attributes of innovations are decisive factors that influence the individual or collective attitude towards an appearing technical or non-technical novelty [1]. Recalling Roger's five stages of the innovation-decision process (*Knowledge of the innovation*, *Persuasion*, *Decision to adopt*, *Implementation* and *Confirmation*), it is obvious that the perception of the innovation attributes and its relative advantage over the existing technology/process is especially relevant to the stages that precede the decision to adopt.

According to Rogers [1], potential adopters thrive for information in order to decrease the uncertainty of consequences that are related to the adoption of the innovation. They further try to understand to which degree the innovation is outperforming existing solutions or alternatives.

The perception of "relative advantage" is supposed to possess the highest influence on the adoption rate. It is defined as "the expected benefits and the costs of adoption of an innovation (p.233)" [1]. Constituting factors are profitability, upfront costs, expected increase in comfort, a decrease in work, financial expenses or time spent, and the instantaneous reward. This assumption is well in line with the observations that the decline in the innovations price accelerates the adoption rate of a technological innovation (e.g. for EV/PV in [2], [3]).

Many political stakeholders or change agents use financial incentives or subsidies to accelerate the adoption rate of an innovation. Such are [1]:

1) Adopter versus diffuser incentives;
2) Individual versus system incentives;
3) Positive versus negative incentives;
4) Monetary versus nonmonetary incentives;
5) Immediate versus delayed incentives.

As stated by Rogers [1], the provision of incentives or subsidies can come along with ethical issues (such as distributional justice). Subsidies are necessarily payments or tax reductions for the benefit of the adopter, while incentives can also include non-financial means such as status improvements or positive reinforcement. For example, in Norway, individual passenger EV are allowed to use bus lanes and can therefore avoid congestion during rush hours [4].

According to [1], individual status and innovation acquisition costs are among the strongest influencing criteria that shape adoption rates. Financial upfront payments or non-financial incentives can increase the relative advantage of one technology over the other and thus, foster adoption. On the other hand, it was argued that complexity of the use or functioning of innovations would negatively correlate with adoption rates.

Figure 5.1. displays the various factors that have been associated to adoption rate changes. Although Rogers does not provide a mathematical framework to predict the diffusion speed and adoption rates of innovations which would be applicable to the modelling of energy technologies, his framework of ranking consumer groups along innovativeness can serve as an allocation formula for a given innovation that spreads inside a population structure [1].

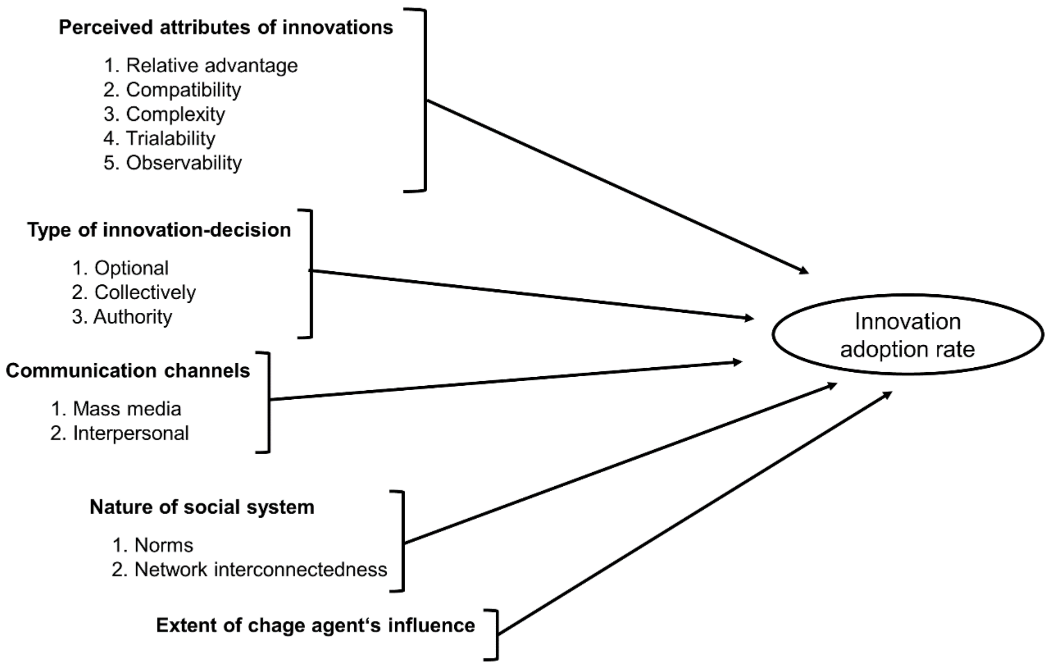


Figure 5.1. Web of factors that influence the innvoation adoption rate (inspired by [1]).

Various studies found that incentive design schemes have, mostly through the decrease of investment and operational costs, accelerated the uptake of DER. Such link has been modelled and documented in [2], [3], [5], [6]. Likewise, an increasing number of works has been dedicated to study the technology adoption process with increasing detail of consumer choice and adopter characterizations. Latter has been modelled using spatial and temporal, or other non-spatial models.

In the following, common mechanisms that intend to increase the spread of a given energy technology are discussed. In this work, incentive design is defined as a set of financial and non-financial instruments that policy makers can use to increase the attractiveness of a certain technology. Incentive design combinations (IDC), instead, are a mix of several incentive designs covering a multitude of different technologies. A short introduction of most frequent incentive designs per technology of interest is provided below.

Incentive designs for energy technologies

Incentive designs, or, as in [7], [8], support mechanisms or support policies, are policy initiatives aimed to increase the attractiveness of a certain innovation. Such incentive designs can be divided into direct and indirect support methods [7]. Direct methods can be further divided into quantity- and price-driven methods [7], [9]. An exemplarily overview of such direct and indirect support methods for the case of renewable energy technologies is shown below (Table 5.1.).

Table 5.1. Renewable energy technologies support scheme classification [9].

Method type	Specific incentive design
Direct incentives	
a) Quantity-driven	Green certificates Tendering schemes
b) Price-driven	Feed-in-tariffs (FITs) Capital grants Fiscal incentives Green loans
Indirect incentives	R&D subsidies Net metering Standards (RES share in new housing)

Here, the aim of direct incentives is to increase the attractiveness of the technology under question. Direct means may include financial support to technology acquisition (capital grants, fiscal incentives, green loans), as well as measures that reduce operational costs (feed-in-tariffs) or investment uncertainties (green certificates, tendering schemes). Latter represent a quantity-driven incentive, where in the case of renewable energy technologies, a fixed amount of capacity is procured, leaving the price-building process to market forces.

On the other hand, indirect methods have mostly no tangible impact on the attractiveness of a given technology but tend to decrease the attractiveness of other alternatives or provide intangible support that may increase the technologies' attractiveness at a later time step (e.g. R&D subsidies).

While this incentive characterization has been presented for the case of renewable energy technologies, it can be applied to programs for electric mobility or electrified heating adoption as well.

In the following, common incentives for renewable energy technologies and electrified mobility and heating in European markets are presented.

Renewable energy technologies

According to [6], the main four instruments that are used to promote renewable energy technology adoption in Europe are feed-in tariffs, feed-in premiums, green certificates and investment grants. Such options have been detailed in [10] and are introduced below:

- *Feed-in tariffs (FITs)*. Under this incentive design, deployment of RES is fostered providing a guaranteed, regressive remuneration to power plant owners for electricity fed into the system. This mechanism has been attractive as it is an investment freed from any market-risk (e.g. selling price volatility)

- *Feed-in premiums (FIPs)*. Feed-in premium, on the other hand side, provides RES power plant owners with a market premium that is paid on top of the market revenue. Different design options include fixed, floating, cap or a floor type premium. Importantly, under this scheme, less variable generators with a stronger generation control or less variable input sources (biomass, hydro) can respond to short-term market price signals.

- *Green Certificates (GCs)*: A market-based support mechanism where market actors (such as TSOs, generation owners or suppliers are obliged to buy a predetermined number of certificates) in order to comply with a specific renewable energy target. Parties that install and operate renewable generation falling under the GC scheme are, in return, remunerated once they sell the certificate, thus receiving additional income over the market revenue.

- *Investment grants or externality accounting*. These incentive design options either account for additional externalities that have not been integrated in the price formation process (e.g. CO_2 tax) or significantly reduce the upfront price of the desired technology. Investment grants can be differentiated along technological maturity and efficient values. However, little control of the adoption rate may result in over- or under-investment.

A detailed overview of all current policy instruments that support the adoption of renewable energy technologies is provided in [7]–[9]. An overview of current renewable energy incentive designs in Europe is provided in [6].

Electrified mobility and heating applications

With rising adoption of electric vehicles, research has been dedicated to the study of EV incentive designs and the effectiveness of certain policy measures. Given its per-capita share of private passenger EV, Norway is the current frontrunner of EV adoption [11], [12]. It provides a wide array of direct and indirect incentives that increased the attractiveness of EV in Norway. Such are:

- Direct incentives: Purchase tax, reducing the price of EV almost to conventional cars. Exemption from value added tax (VAT), vehicle registration tax as well as a reduced vehicle licensing fee;

- Indirect incentives: Allowance to use bus lanes, free parking on most municipal parking spots, exemption from ferry fees and road tolls.

Similar measures have been applied across most European countries. A detailed, country-wise overview of incentive designs for electric mobility can be found in [13]. A global review on policy instruments that aim to foster electric mobility is provided in [14].

As in the mobility sector, European legislations foster the use of energy efficient appliances in houses and apartments through the setting of strong standards. The respective, recently amended directive [15], states that all newly constructed buildings should reduce net energy consumption to nearly zero by the end of 2020. For new, public buildings, such situation should be reached even from 2018 onwards. However, the encouragement of electrified heating and cooling appliances is limited to the compliance with building codes, with little financial support or indirect, additional incentives.

In the following, a condensed overview of Portuguese support schemes for EV, PV modules and electrified heating and cooling appliances is provided. The focus lies in the residential sector for all technologies concerned.

With regard to EV, Portugal has provided for several years attractive conditions to early EV adopters [12], [16], that lead to EV adoption shares that are among the highest in Europe [12], [13]. Currently, direct incentives include purchase subsidies (e.g. 2,250 Euros per BEV), while indirect incentives include free charging at public EV charging stations or parking in some municipalities.

On the other hand, the application of a feed-in-tariff scheme represents the sole direct incentive for the acquisition and installation of PV modules in Portugal. Indirect incentives include the legislation for self-consumption as well as guaranteed acquisition of the generated energy by the DSO. For larger producers, that indirect incentive is replaced by a preference within the merit order process of wholesale markets.

Table 5.2. Incentive designs in Portugal: The case of EV, PV and electrified heating and cooling.

DER/ Incentive types	Electric vehicles	Photovoltaic modules	Electrified heating/cooling systems
Direct incentives	• Purchase subsidies, e.g. 2,250€ per BEV and 1,125 € per PHEV • Exemption/reduction from taxes (ownership tax, circulation tax, VAT)	• Feed-in tariff (existing contracts)	-
Indirect incentives	• Free parking (e.g. in Lisbon) • Free or subsidized charging rates • New vehicle standards set maximum emission levels per fleet	• Self-consumption regulation • Priority in merit order • New building standards	• New building standards

For electrified heating and cooling appliances, no direct incentives could be found. Instead, indirect incentives include requirements for newly built/refurbished buildings minimum coverage of endogenous demand with RES (which is applicable in case of PV, solar thermal or heat pumps).

It is noteworthy that, in principle, most indirect incentives root back to standards and specifications set in European directives. Such directives include the Regulation (EU) No 443/2009 setting emission performance standards for new passenger cars, the Renewable Energy Directive (revised, 2018/2001/EU), the Energy Performance of Buildings Directive ((2018/844/EU).

Incentive designs and distributional justice

Alongside increasing adoption of DER in power systems across the world, a growing amount of research has been dedicated characterizing early adopter households. A majority of such socio-demographic characterizations found young, well-educated, materially better-off households being among the first to adopt DER [17]–[21] (compare Chapter 2 as well). These population groups were also the principal recipients of financial support through governmental programs [22].

Furthermore, research has been directed towards the analysis of potential barriers that may hinder a widespread adoption of DER [23]–[25]. Outcomes suggest that lack of information, consumer perceptions, investment costs or unattractive incentive schemes compared to other technological alternatives represent hindrances to a wide adoption of DER. Given that consumer groups possess different levels of access to information and financial resources, studies started to look at participation and benefit allocation of household groups in DER incentives [26]–[29]. These works opened the path to further investigations on the link of incentive designs in place to distributional justice and the socio-economic impact of DER employment [22], [26]–[30].

The work in [4] and [31] highlight the importance of EV incentives to the adoption behaviour in Norway. Both studies find that both direct incentives (various tax exemptions) and indirect incentives (allowance to use bus lanes, access to charging infrastructure) have been responsible for fostering EV adoption. The work of [4] delineated population groups and their responses to the same incentives. While the Norwegian studies investigate on EV adoption behaviour to a specific constrained set of incentives in place, several studies analyse in large the interaction of governmental incentive design schemes and macro-scale DER diffusion (e.g. [2], [5], [32]).

The effects of current PV incentive design (FITs) for England and Wales on social equity has been investigated in [26]. Crossing spatial census data with domestic PV installation records, authors found moderate to high level inequality in the access to benefits provided under FITs. Other contributions addressed complementary aspect,s such as the distributional effects of network charges in Portugal [33] and the cost-benefit allocation under net-metering from consumers to the prosumers using power flow simulations in synthetic US distribution networks [30]. In addition, studies have been directed to study DER incentive's design using system dynamic modelling [28] or time series of the Spanish power system [29]. A similar study has been conducted comparing the cross-subsidization by non-adopters of PV or air-conditioner (AC) systems in Australia [27]. Like the previously cited works, this study concludes that non-adopters tend to subsidize early DER adopters, estimating cross-subsidy payments of 300-350 Australian dollar per non-adopter.

The study presented in [34] represents the only attempt to link DER incentive designs to spatial patterns of technology adoption. Conducting a spatial adoption forecast for residential PV in a mid-sized Brazilian city, the model is constrained to one technology (roof-top PV) and a test-network. Furthermore, it does not assess the grid impact of a set of DER incentive designs.

Likewise, there is a growing amount of research dedicated to analyse the uptake of DER onto electrical distribution networks [35]–[40]. Although some authors rely on spatiotemporal adoption forecasting models (e.g. [39], [42]) for such studies, most neglect underlying socio-demographic structure and adoption likelihood structures.

Granqvist and Grover [41] illuminate the distributional justice aspects for cross-subsidies in energy infrastructures from an ethical perspective. They differentiate four principles for distributive fairness, stretching on the fact that the perspective followed by each principle has put different burdens on consumer groups. The principles are polluter pays, ability to pay, beneficiary pays, and grandfathering.

Finally, it should be noted that the scope of this chapter is constrained to distributional aspects of DER adoption and does not address large scale market distortions that incentive designs may introduce (see [7], [8] for further details).

Concluding, studies suggest that subsidizing the acquisition of DER potentially leads to social inequalities through two mechanisms:

- Early DER adopters tend to benefit from financial incentive schemes. These reduce DER investment costs of early adopters which are, however, borne by all taxpayers. Therefore, recipients of DER incentives profit from typically lower total system costs (investment and running costs) along the products' lifetime [26]. If DER incentives are not equally distributed, social inequality may increase [1], [22].

- In addition, for the case of power systems, early DER adopters also tend to pay a disproportionally lower contribution to network investments that are needed to accommodate these resources [28], [29]. As a result of this, late adopters' risk to cross-subsidize early DER adopters through their participation in rising costs that may be lower for early adopters (e.g. as with net metering tariffs).

Given first evidence that the design of incentive schemes may have strong effects on distributional justice through asymmetrical benefit allocation, policy makers and researchers started reviewing DER incentive schemes. First programs that intend to enable an equally strong participation of low- and medium-income groups in DER incentive programs have been already designed [42], [43]. In this light, the current chapter investigates on three research questions from a system perspective:

- How can one model the effect of different DER incentive designs on technology diffusion patterns?
- What is the likely effect of different DER incentive designs on electricity network planning and investments?
- Given previous outcomes, how could optimal incentive designs be identified that both minimize system-level expansion costs while diminishing adoption asymmetries across population groups?

In this chapter, the previously introduced spatiotemporal DER adoption model is used to obtain insights into policy and technical aspects of incentive designs.

As main contribution, this chapter illuminates system-wide effects of various DER incentive design combinations for EV and PV on distribution network expansion costs and adoption asymmetries. Such insights possess a high utility both for governmental policy makers or distribution network companies. While latter can use the methodology to screen their service area and anticipate network expansion, policy makers can use the model to assess large-scale social and economic impacts of current and potential future DER incentive designs.

Eventually, the overarching goal is to identify a way to orchestrate DER incentive schemes so that, on a system level, network expansion costs and adoption inequality are minimized.

It should be highlighted that the present analysis does not take in account, in line with previous studies, the role of cultural differences across societies in technology adoption processes. While the adoption behaviour of unknown innovations is obviously linked to the tendency in each society of risk-taking or risk avoidance, differences in technology uptake tend to be explained by rational profit-seeking behaviour of individuals, only [9].

5.2 COMPREHENSIVE INCENTIVE DESIGN ANALYSIS

Overall model architecture

Given the interdisciplinary character of the presented methodology, its building blocks cover multiple disciplines that relate to the topic addressed in this chapter (Figure 5.2), such are:

- Social science and marketing: DER adopter characterizations based on census data.
- Energy engineering: Spatiotemporal DER adoption forecasting techniques.
- Power system planning: Impact assessment of DER and their effects onto the planning and operation of electricity distribution networks.
- Energy policy and law: DER incentive designs and their impact to distributional justice.

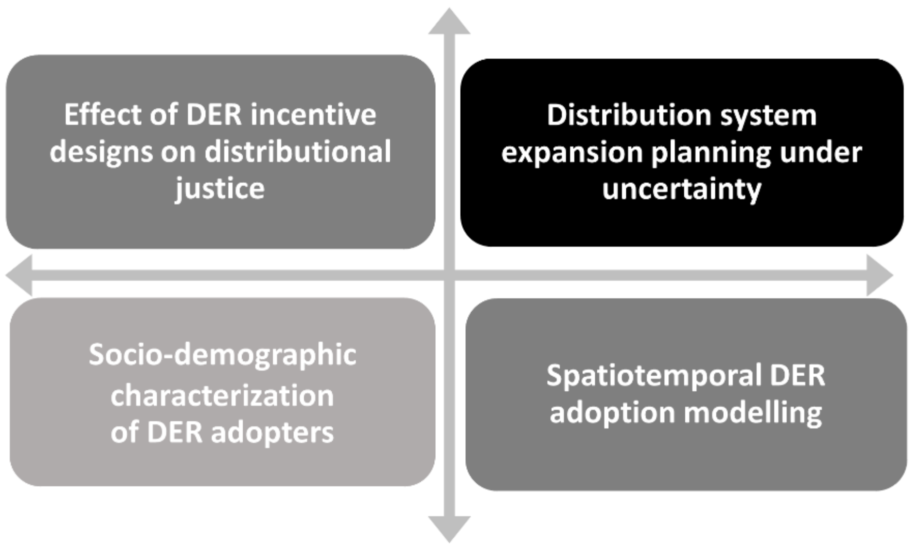


Figure 5.2. Research areas addressed in this chapter

The developed methodology consists of six, linked subroutines (Figure 5.2):

- Steps i) – v) are identical to the ones described in Chapter 4 and include the geolocation of DER adopters, the construction of HV/MV transformer service areas, net-load analysis and the calculation of transfomer loadings.

  In this chapter, the one difference lies in the activation order of DER, which is driven by synthetic incentive designs for EV and PV. Such are bundled to incentive design combinations (***IDC***). Another difference is the consideration of peak-load reduction potentials through optimized use of EV-PV, in case certain conditions are met.

- Step vi), instead, compares the estimated network expansion costs with adoption pattern asymmetries. Latter is analyzed using the Information-Theoretic inequality index Theil's T (***TT***). Eventually, for each EV-PV incentive design combination, the estimated network expansion cost (considering peak-load reduction potentials) and the impact adoption asymmetries are compared.

In the following, a detailed descpription of the three new aspects (synthetic incentive designs, peak-load reduction potential and adoption asymmetries) covered in Chapter 5 (compared to the approach presented in Chapter 4), is provided.

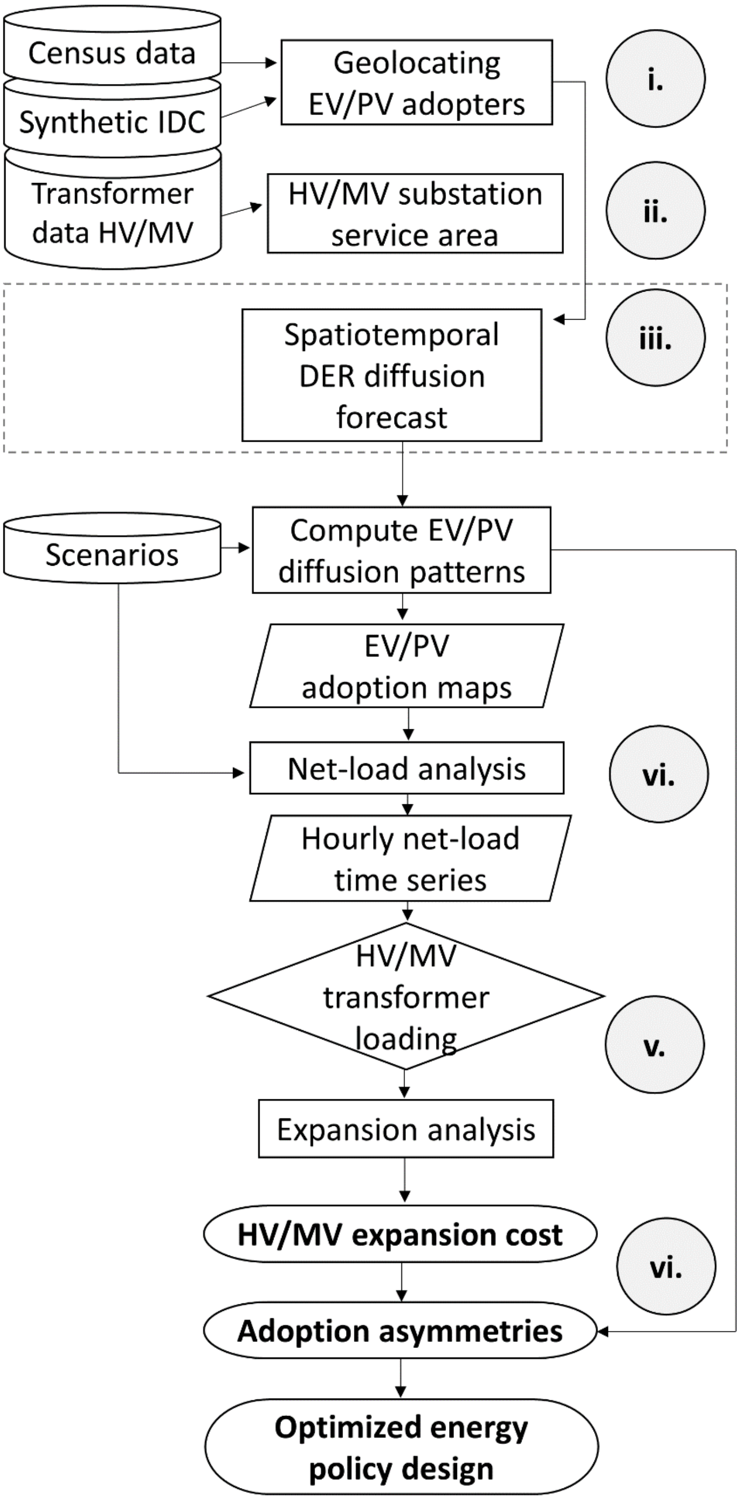


Figure 5.3. Developed methodology

In Chapter 3, it has been shown how innovativeness scores are used to build the cellular activation order in spatiotemporal DER adoption forecasts. The main idea behind innovativeness was that population subgroups or individuals can be ranked according to their preference to adopt a given innovation. That way, innovativeness is an endogenous characteristic of individuals or population subgroups that allows comparing adoption behaviour in a social system [1]. While in Chapter 3, innovativeness scores were inferred using real DER adoption observations intersected with census data, in this chapter, innovativeness scores are synthetically built for each spatial census unit. Using outcomes of first studies that reported socio-demographic census variables with a causal link to DER adoption (Figure 5.3), three different incentive designs for both EV and PV have been constructed. Mainly two types of studies have been informative for this process:

1) Research dedicated to analyse the responsiveness of population subgroups to DER incentive schemes [4], [18], [44], [45].
2) Research that investigated the relation of socio-demographic population characteristics and DER adoption behaviour using census data [4], [17], [20], [39], [44], [46]–[49].

Using the outcomes of such studies, the adoption positive response of a certain population subgroup to a future, synthetic incentive design can be modelled. As presented in Chapter 3, a predefined, numeric adoption factor is set for each socio-demographic variable, representing the adoption preference of a given census cell to EV/PV technology (Figure 5.3).

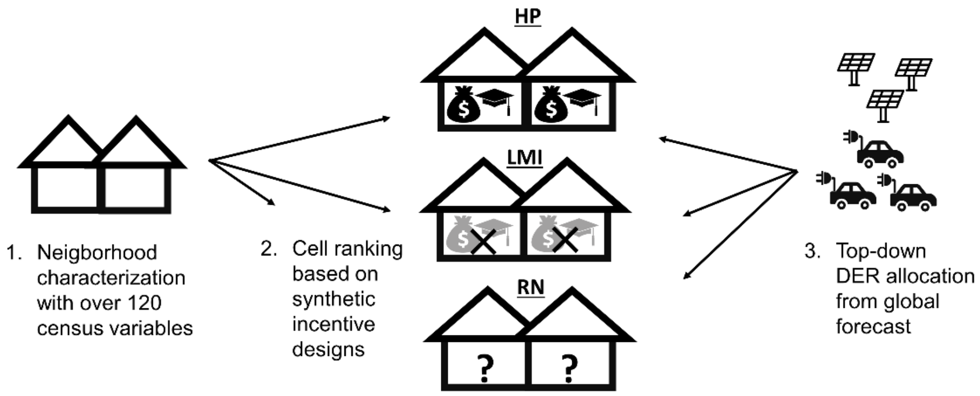


Figure 5.4. Adopter response model using synthetic incentive designs with household characterization based on selected census criteria.

It should be noted that modelling technology adoption using synthetic innovativeness scores follows a black-box approach, where assuming certain socio-demographic census variables triggers EV/PV adoption stronger than other variables.

As described in Chapter 3, innovativeness scores (***IS***) are constructed using over 120 socio-demographic criteria (***c***=120) of the census data-set containing roughly 17,000 spatial census tracts (***r***=17,337).

$$IS = \begin{pmatrix} x_{1,1} & \cdots & x_{1,c} \\ \vdots & \ddots & \vdots \\ x_{r,1} & \cdots & x_{r,c} \end{pmatrix} \times \begin{pmatrix} af_1 \\ af_2 \\ \cdots \\ af_c \end{pmatrix} \tag{5.1}$$

Here, the census dataset is multiplied with a predefined vector of adoption-influence (***af***). Under each incentive design combination (IDC), this vector is different, considering literature findings on the relation of socio-demographic census variables and DER adoption preference.

In this chapter, we build three distinct EV/PV incentive designs using different innovativeness scores (compare also Figure 5.3):

- A low-or-medium income (LMI), which allocates stronger weights to census variables that have been associated with groups with lower average education levels, smaller apartment sizes and housing renting.
- A high performance (HP) incentive design, where preference is given to households/individuals that possess privileged access to education and above-average financial resources.
- A completely randomized allocation (RN), which intends to mimic equal chance to adopt DER across all population subgroups and is used for benchmarking purpose.

For the construction of HP incentive scheme, high adoption-influence values where linked to criteria mentioned in [49]. Furthermore, HP adopters were differentiated into HP-EV and HP-PV adopters, allocating higher weights to parking space at the residency (HP-EV) and independent-standing, and occupier-owned family houses (HP-PV), respectively. LMI adopters were designed identical for EV and PV, linking higher weights to lower education levels, higher unemployed rates and households living in small apartment. In the following, the weight allocation process to the census variables is further explained (compare also Eq. 5.2). Eventually, the causal “incentive design - DER adoption” link is modelled using predefined vectors with census variable weights (af_n). These weight vectors intend to numerically discriminate the adoption propensity (expressed as numeric value) of adoption favouring criteria ($\boldsymbol{afc}$) compared to all other census criteria ($\boldsymbol{occ}$).

$$af_n = \begin{Bmatrix} 1.00 \ \ for\ n \in \{afc\} \\ 0.05 \ \ for\ n \in \{occ\} \end{Bmatrix} \tag{5.2.}$$

While the numerical values are fixed ex ante under the current methodological framework, it is expected that under increasing EV/PV adoption observations, $\boldsymbol{af_n}$ can be eventually substituted by values derived from enhanced inference tools (such as in [39], [44]).

Under the scope of the presented analysis, nine EV-PV incentive design combinations were assessed, crossing HP, LMI and RN incentive designs for both EV and PV. An overview of the nine incentive design combinations analysed is shown in Table 5.3.

Table 5.3. Incentive design combinations.

IDC	EV	PV
1	HP	HP
2	LMI	HP
3	RN	HP
4	HP	LMI
5	LMI	LMI
6	RN	LMI
7	HP	RN
8	LMI	RN
9	RN	RN

Synergetic use of EV and PV for peak-load reduction

If a sufficiently large capacity of DER is present within a HV/MV transformer service area, potential synergies might be exploited. First studies investigated the potential interplay of optimzed use of HVAC [50], EV [36] or coordinating PV and battery systems for peak-load reduction [51], [52]. In this Chapter, we will consider outcomes of studies on optimized PV-battery utilization, as the presented case study investigates EV-PV only. However, the presented approach might be extended to analyze any other DER interplay.

Recent studies investigated the synergetic use of residential batteries bundled with PV systems to reduce household’s peak demand. For example, one study suggested that an aggregated PV capacity of 1 MW can achieve peak-load reductions of 6% - 51% [52] or 8% - 32% [51]. The studies considered a battery (in kWh) to PV system capacity (in kW, nameplate) ratio of roughly two. Interestingly, the above-cited studies were conducted on multiple neighbourhood level (e.g. 99 households in [51]), which is well in line with the scope of the analysis. In fact, under the Portuguese case study, even 1,000 households might be connected to a HV/MV transformer, with additional peak-shaving benefits. Previous findings are integrated in the presented model as follows:

$$\left\{\begin{array}{c} if\ (N_{PV} \times pcpv > \ 1\ MW) \quad and \\ if\ (N_{EV} \times s \times bs > 2 \times N_{PV} \times pcpv) \\ \rightarrow PNL = 0.94 \times PNL_{t0} \\ \\ if\ (N_{PV} \times pcpv > \ 10\ MW) \quad and \\ if\ (N_{EV} \times s \times bs > 2 \times N_{PV} \times pcpv) \\ \rightarrow PNL = 0.80 \times PNL_{t0} \end{array}\right\} \tag{5.3}$$

Here, the number of EV and PV adopters connected to each HV/MV transformer are $\boldsymbol{N_{EV}}$ and $\boldsymbol{N_{PV}}$, where $\boldsymbol{s}$ is again the simultaneity rate of EV connected, $\boldsymbol{bs}$ the typical battery size (24 kWh as in [53]) and $\boldsymbol{pcpv}$ is the per capita capacity of PV (0.4, as derived in Chapter 3 and 4). $\boldsymbol{PNL}$ is the peak-net-load of each HV/MV transformer.

In case at a given HV/MV transformer, the installed PV potential exceeds 1 MW and the ratio of the aggregated EV battery storage ($\boldsymbol{N_{EV}}$ x $\boldsymbol{s}$ x $\boldsymbol{bs}$) to the installed PV capacity is at least two, then a peak reduction of 6% is assumed. In a high PV penetration scenario, where the installed PV capacitiy exceeds 10 MW, a peak-reduction potential of 20% is assumed (Eq. 5.3).

Assessing the asymmetry of technology adoption patterns

In order to estimate the impact of DER incentive designs on social welfare distribution, DER adoption patterns are spatially assessed using an Information-Theoretic inequality metric.

In the past decades, several inequality measures (indices and ratios) have been proposed, mainly within social and economic sciences. Mostly these measures are applied to questions of economic development. Popular inequality metrics are the Gini-coefficient, the 20/20 ratio, Atkinson's inequality measure, the Hoover index or Theil index (Theil's T) [54]. As explained in [55], Theil's T (TT) is commonly preferred over the former ones, as it satisfies all desirable properties of inequaltiy measures. Such properties are the independence to population size, symmetry and decomposability. TT has been widely applied, for example, in the assessment of income equality in European countries [56].

This chapter presents the first application of an inequality metric, namely TT, to DER adoption patterns. The main idea behind that is to investigate the equality/inequality of consumer participation in DER incentive programs. TT is caculated as follows [55]:

$$TT = \frac{1}{NO} \sum_{i=1}^{NO} \frac{y_i}{\bar{y}} ln\left(\frac{y_i}{\bar{y}}\right) \tag{5.4}$$

Here, $\bar{y}$ is the mean value of the variable concerned – in the presented case, the DER adoption ratio – while $\boldsymbol{NO}$ is the total number of observations $\boldsymbol{y_i}$. In the presented approach, we use TT to assess the distribution of DER across all HV/MV substation transformers. In order to account for the differences of individuals connected to each HV/MV transformer, DER adoption values have been set in relation to population within each HV/MV service area. The detailed analysis of population structure across all HV/MV service areas and the effect of each IDC onto various population subgroups lie outside the scope of this work.

For the analysis of change in inequality over time, we define the coefficient of change in Theil's T ($\boldsymbol{CT}$) as the inequality of DER adopted with relation to the inequality of population distributed under each HV/MV transformer. As TT is relying on the natural logarithm, a direct comparison of TT for DER adoption shares at the base year and after a 20-year horizon is intractable, given that multiple HV/MV service areas have had zero DER adopters in 2015. Hence, we compare the population shares at each HV/MV transformer under the status quo (year 2015) with the DER adoption state after 20 years of DER uptake given a chosen EV/PV incentive design combination.

Therefore, the change in TT under each incentive design combination is calculated as follows:

$$CT = \frac{(TT_{20} - TT_0)}{TT_0} \tag{5.5}$$

Thus, one can estimate the increase/decrease of DER per capita connected to a HV/MV transformer under each IDC. This way, policymakers, who would desire equal adoption behaviour of DER across all population groups, can check as a first approximation if similar per capita shares within each HV/MV transformer service area are achieved. In other words, a balanced participation of all population subgroups in DER incentive programs would result in the minimum CT value. Thus, the methodology presented in this Chapter provides policy analysts with a tool that can estimate distributional effects of single incentive schemes and incentive design combinations. This is achieved through the analysis of spatial distribution of adoption ratios across all HV/MV transformer service areas.

However, the outcomes of the presented methodology cannot replace sound analysis of the detailed reasons of inequality structures and their change, but rather serve as a starting point to consolidate further empirical and micro-economic evidence.

5.3 Optimized incentive design choice

One major outcome of the presented methodology is the provision of a tool to compare the effects of different policy designs on system expansion costs and adoption asymmetries. Such goals are translated into two objectives: The reduction of network expansion costs, approximated through HV/MV transformer upgrade costs and expressed through $\boldsymbol{TTC}$ as well as the reduction of asymmetries in spatial DER adoption patterns. Latter objective, i.e. its change over time, is expressed through the change in $\boldsymbol{CT}$. In this chapter, we assume that policy makers want to minimize both objectives. Furthermore, the trade-offs between both objectives shall be analysed.

The study of efficient (non-dominated) solutions and the study of the trade-offs between objectives is part of multi-objective optimization and decision making [57]. As we assume that policy makers seek to minimize a function of both system costs and adoption asymmetries, the optimization problem statement can be formulated in the following way:

$$minimize\ f(TTC, CT), TTC, CT \in OS \tag{5.6}$$

A theoretically optimal solution, usually not feasible and thus called "Ideal", that is, the IDC that would be outperforming all other IDC, can be identified within the attribute or objective space by independently minimizing $\boldsymbol{TTC}$ or $\boldsymbol{CT}$. The objective space ($\boldsymbol{OS}$) is the result of all possible realizations of the objective values. The decision space ($\boldsymbol{DS}$) is the space defined by all the variables of the problem. In this work, the attribute space ($\boldsymbol{OS}$) and the decision space ($\boldsymbol{DS}$) are discrete, because only nine solutions, in the form of discrete incentive design combinations, will be compared.

In a multiple criteria problem, a decision maker is predominantly interested in discriminating dominated solutions, mostly in the interior of the domain representation in the attribute space, from efficient or non-dominated solutions constituting the Pareto set. Dominated solutions are worse in both objectives than other solutions, therefore they are usually of little interest to decision makers. Pareto dominating solutions are preferred, the ones for which one cannot find another solution better in one criterion and not worse in the other criterion.

If a continuous solution space is considered, the subset of Pareto dominating solutions spans a surface, which is commonly referred to as the Pareto frontier. In order to further assess and compare solutions that lie on the Pareto front, the distance from the Ideal (in the OS) may be calculated. In this work, two common distance metrics have been used for such calculation. Utilized metrics are the Euclidean distance (***ED***) and the Manhattan distance (***MD***). Distances (***D***) between two solutions (including the Ideal) can be calculated with the following equation:

$$D = \left[\left(x_i - x_y\right)^{dp} + \left(y_i - y_y\right)^{dp}\right]^{1/dp} \tag{5.7}$$

In a two-dimensional decision space, each solution can be represented through a unique coordinate pair given both objectives to be minimized (***TTC***, ***CT***). In order to determine the distance from the Ideal, Eq. 5.8. is applied. In the presented formulation, (***dp***) represents the generic distance parameter. The Euclidean distance (metric L_2) is calculated replacing ***dp*** with 2, whereas, the Manhattan distance (metric L_1) can be retrieved by setting ***dp*** equal to one.

The Manhattan metric introduces a linear compensation between objective values, while the Euclidean metric tends to balance smaller objective values by overcompensating in the larger objective values. Other metrics could be adopted, the choice of a metric being intimately linked to the nature of the problem and the decision making process – it is an external decision and not a consequence of the mathematics of the method.

Input data

Compared to the approach presented in Chapter 4, the presented model in this Chapter relies on the following additional data input.

- **Scenario analysis and DER diffusion forecast.** As in Chapter 4, the presented model is applied to a fixed reference scenario considering a 20-year time horizon. As before, global EV/PV adoption forecasts have been extracted from [59], correcting for the current ratio of dispersed PV to overall PV installations in Portugal as stated in [60]. Total EV and PV potentials (TEV, TPV) per census cell and Bass model coefficients ***p, q*** and ***M*** and discretization of the adoption process (4 stages) are chosen as in Chapter 4.

- **HV/MV Substation Service Areas.** As introduced in Chapter 4, HV/MV substation transformer service areas have been approximated using geometrical operations available in GIS [61]. Inaccuracies in the spatial HV/MV service area model (e.g. misalignment) are neglected.

As discussed earlier, the Voronoi diagrams which represent HV/MV service areas represent geometric simplification of the real service areas, disregarding the historically grown structure and interconnectivity of the real-world feeder system. Therefore, the estimated service areas have not been adopted to geographical limiting factors (rivers, mountain, ranges, transport infrastructure), which could be included in future works. However, it should be reminded that eventually, distribution system operators possess detailed knowledge on the service boundaries supplied by each HV/MV substation and can therefore redraw/replace service area boundaries where needed.

All approximated 391 HV/MV transformer service areas are shown in Figure 5.5.

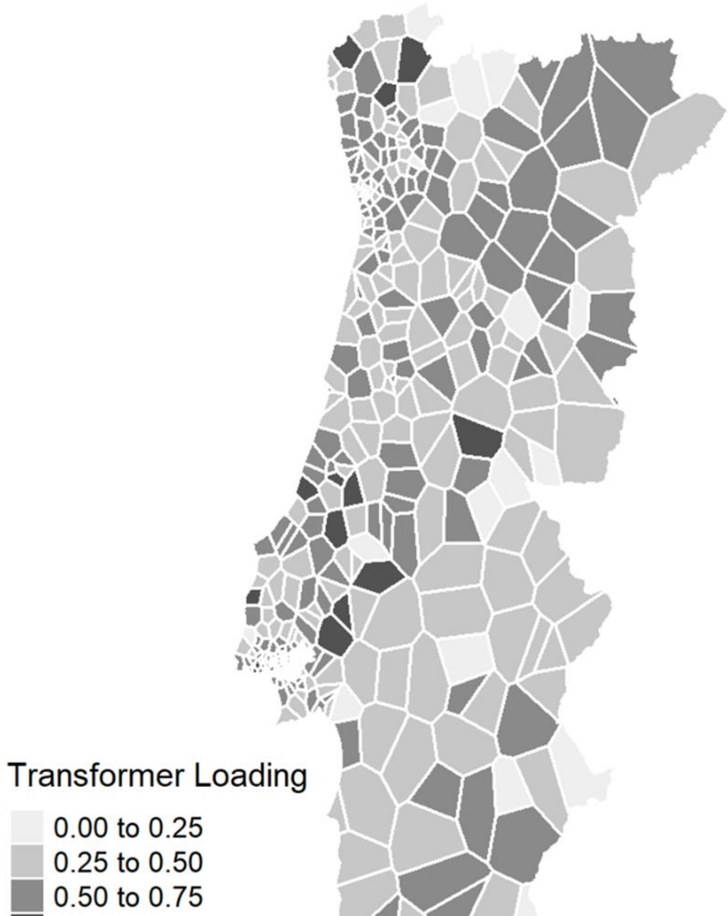


Figure 5.5. Approximation of HV/MV substation service areas.

- **Sensitivities.** The model outcomes are assessed estimating the sensitivities for parameters that positively affect the peak load (and thus, network expansion) under the chosen model architecture (Table 5.3.). Therefore, two charging rates are compared – with 5.89 kW per EV connected (30% with 12kW and 70% with 3.7kW) is assumed under light conditions, while 9.82 kW per EV connected (10% charge with 40kW, 30% with 12kW and 60% with 3.7kW) are considered as upper boundary. These extreme conditions also consider a higher simultaneity rate of 0.75. The PV model remained fixed, as further technology improvements (e.g. other PV module materials) that may bring improved conversion efficiencies lie outside the scope of this work.

  Eventually, a larger battery size (35 kWh) and peak-load reduction potentials are investigated (-20%/ -50%) to estimate the effect of increased demand flexibility.

Table 5.4. Standard model parameters and sensitivities for the EV and PV model.

Technology		Standard parameters	Sensitivity analysis
EV	Car ownership ratio	0.45	0.45
	Charging power (kW)	5.89	9.82
	Battery size (kWh)	25	35
	Simultaneity factor	0.5	0.75
PV	Capacity/capita (kW)	0.4	0.4
	Usable roof fraction	0.3	0.3
	Peak-load reduction	- 6%/- 20%	- 20%/-50%

5.4 NETWORK EXPANSION COSTS UNDER LARGE-SCALE DER ADOPTION

Estimating network expansion costs without DER adoption

As first step, the network expansion costs considering only load growth and neglecting DER adoption are estimated. Such base case scenario assumes a peak-load growth at 0.5% annually, close to the major Portuguese distribution system operator (DSO) investment plan [62]. This major national DSO supplies around 99% of the customers that constitute the Portuguese Continental distribution system.

The load growth is equally applied to all 391 HV/MV substations, which possess an aggregated installed capacity of 15.447 MW. HV/MV substation characteristics (installed capacity, peak-load values and coordinates) have been extracted from [62], [63], while some values have been corrected after consulting with the major national distribution company. Likewise, typical HV/MV transformer upgrade costs (***TC***) were identified in the same DSO's investment reports [62], [63]. Those reports also stated the network expansion threshold applied on Continental Portugal for upgrading HV/MV transformers upgrades.

As a rule, reinforcement is triggered if the transformer peak-load (or peak-net-load ***PNL***) surpasses 90% of the installed transformer capacity (Eq. 5.6).

$$TC = \begin{cases} 2.5\, million\ Euro \ \ for\ PNL \geq \ 0.9 \\ 0.0\, million\ Euro\ for\ PNL < \ 0.9 \end{cases} \tag{5.8}$$

For the base case, model outcomes suggest system expansion costs, approximated through aggregated HV/MV transformer costs, totalling 22.5 million Euros over a 20-year time horizon. It is noteworthy that roughly one third of these costs (7.5 million Euros) is invested in the first 10 years. The remaining 15 million Euros are needed for the second planning horizon. The summed investment costs and investment timing under the base case (only load growth) are further considered in the comparison of investment requirements under different incentive design combinations (Figure 5.6.), which is detailed in one of the following subchapters.

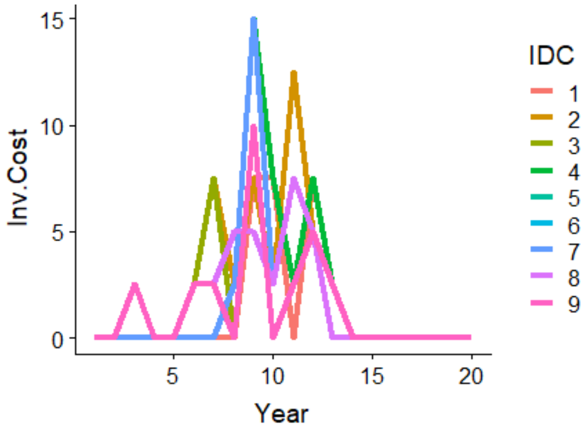


Figure 5.6. Investment time series of each incentive design combination.

Please note that the investment cost used in this chapter is slightly above the values used in Chapter 5 and reflects the evolution of costs under the most recent DSO development plan.

Mapping HV/MV transformer expansion under IDC

A major outcome of the presented model are map stacks that display HV/MV transformer loadings under different energy policy designs. In this way, the presented case study enlightens spatial HV/MV expansion patterns under different EV and PV incentive design combinations.

The presented model computes 180 maps in an automated way that correspond to the combination of 20 analysed years and 9 IDC. These maps can be read and combined with DSO's asset management system, exploiting a GIS-based decision support tools used by network planners.

Like in the base case (only load growth), analysis outcomes under DER adoption suggest again high investments during years 7-12 ahead of the base year (2015). This is expected as the global DER forecast suggests the largest share of DER being adopted within this time horizon. It is noteworthy that, under most IDC, a major network expansion is foreseen, expressed through investments of 10 - 15 million Euros in years 9 and 10 (Figure 5.6.). On the other hand, under certain incentive designs, more concentrated investment requirements can be observed (e.g. for IDC2 or IDC6). In contrast, other IDC show investments needs being stretched across several years (IDC8 and IDC 9). Latter IDC might be favoured by network operators and regulation agencies which would probably allow for a more constant spending (and thus, more distributed weight on the network tariff) over a given regulation period.

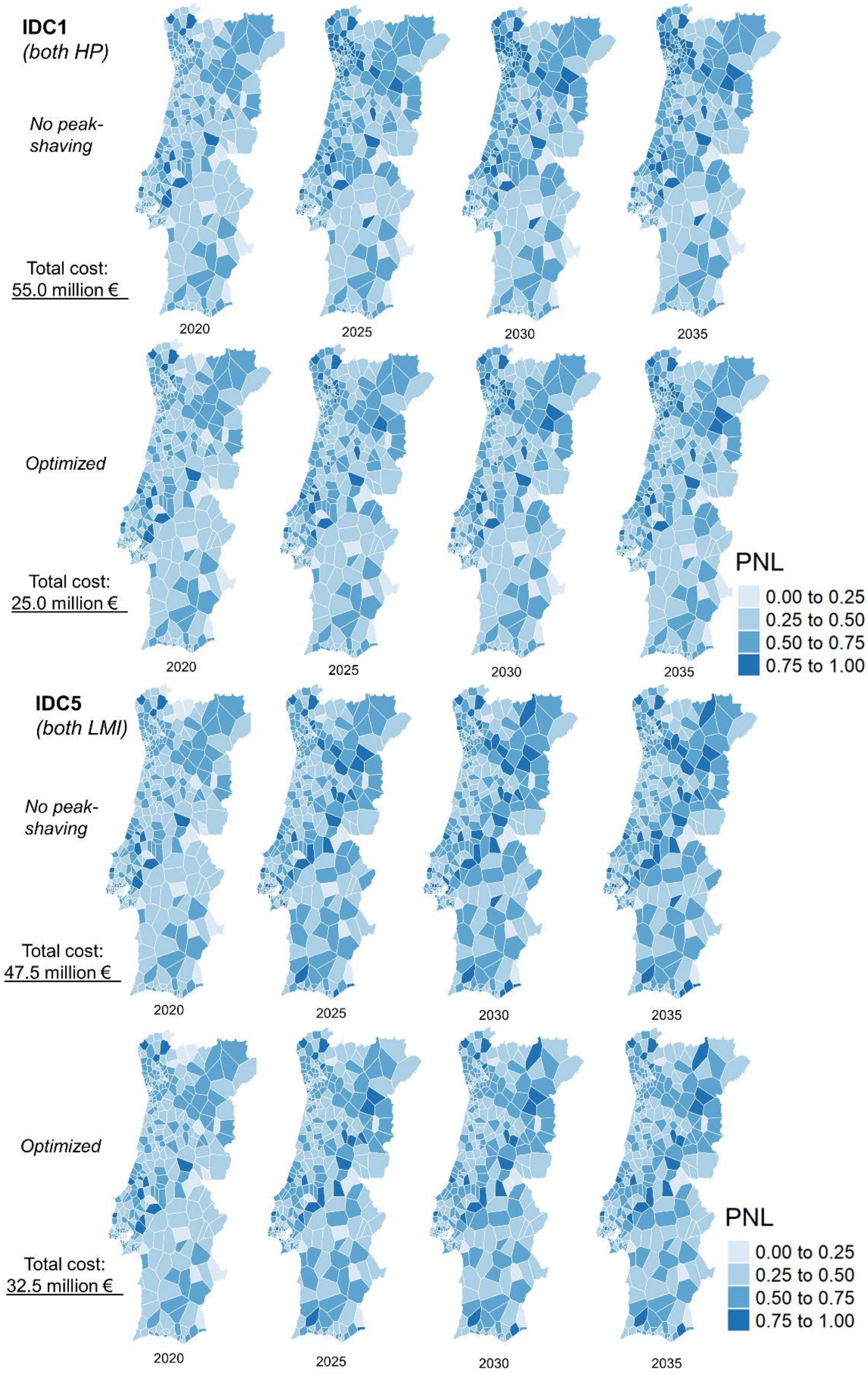


Figure 5.7. Spatial network expansion patterns for High-performer EV – High-performer PV (IDC1) and Low-Medium-Income EV – Low-Medium-Income PV (IDC5) incentive designs.

Mapping HV/MV transformer loadings reveals a trend towards higher peak-net-loads along the adoption process (Figure 5.7). Under IDC1, higher HV/MV transformer loadings are observed along the coastline. On the other hand, stronger loadings in the interior areas of Continental Portugal (e.g. areas close to the continental Spanish border) can be observed under IDC5. Latter can be explained by a stronger geographical concentration of household shares classified as LMI (IDC5 is the LMI-EV and LMI-PV incentive design combination). Likewise, transformer loadings under IDC1 are in line with expectations as high-income households with above-average education levels (HP groups) spread stronger along the coast lines that accommodate most urban areas with a higher access to higher education institutions.

In the following, the impact of PV and battery induced peak-load shaving on network expansion costs is analysed (Figure 5.7.). Here, results for the two extreme incentive designs (IDC1 – HP-HP and IDC5 – LMI-LMI) are displayed. As a general outcome, results suggest that reduced peak-loads (and thus, reduced investment), can be achieved for both cases (IDC1 and IDC5). Results further show that higher investment reduction is achieved under orchestrated incentive designs that target HP groups (IDC1). Under this IDC, global network expansion costs can be reduced by more than 50% (from 55 to 25 million Euros). Likewise, a potential network expansion cost reduction of 15 million Euros (from 47.5 to 32.5 million Euros) can be realized under LMI-oriented DER incentive designs (IDC5).

One explanation of the differences in IDC1 and IDC5 may be that HP adopters mostly locate in urban HV/MV service areas along the coast-line. Such areas tend to be densely populated and therefore rapidly reach the expansion thresholds (***PNL*** of 0.9) with EV adopters driving peak-net-loads upwards. Therefore, such zones would possess relatively higher reduction potential under HP schemes that consider peak-load reduction programs.

Another explanation could lie in a higher concentration of HP adopters in HV/MV service areas with a lower hosting capacity for added EV charging. However, an extensive, case-by-case analysis of load growth patterns and evolution ***PNL*** in all 391 HV/MV service areas lies outside the scope of this analysis and remains future work.

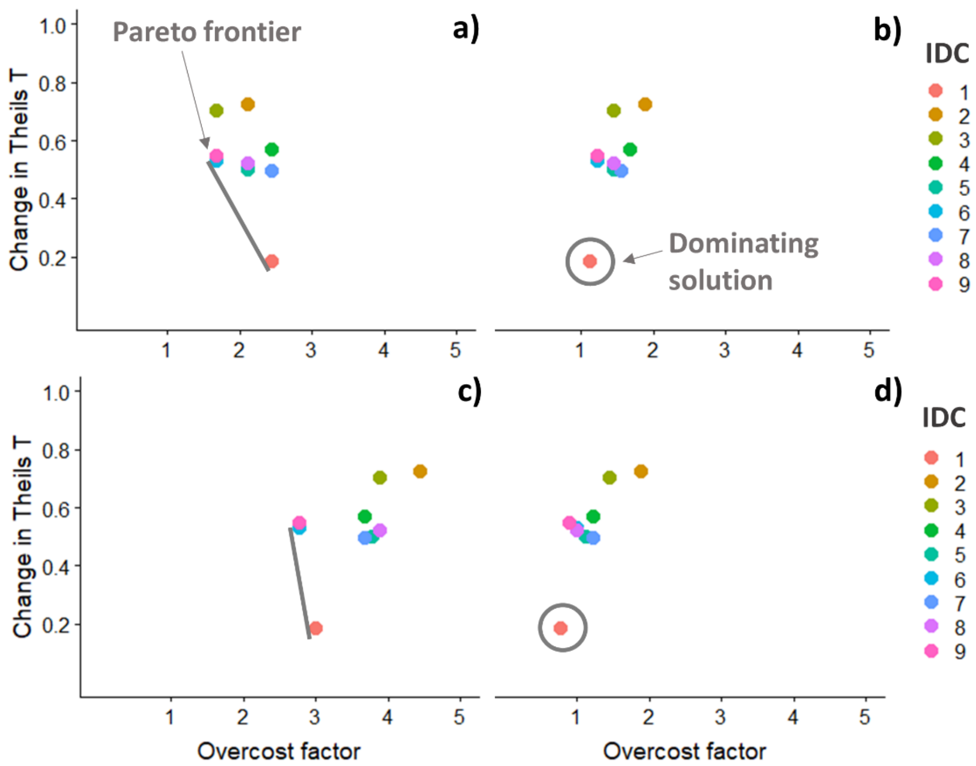


Figure 5.8. Mapping of IDC under different scenarios for modeling year 20.

5.5 Adoption asymmetries against expansion costs

The estimation of network expansion costs and adoption asymmetries further allows to visualize the analysed IDC in a Cartesian coordination system (Figure 5.8). Given that value pairs of CT and overrun cost estimates are available at this step, trade-offs can now be calculated. The overcost factor is defined as the ratio of the aggregated expansion costs of each IDC for 20 modelling years divided by expansion costs under baseline conditions (load growth only). In addition, the Pareto frontier, which is the line delineating the most efficient IDC under the criteria considered, can be plotted.

The nine IDC are further analysed considering four sensitivities. As shown in Fig. 5.8., Case a) displays a situation of non-controlled EV charging (dumb charging). On the other hand, Case b) shows model outcomes for peak-shaving using PV and batteries (compare also Eq. 5.3). Cases c) and d) show results assuming an increased average charging rate of 9.82 kW EV and 35 kWh battery size (c), whereas Case d) analyses a potential HV/MV transformer peak-load reduction of -20% respectively -50%.

For all four sensitivities considered (normal conditions, optimized with EV-PV bundling, high EV charging, high PV & EV battery coupled peak-load shaving), absolute performance and their relative differences strongly vary among all IDC for each scenario analysed. Explanations may be found in the difference of spatial HP group and LMI group distributions.

Under the base cases (a) and (c), IDC1 and IDC6 are the dominating of Pareto-efficient solutions. They are dominating the remaining IDC in terms of both asymmetries and overrun costs. According to the model outcomes, IDC1 (HP groups-oriented EV and PV incentives) introduces the lowest asymmetries into DER adoption patterns. IDC1 also dominates all other IDC under sensitivities b) and d), which correspond to situations under light and strong peak-load reduction. In(a), IDC5 (LMI groups-oriented EV and PV incentives) is typically dominated by IDC6. Under sensitivities a) and c), IDC9 (double random allocation) introduces smaller overrun costs than IDC1 but is dominated by IDC6.

Future work could consider the influencing role of initial transformer loadings (t=0) or spatial correlation metrics to analyse the spatial evolution of ***PNL*** over time. In addition, established geostatistical metrics (e.g. Moran's I) could be used to link DER adoption patterns to system expansion needs.

Table 5.5. IDC rankings across scenarios for the original census dataset.

IDC	1	2	3	4	5	6	7	8	9
Scenarios a) - d)									
Euclidean	1	9	8	7	3	2	4	6	5
Manhattan	1	9	8	7	4	2	5	6	3

Eventually, in order to rank the analysed IDC, a single criterion is developed for each policy design. In other words, the IDC's performance in terms of system expansion overrun costs and distributional effects is assessed through an aggregation of both into a single metric. While the decision maker can in practice allocate predefined weights to each criterion that reflect the preference of one criterion over the other, an equal-weight scheme has been used in this analysis. Therefore, the search for the IDC that has the minimum expansion overrun cost as well as introduces least DER adoption pattern asymmetries is conducted using the Euclidean and Manhattan distance to the Ideal. Here, the Ideal is defined as the virtual point with the lowest overrun cost and asymmetry change observed for a set of IDCs, that is, the point where the change in Theil's T and expansion overrun cost are minimal. Distances are averaged for all four sensitivities considered.

Outcomes in Table 5.5. suggest that IDC1 is clearly the incentive design combination that is the closest to the Ideal under the defined criteria and distance functions. In other words, the compromise set resulting from choosing any metric from L_1 to L_2 is comprised of a single solution, IDC1, which may therefore be considered as a very robust solution regarding the choice of metric (the decision does not change, regardless o metric choice), so metric choice is a non-issue. This would not be the case if the closest-to-Ideal solution would have changed with a change in metric, leading to the need of further considering a diversity of aspects in the decision-making process in this specific problem, in order to build arguments in favour of one choice or the other.

The examination of the second-best solutions is also usually revealing. For instance ID6, which is EV incentives for RN consumer groups and LMI-targeting PV incentives, is the second closest to the Ideal. Furthermore, depending on the distance function, IDC9, which represents randomized EV and PV adoption across all consumer groups, can rank third position. These outcomes confirm again that under an equal weighting of network expansion over-cost and DER asymmetry changes, IDC targeting HP for both EV/PV or randomized adoption may be attractive both to energy policy designers and network planners. While former comes unexpected (as discussed above), latter (IDC6) is well in line with the intuition that maximally different spatial adoption patterns of DER, that is, non-overlapping EV and PV adoption patterns, would not allow to exploit potential synergies (e.g. peak-load shaving). Instead, almost uniform DER patterns may, in return, reduce network expansion costs..

Robust incentive design choice through census data permutation

Under the current socio-demographic population structure and its spatial distribution IDC1, followed by IDC6 and then by IDC5 and IDC9, are attractive policy choices to decision maker. However, the question arises if this holds for other population compositions or different spatial distributions of LMI and HP consumer groups. Therefore, the impact of changes in the underlying socio-demographic population structure on optimal incentive designs has been analysed.

Table 5.6. IDC rankings across scenarios and permutated census datasets.

IDC	1	2	3	4	5	6	7	8	9
Permutated. census data									
Euclidean	1	7	9	8	4	3	5.5	5.5	2
Manhattan	1	7	9	8	4	3	5.5	5.5	2

In order to assess the influence of structural changes in the population structure, the allocation key that links census cells to HV/MV transformers is randomly modified. In total, 20 of such permutations are computed, while comparing the performance and ranking of all nine IDC. Outcomes show that under such conditions, IDC1 would be again the optimal incentive design under which both expansion overrun cost and adoption asymmetries are minimized. In addition, IDC9 (the randomized distribution of DER) becomes a high-ranked incentive scheme combination, in position 2 right after IDC1. Such ranking is unchanged for both Euclidean and Manhattan distances (Table 5.6.).

A comparison of the IDC rankings across all 20 scenarios shows that indeed, IDC1 and IDC9 are the highest ranked IDC, with a mean rank of 3.7 and 4.15 respectively, if Euclidean distance is considered. The same analysis using Manhattan distance reveals a very similar pattern with only minor deviations in mean rank values <10%. However, retrieved mean rank vaules also suggest that IDC1 and IDC9 are not always the outperforming policy options. Both IDC are among Top-3 ranks in half of the simulations, although there are a few permuations where these two IDC would obtain ranks 4-9. This highlights the outcomes' reliance on the distribution of population groups in each case study. Policy makers or decision makers may extend the presented studies to a Monte-Carlo approach, realising more extensive permutations and decision framework coupled with risk analysis and regret value optimization. Such studies are outside the scope of this thesis and represent future work.

Furthermore, the outcomes of 20 census data permutations (assuming equal probability of realization) show that IDC 2, IDC3, IDC4, IDC7 and IDC8 shall be of little interest to energy policy makers that envision reduced expansion overrun cost and low DER adoption asymmetries (Figure 5.9).

All such incentive design combinations are dominated options due to an uneven attribution of both EV and PV within all HV/MV service areas which tend to lead to relatively pronounced global system expansion costs.

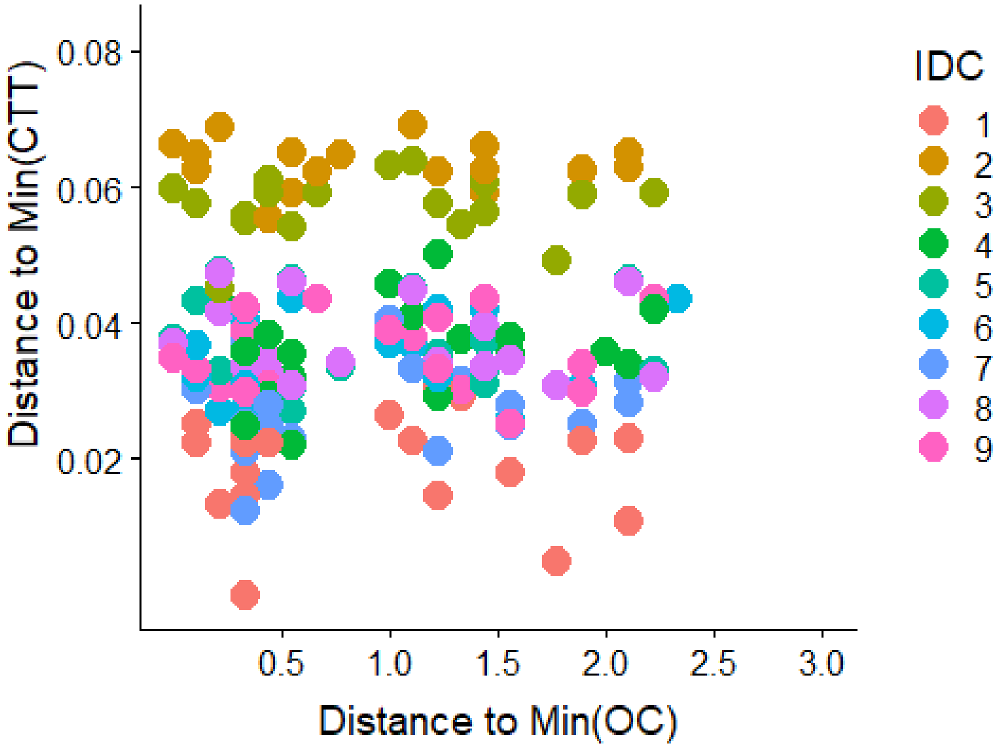


Figure 5.9. Solution space under 5 permutated census datasets.

While the presented model allows to simulate effects of incentive designs on network expansion costs and DER adoption asymmetries, a real-world implementation of the methodology would require further improvement.

One vision is that distribution companies, governmental agencies or regulatory authorities might further improve the accuracy of the spatiotemporal DER adoption forecasting model. This might be achieved by replacing the weight vectors (innovativeness scores) with methods that exploit recent advances in data science. Furthermore, the spatial simulation model could be extended to include neighbourhood interaction, that is, the way DER adoption is influenced by adjunct households.

In addition, extensions could include a more detailed expansion cost model that would consider different transformer sizes and costs as well as other network elements (e.g. lines) that drive expansion costs.

Finally, the impact of various IDC on distributional justice aspects shall be further analysed, breaking down the impact of DER adoption on different population subgroups (e.g. HP, LMI) while considering different dimensions on distributional fairness [41].

Chapter conclusions

This chapter presented a methodology to assess the impact of future EV/PV incentive design combinations onto network expansion costs and the dimension of distributional justice. After reviewing recent studies on the interplay of technology incentive designs and electricity network planning and incentive designs in place, the chapter introduces nine IDC. Such combinations represent a mix of three different incentive designs that are applied to both EV and PV: Incentives that trigger adoption of high-educated, above-average income groups (***HP***), lower education levels and low-medium income groups (***LMI***) and randomized adoption (***RN***). These synthetic incentive designs allow to analyse the effects that result from different spatial adoption patterns under large-scale DER adoption. This way, outcomes allow decision makers to weight estimated network expansion costs against DER adoption asymmetries. The analysis of trade-offs, and the identification of Pareto optimal solutions, further allows to discard incentive designs that would be dominated by other solutions. The research outcomes presented in this chapter can be summarized in the following way:

- DER adoption is accelerated by incentive designs in place.
- Orchestrated incentive design schemes could facilitate the realization of prosumer paradigm.
- A system-wide study of the grid investment needs of various DER incentive designs that are currently discussed among researchers and policymakers is presented.
- The analysis is based on the spatiotemporal DER adoption forecasting model that uses high-resolution census data.
- The presented methodology allows to simulate adoption patterns under different, synthetic incentive designs (high-performer, low-medium income, randomized).
- Outcomes from simulating various EV/PV IDC suggest that estimated network expansion costs can be reduced without adverse effects on DER adoption asymmetries.
- Eventually, the presented work opens a path to analyse network expansion cost estimations and distributional effects of DER incentives based on spatial adoption patterns.

Chapter references

[1] E. M. Rogers, *Diffusion of innovations*. 2003.

[2] L. M. M. Benvenutti, A. B. Ribeiro, F. A. Forcellini, and M. U. Maldonado, "The Effectiveness of Tax Incentive Policies in the Diffusion of Electric and Hybrid Cars in Brazil," *41st Congr. Latinoam. Din. Sist. São Paulo*, Sao Paulo, 2016.

[3] D. Diamond, "The impact of government incentives for hybrid-electric vehicles: Evidence from US states," *Energy Policy*, vol. 37, no. 3, pp. 972–983, 2009.

[4] K. Y. Bjerkan, T. E. Nørbech, and M. E. Nordtømme, "Incentives for promoting Battery Electric Vehicle (BEV) adoption in Norway," *Transp. Res. Part D Transp. Environ.*, vol. 43, pp. 169–180, 2016.

[5] A. A. Radomes and S. Arango, "Renewable energy technology diffusion: An analysis of photovoltaic-system support schemes in Medellín, Colombia," *J. Clean. Prod.*, vol. 92, pp. 152–161, 2015.

[6] Council of European Energy Regulators, "Status Review of Renewable Support Schemes in Europe for 2016 and 2017 Public report," Brussels, 2018.

[7] C. Batlle, I. J. Pérez-Arriaga, and P. Zambrano-Barragán, "Regulatory design for RES-E support mechanisms: Learning curves, market structure, and burden-sharing," *Energy Policy*, vol. 41, pp. 212–220, 2012.

[8] S. C. Huntington, P. Rodilla, I. Herrero, and C. Batlle, "Revisiting support policies for RES-E adulthood: Towards market compatible schemes," *Energy Policy*, vol. 104, no. August 2016, pp. 474–483, 2017.

[9] L. De Boeck, S. Van Asch, P. De Bruecker, and A. Audenaert, "Comparison of support policies for residential photovoltaic systems in the major EU markets through investment pro fi tability," *Renew. Energy*, vol. 87, pp. 42–53, 2016.

[10] Council of European Energy Regulators, "Key support elements of RES in Europe : moving towards market integration CEER report," Brussels, 2016.

[11] P. Haugneland, E. Hauge, "Norwegian electric car user experiences 2014," *EVS28 International Electric Vehicle Symposium and Exhibition*, Korea, 2014.

[12] Electric Vehicle Initiative (IEA/OECD), "Global EV Outlook 2017," Paris, 2015.

[13] European Alternative Fuels Observatory, "Electric vehicle statistics," 2017. [Online]. Available: http://www.eafo.eu/vehicle-statistics/m1. [Accessed: 01-Jan-2017].

[14] T. Lieven, "Policy measures to promote electric mobility - A global perspective," *Transp. Res. Part A Policy Pract.*, vol. 82, pp. 78–93, 2015.

[15] European Parliament; European Council, "DIRECTIVE (EU) 2018/844 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency," *Off. J. Eur. Union*, no. May, 2018.

[16] European Alternative Fuels Observatory, "Portugal - Country detail incentives," 2018. [Online]. Available: https://www.eafo.eu/countries/portugal/1749/incentives.

[17] T. D. Chen, "Where are the electric vehicles? A spatial model for vehicle-choice count data," *J. Transp. Geogr.*, vol. 43, pp. 181–188, 2015.

[18] V. Vasseur and R. Kemp, "The adoption of PV in the Netherlands: A statistical analysis of adoption factors," *Renew. Sustain. Energy Rev.*, vol. 41, pp. 483–494, 2015.

[19] C. Lee, "Influence of local environmental , social , economic and political variables on the spatial distribution of residential solar PV arrays across the United States," *Energy Policy*, vol. 47, pp. 332–344, 2012.

[20] J. Saarenpää, M. Kolehmainen, and H. Niska, "Geodemographic analysis and estimation of early plug-in hybrid electric vehicle adoption," *Appl. Energy*, vol. 107, pp. 456–464, 2013.

[21] D. S. Noonan, L.-H. C. Hsieh, and D. Matisoff, "Economic, sociological, and neighbor dimensions of energy efficiency adoption behaviors: Evidence from the U.S residential heating and air conditioning market," *Energy Res. Soc. Sci.*, vol. 10, pp. 102–113, 2015.

[22] F. Bartiaux, L. Schmidt, A. Horta, and A. Correia, "Social diffusion of energy-related practices and representations: Patterns and policies in Portugal and Belgium," *Energy Policy*, vol. 88, pp. 413–421, 2016.

[23] O. Egbue and S. Long, "Barriers to widespread adoption of electric vehicles: An analysis of consumer attitudes and perceptions," *Energy Policy*, vol. 48, no. 2012, pp. 717–729, 2012.

[24] E. Karakaya and P. Sriwannawit, "Barriers to the adoption of photovoltaic systems: The state of the art," *Renew. Sustain. Energy Rev.*, vol. 49, pp. 60–66, 2015.

[25] M. Yaqoot, P. Diwan, and T. C. Kandpal, "Review of barriers to the dissemination of decentralized renewable energy systems," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 477–490, 2016.

[26] D. Grover and B. Daniels, "Social equity issues in the distribution of feed-in tariff policy benefits: A cross sectional analysis from England and Wales using spatial census and policy data," *Energy Policy*, vol. 106, no. March, pp. 255–265, 2017.

[27] R. Passey, M. Watt, A. Bruce, and I. Macgill, "Who pays , who benefits ? The financial impacts of solar photovoltaic systems and air-conditioners on Australian households," *Energy Res. Soc. Sci.*, vol. 39, no. November 2017, pp. 198–215, 2018.

[28] M. Kubli, "Squaring the sunny circle? On balancing distributive justice of power grid costs and incentives for solar prosumers," *Energy Policy*, vol. 114, no. November 2017, pp. 173–188, 2018.

[29] C. Eid, J. Reneses Guillén, P. Frías Marín, and R. Hakvoort, "The economic effect of electricity net-metering with solar PV: Consequences for network cost recovery, cross subsidies and policy objectives," *Energy Policy*, vol. 75, pp. 244–254, 2014.

[30] A. Picciariello, C. Vergara, J. Reneses, P. Frías, and L. Söder, "Electricity distribution tariffs and distributed generation: Quantifying cross-subsidies from consumers to prosumers," *Util. Policy*, vol. 37, pp. 23–33, 2015.

[31] A. C. Mersky, F. Sprei, C. Samaras, and Z. S. Qian, "Effectiveness of incentives on electric vehicle adoption in Norway," *Transp. Res. Part D Transp. Environ.*, vol. 46, pp. 56–68, 2016.

[32] H. Duan, G. Zhang, S. Wang, and Y. Fan, "Peer interaction and learning: Cross-country diffusion of solar photovoltaic technology," *J. Bus. Res.*, vol. 89, no. August 2017, pp. 57–66, 2018.

[33] R. Prata and P. M. S. Carvalho, "Self-supply and regulated tariffs : Dynamic equilibria between photovoltaic market evolution and rate structures to ensure network sustainability," *Utilities Policy*, vol. 50, no. December 2017, pp. 111–123, 2018.

[34] J. Villavicencio Gastelu, J. D. Melo Trujillo, and A. Padilha-Feltrin, "Hierarchical Bayesian Model for Estimating Spatial-Temporal Photovoltaic Potential in Residential Areas," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 971–979, 2018.

[35] R. Hemmati, R. A. Hooshmand, and N. Taheri, "Distribution network expansion planning and DG placement in the presence of uncertainties," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 665–673, 2015.

[36] J. Pecas Lopes, F. J. Soares, and P. M. R. Almeida, "Integration of Electric Vehicles in the Electric Power System," *Proc. IEEE*, vol. 99, no. 1, 2011.

[37] A. Rodriguez-Calvo, R. Cossent, and P. Frías, "Integration of PV and EVs in unbalanced residential LV networks and implications for the smart grid and advanced metering infrastructure deployment," *Int. J. Electr. Power Energy Syst.*, vol. 91, pp. 121–134, 2017.

[38] Y. Mu, J. Wu, N. Jenkins, H. Jia, and C. Wang, "A Spatial – Temporal model for grid impact analysis of plug-in electric vehicles," *Appl. Energy*, vol. 114, pp. 456–465, 2014.

[39] R. Bernards, J. Morren, and H. Slootweg, "Development and Implementation of Statistical Models for Estimating Diversified Adoption of Energy Transition Technologies," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1540–1554, 2018.

[40] M. De Gennaro, E. Paffumi, H. Scholz, and G. Martini, "GIS-driven analysis of e-mobility in urban areas: An evaluation of the impact on the electric energy grid," *Appl. Energy*, vol. 124, pp. 94–116, 2014.

[41] H. Granqvist and D. Grover, "Distributive fairness in paying for clean energy infrastructure," *Ecol. Econ.*, vol. 126, pp. 87–97, 2016.

[42] J. J. Cook and L. Bird, "Unlocking Solar for Low- and Moderate-Income Residents: A Matrix of Financing Options by Resident, Provider, and Housing Type," Golden (CO), 2018.

[43] J. Espino and V. Truong, "Electric Carsharing in Underserved Communities Underserved Communities," Berkeley, 2015.

[44] T. Islam and N. Meade, "The impact of attribute preferences on adoption timing: The case of photo-voltaic (PV) solar cells for household electricity generation," *Energy Policy*, vol. 55, pp. 521–530, 2013.

[45] S. Hardman, E. Shiu, and R. Steinberger-Wilckens, “Comparing high-end and low-end early adopters of battery electric vehicles,” *Transp. Res. Part A Policy Pract.*, vol. 88, pp. 40–57, 2016.

[46] S. A. Robinson and V. Rai, “Determinants of spatio-temporal patterns of energy technology adoption: An agent-based modeling approach,” *Appl. Energy*, vol. 151, pp. 273–284, 2015.

[47] P. Plötz, U. Schneider, J. Globisch, and E. Dütschke, “Who will buy electric vehicles? Identifying early adopters in Germany,” *Transp. Res. Part A Policy Pract.*, vol. 67, pp. 96–109, 2014.

[48] B. M. Sopha, C. A. Klöckner, and E. G. Hertwich, “Adoption and diffusion of heating systems in Norway: Coupling agent-based modeling with empirical research,” *Environ. Innov. Soc. Transitions*, vol. 8, pp. 42–61, 2013.

[49] N. Ameli and N. Brandt, “Determinants of Households' Investment in Energy Efficiency and Renewables,” *Environmental Research Letters*, vol. 10, 044015, 2014.

[50] K. X. Perez, M. Baldea, and T. F. Edgar, “Integrated HVAC management and optimal scheduling of smart appliances for community peak load reduction,” *Energy Build.*, vol. 123, pp. 34–40, 2016.

[51] R. L. Fares and M. E. Webber, “The impacts of storing solar energy in the home to reduce reliance on the utility,” *Nat. Energy*, vol. 2, no. January, 2017.

[52] W. L. Schram, I. Lampropoulos, and W. G. J. H. M. van Sark, “Photovoltaic systems coupled with batteries that are optimally sized for household self-consumption: Assessment of peak shaving potential,” *Appl. Energy*, vol. 223, no. April, pp. 69–81, 2018.

[53] J. F. Soares, P. Rocha Almeida, and J. Pecas Lopes, “Quasi-real-time management of Electric Vehicles charging,” *Electr. Power Syst. Res.*, vol. 108, pp. 293–303, 2014.

[54] UN, “Inequality Measurement: Development Issues No. 2,” *United Nations*, no. 2, p. 2, 2015.

[55] World Bank Institute, “Chapter 6. Inequality Measures,” *Introd. To Poverty Anal.*, pp. 95–105, 2005.

[56] J. Novotný, "On the measurement of regional inequality: Does spatial dimension of income inequality matter?," *Ann. Reg. Sci.*, vol. 41, no. 3, pp. 563–580, 2007.

[57] M. Emmerich and A. Deutz, "A tutorial on multiobjective optimization: fundamentals and evolutionary methods," *Natural Computing*, vol.17, pp. 585–609, 2018.

[58] R. Shahid, S. Bertazzon, M. Knudtson, W. Ghali, "Comparison of distance measures in spatial analytical modeling for health service planning," *BMC Health Services Research*, vol. 9 (200), 2009.

[59] ENTSO-E, "TYNDP 2018: Scenario Report," Brussels, 2018.

[60] Portuguese General Direction of Energy and Geology (DGEG), "Rapid Statistics - Renewables No.168," Lisbon, 2018.

[61] E. Gómez-Rubio, Virgilio; Bivand, Roger; Pebesma, *Applied Spatial Data Analysis with R*, Second Edi. 2013.

[62] EDP Distribuição, "Distribution Network Development and Investment Plan 2017-2021 (in Portuguese)," Lisbon, 2016.

[63] EDP Distribuição, "National distribution grid - High Voltage and substations," 2016. [Online]. Available: http://edp-distribuicao-rede.wntech.com/Nut/4. [Accessed: 05-Sep-2017].

6 Conclusions

This final section provides a condensed overview of the innovations presented in this thesis. In addition, answers to the research questions initially stated are presented, whereas newly emerging, open questions and future research avenues are discussed.

6.1 INNOVATIONS OF THIS THESIS

The thesis presented a holistic framework to characterize, compare and predict technology adoption patterns in space and time. After introducing a spatiotemporal DER adoption forecasting model, a rigorous comparison of various models to represent technology adoption dynamics in electricity network planning and energy policy studies were presented. Eventually, a system-level analysis on policy designs was constructed on previous findings.

The thesis presented innovations on a broad range of subjects. Such are:

- The **development of a holistic technology adopter analysis, merging spatial and non-spatial analysis** tools. In Chapter 2, DER adopters have been assessed using spatial autocorrelation metrics (Moran's I) and census-based inference algorithms (Information Gain Ratio). The innovative approach allowed to characterize adoption drivers and incentive designs while linking them to spatial adoption patterns.

- Building on the developed inference tools, Chapter 3 introduced a **census data-driven spatiotemporal DER adoption forecasting model**. Relying exclusively on population data and DER observations, the forecasting model can also easily incorporate different policy designs. That way, the proposal allows bridging the gap of spatial analysis, energy policy design and network planning.

- Merging locational information of HV/MV transformers and tabular information on transmission-to-distribution connectivity, Chapter 4 **presents a fast and automatable way to approximate transmission service areas** through HV/MV HV/MV distribution polygons. The spatial distribution and transmission network models are useful inputs for capacity expansion screening, uncertainty assessment or the spatial decomposition of scenario forecasts.

- A **straightforward way to estimate the trade-offs between distributional justice and network expansion costs** was presented in Chapter 5. The approach numerically assesses adoption asymmetry patterns, using an Information Theoretic inequality measure, and allows decision makers to weight economic costs against inequality reduction measures.

The innovations of this thesis are complemented by further, principal research outcomes that are detailed in the following section.

6.2 PRINCIPAL RESEARCH OUTCOMES

The research that culminated into this thesis has brought a variety of outcomes that respond to the initial research questions (compare Chapter 1). While a detailed overview of retrieved results is annexed to each chapter, a condensed summary of the responses to the stated research questions is provided in the following.

Chapter 2:

How can one describe and compare technology adoption patterns?	Technology adoption may be described by diffusion patterns, which can be represented using non-spatial analysis tools together with adopter data, crossed with census or survey information. Given that such information allows characterizing population groups, adoption drivers can be inferred. Convenient approaches to infer such drivers are algorithms that combine multi-linear regression and relative importance, artificial neural networks and e.g. the Olden approach or information theoretic measures such as the Information Gain Ratio. While such inference methods allow estimating the propensity of consumer groups to adopt a certain technology, the technologies' distribution in space is fundamental to fully understand diffusion phenomena. Given that population groups distribute heterogeneously in space, different strengths in adoption drivers result in varying DER adoption patterns. Therefore, analyses that combine spatial and non-spatial methods are preferred. Eventually, the analysis outcomes are strongly affected by data aggregation. Therefore, decision makers shall carefully select data aggregation levels. Given the scope of analysis, such aggregation levels shall preferably be closest to ground truth of observations.

Chapter 3:

How to predict technology adoption patterns in space and time?

Spatiotemporal DER adoption models are a recent, emerging family of DER adoption models. Main model typologies are spatial regression models, simulation-based models and agent-based models. The latter two model types are more computationally demanding but allow modelling dynamic behaviour at cellular level.

While agent-based DER adoption models usually rely on cell lattice data, simulation-based DER adoption models allow to flexibly adjust to the spatial data substrate. On the downside, neighbourhood interaction is typically constrained to strong simplifications.

Mostly, such models incorporate detailed data on socio-economic population structure, which can be found in census data or in geocoded, empirical surveys. In addition, some spatiotemporal DER adoption models build their prediction on data-driven analysis of spatial DER adoption patterns.

Which components do spatiotemporal technology adoption models typically consist of?

All spatiotemporal technology adoption models, such as DER adoption models, consist of three main components.

First, technology quantities (e.g. EV) for a given time horizon are predicted using a global stock forecasting module.

Then, the aggregated amount of a defined technology is then fed into a cellular adopter module. Such module determines the uptake behaviour inside a given spatial cell, considering population, propensity of population groups, and maximum adoption levels. The uptake is typically modelled through a S-curve model, considering discretized adoption states.

Finally, the temporal dynamics are translated into spatial distributions with the help of an adoption pattern mapping module. Such module usually sequences cells along a determined adoption order.

Along which criteria can technology adoption models be categorized?	Spatiotemporal adoption forecasting models are the most recent added category of technology adoption models that received major attention throughout this thesis. However, a variety of representations of DER adoption dynamics could be identified. Such models have been built on time series, Bass diffusion, combined market penetration or macroeconomic approaches. All technology adoption models can be categorized along their spatial and temporal axis. Depending on the character of the technology adoption model, one can separate non-spatial and non-temporal models, spatial and non-temporal models, non-spatial and non-temporal models and spatial and temporal (herein named "spatiotemporal") models.

Chapter 4:

What are the effects of different technology representations on electricity network planning?	Currently, transmission and distribution planners rely on mostly very simplified approaches to represent DER adoption dynamics. Such include extrapolations based on installed busbar capacities, peak-load or equal/random allocations across the networks. As such approaches do not consider the underlying population characteristics and the adoption propensities by different social groups and structures, such representations add additional uncertainty to the planning process. Model outcomes presented in Chapter 4 show that such simplified approaches may underestimate the impact of DER on the peak load in early adoption phase, while overestimate their effect for late adoption phases. Especially in transmission studies, simplified DER allocation techniques result in large uncertainties at T/D boundary that can reach DER forecast variations of over 100 MW per transmission service area (e.g. for the example of residential PV module capacities).

Chapter 5:

How do different policy designs affect system expansion costs and distributional effects?	The model proposed in Chapter 5 shows a way to model spatial DER adoption patterns for different policy designs. This, eventually, allows comparing system expansion costs with adoption asymmetry patterns. The results obtained suggested that, due to the geographical dispersion of population groups, different incentive designs lead indeed to different spatial DER adoption patterns. In return, this does result, depending on the spatial concentration of DER adopters, in a variety of system expansion cost estimates and adoption asymmetries.
Can orchestrated incentive designs reduce such costs/ distributional effects?	The results suggest that orchestrated incentive designs, depending on the decision makers' weight allocation, can either reduce system expansion costs or effects of DER adoption asymmetry. In fact, outcomes of the analysed case study suggest that current incentive design schemes lead to relatively concentrated adoption patterns. On the contrary, policy designs that favour randomized DER adoption or uptake through low- and medium-income groups may reduce system expansion costs.

Finally, all the above-listed research outcomes together show how, on global level, the representation and modelling of technology diffusion dynamics in electricity network planning and policy design studies can be enhanced to address emerging challenges.

6.3 FUTURE WORK

The innovations brought by this thesis, tested and applied to a set of case studies, not only produced valuable results and insight, but also opened the way to further improvement and model extensions. A short overview of current model limitations and a flavour of such foreseen extensions is provided below.

Throughout this thesis, we assessed the adoption dynamics and, in particular, their representation in electricity network studies and policy design. While this thesis has covered how incentive designs and different adoption drivers of population groups drive the *spatial diversity* of DER utilization, more focus may be put on the *temporal* utilization patterns of DER.

In fact, while adoption patterns have been spatially differentiated, the underlying models were based on deterministic load, EV charging and PV generation profiles that assume similar user behaviour across all HV/MV substation service areas. Consequently, the extension to a probabilistic simulation-model represent a logical, following step. This would both include the modelling of stochastic time series (e.g. different EV arrival times, PV generation patterns) as well as stochasticity in spatial adoption behaviour. Given an increased amount of DER adopter observations, the stochastic representation should further include the interaction of neighbouring DER adopters, modelling what has been called the "peer effects" of DER adoption. For example, the adoption of the same technology in the neighbourhood would likely increase the adoption likelihood of that same (or other) technologies in the neighbouring cells.

In addition, it can be argued that, while incentive designs drive spatial DER adoption patterns, electricity tariffs largely steer temporal DER utilization behaviour. Future studies should, from a system view, compare both instruments (incentive designs vs. tariffs) in their effectiveness to reduce system expansion costs or maximize self-consumption. Furthermore, their comparison to other "non-wires" alternatives such as demand side management represent a very promising and timely research direction.

In fact, this thesis has largely focused on the interaction of EV/PV technologies in residential environments. While this has been in line with extensive literature that suggest potential synergies between them, further technologies may be included in the future. As one example, the current electrification of the heating sector may require the integration of HVAC/heat pumps adoption behaviour into future model extensions. Furthermore, household-level battery employment could be further added to the scope of analysis. While such additional technologies increase the complexity of scenario building (in case, different uptake paths for each technology shall be maintained), their integration in the presented model framework is straightforward.

Finally, four different research pathways are outlined, suggesting potential evolutions of the methodological framework of this thesis under different application focusses.

A self-adaptive, integrated DER adoption forecasting model

One foreseen evolution of the presented modelling framework may further enhance the presented spatiotemporal DER adoption forecasting model. However, model improvement shall be driven by model testing and validation, which require additional observations. Given that the current level of adoption of the technologies included in the frame of this work is still low, calibration was based mainly on aggregated rather than cell level. This drawback is strongly linked to the difficulties of error calculation for highly unbalanced data-sets, which represents a research area by itself.

Crossing smart-meter data, surveys or mobile application data with high-resolution satellite imagery may eventually result in a household-level representation in space. While such development raises questions of personal data confidentiality and protection, resulting data-sets allow for an improved validation of forecasting models.

3D load modelling for accurate network planning

One of the major contributions of this thesis is an improved representation of residential load patterns under DER adoption dynamics. A straightforward extension lies in the inclusion of commercial and industrial utilization patterns. However, the inclusion of these sectors poses, as most greenfield planning exercises, additional challenges to the planning exercise. Information on the magnitude and location of industrial or commercial activities is seldom available and hard to retrieve. In addition, access to some information on energy consumption may be constrained due to the risk of losing a competitive advantage of a specific business process.

However, with the help of typical load profiles and increasingly available open geodata and remote sensing imagery, different activity levels (industry, commercial, residential, agriculture) can be merged to create spatiotemporal, 3D net-load models.

In addition, detail of network analysis could be increased, using, if available, real network information (including lines) or relying on reference network models (as the RNM developed by IIT Comillias). In fact, such models are appealing as they can both create greenfield or bownfield network models under absence of the real network. Such synthetic network models can be complemented with the outputs of the spatiotemporal diffusion model and satellite information on the exact consumer location and class.

Optimized policy design choice with human-machine interfaces

The approach presented in Chapter 5 allows understanding the effects of policy designs. Thus, the provided tools can support a careful weighting of decision makers objectives such as decreased system expansion costs or adoption asymmetries for improved energy policy design.

Future extensions of the methodological backbone may enable to estimate the global trade-off of centralized, state-level welfare gains against distributed household level (residential, commerce, industry) welfare gains, under different policy designs, considering expansion costs, welfare gains, welfare distribution, and total expansion cost

In addition, the creation of an interactive, human-machine interface that allows energy policy makers to adjust their objective weightings after understanding the likely impacts of a specific policy design can be envisaged.

Transferable model to other technologies and infrastructures

Finally, the model framework developed possesses a high potential to be transferred to other contexts. On the one hand, the modelling framework is suitable to be applied to other technology diffusion processes that are embedded in social systems such as heat pumps, electricity storage systems, energy efficiency programs, or HVAC systems.

On the other hand, the models could be applied together with rural electrification planning, to estimate peak load evolution under the adoption of electrified cooking systems, refrigerators or other household appliances. Finally, the use of the developed models might go beyond challenges of the energy sector, extending to any other infrastructure planning exercises that require the estimation of demand through consumers' characterisation.

Annex

ANNEX I: Census data of Continental Portugal

a)

The presented models used the census information provided by the Portuguese National Institute for Statistics (INE). This institution publishes every 10 years detailed census data-sets with extensive population information. For the smallest data resolution considered (subsections), around 280,000 units with over 120 socio-demographic criteria are reported. In addition, the census data-set contains spatial polygons that can be manipulated to retrieve different administrative levels (e.g. region, municipality, neighborhood, section and subsection). The polygons can be linked to the population characteristics. Municipal polygons are shown above (a).

The data for the latest survey from 2011 is accessible online (http://mapas.ine.pt/download/index2011.phtml, last accessed on 4th of September 2019).

In our work, unpopulated cells have been removed. A stable population with unchanged population characteristics has been assumed.

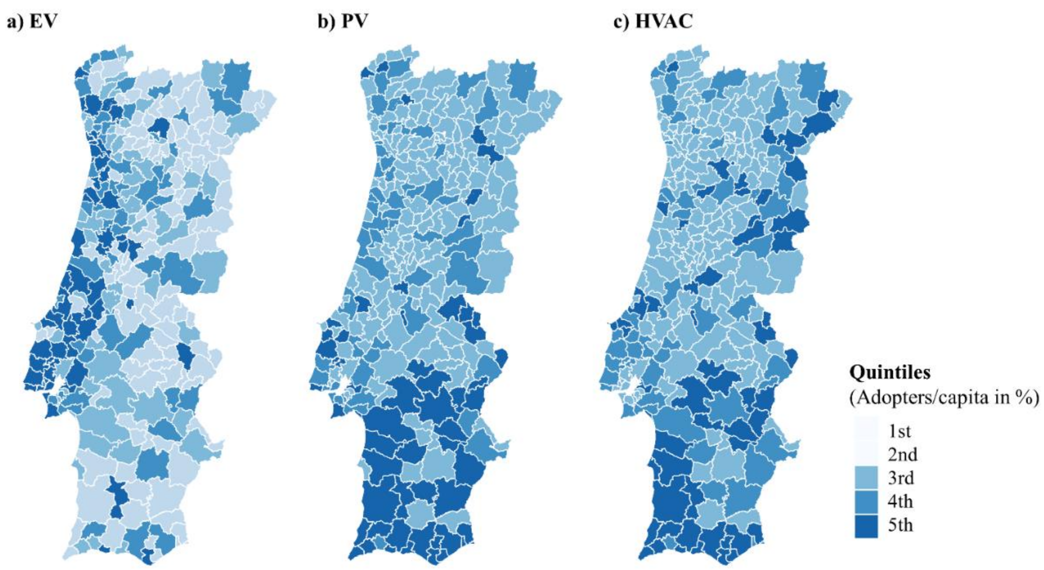
ANNEX II: Georeferenced DER adopter data-set

EV, PV and HVAC adopter data has been retrieved from CeiiA and ADENE, the Portuguese Energy Agency. Provided data included the position of adopters and, for PV and HVAC, further technology characteristics and the sector of its user (industrial, commercial, residential). For PV and HVAC, other than residential users have been removed from the data-set. On the other hand, EV entries have been eliminated if several EV have been registered under the same address. Investigations using Google maps suggested that most of these locations were car selling offices or commercial users.

In general, it was assumed that all individual observations correspond to one household each. Furthermore, it was expected that one household had one appliance registered. Multiple entries for the same address have been removed.

All DER adopters have been georeferenced. In the following, resulting spatial point data-sets were intersected with Continental Portugal to exclude observations for Portuguese islands. The final data-set included 2,632 EV, 474 PV and 2111 HVAC adopter households (as shown in a)).

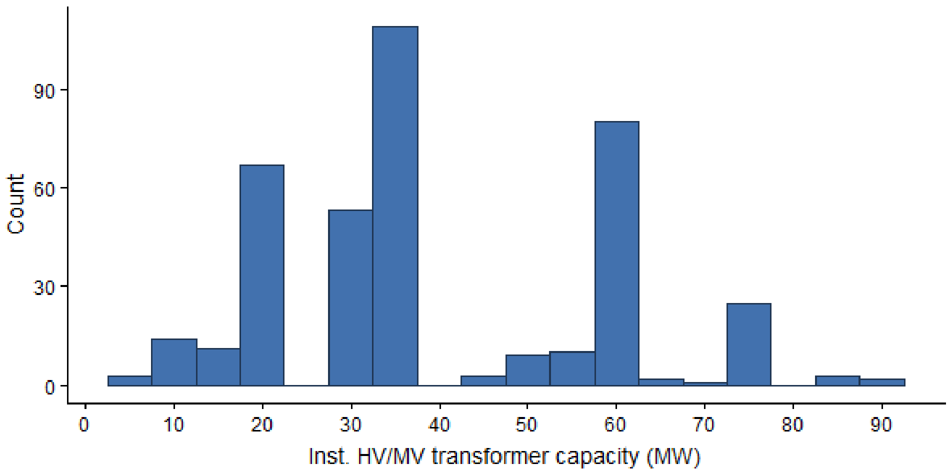
a)



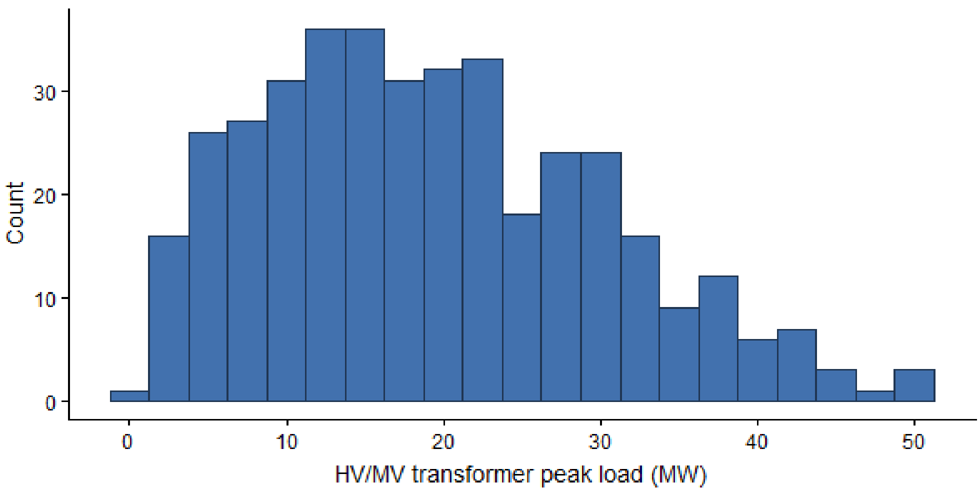
ANNEX III: Spatial distribution and transmission network representation

EDP Distribuição manages a major part of the Portuguese Continental distribution system, supplying more than 6 million clients. Consumers are supplied through 392 HV/MV substations that possess an aggregated, installed capacity of 15.447 MW. One substation that is not actively used has been excluded. The figures below show histograms of the HV/MV transformer capacities (a) used throughout the continental distribution system, peak load occurrences (b) and the maximal loading for each HV/MV transformer (c).

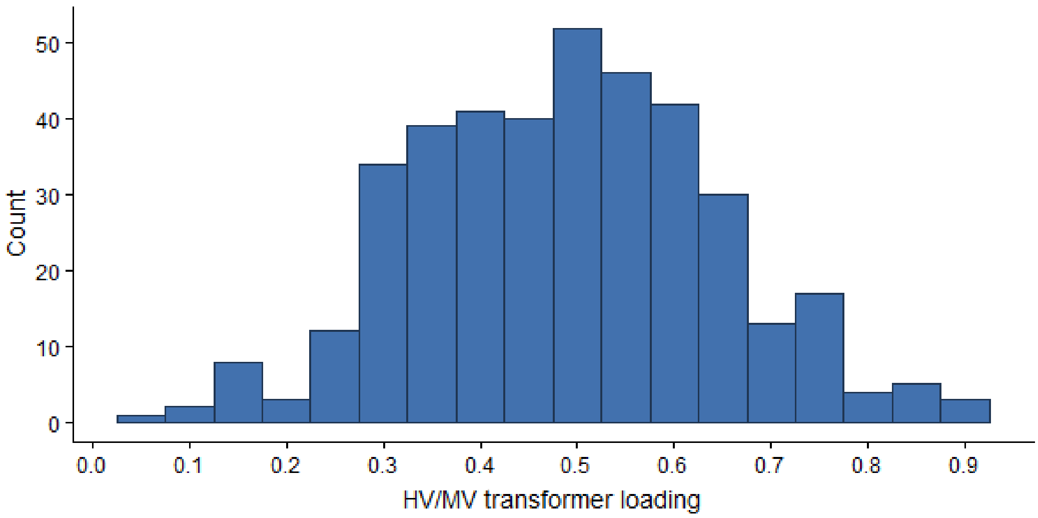
a)



b)



c)



HV/MV substation transformer characteristics (geographical coordinates, peak-load, installed capacity) have been retrieved from the EDP Distribuição webite and latest system planning reports , whereas some values have been corrected in discussion with the distribution company. Natural peak-load growth is estimated at 0.5% annually in line with the estimates provided in the latest investment report of EDP Distribuição (PDIRD-E 2018) .

The average transformer upgrade cost estimation (***TC***) has been retrieved through a comparison of past upgrades that are listed in the same investment report. A conservative expansion cost of 2.5 million € per upgrade is obtained.

Likewise, the companies typical network expansion threshold that is applied to HV/MV transformer expansion has been retrieved. In the Portuguese distribution system, typical expansion threshold for HV/MV transformers is a maximal loading of 0.9. In other words, transformer's replacement or capacity addition is triggered once peak-load (or PNL – peak-net-load) surpasses 90% of the installed HV/MV transformer capacity. Therefore, expansion can be modelled with the following equation:

$$TC = \begin{Bmatrix} 2.5\, million\ Euro \quad for\ PNL \geq\ 0.9 \\ 0.0\ \ million\ Euro\ \ for\ PNL <\ 0.9 \end{Bmatrix} \tag{1}$$

It is important to note that technical problems in LV networks are neglected. LV networks are modelled as a single net-load downstream the HV/MV substation. Substation feeder reconfigurations are not considered as well.

The approximated HV/MV transformer service areas and the respective installed capacities (in MW) are shown below (d). The approximated substation service areas are equivalent to the Voronoi diagram for all continental substations, resulting from purely geometrical calculations. However, in reality, we expect that DSOs possess detailed knowledge on the areas served by each transformer.

d)