UNIVERSITÀ DEGLI STUDI
DI TRENTO

DEPARTMENT OF INFORMATION ENGINEERING AND COMPUTER SCIENCE
ICT International Doctoral School

Development of innovative tools for multi-objective optimization of energy systems

Md Shahriar Mahbub

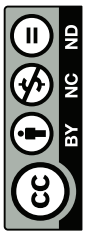
Advisor

Dr. Luigi Crema

Applied Research on Energy Systems

Fondazione Bruno Kessler: FBK

January 27, 2017



This work is distributed under the Creative Commons 4.0 license. You are free share — copy and redistribute the material in any medium or format under the following conditions: i) you must give appropriate credit, provide a link to the license, and indicate if changes were made; ii) you may not use the material for commercial purposes; and iii) if you remix, transform, or build upon the material, you may not distribute the modified material. Please see `https://creativecommons.org/licenses/by-nc-nd/4.0/` for further details.

Examination Committee Members:

Poul Alberg Østergaard
Professor
Department of Development and Planning, Aalborg University, Denmark.

Mauro Brunato
Assistant professor
Information Engineering and Computer Science Department, University of Trento, Italy.

Luca Di Gaspero
Assistant professor
Computer Engineering, University of Udine, Italy.

Abstract

From industrial revolution to the present day, fossil fuels are the main sources for ensuring energy supply. Fossil fuel usages have negative effects on environment that are highlighted by several local or international policy initiatives at support of the big energy transition (including the conference of the parties, the Climate change action plan, the EU energy work programme and so on). The effects urge energy planners to integrate renewable energies into the corresponding energy systems. However, large-scale incorporation of renewable energies into the systems is difficult because of intermittent behaviors, limited availability and economic barriers. It requires intricate balancing among different energy producing resources and the syringes among all the major energy sectors. Although it is possible to evaluate a given energy scenario (complete set of parameters describing a system) by using a simulation model, however, identifying optimal energy scenarios with respect to multiple objectives is a very difficult to accomplished. In addition, no generalized optimization framework is available that can handle all major sectors of an energy system.

In this regards, we propose a complete generalized framework for identifying scenarios with respect to multiple objectives. The framework is developed by coupling a multi-objective evolutionary algorithm and EnergyPLAN. The results show that the tool has the capability to handle multiple energy sectors together, moreover, a number of optimized trade-off scenarios are identified. The framework opens a door for policy makers to optimize corresponding energy systems in terms of multiple objectives and choose the appropriate one for his/her respective region.

Since the proposed framework is computationally costly we make significant improvements by suggesting different techniques. Firstly, a technique is proposed that exploits given domain knowledge by incorporating the knowledge into different phases of the algorithm to improve algorithmic efficiency. Secondly, a convergence detecting criterion is proposed to stop the algorithm at the right moment to prevent wasting computational resources. The results show that a significant improvement can be achieved by utilizing domain knowledge and a stopping criterion. Finally, an algorithmic modification is proposed to explore particular targeted regions of a Pareto-front in a very straight forward way. The exploration of targeted regions is really important for energy domain as policy makers want to find scenarios that fulfill certain goals in terms of different objectives (e.g., 20–30% CO_2 emissions reduction with respect to a particular scenario). It is demonstrated that the modified algorithm is capable of finding solutions in user-defined multiple regions for both benchmark and real-world problems.

The framework is applied for identifying optimized scenarios for two Italian Alpine valleys into two different contexts. In the first circumstance, optimized scenarios are identified by considering recent energy demands of the valley. Whereas

long-term energy planning is performed by identifying future optimized scenarios that consider projected energy demands for another valley.

The comparison of current operating scenario (of the valley) with optimized scenarios shows that much better scenarios can be achieved in terms of different aspects. The results demonstrate that economically attractive, environmentally friendly and less dependent energy scenarios can be obtained by introducing more renewable energy into the electrical sector and modifying thermal sector by introducing heat pumps and biomass boilers. The modification of the transport sector by introducing electric cars is not economically viable under the current market conditions.

The results of the second phase show that the framework can be also applied for long term planning of a system. Optimized scenarios for the valley "Val di Non" are identified for three different time periods as the community wants to reach distinct CO_2 reduction goals within the periods. Moreover, a new technique is proposed for selecting transient scenarios. The results show that the thermal sector would be transformed within first two periods. Afterwards, the changes from tradition cars to electric cars would be taken place.

The proposed framework and the corresponding improvements make it possible to provide a complete tool for policy makers for designing optimized energy scenarios. The tool can be able to handle all major energy sectors and can be applied in short and long-term energy planning.

Keywords: Energy system optimization, energy scenario design, energy transition, multi-objective optimization, multi-objective evolutionary algorithm, incorporating domain knowledge, robust stopping criteria.

List of published papers:

Journal papers:

[1] M. S. Mahbub, M. Cozzini, P. A. Østergaard, and F. Alberti, "Combining multi-objective evolutionary algorithms and descriptive analytical modelling in energy scenario design," *Applied Energy*, vol. 164, pp. 140 – 151, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261915014920

The text is substantially verbatim in chapter 3 except introduction and conclusion; the sections have been modified to ensure a better flow of the text in the thesis. The section related to multi-objective evolutionary algorithms is removed as the details of the algorithms is described in chapter 2 in details. The analyses regarding selecting energy scenarios are included in section 3.5 compared to the published journal version.

[2] M. S. Mahbub, M. Wagner, and L. Crema, "Incorporating domain knowledge into the optimization of energy systems," *Appl. Soft Comput.*, vol. 47, no. C, pp. 483–493, Oct. 2016. [Online]. Available: http://dx.doi.org/10.1016/j.asoc.2016.06.013

The text is verbatim apart from the introduction and the conclusion and mentioned additions in chapter 4 and chapter 6; section 4.2.2.2 and 4.2.3 are extensively modified, paragraph 4.2.2.2 and 4.2.2.2 are added, moreover, new results are included in paragraph 4.4.3 compared to published version. Results regarding integrated approach are included in chapter 6. However, the introduction and the conclusion of the chapter are written independently to the published copy.

[3] M. S. Mahbub, D. Viesi, and L. Crema, "Designing optimized energy scenarios for an italian alpine valley: the case of giudicarie esteriori," *Energy*, vol. 116, Part 1, pp. 236 – 249, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544216313317

The text is verbatim without introduction and conclusion in chapter 7; the introduction and conclusion have been modified to ensure a better flow of the text in the thesis.

Conference papers:

[1] M. S. Mahbub, T. Wagner, and L. Crema, "Improving robustness of stopping multi-objective evolutionary algorithms by simultaneously mon-

itoring objective and decision space," in *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '15. New York, NY, USA: ACM, 2015, pp. 711–718. [Online]. Available: http://doi.acm.org/10.1145/2739480.2754680

The text is substantially verbatim apart from the introduction and the conclusion and mentioned additions in chapter 5. Section 5.3 is modified for better explanation compared to published version.

[2] M. S. Mahbub, M. Wagner, and L. Crema, "Multi-objective optimisation with multiple preferred regions," in proceedings of *Artificial Life and Computational Intelligence: Third Australasian Conference, ACALCI 2017*. Springer, pp. 241–253 [online] http://dx.doi.org/10.1007/978-3-319-51691-2_21

The text presented in chapter 9 is revised versions of the articles. Most of the sections of the chapter are paraphrased and extended for better understanding to the readers.

[3] M. S. Mahbub, "A domain knowledge-based multi-objective evolutionary algorithm for optimizing energy systems," in *MENDEL 2014: 20th International Conference on Soft Computing*, vol. 20. [Online]. Available: mendel-conference.org

The text of the paper is extensively revised and adopted in chapter 4.

Acknowledgements

I greatly thank almighty Allah for giving me the inspiration, patience, time, and the ability to finish the thesis. The thesis is not only the result of my works but also as a consequence of all the collaborations with a number of people. All the discussions, debates and thought experiments with my colleagues encourage me to challenge state of the art and push the boundary of the knowledge a bit further in the domain. I am very grateful to all of them for collaborating with me.

During this time I not only learn about scientific approaches for solving problems but also other perspectives of life. The ride of my PhD journey was not smooth all along, I had a bumpy ride at the beginning; I learned many different aspects of life.

First of all, I would like to thank my supervisor, *Dr. Luigi Crema*, for his continuous help and support throughout the period. He gave me the freedom to spend my time on my ideas. The supports and encouragements given by him at my difficult time were really important for me.

I would like to thank – *Mauro Brunato, Poul Alberg Østergaard and Luca Di Gaspero* for agreeing to be committee members of my PhD examination. They have provided valuable feedback that enriches my thesis.

I take the opportunity to thank all of my lab mates; *Matteo, Mattia, Simone and Diego*, who made me feel that I am just one of them. Special thanks to *Diego* who have helped me collect all the data about local territories, without his help the achievements of the thesis would be largely limited. Moreover, we have spent many hours to discuss about how to model the energy systems for local territories.

I would like to thank all of my co-authors of my research papers –*Markus Wagner, Luigi Crema, Marco Cozzini, Poul Alberg Østergaard, Tobias Wagner and Diego Viesi* for their supports and valuable comments. Without their proper suggestions, it would be very difficult to publish research articles. Special thanks to *Markus Wagner* for his continuous support during this period. He has helped me almost in every aspect of my PhD for last two years.

My PhD would not be possible without the support of my wife, *Nusrat Sharmin*; herself being a PhD student she knows how much frustrating it may feel sometimes. It would not be possible to overcome these frustrating moments without her support and care.

Finally, I want to thank my parents and parents-in-law for their love and for always being there with me. I would like to thank my bother, brother-in-law, all of my relatives and friends (in Trento and back home) for being helpful.

Thanks to all of you.

Shahriar

Contents

1 Introduction **1**
- 1.1 Motivation . . . 2
- 1.2 Research goals . . . 3
- 1.3 Contributions . . . 3
- 1.4 Structure of the thesis . . . 5

2 Multi-objective optimization **8**
- 2.1 Introduction . . . 8
- 2.2 Multi-objective optimization problems . . . 9
- 2.3 Objectives in MOOP . . . 9
- 2.4 Dominance and Pareto-optimality . . . 10
- 2.5 Classical methods . . . 12
  - 2.5.1 Lexicographic method . . . 12
  - 2.5.2 Weighted sum method . . . 12
  - 2.5.3 ϵ-Constraint method . . . 13
  - 2.5.4 Weighted Tchebycheff method . . . 13
  - 2.5.5 Summary of classical methods . . . 13
- 2.6 Multi-objective evolutionary algorithms . . . 14
  - 2.6.1 NSGA-II . . . 15
  - 2.6.2 SPEA2 . . . 17
- 2.7 Constraint handling . . . 19
  - 2.7.1 Ignoring infeasible solutions . . . 19
  - 2.7.2 Penalty function method . . . 19
  - 2.7.3 Constrained tournament selection . . . 20
- 2.8 Conclusion . . . 20

3 Proposed framework **22**
- 3.1 Introduction . . . 22
- 3.2 Methodology . . . 25
  - 3.2.1 EnergyPLAN . . . 25
  - 3.2.2 Computational demand . . . 27
  - 3.2.3 Combination of EnergyPLAN and multi-objective evolutionary algorithms . . . 28
- 3.3 Case study . . . 29
  - 3.3.1 Description of Aalborg scenario . . . 30
  - 3.3.2 Choice of decision variables, objectives and constraints of an energy system to optimize . . . 31
- 3.4 Simulation and results . . . 33

3.4.1 Pareto-front obtained with MOEA . . . 34
3.4.2 Capacity trends . . . 35
3.5 Compromised scenarios . . . 40
3.6 Conclusion . . . 41

4 Incorporating domain knowledge into the framework 43
4.1 Introduction . . . 43
4.2 Incorporating domain knowledge . . . 44
4.2.1 Domain knowledge related to energy systems . . . 44
4.2.2 Smart initialization . . . 45
4.2.3 Smart mutation . . . 51
4.3 Test problem: Aalborg energy system . . . 53
4.4 Comprehensive experiments and results . . . 54
4.4.1 General experimental settings . . . 54
4.4.2 Evaluation metrics . . . 55
4.4.3 Influence of smart initialization . . . 55
4.4.4 Influence of smart mutation . . . 60
4.5 Conclusion . . . 61

5 Development of a robust stopping criterion 63
5.1 Introduction . . . 63
5.2 State of the art . . . 64
5.3 Proposed methodology . . . 65
5.4 Experiments and results . . . 70
5.4.1 Experimental settings . . . 70
5.4.2 Evaluation metrics . . . 71
5.4.3 Discussion . . . 71
5.5 Conclusion . . . 75

6 An integrated approach to energy system optimization 76
6.1 Introduction . . . 76
6.2 Experiments and results . . . 76
6.3 Conclusion . . . 79

7 A practical application of the framework: the case of Giudicarie Esteriori 80
7.1 Introduction . . . 83
7.2 The study area: Giudicarie Esteriori . . . 83
7.2.1 Location and population . . . 83
7.2.2 Electricity production and demand . . . 84
7.2.3 Thermal energy demand . . . 85
7.2.4 Energy demand for transport . . . 87
7.3 Methodology . . . 87
7.3.1 Decision variables and constraints . . . 88
7.4 Objectives . . . 89
7.5 Simulation and results . . . 90
7.5.1 Results of the reference scenario . . . 90
7.5.2 Results of optimized scenarios . . . 90
7.5.3 Best annual cost scenarios . . . 92

7.5.4 Target CO_2 emissions scenarios . . . 95
7.5.5 Target ESD scenarios . . . 97
7.5.6 General discussion . . . 99
7.6 Conclusion . . . 100

8 Long-term energy planning with multi-objective optimization 101
8.1 Introduction . . . 101
8.2 The studied area . . . 102
8.3 Reference scenario . . . 103
8.3.1 Overview of the energy system . . . 103
8.3.2 Electrical demand and production . . . 103
8.3.3 Thermal demand and production . . . 104
8.3.4 Energy demand for transportation . . . 105
8.4 Parameters for identifying future optimized scenarios . . . 105
8.4.1 Future projections . . . 105
8.4.2 Decision variable, objectives and constrains . . . 107
8.5 Identifying future optimized scenarios . . . 109
8.5.1 Simulation and results . . . 109
8.5.2 Results of reference scenarios . . . 109
8.5.3 Result of future optimized scenarios . . . 110
8.6 Emission target scenarios . . . 112
8.6.1 Target scenarios for 2020 . . . 113
8.6.2 Target scenarios for 2030 . . . 114
8.6.3 Target Scenarios for 2050 . . . 114
8.7 Target scenarios for 4DS case . . . 117
8.8 Transition from one scenario to another . . . 117
8.8.1 Methodology . . . 117
8.8.2 Results . . . 118
8.9 Conclusion . . . 120

9 Multi-objective optimization with multiple preferred regions 122
9.1 Introduction . . . 122
9.2 Basic principles . . . 123
9.3 Preferred regions for different MOEAs . . . 123
9.3.1 Ideas adopted in pNSGAII . . . 124
9.3.2 pAGE . . . 127
9.4 Experimental study . . . 129
9.4.1 Evaluation metrics . . . 129
9.4.2 Experimental setup . . . 129
9.4.3 Results and discussion . . . 130
9.5 pNSGAII on energy system optimization problem . . . 132
9.6 Conclusions . . . 133

10 Conclusion 135
10.1 Thesis summary . . . 135
10.2 Future work . . . 136

Appendices 138

A **139**
A.1 Local PV productivity and maximum reasonable PV capacity 139
A.2 Assessment of the local sustainable wood resource 139
A.3 Cost and efficiency related data 140

B **142**
B.1 Parameters regarding the energy system optimization framework for VdN 142
B.2 Results for 4DS case for VdN 142
B.3 Projected investment, operational and maintenance costs, lifetimes and efficiencies 144

List of Figures

2.1 Decision and objective spaces. 9
2.2 Example of Pareto-front and dominance relationship. 10
2.3 An example of Pareto-optimal set and Pareto-front. 11
2.4 An example of ideal point (Z^*) in two-dimensional space. 11
2.5 Illustration of NSGA-II procedure. 16
2.6 Crowding distance calculation for i^{th} individual. 17

3.1 Integration of a generic MOEA and EnergyPLAN. 29
3.2 Primary energy demand for Aalborg Municipality in the 2007 reference situation as well as in the 100% renewable energy scenario. . . . 30
3.3 Pareto-front for deterministic PP capacity simulation. 34
3.4 PV and total wind capacities with respect to CO_2 emissions. 36
3.5 On- and offshore wind capacity trends with respect to CO_2 emissions. 36
3.6 Energy production vs. installed capacity for on- and offshore wind. Production per unit capacity evidently decreases for large installations, due to excess electricity availability. 37
3.7 CHP and HP capacities with respect to CO_2 emissions for the optimization of the energy system of Aalborg. HPs are used in three cases; individual houses, the central DH system and local DH system. The HPs here are the HPs used in the central DH system. 38
3.8 PP capacity with respect to CO_2 emissions for the optimization of the energy system of Aalborg. 38
3.9 Flow chart for deciding PP capacity. 39
3.10 Trend of PP with respect to CO_2 emissions in deterministic PP capacity simulation. 40
3.11 Pareto-front and the 10 best-compromised scenarios (marked with blue color) for Aalborg energy system. 41

4.1 Example of probability distributions. 46
4.2 Probability distributions for polynomial and modified polynomial mutation. 51
4.3 Flowchart for proposed mutation technique. 53
4.4 Pareto-fronts generated by SPEA2 with random initialization (RI) and with smart initialization (SI). 56
4.5 Boxplots for four metrics comparing random initialization (RI) with smart initialization (SI) in NSGA-II and SPEA2. 57
4.6 Development of the indicators during the optimisation runs for SPEA2. 58

4.7 Boxplots for four metrics comparing smart initialization with the default (SI_d) and the alternate approach with different parameters' settings (SI_{A1}, SI_{A2}) for NSGA-II and SPEA2. 59
4.8 Pareto-fronts generated by SPEA2 with polynomial mutation (PM) and with smart mutation (SM). 60
4.9 Boxplots for four metrics comparing PM with SM for NSGA-II and SPEA2. 61

5.1 Example of Hausdorff distance. 66
5.2 Normalized AHD and diversity are represented by black and blue trends. Three vertical lines with three different colors indicate the stopping generations for three different criteria. 68
5.3 Pareto-fronts extracted by NSGA-II for two consecutive generations. . 69
5.4 Boxplots for *stopping generation*, HV_d and eps_d on the different problems for NSGA-II. 73
5.5 Boxplots for *stopping generation*, HV_d and eps_d on the different problems for SPEA2. 74

6.1 An integrated approach to energy system optimization model. 77
6.2 Pareto-fronts generated by generic SPEA2 and SPEA2_Int. 77
6.3 Boxplots for four metrics comparing NSGA-II and SPEA2 with integrated approach. 78

7.1 The Province of Trento and the area of Giudicarie Esteriori supplied by CEIS 84
7.2 Monthly electricity demands and productions from different plants. . 85
7.3 Monthly thermal energy demand for two sub-sectors in Giudicarie Esteriori. 86
7.4 Pareto-front for Giudicarie Esteriori energy system optimization problem. 92
7.5 Capacities of different technologies (except transportation) for the reference scenario (RS) and for 15 best scenarios in terms of annual cost. X-axis represents scenarios and Y-axis represents capacities (in kW). Each vertical grey line represents a particular scenario. 94

8.1 The Province of Trento and the area of Val di Non. 103
8.2 Projected demands for three sub-section including reference scenario (in red). 106
8.3 Reference and projected national electricity grid mix for different years. 108
8.4 Comparison of all Pareto-fronts. 111
8.5 Target ranges and scenarios for different time periods for 2DS case. . 112
8.6 Capacity of different decision variables for 10 target scenarios for 2020:2DS. 113
8.7 Heat production (in percentage) from different technologies for 2020:2DS. 114
8.8 Number of Cars (ICE and EV) for 10 target scenarios for 2030:2DS. . 116
8.9 Heat production shares from different technologies for 2050:2DS. . . . 116
8.10 Transition 2020 $\rightarrow$ 2030 $\rightarrow$ 2050 shown in objective space 121

9.1 Reference points, weights in the objective space, and preferred regions. 124
9.2 Association of solutions with regions. 126

9.3 Pareto-fronts obtained by pAGEoffline and pNSGAII on ZDT1 problem. 131
9.4 Comparison of our pNSGAII with their original variants on the ZDT functions with $m = 2$. The Regions 1–3 are defined in Section 9.4.2. Shown are the means and standard deviations of 100 independent runs. From top to bottom, we show the results of epsilon, IGD and HV. Within each block of four markers, we first show our pNSGAII with ● , then with ○ and □ the original algorithm with two population sizes ($\mu = 30$ and $\mu = 100$, $EF = 12,000$ each), and then with ⬠ the original algorithm with twice the evaluation budget ($\mu = 100$, $EF = 24,000$). 132
9.5 Comparison of our pAGE with their original variants on the ZDT functions with $m = 2$. The Regions 1–3 are defined in Section 9.4.2. Shown are the means and standard deviations of 100 independent runs. From top to bottom, we show the results of epsilon, IGD and HV. Within each block of five markers, we first show our pAGE variants with ○ (pAGEoffline with × and pAGEonline with +), then with ○ and □ the original algorithm with two population sizes ($\mu = 30$ and $\mu = 100$, $EF = 12,000$ each), and then with ⬠ the original algorithm with twice the evaluation budget ($\mu = 100$, $EF = 24,000$). 133
9.6 Comparison of our pMOEAs ($\mu = 30$, $EF = 50,000$) on a subset of the DTLZ functions with $d = 3$. ● denotes pNSGAII and ○ denotes our pAGEoffline (×) and pAGEonline (+) variants, and □ and ⬠ denote the original AGE algorithm with $\mu = 100$, $EF = 50,000$ and $\mu = 150$, $EF = 49,950$ respectively. Typically, our pMOEAs achieve significantly better approximations of the preferred regions. 134
9.7 Result of the energy system optimization. Shown is the final solution set computed by pNSGAII. 134

List of Tables

3.1 Positioning of papers in the thesis with respect to comprehensive energy system modelling and multi-objective optimization. 23
3.2 Electricity demand in Aalborg Municipality in the 2007 reference and in the 100% RE scenario. Any demands for Transport, Hydrogen, and Biogas production are insignificant and included in the category "Other electricity demand" in 2007. In 2007, also industry is included in "other". Thus, the 117 GWh indicated is additional to what is included in the category "Other". 31
3.3 Electricity production in Aalborg Municipality in the 2007 reference and in the 100% RE scenario. 32
3.4 Final heat demand in Aalborg Municipality in the 2007 reference and in the 100% RE scenario. Final DH demands include grid losses. . . . 32
3.5 Lower and upper bounds for the decision variables used in this simulation. 34

4.1 A real-life domain-knowledge representation example. 45
4.2 Combination of all β values when $d = 2$ and $b = 3$. 47
4.3 Lower and upper bounds of different decision variables for Aalborg energy system problem. 54
4.4 General parameter settings for NSGA-II and SPEA2. 55
4.5 Domain knowledge related to different decision variables and objectives for Aalborg energy system problem. 55
4.6 Mean and standard deviation (in subscript) for different quality indicators. 57
4.7 Mann-Whitney U-tests: p-values for different metrics when comparing smart initialization (SI) with the common random initialization (RI). 58
4.8 Parameters settings for examine alternate approach. 59
4.9 Mann-Whitney U-tests: p-values for different metrics when comparing alternate approach with default approach. 60
4.10 Mann-Whitney U-tests: p-values for different metrics when comparing our smart mutation (SM) with the common polynomial mutation (PM). 61

5.1 Parameter settings for NSGA-II and SPEA2. 70
5.2 Standard Maximum allowed number of generations [171], [54]. 71
5.3 Mean and standard deviation for parameters for different problems. . 72

6.1 Mean and standard deviations of four different metrics 78

6.2 Mann-Whitney U-tests: p-values for different metrics when comparing our integrated approaches (Int) with the generic approaches. . . . 79
6.3 Number and percentage of saved function evaluations for different MOEAs. 79

7.1 Number of different electricity producing plants and corresponding installed capacities (kW) of CEIS. 84
7.2 Energy mix for thermal energy in Giudicarie Esteriori. 87
7.3 Characteristics (average annual distance travelled, fuel efficiency and LCV) of petrol and diesel cars applied in this study. 87
7.4 Transport energy demand for the municipalities situated in Giudicarie Esteriori. 88
7.5 The reference scenario of Giudicarie Esteriori in terms of four objectives. 91
7.6 Reference scenario (RS) and best 15 scenarios in terms of Annul cost (AC). 93
7.7 Target scenarios in terms of reduction of CO_2 emissions. 96
7.8 Target scenarios in term of ESD reduction. 98

8.1 Breakdown of the thermal energy demands by sub-sectors and by energy sources. 104
8.2 Projected demands for three sub-sectors for three different years. . . . 106
8.3 Projected Fuel prices. 107
8.4 Projected grid electricity prices. 107
8.5 Reference and projected national electricity grid mix in percentage for different years (2008 and renewable shares are from [69]). 108
8.6 Results of reference scenarios for different time periods. 110
8.7 Comparison of target scenarios with reference scenarios for 2020:2DS. 115
8.8 Comparison of target scenarios with reference scenarios for 2030:2DS. 115
8.9 Comparison of target scenarios with reference scenarios for 2050:2DS. 115
8.10 Energy transition 2020 $\rightarrow$ 2030 for all 2020 scenarios. 119
8.11 Energy transition 2030 $\rightarrow$ 2050 for all 2030 scenarios. 120
8.12 Transition 2020 $\rightarrow$ 2030 $\rightarrow$ 2050 for all the 2020 scenarios. 120

9.1 Configurations in terms of population size μ and evaluation budget EF to test the efficiency of the interval-based preferences. 129
9.2 Parameter settings 130

A.1 Parameters for assessing sustainable wood resource. 140
A.2 Investment cost, lifetime, fixed O&M cost 140
A.3 Variable O&M cost 140
A.4 Generation efficiency 141
A.5 Fuels' prices and additional cost 141

B.1 Parameters for framework used for finding optimized scenarios for VdN 142
B.2 Domain-knowledge related to decision variable and objectives 142
B.3 Comparison of target scenarios with reference scenarios for 2020 and 4DS case 143
B.4 Comparison of target scenarios with reference scenarios for 2030 and 4DS case 143

B.5 Comparison of target scenarios with reference scenarios for 2050 and 4DS case 143
B.6 projected investment cost for different technologies for different targeted years. 144
B.7 Projected operational and maintenance cost for different technologies for targeted years. 144
B.8 Projected lifetime for different technologies for targeted years. 145
B.9 Efficiencies for different technologies for targeted yeas. 145

Chapter 1

Introduction

> *You see that pale, blue dot? That's us. Everything that has ever happened in all of human history, has happened on that pixel. All the triumphs and all the tragedies, all the wars, all the famines, all the major advances... it's our only home. And that is what is at stake, our ability to live on planet Earth, to have a future as a civilization. I believe this is a moral issue, it is your time to seize this issue, it is our time to rise again to secure our future.*
>
> — Al Gore

Energy is one of the key components for the development of modern society. The society can not be evolved without proper supply of energy. Energy supply is however predominantly based on fossil fuels, which have several negative consequences on the environment [162]. This harmful impact encourages the use of renewable energy resources (RES) within the energy system to develop a sustainable energy system. The design of future energy scenarios[1] with a correct balance between fossil fuels and RES is hence a very important topic to energy planners worldwide.

In fact, though RES are desirable for the reasons mentioned above, their exploitation on large scale involves other issues, like fluctuating behavior, limited availability, and economic or financial obstacles. These difficulties can be addressed – by introducing, e.g., proper control strategies, efficient couplings between different resources, and supporting policies – but they increase the complexity of the resulting energy systems, requiring the analysis of many variables. Identifying viable configurations – parametrized for example in terms of type and capacity of energy generation technologies, for given demand conditions – can hence be a hard task for energy planners [38, 39].

In order to solve the problem of integration of RES into a system, two optimization phases can be considered [141], i.e., (i) operational optimization and (ii) capacity/sizing optimization[2]. While the day-to-day operations of resources of a given energy system are optimized in the first phase, the second phase is mainly concerned with the design of future energy scenarios to integrate renewable ener-

[1] In this thesis, we use the terms energy scenario and energy system interchangeably, referring to the complete set of parameters (e.g., generation capacities for given technologies) describing an energy system configuration.

[2] In this thesis, we refer capacity optimization of an energy system as optimization of an energy system.

gies. For the first phase, many optimization models such as energy system simulation models are available (e.g., see the comprehensive review article by Connolly et al. [44]). Although a notable number of literatures are available regarding first phase, very limited research work could be found for later stage. Moreover, optimizing an energy system can be formulated as either single-objective or multi-objective optimization problem. It is however practical to formulate the problem as multi-objective optimization problem since there are more than one potential objectives (probably conflicting) exists when considering the complexity of an energy system.

1.1 Motivation

We have stated earlier that it is absolutely necessary to introduce renewable energy sources to mitigate harmful effects on environment from fossil fuels. The ultimate goal is to reach 100% renewable and sustainable energy system (can be reached by introducing renewable energies or by introducing more cleaner technologies). However, it is not easy to reach the goal overnight, each energy system has to pass through a number of transition phases. The transition from a fossil fuels' based energy system to a renewable energy system has to be smooth - gradual introduction of renewable energy. Therefore, each transition phase requires well designed optimized scenarios that have optimal balances between different types of renewable energy resources and fossil fuels' based generation technologies. Nevertheless, finding the optimal balance is not an easy task. Because of intermittent behaviors of renewable energy and complex interactions among different energy sectors (e.g., interaction between electrical and thermal sector), it is very difficult to find the optimal amount of renewable energy within a system. Therefore, capacities optimization (finding viable configurations for different energy generating technologies) is one of the most important aspects to design an energy system. Moreover, the problem has to solve in a multiple objectives optimization manner as most of the time a real-world energy system problem has more than one objectives to consider.

Designing energy scenario is an important research topic for last decades. A large number of private and public institutes, universities and other organizations of respected regions put forward resources to design energy scenario for the regions. A number of literatures can be found that focus on designing energy systems ranges from city [136, 112], state [85, 86] to country [98, 102, 80]. Therefore, our research is a step forward towards the goal of automatic identification of optimized scenarios. The research aims to propose a generalized user-friendly framework that can be used by researchers/policy makers. The model should have two properties to be a generalized framework: i) flexible enough to be used for any region, ii) can handle most important energy sub-sectors. These two properties will provide the policy makers a wide range of flexibility. Policy and decision makers of any region can use the framework, at the same time, having the possibility to deal with all possible energy sub-sectors.

Moreover, the policy maker has to deal not only with short term planning (which includes the scenario that handles current demands of a region) but also with a long-term planning. The long planning includes identification of optimized scenarios for different future time frames (i.e., for all the transition phases/stages) by considering future energy demands. Moreover, the selection of particular scenarios from those optimized scenarios for each transition phase is another important aspect of long-

term energy planning [3]. The framework could provide a support to identify future scenarios and helps policy makers to make plans for future.

1.2 Research goals

Generally the energy scenarios are designed manually [136, 20, 40, 108, 127]. The scenarios are introducing manual dataset to fulfill certain targets such as reaching to a specific goal of CO_2 emissions reduction or reaching to a certain percentage in terms of introduction of renewable energy within a system. In the case of manual approach, the possibility of improving those objectives may not be explored or the optimal combination of technologies to reach the target may not be found.

It has been found that optimization tools are applied for optimizing particular energy sector (mainly electrical sector) [19, 68]. However, this approach is not enough because every sectors of a future energy system have to be interconnected [104, 107][4]. Synergies among different energy sectors provide a lot of flexibility for an introduction of renewable energies [5]. Nevertheless, optimizing this kind of integrated systems is really difficult by using a typical approach because of the complex interactions and discontinuous nature of the system. To our best knowledge, no framework is available to handle the optimization problem of interconnected energy systems.

The primary research goal of the thesis is to develop a generalized framework. The computational model has to be versatile enough to deal with the complexity of "smart energy system", as the same time, scalable enough to tackle energy systems of different sizes. Moreover, the model need to handle multiple objectives since practical energy system problem generally has more than one objective.

The secondary goal is to improve the efficiency of the framework in terms of finding better solutions in shortest possible time. The preliminary idea is to take advantage of available domain knowledge regarding an energy system and use the knowledge to find optimal scenarios efficiently. Our goal is to provide a novel technique to integrate domain knowledge into the proposed framework. Moreover, we would like to develop new methods that integrate user preferences (i.e., target ranges) in a very simple way within the optimization phase to identify optimal scenarios that fulfill the particular target.

The final goal is to explore the capability of the framework of finding optimized scenarios by applying it on the real-world energy system optimization problems.

1.3 Contributions

In this thesis, we have proposed a framework based on a multi-objective evolutionary algorithm and EnergyPLAN [111] that provides us a novel way to optimize interconnected energy systems. The proposed framework is applied on a test problem and a number of optimized scenarios are identified with respect to two conflicting objectives. The comparison of found optimized scenarios with a manually configured scenario (designed by energy experts) shows that it is possible to find a better

[3] We call transient scenarios.

[4] The system is called "smart energy system" [119, 107].

[5] Considering that when a system produces excess electricity from photovoltaics or wind power, it is possible to turn off the boilers fueled by fossil fuels and turn on the heat pumps.

scenario than manually configured one. Moreover, a set of trade-off scenarios is identified as multiple objectives are considered. It also demonstrates that the framework is very convenient in terms of usability (i.e., scalable and fast). The framework can handle a reasonable number of decision variables (capacities of different technologies) and able to find the solutions in an acceptable amount of time.

To make the framework efficient, we propose an innovative approach to incorporate domain knowledge regarding energy system into *initialization* and *mutation* phases of the algorithm. Simple domain knowledge such as *increasing wind power capacity reduce emissions of a system* is incorporated by exploiting different aspects of probability distributions. In the initialization phase, rather than drawing individuals from a uniform distribution, the individuals are drawn from proposed distributions that utilize domain knowledge. A similar approach is suggested for mutation phase. The results show a significant improvement can be achieved in terms of quality of identified solutions.

Generally, a multi-objective evolutionary algorithm is stopped by user defined parameters (i.e., number of function evaluations). Without prior knowledge about a problem, the parameter is difficult to specify. Therefore, a robust stopping criterion is proposed to stop the algorithm automatically to save valuable computational resources used by function evaluations. The criterion is based on simultaneous monitoring of two spaces; i.e., decision and objective spaces. The criterion is tested on different benchmark problems against state of the art approaches and it is demonstrated that our approach stops an algorithm more reliably than other methods. Finally, the proposed criterion is integrated within our proposed framework and tested against a real-world energy system optimization problem and the results show approximately 22% computational time is saved while achieving a comparable performance with respect to default approach.

The framework is applied to identify scenarios for two local systems in two different situations. In the first case, the model is applied for short term energy planning where optimized scenarios are identified for near future (based on the current demands). In the other case, long-term energy scenario planning is performed using the model where optimized scenarios are identified for different time periods. In both cases, a number of optimized scenarios are identified. From the optimized scenarios, different target scenarios[6] are determined. An approach is proposed for selecting target scenarios from the optimal scenarios that are diverse (i.e., dissimilar to each other in terms of capacities of resources). Diverse target scenarios offer a wider view to the decision makers. Moreover, a novel technique is proposed to select the transient scenarios form the targeted scenarios.

In the first case, the comparison of the current scenario with the identified scenarios shows that economically attractive, environmentally friendly and less dependent scenarios can be achieved. It also reveals that increasing photovoltaics capacity, maximizing local biomass usage though biomass boiler and partial electrification of thermal sector through heat pumps could help the community to have a green and sustainable energy system. Moreover, the results also clarify that transformation of the transport sector is not economically viable under current market conditions. By analyzing the results of the second case ("Val di Non" energy system), it has been found that it is required to transform thermal and transportation sectors to

[6]Target scenarios are the scenarios that fulfill specific goals on different objectives (e.g., 20% emissions reduction with respect to the 2012 system).

reach certain emissions reduction goals. However, transformations of different sectors could take place in different time periods. The results show that it is enough to reshape the thermal sector within the period of 2020–2030, whereas, in the period of 2030–2050, the transformation of transport sector would be spontaneously taken place.

As we have just stated that target scenarios are important to policy maker and typical multi-objective evolutionary algorithms are not able to explore given target regions in a straight forward way, therefore, we propose an algorithmic modifications to identify solutions in particular regions. The basic idea is to generate more individuals within the given targeted regions as the population evolves. To generate more individuals, we suggest modifications of parent selection and ranking procedures of generic algorithms. The modified procedures favor the individuals that are near to or within the preferred regions and eventually that leads to explore and generate individuals around the regions. The modified algorithms perform better than the generic ones on six benchmark problems. In addition, it is demonstrated that the algorithms can explore given regions for energy system optimization problem which may help policy makers to identify scenarios on the particular targeted regions.

1.4 Structure of the thesis

In the following, we will have a brief discussion of the remaining chapters:

Chapter 2 - Multi-objective optimization: The chapter discusses all the basic concepts and definitions related to multi-objective optimization. Firstly, some classical methods for solving multi-objective optimization problems are presented. Afterwards, we discuss the details of the two most widely used evolutionary algorithms of the domain. Finally, the chapter is concluded by presenting some constraints handling techniques.

Chapter 3 - Proposed framework: This chapter presents the details of the proposed framework. The framework is a coupling between EnergyPLAN and non-dominated sorting genetic algorithm (NSGAII). The framework is tested on Aalborg energy system to identify optimal solutions. A comparison is made with a manually configured system. The results show a large number of optimized scenarios can be identified in a single run. Capacity trends of different decision variables (technologies) regarding the optimized scenarios are discussed. Finally, a simple technique to determine compromised scenarios finishes the chapter.

Chapter 4 - Incorporating domain knowledge into the framework: The chapter presents the approaches for incorporating energy system related domain knowledge into different phases of the multi-objective optimization algorithm. We present two techniques: *smart initialization* and *smart mutation* that utilize domain knowledge to find better optimized solutions. The techniques are tested individually on Aalborg energy system problem and results show a significant improvement is achieved by incorporating domain knowledge.

Chapter 5 - Development of a robust stopping criterion: Development of a robust stopping criterion for multi-objective evolutionary algorithms is pre-

sented in the chapter. The proposed criterion is developed based on simultaneous monitoring of two spaces: *decision* and *objective* spaces of a problem. *Average Hausdorff distance* and *diversity* metrics are considered to monitor objective and decision spaces, respectively. The technique is evaluated on six benchmark problems against state of the art techniques. The results show that the proposed method can stop multi-objective evolutionary algorithms more robustly than other approaches.

Chapter 6 - An integrated approach to energy system optimization: The chapter is dedicated to integrate all the techniques developed in other chapters (smart initialization, smart mutation and a robust stopping criterion) into energy system optimization framework. The integrated approach is applied on a test problem. The approach can save a significant amount of computational resource while providing similar results as the default approach.

Chapter 7 - A practical application of the framework: the case of Giudicarie Esteriori: In this chapter, we have demonstrated the capability of the framework by finding out optimized scenarios for an Italian Alpine valley. All the energy sectors of the current energy system of the valley are analyzed (reference scenario). Afterwards, by applying the proposed framework a larger number of optimized scenarios are identified. Identified scenarios are categorized into different groups according to different ranges of target objectives' values. The categorization narrows down the number of scenarios for decision makers that helps them to make decisions. Finally, the chapter is concluded by providing suggestions regarding different technological implementation possibility for the valley.

Chapter 8 - Long term energy planning with multi-objective optimization: The chapter demonstrates how the proposed tool can be applied in the case of long-term energy planning for a region (identifying future optimize scenarios by considering future demands). The chapter starts by introducing a reference scenario for an Italian valley named *"Van di Non"*. Afterwards, optimized energy scenarios are identified for different time periods (i.e., 2008–2020, 2020–2030 and 2030–2050). A new technique is suggested to find out target scenarios (in terms of CO_2 emissions reduction) from the optimized scenarios for different periods. The new technique ensures to provide diverse scenarios to the decision makers. Sometimes target scenarios are not enough for decision makers as they want to select transient scenarios for different periods. Therefore, a novel technique is proposed to determine transient scenarios from the given target scenarios. The technique is applied on the identified target scenarios of the valley and found transient scenarios are presented.

Chapter 9 - Multi-objective optimization with multiple preferred regions: The typical goal in multi-objective optimization is to find a set of good and well-distributed solutions. However, it has recently become popular to focus on specific regions of the objective space, e.g., due to market demands or personal preferences. In this chapter, we contribute to the set of algorithms that consider preferences. In particular, we propose the easy-to-use concept of "preferred regions" that can be used by laypeople, we explain algorithmic modifications of NSGAII and AGE, and we validate their effectiveness on benchmark

problems. The results show an improvement over typical algorithms. Finally, we applied our modified algorithm on energy system optimization problem. The proposed modifications is capable to explore different regions simultaneously.

Chapter 10 - Conclusion: The thesis is concluded by providing a brief summary and potential future research directions.

Chapter 2

Multi-objective optimization

Theory provides the maps that turn an uncoordinated set of experiments or computer simulations into a cumulative exploration.

— David E. Goldberg

2.1 Introduction

Mathematically, optimization problems can be conveniently formulated as minimization or maximization problems. The most familiar case is that of single-objective optimization (SOO), where the objective is represented by a scalar function which has to be "extremized" (i.e., minimized or maximized).

For real-world problems, however, it is quite common to encounter the presence of multiple conflicting objectives which have to be optimized simultaneously. For example, this situation is typical in multi-criteria decision analysis (MCDA). The presence of multiple objectives then requires the introduction of a vector function, whose components are the quantities to be extremized. While for scalar values a natural ordering is available, for vector values it is not obvious how to decide what is "higher" or "lower".

To solve this issue, multi-objective optimization (MOO) typically relies on the concept of "dominance", which allows to identify the set of optimal solutions as the set of non-dominated solutions. A solution is said to dominate another solution if it is strictly better in at least one objective, while at the same time not being worse in all the other objectives. Mathematically, the dominance relation yields a strict partial ordering within the objective space [51].

The chapter organizes as follows. In the following few sections, we will formalize different concepts related to multi-objective optimization. Section 2.5 discuss some of the classical methods for solving a multi-objective optimization problem (MOP). We describe the basic concept of multi-objective evolutionary algorithms in section 2.6. The well-known constraints handling techniques are briefly described in section 2.7.

2.2 Multi-objective optimization problems

A multi-objective problem (MOOP) can be mathematically formulated as follows:

$$\begin{aligned} \text{Minimize/Maximize} \quad & f_m(X), && m = 1, 2, \ldots, M \\ \text{subject to} \quad & g_j(X) \geq 0, && j = 1, 2, \ldots, J \\ & h_k(X) = 0, && k = 1, 2 \ldots, K. \end{aligned} \tag{2.1}$$

Here, X is a solution, a vector with n number of decision variables, $X = (x_1, x_2, \ldots, x_n)^T$. Each decision variable can have a bound called upper and lower bounds, $x_i^l \leq x_i \leq x_i^u$. $f_1(X), f_2(X), \ldots, f_M(X)$ are the M number of objectives that need to be maximized/minimized. $f_m(X)$ formed a M dimensional space called *objective space*, at the same time, X constitutes a n dimensional space called *decision space*. Figure 2.1 shows a mapping of a solution from 3 dimensional *decision space* to a 2 dimensional *objective space*.

Most of the time optimization problems face some restrictions because of physical or environmental limitations. The solution that satisfy the restrictions are called feasible solution, otherwise, it is called in-feasible solution. These restrictions are generally called *constraints*. $g_j(X)$ is called inequality constraints and $h_k(X)$ is called equality constraints.

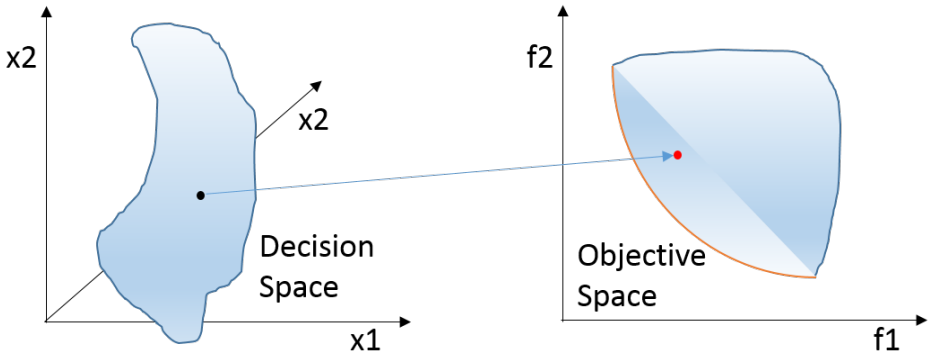


Figure 2.1: Decision and objective spaces.

2.3 Objectives in MOOP

Generally objectives in a MOO problem are conflicting in nature. By considering the nature of conflicting objectives, it is impossible to distinguish among some solutions which one is better than others. These solutions are called optimal solutions. In the case of multi-objective optimization, usually the set of optimal solutions contains more than one solutions. The set projected in objective space is formally called Pareto-optimal front [1]. As all the optimal solutions are equally important, in a ideal case, it is vital to find out as many such solution as possible. Therefore, the algorithms for solving multi-objective optimization problems should have two major goals:

[1] We will provide formal definition of Pareto-optimal front later. The front is also called Pareto-front.

- Find the optimal set as close as possible to true Pareto-front.
- The set should be as diverse as possible.

The first goal is very similar to any single objective optimization algorithm. The found solutions has to be close to true optimal solutions.

However, the second desirable property is explicitly related to the multi-objective optimization. As it is not possible to identify all the optimal solutions, therefore, it is absolutely necessary to have diversity property of the set [2].

2.4 Dominance and Pareto-optimality

In this section, we will formally define some notations related to multi-objective optimization. Without loss of generality, we are considering only minimization problem.

Pareto dominance If a solution S is said to dominate another solution $\bar{S}$ (mathematically, $S \preceq \bar{S}$) if and only if the both the following statements are true:

1. S is no worse than $\bar{S}$ in all objectives, $\forall m \in \{1, \ldots, M\}, f_m(S) \leq f_m(\bar{S})$.
2. S is strictly better than $\bar{S}$ in at least one objective, $\exists m \in \{1, \ldots, M\}, f_m(S) < f_m(\bar{S})$.

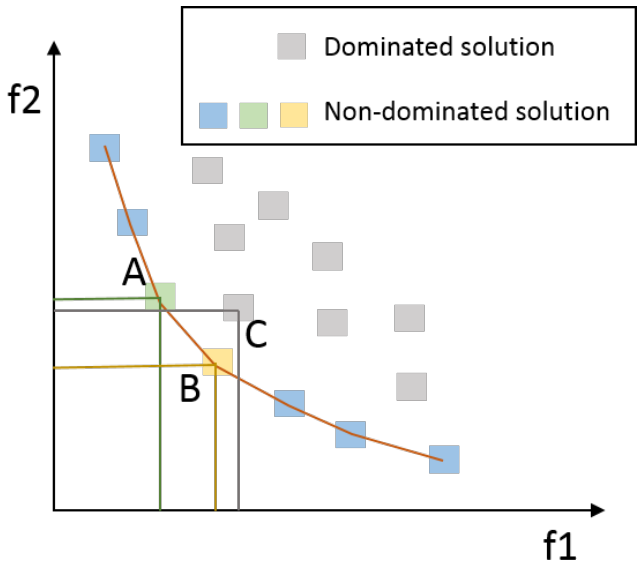


Figure 2.2: Example of Pareto-front and dominance relationship.

Figure 2.2 illustrates an example of Pareto dominance. All the solutions presented in gray color are dominated solutions. The gray solutions are dominated by other colored solutions. Considering solution C and B, solution B has less function values (both for $f1$ and $f2$) than solution C. Therefore, B dominates solution C. However, if we consider solution A and B, A is better than B on function $f1$ and B is better on function $f2$. It is not possible to say who dominates whom; therefore, solution A and B are non-dominated to each other.

[2]Diversity can be measure both in objective and decision spaces. However, usually it refer to the diversity in objective space.

Pareto-optimal set & Pareto-front For a particular problem, the set of non-dominating solutions in decision space is called *Pareto-optimal set*. However, when the Pareto-optimal set is projected in objective space, the set is formally called *Pareto-front*. Figure 2.3 illustrates the Pareto-optimal set and Pareto-front. The set with dark-blue solutions (left side of the figure) represents Pareto-optimal set and the set with red (right side of the figure) solutions represents a Pareto-front of the Pareto-optimal set.

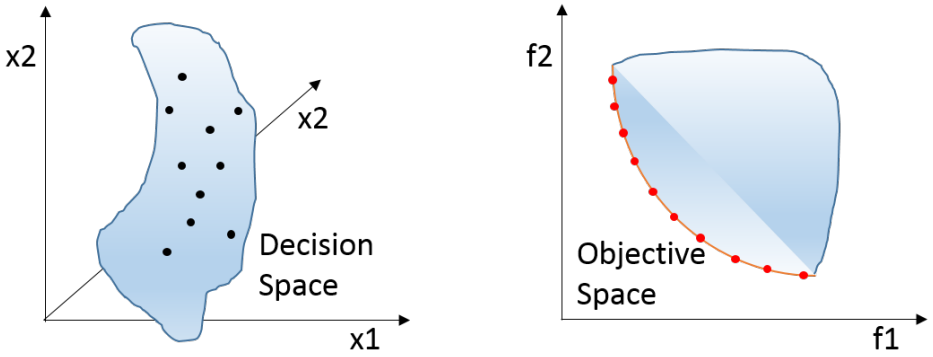


Figure 2.3: An example of Pareto-optimal set and Pareto-front.

Ideal objective vector Ideal vector (also called ideal point or utopia point) is a vector (point) in objective space, constructed by considering all the minimum objective values for M objectives [115].

Mathematically, Ideal vector (Z^*) is a vector such that for each $m = 1, 2, \ldots, M$, $Z^*_m = \text{argmin}_x\{f_i(Q)|Q \in X\}$.

In general, ideal point/vector is unattainable; that means the point corresponds to a non-existent solution. In other words, the solution actually does not exist. Figure 2.4 shows an example of an ideal point. In the two-dimensional space, the ideal point is defined by taking minimum of function $f1$ and $f2$ of the Pareto-front.

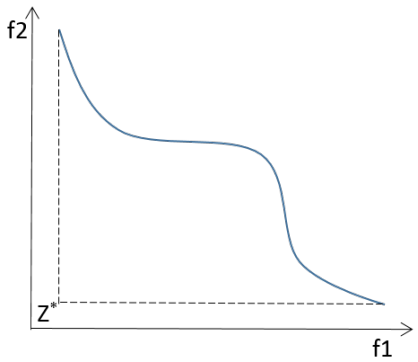


Figure 2.4: An example of ideal point (Z^*) in two-dimensional space.

Compromised solution Compromised solution is a Pareto-optimal solution that has to be as close as possible to the ideal point. Therefore, within all the Pareto-optimal solutions, the closet solution to the ideal point is called compromised solution. The *closeness* can be defined as the minimization of Euclidean distance from

the ideal point to a solution. Mathematically,

$$\sqrt{\sum_{m=1}^{M}\Big(f_m(X) - Z_m^*\Big)^2} \tag{2.2}$$

Please note that if the objectives are in different units, Euclidean distance becomes insufficient to represent closeness. Therefore, it is required to normalized objectives' values so that the values become unitless.

2.5 Classical methods

In this section, we are going to discuss some of the most frequently used conventional methods for solving a multi-objective optimization problem. We call these methods as classical methods.

2.5.1 Lexicographic method

Lexicographic method is based on the minimization of objectives based on importance [114]. Considering that the decision makers provide importance/ranking to the objectives, and, the minimization is performed based on the ranking of objectives one after another. In first iteration 1^{st} ranked objective is minimized, in the subsequent iteration, next ranked objective is minimized by keeping 1^{st} ranked objective at least as minimum as previous iteration. Arithmetically,

$$\begin{aligned} \text{Minimize} \quad & f_m(X), \\ \text{subject to} \quad & f_j(X) \leq f_j(X_j^*), \quad j \in \{1, 2, \ldots, i-1\} \quad \& \quad i > 1 \quad \& \quad i \in \{1, 2, \ldots, M\} \\ & g_j(X) \geq 0, j \in \{1, 2, \ldots, J\} \\ & h_k(X) = 0, k \in \{1, 2 \ldots, K\}. \end{aligned} \tag{2.3}$$

Where i represents the position of ranked objectives, $f_j(X_j^*)$ represents the j^{th} objective value found in j^{th} iteration.

The approach is simple, however, it is very difficult to rank the objectives for a real-world problem. In an ideal world, all the objectives of a problem are equally important.

2.5.2 Weighted sum method

This is probably one of the most frequent used classical method in multi-objective optimization domain, also called linear aggregating functions. This method is very easy to understand and use. The idea is to convert a multi-objective problem into a single objective problem by providing weights to the objectives [51]. Mathematically,

$$\begin{aligned} \text{Minimize} \quad & F(X) = \sum_{m=1}^{M} w_m f_m(x), & \\ \text{subject to} \quad & g_j(X) \geq 0, & j = 1, 2, \ldots, J \\ & h_k(X) = 0, & k = 1, 2 \ldots, K. \end{aligned} \tag{2.4}$$

Where, w_m is the weight associated with each objective and $\sigma_{m=1}^{M} w_m = 1$. In addition, $F(X)$ can be seen as summation of weighted normalized objectives.

The approach is simple, however, it is extremely difficult to assign weights. It obviously depends on decision makers which objectives are given emphasis on. A priori selection of weights can not guarantee to provide a good solution. However, by changing the weights, it is possible to get new a solution every time. In that case, the whole procedure has to perform again to get a new solution.

2.5.3 ϵ-Constraint method

Haimes et al. [81] proposed a new approach in order to solve multi-objective optimization problems. Like the others, the approach also converts a multi-objective problem to a single objective problem with some additional constraints on objectives. Mathematically,

$$\begin{aligned} \text{Minimize} \quad & f_z(X), \\ \text{subject to} \quad & f_m(X) \leq \epsilon_m, \quad m \in \{1, 2, \ldots, M\}/\{m = z\} \\ & g_j(X) \geq 0, \qquad\qquad j \in \{1, 2, \ldots, J\} \\ & h_k(X) = 0, \qquad\qquad k \in \{1, 2 \ldots, K\}. \end{aligned} \tag{2.5}$$

It is clear from the formulation that one objective (z^{th}) is minimized and others are constrained by different constants (ϵ_m). The idea is to optimize one objective where the others are fixed. In addition, by providing different values of ϵ_m, different $f_z(X)$ values can be obtained. Therefore, by providing appropriate ϵ_m values, different points of non-convex objective space can be explored. However, without knowing the shape of the Pareto-front, it is difficult to assign appropriate ϵ_m values.

2.5.4 Weighted Tchebycheff method

Instead of using linear combination of weighted objective (i.e., weighted sum), a different approach is developed by Tchebycheff [114]. Following is the formalization of the approach:

$$\begin{aligned} \text{Minimize} \quad & T(X) = max_{m=1}^{M} w_m |f_m(x) - z_m^*|, \\ \text{subject to} \quad & g_j(X) \geq 0, \qquad j = 1, 2, \ldots, J \\ & h_k(X) = 0, \qquad k = 1, 2 \ldots, K. \end{aligned} \tag{2.6}$$

Where z_m^* is the component of ideal objective vector, Z^*. The advantage of the approach is that it can explore each and every Pareto-optimal solution. The approach has also some disadvantages. Considering that each objective may have different bounds, hence, it is advisable to normalize all the objectives to have value in a uniform scale. However, it is required advanced knowledge about maximum and minimum values for all the objectives.

2.5.5 Summary of classical methods

All the methods described here based on the principle of converting convert a multi-objective optimization problem into a single objective optimization problem. Different methods use different approaches for the conversion. In the following, some drawbacks related to the classical methods can be identified:

- Only one Pareto-optimal solution can be identified in a single run.
- Some classical algorithms suffer for not finding all the Parto-optimal solutions for a non-convex problem.
- All the algorithms require some user defined parameters.

The most obvious disadvantage of classical approaches when comparing multi-objective evolution algorithms (discussed in next section) is that a multi-objective evolutionary algorithm can identify multiple solutions in a single run without using any additional problem specific parameters.

2.6 Multi-objective evolutionary algorithms

The concept of evolutionary algorithm is first proposed by Holland in 1970 [94]. An evolutionary algorithm (EA) is a meta-heuristic optimization algorithm inspired from nature and originally developed for single-objective optimization problems. The algorithm mimics the idea of natural reproduction and selection.

In the reproduction phase, the ideas of genetic 'crossover' and 'mutation' are adopted in the algorithm. Crossover corresponds to an operator that generates new individuals (called offspring) by somehow combining the information (i.e., decision variable or gene) contained within a pair of parents. Unlikely crossover mutation is instead a random change that happens in an individual, mostly at the gene level. Parents are selected from the population based on some *parent selection methods*. Most of the methods prefer fitter individuals as parents. However, different parent selection methods can be found in literature [51].

Crossover and mutation play important roles in EA. Crossover generates a offspring similar to the parent, hence, generated offprings search the solution space close to parents. It is expected that better genes will be preserved in the population (converge towards overall good solutions) by using crossover. On the other hand, mutation introduce genetic diversity which helps the solution to escape from local optima. The application rate of these two operators is controlled by two user-defined parameters called crossover probability and mutation probability.

After reproduction phase, it is required to select finite number of individuals from a combined population (a combination of parent and offspring populations) for the next generation. Therefore, the concept of 'survival of the fittest' (i.e., natural selection) is implemented to ensure that only the best individuals have the chance to reproduce in the next generation. In this way, new generations are produced in a loop until a defined stopping criterion is met [51, 175, 36]. The individuals contained in the last generation are considered as final solutions.

David Schaffer was the first to implement a real multi-objective evolutionary algorithm (MOEA) in 1984, called VEGA (vector evaluated genetic algorithm) [153]. VEGA was most straight forward implementation of single-objective genetic algorithm to find non-nominated solutions. The reproduction and selection processes are performed in an independent loop based on each objective. The proposed algorithm performed better in some cases, however, it failed to achieve finding well spread Pareto-front because of biases towards the solutions with good individual objective values [51].

Algorithm 1 NSGA-II

1: P_t, Q_t ▷ Parent and offspring population at generation#t, respectively
2: N ▷ Population size
3: SG ▷ Stopping generation
4: s ▷ An individual or a solution
5: $P_0 \leftarrow \emptyset, Q_0 \leftarrow \emptyset$
6: Initialize population P with N random individuals ($P_1 \leftarrow \cup_{j=1}^{N} s_j$)
7: $t \leftarrow 0$
8: **while** $t \leq SG$ **do** ▷ Until the algorithm reaches to stopping generation
9: $t \leftarrow t + 1$ ▷ Increment generation counter
10: Generate N offsprings using crossover and mutation, $Q_t \leftarrow \cup_{j=1}^{N} s_j$
11: $R_t \leftarrow P_t \cup Q_t$ ▷ Merge parent and Offspring population into R_t
12: Identify all fronts (F) from R_t using non-dominated sorting procedure ($F = F_1, F2, \ldots$)
13: $P_t \leftarrow \emptyset$ and $i \leftarrow 1$
14: **while** $|P_t + F_i| \leq N$ **do** ▷ Until the parent population is filled with N solutions
15: $P_t \leftarrow P_t \cup F_i$ ▷ Insert i^{th} front into parent population
16: $i \leftarrow i + 1$
17: **end while**
18: Sort F_i according to crowding distance
19: $P_t \leftarrow P_t \cup F_i[1 : (N - |P_t|)]$ ▷ Take first $(N - |P_t|)$ solutions from F_i and insert the solutions into parent population
20: **end while**
21: **return** P_t

Afterwards, no significant step forward had been made until Goldberg suggested a new non-dominating storing procedure in 1989, which kicked-off the MOEA field. He suggested to assign more non-dominating solutions in a population by using domination concept. To deal with diversity, he also suggested a niching strategy. By following his suggestions, at least three independent groups of researchers developed three different MOEAs: *multi-objective genetic algorithm*, *niched Pareto genetic algorithm*, *non-dominated sorting genetic algorithm*. These algorithms are differ from each-other by the way fitness is assigned to the individuals. In the following sections, we will describe two well-known multi-objective evolutionary algorithms.

2.6.1 NSGA-II

Elitist non-dominated sorting genetic algorithm (named: NSGA-II) was proposed by Deb and his student in 2001 [53]. NSGA-II is a improved version of non-elitist non-dominated sorting genetic algorithm [157] developed by Deb and his student in 1995. NSGA-II was critically modified over NSGA mainly in three aspect: i) proposing better non-dominating sorting procedure, ii) introducing the elitist approach (when selecting solutions for next generation), and iii) better methods for diversity preservation. Algorithm 1 provides the details of NSGA-II.

NSGA-II starts with a population P_t that contains N number of solutions (step # 6); using the genetic operators (i.e., crossover and mutation), N number of children

are generated (step # 10). Afterwards, the parent and child population are merged (step # 11). The merged population (R_t) is sorted by applying non-dominated sorting procedure (step # 12). Non-dominated sorting procedure classifies the solutions of the population into different ascending non-dominated levels/fronts [3] based on domination property of the solutions. The parent population for next generation will be filled by the different identified fronts. The population is filled by one front at a time (when an entire front can be accumulated into parent population, all the solutions of the front are copied into the parent population) (step # 15). Most of the time, it is not possible to copy all the solutions from the last front. The situation is illustrated in Figure 2.5. The entire front # 4 of the figure can not copied to the parent population. Therefore, some solutions will be accumulated until the parent population is filled. However, rather than copied random solutions from the front, first few solutions are copied after performing a ranking among the solutions of the front. A technique called *crowding-distance* is used to rank the solutions (step # 18). Please note that other part of the front # 4 and other trialing fronts are rejected to enter into next generation.

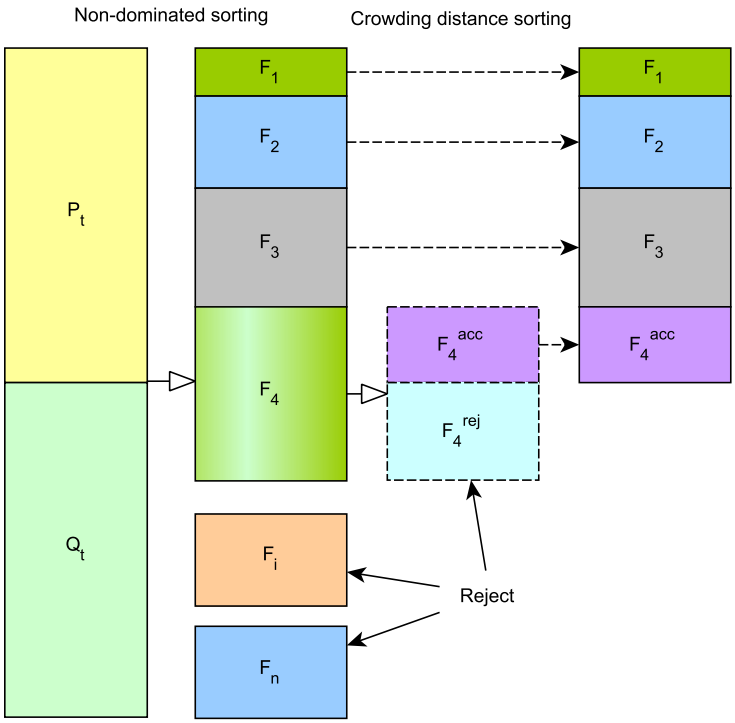


Figure 2.5: Illustration of NSGA-II procedure.

Crowding distance Crowding distance is a metric that is used to identify the density of a solution with respect to the surroundings solutions. The metric may have little effect on the early stage of evolution as there exist many non-dominating solutions. However, in the later stage of the evolution, it is more probable that only a single front exists with more than N number of solutions. Therefore, crowding

[3] The set of best non-dominated solutions are called non-dominated solutions of level# 1/front# 1 and the next best is called front 2 and so on.

distance becomes handy to deal with this kind of situation. In addition, it ensures that well-distributed solutions (over objective space) is passes to the next generation.

The crowding distance of a solution is defined as the distance of two surrounding solutions on either side of the solution along each objective. According to the authors of NSGA-II [53], crowding distance of a solution is calculated by considering each objective independently, and sum the distances between the surrounding solutions. Figure 2.6 illustrates the crowding distance of i^{th} solution, where $(i+1)^{th}$ and $(i-1)^{th}$ solutions are considered as surrounding solutions.

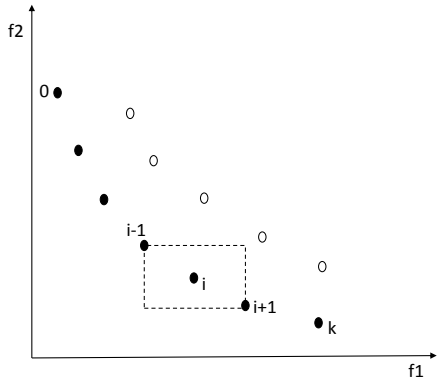


Figure 2.6: Crowding distance calculation for i^{th} individual.

2.6.2 SPEA2

Zitzler and Thiele proposed a new approach in 1998, called Strength Pareto Evolutionary Algorithm (SPEA) [178]. In 2001, Zitzler et al. [180] proposed SPEA2, an improved version of SPEA. SPEA2 has some differences with respect to NSGA-II. Firstly, SPEA2 introduces an external archive to save non-dominated solutions. Secondly, a new technique was suggested for assigning fitness to solutions. Thirdly, a truncation approach was proposed to reduce the number of solutions in external archive. Algorithm 2 shows the steps of SPEA2.

The algorithm starts with N number of randomly initialized solutions stored in P_0 (step # 7). Afterwards, all the solutions contained within P_t and E_t are assigned fitness according to the proposed approach. All the non-dominating solutions extracted from P_t and E_t will be stored in E_{t+1} (step # 10). The number of solutions in the archive can be higher or lower than the fixed size of the archive. Therefore, step # 12 and step # 14 of the algorithm are related to adjusting the number of solutions in E_{t+1}. The entire process of adjusting the number of solutions is called *environmental selection*. Afterwards, binary tournament selection is performed on the archive to select the parents (step # 16). Using the genetic operators, N number of offsprings are generated; finally the offsprings are copied to P_{t+1}. Except the initialization step all other steps are inside a loop and the loop continues until the stopping criteria is met.

Fitness assignment Each of the solution in P_t and E_t is assigned a strength value based on number of solutions it dominates. Subsequently, the solution is assigned a raw fitness which is calculated by summing all the strength values of

Algorithm 2 SPEA2

P_t ▷ Parent population at generation#t
E_t ▷ External archive at generation#t
N, $\bar{N}$ ▷ Population size, external archive Size
SG ▷ Stopping generation
s ▷ An individual or a solution
$P_0 \leftarrow \emptyset$, $E_0 \leftarrow \emptyset$
Initialize population P with N random individuals ($P_1 \leftarrow \cup_{j=1}^{N} s_j$)
while $t \leq SG$ **do** ▷ Until the algorithm reaches to stopping generation
    Calculate fitness for all the individuals of P_t and E_t
    Find all the non-dominated solutions from P_t and E_t and copied to E_{t+1}
    if $|E_{t+1}| > \bar{N}$ **then**
        Reduce the size of E_{t+1} using truncating technique
    else
        Fill E_{t+1} from the dominated solutions from P_t and E_t
    end if
    Apply binary tournament selection on E_{t+1} to fill the mating pool
    Using crossover and mutation, N number of offsprings are generated
    Copy the offsprings to P_{t+1}
    $t \leftarrow t + 1$
end while
return E_t

dominators (in P_t and E_t) of the solution[4]. In the case of presence of many non-dominating individuals in population, it becomes difficult to distinguish among the individuals. Therefore, an additional value based on destiny estimation is added to the raw fitness value. In addition, the destiny estimation is calculated by adopting k^{th} nearest neighbor method. The details can be found in [180].

Environmental selection When all the non-dominating solutions are copied in external archive, any of the three situations can happen: i) the number of non-dominating solutions are exactly equal to the size of the archive, ii) the number of solutions are less than the size of the achieve, or ii) the number of solutions are grater than the size. In the case of first condition, no additional operation is required; however, this is very unlikely condition to happen. To handle the second case, best ($\bar{N} - |E_{t+1}|$) dominated solutions from P_t and E_t are copied to archive. The solutions are selected based on fitness values (by sorting the solutions according to the fitness values). For the third case, a loop continue eliminating a solution at a time until the number of solutions in E_{t+1} becomes $\bar{N}$. The elimination is performed based on the distance of a solution and its k^{th} nearest neighbor. The solution with smallest distance with its neighbor will get eliminated in each phase.

[4]Raw fitness value needs to minimize; therefore, an individual with 0 raw fitness value refers non-dominating solution.

2.7 Constraint handling

Very few real-world problems are unconstrained optimization problems. Most of the times, they come up with a number of constraints. Constraints divide the entire search region into two different regions – feasible and infeasible regions. Feasible region contains all the solutions those are acceptable under the given constraints. And the goal is to find the Pareto-front within the feasible region.

There are two types of constraints – equality and inequlity constraints. In Equation 2.1, J and K numbers of inequality and equality constraints are formulated. Please note that within the formulation, no *"less-than-equal-to"* ineqality constraints are considered. Howevere, it is easily possible to convert *"less-than-equal-to"* to *"grater-than-equal-to"* by multiplying left side of the constraints by -1.

In the few following sub-sections, we will discuss some of the constraints handling methods used in MOEA.

2.7.1 Ignoring infeasible solutions

The most simplest way to handle constraints is to avoid the infeasible solutions [51]. The approach is easy to implement, however, it is difficult to find feasible solutions for real-world problems at the beginning of the evolution. Therefore, a guiding approach is required for generating feasible solutions. The approach simply compares the infeasible solutions with each other to understand which one may produce a feasible solution. Most of the time, the *overall constraint violation* [51] metric is used for the comparison stated above.

2.7.2 Penalty function method

Another popular method used for constrain handling is penalty function. The general idea is to convert a constrained problem to an unconstrained problem. The conversion is done by adding (or subtracting) a value (depending on how much the constraints is violated) to (or from) the objective function. The general formulation of penalty function method is as follows:

$$F_m(X) = f_m(X) + P_m\Phi(X) \tag{2.7}$$

Where F_m is the new objective function by considering constraints violations. $\Phi(X)$ is the term that calculates overall constraint violation by the X individual. P_m is called *penalty factor*. As the original objectives are in different scale in real-world problems, the *penalty factor* should be adjusted differently for different objectives. Generally, the parameter is very difficult to set, however, most of the time a static value is chosen for the parameter [51]. It is clear from the equation that if $\Phi(X)$ is zero, the F_m term becomes same as f_m. However, in the case of constraint violation, penalty (according to overall constraint violation) are added to the original objective function.

To calculate the overall constraint violation, fist of all, constraint violation for each constraint ($j = 1, 2, \ldots, J$) is calculated.

$$\phi_j(X) = \begin{cases} |g_j(X)|, & \text{if } g_j(X) < 0 \\ 0, & \text{otherwise.} \end{cases} \tag{2.8}$$

And the overall constrain violation is the summation of all the constraint violations (calculated in Equation 2.8).

$$\Phi(X) = \sum_{j=1}^{J} \phi(X). \tag{2.9}$$

The approach is easy to understand and implement. However, the difficult part is to setup appropriate values for penalty factors. Study shows that if the factor is chosen correctly, MOEAs perform well, otherwise an infeasible solutions or poorly distributed set can be achieved.

2.7.3 Constrained tournament selection

Deb [50] proposed the method which is based on binary tournament selection [51]. The basic idea of binary tournament selection is that a better solution is chosen within randomly picked two solutions. However, in the proposed approach feasible and infeasiable solutions are considered to choose a better solution. Constrained tournament selection simply prefers a feasible solution over an infeasible solution. Therefore, when comparing two solutions, the following rules are strictly followed:

1. Feasible solution is chosen over infeasible solution.
2. Within two feasible solutions, the one is chosen which has better objective value.
3. Within two infeasibale solutions, the one is chosen which violets less overall constraints.

To handle step # 2, it is difficult to define a single objective value as there are more than one objectives exist in the case of multi-objective optimization. To deal with the problem, the concept of domination is employed. When the solutions are in different non-dominating fronts, the solution will be chosen from better non-dominating front. However, when two solutions belongs to the same front, crowding distance is used to break the tie. Less crowded solution is preferred over more crowded solution.

The approach does not require any extra computational demand. The approach is generalized and can be used for any MOEA algorithms those are based on the domination concept.

2.8 Conclusion

A multi-objective optimization problem is a kind of optimization problem where more than one objectives exist that need to be optimized (minimized or maximized). The conflicting nature of the objectives forced to have a set of non-dominating individuals as solutions to a multi-objective optimization problem. Some classical methods require to convert a multi-objective problem into a single objective problem and solve the problem using some single objective optimization algorithms. The classical methods face different problems when solving multi-objective problems; the most difficult problem is that the algorithms need to run multiple times to generate

multiple solutions by changing the parameters. On the other hand, MOEAs are inspired from nature can deal with MOO problems because of inherited nature of the algorithms. In addition, MOEAs can deal with high-dimensional, discontinuous and multi-model MOO problems.

Real-world optimization problems are mostly bounded with constrains. There are many ways to handle the constraints. One of the most well-known method is *penalty function method*. A penalty is added (for minimization case) to the function value for violating constraints. However, the user defined parameters are difficult to set and the parameters are varied depending on the nature of the problem. In this regards, constrained tournament selection was proposed which is parameter less and well suited for evolutionary algorithms.

In the following chapter, we formulate the energy system optimization problem as a constrained multi-objective optimization problem. Afterwards, a framework is proposed to deal with the problem. The framework is tested against a real-world energy system problem and satisfactory results are achieved.

Chapter 3

Proposed framework

The Earth is the only world known so far to harbor life. There is nowhere else, at least in the near future, to which our species could migrate... Like it or not, for the moment the Earth is where we make our stand.

— Carl Sagan

3.1 Introduction

Nowadays, renewable energy is a cornerstone of an energy system; the necessity of introducing renewable energy is required to mitigate harmful effects of our environment. Although an introduction of large scale renewable energy is preferable - it poses some issues mainly because of intermittent behaviors and economical barriers. By developing proper control strategies and by interconnecting different energy resources the issues of renewable energy can be addressed; as a down side, systems' complexity is hugely increased. Hence identifying viable system configurations (i.e., type and capacity of energy generation technologies for a given demand) is getting extremely difficult.

To face this challenge, different approaches are possible. From a quantitative viewpoint, in order to provide a reliable background for the design of future energy systems, two ingredients appear to be crucial: the simulation model used to analyze the behavior of the considered configurations and the optimization method used to identify the most convenient parameters.

While several solutions are available in the literature in terms of these two ingredients, their coupling in the context of energy scenario design is still far from being fully satisfactory. In practice, either advanced optimization algorithms are applied to sectorial models, or more comprehensive models are optimized with simplified

M. S. Mahbub, M. Cozzini, P. A. Østergaard, and F. Alberti, "Combining multi-objective evolutionary algorithms and descriptive analytical modelling in energy scenario design," *Applied Energy*, vol. 164, pp. 140 – 151, 2016. DOI: `http://dx.doi.org/10.1016/j.apenergy.2015.11.042`

The text is substantially verbatim in chapter 3 except introduction and conclusion; the sections have been modified to ensure a better flow of the text in the thesis. The section related to multi-objective evolutionary algorithms is removed as the basics of the algorithms is described in chapter 2 in details. The analyses regarding selecting energy scenarios is included in section 3.5 compared to the published journal version.

methods. Several examples of the first case are reviewed in [19], where distributed energy resources (DER) are considered, and in [68] where hybrid renewable energy Systems (HRES) models are considered: while detailed models and optimization algorithms are used, these works typically analyze small systems or limited energy sectors (typically focusing on electricity only). Concerning the second case, much less literature is available. Some notable examples are found in [45, 57, 140]: here the used models allow to include electric energy, thermal energy, and transportation, but only single-objective optimization is considered. Koroneos et al. [95] performed a case study on the Greek island Lesvos, investigating the penetration of renewables by applying multi-objective optimization (in terms of costs and CO_2 emissions). However, the energy system (including electric and thermal energy, but no transportation) is represented with a simplified and specifically developed model, not immediately generalizable to other cases, and no details about the optimization method are provided. In Table 3.1, a simple classification of the screened papers is reported.

Table 3.1: Positioning of papers in the thesis with respect to comprehensive energy system modelling and multi-objective optimization.

Papers	Simultaneously including electrical, thermal and transportation sectors	Multi-objective
Alarcon-Rodriguez et al. [19] (review including about 80 papers)	No	Yes
Fadaee and Radzi [68] (review including about 50 papers)	No	Yes
Koroneos et al. [95]	No	Yes
Pina et al. [140]	Yes	No
Dong et al. [57]	Yes	No
Cormio et al. [45]	Yes	No

We therefore propose a step forward in this direction by coupling advanced optimization techniques (multi-objective evolutionary algorithms) to a fairly detailed and comprehensive energy system simulation model (EnergyPLAN). Our choices concerning the model and optimization algorithms are motivated as follows.

A wide literature about energy simulation models exists. They can be classified in different ways, depending on their nature (descriptive, analytical, etc.) or on technical aspects. An extended review is contained in [42], which differentiates models mainly in terms of time step, time extent, and modelled energy sectors. From these points of view, two requisites seem to be needed for a complete analysis. First, the intermittency typical of renewable sources requires a fine time step in order to properly evaluate the issues related to this aspect (possible need for energy storages, effects on grid stability, transmission line capacity, etc.). At least an hourly simulation model appears to be necessary to this purpose. Second, issues related to intermittency and supply-demand matching have shown the importance of considering peak shaving strategies exploiting all the possible synergies between different energy sub-systems, for example between electric energy and thermal energy (through, e.g., heat pumps and thermal storages), and between electric energy and transportation (through, e.g., electric vehicles). Hence, a comprehensive model is needed, including

all the three energy sectors mentioned above. Within the large number of available models – HOMER [13, 100], RETScreen [78, 70], H2RES [12, 108], LEAP [14, 150], and TIMES [18, 169], to cite a few – it is typically difficult to find tools satisfying both of these requisites. Either they are not fully comprehensive (e.g., focused on the electric system only) or difficult to extend to large scale. Our choice fallen hence on EnergyPLAN, which satisfies both requisites, is a freely available model, and is already used in several papers [101, 128], as further described in Section 3.2.1. A model such as EnergyPLAN can simulate an energy system yielding its yearly performance (e.g., in terms of aggregate energy consumptions, costs, and emissions) after proper inputs have been provided (e.g., power capacities for different energy production, conversion, and consumption units). On the other side, as capacities are an input of the tool, their choice is left to the user, so that their optimization against specific objectives is typically performed manually.

Concerning optimization of system configurations, again several methods are available in the literature. In this case, the requisites are determined by the following aspects. First, the high number of decision variables which can enter the optimization process gives rise to a very large search space, where advanced optimization techniques are required in order to yield a feasible computational demand. Second, optimization of energy systems needs to deal with multiple criteria, often in mutual contrast. For instance, the ability of an electricity system to balance demand and supply may be in opposition to its efficiency, as higher flexibility typically requires higher consumptions. Consequently, the optimization problem of a large energy system is in general a multi-objective optimization problem with features reminiscent of complexity (e.g., the strong interaction among its many components). Combining these needs, we decided to resort to meta-heuristic optimization algorithms in a multi-objective framework [51] to tackle this task. This goes well beyond the optimization tools embedded in some energy models (e.g., HOMER), which are single-objective and tailored for small systems, where a brute-force search on a discretized design space is possible.

Meta-heuristic algorithms are indeed especially suitable for large and complex search spaces. Among these algorithms, we choose the class of evolutionary algorithms (EAs). EAs (with single or multi-objective optimization) have been applied for solving different energy related problems such as: photovoltaic related problems [77]; wind farm layout (turbine selection and positioning) problems [121, 152, 149, 163]; design and optimization of hybrid stand-alone energy systems [92, 24]; HVAC (heating, ventilation, and air conditioning) systems optimization [96]; and many others. In addition, several of the papers reviewed in [39] and [68] adopt these techniques.

Summarizing, future energy scenario design is a challenging MOO problem and in this work we present a generalized framework to analyze it, coupling for the first time a simulation model able to describe all the relevant energy sectors (electric, thermal and transport) to an advanced multi-objective optimization algorithm. When designing an energy scenario, it is indeed always desirable that the proposed scenario is in a near optimal state, which is not always achievable with analytical optimization [135]. The proposed meta-heuristic approach takes the form of a multi-objective evolutionary algorithm coupled with the EnergyPLAN simulation model. Moreover, EnergyPLAN is not a linear model, containing as it does conditional clauses for variable assignment and conditional procedures, which is an additional

motivation to combine with MOEA. The goal is to be able to identify a set of future energy scenarios taking simultaneously into account multiple criteria - e.g., decrease of CO_2 emissions and decrease of annual costs. The proposed framework assures a significant speed-up of the identification process compared to the manual approach used in the past for scenario design with EnergyPLAN – provided the latter is feasible at all.

As a case study, we apply our framework to the energy system of Aalborg, Denmark. The performed simulations allow to identify the Pareto-optimal solutions with respect to the two contrasting objectives of costs and emissions. The simulation results are compared with a manually-identified future scenario, taken from literature [20]. Though our main purpose is to explain and demonstrate our computational framework, we also offer some qualitative understanding of the observed trends for the decision variables, providing support to the effectiveness of the optimizer.

The remainder of the Chapter is organized as follows. Section 3.2 discusses the methodology and the details of the framework. In Section 3.3, we present the case study being investigated, as well as the choice of decision variables. Section 3.4 presents the simulation results, with two slightly different approaches for the determination of one of the decision variables. Section 3.5 presents a technique to identify compromised scenarios from the Optimized scenarios. Conclusions are given in Section 3.6.

3.2 Methodology

In this section, we will present a framework that coupled a MOEA and EnergyPLAN. In this regard, we will first introduce EnerlyPLAN model. Afterwards, computational requirement for the framework and the framework itself will be presented.

3.2.1 EnergyPLAN

EnergyPLAN [101] is a deterministic, descriptive, analytically programmed computer model for hour-by-hour simulations of a regional or national energy system. It is developed as a freely available tool at Aalborg University, though its use extends beyond the academic environment.

EnergyPLAN is a deterministic model [108] in the sense that a given output will always result in the same output and thus without any effects of randomness or probability profiles. Wurbs [174] describes descriptive models as models that "demonstrate what will happen if a specified plan is adopted" – as opposed to prescriptive plans that "determine the plan that should be adopted to best satisfy the decision criteria". Hence, EnergyPLAN is a simulation model that simulates the behavior of a system with a given configuration, as opposed to e.g. investment optimization models which seek the optimal investments within a certain area of possibility. Our framework, however, introduces the possibility of using EnergyPLAN as a prescriptive model. Finally, EnergyPLAN is analytically programmed [40]; i.e. the programmer has defined a priori which steps the model must take under various circumstances – as opposed to e.g. linear programming where a solver finds the optimal solution based on a set of interrelations and constraints and an objective function. EnergyPLAN encompasses the entire energy system in terms of energy

resources, energy conversion technologies and demands of electricity, heat and fuel for all demands sectors [101]. It has been designed with particular focus on energy systems with interdependencies between sectors, exploitation of synergies, use of district heating (DH) and with high proportions of non-dispatchable renewable energy production.

Most demands are given as annual aggregates combined with annual distribution profiles. Likewise, fluctuating renewable energy sources (RES) and non-dispatchable energy sources in general are given as annual aggregates or installed capacities combined with annual distribution profiles (i.e., hourly distribution profile for a year). All energy conversion units are given as aggregates – e.g. one representation of all heat pumps (HP) in dwellings, one representation of all land-based wind turbines and one representation of all large-scale waste incineration plants. This aggregation combined with analytical programming makes the model very fast with calculation times measured in seconds [101].

Dispatchable plants include: (i) condensing mode power plants and combined heat and power (CHP) plants for power production; (ii) boilers, heat pumps, and CHP units for heat production; and (iii) electrolyzers and more for the production of synthetic fuels. Combined with storages for electricity, heat and fuels, as well as with flexible demands, these are the flexibilities in the energy system [101] .

EnergyPLAN has two primary regulation strategies; technical and economic. With the economic regulation strategy, the model seeks to dispatch plants and use import and export of electricity to realize the lowest economic costs. In technical regulation, the user chooses between a set of regulation strategies detailing how the CHP plants are operated; the main regulation strategies are 1) according to heat demands, 2) according to pre-defined high electricity price intervals and 3) in order to seek the most energy efficient balance between heat supply and heat demand and between electricity supply and demand.

The model uses an endogen list of priorities, giving preference to units that are considered fuel efficient. Hence, productions with a use-it-or-lose-it character come in first, while dispatchable non-CHP productions (i.e., condensing mode power production and boiler-based heat production) come in last.

The EnergyPLAN does not endogenously consider uncertainty in input parameters such as weather-given production profiles. Previous work has however demonstrated that energy scenarios showing relatively good performance against weather-given production profiles from one year, also shows relatively good performance against weather-given production profiles from other years. Thus, for planning and scenario design purposes, uncertainty in weather-given production profiles is not critical – whereas it is clearly important for actual operation or dispatch simulation [105].

Based on the system configuration and the choice of regulation strategy, the model provides aggregate annual outputs for fuel demands, productions and demands of all types of units, CO_2 emissions and costs. In addition, the model provides hour-by-hour results for all production and demand technologies as well as storages.

For annual energy systems simulation tools, a 1-hour time step appears to be the norm. In a survey of computer tools for energy systems models, nine out of thirteen time step simulation tools employ hourly analyses [42]. Norms do not necessarily reflect optimality, but the choice of a time step no longer than one hour is

a compromise between data availability, processing time and the facility to capture variations in demands and productions. In a comparison between dynamic simulations and EnergyPLAN simulations, Pillai et al. [139] found that "The EnergyPLAN tool needs to be taken conservatively if used for energy system analyzes of islands and 'islandable' systems" due to intra-hour dynamics. Essentially, EnergyPLAN simulates steady-state hourly situations, however in simulations, intra-hour requirements may be satisfied by stipulating minimum absolute productions and minimum relative productions shares from units which in a given scenario is deemed grid stabilising. In addition to this, in future high-RES energy systems it must be expected that RES-based production units will function as virtual power plants, thus with the ability to supply short-circuit power without tripping, and with the ability to stabilise frequency and voltage through control of reactive and active power (see e.g. [35] for developments in the field of wind farms or [62] for the role of smart grids to enable renewable energy integration).

In a review, Prasad et al. [142] identify the EnergyPLAN model as being appropriate for "long-term planning in small developing island countries" and in another review article, Connolly et al. [42] list the model as being one of the few models capable of modelling 100% renewable energy systems encompassing electricity, heat and transport sectors. These are in fact two main applications of the model. Additionally, EnergyPLAN is furthermore versatile as it may be run in command line mode, enabling scenario file generation, simulation and result harvesting from outside the actual EnergyPLAN simulation environment.

The model has been used on various national systems including Denmark [98, 102], Norway [80], Hungary [151], Hong Kong [109], Croatia [33], Mexico [127] [51], Ireland [40, 43], Italy [72] and Jordan [137, 125]. It has also been applied on local areas in Denmark [136], China [86, 85] and Croatia [108], as well as in more specific simulations of the impact of particular elements of the energy system such as flexible demand [97], transport systems [106, 145], heat pumps [84], heat savings [158], biomass availability restrictions [118, 99] and district heating [41, 160]. The model has also been applied to analyze different pricing systems for wind power, as well as for generating input for more detailed electro-technical system analyzes [139, 160, 130, 21, 131].

3.2.2 Computational demand

As explained above, the goal of our framework is to identify the best configurations for the energy system. In the quest for these optimal energy scenarios, the search space to be explored can be huge, as it depends on a large number of decision variables and on all their possible values. For continuous variables, considering a combinatorial approach would not make sense at all, but it is worth emphasizing that the number of possible combinations would easily explode even assuming a reasonable discretization.

A simple way to discretize decision variables is to restrict them to multiples of the accuracy required for the solution, which in practical applications is always finite. For explanatory purposes, we then consider that the optimal value of the i^{th} resource capacity x_i has to be determined with an accuracy Δx_i and that the feasible range is $(0, x_{i,max})$. Then, we restrict the admissible values to the set $\{0, \Delta x_i, 2\Delta x_i, \ldots, m_i \Delta x_i\}$, where $m_i = round(x_{i,max}/(\Delta x_i))$. For the sake

of simplicity, we also assume $m_i = m$ for each resource. Hence, for n resources, there are in principle m^n possible scenarios in total. Even for a quite rough accuracy of 10% with respect to each maximum range value (i.e., $m = 10$), in order to fully explore $n = 6$ resources as done in this paper (onshore wind, offshore wind, PV, conventional power plants, CHP, HP), 10^6 configurations should be compared. EnergyPLAN, considered a very fast tool in the domain of energy systems simulation, takes approximately 3 seconds to evaluate a scenario on a typical single core processor, so that the investigation of such a number of combinations would take hundreds of hours of single-core computational time. With more demanding accuracy requirements or a larger number of decision variables, the combinatorial problem would quickly become intractable even with parallelization on large cluster computing architectures.

While limiting the number of decision variables and exploiting other scenario reduction techniques can always be useful, the MOEA ability of handling large search spaces is highly valuable in this context.

Considering 6 decision variables, 100 individuals and 100 generations to compute, a MOEA is able to generate a well-structured Pareto-optimal front in a few hours on a common workstation. Increasing the number of decision variables requires computing more generations. It is also advisable to run the MOEA multiple times and to combine the fronts found by each run to generate a more appropriate Pareto-optimal front. The topic is further discussed in Section 3.4.1.

3.2.3 Combination of EnergyPLAN and multi-objective evolutionary algorithms

The integration of EnergyPLAN with a MOEA provides a complete framework for energy scenario design. The corresponding scheme is described in Figure 3.1. On the upper part of the figure, the general steps of a MOEA are shown. Depending on which MOEA is used, implementation details regarding genetic operators (selection, crossover, mutation) and ranking procedures (anyway based on the dominance concept) can be different. EnergyPLAN comes into play when an individual needs to be evaluated. The lower part of Figure 3.1 shows the corresponding procedure. All required distributions (e.g., electricity demand, thermal demand, solar, on- and offshore wind availability) and relevant costs (e.g., investment costs for PV and wind or fuel costs) are fixed[1] inputs of EnergyPLAN, as they do not change during the algorithm evolution. Conversely, at each generation the MOEA provides new values of decision variables to EnergyPLAN, which in turn simulates the scenario and calculates all output parameters. The necessary objective parameters are fed back to the MOEA. The loop continues until a termination condition is met, which in our case is given by the number of specified generations. After completion of all the generations, a Pareto-optimal front is generated by the MOEA.

As regards implementation, the jMetal [58] framework is used. This is an

[1]In the simulations, transitions are not modelled but rather optimal system's configurations are determined. Learning effects in terms of investment cost reductions are therefore not endogenously modelled, but applied costs must reflect the expected costs at the time of investment. In the Aalborg scenario, costs stem from the national Danish catalogue Technology Data for Energy Plants developed for energy planning and scenario making-purposes by the Danish Energy Agency together with the Danish transmission System Operator (TSO) Energinet.dk.

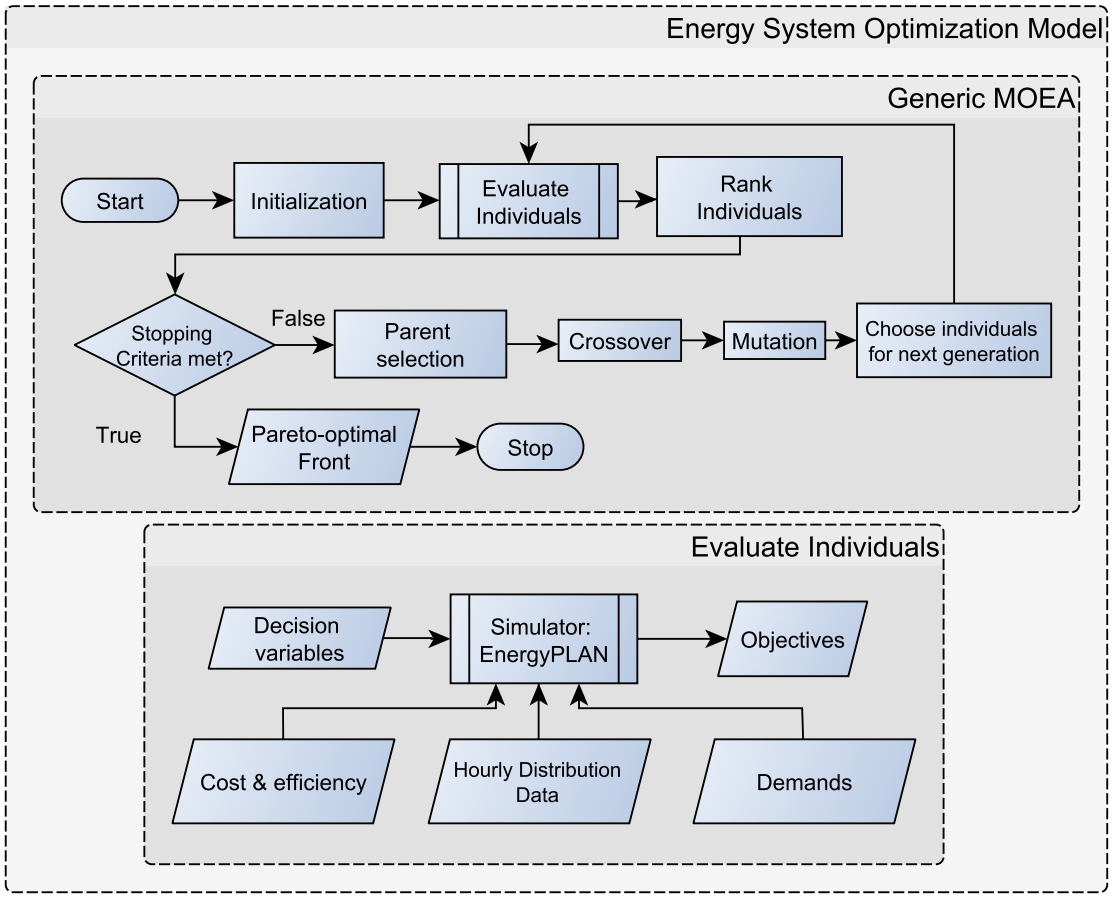


Figure 3.1: Integration of a generic MOEA and EnergyPLAN.

object-oriented JAVA-based framework for multi-objective optimization with metaheuristics. Using the advantages of this framework, the described scheme is implemented[2]. The MOEA generates an input file for EnergyPLAN, executes EnergyPLAN in command-line mode and reads main results from an EnergyPLAN output file. The output is then used to evaluate the fitness functions and proceed to the next iteration.

3.3 Case study

To test the performance of our framework, we perform simulations on the dataset of Aalborg, Denmark. We have chosen this municipality because of the extensive studies already carried out on its energetic performance, making all the required data readily available. Moreover, a possible future scenario to improve the Aalborg energy system was already proposed in the past by a pool of experts [20], thereby providing a valuable chance to compare the results of our automatic framework with a "manual" work.

[2]Source code: https://github.com/shaikatcse/EnergyPLANDomainKnowledgeEAStep1

3.3.1 Description of Aalborg scenario

Aalborg is the fourth largest city in Denmark and the main city in the municipality bearing the same name. The municipality has 200000 inhabitants distributed among the city, smaller towns and rural areas and spans more than 1000 km^2. In the current situation, energy supply is based on a mixture of (see Figure 3.2):

- Wind power.
- Electricity and district heating from coal, natural gas and biogas fired power plants.
- Individual heating (outside district heating areas) based on oil, biomass and natural gas boilers.
- A transport sector almost exclusively based on fossil fuels.

The Aalborg scenario was developed by researchers at Aalborg University for Aalborg Municipality as part of a political process to indicate ways for Aalborg to transform its energy system towards a 100% renewable energy system based on locally available renewable energy sources. The reports are both in Danish [132, 129], however main results have also been published internationally [20] and have been the starting point for further analyzes [161, 134].

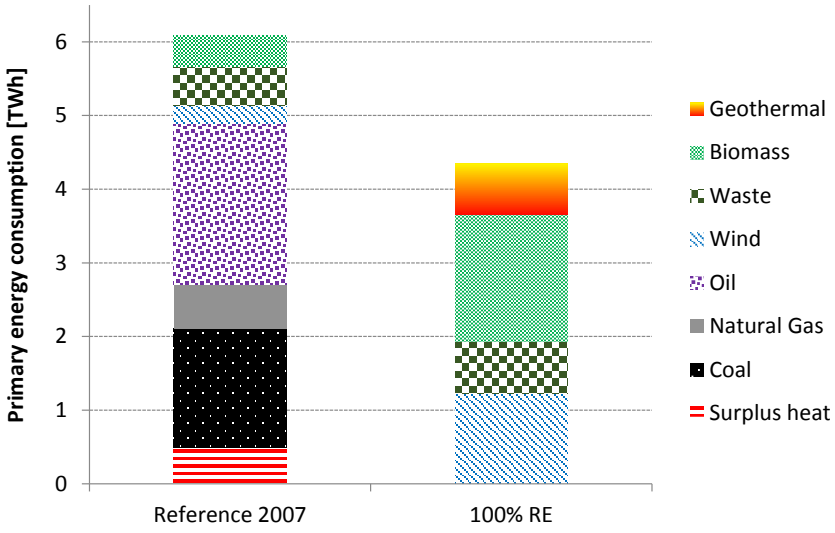


Figure 3.2: Primary energy demand for Aalborg Municipality in the 2007 reference situation as well as in the 100% renewable energy scenario.

In the process, eight distinct scenarios were designed from which one was further developed into the final scenario. EnergyPLAN was applied in the process, but as a simulation model without endogen investment optimization, all refinement was conducted iteratively outside the model. General characteristics of the transition are:

- Strong energy savings within heat and electricity.

- Phasing out of small-scale CHP plants, fossil-fuel vehicles and individual boilers for heating.
- Transition to electricity use where possible within industry, transport and heating.
- Expansion of the central district heating network to connect to outlying networks.
- Exploitation of geothermal energy for district heating.
- Increased use of heat pumps for both district heating systems and individually heated houses.
- Increased biomass use within industry, transportation and large-scale CHP.
- An electricity system which is in balance on annual basis, while on hourly basis import/export is applied.

In spite of significant electricity savings, the overall electricity demand will in fact increase as other sectors are converted to electricity as shown in Table 3.2.

Table 3.2: Electricity demand in Aalborg Municipality in the 2007 reference and in the 100% RE scenario. Any demands for Transport, Hydrogen, and Biogas production are insignificant and included in the category "Other electricity demand" in 2007. In 2007, also industry is included in "other". Thus, the 117 GWh indicated is additional to what is included in the category "Other".

[GWh]	2007 Reference	100% RE Scenario
Electrical Heating	5	0
Heat pumps	1	120
Industry	-	117
Transport	-	460
Hydrogen	-	192
Biogas Production	-	9
Other electricity demand	1047	573
Total electricity demand	1053	1470

Electricity supply will shift from CHP units based on natural gas and coal to wind power, supplemented by smaller fractions of production based on waste and biomass incineration plants and a highly efficient CHP plant (see Table 3.3).

Heat supply will almost exclusively come from central district heating, and where this is not feasible, heat pumps will be applied, as shown in Table 3.4. Overall heat losses are reduced through insulation.

3.3.2 Choice of decision variables, objectives and constraints of an energy system to optimize

There are plenty of decision variables to choose from when optimize an energy system. A handful of decision variables are not only easy to optimize, but also

Table 3.3: Electricity production in Aalborg Municipality in the 2007 reference and in the 100% RE scenario.

[GWh]	2007 Rference	100% RE Scenrio
Small-scale CHP	90	0
Large-scale CHP	670	150
Waste incineration CHP	110	100
Wind power	240	1230
Total	1110	1480

Table 3.4: Final heat demand in Aalborg Municipality in the 2007 reference and in the 100% RE scenario. Final DH demands include grid losses.

[GWh]	2007 Reference	100% RE scenario
Individual oil	78	0
Individual gas	43	0
Individual biomass	194	0
Individual HP	3	94
Individual electric	5	0
Individual solar	0.5	0.5
Bioler DH	38	0
Local CHP or HP DH	138	17
Central DH	1730	1344

help to understand the behavior or trend of a variable with respect to others. In the following, we divided a range of possible decision variables into two main categories – *production unit investments* and *flexible technologies.*

Production unit investments Within this category, the main decision variables are the capacities of the units. In order to select the most significant ones, two main arguments were taken into account: (i) the options given by local conditions (which mainly offer wind and PV) and (ii) the fact that the combination of fluctuating RES and dispatchable units (CHP / conventional power plant) is a core issue in future high-RES systems. Hence, the following decision variables were chosen for this category:

a) Onshore wind capacity.

b) Offshore wind capacity.

c) PV capacity.

d) CHP capacity.

e) Conventional power plant (PP) capacity.

Flexible technologies These are the technologies assisting the energy system to establish continuous balance between supply and demand. Among these, we choose:

a) Heat pump (HP) capacity – efficient conversion from electricity to heat. It is worth pointing out that, when shifting from CHP to HP, an extra electricity demand is introduced.

As for performance indicators or objectives to be minimized, two are selected which are at the center point of energy planning and policy:

a) CO_2 emissions.

b) Annual costs.

The optimization takes place under a number of constraints:

- Transmission line capacity of export/import: 160 MW. This constraint enforces the system to produce enough electricity so that it does not require to import more than 160 MW.
- Heat balance: This constraint enforces the system to produce exactly the amount of heat necessary to meet the heat demand.
- Grid stabilization: More than 30% of power production in all hours must come from units able to supply grid support (see [133] for details on grid stability in EnergyPLAN).

3.4 Simulation and results

In the simulation, we optimize the values of all the mentioned six decision variables with respect to the two objectives and under the three constraints specified in the previous sub-section. All other parameters are identical to the 100% RE Scenario. Table 3.5 shows the lower and upper bounds in MW used for the different decision variables.
The MOEA optimization is conducted using the following parameters:

- Population size: 100
- Number of generations: 100
- Crossover probability: 0.9
- Mutation probability: 1/number of decision variables
- Number of runs: 4

The parameters are selected to provide enough time to the MOEA to converge. However, sometimes it is possible for a MOEA to trap in local optima. Therefore, multiple runs are performed to get a well-approximated Pareto-front [28].

An important remark is in order concerning the PP capacity. In our model, no annualized investment costs are associated with PP, resembling a situation where these plants are close to end of life. In the absence of explicit costs, PP capacity gives rise to degenerate solutions. On the other hand, PP capacity can be deterministically fixed by calculating the minimum value which allows to satisfy the demand. With this second approach, only 5 free variables are left. In order to provide a

Table 3.5: Lower and upper bounds for the decision variables used in this simulation.

Decision variable	Capacity	
	Lower bound [MW]	Upper Bound [MW]
CHP	0	1000
HP	0	1000
PP	0	1000
Onshore wind	0	1500
Offshore wind	0	1500
PV	0	1500

better exemplification of our framework, we implemented both approaches: a full optimization with 6 decision variables, yielding degeneracy in the value of PP, and a partially deterministic optimization with 5 decision variables and a properly fixed PP capacity. The two approaches clearly yield the same Pareto-front in the objective space. The only difference will be given by the value of the PP capacity associated with a given Pareto point. These aspects will hence be explained in more detail when discussing the capacity trends.

3.4.1 Pareto-front obtained with MOEA

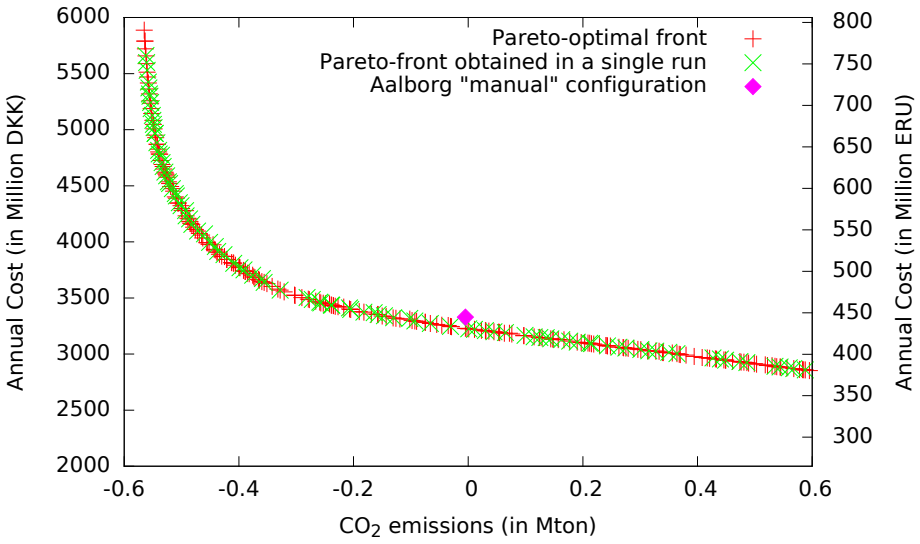


Figure 3.3: Pareto-front for deterministic PP capacity simulation.

The Pareto-optimal front obtained from the simulations is shown in Figure 3.3, where the horizontal axis refers to CO_2 emissions (in million tons) and the vertical axis to annualized costs (in million DKK). In the right side of the figure, a secondary axis is added to show the cost in million Euro.

The pink diamond corresponds to the scenario manually configured by experts [20], as described in the previous section. It is interesting to note that the manual configuration is very close to the Pareto-front, which is formed by all the other

points (+), each of them representing a scenario optimized based on CO_2 emissions or annual costs with a different configuration of the decision variables. Starting from the left side, we have scenarios that are more costly but less polluting, while moving to the right we get increasingly more polluting scenarios with lower costs. This clearly shows the additional benefit of running a multi-objective optimization: along a large range of costs, several Pareto-optimal configurations are obtained simultaneously. This opens up the possibility for the planner to choose from a whole set of optimized energy scenarios.

The numerically obtained Pareto-front is given by the series represented by red '+' sign. The front is a combination of four fronts found by four separate runs. Due to random effects in the choice of the initial population and in the application of genetic operators, results of different runs may differ. The green-cross series (×) represents the front found in the 1st run. There is a good agreement between the combined Pareto-optimal front of all four runs and the front found in the 1st run, except that the 1st run did not generate some leading points on the left side and obviously provided a less dense set. This shows that it is advisable to run the MOEA multiple times in order to obtain an accurate Pareto-front. Of course, the most convenient number of runs depends on the number of decision variables and on the computational time required by the evaluation of each configuration.

Note that some scenarios have negative CO_2 emissions. In the simulation, we not only consider the CO_2 emissions generated within the system, but also the emissions generated or avoided by importing or exporting electricity. Carrying out the analysis in island mode [135], the calculation only requires to consider CO_2 production within the system. However, in practice, most energy systems should be analyzed in connected mode, where a system is connected to the outside world for importing or exporting electricity. To calculate CO_2 emissions in connected mode, EnergyPLAN assumes that all the imported electricity is generated from condensing-mode plants; therefore, the corresponding CO_2 emissions are added to the CO_2 emissions of the internal system. It is clear that the negative emission systems/scenarios export some "green" electricity generated within the system to the outside world.

3.4.2 Capacity trends

Besides showing the resulting Pareto-front, it is interesting to discuss the values of the decision variables in the various Pareto-optimal configurations. In this regards, a short discussion will be presented in the followings. The corresponding configurations, as evident from the above figure, can be parametrized by a single objective, e.g., CO_2 emissions. We therefore present the trends of the different capacities as a function of emissions for Pareto-optimal configurations.

RES capacities As a starting point, we present the trends for PV and for the overall wind capacity (on- plus offshore) on Figure 3.4. One clearly recognizes an increase in the installed power for low emissions. Wind is generally favored, as expected for the Danish climate. Moreover, the trends are rather smooth, showing a strict correspondence between reduction in emissions and increase in RES capacity. It is worth pointing out that this relation appears to be linear at high emissions, but not at low emissions. Here, a given increase in the installed capacity yields a lower

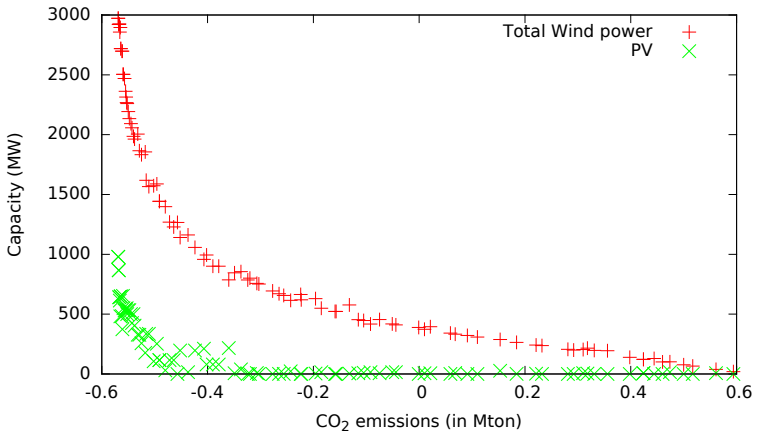


Figure 3.4: PV and total wind capacities with respect to CO_2 emissions.

effect on emission reduction. This is related to the excess electricity availability occurring in high RES scenarios (i.e., the instantaneous RES capacity occasionally exceeds the sum of internal demand and transmission line capacity for export, so that not all the potentially available energy can be exploited).

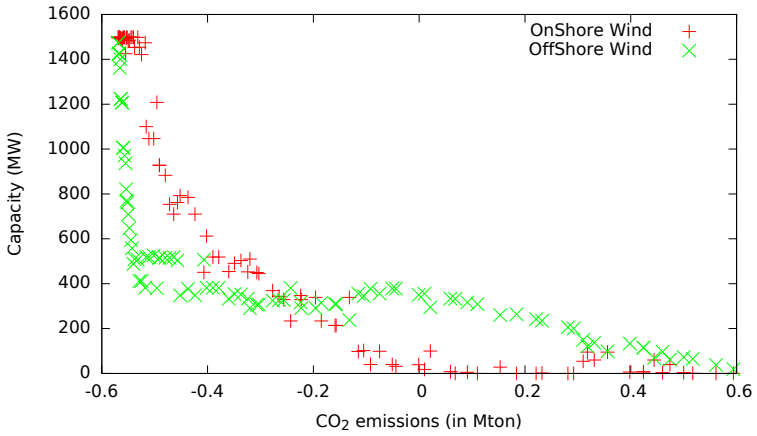


Figure 3.5: On- and offshore wind capacity trends with respect to CO_2 emissions.

To get a better insight into these effects, we discuss separately the behavior of wind capacities. Figure 3.5 shows the trends for the two kinds of considered wind power (on- and offshore). It is clear from the figure (Figure 3.4) that, while the total wind capacity smoothly increases as CO_2 emissions decrease, the behavior of the single wind technologies is more complicated. At the lowest CO_2 emission value, on- and offshore wind capacities are at their maximum. However, above about -0.2 Mt of emissions, offshore wind is favored over onshore wind and vice versa below.

Hereafter, we present a qualitative explanation of the RES behavior in the Pareto-optimal scenarios. First of all, we note that with two objectives each point on the Pareto-front can be obtained by fixing the value of one objective (at least within a proper range) and then optimizing the other. As an example, we assume to fix the emission value and to minimize the costs. We then discuss the trends occurring at different emission levels.

At a high emission level, the RES share is low and all the renewable electricity is

absorbed by the internal demand (no excess electricity production). In this situation, the intermittency of RES does not play any role: everything can be understood on the basis of yearly balances. The minimization of RES costs for a given yearly RES production is obtained by favoring the RES technology with the lowest costs per unit energy. RES costs are typically mainly fixed by the installed capacity. The corresponding investment costs can be converted into annualized costs per unit capacity considering the operational lifetime of a plant. Costs per unit energy are then obtained by taking into account the yearly operational time. For the case of Aalborg, the technology with the lowest costs per unit energy, when fully exploited, is offshore wind, which then saturates the RES share at high emission levels, forcing onshore wind and PV to zero.

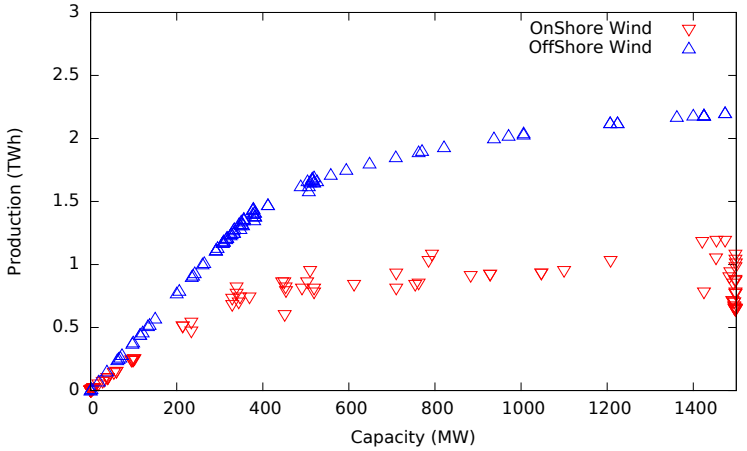


Figure 3.6: Energy production vs. installed capacity for on- and offshore wind. Production per unit capacity evidently decreases for large installations, due to excess electricity availability.

For lower emission levels, the RES share must be higher. Below a certain emission threshold, depending on the hourly distribution of demand and RES availability, excess electricity production occurs. In this situation, RES intermittency becomes significant. If the surplus even exceeds the transmission line capacity for export (at least occasionally), then not all the theoretically available electricity can be exploited, so that, for given costs per unit capacity, costs per unit energy increase. Here, it is crucial to consider the instantaneous (hourly, for EnergyPLAN) matching between the RES availability and the demand. In the extreme case, it can happen that an initially favored RES cannot provide more exploitable energy, no matter how large its capacity, as its distribution does not cover some demand periods. It is then clear that, if available, another RES with a non-zero distribution during these periods must be used. Similar effects give rise to optimal solutions with a mix of RES technologies. From a different point of view, one could say that the cost per unit energy of the different RES sources changes in the case of excess capacity (i.e., when not all the theoretically available energy can be exploited), modifying the relative convenience of the different sources during some periods. The effect of incomplete exploitation at high RES shares is evident in the above figure (Figure 3.6), plotting production as a function of capacity.

As a final comment, we note that, to increase the overall complexity, whenever there is excess capacity and more than one RES, one needs to establish a prioritiza-

tion strategy. The latter will affect the actual production from a certain RES and hence its costs per unit energy.

We point out that it would be very complicated to analyze in detail these effects, which necessarily involve the actual hourly distributions. While not able to provide a full insight into the various driving forces at play, an automatic optimizer is hence a needful tool to quickly obtain the desired solutions.

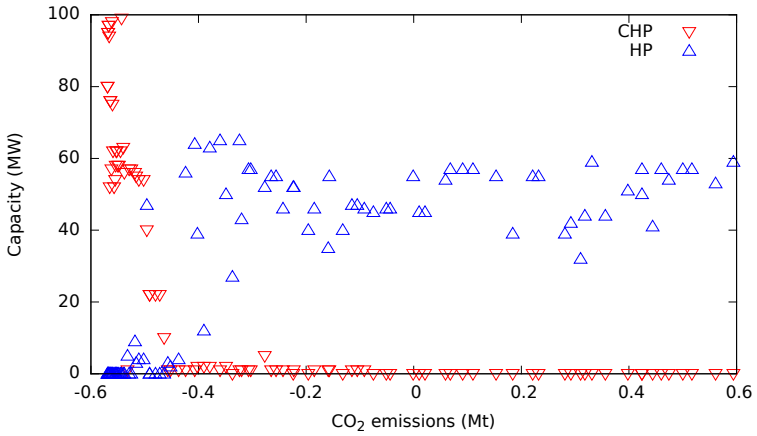


Figure 3.7: CHP and HP capacities with respect to CO_2 emissions for the optimization of the energy system of Aalborg. HPs are used in three cases; individual houses, the central DH system and local DH system. The HPs here are the HPs used in the central DH system.

Non-RES capacities Figure 3.7 shows the CHP and HP capacities with respect to CO_2 emissions. CHP and HP change rather abruptly (and with opposite directions, suggesting they compensate each other) between the minimum value of CO_2 emissions and about -0.45 Mt; for CO_2 emissions above -0.45 Mt they are constant on average. CHP reaches almost 0 MW and HP is around 50 MW capacity on average. The apparent volatility of the HP capacity is discussed later.

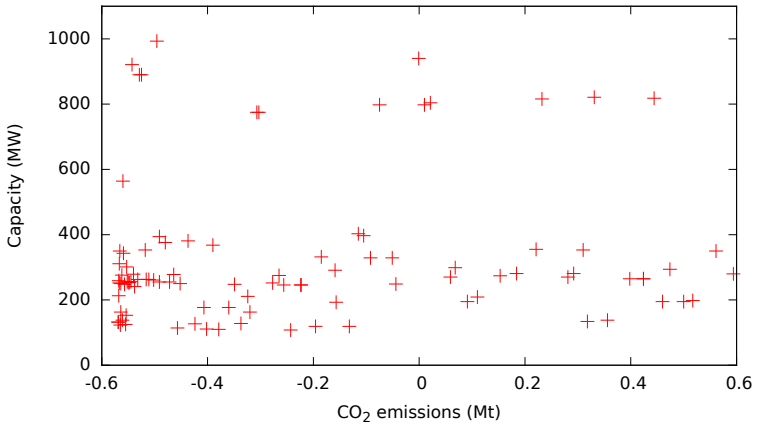


Figure 3.8: PP capacity with respect to CO_2 emissions for the optimization of the energy system of Aalborg.

As anticipated above, for the determination of the PP capacity two different approaches were followed. First, we present the results obtained including the PP capacity among the decision variables. The PP behavior is reported in Figure 3.8, where the capacity appears to be roughly constant on average, but with significant fluctuations. This is in agreement with the absence of a cost associated to PP capacity. Costs are only taken into account for fuel consumptions, which are however independent of capacity when the latter is large enough. Hence, it can happen that the PP capacity is quite large, but there is no significant production from PP. From the point of view of the optimization, the MOEA does not have any motivation to lower the capacity of PP, as it focuses on the minimization of cost and CO_2 emissions.

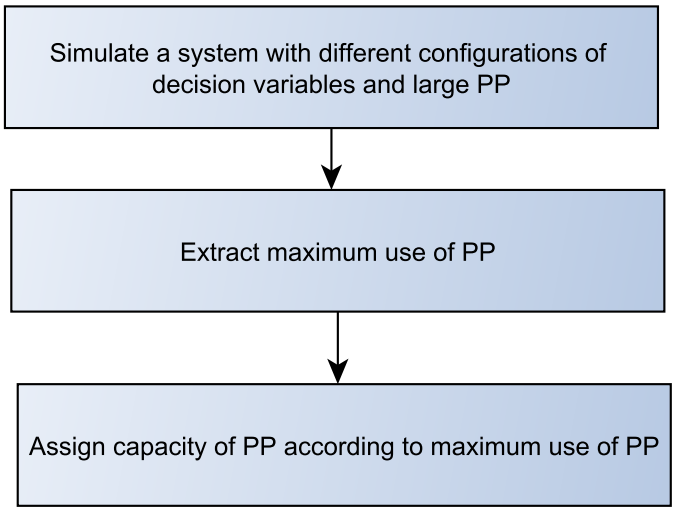


Figure 3.9: Flow chart for deciding PP capacity.

The approach described before exemplifies well a degenerate case. In our second approach, we remove this degeneracy by fixing the PP capacity deterministically. The approach is split in two steps. In the first step, a scenario starts with a very large capacity of PP (for our case, 1000 MW; it can be larger or smaller depending on the total electricity demand). In the internal technical optimization of EnergyPLAN, PP is the least prioritized unit; hence, the production of PP depends on the configuration of the other electricity-producing units. As an example, if other units (e.g., onshore wind, offshore wind, PV) are large enough to meet the electricity demand in a given hour, then the PP is not used even if there is a large PP capacity. Therefore, in the second step, the actual capacity of PP is decided by setting it to the maximum use of the PP for the obtained scenario. Figure 3.9 illustrates the process. The described process is applied in the evaluation phase of the MOEA (see Figure **??**), that is, PP is set to the minimum possible value at each iteration. This would not be necessary from the point of view of the MOEA, but it yields the possibility to follow the evolution of the variable after each generation.

Figure 3.10 shows the new trend of PP capacity with respect to CO_2 emissions. The new trend is much smoother than the trend of Figure 3.8. The new trend shows an increasing pattern when the CO_2 emissions of the system are increased. The optimization process is clearly much more efficient when this second approach

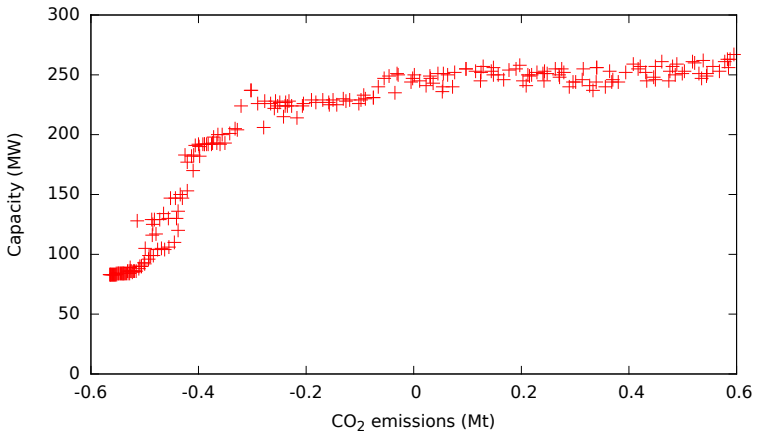


Figure 3.10: Trend of PP with respect to CO_2 emissions in deterministic PP capacity simulation.

is used, as only 5 decision variables are actually optimized. On the other side, in a model including investment costs for PP (perceived as being beyond boundary in EnergyPLAN), it would be necessary to resort to the full optimization of the 6 variables.

Summary about capacity trends As some of the shown trends can be explained qualitatively, while others are more elusive. Moreover, we pointed out as it can happen that multiple configurations have very similar performances (or even identical, in the presence of a degeneracy).

The Pareto-optimal front is a collection of different configurations, possibly resulting from different runs. If one of the decision variables has little influence on the overall performance (or if two variables can be exchanged to give a similar overall result), then it can occur that even for close configurations on the Pareto-front such variable changes significantly and randomly. As an example, if one considers the trends of CHP and HP on Figure 3.7 when emissions are very low (negative), some exchange configurations between CHP and HP can be found. At -0.5 Mt CO_2 emission, one finds a configuration with almost the same amount of capacity (40 MW for CHP and 45 MW for HP). On the left side of the position, there are multiple configurations where HP capacity is low and CHP capacity is high. In conclusion, it is not possible to easily predict and interpret all the trends. Multiple configurations (systems) with similar performance (i.e., similar objective values) can give rise to scattered behaviors, even more complicated to interpret than those discussed for RES. This discussion emphasizes the importance of having access to numerical optimization for the identification of all possible optimal scenarios in energy systems.

3.5 Compromised scenarios

Despite giving the decision makers a very large set of scenarios, there are several ways to focus on some particular scenarios. We would like to discuss one method here to identify some best compromised scenarios. The best compromised solution

is a solution within all solutions of a Pareto-front that is closest to the ideal point [3]. Section 2.4 describe the details about the compromised solution and the ideal point. We will present 10 best compromised solutions (i.e., 10 solutions that have minimum distances from the ideal point) for Aalborg energy system.

Firstly, it is required to normalized as described in section 2.4 because the optimization functions are in different units. Considering that there are n number of scenarios in Pareto-front and Z^* is the ideal point (where z_j refer to value of j^{th} objective value), the normalized j^{th} objective value of i^{th} scenario ($s^i_{j_{nor}}$) will be:

$$s^i_{j_{nor}} = \frac{s^i_j - z_j}{\max_{1 \leq k \leq n} s^k_j - z_j} \tag{3.1}$$

Using the formulation, all the objective values are normalized in a scale between $[0, 1]$. After normalization, 10 scenarios are identified those have least Euclidean distances from the ideal point. Figure 3.11 shows 10 compromised solutions (marked in blue), the ideal point is marked by green colored star.

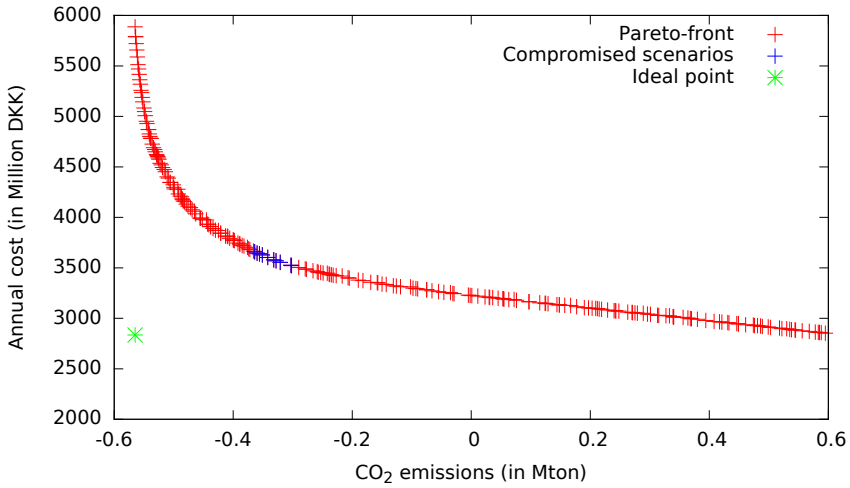


Figure 3.11: Pareto-front and the 10 best-compromised scenarios (marked with blue color) for Aalborg energy system.

3.6 Conclusion

In this chapter, we have presented a new simulation framework combining a multi-objective evolutionary algorithm and a descriptive analytical energy model. We have shown that it is possible to automatically identify a set of energy scenarios which optimize conflicting objectives and satisfy given constraints. For the case study of the Aalborg energy system, it was confirmed that the manual configuration available from the literature is quite close to optimality; at the same time, the presented framework allowed to obtain the full Pareto-front under the chosen decision variable bounds. This illustrates a broad choice of optimal scenarios available to decision makers, corresponding to different possible trade-offs between costs and emissions.

[3] An ideal point is a point on objective space where each coordinate is taken from minimum value of the corresponding objective of all the solution in Pareto-front.

Even for a small number of decision variables, the identification of a Pareto-optimal configuration with manual search is evidently a hard task, not to mention the impracticality of locating by hand multiple optimal scenarios to represent the Pareto-front in a satisfactory way. Combinatorial approaches, which could be introduced on the discretized space obtained by approximating decision variables within a certain accuracy, explode exponentially with the space dimensionality, i.e., with the number of decision variables themselves. Also classical optimization methods often encounter difficulties in this context, for example due to the multi-modal and degenerate nature of this type of problems. All in all, this supports the use of meta-heuristics techniques. In particular, we successfully used the NSGA-II as a MOEA – a readily available optimizer whose efficiency in solving complex problems is well known from literature.

We have presented some criteria for the selection of decision variables. In spite of the optimizer efficiency, it is of course convenient to restrict as much as possible the number of degrees of freedom, as it strongly affects the complexity of the problem and hence the performance of any search method. Even with a reasonable number of decision variables and on a common workstation, the corresponding computational time is expected to be of the order of a few hours. Carrying out multiple runs, which is always advisable to confirm the quality of the obtained front, would similarly be reasonable under these conditions.

Finally, a simple techniques is adopted to identify some compromised scenarios that may help decision markers for selecting scenarios.

In the next chapter, we will propose some techniques that utilize domain knowledge regarding energy systems to improve the efficiency of the framework. Domain knowledge will be integrated into different phases of the framework. It is expected that the framework will be more efficient in terms of finding optimal solutions.

Chapter 4

Incorporating domain knowledge into the framework

The key to artificial intelligence has always been the representation.

— Jeff Hawkins

4.1 Introduction

Energy plays a key factor in the advancement of humanity. As energy demands are mostly met by fossil fuels, the world-wide consciousness grows about their negative impact on the environment. Therefore, it becomes necessary to design sustainable energy systems by introducing renewable energies. Because of the intermittent availability of different renewable resources, the designing of a sustainable energy system should find an optimal mix of different resources. However, the optimization of this combination has to deal with a number of possibly conflicting objectives.

In the previous chapter, we have proposed a multi-objective optimization framework to design energy scenarios. The framework requires to simulate a large number of scenarios to identify optimized scenarios. Generally, simulation models are computationally costly, therefore, we want to optimize energy systems more efficiently. In particular, we want to achieve this by generating a high-quality approximation of the Pareto-front [51] at reduced computational cost. To reach this goal, we investigate the incorporation of domain knowledge related to energy systems into the different phases of a MOEA. Firstly, we propose a smart initialization technique and secondly, incorporate a smart mutation [110]; both exploit domain knowledge. The results clearly show that all these individual methods have an impact to optimizing an system. To the best of our knowledge, this is the first attempt to incorporate energy system domain knowledge into different operators of MOEAs. In this study, Danish Aalborg energy system problem is taken as a test problem to demonstrate

M. S. Mahbub, M. Wagner, and L. Crema, "Incorporating domain knowledge into the optimization of energy systems," *Applied Soft Computing*, vol. 47, no. C, pp. 483–493, Oct. 2016. DOI: http://dx.doi.org/10.1016/j.asoc.2016.06.013
The text is verbatim apart from the introduction and the conclusion and mentioned additions in chapter 4 and chapter 6; section 4.2.2.2 and 4.2.3 are extensively modified, paragraph 4.2.2.2 and 4.2.2.2 are added, moreover, new results are included in paragraph 4.4.3 compared to published version.

the feasibility of our approach. It is a well-understood problem, and the details are readily available as we have described in previous chapter.

The Chapter is organized as follows. Most of Section 4.2 discusses how domain knowledge is represented and how it can be incorporated into a MOEA through problem-specific initialization. A proposed smart mutation techniques is described in Sections 4.2.3. Then, we describe the details of all the performed experiments and the corresponding results in Section 4.4. Finally, we draw our conclusions in Section 5.5.

4.2 Incorporating domain knowledge

In general, a typical MOEA cannot perform well for all classes of problems, as this would be contradictory to the No Free Lunch Theorem [26]. According to this theorem, the average performance of an algorithm over all possible classes of problems is constant. Hence, the good performance of an optimization algorithm on one class of problems is balanced out by the bad performance of the algorithm on another class of problems. However, this also means that problem-specific algorithms with above-average performance are possible. Bonissone et al. [26] define two different ways to achieve this by incorporating domain knowledge: *implicitly* and *explicitly*. Encoding, design of data structures and constraints representation are categorized as an implicit incorporation of domain knowledge. Our proposed approaches mainly focus on the explicit incorporation (i.e., smart seeding of initial population, mutation exploiting domain knowledge) for the energy system optimization problem. In the following sections, we will discuss how we represent domain knowledge of energy systems and how we incorporate this knowledge into *initialization* and *mutation* phases.

4.2.1 Domain knowledge related to energy systems

Typically, experts of a field can provide detailed domain knowledge about the field, and laypeople can at times provide very basic knowledge or "rules of thumb". In the following, we will use such basic knowledge.

We described in previous chapter that the main focus of designing an energy system is to minimize *CO_2 emissions* and the *total annual cost*. In this context, it is obvious that some decision variables have influence on some objectives. For example, increasing the capacities of renewable resources can reduce CO_2 emissions. At the same time, decreasing the capacities of renewable sources can reduce the annual cost as no additional investment cost is required. To encode such basic knowledge, we will use the following method: we mark each decision variable with *true*, *false* or *null* for each objective. *True* and *false* indicate *increasing* and *decreasing*, respectively, of the value of a decision variable to minimize an objective. *Null* indicates that there is no domain knowledge available for the decision variable for the objective.

A real-life example of DK representation in the context of energy systems is given in Table 4.1. $W_{C_{\text{off}}}$ and $W_{C_{\text{on}}}$ represent off- and on-shore wind power capacities, respectively. PV_C is the photovoltaic's and PP_C is the power plant's capacity. C_S, O_S and NG_S represent the coal, oil and natural gas shares, respectively, to fire power plants. Finally, $DK_{O_{EM}}$ and $DK_{O_{AC}}$ are the domain knowledge associated with the minimization of emission and annual cost, respectively. The first row of the

table can be interpreted as increasing (i.e., off-shore, on-share wind, PV capacities; natural gas share to fuel power plants) and decreasing (i.e., coal share) of some decision variables that could minimize the objective emissions. The second row can be interpreted in similar fashion.

Table 4.1: A real-life domain-knowledge representation example.

DVs / DK_{Obj}	$W_{C_{off}}$	$W_{C_{on}}$	PV_C	PP_C	C_S	O_S	NG_S
$DK_{O_{EM}}$	true	true	true	null	false	null	true
$DK_{O_{AC}}$	false	false	false	null	true	null	false

4.2.2 Smart initialization

Most experimental studies on MOEAs use random initialization to initialize the population. There, the initial values of the decision variables are drawn from a uniform distribution within the lower and upper bounds of the variables. In practice, however, optimizers can typically be seeded with good candidate solutions either previously known or created according to some problem-specific method. This seeding has been studied extensively for single-objective problems. For multi-objective problems, however, very little literature is available on the approaches to seeding and their individual benefits and disadvantages.

Friedrich and Wagner [73] provide a recent overview on MOEA seeding, and also a comprehensive study on 48 artificial MOO problems for five different MOO algorithms. Their generic seeding approaches are based on linear combinations of the objectives, and the individual seeds are computed by a single-objective optimizer that solves a particular linear combination. These linear combinations are quite "evenly spread out" in order to achieve an unbiased but close initial population. They observe that some problems benefit significantly from their seeding strategies, while others profit less. The advantage of seeding also depends on the examined algorithm.

In contrast to their problem-independent seeding approaches, our method initializes a decision variable by using actual domain knowledge. We enable laypersons to provide some domain knowledge in a very basic manner. For example, "decreasing coal share could minimize the objective emissions" is encoded as "false" in Table 4.1. Our method in the following section then translates this basic knowledge into a diverse set of seeds. We call this process *smart initialization*, and we describe the details in the following.

4.2.2.1 Methodology

The initial value of a decision variable can be defined as follows:

$$dv_i = lb_i + (ub_i - lb_i) * \bar{\delta} \tag{4.1}$$

where lb_i and ub_i are the lower and upper bounds of the decision variable, respectively. We calculate $\bar{\delta}$ from the probability distributions described below.

Probability distribution The following proposed probability distribution is used when a decision variable is needed to be initialized with higher values:

$$p(\delta) = (\beta + 1)\delta^{\beta} \tag{4.2}$$

And the following probability distribution is used when a decision variable is needed to be initialized with lower values:

$$p(\delta) = (\beta + 1)(1 - \delta)^{\beta} \tag{4.3}$$

These distributions are valid for $\delta \in (0, 1)$. β takes a non-negative value that is used to control the shape of the distribution.

Note that a uniform distribution cannot be used to incorporate any domain knowledge. While Gaussian distributions can be used, it is not as easy as with ours to incorporate domain knowledge. We will see in the following how we can control the shape of the distribution in order to initialize values closer to either the lower or upper bounds of the decision variables.

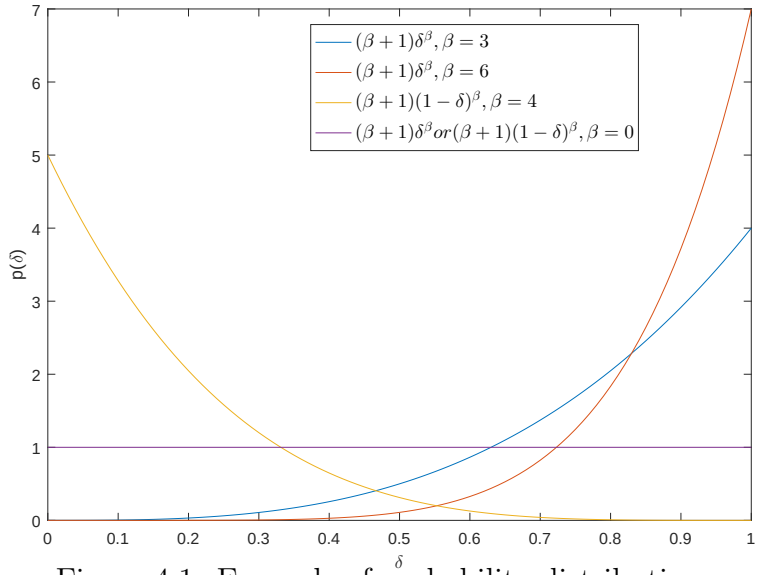


Figure 4.1: Example of probability distributions.

Figure 4.1 illustrates four examples with two different distributions and different β values. The blue and red curves are the plots of Equation (4.2), however, the red curve is steeper than the blue curve due to the larger β value. Hence, with larger β value, it is more probable that the initial value will be closer to the upper bound of the decision variable. For lower β values, the distribution will be flatter, hence, increasing the probability of an initial value to fall within the entire range between the upper and lower bounds, rather than near the upper bound. Eventually, when $\beta = 0$ the distribution will be equivalent to the *uniform distribution* (i.e., purple line in Figure 4.1) that is commonly used for random initializations. Additionally, the orange curve is an example of a distribution based on Equation (4.3). In summary, by controlling the β value and by using the appropriate distribution, one can have a better control over the initial value of a decision variable than with random initialization. Note that, while we could manually set exact values for the initial values, domain knowledge is often a bit "fuzzy", and our distributions allow for random but biased variations of the initial values.

Now, the analytical formula for $\bar{\delta}$ (used in Equation (4.1)) can be found by calculating the inverse of the cumulative distribution function [75](also known as inversion method [56])[1].

$$\int_0^{\bar{\delta}} p(\delta)d\delta = t \tag{4.4}$$

where t is a random number within $[0, 1]$ drawn from a uniform distribution. When Equation (4.2) is used as probability distribution, the analytical formula for $\bar{\delta}$ is

$$\bar{\delta} = t^{\frac{1}{\beta+1}}. \tag{4.5}$$

When lower values are preferred (Equation (4.3)), the formula is

$$\bar{\delta} = 1 - (1-t)^{\frac{1}{\beta+1}} \tag{4.6}$$

4.2.2.2 Generating initial population

As it is not possible to know in advance which combinations of β values for the decision variables of an individual will be suitable to generate the initial biased populations, we will consider all possible combinations. If the problem at hand has d decision variables and if we consider b different values for β, then the total number of combinations is $n_c = b^d$. Considering an example where $d = 2$ and $b = 3$, therefore, $\beta \in [0, 1, 2]$ and total number of combination is 9. Table 4.2 shows all the combinations for β values. Moreover, for each of the o objectives, there are n_c combinations of β values, and for each combination we can generate k individuals to achieve some diversity among the biased solutions. Hence, we generate $n_{DK} = o * n_c * k$ individuals in total.

Table 4.2: Combination of all β values when $d = 2$ and $b = 3$.

β_{d_1,d_2}	β_{d_1,d_2}	β_{d_1,d_2}
(0,0)	(1,0)	(2,0)
(0,1)	(1,1)	(2,1)
(0,2)	(1,2)	(2,2)

Most evolutionary algorithms, however, only work with a population that is significantly smaller than n_{DK} for real-world problems. For example, with $o = 3$, $b = 3$, $d = 5$, and $k = 10$ the initial population size would be $n_{DK} = 7290$, which is about two orders of magnitude above what is typically used in studies. A population of this size tends to slow down the actual optimization algorithm as the computational complexity of the algorithm is typically dependent on the population size. To solve this problem, we provide a procedure for reducing the number of individuals down to a fixed number. Algorithm 3 presents our overall process of generating initial populations that uses domain knowledge.

Reducing the number of individuals Due to combinatorial explosion, our previously proposed methodology generates many individuals. As we want to reduce the number of individuals, we adapt the concept of *decision space diversity* to cover the space efficiently using a fixed number n_p of individuals.

[1] A process for generating random variates from a specific probability distribution.

Algorithm 3 Algorithm for generating initial population

Require: DK_{o_i} – an array contains domain knowledge regarding i^{th} objective (e.g., 1^{st} row of Table 4.1)

```
for all domain-knowledge arrays DK_{o_i} do
    for all combinations (c) from n_c number of combinations do
        Generate k individuals using Algorithm 4
    end for
end for
Reduce number of individuals from n_DK to n_p
```

Algorithm 4 Algorithm for generating a single individual

Require: c – a combination for β values

```
for all dv_j do
    if DK_{o^i}^{dv_j} is true then
        Use Equations (4.5) and (4.1) to generate the value of the decision variable; value of β is taken from the combination c
    else if DK_{o^i}^{dv_j} is false then
        Use Equations (4.6) and (4.1) to generate the value of the decision variable; value of β is taken from the combination c
    else if DK_{o^i}^{dv_j} is null then
        Generate a random value within upper and lower bound from uniform distribution (i.e., use Equations (4.5) or (4.6) with β = 0)
    end if
end for
```

Ulrich [167] categorizes diversity measurements depending on the way they are calculated. The categorization is based on (i) relative abundances, (ii) taxonomy, (iii) aggregating the dissimilarities, and (iv) utility of solutions. In the first class (relative abundance), diversity is calculated by measuring the relative abundance of each solution within a population set, with one example being the Shannon entropy [117] metric. The metrics in the second class use the path length within a taxonomy tree, where the solutions are arranged in a tree that reflects the taxonomic classification of the solutions. Clustering metrics belong to this class, and the calculation of metrics from this class typically suffers from high computational cost. The metrics of the third kind are computed by summing up all the dissimilarities between all the individuals. For example, Shir et al. [155] use a metric from this class to enhance diversity in a MOEA. The last class is based on measuring the utility of solutions. Solow and Polasky [156] proposed a metric that uses a utilitarian view on solutions. There, each solution has a pre-defined utility value, and the key idea is that the total utility of a population does not increase by adding duplicate individuals. Ulrich [167] has shown that this Solow and Polasky metric (a metric of the forth class) is the only one to fulfill all the three basic requirements of a diversity measure (i.e., monotonicity in varieties, twinning and monotonicity in distance). In our proposed approach, we will use the Solow and Polasky metric to measure the diversity of an initial population.

Let us consider a population P with k individuals $(A_1, A_2, \ldots, A_k)$, and let $d(A_i, A_j)$ be the Euclidean distance in the decision space between A_i and A_j (i.e., $d(A_i, A_j) = \sqrt{\Sigma_{l=1}^{d}(dv_{A_i}^l - dv_{A_j}^l)^2}$, where $dv_{A_i}^l, dv_{A_j}^l$ are the l^{th} decision variable of A_i and A_j individuals, respectively). Then each element $m_{i,j}$ of a $k \times k$ matrix M can be defined as $m_{i,j} = exp(-\theta * d(A_i, A_j))$, where θ is a normalizing parameter

between distance and number of individuals. Finally, the Solow and Polasky metric is the summation of all the elements of the M^{-1} matrix. Algebraically, it can be written as follows:

$$SP(P) = vM^{-1}v^T \tag{4.7}$$

where $v = (1, 1, \ldots, 1)$ is a row vector of size $1 \times k$ and v^T represents the transpose of v. Intuitively, the metric measures the number of different individuals present in the population. The individuals that are close to each other are considered as the same individuals, which can be adjusted via the value of θ.

The initial population P_{DK} from our previous approach contains $n_{DK} = o * n_c * k$ individuals. As we need to reduce the number of individuals from initial population, the idea is to select a subset P_I of size n_p from P_{DK} that maximizes the population's diversity. The problem is formulated as follows:

$$\underset{P_I \subseteq P_{DK}}{\operatorname{argmax}} SP(P_I) \tag{4.8}$$

As it is not practical to consider all possible subsets because of combinatorial explotion, we use the greedy approach proposed in [167]. This approach iteratively removes the individual that contributes the least to the Solow and Polasky metric. The iterative process stops when the population size reaches n_p. In the following the approach is described in details.

Calculate contribution of a solution to Solow-Polasky metric Considering M is a symmetric matrix because of symmetry of distance measurement (i.e., $d(A_i, A_j) = d(A_j, A_i)$. Therefore, M can be partitioned into following form:

$$M = \begin{pmatrix} G & f \\ f^T & e \end{pmatrix}, M^{-1} = \begin{pmatrix} \bar{G} & \bar{f} \\ \bar{f}^T & \bar{e} \end{pmatrix} \tag{4.9}$$

where f and $\bar{f}$ are column vectors and f^T, $\bar{f}^T$ are transpose of the two vectors, respectively. Moreover, e and $\bar{e}$ are single elements. Using block matrix inversion method [25],

$$G^{-1} = \bar{G} - \frac{1}{\bar{e}} \bar{f} \bar{f}^T \tag{4.10}$$

Please note that $SP(P)$ is the summation of all the element of M^{-1} matrix. In addition, please consider that we want to calculate the contribution of the individual that is related to last row and column of M. Hence, our goal is to calculate $\sum(M^{-1}) - \sum(G^{-1})$, where $\sum(M^{-1})$ refers to the summation of all the elements in M^{-1}. The contribution can be calculated as:

$$\begin{aligned} \Sigma(M^{-1}) - \Sigma(G^{-1}) &= \Sigma(\bar{G}) + 2\Sigma(\bar{f}) + \bar{e} - \Sigma(\bar{G}) + \frac{1}{\bar{e}}\Big(\Sigma(\bar{f})\Big)^2 \\ &= \frac{1}{\bar{e}}\left[(\bar{e})^2 + 2\bar{e}\Sigma(\bar{f}) + \Big(\Sigma(\bar{f})\Big)^2\right] \\ &= \frac{1}{\bar{e}}\Big(\Sigma(\bar{f}) + \bar{e}\Big)^2 \end{aligned} \tag{4.11}$$

The term $\frac{1}{\bar{e}}\Big(\Sigma(\bar{f}) + \bar{e}\Big)^2$ is the contribution from the element of the last column's or last row's M^{-1}. The least contributed element is identified by comparing all the contributions.

Alternate approach for maximizing population diversity The previous approach suffers the problem of lager computational time when number of decision variables are increased. Therefore, an alternative diversity maximization approach is proposed later. This alternate approach is effective when number of decision variables are increased a lot, hence, the number of individuals in initial population are very large (as $n_{DK} = o * n_c * k$ and $n_c = b^d$, please see Section 4.2.2.2). Therefore, inverting the large M^{-1} will take considerable amount of time, in addition, the completion of the phase (i.e, reducing number of individuals) will take a signification amount of time compared to others phases of optimization. As oppose to previous approach, an alternative approach could be – starting with a small population (i.e., matrix) and add the individual to the population that contributes most to increase the diversity of the population.

Now considering Q is the smaller matrix and we want to calculate the contribution of adding a new individual (n_i), therefore, the new matrix becomes as follows:

$$U = \begin{pmatrix} Q & r \\ r^T & s \end{pmatrix} \tag{4.12}$$

Where r is the distance vector that contain all the distances from newly added individual (n_i) to other individuals. The goal is to calculate $\sum(U^{-1}) - \sum(Q^{-1})$ without inverting U. Form bordering method [25], it is easily possible to find U^{-1}.

$$U^{-1} = \begin{pmatrix} Q^{-1} + \frac{1}{v}Q^{-1}rr^TQ^{-1} & -\frac{1}{v}Q^{-1}r \\ -\frac{1}{v}Q^{-1}r & \frac{1}{v} \end{pmatrix}, \tag{4.13}$$

where $v = s - r^TQ^{-1}r$. Now, considering $W = [w_1, w_2, \ldots, w_n]^T = Q^{-1}r$, then U^{-1} becomes:

$$U^{-1} = \begin{pmatrix} Q^{-1} + \frac{1}{v}WW^T & -\frac{1}{v}W \\ -\frac{1}{v}W^T & \frac{1}{v} \end{pmatrix} \tag{4.14}$$

Therefore, $\sum(U^{-1}) - \sum(Q^{-1})$ is the sum of all the elements of

$$\frac{1}{v}\begin{pmatrix} WW^T & -W \\ -W^T & 1 \end{pmatrix} \tag{4.15}$$

Finally by simplifying,

$$\begin{aligned} \Sigma(U^{-1}) - \Sigma(Q^{-1}) &= \frac{1}{v}\left[\Sigma(WW^T) - 2\Sigma(W) + 1\right] \\ &= \frac{1}{v}\left(\phi^2 - 2\phi + 1\right) = \frac{1}{v}(\phi - 1)^2 \end{aligned} \tag{4.16}$$

Where $\phi = \sum_{i=1}^{n} w_i$. By comparing all these terms, it is possible to find which individual contributes most to increase diversity of the previous population.

In this approach, the process starts with a randomly selected individual and continue adding individuals (in each iteration an individual is added to the population) until the population size reaches n_p. This approach requires less time compared to other approach, however, other approach is more efficient to maximize diversity of the population; as it starts with whole population and eliminates individuals one-by-one.

4.2.3 Smart mutation

While we expect the smarter initialization to give the optimization a "head start", we also want to speed up the actual optimization process itself. We developed the idea of *smart mutation*, which is based on the same concept of domain knowledge utilization as is our initialization strategy. In this proposed technique, the polynomial mutation as proposed by Deb [51] is modified to create two additional mutation operators called *renewable energy favor mutation (REFM)* and *conventional energy favor mutation (CEFM)*. Polynomial mutation mutates an offspring so that the values of the decision variables can be increased or decreased depending on a randomly generated value. On the other hand, both REFM and CEFM use the same domain knowledge representation as described in Section 4.2.1. For example, the first and second row of Table 4.1 can be used for REEM and CEFM, respectively. In addition, REEM mutates the decision variables to be increased or decreased based on domain knowledge represented by the first row of Table 4.1 and in the same way, the second row can be interpreted for CEFM. In the following, we will briefly outline the details of smart mutation.

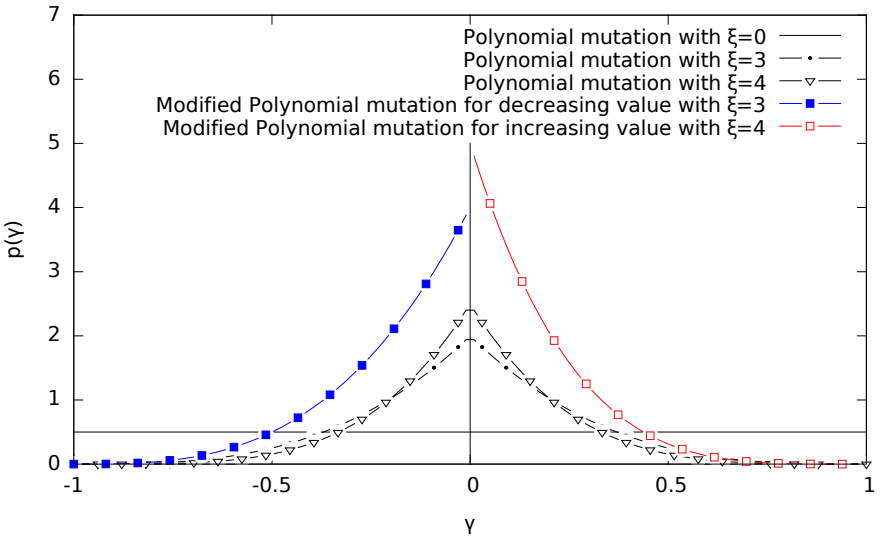


Figure 4.2: Probability distributions for polynomial and modified polynomial mutation.

Polynomial mutation In polynomial mutation, a probability distribution is used to perturb a decision variable of an offspring solution.

$$\bar{\mu} = \mu + (\mu^U - \mu^L)\bar{\gamma} \tag{4.17}$$

where $\bar{\mu}$ is the mutated offspring decision variable, μ is the offspring decision variable that will be perturbed. μ^U and μ^L are the upper and lower bounds of μ, respectively. $\bar{\gamma}$ will be calculated from the following probability distribution.

$$p(\gamma) = 0.5(\xi + 1)(1 - |\gamma|)^{\xi} \tag{4.18}$$

where ξ is a parameter that takes only non-negative values, called distribution index. This distribution is valid for $\gamma \in (-1, 1)$. Three black lines of Figure 4.2 illustrate the distributions with different distribution index values.

Followings are the analytical formulas for $\bar{\gamma}$ (in Equation (4.17)), calculated from the probability distribution given by Equation (4.18). Equation (4.19) is valid when no upper and lower bounds of a variable are specified.

$$\bar{\gamma} = \begin{cases} (2r)^{\frac{1}{\xi+1}} - 1, & \text{if } r \leq 0.5 \\ 1 - [2(1-r)]^{\frac{1}{\xi+1}}, & \text{otherwise} \end{cases} \tag{4.19}$$

Where r is a random number within a range of 0 to 1, drawn from a uniform distribution.

When an upper and lower bounds of a parent variable are defined, the calculation of $\bar{\gamma}$ is changed in the following way so that the new value will be within specified range.

$$\bar{\gamma} = \begin{cases} [2r + (1-2r)(1-\gamma^{max})^{\xi+1}]^{\frac{1}{\xi+1}} - 1 & \text{if } r \leq 0.5 \\ 1 - [2(1-r) + 2(r-0.5)(1-\gamma^{max})^{\xi+1}]^{\frac{1}{\xi+1}}, & \text{otherwise} \end{cases} \tag{4.20}$$

Where $\gamma^{max} = \min\{(\mu - \mu^L), (x^U - x)\}/(\mu^U - \mu^L)$.

Modified polynomial mutation to increase value When it is required to increase the value of a decision variable (e.g., require for REFM and CEFM), a modified probability distribution function is used instead of Equation (4.18).

$$p(\gamma) = 0.5(\xi + 1)(1 - \gamma)^{\xi} \tag{4.21}$$

The distribution is valid for $\gamma \in (0, 1)$. This distribution insures that the value of an offspring decision variable can only be increased or stay same as before. The black line with downward triangles in Figure 4.2 illustrates the probability distribution used in a polynomial mutation with distribution index 4. Whereas, the red line with empty rectangles in the figure shows the modified distribution with same distribution index. The closed/analytical form of $\bar{\gamma}$ for the above distribution is as follows:

$$\bar{\gamma} = 1 - [1 - r + r(1-\gamma^{max})^{\xi+1}]^{\frac{1}{\xi+1}} \tag{4.22}$$

Where r is a random number within a range of 0 to 1, drawn from a uniform distribution. And $\gamma^{max} = \min\{(\mu - \mu^L), (\mu^U - \mu)\}/(\mu^U - \mu^L)$. This closed form of $\bar{\gamma}$ is used in Equation (4.17) to calculate the new mutated value.

Modified polynomial mutation to decrease value In the following, the modified probability distribution and the closed form of $\bar{\gamma}$ are presented, when the value of a decision variable of an offspring needs to be decreased:

$$p(\gamma) = 0.5(\xi + 1)(1 + \gamma)^{\xi}, \text{ valid for } \gamma \in (-1, 0) \tag{4.23}$$

$$\bar{\gamma} = [r + (1-r)(1-\gamma^{max})^{\xi+1}]^{\frac{1}{\xi+1}} - 1 \tag{4.24}$$

At the beginning of an optimization, the probabilities of the problem-specific mutations (i.e., REFM and CEFM) are maximal and these decrease over time. As generations pass, the probability of generic mutation (i.e., polynomial mutation) is increased. This dynamic adjustment of the mutation probabilities ensures that modified mutations are used more in the early stage to better explore the search space (generates some extreme individuals) and polynomial mutation is used more in the later stage to maintain the usual exploration and exploitation behavior of a generic mutation.

Mutation probability The mutation technique presented in the thesis, specially designed to deal with the optimization problem of an energy system, is a combination of three different mutations. In a single mutation process, only one mutation is applied among three mutations (REFM, CEFM and polynomial mutation) . The selection of a mutation among three mutations depends on a random number drawn from a uniform probability distribution. Figure 4.3 illustrates the overall mutation process (P_{REFM} is the probability of applying the mutation favoring renewable energy and P_{CEFM} is the probability of applying the mutation favoring conventional low cost energy sources). The probability of P_{REFM} and P_{CEFM} are equal and defined by the following formula.

$$P_{REFM} = P_{CEFM} = \frac{1 - \frac{g_c}{g_{max}}}{2} \tag{4.25}$$

Where g_c and g_{max} is the current generation number and maximum generation number of a run, respectively. At the beginning of a run, the probability of applying these two mutations are high and the probability is linearly decreased when generations pass. At the initial stage of the search, we want to explore the search space as much as possible by using two specially designed mutation operators. In the later stage of the evaluation the probability of of applying generic mutation is increased.

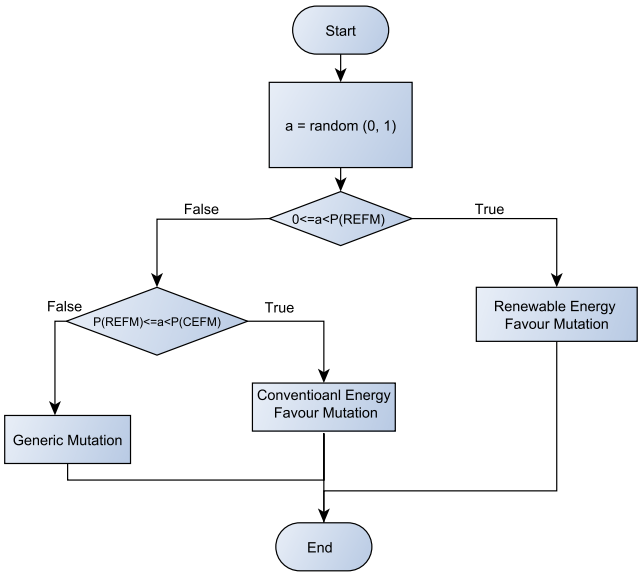


Figure 4.3: Flowchart for proposed mutation technique.

4.3 Test problem: Aalborg energy system

To study the effect of our proposed techniques to use of domain knowledge in initialization and mutation, we choose the Aalborg energy system problem [20] as a test problem [2]. The reason behind choosing this particular problem is that it

[2]Our framework can be applied to identify future optimization scenarios for other cities, states and countries. With the availability of specific domain knowledge, we conjecture that our proposed improvements can solve these problems more efficiently as well.

is extensively analyzed, therefore, all necessary data is available and described in Section 3.3.1.

We optimize five decision variables (i.e., capacities in MW of: combined heat and power (CHP), heat pump (HP), on-shore wind (ONW), off-shore wind (OFW) and photo-voltaic (PV)) in order to minimize CO_2 emissions (EM) and annual cost (AC). The lower and upper bounds for the decision variables are presented in Table 4.3. The computational cost of about three seconds per evaluation through EnergyPLAN make this a challenging real-world problem[3].

Table 4.3: Lower and upper bounds of different decision variables for Aalborg energy system problem.

DVs / Bounds	dv_{CHP}	dv_{HP}	dv_{ONW}	dv_{OFW}	dv_{PV}
Lower (MW)	0	0	0	0	0
Upper (MW)	1000	1000	1500	1500	1500

4.4 Comprehensive experiments and results

We divide our experiments into two parts. The first part reports the impact of our *smart initialization* by comparing it against the commonly used *random initialization*. In the second part, we investigate the influence of the *smart mutation* by comparing it with the commonly used *polynomial mutation*. The results of smart mutation and initialization will be reported by adopting these techniques into NSGA-II [53] and SPEA2 [180], separately.

4.4.1 General experimental settings

In this section, we present all the experimental settings that are used for all three phases of the experiments. Specific experimental settings related to a specific phase will be presented in the corresponding section.

All the proposed methodologies are implemented in jMetal [59][4]. Table 4.4 shows the general parameter settings that are used in the experiments. Additionally, we use simulated binary crossover [51], polynomial mutation [51] and binary tournament selection [51] for both algorithms. For each phase of the experiments, the algorithms are run independently 30 times to facilitate a statistical analysis of the results. Table 4.5 presents the domain knowledge used in the experiments to optimize Aalborg energy system. This knowledge is based on intuitive understanding. The first row of the table describes that the increasing capacities of CHP, HP On-, off-shore wind and PV can decrease the emissions of Aalborg energy system. In addition, the second row represents that the decreasing of the capacities can decrease the annual cost of the system.

[3]The simulation was performed on a machine having two 2.6 GHz CPUs with six cores each and 96 GB RAM.

[4]Our complete source code is available online: `https://github.com/shaikatcse/EnergyPLANDomainKnowledgeEAStep1`.

Table 4.4: General parameter settings for NSGA-II and SPEA2.

	NSGA-II	SPEA2
Population size	100	100
Archive Size	–	100
Crossover probability	0.9	0.9
Mutation probability	0.1	0.1
Distribution index	10	10
Maximum evaluations	7000	7000

Table 4.5: Domain knowledge related to different decision variables and objectives for Aalborg energy system problem.

DVs / DK_{Obj}	dv_{CHP}	dv_{HP}	dv_{ONW}	dv_{OFW}	dv_{PV}
$DK_{o^{EM}}$	true	true	true	true	true
$DK_{o^{AC}}$	false	false	false	false	false

4.4.2 Evaluation metrics

Evaluation metrics are necessary to measure the quality of the found set of trade-off solutions. To evaluate our proposed initialization method, we will use four metrics that are commonly used, namely the hypervolume (HV) [177], the inverted generational distance (IGD) [37], the additive epsilon approximation [181], and the spread [53]. The HV metric measures the volume covered by a set of solutions in the objective space with respect to a pre-defined reference point. IGD is the average distance of all the solutions in the true Pareto-front (tPF) to the nearest solution of given set of solutions. The concept of epsilon approximation is that one determines the minimum distance a found solution set in the objective space needs to be translated to in order to dominate the tPF. Deb at al. [53] introduced an indicator, called spread, to understand how well the front is distributed. Typically, a tPF (or an approximation thereof) is required to calculate the values of all the metrics except for the hypervolume. The reference point can be chosen either arbitrarily or based on a reference set: assuming we are minimizing, the coordinates of the point can be the maximum values per objective that the reference set in the objective space attains.

As the tPF of Aalborg energy system problem is not known in advance, we merge all the found sets of solutions (i.e., solution sets of 240 different individual runs), and we take only the non-dominated solutions from the merged fronts. We use this approximation of the tPF to calculate all the metric values. Please note that higher HV values indicate better results, whereas lower values for the other metrics indicate better results.

4.4.3 Influence of smart initialization

To investigate the impact of our smart initialization (SI), we will compare it with random initialization (RI) that is typically used in MOEA experiments. Random initialization is the process of selecting the value of a decision variable uniformly at random within the lower and upper bound of the decision variable. For the smart

initialization process, we set $\theta = 6.0, k = 4$ and $\beta \in [0, 1, 2]$ based on preliminary experiments.[5]

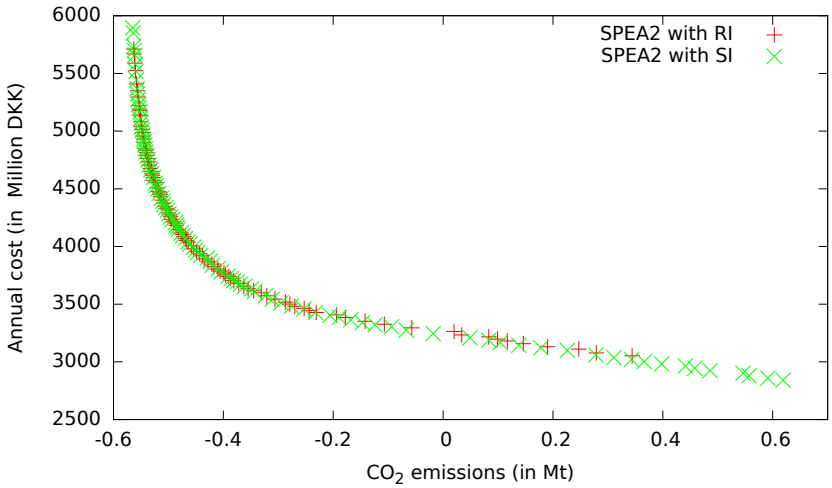


Figure 4.4: Pareto-fronts generated by SPEA2 with random initialization (RI) and with smart initialization (SI).

Figure 4.4 presents two Pareto-fronts for Aalborg energy system problem, generated by the SPEA2 algorithms; the red one is generated using RI and the green one uses SI. The X-axis represents CO_2 emissions in million tons and the Y-axis represents annual cost in million Danish Krone (DKK) for Aalborg energy system. It is noticeable that some of the solutions (i.e., scenarios) have negative emissions. This is simply because these scenarios export electricity generated by green sources from within the system to outside partners of the system. The net amount of emissions of the system is adjusted by the electricity generation mix of imported electricity. Details of this aspect can be found in the article by Lund [103].

Moreover, the green set of solutions clearly has a better spread, hence, produces more optimized energy scenarios towards the corners that can be interesting to energy planners. In addition, sometimes it produces better solutions than some of the solutions of the Pareto-front generated by using RI. However, as previously mentioned, it is required to perform statistical analyzes to understand the performance of the algorithms. Therefore, Figure 4.5 shows the results as boxplots for the four different evaluation metrics. The means and standard deviations of four metrics for two MOEAs are presented in Table 4.6. The dark gray shade indicates better results. It is very clear from the figure and the table that SI outperforms RI on all the metrics for the two algorithms. Moreover, we perform Mann-Whitney U-tests [113][6] on all the metric values of the different runs. The test is performed to test the null hypothesis against an alternative hypothesis to determine whether two samples come from same population or not. In our context, we want to reject the null hypothesis, as the evaluation metric values of applying SI and RI should be significantly different to each other. We consider that the null hypothesis will be rejected if the corresponding p-value is less than 0.05. Table 4.7 presents p-values

[5]We conjecture that additional performance gains are possible, however, a tuning of these parameters is beyond the scope of this thesis.

[6]A non-parametric statistical test, also known as Wilcoxon rank-sum test.

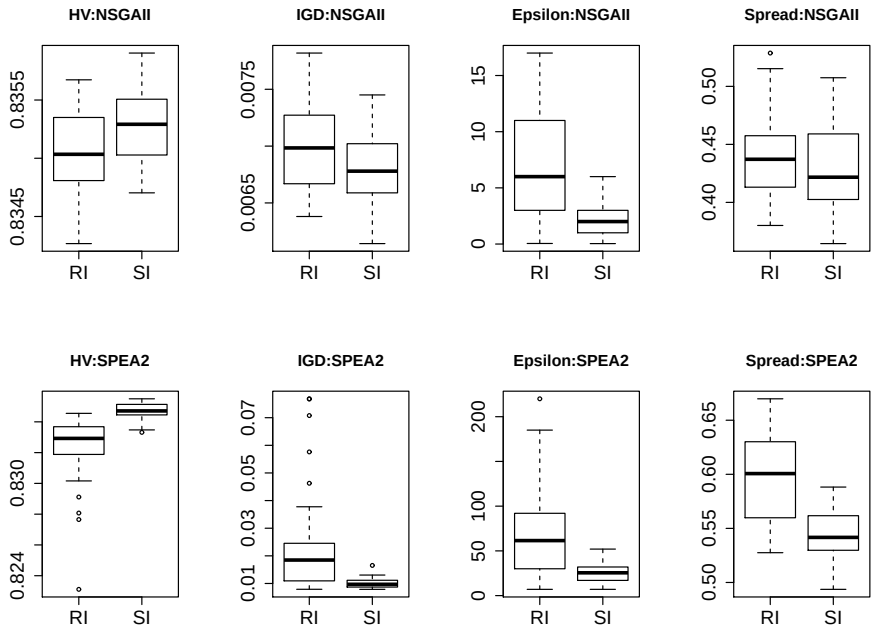


Figure 4.5: Boxplots for four metrics comparing random initialization (RI) with smart initialization (SI) in NSGA-II and SPEA2.

for all the metrics of the two different MOEAs. All the p-values except *spread and IGD for NSGA-II* are less than 0.05. The test results shows that there are statistically significant differences for the metrics (except spread for NSGA-II) when SI is used instead of RI. In summary, for most of the metrics, significant amounts of improvements are achieved by using our proposed smart initialization technique.

Table 4.6: Mean and standard deviation (in subscript) for different quality indicators.

		Mean and standard deviation	
Algorithm	Evaluation metrics	Random Initialization	Smart Initialization
NSGA-II	HV	$8.35e-01_{3.6e-04}$	$8.35e-01_{2.9e-04}$
	IGD	$7.01e-03_{4.0e-04}$	$6.81e-03_{3.5e-04}$
	Epsilon	$6.70e+00_{4.8e+00}$	$2.37e+00_{1.5e+00}$
	Spread	$4.38e-01_{3.6e-02}$	$4.29e-01_{3.5e-02}$
SPEA2	HV	$8.32e-01_{2.4e-03}$	$8.35e-01_{5.1e-04}$
	IGD	$2.51e-02_{2.0e-02}$	$1.00e-02_{1.8e-03}$
	Epsilon	$7.06e+01_{5.5e+01}$	$2.63e+01_{1.2e+01}$
	Spread	$5.98e-01_{4.2e-02}$	$5.45e-01_{2.3e-02}$

While it is necessary to compare the final solutions, we are also interested in the actual effect that SI has on the optimization. As SPEA2 benefits more than NSGA-II from the use of SI, we are showing the development of the indicator values over time in Figure 4.6. As we can see, our smart initialization strategy results in significantly better starting points for the optimization process than random initialization does. In addition, SI also appears to provide better populations for the subsequent optimization, as the progress over time is "steeper" in comparable parts of the optimization. For example in the case of additive epsilon approximation, it takes six generations to improve from approximately 380 to 200 when SI is used, whereas twice as much time is needed when random initialization is used. Similar

Table 4.7: Mann-Whitney U-tests: p-values for different metrics when comparing smart initialization (SI) with the common random initialization (RI).

	p-values	
Evaluation metrics	Compare NSGA-II: With SI and RI	Compare SPEA2: With SI and RI
HV	0.02247	$9.083e^{-12}$
IGD	0.087783	$2.2e^{-16}$
Epsilon	$5.92e^{-05}$	$1.886e^{-04}$
Spread	0.3986	$2.841e^{-07}$

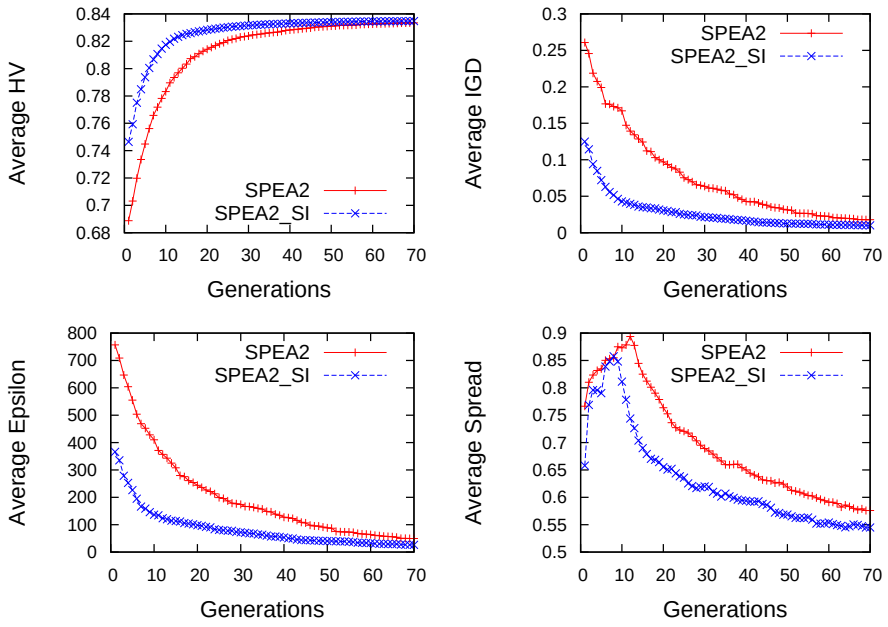


Figure 4.6: Development of the indicators during the optimisation runs for SPEA2.

trends can be observed for development of the other indicator values. We conjecture that the bias of the initial populations as generated by SI has favorable properties that are exploited in the subsequent optimization, whereas RI with its uniform initialization is not problem-specific.

Result of the alternate approach for diversity maximization In this experiment, we want to compare the optimization results when the default and an alternate technique (described in Section 4.2.2.2) are used for initialization. Table 4.8 presents the parameters related to initialization phase. First row shows the same parameter settings presented before (i.e., last experiment of smart initialization). Second and third rows show the parameter settings for the alternate technique. Please note that the parameters for first and second row are identical. Only the reducing initial population approaches are different. However, the last row shows different parameter settings. Please note that with the parameter settings of last row, a very large number of initial individuals is produced (i.e., in total 10^6 number of initial individuals are produced). It is practically impossible to use the default approach with this parameters because of computation time. However, with the help of alternate approach, it is possible to reduce that large number of individuals in several seconds.

Table 4.8: Parameters settings for examine alternate approach.

Experiment name	Reducing approach	θ	k	β
SI_d	Default approach	6.0	4	$[0, 1, 2]$
SI_{A1}	Alternate approach	6.0	4	$[0, 1, 2]$
SI_{A2}	Alternate approach	6.0	5	$[0, 1, 2, \ldots, 8, 9]$

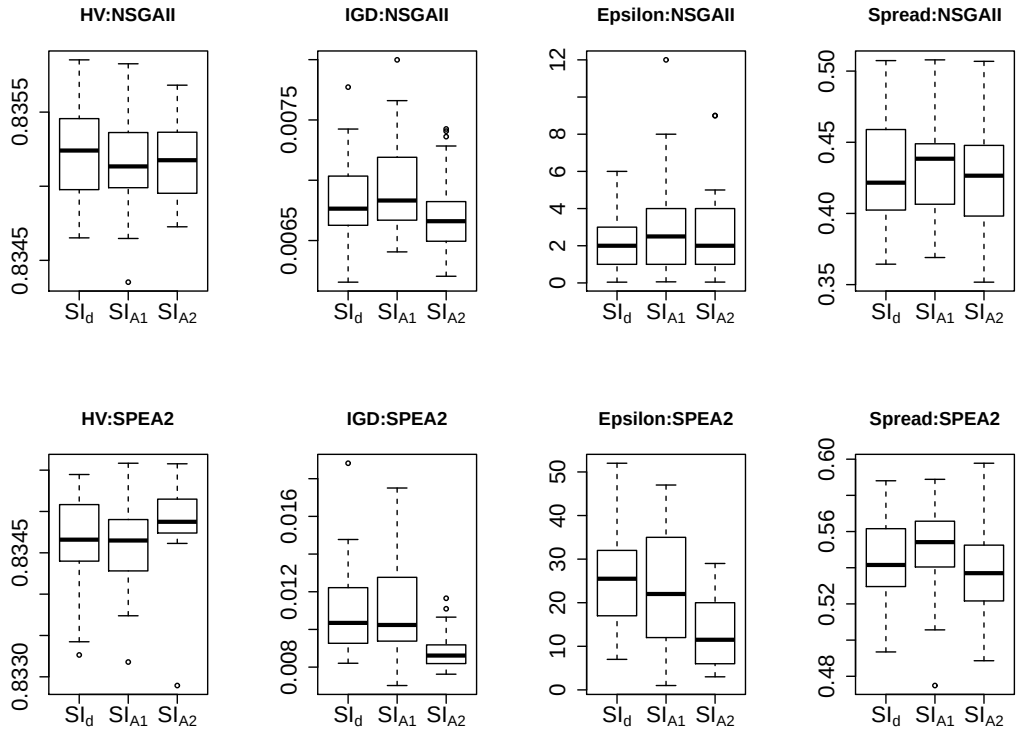


Figure 4.7: Boxplots for four metrics comparing smart initialization with the default (SI_d) and the alternate approach with different parameters' settings (SI_{A1}, SI_{A2}) for NSGA-II and SPEA2.

Figure 4.7 presents a comparison of the results obtained by using the default approach and the alternate approach with two different parameters settings reported in Table 4.8. In addition, Table 4.9 presents the results of Mann-Whitney U-tests. When comparing SI_d and SI_{A1} for NSGAII and SPEA2, SI_d performs slightly better, however, the resultant samples are not statistically different (from Table 4.9). Therefore, it can be concluded that alternate approach with identical parameter has very similar results. When comparing SI_d with SI_{A2} for NSGAII, SI_{A2} perform marginally better. Still there is no statistically significant differences between the results. Only, the significant difference is observed for IGD and epsilon for SPEA2. By considering the results, it can be concluded that the alternate approach produce similar results of the default approach. However, the default approach produces minor improved results. Therefore, it is recommend to use the default approach when the computational cost is under manageable range. Alternate approach should be employed when a problem has a large number of decision variables.

Table 4.9: Mann-Whitney U-tests: p-values for different metrics when comparing alternate approach with default approach.

Evaluation metrics	p-values			
	NSGAII		SPEA2	
	SI_d, SI_{A1}	SI_d, SI_{A2}	SI_d, SI_{A1}	SI_d, SI_{A2}
HV	0.4147	0.3898	0.5229	0.07228
IGD	0.2861	0.7412	0.9824	5.833e-05
Epsilon	0.2843	0.5033	0.433	2.652e-05
Spread	0.6865	0.592	0.1774	0.2601

4.4.4 Influence of smart mutation

In this section, we investigate the impact of smart mutation on the Aalborg energy system problem. We compare polynomial mutation (PM) with our smart mutation (SM) using the same evaluation metrics as above for NSGA-II and SPEA2 as the underlying MOEAs. For each MOEA, the same initial population is used for each run to ensure a fair comparison between PM and SM. By using the same initial populations we cut down any advantage gain by a MOEA in the initial phase. Therefore, the settings ensure that if any MOEA performs better than other, then this is strictly due to the use of specific mutation.

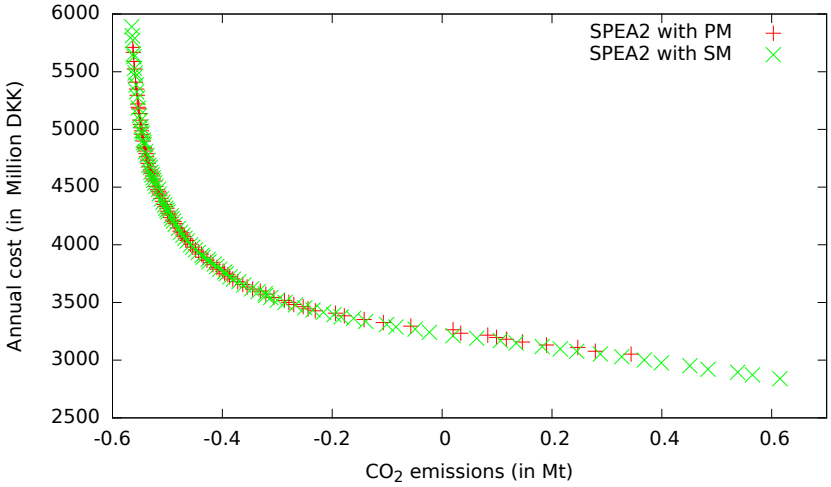


Figure 4.8: Pareto-fronts generated by SPEA2 with polynomial mutation (PM) and with smart mutation (SM).

Figure 4.8 shows two Pareto-fronts generated by SPEA2 using PM and SM. The Pareto-front with SM has better spread and produces better solutions. Figure 4.9 presents the comparison of PM and SM as boxplots. For all the metrics, SM performs better than PM. However, the statistical test (Table 4.10) shows that not all the metrics are not statistically significantly different for NSGA-II. Nevertheless, significant improvement is achieved for SPEA2. From the boxplots and the statistical test, it can be concluded that SM provides a good performance improvement for both MOEAs.

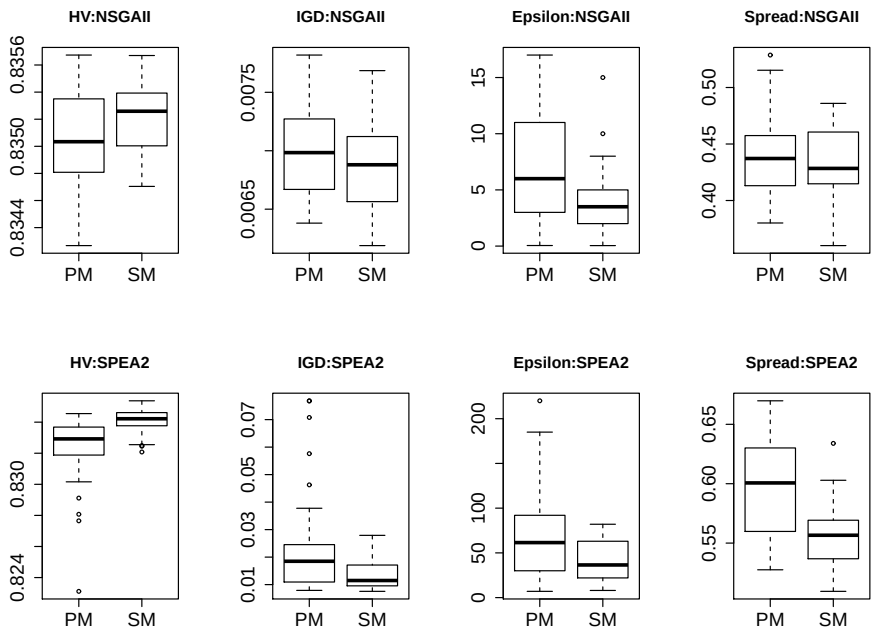


Figure 4.9: Boxplots for four metrics comparing PM with SM for NSGA-II and SPEA2.

Table 4.10: Mann-Whitney U-tests: p-values for different metrics when comparing our smart mutation (SM) with the common polynomial mutation (PM).

	p-value	
Evaluation metrics	Compare NSGA-II: With SM and PM	Compare SPEA2: With SM and PM
HV	0.0906	$1.51e^{-05}$
IGD	0.2539	$1.20e^{-16}$
Epsilon	0.01664	0.07721
Spread	0.7082	$1.34e^{-04}$

4.5 Conclusion

To accurately plan a system with an introduction of more renewable energies, complex and computationally costly simulations are typically used to assess configurations according to different objectives, for example, based on their cost and their emissions.

General purpose multi-objective evolutionary algorithms are often used to solve such problems, however, the simulation cost result in time-consuming optimizations. In this chapter, we present different techniques to improve solution quality of an optimization. First and foremost, we incorporate basic domain knowledge about energy systems into different operators of such algorithms in order to increase the solution quality.

Our results on the Danish Aalborg energy system problem reveal that our problem-specific approaches achieve significant improvements over generic state of the art approaches. It is noteworthy that this was achieved with rather basic domain knowledge. It remains to be seen how much solution quality can be improved further by

using more detailed knowledge.

As we see the function evaluations for optimizing an energy system is computationally costly, next chapter will be dedicated to the development of a robust stopping criterion. The goal is to develop a technique that stops a MOEA automatically when convergence is detected.

Chapter 5

Development of a robust stopping criterion

You don't understand anything until you learn it more than one way.

— Marvin Minsky

5.1 Introduction

Virtually all the engineering fields have to deal with optimization problems. Most practical problems encounter multiple conflicting objectives; therefore, the problems need to be considered as multi-objective optimization problems. Recently, multi-objective evolutionary algorithms (MOEA) become a major technique for optimizing real-world optimization problems [171]. A MOEA should find a Pareto-front by using minimum computational cost. Since function evaluations (FE) of most of the real-world optimization problems are costly (e.g., function evaluation of energy system optimization take around 3 seconds), FE occupy the lion's share of total computational cost of an optimization algorithm. Therefore, finding an appropriate stopping criterion is an important task to minimize computational cost. To save wasteful function evaluations, it is necessary to spot the stagnation or convergence of an algorithm. Almost all MOEAs are stopped after certain number of function evaluations [171]. However, specifying this number for a practical problem without any prior knowledge, is quite difficult. Therefore, in last few years some techniques for detecting convergence of a MOEA have been proposed [171, 164, 30, 76, 147, 116].

All the proposed techniques are based on investigating an objective space of a problem. However, our approach is based on simultaneous monitoring of an objective and a decision space. Investigating a decision space with an objective space is

M. S. Mahbub, T. Wagner, and L. Crema, "Improving robustness of stopping multi-objective evolutionary algorithms by simultaneously monitoring objective and decision space," in *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, GECCO '15, ACM, 2015, pp. 711–718. DOI: `http://doi.acm.org/10.1145/2739480.2754680`
The text is substantially verbatim apart from the introduction and the conclusion and mentioned additions in chapter 5. Section 5.3 is modified for better explanation compared to published version.

important to efficiently detect the convergence of a MOEA. Simultaneous stabilization of indicators on the two spaces provides us more robust[1] stopping criterion as it is shown later that an algorithm runs longer when a decision space indicator is included with an objective space indicator. The approach has another major benefit. It does not require any thresholding parameters that require prior knowledge about a problem. Finally, the convergence is detected by analyzing the trends of the two indicators. A two-sided t-test on the slop coefficients is used to detect the stagnation of linear trends.

In this chapter, our developed method is tested on the theoretical multi-objective test problems found in literature against state of the art techniques. The results show that our method perform reliably than other techniques. Finally, the capability of the stopping methods on real-world optimization problem is demonstrated in the next chapter.

The remainder of the chapter is organized as follows. In section 5.2, we briefly describe the state of the art techniques of stopping criteria of MOEAs. Section 5.3 presents the details of our proposed method. Experimental results and corresponding discussion are reported in section 5.4. Conclusion and future work will be described in section 5.5.

5.2 State of the art

Deb and Jain [52] was first to propose two metrics, *convergence* and *diversity*, that can be monitored online. The stopping decision of a MOEA depends on the visual inspection of the metrics. At that time, no automatic convergence method was proposed. Inspired from the technique of a stopping criterion of a single objective evolutionary algorithm (EA), Redenko and Schoenauer [147] proposed a stability measure. The measurement is based on density of non-dominated solutions. By studying the dynamics of NSGA-II [53], the authors experimentally showed that the algorithm converge when maximum crowding distance [53] is reached. The user needs to provide a threshold and the algorithm stops when the standard deviation of maximum crowding distance falls below the threshold for pre-defined number of generations.

Martí et al. [116] proposed a stopping method called MGBM criterion. The authors proposed an indicator named *mutual dominance rate (MDR)* that is basically a measurement of how many non-dominate solutions of one generation dominate the non-dominate solutions of consecutive generation. They used simplified Kalman filer to gather evidences about when to stop. A MOEA is stopped when a-posteriori estimation of MDR falls below a pre-defined threshold.

Goel and Stander [76] uses an external achieve to propose a new indicator, named *consolidation ratio(CR)*. The archive keeps non-dominated solutions and in each generation the archive is updated. Additionally, CR is a ratio between the number of solutions of previous generation that are still present in archive and size of the archive. Two different stopping criteria are proposed: *fixed threshold approach* and *utility-function based approach*. In fixed threshold approach, a MOEA is stopped when the CR falls below than the pre-defined threshold. For the next approach, the

[1] In this context of stopping criteria of a MOEA, robust means the consistency of stopping decision when dealing with stochastic nature of a MOEA.

utility of evolving extra generation is calculated and an algorithm terminates when the utility falls below the utility threshold.

Bui et al. [30] suggest a stability measure named *dominance-based quality (DQ)* based on dominance relation of a solution and neighbouring solutions. The basic idea is that the number of non-dominated neighbouring solutions will be decreased over time as a solution moves towards a Pareto-front. To measure the local dominance of a solution, the authors use additional function evaluations and a Monte Carlo simulation approach. The authors do not provide any suggestions about the convergence value of DQ. However, it is obvious that DQ = 0 is a powerful stopping criterion and with the help of visual inspection, the stagnation of an algorithm can be identified.

Trautmann et al. [165] and Wagner et al. [171] apply a statistical approach to solve the problem. In [171], the authors propose two different statistical tests on three performance indicators (PI) (i.e., hypervolume, R2 and additive epsilon) to detect convergence. One of the test is a *one-sided χ^2-variance test* for measuring the significance decrease of the variances of PIs than pre-defined thresholds. Another proposed test is a *two-sided t-test* for detecting stagnation by analyzing slope coefficient of different PI trends. A MOEA is stopped when any of the two tests are able to detect convergence (i.e., p-value lower/above the critical level depending on the test).

Roche et al. propose a stopping criterion based on the detection of loss of population diversity in the decision space [146]. The authors propose a formula to determine the time, when the diversity falls below a threshold. The authors also propose a statistical procedure to confidently stop an EA. However, without prior knowledge about decision space of a problem, it is very difficult to set the threshold. The method is validated only on single objective problems. In contrast, as we are dealing with MOEAs (population consists of a set of solutions covering a Pareto-front, not converging towards a single solution), the population always has certain diversity. Hence, in our context, it is more important to detect the moment in time, when the solutions covering decision space are no longer moved. Therefore, it is good to detect stagnation of the diversity value instead of the value falling below a certain threshold.

5.3 Proposed methodology

Most of the studies we have discussed above require an objective space parameter falls below user-defined parameters to stop. Additionally, some recent studies also address the stagnation of more than one objective space indicators using statistical tests. In contrast to these approaches, our proposed method is based on simultaneous monitoring of two metrics; one on objective space and another on decision space. To get a better stopping/convergence method, our primary assumption is that these two spaces (i.e., objective and decision) should be stabilized concurrently. Considering stagnation of some indicators on objective space (e.g., hypervolume [177], epsilon [181] and so on) does not mean that an algorithm could not improve later. The individuals of a population may contain enough diversity to generate better children in later generations. Therefore, by concurrently monitoring of two spaces, more robust stopping criterion could be developed. *Average Hausdorff distance* [154] and *diversity* metrics are used to monitor an objective and a decision space, respectively.

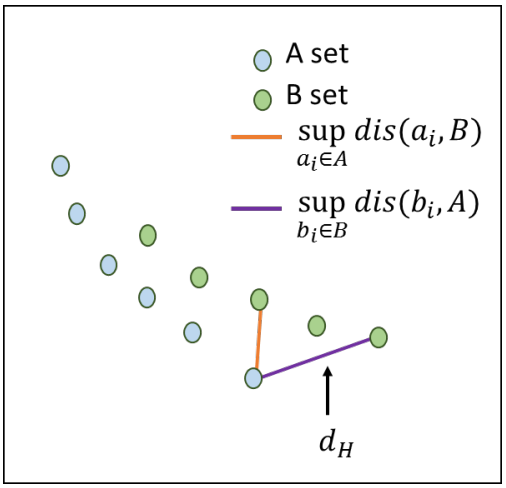


Figure 5.1: Example of Hausdorff distance.

In the field of MOEA, diversity usually refers population diversity in an objective space; however in our approach, we use the term *diversity* to refer decision space diversity.

Average Hausdorff distance Classical Hausdorff distance is a widely used metric to measure the distance between two sets. Let $A = \{a_1, a_2 \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_m\}$ are two finite sets. Hausdorff distance (d_H) is defied as:

$$d_H(A, B) = max(sup_{a_i \in A} dis(a_i, B), sup_{b_i \in B}(b_i, A)) \tag{5.1}$$

Where $dis(a_i, B) = \inf_{b_i \in B} \|a_i - b_i\|$ and $\|.\|$ refers norm. An example of Hausdorff distance is depicted in Figure 5.1. However, the original version of Hausdorff distance is not a suitable metric for measuring the convergence of multi-objective optimization problem [154]. The main reason is that a set with an outlier is largely penalized that does not compatible with the stochastic nature of MOEAs.

Therefore, Average Hausdorff distance (AHD) is proposed by Schutze et al. in [154]. The metric is a modified version of Hausdorff distance adopted for multi-objective optimization problem by applying generational distance (GD) and inverted generational distance (IGD). Let $A = \{a_1, a_2 \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_m\}$ are two finite sets. GD is defined as follows:

$$GD(A, B) = \frac{1}{n}\left(\sum_{i=1}^{n} dis(a_i, B)^p\right)^{\frac{1}{p}} \tag{5.2}$$

Where $dis(a_i, B) = \inf_{b_i \in B} \|a_i - b_i\|$ and $\|.\|$ refers norm. The indicator suffers a problem when a candidate set could have many similar solutions, hence reducing the GD value. Therefore, it becomes difficult to compare the set with different number of solutions. To avoid the unwanted problem, the authors of [154] propose a slightly modified version of GD by talking power mean of average distances.

$$GD(A, B)_p = \left(\frac{1}{n}\sum_{i=1}^{n} dis(a_i, B)^p\right)^{\frac{1}{p}} \tag{5.3}$$

The same approach is applied on IGD [154].

$$IGD(A,B)_p = \left(\frac{1}{m}\sum_{i=1}^{m} dis(b_i,A)^p\right)^{\frac{1}{p}} \tag{5.4}$$

Finally, the average Hausdorff distance is defined as follows:

$$AHD_p(A,B) = max(GD_p(A,B), IGD_p(A,B)) \tag{5.5}$$

The metric is particularly interested to us, since the metric considers the averaged distances between between two sets without penalized by outliers.

Diversity There are many possible techniques to measure genetic diversity of a population. The most popular two methods are Hamming distances among all pairs of chromosomes and Shannon entropy on gene frequencies [173]. The authors of [173] argues that the underlying mechanism of all the diversity measurements is very similar. Fundamentally, population diversity is a measurement of how individuals of a population are different from one to another.

Considering P is a population of n individuals and $\{x_{i,1}, x_{i,2}, \ldots, x_{i,l}\}$ are l numbers of chromosomes (decision variables) of an arbitrary i$^{\text{th}}$ individual. The diversity for a chromosome (k) is the all possible distances among all the individuals of the population for the chromosome and can be written in the following way [173]:

$$Div_k^2(P) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} D^2(x_{i,k}, x_{j,k}) \tag{5.6}$$

where D^2 is Euclidean distance based on L_2 norm. To get all possible pairs diversity for all the chromosomes, the Equation (5.6) becomes

$$Div^2(P) = \frac{1}{2}\sum_{k=1}^{l}\sum_{i=1}^{n}\sum_{j=1}^{n} D^2(x_{i,k}, x_{j,k}) \tag{5.7}$$

By using some algebraic manipulation, we get the diversity in the following form:

$$Div^2(P) = n^2\sum_{k=1}^{l}\left(\overline{x_k^2} - \overline{x_k}^2\right) \tag{5.8}$$

Where $\overline{x_k} = \frac{1}{n}\sum_{i=1}^{n} x_{i,k}$ and $\overline{x_k^2} = \frac{1}{n}\sum_{i=1}^{n} x_{i,k}^2$. Finally, since diversity of a population should not be increased by adding same individual and same chromosome over and over again, it is defined as follows [173]:

$$Div(P) = \frac{1}{l}\sqrt{\sum_{k=1}^{l}\left(\overline{x_k^2} - \overline{x_k}^2\right)} \tag{5.9}$$

One can wonder why do not we use the Solow and Polasky diversity metric that was proposed in previous chapter. Please note that our primary goal is to detect the movements of individuals in decision space though the metric. Movements can be detected by any diversity metrics, therefore, in principle Solow and Polasky metric

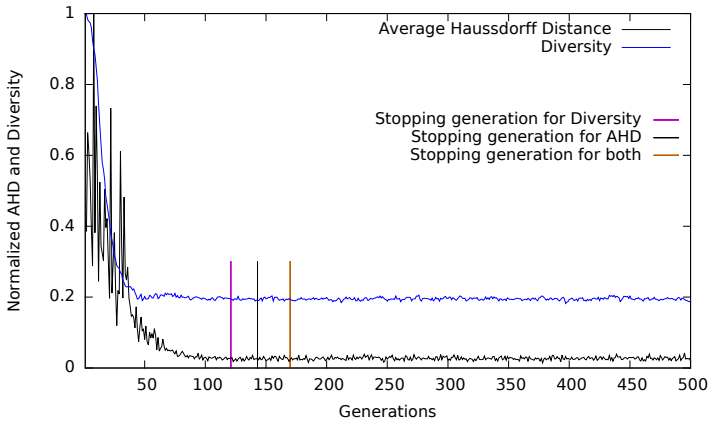


Figure 5.2: Normalized AHD and diversity are represented by black and blue trends. Three vertical lines with three different colors indicate the stopping generations for three different criteria.

could be used. However, as we see in the last chapter that Solow and Polasky metric is computationally expensive; the proposed diversity metric is chosen because of low computational cost.

Figure 5.2 illustrates the values of two metrics for ZDT1 [179] problem solved by NSGA-II. In the figure, the values for AHD and diversity for each generation is normalized to illustrate two metrics in the same plot. The figure shows that after some time, each metric is stabilized. There are still some random fluctuation, however no definite increasing or decreasing trends can be found.

One can argue that the AHD is still not zero, therefore there is still a movement in objective space. The movement of individuals are still there, however, it is not so significant. The reason for these insignificant movement are explained by diversity controlling mechanism of MOEAs. Most classical MOEAs such as NSGA-II [53], SPEA2 [180]; have a mechanism to maintain diversity in the objective space (e.g., crowding distance, archive truncation). When the trend of AHD remains almost stable, it means that some individuals are removed from a crowding region and some are added in less crowding region by objective space diversity controlling mechanism of a MOEA. In addition, in this stage of an algorithm, this process continues repeating. Figure 5.3 illustrates the Pareto-fronts of 140^{th} and 141^{th} generation of same problem (i.e., ZDT1) solved by NSGA-II. The region covered by a rectangle illustrates the behavior discussed before. A zoomed view of the region is illustrated in the upper part of the figure. By carefully examine the area, one can see deletion of an individual of 140^{th} generation marked by '+' and addition of an individual marked by '□' for 141^{th} generation. Therefore, if this kind of stable trend (Figure 5.2) continue for long time, we conclude that significant improvement may not be possible. On the other hand, when the diversity becomes nearly stable, individuals in the decision space become stabilized.

To detect the trends of stability, a regression analysis is performed. More specifically, a two-sided t-test is carried out to check the significance of decreasing linear trend [171]. Using the test, it is possible to measure the slope (β) of some points using least-square method [171]. Moreover, a statistical hypothesis test (i.e., t-test) is carried out to determine the significance of error on β. As we want to detect the stability of the trends (i.e., $\beta = 0$), the statement for the hypothesis test can be

expressed as:

$$H_0 : \beta = 0 \text{ vs } H_1 : \beta \neq 0 \tag{5.10}$$

Consequently, we try to find out the generation where the null hypothesis can not be rejected anymore. Figure 5.2 shows the stopping generations (three vertical lines) when the test is applied individually and concurrently on different metric (i.e., AHD and diversity). It is clear from preliminary results that each metric stabilizes in different time. Moreover, by considering two metric simultaneously, the algorithm stops later. Hence, it is possible to detect convergence more robustly by monitoring an objective and a decision space concurrently.

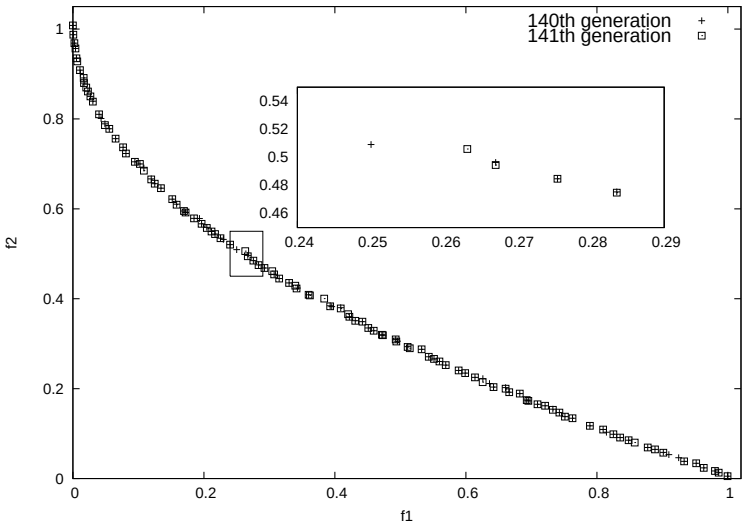


Figure 5.3: Pareto-fronts extracted by NSGA-II for two consecutive generations.

Algorithm 5 Algorithm for detecting stopping generation

Require:

α ▷ Significance level for statistical test
$nGenLT$ ▷ Number of previous generations considered for calculating slop of linear regression model
$nGenUnCh$ ▷ Number of previous generations the p-values remain unchanged ($pVal_j > \alpha, j \in \{AHD, Div\}$)
$MaxGen$ ▷ Maximum number of allowed generations
$i = 0$
$pAHD[] \leftarrow 0, pDiv[] \leftarrow 0$
repeat
  $i \leftarrow i + 1$
  $AHD[i] \leftarrow$ AHD between Pareto-front of i^{th} and $(i-1)^{th}$ generation
  $Div[i] \leftarrow$ Diversity of i^{th} generation
  if $i > nGenLT$ **then**
    $pAHD[i] \leftarrow pValue$ for previous $nGenLT$ of AHD $(\forall_z \in \{i, i-1, \ldots, i-nGenLT\} : AHD[z])$
    $pDiv[i] \leftarrow pValue$ for previous $nGenLT$ of Div $(\forall_z \in \{i, i-1, \ldots, i-nGenLT\} : Div[z])$
  end if
until $i \geq MaxGen$ **or** $(\forall_j \in \{i, i-1, \ldots, i-nGenUnCh\} : pAHD[j] > \alpha \wedge pDiv[j] > \alpha\ )$

Algorithm 5 presents the details of the proposed method. There are four intuitive parameters that need to be specified. The user has to provide a significance level (α) for the statistical test. Mainly two values are found: 0.05 (standard) and 0.1 (conservative) [171]. A user needs to specify the number of previous generations ($nGenLT$), for which AHD and *diversity* values are considered in order to estimate the slop of the liner regression model. The next parameter ($nGenUnCh$) is the number of preceding generations for which no significant improvement can be obtained

(H_0 can not be rejected, $p - value > \alpha$). Finally, the last parameter ($MaxGen$) is the number of maximum allowable generations. AHD and diversity are calculated for each generation (step #9 and #10). After first $nGenLT$ generations, p-values for AHD and diversity are calculated (step #12 and #13). To calculate p-values for corresponding metric previous $nGenLT$ metric values are considered. A MOEA will be stopped when either the algorithm already evolved for $MaxGen$ generations, or the p-values for both metrics (AHD and *diversity*) are larger than α for previous $nGenUnCh$ generations (step #15).

5.4 Experiments and results

A number of experiments have been conducted to analyze the performance of proposed MOEA stopping criterion. For the experiments, we have selected two most widely used MOEAs (i.e., NSGA-II [53] and SPEA2 [180]) and six benchmark problems (***ZDT1, ZDT2, ZDT3, ZDT4, DTLZ2, DTLZ5***) [54, 179] depending on the different characteristics of solution Pareto-fronts. All the problems of ZDT family are bi-objective problems and DTLZ2, DTLZ5 are three objectives problems. Our proposed approach will be compared with individual monitoring of each space. Additionally, it will be compared with state-of-the-art stopping criteria, called online convergence detection (OCD) [171]. Finally, a comparison with standard budget FE recommendation, will be presented.

5.4.1 Experimental settings

The proposed methodology is implemented in jMetal [59], Table 5.1 shows the standard parameter settings for our experiments. Moreover, we have used simulated binary crossover [51], polynomial mutation [51] and binary tournament selection [51] for both of the algorithms. Additionally, we set $nGenLT = 30$, $nGenUnCH = 10$ and $\alpha = 0.05$ for the experiments. $p = 2$ is used to calculate AHD. For each problem, the algorithms independently run *30* times. Moreover, for each run, we let the algorithms evolve *500* generations. The reason of letting the algorithms to run more than standard $MaxGen$ (Table 5.2) is that we want to investigate when exactly the proposed stopping method is activated. Finally, we have used the following parameters in the variance test of OCD for the three indicators: $\epsilon_{HV} = 1e^{-4}$, $\epsilon_R = 1e^{-6}$ and $\epsilon_{Epsilon} = 2e^{-4}$.

Table 5.1: Parameter settings for NSGA-II and SPEA2.

	NSGA-II	SPEA2
Population size	100	100
Archive Size	–	100
Crossover probability	0.9	0.9
Mutation Probability	1/NDV [a]	1/NDV[a]
Distribution Index	20	20

[a] Number of decision variables

5.4.2 Evaluation metrics

To evaluate our proposed method, we have defined four metrics/indicators. The first indicator is the difference between hypervolume (HV) [177] achieved by different methods ($HV_i : i \in \{AHD, Div, AHD + Div, OCD\}$ and HV achieved by standard $MaxGen$ (i.e., $HV_i - HV_{std}$), we call the difference, HV_d. In the same manner differences between epsilons [181] ($eps_d = eps_{std} - eps_i : i \in \{AHD, Div, AHD+Div, OCD\}$) will be computed. The Pareto reference sets provided by jMetal framework [2] are used to calculate epsilon. Negative HV_d value means smaller HV is achieved when an algorithm stops by a method than standard $MaxGen$ and vice versa. Similarly, negative eps_d means epsilon is larger when an algorithm stops by a method than standard $MaxGen$ and vice versa. Therefore, HV_d and eps_d values closer to 0 indicate better approximation of a Pareto-front. The standard stopping generation for different problems can be found in Table 5.2. Additionally, we want to study the percentage of HV of true Pareto-front [122]; obtained by an algorithm when the algorithm stops by the combined approach ($AHD + Div$). The calculation is easily done by taking percentage between HV achieved by an algorithm and HV of true Pareto-front(i.e., $HV_{per} = \frac{HV_{ADH+Div}}{HV_{tPF}}$). We use Pareto-reference sets provided by jMetal framework as true Pareto-front. We will also report average number of saved function evaluations ($NoSFE$) for each problem. The $NoSFE$ will be calculated by $(MaxGen - Gen_{AHD+Div}) * PS$, where $Gen_{AHD+Div}$ is the stopping generation number by combined method; PS is the population size.

Table 5.2: Standard Maximum allowed number of generations [171], [54].

Problem	ZDT family	DTLZ2	DTLZ5
Standard *MaxGen*	200	300	200

5.4.3 Discussion

Figure 5.4 and 5.5 present comparisons of the four different stopping approaches with respect to stop generation (StopGen), HV_d and eps_d on the different test scenarios. Each row of the figures presents a problem and each column presents the performance with regard to a particular indicator. Horizontal lines in the plots of the first column indicate the standard $MaxGen$. Moreover, horizontal lines in the plots of the second and third columns are drawn to indicate the zero level for the corresponding indicators.

It is clear from the figures that for all the problems and algorithms, the combined approach runs longer than the individual approaches. For example, by considering *AHD* metric alone for ZDT4 problem or by considering only *diversity* metric for ZDT1 problem for SPEA2, the algorithm stops prematurely. It is interesting to point out that there is no clear winner between AHD and Div. Therefore, relying on a particular space may detect converge prematurely. This clarifies the fact that the simultaneous monitoring of both spaces is more reliable to detect the convergence of an algorithm. Finally, HV_d and eps_d values get closer to the horizontal lines (with respect to individual monitoring approach, except OCD) also ensure that

[2] `http://jmetal.sourceforge.net/problems.html`

Table 5.3: Mean and standard deviation for parameters for different problems.

Problem	Evaluation metrics	NSGA-II mean	NSGA-II SD	SPEA2 mean	SPEA2 SD
ZDT1	NoSFE	5527	2382	643	1153
	HV_d	-5.89E-03	3.54E-03	-6.24E-04	1.19E-03
	eps_d	-3.06E-03	3.33E-03	-5.40E-04	1.21E-03
	HV_{per}	0.979285	0.005466	0.986092	0.001775
ZDT2	NoSFE	2793	1887	0	0
	HV_d	-2.68E-03	2.04E-03	0.00E+00	0.00E+00
	eps_d	-1.63E-03	2.38E-03	0.00E+00	0.00E+00
	HV_{per}	0.966308	0.006051	0.957989	0.042851
ZDT3	NoSFE	5077	2427	1427	2509
	HV_d	-4.36E-03	2.88E-03	-3.46E-03	1.40E-02
	eps_d	-4.71E-03	4.35E-03	-5.37E-03	2.38E-02
	HV_{per}	0.984483	0.00566	0.983046	0.027052
ZDT4	NoSFE	777	1864	953	2582
	HV_d	-2.30E-02	6.62E-02	-4.03E-02	1.29E-01
	eps_d	-2.04E-02	6.17E-02	-7.41E-02	2.42E-01
	HV_{per}	0.925081	0.109336	0.874061	0.185557
DTLZ2	NoSFE	18323	2410	15173	6152
	HV_d	-3.13E-03	7.17E-03	-3.85E-03	4.81E-03
	eps_d	3.52E-03	3.21E-02	-2.45E-03	1.18E-02
	HV_{per}	0.790937	0.011261	0.857519	0.008931
DTLZ5	NoSFE	3113	3467	2767	4192
	HV_d	-8.90E-05	2.66E-04	-4.36E-04	7.50E-04
	eps_d	-2.66E-04	1.83E-03	-6.47E-04	1.58E-03
	HV_{per}	0.969181	0.003113	0.968569	0.00784

the combined approach has a better performance with regard to the Pareto-front approximation.

The figures also illustrate that the algorithms run much longer for OCD than for the combined approach on all ZDT problems. In addition, the algorithms run longer than standard MaxGen. Consequently, the HV_d and eps_d values pass the horizontal lines having positive values implies that the algorithms with OCD better approximate Pareto-fronts than the algorithm with standard stopping criteria (standard MaxGen). However, it is straight forward to understand because the algorithms that run longer generations would have better approximation of Pareto-fronts. In contrast, on DTLZ problems, the algorithms with OCD stop earlier than our approach. However, using our method the algorithms stop more reliably with well-approximated Pareto-fronts. Therefore, the simultaneous monitoring approach is more consistent, whereas OCD is not so stable when confronted with different problems.

In this paragraph, we will compare our method ($AHD + Div$) with standard budget recommendations (that means MaGen values are set according to Table 5.2). Table 5.3 reports the means and standard deviations of the respective indicator values. We have found very well approximated Pareto-fronts for all problems and for both algorithms, except on *ZDT4*. The means and standard deviations of HV_d and eps_d are very low for all these problems.The good quality Pareto-fronts are found while also obtaining reasonably good *NoSEF*s. The only exception is *ZDT2* for SPEA2, where the proposed method does not converge before $MaxGen$

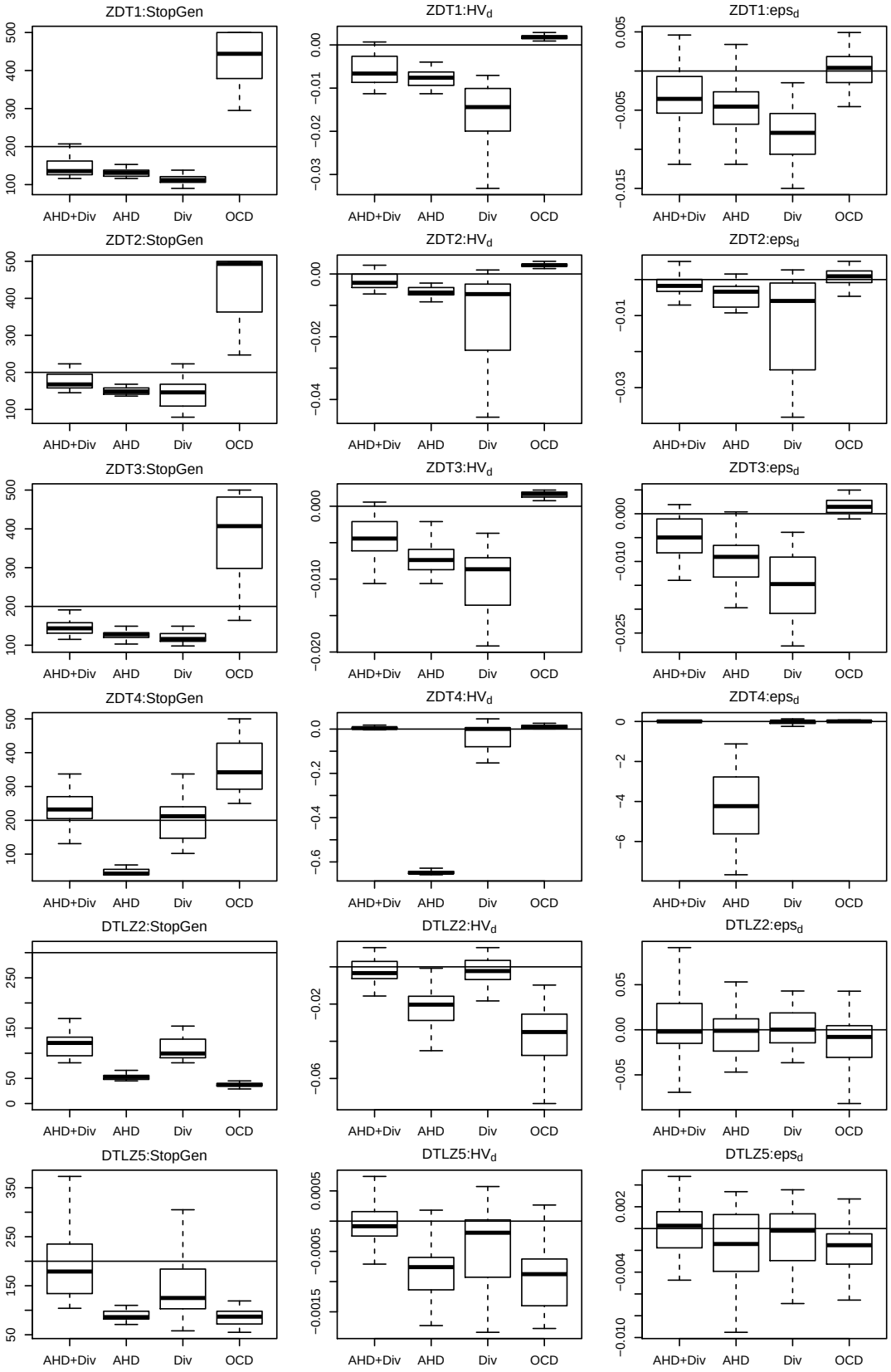


Figure 5.4: Boxplots for *stopping generation*, HV_d and eps_d on the different problems for NSGA-II.

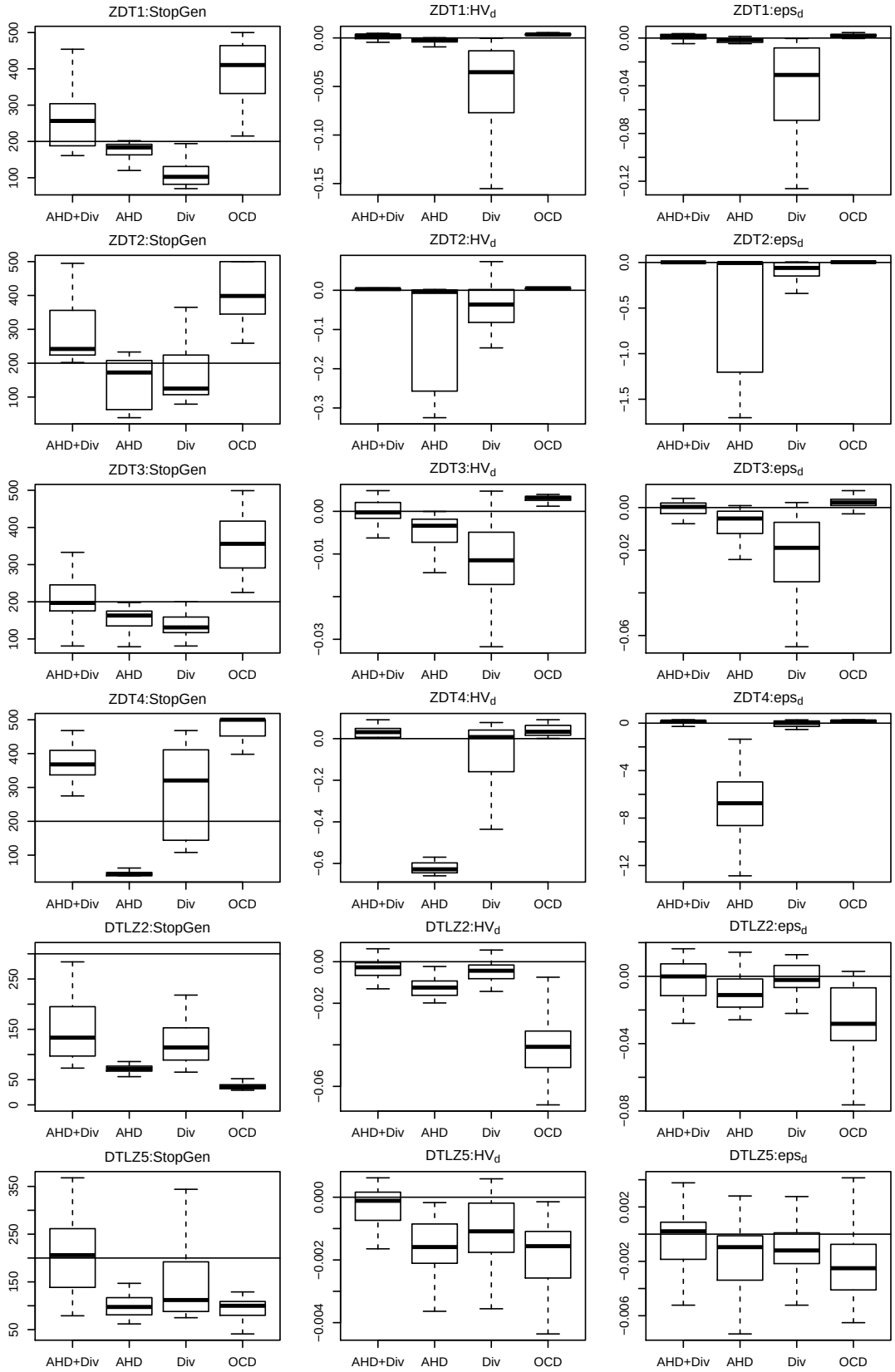


Figure 5.5: Boxplots for *stopping generation*, HV_d and eps_d on the different problems for SPEA2.

(standard budget). Very good percentages of true HVs (HV_{per}) are obtained. HV_{per} has reached over 95% for all the problems, except *DTLZ2* and *ZDT4*. HV_{per}s are considerably low (0.7916) for NSGA-II and SPEA2 (0.8557) on the *DTLZ2* problem. However, please note that the average HV_{per}s are 0.7976 and 0.8657 for standard $MaxGen$ for NSGA-II and SPEA2, respectively. Therefore, there is almost no difference in terms of achieved HV_{per} when comparing our method with standard $MaxGen$. Nevertheless, sometimes the method detects convergence prematurely on the *ZDT4* problem. It should be noted that the *ZDT4* problem has 21^9 local Pareto-fronts. Finally, in comparison to standard budget recommendation, on average our proposed method saves 29% and 17% of function evaluations for NSGA-II and SPEA2, respectively.

5.5 Conclusion

In this chapter, we have shown that the simultaneous monitoring of the objective and decision space of a problem, can improve the robustness of stopping MOEA. Two metrics have been proposed for the different spaces. The stagnation has been detected by a two-sided t-test on the regression coefficients estimated from the matrices in the two spaces.

We have validated our method by investigating the performance on six different benchmark problems for two well-established MOEAs. Our proposed method performs consistently well for all problems, also compared to state of the art stopping criteria. Additionally, on average a decent number of function evaluations has been saved compared to fixed budget recommendations without losing significant approximation accuracy.

In the next chapter, the developed stopping method will be employed within the energy system optimization framework. Moreover, the efficiency of the framework will be improved by incorporating all the individual components developed in previous chapter.

Chapter 6

An integrated approach to energy system optimization

Science is what we understand well enough to explain to a computer. Art is everything else we do.

— Donald Knuth

6.1 Introduction

It is very importation for an optimizer to identify solutions in a efficient manner; both in terms of quality of the solutions and in reduced computational cost. Some of the individual components, such as *smart initialization* and *smart mutation* where domain knowledge about energy system is incorporated), are developed and tested individually (in Chapter 4) which improve the efficiency of energy system optimization framework. Finally, a robust stopping criterion is proposed in the previous chapter that stops a MOEA automatically when the algorithms converge. In this chapter, we will integrate all the components developed previously into a single framework. Figure 6.1 presents the typical steps for the framework proposed in Chapter 3. However, three yellow-blue boxes show modified steps where developed components are integrated into the framework. The results show that our integrated approach yields better Pareto-front in less time than generic approach.

6.2 Experiments and results

We want to compare the integrated approach (i.e., smart initialization, smart mutation and stopping criterion are combined) with a typical approach (i.e., random initialization, polynomial mutation and fixed maximum generations) for NSGA-II and SPEA2 on the Aalborg energy system problem.

M. S. Mahbub, M. Wagner, and L. Crema, "Incorporating domain knowledge into the optimization of energy systems," *Applied Soft Computing*, vol. 47, no. C, pp. 483–493, Oct. 2016. DOI: `http://dx.doi.org/10.1016/j.asoc.2016.06.013`
The text is verbatim apart from the introduction and the conclusion and mentioned additions in chapter 4 and chapter 6; results regarding integrated approach are included in chapter 6. However, the introduction and the conclusion of the chapter are written independently to the published copy.

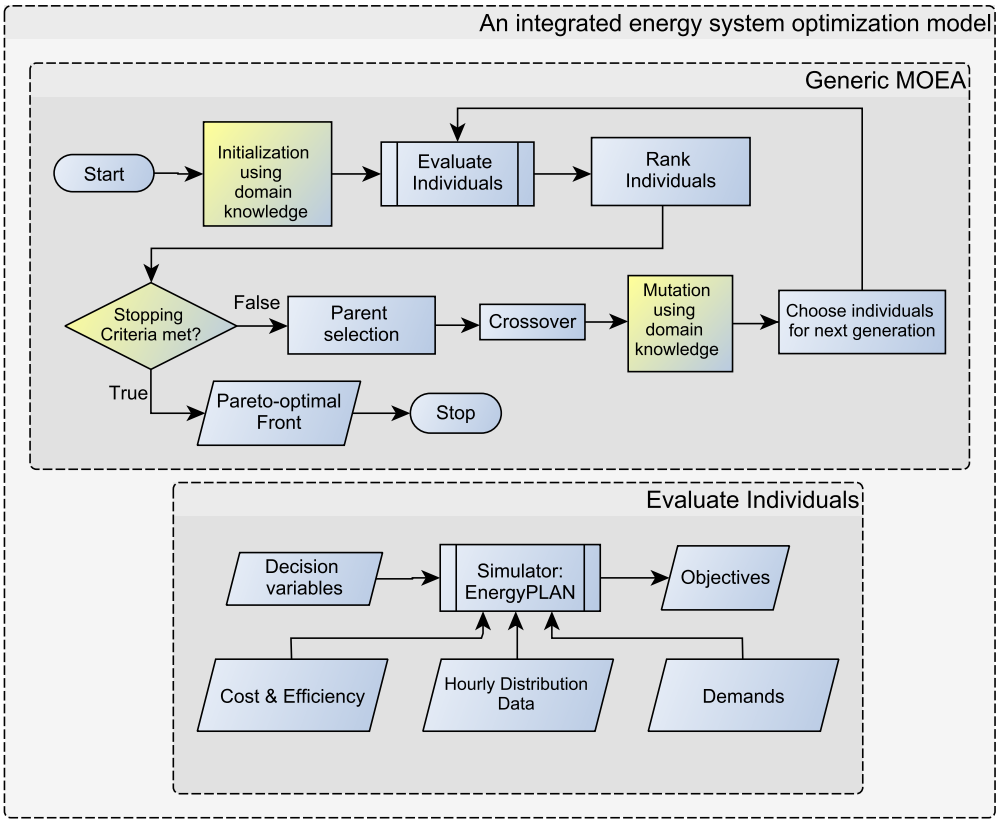


Figure 6.1: An integrated approach to energy system optimization model.

For this experiment, we use the following parameters for the stopping criterion: $nGenLT = 20$, $nGenUnCH = 5$, $\alpha = 0.05$ and $MaxGen = 70$ [1]. The values are chosen in such a way that the stopping happens early, which we prefer as we are using costly simulations. All other parameters remain unchanged as it was in Chapter 4. Each algorithm runs 30 times independently for the experiment.

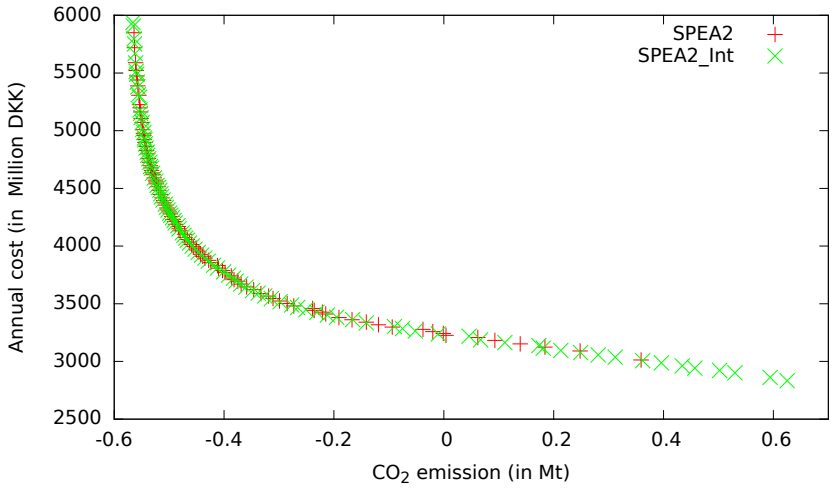


Figure 6.2: Pareto-fronts generated by generic SPEA2 and SPEA2_Int.

[1] Please see Chapter 5 for the explanation of the parameters.

First, Figure 6.2 presents an example of the comparison of two Pareto-fronts generated by the generic SPEA2 and by SPEA2 with the integrated approach (SPEA2_Int). Again, SPEA2_Int clearly performs better than the generic approach. In addition, Table 6.1 shows the results of the four different MOEAs, where a darker cell color indicates a better result. Figure 6.3 shows the results as boxplots to compare NSGA-II and NSGA-II with the integrated approach (NSGA-II_Int), and similarly for SPEA2. Table 6.2 presents the *p-values* for all the metrics. From the boxplots and Table 6.1, it is clear that our integrated approach performs very similarly to a typical approach for NSGA-II (*p-values* show there is no statistically significant differences), however, it outperforms a typical approach on every chosen metric for SPEA2 (*p-values* show there is statistically significance differences). Nevertheless, it should be noted that our approach achieves the similar results for NSGA-II with fewer function evaluations (Table 6.3); on average almost one-forth of total function evaluations (Table 6.3) are saved. On the other hand, for SPEA2, a significant improvement of the metric values are obtained, compared to the typical method. Hence, it can be concluded that our approach has a positive impact on MOEAs.

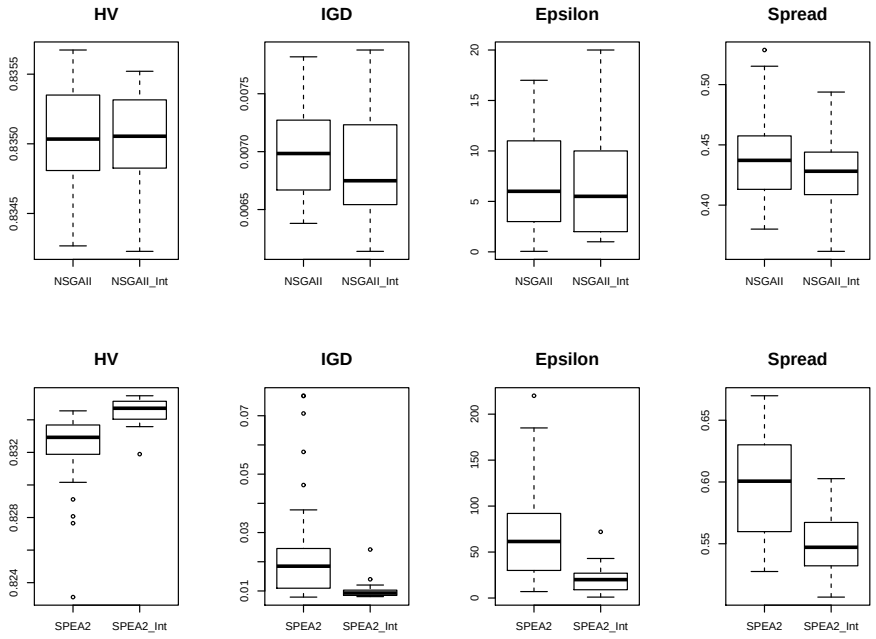


Figure 6.3: Boxplots for four metrics comparing NSGA-II and SPEA2 with integrated approach.

Table 6.1: Mean and standard deviations of four different metrics

Metric	NSGA-II	NSGA-II_Int	SPEA2	SPEA2_Int
HV	$8.35e-01_{3.6e-04}$	$8.35e-01_{3.5e-04}$	$8.32e-01_{2.4e-03}$	$8.35e-01_{7.5e-04}$
IGD	$7.01e-03_{4.0e-04}$	$6.85e-03_{4.3e-04}$	$2.51e-02_{2.0e-02}$	$9.96e-03_{2.9e-03}$
Epsilon	$6.70e+00_{4.8e+00}$	$6.57e+00_{4.9e+00}$	$7.06e+01_{5.5e+01}$	$2.03e+01_{1.4e+01}$
Spread	$4.38e-01_{3.6e-02}$	$4.27e-01_{3.1e-02}$	$5.98e-01_{4.2e-02}$	$5.51e-01_{2.5e-02}$

Lastly, we list in Table 6.3 the total number and average percentage of saved function evaluations for each MOEA for 30 runs. Note that all the simulations

Table 6.2: Mann-Whitney U-tests: p-values for different metrics when comparing our integrated approaches (Int) with the generic approaches.

	p-value	
Evaluation metrics	Compare NSGA-II: With integrated and generic approach	Compare SPEA2: With integrated and generic approach
HV	0.7191	$1.63e^{-09}$
IGD	0.1727	$2.20e^{-16}$
Epsilon	0.8704	$6.932e^{-06}$
Spread	0.3136	$4.106e^{-09}$

Table 6.3: Number and percentage of saved function evaluations for different MOEAs.

MOEA	Total saved function evaluation	Average percentage of saved function evaluations
NSGA-II_Int	45400	22%
SPEA2_Int	27400	13%

without our stopping criterion technique use 7,000 function evaluations (i.e., 70 generations, Section 4.4.1). We calculate the number of saved function evaluations (SFE) for each run by $SFE = (70 - Gen_{sc}) * PS$, where Gen_{sc} is stopping generation when the the stopping criterion is activated and PS is the population size (in the experiment, $PS = 100$). Then, we sum up all SFE values of the 30 runs to find the total SFE. The average percentage of SFE is calculable by $\frac{SFE^T * 100}{FE^T}$, where SFE^T is the total SFE and $FE^T = 210,000$ as this is the total number of function evaluations required (for 30 runs) when the stopping criterion is not used. According to the results, the modification of NSGA-II has saved the largest number of function evaluations. On average 22% function evaluations are saved, while also yielding a similar performance of the final solutions (see Table 6.1 and Figure 6.3).

The poor results of SPEA2 can be also explained by the number of saved evaluations. According to the table, fewer evaluations are saved compared to NSGA-II. Therefore, it is clear that SPEA2 is not able to converge as quickly as NSGA-II on this problem.

6.3 Conclusion

In this chapter, we incorporate all the developed components of Chapter 4 and 5 under a single framework. The framework is applied on the Aalborg test problem. The results show a significance improvement can be achieved in term of evaluation metrics (HV, IGD, Epsilon ans spread for SPEA2) and number of saved function evaluations (22% for NSGAII with respect to usual number of function evaluation). It is now a complete framework where domain knowledge about energy system is applied to increase the efficiency of algorithm, at the same time, a robust stopping criterion is incorporated to detect the convergence, hence, saving a lot of valuable computational time.

In the following chapter we will describe a practical application of the framework. The framework will be applied to identify multiple optimized scenarios for the studied area.

Chapter 7

A practical application of the framework: the case of Giudicarie Esteriori

Global warming is no longer a philosophical threat, no longer a future threat, no longer a threat at all. It's our reality.

— Bill McKibben

Nomenclature

Q_{sh} Space heating demand (kWh/year)

n_{dw} Number of dwellings in a specific time period

$\overline{Q_{sh}^{TR}}$ Average dwelling space heating demand at the age of construction (kWh/m^2)

DD_{Mun} Local municipality degree day

DD_{TR} Degree day in the city of Trento

$\bar{S}$ Average living area of the dwellings (m^2)

Q_{sh}^{h} Hourly space heating demand (kWh)

T_a Conventional ambient temperature ($°C$)

T_{out}^{h} Hourly outdoor temperature ($°C$)

Q_{HSW} Yearly hot sanitary water heating demand (kWh/year)

ρ_w Volumetric mass of water (Kg/m^3)

c_w Specific heat of water ($J/(Kg * K)$)

V_w Volume of daily required hot and sanitary water (m^3)

θ_{HSW} Hot sanitary water temperature ($°C$)

θ_t Tap water temperature ($°C$)

ND Number of days in a year

n_{in} Number of inhabitants

Q_{tr} Yearly transportation energy demand (kWh/year)

Q_{tr}^{pc} Yearly transportation energy demand for petrol cars (kWh/year)

Q_{tr}^{dc} Yearly transportation energy demand for diesel cars (kWh/year)

n_{pc} Number of cars fueled by petrol

n_{dc} Number of cars fueled by diesel

d_{pc} Average yearly distance traveled by a petrol car (km/year)

d_{dc} Average yearly distance traveled by a diesel car (km/year)

f_{pc} Average fuel efficiency of a petrol car (km/l (per liter of fuel))

f_{dc} Average fuel efficiency of a diesel car (km/l (per liter of fuel))

LCV_{pc} Lower calorific value for petrol (kWh/l)

LCV_{dc} Lower calorific value for diesel (kWh/l)

el_i Yearly electricity imported (kWh/year)

el_e Yearly electricity exported (kWh/year)

el_d Yearly electricity demand (kWh/year)

PE_i Yearly primary energy imported (kWh/year)

PE_d Yearly primary energy demand (kWh/year)

el_{lp} Yearly local electricity production (kWh/year)

LPG_i Yearly liquid petroleum gas imported (kWh/year)

NG_i Yearly natural gas imported (kWh/year)

Oil_i Yearly oil imported (kWh/year)

BM_c Yearly local biomass consumption (kWh/year)

H_{HP} Yearly heat produced by heat pumps (kWh/year)

COP Coefficient of performance

PEF_i^{el} Primary energy factor for imported electricity

PEF_{lp}^{el} Primary energy factor for locally produced electricity

el_d^{con} Electrical demand by considering electrification of all the sectors (kWh)

PV_{LP} Local producibility of photovoltaics (kWh/kW)

elC_{eff} Electric car efficiency (kWh/km)

SWR Yearly sustainable wood resource (GWh/year)

S_{tot} Total surface area (Hectare (ha))

S_{for}^{p} Percentage of surface area covered by forest

S_{for} Surface covered by forest (ha)

P_{SWR} Potentiality of sustainable use of the wood resource (MWh/(ha*year))

7.1 Introduction

There is an international effort to support multiannual programs for the energy transition from a fossil fuel-based society to a different energy mix and integration. In particularly in the last decade, the European Union has initiated several plans that promote new technologies, introduce renewed infrastructure and improve energy efficiency. All the efforts will help the transition of Europe to a sustainable, low-carbon and environmental friendly economy. Moreover, the European Union has autonomously set targets for climate and energy for the year of 2020 [65], 2030 [67] and 2050 [64]. Member states are required to contribute to achieve these goals with internal plans and strategies.

The challenges of minimizing energy costs, decreasing the dependency on foreign resources, reducing carbon footprint and integrating renewable energy sources, require a detailed analysis of many possible energy scenarios. Moreover, many renewable energy sources, such as solar, wind and water, are dependent on weather; therefore, it is necessary to find the correct mix among all the available resources to model potential energy scenarios. Usually, the appropriate sustainable scenarios do not include a predominance of one resource over others, rather a suitable mix, based on local resources and demands is sought.

In this study, the analysis of the energy system of Giudicarie Esteriori (a small Italian Alpine valley) is carried out by EnergyPLAN [101]. As optimizing an energy system is a multi-objective optimization problem, we use our developed framework to identify optimized scenarios.

Afterwards, the identified optimized scenarios are categorized into different target groups. The identification of optimized scenarios and the analysis of some target scenarios are carried out for providing a whole range of different scenarios (in terms of different objectives) to the policy makers of the studied area. These results could be used by local policy makers to make a sustainable energy action plan (SEAP), reducing the local emissions under the framework of the covenant of mayors initiative [5].

The remainder of the chapter is organized as follows. In section 2, the analysis of electrical, thermal and transport sectors of Giudicarie Esteriori is presented. Section 3 describes the applied framework in brief; moreover, considered decision variables, constraints and objectives are discussed. Section 4 presents the results and the corresponding discussion about the results. Finally, the chapter is concluded by providing some potential directions for the decision makers.

7.2 The study area: Giudicarie Esteriori

7.2.1 Location and population

The studied area, an Italian Alpine valley named Giudicarie Esteriori, is located in the province of Trento (Figure 7.1). The area is characterized by the presence

M. S. Mahbub, D. Viesi, and L. Crema, "Designing optimized energy scenarios for an italian alpine valley: the case of giudicarie esteriori," *Energy*, vol. 116, Part 1, pp. 236 – 249, 2016. DOI: http://dx.doi.org/10.1016/j.energy.2016.09.090
The text is verbatim without introduction and conclusion in chapter 7; the introduction and conclusion have been modified to ensure a better flow of the text in the thesis.

of a cooperative called "Consorzio Elettrico Industriale di Stenico (CEIS)". CEIS was established in 1905 with the aim of contributing to the economic and social improvements of the people living in the area, through the production and distribution of electricity. CEIS is one of the 77 Italian historical electric cooperatives. A considerable amount of economic saving (up to 40% compared to the national average electricity price [32]) is possible to the member of CEIS, as CEIS has the ownership of the electricity production plants and the electrical distribution grid.

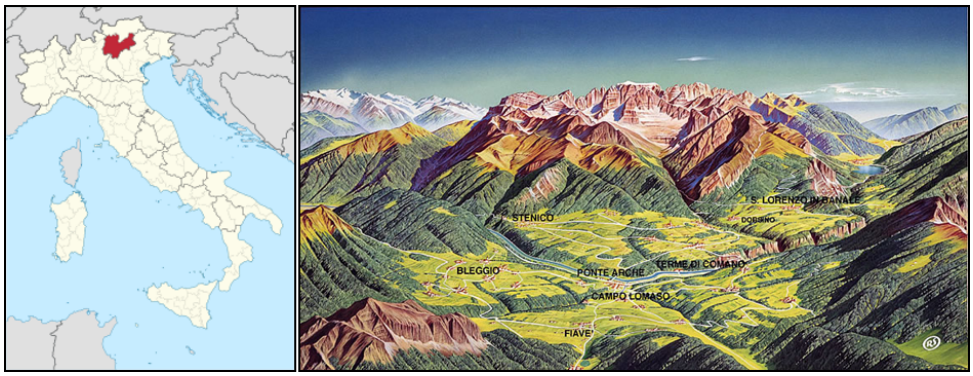


Figure 7.1: The Province of Trento and the area of Giudicarie Esteriori supplied by CEIS

The supplied area of CEIS has a surface of 248 km^2 and includes 8,426 citizens and 3,536 families. At 31st December 2013, the shareholding structure was made of 3,425 members who represent about 80% of households served by CEIS [32]. In the next few sub-chapters, we analyze energy demands and productions of the reference scenario (RS). The RS characterizes the state of the energy system in the year 2013.

7.2.2 Electricity production and demand

CEIS produces electricity using only renewable energy sources; it has the ownership of 1 hydropower plant (4.0 MW) and 5 centralised photovoltaic (PV) plants (1.0 MW). In addition, the diffusion of residential PV panels has been greatly increased (442 PV plants, 6.5 MW, at 31st December 2013) in recent year because of high national grants. Two CEIS members own two biogas plants (0.5 MW), producing renewable electricity from cow waste [32]. Part of the electricity produced by residential PV panels and biogas is used for self-consumption and the remaining is injected into the CEIS grid. The number of plants and corresponding installed capacities (kW) for the reference year (2013) [32] is reported in Table 7.1. As an

Table 7.1: Number of different electricity producing plants and corresponding installed capacities (kW) of CEIS.

Renewable energy sources	Number of Plants	Installed capacity
Biogas	2	500
PV	447	7515
Hydropower	1	4000

agreement with the cooperative, the hourly electrical production and demand data

are made available to the authors. Using the available data, the comparison of electrical production (monthly) from different plants (total: 29.4 GWh) and electrical demand (total: 26.2 GWh) is illustrated in Figure 7.2 (the demand includes domestic sector, industry, services, agriculture etc.). The hydro plant has the highest production among all other technologies and it is characterized by a major peak in spring and a secondary peak in autumn. PV production has one major peak in summer. On the other hand, electrical demand remains steady over the time, however, some small monthly variations are observed. The total monthly electrical production exceeds the electrical demand between the months of April to November, while between December to March, a considerable amount of electricity import is required from the national grid (purchased by the vendor Trenta S.p.A.). The lack of electrical balancing between local production and demand is even more visible at the daily and hourly level; indeed, the daily production profile does not match two peaks of the demand (i.e., morning and evening). While hydro production is moderately regulated during low-production months (using a very small hydro storage reservoir), such activity is not implemented for PV production. The large variable and non-programmable power flows make the management (of the system) challenging.

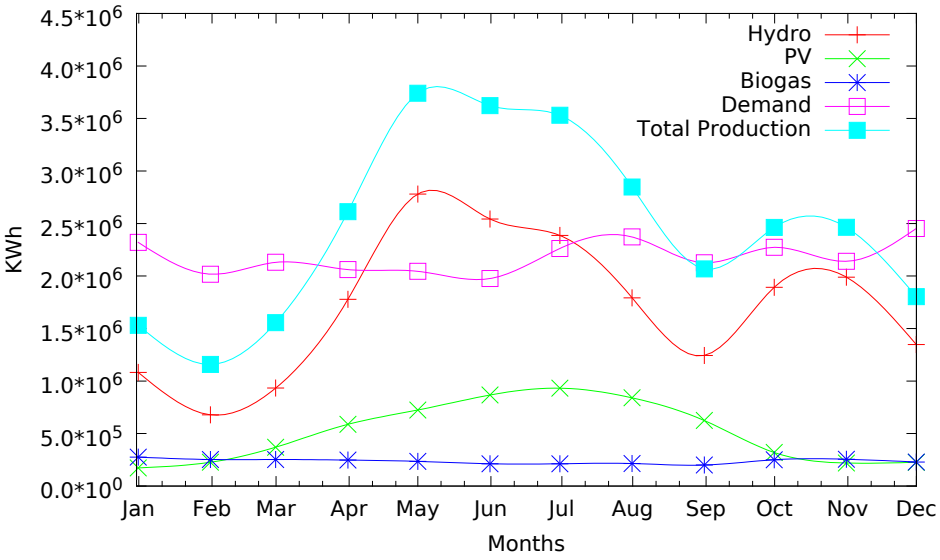


Figure 7.2: Monthly electricity demands and productions from different plants.

The area served by the cooperative is not an energy island system; the local electrical gird is connected with the national grid. During the reference year the total electricity import reached about 4.6 GWh and the export reached about 7.8 GWh, and maximum peak of importing is about 3.2 MW, while maximum peak of exporting is about 7 MW.

7.2.3 Thermal energy demand

Because of the lack of monitoring facilities, it is not possible to get the real thermal consumption data for the reference scenario. Therefore, the thermal demand is estimated. The total thermal demand is divided into two sub-sectors: space heating and heating requirement for hot sanitary water (HSW).

The yearly space heating demand (Q_{sh}) can be estimated at the municipality level by using the following formula:

$$Q_{sh} = n_{dw} * \overline{Q_{sh}^{TR}} * \left(\frac{DD_{Mun}}{DD_{TR}} \right) * S \tag{7.1}$$

The demand for space heating is considered variable throughout a year, based on the outdoor temperature (our data is collected from the weather station "Meteotrentino T0414 San Lorenzo in Banale (Pergoletti)" [15]) and by considering a conventional ambient temperature of $20°C$. The hourly space heating demand (Q_{sh}^{h}) is evaluated at municipality level by the following formula:

$$Q_{sh}^{h} = (Q_{sh}) * \frac{1}{\sum_{h=1}^{8760}(T_a - T_{out}^{h})} * (T_a - T_{out}^{h}) \tag{7.2}$$

The heating demand for HSW (Q_{HSW}) is considered constant throughout a year, with an average HSW usage of $\frac{65l}{(day*person)}$ tap water. In addition, tap water temperature is considered to increase from $10°C$ to $40°C$. Therefore, the yearly HSW heating demand is calculated at the municipality level by:

$$Q_{HSW} = \rho_w * c_w * (V_w * (\theta_{HSW} - \theta_t)) * ND * n_{in} \tag{7.3}$$

The total yearly thermal demand in Giudicarie Esteriori is estimated to be 55.83 GWh. Space heating demand is very high compared to heating required for hot water (monthly demands are illustrated in Figure **??**reaching its peak in the middle of winter (December — March). Hot water heating demand remains constant throughout the year.

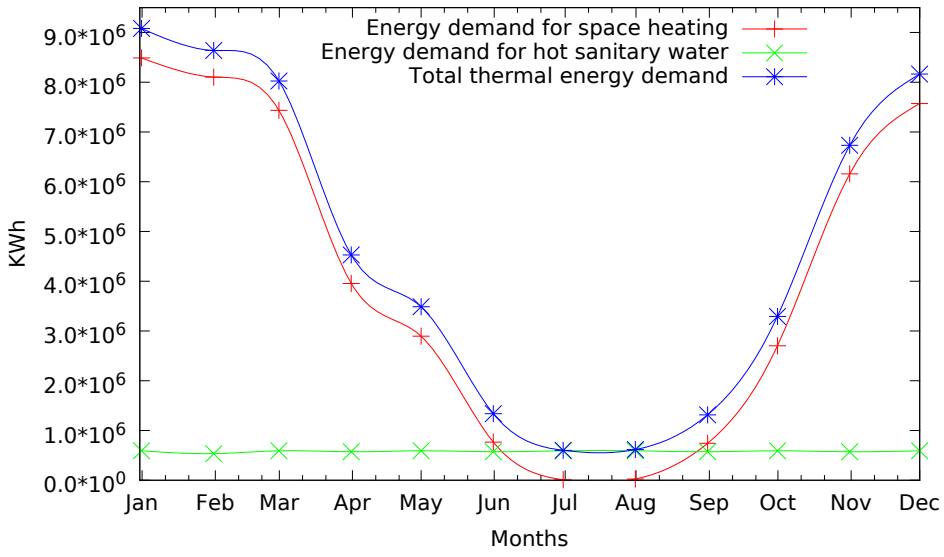


Figure 7.3: Monthly thermal energy demand for two sub-sectors in Giudicarie Esteriori.

The energy mix and generation technologies for thermal energy are identified by considering municipality statistical data provided by the province of Trento [17] and questionnaires collected from local citizens. Almost all the dwellings use individual boilers to meet the corresponding demands, however, different types of individual

boilers are in use (please see Table 7.2 for thermal energy mix). The main fuels for the individual boilers are wood, oil and liquid petroleum gas (LPG). The use of wood-fired individual boilers is highly widespread; thanks to the large local surface covered by forest and the low cost of wood in the market. In comparison with the electrical part, thermal part has a lower share of renewable sources (i.e., only wood) and a large share of imported fossil fuels (oil and LPG).

Table 7.2: Energy mix for thermal energy in Giudicarie Esteriori.

Thermal energy source	Thermal energy demand (kWh)	Peak Power (kW)
Individual wood boiler	29,694,971	9,549
Individual oil boiler	19,004,782	6,126
Individual LPG boiler	7,126,793	2,342
Total	55,826,546	—

7.2.4 Energy demand for transport

The analysis of transport energy demand in Giudicarie Esteriori is conducted at municipality level based on data collected from Automobile Club d'Italia [23] and Unione Petrolifera Italiana [168] for the reference year 2013. The transport demand of the area is calculated by considering the demands of petrol and diesel cars (the average annual distance travelled, fuel efficiency and lower calorific value (LCV) applied in this study are presented in Table 7.3).

The yearly transport energy demand (Q_{tr}) is calculated by the following formula:

$$Q_{tr} = n_{pc} * d_{pc} * f_{pc} * LCV_{pc} + n_{dc} * d_{dc} * f_{dc} * LCV_{dc} \tag{7.4}$$

Table 7.3: Characteristics (average annual distance travelled, fuel efficiency and LCV) of petrol and diesel cars applied in this study.

Type of fules	Distance travelled (Km/year)	Fuel efficiency (km/l)	LCV (kWh/l)
Petrol	7,250	15.5	8.86
Diesel	13,400	18.2	10.12

Based on the formula, transport demand for each municipality is calculated and reported in Table 7.4. By summing all the municipalities' demands, it is calculated that the total yearly petrol and diesel demands are about 11.45 GWh and 15.60 GWh, respectively.

7.3 Methodology

We use the very similar framework to optimize energy system of the cooperative. As an optimization algorithm SPEA2 [180] is used in the place of NSGAII [53].

Table 7.4: Transport energy demand for the municipalities situated in Giudicarie Esteriori.

Municipality	n_{pc}	n_{dc}	Q_{tr}^{pc} (kWh)	Q_{tr}^{dc} (kWh)	Q_{tr} (kWh)
Bleggio Superiore	511	387	2,117,683	2,883,533	5,001,216
Comano Terme	937	710	3,883,109	5,290,202	9,173,312
Dorsino	135	102	559,446	760,001	1,319,467
Fiavè	385	291	1,595,515	2,168,238	3,771,203
San Lorenzo in Banale	385	292	1,595,515	2,175,689	3,771,203
Stenico	410	311	1,699,119	2,317,258	4,016,377
Total	2762	2094	11,446,263	15,602,371	27,048,634

We prefer SPEA2 over NSGAII because we need to optimize four objectives and SPEA2 has better perform better when more than three objectives need to optimize. Moreover, SPEA2 is also preferable over MOEA/D [176] because it needed fewer algorithmic parameters.

The section also describes the decision variables (i.e., technologies that considered for optimization) and the constraint of the energy system are described in detail. Finally, the objectives for the optimization are formalized and described.

7.3.1 Decision variables and constraints

In this sub-section, we discuss the technologies that are used as decision variables to design optimized energy scenarios for the studied area:

- **PV capacity:** It is still possible to increase the capacity of PV. The maximum reasonable PV capacity is decided by dividing the annual maximum electrical consumption (by assuming that the energy system is completely electrified: existing electrical demand, existing thermal demand covered by ground source heat pumps, existing transport demand covered by electric cars) with the local PV productivity. The PV productivity can be calculated by dividing total yearly PV production by peak power of PV capacity. The calculated maximum reasonable PV capacity is approximately 42 MW (calculation is performed in Appendix A.1). Specifying a maximum capacity is necessary because the optimizer will find the optimal scenarios within the specified maximum limit. A large maximum limit is specified to let the optimizer find optimal scenarios within the range.

- 2. All the existing heat production technologies **(individual wood, oil and LPG boilers)** are considered as decision variables. In addition, **ground source heat pumps (GSHP)** are also considered. Heat pumps are a well-consolidated technology that can use ground, water or air as sources of heat or cold; the need of an electric compressor for running the thermodynamic process implies an electrical consumption.

- 3. Another technology that is chosen as a decision variable is **wood organic rankine cycle micro cogeneration (wood ORC mCHP)**. Organic ranking cycle allows to suit low grade thermal power to simultaneously provide

electric and thermal power thanks to an organic working fluid such as refrigerant (e.g., R245fa) [22].

- 4. The transport sector could be radically transformed by increasing the use of an alternative energy carrier such as electricity. The transition from fossil fuels cars to electric cars could lead very high benefits in terms of CO_2 emissions and achieving energy independence. In this study, a night charging profile is considered as it is assumed that all electric cars will be charged from 21:00 to 4:00. The cost associated with recharging infrastructure is not considered.

In the optimization process, one constraint is considered: the total wood consumption has to be less than nearly 57 GWh/year for Giudicarie Esteriori. The calculation is based on the assessment of local sustainable wood resource presented in the Appendix A.2.

7.4 Objectives

In this study, four objectives are considered. In addition, these objectives need to be minimized to design better energy scenarios.

1. **CO_2 emissions**: Minimizing CO_2 emissions is one of the most important aspect in terms of the environment. The quantity of CO_2 emissions of a given energy system is calculated by EnergyPLAN [101].

2. **Annual Cost**: Annual cost is a parameter that measures the economical aspect of an energy system. Annual cost is calculated by summing four different yearly costs: annual investment cost, variable operational and maintenance (O&M) cost, fixed operational and maintenance cost, and additional electric grid cost. Annual investment, variable O&M and fixed O&M costs are calculated by EnergyPLAN [101]. Additional electric grid cost includes all other extra costs (i.e., general system cost, grid and metering cost and taxes) that are incurred by consuming electricity.

3. **Load following capacity (LFC)**: LFC is a technical parameter that measures how much electricity production follows electricity demand over a period (e.g., yearly). LFC is formulated as follows:

$$LFC = \frac{el_i + el_e}{el_d} \tag{7.5}$$

   Please note that el_i and el_e are output parameters of EnergyPLAN [101]. It is obvious that when import and export are low, the load following capacity is getting low. In addition, import and export are low when an energy system can produce electricity as required. This parameter is particularly relevant for energy systems with a high fraction of electricity produced from renewable sources, because of intermittent behavior.

4. **Energy system dependency (ESD)**: ESD measures how much an energy system depends on foreign import of energy. The dependency is a ratio between primary energy imported (i.e., supply from outside of the system) and

total primary energy demand of the system. Minimizing the dependency implies that less energy is imported. Being energy independent is an important consideration for a community and its policy makers, since it could indeed reduce energy costs, increase local employment and reduce dependency from external energy markets that are not directly controllable. The general equation for ESD is:

$$EDS = \frac{PE_i}{PE_d} \tag{7.6}$$

However, a specific formula for Giudicarie Esteriori is as follows:

$$ESD = \frac{(el_i * PEF_i^{el} + LPG_i + oil_i)}{(el_{lp} - el_e) * PEF_{lp}^{el} + el_i * PEF_i^{el} + LPG_i + oil_i + BM_c + (H_{HP} - H_{HP}/COP)} \tag{7.7}$$

The primary Energy factor for locally produced electricity (PEF_{lp}^{el}, used in the aforementioned equation) is calculated by following formula:

$$PEF_{lp}^{el} = \frac{PE_{lp}^{el}}{el_{lp}} = \frac{\sum_{t \in T} el_p^t * PEF^t}{el_p^t} \tag{7.8}$$

Where PE_{lp}^{el} is the total primary energy required for local electricity production, el_p^t is the electricity production t different technologies, PEF^t is the primary energy factor for the technology (t). T is the set of all the electricity-generating technologies in Giudicarie Esteriori.

7.5 Simulation and results

In the first sub-section, the results regarding the reference scenario are presented. A simulation is conducted to identify optimized scenarios, and we report the results in the subsequent sub-sections. Afterwards, some of the optimized scenarios are specially identified and compared with respect to the reference scenario. Finally, a general discussion about different technologies concludes the section.

7.5.1 Results of the reference scenario

After collecting all the data about demands and productions (electric, thermal, and transport) that characterize the Giudicarie Esteriori energy system (Chapter 7.2), it is now possible to model the reference scenario using EnergyPLAN, in order to obtain the numeric values for four considered objectives. The technical and economical parameters used for the simulation can be found in Appendix A.3. The numeric values for four objectives are reported in Table 7.5. In the table, kEuro and kt represent thousand euro and kilo ton, respectively.

7.5.2 Results of optimized scenarios

After the completion of the simulation, 401 optimized scenarios are identified. As we have four objectives to optimize, we need to visualize four dimensional Pareto-front, which is a difficult task. However, using a three-dimensional view with different

Table 7.5: The reference scenario of Giudicarie Esteriori in terms of four objectives.

Economical aspect	
Total variable cost	11,782 kEuro
Fixed operational cost	713 kEuro
Additional cost	2,780 kEuro
Investment cost	0 kEuro
Total Annual cost	**15,275 kEuro**
Environmental aspect	
CO_2 emissions	13.092 kt
Technical aspect	
Import	4.63 GWh
Export	7.87 GWh
Load following Capacity (LFC)	**0.48**
Political aspect	
Energy System Dependency (ESD)	**0.51**

colors (i.e., 4th objective is presented as color), it is possible to visualize the front (see Figure 7.4). X, Y and Z-axis correspond to CO_2 emissions (kt/year), annual cost (KEuro) and load following capacity, respectively; the different colors represent different values of energy system dependency. Please note that some of the optimized scenarios have negative emissions, since the emissions are emitted from not only within the system, but also outside of the system (due to importing/exporting electricity). As some scenarios in the Pareto-front export a large amount of green electricity, the net emissions of the systems get lower and eventually even become negative (please see Chapter 3 for more explanation).

The CO_2 emissions of the optimized scenarios range from -26.14 to 9.40 kt, the annual costs range from 13456 to 39626 kEuro, load following capacities extend from 0.38 to 1.88, and energy dependency range from 0.11 to 0.36; these are shown as spheres in Figure 7.4. In comparison to the reference scenario (shown as a gray cube at the bottom of Figure 7.4) all the identified scenarios emit less emissions than reference one. Most of the scenarios (i.e., 375 out of 401) are costlier than the reference scenario. Finally, all the identified scenarios are better in terms of energy system dependency, however, only 54 scenarios are superior to the reference scenario in terms of load following. Overall, 13 scenarios are identified that are superior to the reference scenario in all four objectives. It can be concluded from the results that it is straightforward to make improvements with respect to emissions and dependency of the considered studied area. However, improvements for the other objectives (i.e., annual cost and load following) appear a bit more difficult to achieve.

It is very difficult to study all the 401 optimized scenarios individually. Therefore, we present some significant scenarios and their trends in terms of decision variables in three different categories. In the first category, some best scenarios in terms of annual cost will be presented. Afterwards, some target scenarios in term of CO_2 emissions and energy dependency reduction will be presented. To present the target scenarios, some groups are defined. Target groups are defined in terms of CO_2 emissions reduction that involve a specific range of reduction percentage with respect to the reference scenario (RS) (e.g., 40 to 45% emissions reduction with respect to the RS emissions). We present the 3 least costly scenarios for each target group.

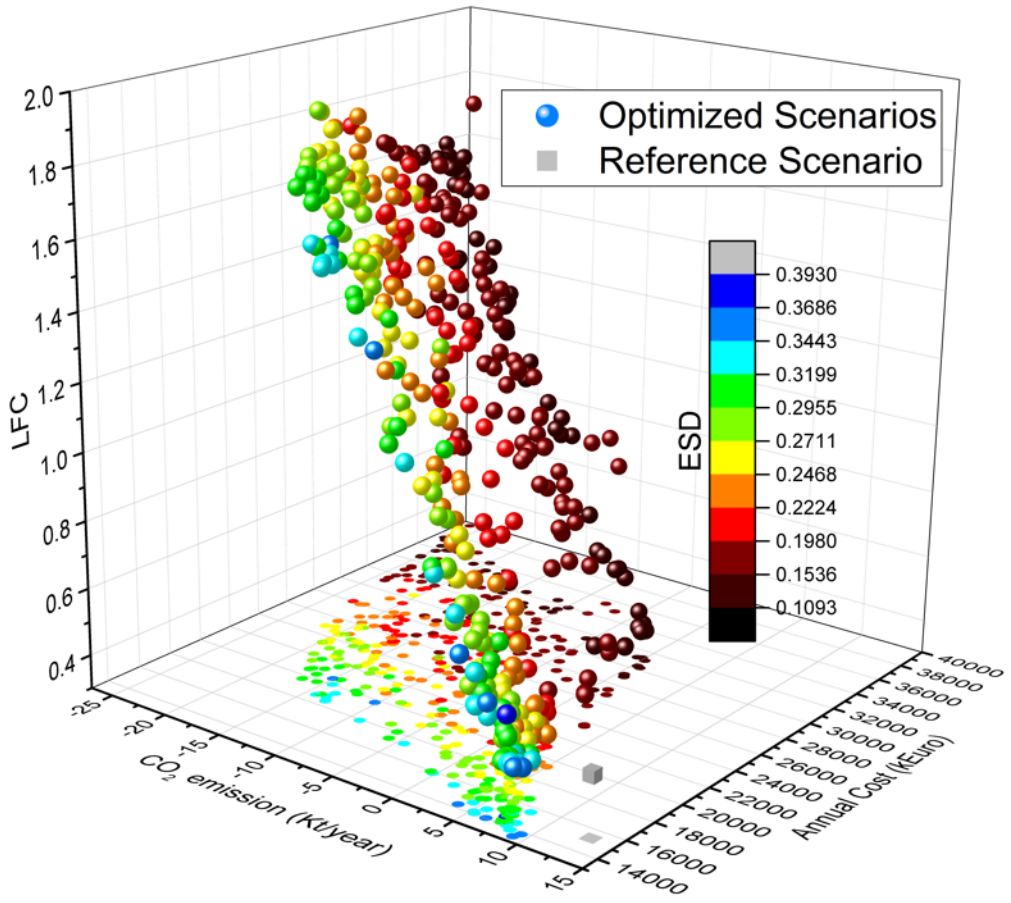


Figure 7.4: Pareto-front for Giudicarie Esteriori energy system optimization problem.

We use a similar approach for the energy dependency target scenarios. Finally, a general discussion about the implementation possibilities for the different addressed technologies (i.e., decision variables) concludes this section.

7.5.3 Best annual cost scenarios

One of the most interesting objectives of the optimization is annual cost, as the parameter is the most attractive to the policy makers of their respective communities. All the numeric values for the decision variables, the objectives and the amount of electricity exchanges of best 15 scenarios in terms of annual cost are reported in Table 7.6; first row of the table presents the reference scenario and other rows of the table present scenarios in ascending order with respect to annual cost (named as AC1, AC2, ..., AC15). Third column of the table shows the reduction of the cost in percentage with respect to the reference scenario. Moreover, a visual comparison of the capacities of different technologies of the scenarios is illustrated in Figure 7.5.

Based on Table 7.6 we find that consistent economic savings (from 4.8 to 11.9%) with respect to the reference scenario can be achieved. At the same time, it is interesting to point out that minimized annual cost scenarios also have consistence reductions of CO_2 emissions (from 28.1 to 70.9%). In other words, several energy

Table 7.6: Reference scenario (RS) and best 15 scenarios in terms of Annul cost (AC).

Scn.	AC (KEuro)	red(%)	CO_2 Emissions (kt)	LFC	ESD	Capacity						Number of car		
						PV (kWe)	Wood ORC mCHP (kWe)	GSHP (kWth)	Oil boiler (kWth)	LPG boiler (kWth)	Wood boiler (kWth)	Petrol cars	Diesel cars	Elec Cars
RS	15275	0	13.09	0.48	0.51	5000	0	0	9155	3431	14306	2762	2094	0
AC1	13456	11.9	9.16	0.49	0.36	5261	151	7008	97	406	15657	2759	2091	6
AC2	13800	9.69	9.40	0.48	0.37	5224	110	5558	1579	412	16404	2761	2093	2
AC3	13829	9.5	8.48	0.49	0.34	5363	320	6195	139	280	16706	2693	2042	121
AC4	13836	9.4	8.50	0.48	0.34	5015	44	5216	597	132	18278	2669	2023	164
AC5	14041	8.1	6.351	0.62	0.36	9765	82	7334	67	732	14973	2759	2092	5
AC6	14047	8	7.99	0.47	0.33	5057	204	3523	621	1020	19665	2760	2092	4
AC7	14079	7.8	5.38	0.6	0.32	8786	30	3949	860	48	20021	2755	2089	12
AC8	14145	7.4	7.42	0.47	0.32	5204	773	3661	559	441	19244	2761	2093	2
AC9	14196	7.1	5.19	0.6	0.32	8786	211	3738	860	48	20066	2755	2089	12
AC10	14243	6.8	7.819	0.47	0.33	5023	687	3654	646	831	18908	2757	2090	9
AC11	14296	6.4	3.81	0.72	0.35	12819	87	7072	97	161	15903	2759	2091	6
AC12	14372.2	5.9	7.274	0.47	0.31	5170	592	3738	257	464	19679	2657	2014	185
AC13	14398	5.7	7.67	0.47	0.32	5158	1269	4970	101	522	16918	2721	2063	72
AC14	14458	5.3	8.26	0.47	0.34	5092	1114	4739	590	1037	16492	2757	2090	9
AC15	14546	4.8	7.40	0.46	0.31	5015	1175	4042	597	132	18345	2669	2023	164

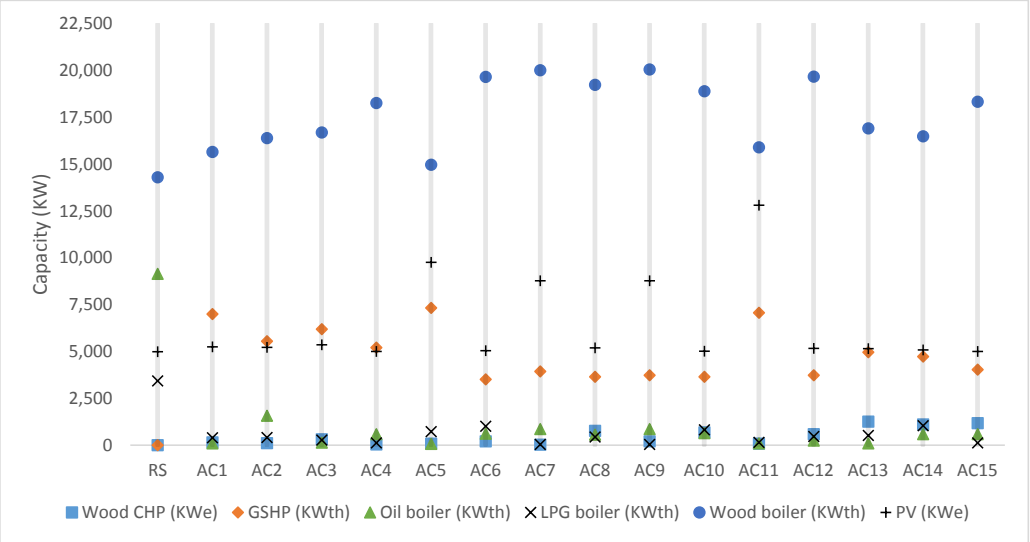


Figure 7.5: Capacities of different technologies (except transportation) for the reference scenario (RS) and for 15 best scenarios in terms of annual cost. X-axis represents scenarios and Y-axis represents capacities (in kW). Each vertical grey line represents a particular scenario.

scenarios are greener and cheaper than the reference one. Additionally, energy dependencies of the scenarios are highly reduced; these scenarios import 14 to 20% less energy, mainly thanks to the increasing use of local wood and the partial electrification of the thermal sector (through heat pumps). Load followings in most cases are very close to the value of the reference scenario, ensuring electrical grid stability. Electrical import increases from 4.63 GWh/year (of RS) to a range of 6.76 – 9.68 GWh/year, and electrical export moves from 7.87 GWh/year to a range of 6.48 – 14.87 GWh/year depending on scenarios (last two columns of Table 7.6). Note that there are no specific import or export trends can be found related to cost of the system. The reason of not getting any particular trend is that the cost related to importing and exporting electricity is rather small compared to other costs, and therefore no significant effect on cost can be observed.

According to the trend of micro CHP illustrated in Figure 7.5 (also illustrated the trends of other decision variables of 15 scenarios), the suggested introduction of micro CHP appears often negligible (30 – 1,269 kW_{el} for a yearly production of 0.02 – 0.89 GWhel). Instead, an increase of PV capacity is suggested from 5,015 kW_{el} (0.3% increases compared to the reference scenario; yearly production of 6.13 GWh) to 12,819 kWe (156.38% increases; yearly production of 15.68 GWh). Although small modifications are suggested in the electrical sector, the thermal sector is highly transformed; all individual oil and LPG boilers' capacities reach nearly zero. In other words, it is suggested to lower the oil and LPG boilers' capacities because of high fuel costs compared to other alternative energy carriers. On the other hand, the use of local wood and the electrification of the thermal sector have gained importance. Most of the thermal demands are covered by individual wood boilers (i.e., compared to the reference scenario, individual wood boilers' capacities for all 15 scenarios show higher trends), moreover, total wood consumption increases from 39.59 GWh/year (of RS) to a range of 41.75 – 56.76 GWh/year (close to the given constrain of

57 GWh/year, see Chapter 7.3.1). Even though the micro CHP does not appear economically attractive, an extensive introduction of heat pumps is observed for the cost effective scenarios. Thermal capacities of heat pumps range from 3,523 to 7,334 kWth. In the optimized scenarios, individual wood boilers cover 56 – 73% of total thermal demands (was 53% for the RS) while heat pumps cover 20 – 41% (was 0% of the reference scenario).

Concerning the transport sector, the number of introduced electric cars are almost negligible (from 2 to 185) [1]. Indeed, the investment necessary to replace fossil fuels based cars with electric cars is economically unattractive under the current market conditions.

7.5.4 Target CO_2 emissions scenarios

Annual cost is not always the only objective to consider, and many communities commit themselves to reach particular environmental targets, for example, as decided by the Covenant of Mayors [5]. In this regard, this study suggests several scenarios with ambitious CO_2 emissions reduction (Table 7.7) for the policy makers. In particular, some target groups are defined based on percentage reduction of CO_2 emissions with respect to the emissions of the reference scenario (each group has an interval size of five percent, e.g., 30 to 35%, 40 to 45% reduction and so on). Target groups start with a group of 30 to 35% reduction and end with a group of 100 to 105% reduction. Afterwards, we present three less costly scenarios (ascending order) for each target group (e.g., the three less costly scenario in a target group of 40 to 45% reduction are named as EM40A, EM40B and EM40C). Please note that for the first category/group (i.e., 30 to 35%) only one scenario is identified (i.e., only one scenario of the group is identified by the optimizer).

All the scenarios of the 1st and 2nd groups are less costly than the reference scenario (RS). Even some scenarios in different target groups (until the group of 70-75%) have less annual cost than the reference scenario. Energy dependency is highly reduced in all proposed scenarios through the reduction of fossil fuels import and by increasing the use of local resources. Load following capacities of the first two groups are close to the value of the reference scenario. In all other cases, the high fraction of non-programmable renewable energy production (i.e., PV) leads to poor load following capacity as it is required to export excess electricity [2]. Electrical imports shift from 4.63 GWh/year of the reference scenario to a range of 2.58 – 9.68 GWh/year; electrical exports move from 7.87 GWh/year to a range of 6.51 – 20.44 GWh/year depending on different scenarios. The exports have a significant impact on emissions reduction for the energy system. Less emitting scenarios export more electricity since the system exports completely green electricity, and thus reducing the emissions. However, there is much lesser impact on emissions of the system for importing electricity.

The electrical resource mix shows gradual increase of PV capacity with respect to the target group, in particular, the highest target scenarios suggest a large introduction of this technology. The trend of micro CHP is not stable, and it appears to

[1] As integer values are used for decision variables for optimizing the system, the number of electric cars can have any integer value between upper and lower bounds depending on different scenarios.

[2] It may require an expensive adaption of the grid transmission capacity.

Table 7.7: Target scenarios in terms of reduction of CO_2 emissions.

Scn.	CO_2 Emissions (kt)	red (%)	AC (kEuro)	LFC	ESD	Capacity						Number of cars		
						PV (kWe)	Wood ORC mCHP (kWe)	GSHP (kWth)	Oil boiler (kWth)	LPG boiler (kWth)	Wood boiler (kWth)	Petrol Cars	Diesel Cars	Elec Cars
RS	13.092	0.0	15,275	0.48	0.51	5,000	0	0	9,155	3,431	14,306	2,762	2,094	0
EM30A	9.162	30.0	13,456	0.49	0.36	5,261	151	7,008	97	406	15,657	2,759	2,091	6
EM40A	7.819	40.3	14,243	0.47	0.33	5,023	687	3,654	646	831	18,908	2,757	2,090	9
EM40B	7.420	43.3	14,145	0.47	0.32	5,204	773	3,661	559	441	19,244	2,761	2,093	2
EM40C	7.274	44.4	14,372	0.47	0.31	5,170	592	3,738	257	464	19,679	2,657	2,014	185
EM50A	6.454	50.7	15,114	0.52	0.32	7,380	2,618	6,902	91	76	12,449	2,738	2,076	42
EM50B	6.360	51.4	15,827	0.44	0.28	5,123	3,323	4,166	364	172	15,128	2,632	1,995	229
EM50C	6.351	51.5	14,041	0.62	0.36	9,765	82	7,334	67	732	14,973	2,759	2,092	5
EM60A	5.198	60.3	14,196	0.60	0.32	8,786	211	3,738	860	48	20,066	2,755	2,089	12
EM60B	4.727	63.9	15,400	0.57	0.30	8,625	2,414	4,678	56	737	15,467	2,755	2,089	12
EM60C	4.620	64.7	16,426	0.51	0.28	7,612	4,671	5,094	276	123	11,851	2,754	2,088	14
EM70A	3.813	70.9	14,296	0.72	0.35	12,819	87	7,072	97	161	15,903	2,759	2,091	6
EM70B	3.654	72.1	15,078	0.61	0.29	9,765	1,865	4,282	67	146	17,463	2,754	2,088	14
EM70C	3.501	73.3	15,588	0.64	0.31	10,768	2,201	5,053	427	253	15,337	2,711	2,055	90
EM80A	2.414	81.6	16,918	0.61	0.27	10,039	3,628	3,626	44	566	15,407	2,577	1,954	325
EM80B	2.149	83.6	17,058	0.60	0.27	10,053	4,806	3,843	44	506	13,374	2,717	2,060	79
EM80C	2.128	83.7	18,225	0.53	0.24	8,612	6,582	3,683	147	75	11,276	2,609	1,978	269
EM90A	0.989	92.4	16,235	0.74	0.29	13,095	2,632	3,533	773	28	16,847	2,718	2,060	78
EM90B	0.796	93.9	18,792	0.62	0.25	11,545	7,429	5,411	20	270	7,349	2,718	2,060	78
EM90C	0.735	94.4	18,224	0.70	0.28	12,754	3,893	3,872	315	902	14,032	2,407	1,825	624
EM100A	-0.022	100.2	19,813	0.59	0.22	10,067	9,488	3,670	42	110	7,009	2,719	2,062	75
EM100B	-0.231	101.8	15,875	0.86	0.32	17,414	1,090	6,602	83	19	15,259	2,663	2,019	174
EM100C	-0.500	103.8	15,966	0.88	0.31	16,709	52	4,064	392	283	20,046	2,551	1,934	371

be an expensive alternative. Some scenarios include it with relatively high capacity, others have less. However, it is clear that the large introduction of micro CHP can be beneficial to reduce energy dependency and load following capacity (e.g., scenario EM100A compared to EM100B).

As expected, in the thermal sector, individual oil and LPG boilers are included at only low overall capacity. However, the capacity of individual wood boilers is typically increased. Interestingly, wood boilers and micro CHP appear to be somewhat exchangeable, as the boilers' capacity decreases when the CHP capacity increases. In addition to this, heat pumps are extensively introduced.

Again, the introduction of electric cars is barely noticeable for these scenarios (i.e., from 2 to 371). This means that it does not require a radical transformation of transport sector even to reach 100

7.5.5 Target ESD scenarios

Scenarios with minimized energy dependency suggest which mix of technologies can maximize the use of local energy resources while minimizing the energy import. This study suggests several scenarios with high dependency reduction. The best identified scenario has an ESD value of 0.11, this means that the system needs only 11% external energy resources to cover all the local energy demands (electricity, thermal and transport). The scenario is characterized by the comprehensive introduction of micro CHPs and electric cars. Our interpretation is that micro CHPs use local resources (i.e., wood), in addition, electric cars may use the electricity produced locally (not the imported electricity). However, the less energy dependent scenarios are very costly and may not be attractive to the policy makers, unless they would come with significant benefits, e.g., in terms of local employment. Therefore, we present some target groups of energy dependency reduction in a similar way presented before.

We report the three least costly scenarios for each target group in Table 7.8. The target groups are defined by the reduction of energy dependency (with respect to the reference scenario) to a specific range. For example, the first target group represents 0.15–0.17 dependency reduction from the dependency value of the reference scenario (0.51). The subsequent groups represent scenarios (named as ESD20A, ESD20B and ESD20C) with a reduction of 0.20–0.22 and so on. The groups are presented in descending order by dependency values. The reduction of the dependency of a scenario with respect to the reference scenario is reported in the third column of the table. Please note that bottom groups of the table have lower dependency values, which implies an increase independence from foreign resources. In terms of annual cost, the first two groups are highly comparable with the annual cost of the reference scenario, with a significant dependency reduction. However, it becomes increasingly expensive to achieve lower dependency values. CO_2 emissions are highly reduced for all the identified scenarios (from 30 to 155.9% compared to the reference scenario): reducing fossil fuels import (and electricity from the national grid) means reducing emissions. Load following capacities, in most cases for first 3 groups, are close to the reference scenario. Electrical imports move from 4.63 GWh/year (of the reference scenario) to a range of 1.92–9.68 GWh/year, electrical exports shift from 7.87 GWh/year to a range of 6.48–22.65 GWh/year. Again, no specific tend is found for importing electricity with respect to energy dependency (there is no

Table 7.8: Target scenarios in term of ESD reduction.

Scn.	ESD	red	AC (kEuro)	CO_2 Emissions (kt)	LFC	Capacity						Number of Cars		
						PV (kWe)	Wood ORC mCHP (kWe)	GSHP (kWth)	Oil boiler (kWth)	LPG boiler (kWth)	Wood boiler (kWth)	Petrol Cars	Diesel Cars	Elec Cars
RS	0.51	0.000	15,275	13.092	0.48	5,000	0	0	9,155	3,431	14,306	2,762	2,094	0
ESD15A	0.36	0.150	13,456	9.162	0.49	5,261	151	7,008	97	406	15,657	2,759	2,091	6
ESD15B	0.34	0.166	13,829	8.480	0.49	5,363	320	6,195	139	280	16,706	2,693	2,042	121
ESD15C	0.34	0.169	13,836	8.509	0.48	5,015	44	5,216	597	132	18,278	2,669	2,023	164
ESD20A	0.30	0.202	15,400	4.727	0.57	8,625	2,414	4,678	56	737	15,467	2,755	2,089	12
ESD20B	0.30	0.210	15,429	6.665	0.46	5,462	2,789	4,554	278	434	15,171	2,684	2,035	137
ESD20C	0.29	0.212	15,078	3.654	0.61	9,765	1,865	4,282	67	146	17,463	2,754	2,088	14
ESD25A	0.25	0.253	17,112	4.572	0.45	6,007	6,066	4,013	277	109	11,394	2,752	2,086	18
ESD25B	0.25	0.258	16,812	4.945	0.42	5,182	5,923	3,843	90	4	12,156	2,760	2,093	3
ESD25C	0.25	0.261	17,793	4.326	0.44	6,007	6,168	3,910	277	109	11,394	2,583	1,959	314
ESD30A	0.20	0.304	23,238	-2.958	0.68	12,868	12,489	4,010	34	693	1,425	2,467	1,870	519
ESD30B	0.20	0.306	23,533	-5.477	0.81	17,254	9,937	4,128	202	47	5,550	2,075	1,573	1,208
ESD30C	0.19	0.320	23,521	-3.216	0.66	12,303	13,125	3,859	105	25	1,292	2,407	1,825	624
ESD35A	0.15	0.352	30,718	-7.312	0.82	17,663	9,036	3,875	78	397	7,058	10	8	4,838
ESD35B	0.15	0.353	30,556	-6.601	0.79	16,648	9,036	3,875	78	397	7,058	10	8	4,838
ESD35C	0.15	0.359	26,934	0.304	0.43	6,504	10,485	3,482	63	133	5,753	830	629	3,397

relation of exporting on energy dependency, please see the corresponding equation). The effect of importing electricity on energy dependency is far less than the effect of importing oil and LPG, as imported electricity is small compared to total energy imported (by converting the imports into primary energy).

The electrical resource mix shows a moderate PV capacity increase within the first 3 groups in the table, while lower dependency scenarios (i.e., the last 2 groups) make extensive use of this technology. Micro CHPs are negligible only for the first group, whereas for more ambitious target groups, more CHPs are introduced. Individual wood boilers are preferred over the CHP for first few target groups; however, less wood boilers are introduced for lower dependent scenarios. As usual, individual oil and LPG boilers are at very low capacity. Heat pumps are always broadly introduced, it is not only cheap and green but also allows an optimal use of local resources.

The first 3 target groups do not introduce a significant number of electric cars (from 3 to 314). However, in order to reach the higher target groups (in particular to reach last two groups of the table), it is necessary to have a radical transformation of the transport sector through the extensive introduction of electric cars.

7.5.6 General discussion

By analyzing the optimized scenarios, we identified some general conclusions regarding different technologies. For example, there is the possibility to increase the PV capacity, and this is suggested in several cases within the category of "best annual cost scenarios" (Section 7.5.3) and "target dependency scenarios" (Section 7.5.5). PV helps to achieve considerable benefits in terms of CO_2 emissions reduction for those scenarios.

Since the study in Appendix A.2 shows the potential effects of increasing the use of wood, it is interesting to investigate a further diffusion of individual wood boilers and the introduction of wood based micro CHPs. The maximal exploitation of wood induces benefits in terms of emissions and energy dependency. Wood resources appear economically very attractive for individual boilers but not for micro CHP because of the high investment cost. However, the lowest dependent scenarios (Table 7.8) make a large introduction of micro CHP technology. Additionally, the technology helps to reduce load following capacity, and therefore assists to increase the technical stability of the energy system. Another interesting solution for the thermal sector is the use of ground source heat pumps; this technology is extensively introduced for most of the target scenarios.

The transport sector could be radically transformed by increasing the use of electric cars. However, the necessary investment to replace fossil-fuels based cars by electric cars is economically unattractive under the current market. As we have seen, a small number of electric cars are introduced for all the groups related to CO_2 emissions reduction (Table 7.7) and for first three groups of energy dependency reduction (Table 7.8). In order to reach the lower dependency values (i.e., last two groups), it is necessary to have a radical change to the transport sector.

7.6 Conclusion

In this work, all the energy demands and productions data that characterize the energy system of Giudicarie Esteriori for the year 2013 (the reference scenario) are collected. Based on this, a precise hourly analysis of the overall energy flows is performed.

Designing optimized scenarios that minimize energy costs, decrease the dependency on foreign resources, improve environmental impacts and integrate renewable energy sources is a multi-objective optimization problem. We formulate two traditional objectives (i.e., CO_2 emissions and annual cost) and two new objectives (i.e., load following capacity and energy system dependency) to quantify environmental, economic, technical and political aspects of an energy system.

To identify optimized scenarios for Giudicarie Esteriori, our framework based on a multi-objective evolutionary algorithm (i.e., SPEA2) and EnergyPLAN is applied. A significant number of optimized scenarios are identified.

From these optimized scenarios, we select a range of "best scenarios" to emphasise different aspects of designing the energy system. Firstly, some less costly scenarios are presented. The least costly scenario is 11% less costly than the reference. Moreover, all these scenarios reduce CO_2 emissions. At the same time, a reasonable improvement is achieved in terms of load following and energy dependency. Secondly, some scenarios belonging to target groups of emissions reduction are presented. A similar procedure is applied to identify scenarios for energy dependency target groups. Several scenarios provide significant improvements in terms of CO_2 emissions and dependency. In addition, it is even possible to reach zero emissions and a system that needs only 11% of external energy resources to cover all the local energy demand for electricity, thermal and transportation.

From the point of view of the investigated technologies, the optimized scenarios show economically attractive potentials for the reduction of CO_2 emissions and dependency through: (1) increasing the capacity of PV, (2) maximizing the exploitation of wood and use for individual wood boilers, and (3) partial electrification of the thermal sector through heat pumps. The transport sector could be profoundly transformed by increasing the use of electric cars but it is currently not cost effective.

Finally, this kind of study can be performed systematically for the policy makers of other regions by considering the following steps: 1) collection of energy demands and productions' data and elaboration of hourly profiles, 2) analysis of a reference scenario, 3) identification of local available renewable resources and corresponding technologies, 4) application of the framework to identify optimized scenarios, and 5) study of the identified scenarios according to the requirements (may compare with the reference scenario). This type of study will help energy planners to design their systems more efficiently and accurately.

In the next chapter, we will show that how the proposed framework can be applied when performing long-term energy planning for a region.

Chapter 8

Long-term energy planning with multi-objective optimization

I think that the world is in the middle of a huge transition that we have to make to renewable energy. We have to transition away from fossil fuels very, very quickly.

— Josh Fox

8.1 Introduction

In the previous chapter we have demonstrated how optimized scenarios are identified for an energy system with respect to multiple objectives. However, long-term energy planning requires to design multiple scenarios for different time periods. The long term energy planning is necessary for communities to fulfill the European energy and climate targets.

The Covenant of Mayors [5] initiative was launched in 2008 under the measures of European Union climate and energy package [63]. The goal of the initiative is to engage local authorities of towns to achieve EU objectives of lowering green house gases within the year of 2020. The participating authorities had to agree to develop a plan formally called Sustainable Energy Action Plan (SEAP). SEAP includes a set of actions that has to be carries out by the local authorities to achieve the European Unions' energy and climate target. A new revised initiative called *The Covenant of Mayors for Climate & Energy* was initiated in 2015 to develop a new SEAP for 2030 to reduce CO_2 by 40%. Developing the plan is a threefold problem: i) estimating local energy demands for different sectors (by considering future trends), ii) finding local renewable sources, and iii) identifying optimal scenarios that fulfill energy demands by utilizing local renewable sources. The authorities have to make a plan by considering local available resources and energy demands. This requires an extensive analysis of future energy demands for different sectors and a planning for optimal utilization of local renewable resources. As the studied area, named "*Val di Non*", is a signatory local authority for The Covenant of Mayors, it has to prepare a SEAP for different time frames.

Val di Non is an Italian Alpine valley in the north-west of the province of Trento (Figure 8.1). In this study, the energy system of Val di Non is analyzed by using

EnergyPLAN. The goal of the study is to design optimized scenarios to minimize annul cost and CO_2 for three different time periods (i.e., 2008–2020, 2020–2030 and 2030–2050) and compare the scenarios with the reference one (for the year of 2008). Future energy scenarios are designed by the framework proposed in Chapter 3 and by considering future energy demands. After identifying optimized scenarios for different periods, a few scenarios[1] are selected from those optimized scenarios to fulfill specific emission reduction targets (such as scenarios that reduce 50–55% emission with respect to 2008).

Additionally, energy transitions [34] from a state of an energy system to a future optimal state can not be abrupt. It depends on the social and economic status of the region. The region adopts different policies, such as carbon taxation [34], feed-in-tariffs (FIT) [93] and other similar policies, to reach a certain scenarios. Therefore, the scenarios of consecutive periods should not be too different so that the policy makers need not to adopt completely new policies. In this regards, a methodology is proposed for selecting transient scenarios from target scenarios that are suitable for a transition from one time period to next period.

The remainder of the chapter is organized as follows. Section 8.2 presents a territorial description of the studied area. Section 8.3 describes all the details of the reference scenario. All the details of projected fuel prices, demands and electricity grid mix are presented in section 8.4. Section 8.5 presents the simulation results for modeling the reference scenarios and results regarding the identification of optimized future scenarios. Corresponding results for identifying target scenarios (in terms of emission reduction) for different time frames are described in Section 8.6. The methodology of selecting scenarios for transition from one period to another and the results are presented in section 8.8. Finally, the chapter is concluded by proving some future directions.

8.2 The studied area

Val di Non (VdN) has a surface area of 597.12 km^2 and it can be considered as a vast valley plateau crossed almost entirely by the river Noce. Rising 268m from Rocchetta and bordered with Val d'Adige, Val di Non has different natural landscapes. The lower valley of the area is dominated by vast apple orchards and the upper valley is used for grazing. The culture of the apple tree has been increasingly intensified in recent years, becoming the main product of the valley (67.4 km^2 land is used for apple production). The whole valley is surrounded by forest, which covers 238.5 km^2, more than a third of the total area. Many lakes, streams and creeks are found throughout the territory. Among the freshwater lakes, the artificial lake named Santa Giustina is the most important resource of the valley; the lake is used for generating electricity, as water resource for apple production and as a touristic attraction.

The population of the area was estimated (in 2014) as 39,500. The population is distributed over 34 small municipalities and about 70% of the population live in the altitude between 600 and 1200m (while the remaining 30% between 500 to 600m). The demographic trend is upward since 1991 with a peak in the years between 2000–2010. The population has grown 11.8% (in total) with an average of 0.5% per year within the year of 1991–2014. According to a projection the population will reach

[1]The scenarios are called "target scenario".

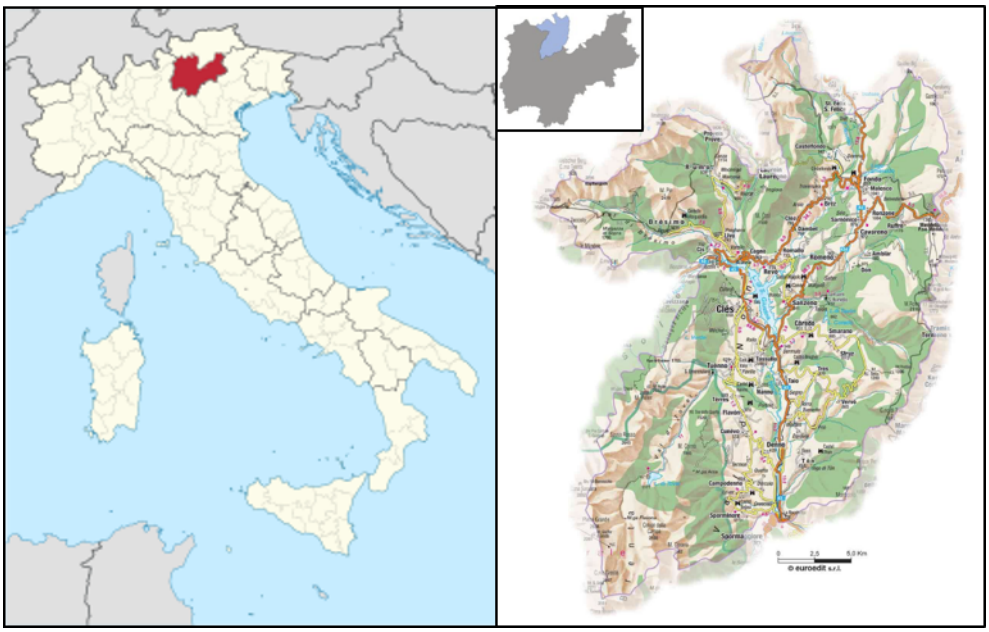


Figure 8.1: The Province of Trento and the area of Val di Non.

under around 49,000 inhabitants in 2045.

8.3 Reference scenario

The energy system of year 2008 is chosen as a reference scenario. The year is chosen as all the data regarding the energy system of "val di Non" is available in that year. In the following sections (8.3.2, 8.3.3 and 8.3.4) electrical, thermal and transport energy demands and productions are calculated and analyzed.

8.3.1 Overview of the energy system

31 municipalities are analyzed to study the energy system of VdN. The total yearly final energy demand for the area is around 632 GWh; in percentage 47% demand is for thermal sector, 26% and 27% are for transport and electrical sectors, respectively. Energy data is partially provided by local distribution system operators (DSO) and partially estimated by the author depending on availability of data.

8.3.2 Electrical demand and production

The yearly electrical demand for VdN (including 31 municipalities) was approximately 163 GWh in the year of 2008[2]. This includes all the sectors from domestic and industrial demands to public lighting demands. The peak demand of electricity is recorded in October which seems unusual in comparison with other valleys.

[2]The data is collected from *SET Distribuzione S.p.A.* (part of Gruppo Dolomiti Energia); electricity distributor in the province of Trento. For the municipality of Cles, distribution data are partially available and are integrated with data from *Azienda Elettrica Comunale di Cles* and *STN Consortium*. For municipalities of Coredo, Smarano, Taio, Tres and Vervò, data is provided by SET.

However, the area is famous for apples and October is apple harvesting time. Many workers come to help harvesting apples that induces some extra electricity demands.

Electricity production of Val di Non is dominated by two kinds of production units: i) hydro-power ii) photovoltaics. The total yearly production from these two sources was around 312 GWh in 2008. The electricity production shows strong fluctuations as hydro-power and photovoltaics are two renewable sources that vary significantly during different months of a year. The installation of PV was very limited in 2008, the installed capacity was only 936 kW (according to Gestore Servizi Energetici) for a total energy production of 1.2 GWh/year. Instead, the hydro-power plant is one of the most important resource of energy for the valley. Located at the heart of Val di Non, the hydro-power plant has an average annual production of around 282 GWh from the main 105 MW plant and 29 GWh from a small secondary 3 MW plant (installed in 2004). The hydro-power electricity production in winter is lower than in other months of a year; the pick production is reached in spring and autumn. The monthly production in 2008 exceeded the monthly demand from April to December; while in the months of January, February and March, it was required to import electricity from the national grid.

8.3.3 Thermal demand and production

Thermal demands are divided into two types of demands: demands for space heating and domestic hot water. Table 8.1 shows the breakdown of thermal demands into different energy sources and different sub-sectors (i.e., residential, industry and other sectors). It is clear from the table that the residential sub-sector has the highest demands and most of the demands are met by oil (41.64%). The reported yearly total thermal demand is approximately 297 GWh. [3].

Table 8.1: Breakdown of the thermal energy demands by sub-sectors and by energy sources.

Energy sources	Residential (GWh)	Industry (GWh)	Municipality (GWh)	Total (GWh)	%
Gas	58.82	25.41	6.17	86.41	19.07
Oil	74.44	44.94	4.39	123.77	41.64
LPG	9.46	-	0.01	9.47	3.19
Wood	68.28	-	3.74	72.02	24.23
Solar	-	-	-	5.59	1.88
Total (GWh)	207.01	70.35	14.31	297.26	

-

The production of thermal energy is covered by individual plants at building level (i.e., wood stoves & boilers, oil, gas and LPG boilers, solar thermal panels). The use of wood stoves & boilers is very widely spread in the area; thanks to the forest and the low cost of wood in the local market. Despite the use of a lot of wood, it can be seen that there is still a strong dependency on fossil fuels (about 74%),

[3]Thermal data are provided by Val di Non Community [4], Dolomiti Reti[6], ISTAT [1]. The methodology for converting yearly data to hourly data described in previous chapter.

whereas renewable energy covers a minor percentage (wood: 24% and solar thermal: 2%).

8.3.4 Energy demand for transportation

The energy demand for transportation is calculated by considering only the cars of the valley. We do not consider other types of vehicles (i.e., trucks, public transport, tractors) since the number of other types vehicles is very small compared to the number of cars. The total number of cars in the area is 19,379 [138]. To calculate the energy demand, it is considered that each car travels around 12,900 km/year [23] on average with an efficiency of 0.684 kWh/km (typical for an internal combustion engine (ICE) car [46]). Therefore, calculated total yearly demand for transportation is approximately 170 GWh. The entire demand is satisfied by fossil fuels: oil, natural gas and LPG.

8.4 Parameters for identifying future optimized scenarios

The main focus of the study is to identify future optimized energy scenarios for the valley. For this purpose, we have divided the time between 2008 to 2050 into three periods (i.e., 2008–2020, 2020–2030 and 2030–2050). For each period, we have identified optimized scenarios by considering corresponding projected demands, prices, efficiencies and other parameters of the energy system. Please note that fuels prices are varied according to different considered cases such as 2^oC and 4^oC (also called IEA 2DS and IEA 4DS). Firstly, optimized scenarios are identified by considering IEA 2DS. In addition, sensitivity analysis has been performed by considering 4^oC cases (please see Section 8.4.1.2). However, future reference scenarios are simulated by considering 6DS case[82, 83] (the details will be discussed in section 8.5.2) In the followings, future projections regarding energy demands, fuels prices and other parameters will be discussed. Finally, the technologies (i.e., decision variables) for designing future scenarios will be presented.

8.4.1 Future projections

Future projections for different parameters are absolutely necessary to design future scenario, as future scenarios have to cope with all new parameters of an energy system (i.e., demands, prices and other parameters). In the following sections, projected energy demands for three sub-sectors are discussed; mainly targeted for the year of 2020, 2030 and 2050. In addition, the prices related to fuels, an investment costs of different technologies are estimated. As the energy system of VdN is not an *energy island* (electricity is exchanged between local grid and nation grid), projecting the future national electricity grid mix is required to calculate emissions of the system. Therefore, in the following we will also present projected future national grid mix for electricity generation.

8.4.1.1 Energy demands

By considering the demand trends presented in [31], electrical and thermal demands are calculated for the year of 2020, 2030 and 2050. However, transport demand is calculated in a different way. It is considered that the distances (in km) traveled by the cars of the valley will be the approximately the same in future. Therefore, the demand is calculated by considering efficiency of ICE car (kWh/km) which will be improved over time.

Table 8.2 and Figure 8.2 show the projected demands for three different sub-sectors. It can observed that there is not much difference found in terms of electrical and thermal demands. The electricity and thermal demands will remain almost constant over the periods. However, the transport demand gradually decreases over time due to improved efficiency of cars.

Table 8.2: Projected demands for three sub-sectors for three different years.

Sectors	Demands (GWh)		
	2020	2030	2050
Electrical	168.17	166.11	170.78
Thermal	253.71	249.85	256.11
Transportation	151.75	137.25	124.25

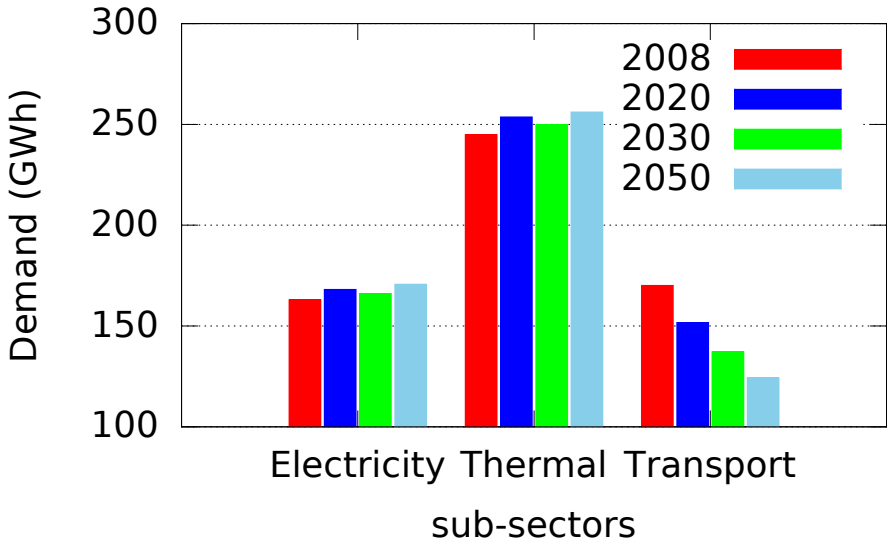


Figure 8.2: Projected demands for three sub-section including reference scenario (in red).

8.4.1.2 Fuel prices

Fuel prices have a significant impact on annual cost of an energy system. In addition, different cases (e.g., 2DS, 4DS) have different impact on prices. In other words, different prices are estimated for considering different cases. In the study, we consider the predictions provided by International Energy Agency (IEA) in Energy

Technology Perspectives 2015 (ETP 2015) [8] and the local prices of fuels [126, 2, 9] to calculate projected prices. Table 8.3 shows the projected prices for different fuels under different considered cases (i.e., 2DS, 4DS and 6DS) for different years. The fuel prices are lowest in the case of 2DS and highest in the case of 6DS, as for 2DS case fuel demands will be low worldwide (due to large adaptation) and opposite for the case of 6DS. For biomass prices are considered the same under all the cases (i.e., 2DS, 4DS and 6DS) since same trends are considered for all the cases.

Table 8.3: Projected Fuel prices.

Fuels	2008 (€/MWh)	2020 (€/MWh)			2030 (€/MWh)			2050 (€/MWh)		
		2DS	4DS	6DS	2DS	4DS	6DS	2DS	4DS	6DS
Diesel for vehicle	133.0	153.09	159.39	162.93	148.72	175.05	195.26	142.88	194.98	234.60
Oil for boiler	122.2	130.23	135.59	138.60	126.51	148.91	166.10	121.54	165.86	199.57
Natural Gas	63.0	84.50	87.48	89.44	80.47	95.37	102.67	72.43	104.04	116.67
Wood/biomass	35.0	37.16	37.16	37.16	39.06	39.06	39.06	43.16	43.16	43.16

Table 8.4: Projected grid electricity prices.

Grid Electricity Price (PUN)	2008 (€/MWh)	2020 (€/MWh)			2030 (€/MWh)			2050 (€/MWh)		
		2DS	4DS	6DS	2DS	4DS	6DS	2DS	4DS	6DS
	87.0	54.75	57.70	58.57	52.34	61.26	64.06	47.82	69.04	76.63

To calculate the variable cost of a system (that contributes to the annual cost), it is also required to consider how much cost is incurred by importing or exporting electricity. Table 8.4 reports the average projected grid electricity prices for different cases and different projected years. Grid electricity price trends are predicted as exact data regarding the trends are not available. A constant increasing or decreasing trends for different cases are considered.

8.4.1.3 National electricity grid mix

The composition of the national electricity grid mix is necessary for analyzing the local system, as the mix is required to calculate CO_2 emissions by considering import and export electricity from th national grid. In this regard it is essential to predict the future energy mixes for electricity generation. The energy mix of 2008 is reported as second row of Table 8.5. The projected renewable energy shares (for 2020, 2030 and 2050) [69] are presented in the second column of the table. The shares of different fossil fuels are calculated by considering the fact that the shares will be decreasing (with respect to share of 2008) according to the increased share of renewable energy (with respect to share of 2008). For example, the renewable energy share is increased by 7.3% from 2008 to 2020. This 7.3% is distributed (by keeping the same ratio of 2008) among fossil fuels' shares and reduced from the shares of 2008. The calculations for 2030 and 2050 are carried out in the same way. Finally, Figure 8.3 illustrates the shares for different years.

8.4.2 Decision variable, objectives and constrains

In this section, the parameters that are used as decision variables to identify future optimized scenarios are described.

Table 8.5: Reference and projected national electricity grid mix in percentage for different years (2008 and renewable shares are from [69]).

Mix / Year	Renewable(%)	Coal(%)	Oil(%)	NGas(%)
2008	28.1	15.6	6.2	50.1
2020	35.4	14.0	5.6	45.0
2030	52.0	10.4	4.1	33.4
2050	85.0	3.3	1.3	10.5

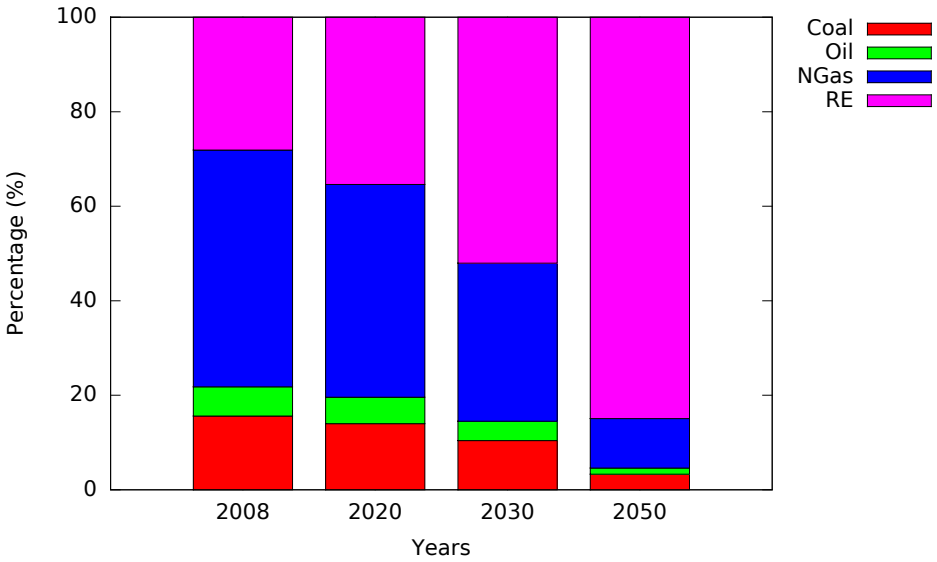


Figure 8.3: Reference and projected national electricity grid mix for different years.

1. PV capacity: There is already some PV plants situated in the valley. In addition, it is still possible to have more PV. It is estimated that the maximum PV capacity can be installed is about 40 MW by considering available roof top area. Hence, the optimized scenarios can have maximum 40 MW of capacity.

2. Existing technologies that are used to meet heat demands (i.e., oil, natural gas and wood boilers) are considered as decision variables. At the same time, ground source heat pump (GSHP) and solar thermal (ST) technologies are considered for investigation.

3. As the valley can offer a lot of wood, we are interested to see how convenient (economically and environmentally) the introduction of wood organic rankine cycle co-generation (wood ORC mCHP) could be for future system.

4. Finally, the transport sector is one of the most emitting sector globally [143]. It is possible to make the sector environmental-friendly by using an alternative energy carrier such as electricity. Therefore, we are interested in investigating the optimal mix between ICE cars and electric vehicle (EV). In this study, a night charging profile (from 21:00 to 4:00) is considered to charge the electric cars.

As already mentioned, there are multiple objectives that need to be optimized. In this study, we select the two most important objectives to minimize: CO_2 emissions and total annual cost. As concern grows about the environment, it is absolutely essential to minimize CO_2 emissions as much as possible. Please note that in this study, total CO_2 emissions are calculated by considering all the emissions of the local system and emission related to importing electricity, excluding electricity export. The reason behind excluding export is that all the local electricity is produced from green sources and the decision makers are only interested in total emissions of the system. At the same time, the costs of the energy system are also an important factor to the community and policy makers, as it is vital for economic aspect. The total cost of a system includes investment cost, operational and maintenance cost, variable cost and additional cost. Except additional cost, other parameters are calculated by EnergyPLAN. Additional cost is the summation of all the extra costs such as grid and metering cost and similar costs.

Finally, there is a good potential for extracting wood in the valley; however, a constraint is set on total collection to maintain a sustainable extraction. 98.84 GWh/year is the maximum limit of wood extraction. The calculation is carried out in the same way as in A.2.

8.5 Identifying future optimized scenarios

As stated in section 1.2 that identifying optimized energy scenarios is a multi-objective optimization problem as there are more than one objectives that need to be optimized. In addition, optimizing an energy system is a discontinuous problem by nature, therefore, an advanced optimization tool is required to handle such a problem. In this regards, the proposed framework in Chapter 6 is employed to deal with the problem. All the corresponding parameters regarding the framework are given in Appendix B.1 (Table B.1 and Table B.2). However, the framework is applied multiple times to identify scenarios for different instances (for different time periods and different cases/IEA perspectives).

8.5.1 Simulation and results

Firstly, the results regarding reference scenarios are presented. Reference scenarios will be used to compare with corresponding optimized scenarios in terms of different energy systems' parameters. Afterward, we present the Pareto-fronts found for three different time periods and two different considered cases (i.e., 2DS and 4DS). Subsequently, some specific emissions target scenarios for different periods will be presented.

8.5.2 Results of reference scenarios

In this section, the results of four reference scenarios (RS) are presented. The first RS is the scenario of the year 2008. In addition, other RS are projected for the year of 2020, 2030 and 2050 by considering that no steps are taken to reduce emissions. Therefore, we consider to have the same technologies of 2008. In other words, all the sectors are kept same as the reference scenario of 2008 in terms of technologies.

As the studied area is not an energy island, we do not need to adjust the capacity of electricity producing technologies for satisfying future electricity demands. Instead, the capacities of technologies of thermal sector are adjusted (increased or decreased by keeping the same ratio of 2008) to meet the projected thermal demands (please see Table 8.2). Efficient ICE cars are used to fulfill the transport demand (please see Table 8.2). Please note that corresponding efficiency values (for different technologies for all sectors) are used to simulate the future reference scenarios.

The considered fuels' cost are estimated according to 6DS case (please see Table 8.3). In addition, the grid projected electricity price (Table 8.4) and grid energy mix (Table 8.3) are also considered according to 6DS case.

Table 8.6: Results of reference scenarios for different time periods.

Year	Reference Scenario	CO_2 emision (kt)	Emission reduction (%)	Annual cost (KEuro)	Cost increased (%)
2008	RS2008	98.09	–	140863	–
2020	RS2020	91.84	6.37	160363	13.84
2030	RS2030	86.05	12.27	167656	19.02
2050	RS2050	82.78	15.60	181145	28.60

Table 8.6 shows the emissions and annual costs for all reference scenarios. It is clearly understood that over time, the emissions of the scenarios will decrease and annual costs will increase. The decreasing of emissions will take place, because of improved efficiencies of boilers and conventional cars; and the greener national energy mix. On the other hand, the annual cost will increase since the fuel prices will be increased in the 6DS case. Fourth and last columns report the percentages of emissions reductions and cost increases for different scenarios with respect to the scenario of 2008. No scenarios can achieve the goals of emissions reduction of Covenant of Mayors, although the costs are increased. Therefore, we need to design some new optimized scenarios that can achieve the goals.

8.5.3 Result of future optimized scenarios

Figure 8.4 presents the comparison of all the Pareto-fronts found by the simulations. The series presented by rectangles represent 2DS case, where different colors present different periods. In addition, the series with filled-triangles represent 4DS case.

Considering general trends of the Pareto-fronts, it is easily understandable from the figure that it will be less costly to introduce renewable energy over time. As an example, a scenario of 2020 has higher annual cost than a scenario of 2030 with same emissions. Considering the 2020 case, it requires more investment cost, hence, more annual cost to get a less emitting scenario. However, considering the 2030 and 2050 cases, less cost is required to reach less emitting scenarios; mostly because of less investment and variable costs. In addition, it will be possible to reach near zero emissions over time (some scenarios of 2050 reach very close to zero emissions).

Considering the Pareto-fronts for 2020, someone may wonder why does not the optimizer identify more less costly scenarios that emit more than 46 ktons (whereas the reference scenario emits around 100 ktons). It is because that behind the particular point (i.e., around 46 ktons for 2020 case), the contradiction between emissions

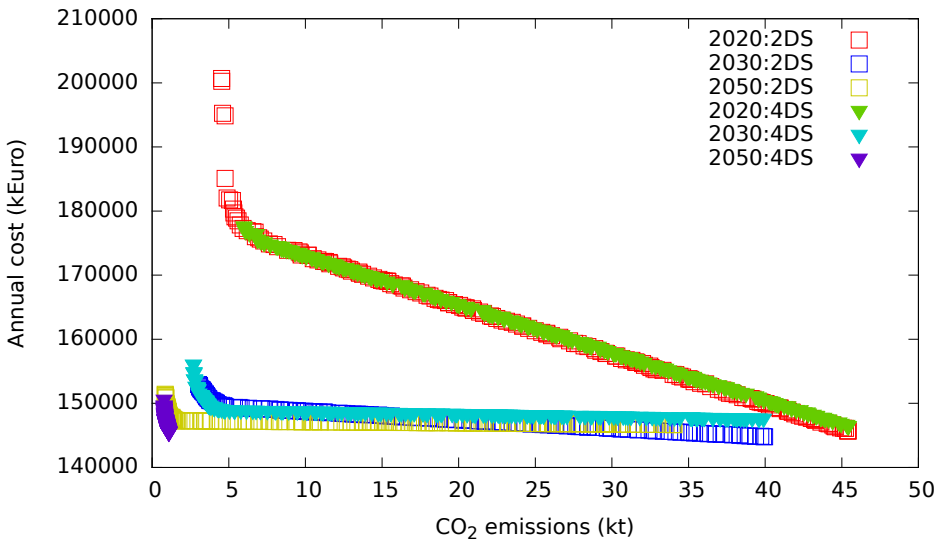


Figure 8.4: Comparison of all Pareto-fronts.

and annual cost breaks down. It means that more emitting scenarios (more than 46 ktons) have more annual costs than the scenarios in the Pareto-front; mainly because of more variable costs. Therefore, those scenarios are dominated by other scenarios in the Pareto-front and these dominated scenarios are not interesting to consider (policy makers are not interested to have costly scenarios that emit more). The same argument is applicable for all other Pareto-fronts.

Now, we will compare Pareto-fronts of different time periods for considered cases (2DS and 4DS). Please note that the only differences between 2DS and 4DS with respect to energy systems parameters are fuel prices (please see Table 8.3) and grid electricity prices (please see Table 8.4). There is virtually no difference between the Pareto-fronts for 2020 for 2DS and 4DS (2020:2DS and 2020:4DS). The reason is that the differences among fuel prices for 2DS and 4DS for 2020 are not very significant (please see Table 8.3). Fuel price for 4DS is little more costly than 2DS. Careful observation of those two Pareto-fronts reveals that more emitting scenarios for 4DS are slightly more costly than 2DS, since more emitting scenarios use more fuels and the use of the fuels increases the variable cost, hence, annual cost. Considering 2030 cases, similar trends as discussed before can be seen as well. The Pareto-fronts intersect around 15kt emissions. Less emitting (less than 15kt) 4DS scenarios are a bit less costly than 2DS scenarios. Please note that all the scenarios are exporting more electricity than importing, because of the big hydro-plant and PV. As the grid electricity price is more for 4DS case, the scenarios of 4DS case earn more than 2DS, therefore, less emitting scenarios are less costly for 4DS case. However, getting towards more emitting scenarios, 4DS optimized scenarios are getting more costly than 2DS; it is because the same reason that the scenarios use more fuels and fuels are costly in 4DS case.

Finally, the most interesting situation can be found for the 2050 scenarios. For 2DS case, we have a Pareto-front with quite a reasonable spread. However, after a certain point, the scenarios are spread almost in a straight line along the X-axis. Hence, many scenarios are found with different emissions but with very similar cost. In addition, for 4DS case, the scenarios are less spread, since there will be

very little contradiction between emissions and annual cost. Therefore, the energy systems with more renewable energies will be less costly than the systems depend on conventional fuels.

8.6 Emission target scenarios

According to European Union's climate action [66], it is recommended to reduce greenhouse gas emissions by 20%, 40% and 80% with respect (wrt) to 1990 emissions levels in 2020, 2030 and 2050, respectively. However, according to the Covenant of Mayors for climate and energy, it is agreed that the reduction of a region will be targeted with respect to recent past year when all energy data of the region is available. In the case of VdN, most energy data is available in 2008, therefore, we will present the target scenarios with respect to year 2008. In the following section, for each time period, we define a reduction range (e.g, 50-55% for 2020) and within the range 10 scenarios will be discussed. If there are more than 10 scenarios are available, we select those 10 scenarios that maximize decision space diversity as proposed in Chapter 4 Paragraph 4.2.2.2. In other words, the scenarios that are different from each other in terms of decision variables will be chosen. On the other hand, we could choose 10 least costly scenarios; however, these scenarios may not differ much in terms of cost as well as decision variables. This is the reason that we propose a scenario selection technique based on maximizing decision space diversity in order to provide a diverse set to the policy makers with a wider range to choose from.

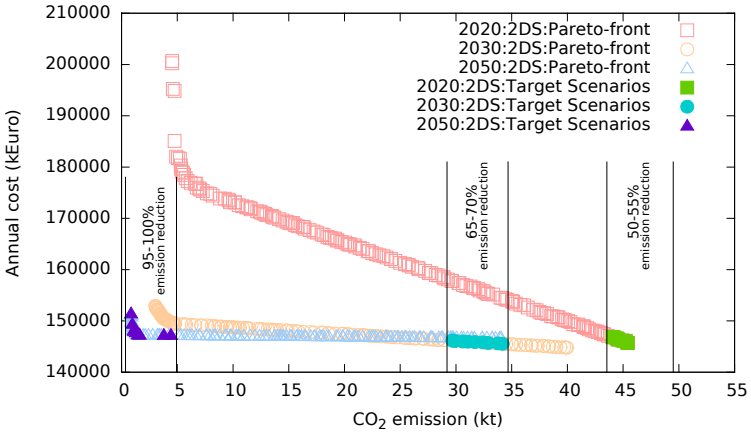


Figure 8.5: Target ranges and scenarios for different time periods for 2DS case.

Figure 8.5 shows the defined emission reduction ranges and corresponding scenarios (on objective space) for different time periods. The faded markers show the same Pareto-front depicted in the previous figure for 2DS case. The ranges are specified by vertical lines and filled markers present the selected target scenarios Please note that there are actually 10, 15 and 58 scenarios within the ranges of 50–55%, 65–70% and 95–100%, respectively and by using the propose method 10 diverse target scenarios are selected within each range. In the following sections, the scenarios for each time frame are analyzed in details.

8.6.1 Target scenarios for 2020

Although it is mentioned that within 2020 emissions should be reduced by 20% with respect to 2008, the Pareto-front (Figure 8.4) shows that the least emitting system achieves 53.69% reduction (i.e., 45.42 ktons of emissions compared to 98.09 ktons in 2008). As mentioned earlier it is not always possible to get scenarios that covers all the ranges (with respect to emissions), thus we decide to present scenarios within the range of 50–55% emissions reduction.

Figure 8.6 presents capacities of 10 selected target scenarios within the defined range. In addition, Table 8.7 reports all the details (annual cost, emissions, capacities of all technologies related to electric and thermal sectors and number of ICE and EV cars) of the scenarios. Moreover, the table shows a comparison with reference scenario (i.e., RS2008 and RS2020) in terms of emissions reductions and cost reduction or increment. The selected target scenarios are 3 to 5% more costly compared to RS2008, however, 8 to 9% less costly when comparing with RS2020 (modeled reference scenario for 2020).

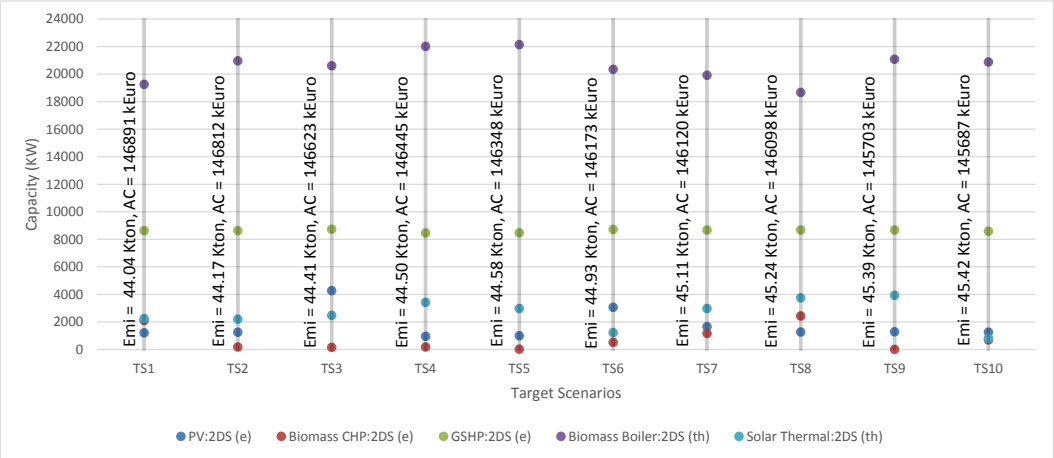


Figure 8.6: Capacity of different decision variables for 10 target scenarios for 2020:2DS.

In terms of decision variables (technologies), little PV introduction is suggested except two scenarios (TS3 and TS5). A reasonable amount of biomass micro CHP is proposed (range: 21 to 2431 kW_e, depending on different scenarios). The small introduction of PV and CHP is explained by the fact that the electricity sector of the territory is already renewable. In contrast to electricity sector, a major change is suggested to the thermal sector. It can be seen from the figure and the table that all the fossil-fuel-based boilers (i.e., oil- and NGas boilers) are recommended to be replaced by heat pumps (around 8.5 MW_e) and biomass boilers (range: around 18.5 to 22.02 MW_{th} depending on different scenarios). Finally, a very small amount of solar tharmal can be introduced (range: 0.8 to 3.9 MW_{th}).

Figure 8.7 shows the annual heat production in percentage in terms of different technologies. GSHP and biomass boilers are the dominant technologies. These two technologies meet most of the heat demands. However, depending on different scenarios, a small amount of heat is produced by biomass CHP and solar thermal.

The transportation sector is slightly transformed. Most of the cars remain as ICE cars. A small numbers of EVs are introduced depending on scenario (limited

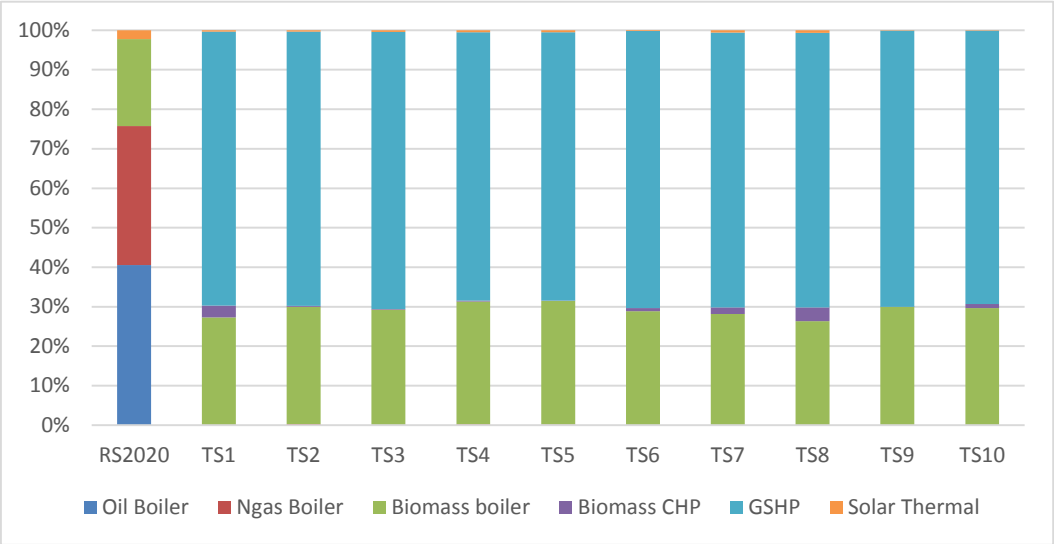


Figure 8.7: Heat production (in percentage) from different technologies for 2020:2DS.

between 7 to 726).

Therefore, it is important to note that the environmental target for 2020 could be reached mainly by transformation of thermal sector with significant intervention for transport sector.

8.6.2 Target scenarios for 2030

Following the 2020 emission range, we present 10 scenarios that achieve 65–70% emissions reduction with respect to RS2008 in 2030. Table 8.8 presents a similar comparison as before for 2030. It should be noted that these scenarios are even less costly than 2020 scenarios (3–4% increase of cost with respect to RS2008 and 12–13% cost reduction with respect to RS2030) because of the reduction of investment cost of renewable technologies. In terms of decision variables regarding electric and thermal sectors, the differences among the scenarios are not so significant. All the variables have very similar values. However, the scenarios are different from each other by the number of EVs are introduced. A moderate number of EVs are introduced; Figure 8.8 presents the mixing of ICE and EVs for those 10 selected scenarios. With higher number of EV, the scenarios have less emissions and vice versa.

In comparison with 2020 scenarios with 2030, there is not much difference can be found in term of suggested technologies in electrical and thermal sector. A further emission reduction could be achieved through the transition in the transport sector. Therefore, it is suggested by analyzing the target optimized scenarios that the transition of transport sector has to begin in this time period (2020 to 2030).

8.6.3 Target Scenarios for 2050

Table 8.9 presents 10 scenarios that can be implemented for 2050 with a target range of 95–100% reduction. It is not possible to reach 100% reduction of emissions as no

Table 8.7: Comparison of target scenarios with reference scenarios for 2020:2DS.

Scn.	CO_2 emissions (ktons)	Red. wrt RS2008 (%)	Red. wrt RS2020 (%)	AC (KEuro)	Inc. wrt RS2008 (%)	Red. wrt RS2020 (%)	Capacities							Number of cars	
							PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kWth)	NGas Boiler (kWth)	Biomass Boiler (kWth)	Solar thermal (kWth)	ICE cars	EVs
RS2008	98.09	-	-	140863	-	-	936	0	0	27463	21275	15983	13310	19379	0
RS2020	91.84	6.37	-	160363	13.84	-	936	0	0	26832	20880	15522	13286	19379	0
TS1	44.04	55.10	52.04	146891	4.28	8.40	1209	2096	8635	1	7	19263	2239	18726	654
TS2	44.17	54.97	51.91	146812	4.22	8.45	1253	174	8639	6	174	20959	2195	18654	726
TS3	44.41	54.73	51.65	146623	4.09	8.57	4275	148	8736	7	21	20613	2477	18910	470
TS4	44.50	54.64	51.55	146445	3.96	8.68	958	173	8463	12	118	22016	3424	18868	512
TS5	44.58	54.55	51.46	146348	3.89	8.74	989	21	8473	12	96	22137	2978	18911	469
TS6	44.93	54.20	51.08	146173	3.77	8.85	3060	515	8719	2	24	20344	1220	19149	231
TS7	45.11	54.01	50.89	146120	3.73	8.88	1652	1167	8679	11	11	19923	2978	19230	150
TS8	45.24	53.88	50.74	146098	3.72	8.90	1262	2431	8680	9	2	18670	3746	19328	52
TS9	45.39	53.73	50.58	145703	3.44	9.14	1279	9	8681	10	1	21082	3933	19328	52
TS10	45.42	53.70	50.54	145687	3.42	9.15	1247	697	8594	10	7	20886	802	19373	7

Table 8.8: Comparison of target scenarios with reference scenarios for 2030:2DS.

Scn.	CO_2 emissions (ktons)	Red wrt RS2008 (%)	Red wrt RS2030 (%)	AC (kEuro)	Inc wrt RS2008 (%)	Red wrt RS2030 (%)	Capacities							Number of Cars	
							PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kWth)	NGas Boiler (kWth)	Biomass Boiler (kWth)	Solar thermal (kWth)	ICE	EVs
RS2008	98.09	-	-	140863	-	-	936	0	0	27463	21275	15983	13310	19379	0
RS2030	91.84	6.37	-	167656	19.02	-	936	0	0	25794	20138	14911	13310	19379	0
TS1	29.70	69.73	67.67	146188	3.78	12.80	1164	258	7531	15	27	20748	903	13742	5638
TS2	29.71	69.72	67.65	146164	3.76	12.82	1164	39	7531	20	21	20964	822	13742	5638
TS3	29.93	69.48	67.41	146092	3.71	12.86	1063	2	7443	3	17	21567	444	13884	5495
TS4	30.62	68.78	66.66	146002	3.65	12.92	1077	64	7431	1	4	21599	367	14271	5108
TS5	31.07	68.33	66.17	145958	3.62	12.94	1100	40	7449	3	17	21491	985	14509	4871
TS6	31.70	67.68	65.49	145882	3.56	12.99	1726	43	7467	5	15	21376	169	14865	4515
TS7	32.42	66.95	64.70	145816	3.52	13.03	1082	318	7494	9	2	20952	731	15252	4128
TS8	32.89	66.47	64.19	145706	3.44	13.09	1071	53	7466	3	5	21389	86	15509	3871
TS9	33.77	65.57	63.23	145611	3.37	13.15	1066	32	7467	18	4	21388	1005	15986	3393
TS10	34.17	65.16	62.79	145531	3.31	13.20	1074	78	7418	4	0	21662	298	16221	3159

Table 8.9: Comparison of target scenarios with reference scenarios for 2050:2DS.

Scn.	CO_2 emissions (ktons)	Red wrt RS2008 (%)	Red wrt RS2050 (%)	AC (kEuro)	Inc wrt RS2008 (%)	Red wrt RS2050 (%)	Capacities							Number of Cars	
							PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kWth)	NGas Boiler (kWth)	Biomass Boiler (kWth)	Solar thermal (kWth)	ICE	EVs
RS2008	98.09	-	-	140863	-	-	936	0	0	27463	21275	15983	13310	19379	0
RS2050	91.84	6.37	-	181145	28.60	-	936	0	0	25880	20227	14923	13286	19379	0
TS1	0.86	99.13	99.07	151431	7.50	16.40	39631	19549	6891	1	6	4242	85998	2	19378
TS2	0.96	99.02	98.95	149408	6.07	17.52	39356	12862	7392	1	28	7327	9158	0	19380
TS3	1.09	98.89	98.81	148157	5.18	18.21	39656	1048	7109	1	1	20447	11671	1	19379
TS4	1.12	98.85	98.78	148046	5.10	18.27	35423	806	7089	7	11	20788	16664	1	19379
TS5	1.29	98.68	98.59	147817	4.94	18.40	34993	33	9299	4	2	7174	249	6	19374
TS6	1.44	98.53	98.43	147340	4.60	18.66	1218	288	7109	3	0	21170	1557	1	19379
TS7	1.55	98.42	98.31	147341	4.60	18.66	2091	239	7470	3	20	18845	583	46	19333
TS8	1.63	98.34	98.22	147282	4.56	18.69	1062	15	9298	1	2	7201	837	2	19377
TS9	3.79	96.13	95.87	147247	4.53	18.71	1019	6	8603	8	24	11695	236	1306	18074
TS10	4.43	95.48	95.17	147232	4.52	18.72	953	35	9298	1	4	7180	730	1663	17717

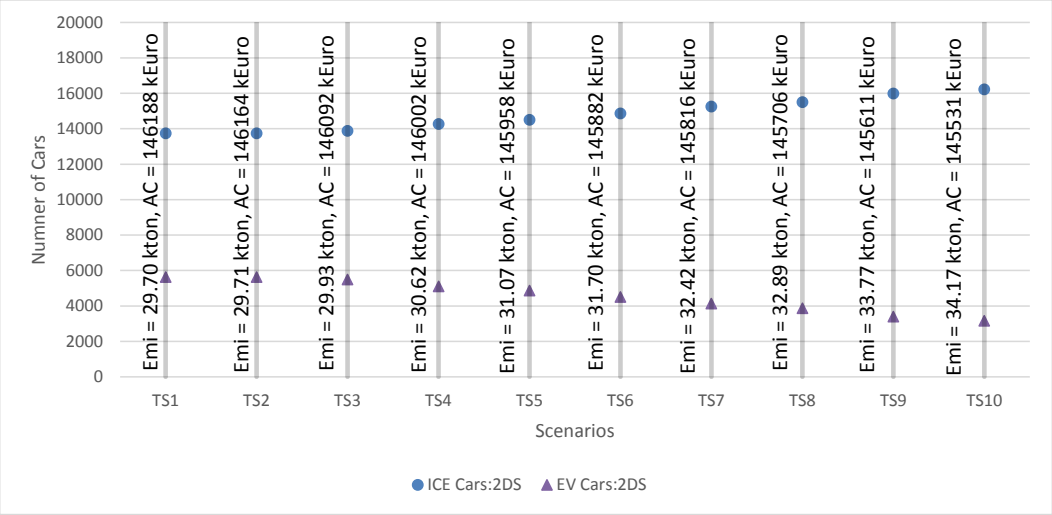


Figure 8.8: Number of Cars (ICE and EV) for 10 target scenarios for 2030:2DS.

scenarios are able to reach the target with the considered technologies [4]. Since no scenarios operate in island mode, therefore, all the scenarios depend on importing some electricity. As the national grid is not completely renewable, there will be always be small amount of emissions by importing electricity to VdN.

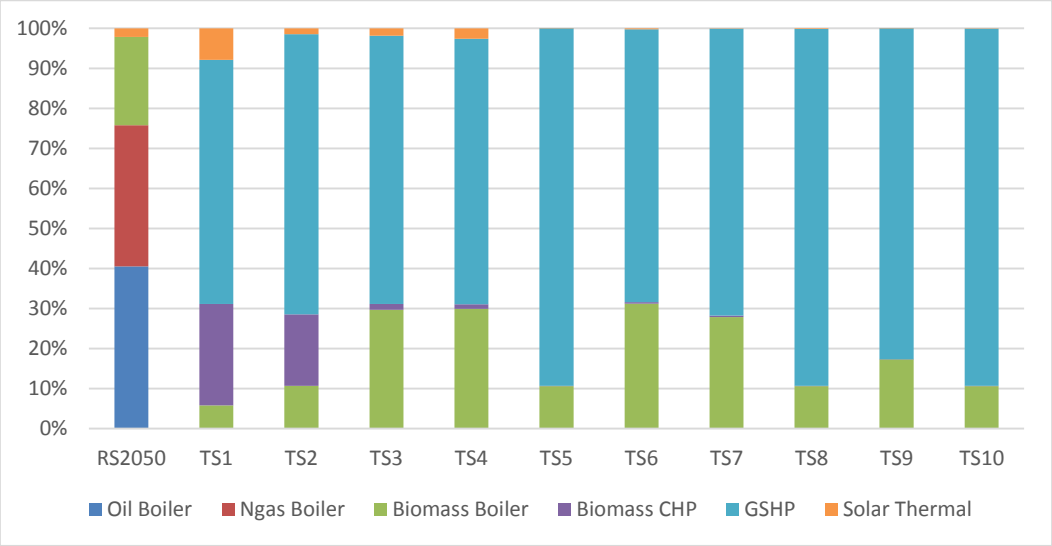


Figure 8.9: Heat production shares from different technologies for 2050:2DS.

Moreover, we have found a diverse set of systems in terms of decision variable/technologies. There are some scenarios where PV reaches to the maximum possible capacity; whereas, some scenarios introduce small amount of PV. Similar patterns can be found for biomass CHP. Biomass CHP reaches very high capacity for near zero emissions scenarios. However, very small CHP is introduced for other scenarios. As we have a constraints on biomass usage, both biomass CHP and boiler compete for the same (limited) resources. When CHP capacity is getting high, boiler capacity is getting low and vice versa. A moderate amount of GSHP introduction

[4]Energy storage is not considered as decision variable.

is suggested (6.8 to 9.2 MW_e), however, when the capacity is relatively low, it is compensated by solar thermal and biomass boilers.

Figure 8.9 presents annual heat production shares with respect to different technologies. It is clear from the figure that GSHP produce most of the required heat to meet the demands. Biomass CHPs and boilers compensate each other in term of heat productions as well. Moreover, solar thermal produces small amount of heat for the selected scenarios.

Finally, the transportation sector is completely transformed by the introduction of electric cars. For most of the scenarios, almost all ICE cars are replaced by electric cars.

In comparison of the 2050 scenarios with 2030, a radical transformation is proposed in the transport sector. Therefore, in this period (2030–2050), the conversion of the sector could lead the community to achieve an almost zero emission target.

8.7 Target scenarios for 4DS case

Table B.3, Table B.4 and Table B.5 present the target scenarios for 4DS case. A very brief discussion about the scenarios is presented. There is not much differences among the scenarios for 2020 and 2030 in comparison with 2DS and 4DS case. However, the scenarios for 4DS case are a little costlier than 2DS because of prices of fuels. For 2050:4DS, PV capacity always reaches maximum. GSHP remain a bit low in capacity compared to 2DS case. In addition, similar behavior can be observed with respect to GSHP, Biomass boiler and solar thermal. Finally, the transport sector is revolutionized by introduction of 10% EV.

8.8 Transition from one scenario to another

In the previous few sections, we have presented some target scenarios for each time period. However, it is not clear which two scenarios should be selected to have a smooth transition from one time period to another (e.g., transition from 2020 to 2030). The transition from one target scenario to another one should be gradual so that no abrupt policy changes are required. Therfore, there should be minimum differences in terms of decision variables/technologies. As we need to satisfy different targets (in terms of emissions) in different time, of course, the scenarios of two different periods will be different. However, the differences should be as small as possible between two selected transient scenarios. Therefore, in next section we will discuss a simple methodology for selecting the scenarios. Moreover, the proposed method is applied on the selected scenarios of the valley and results will be presented in the later sections.

8.8.1 Methodology

Considering $T_1, T_2, \ldots, T_n$ are the n number of sets (i.e., sets containing target scenarios) containing $n_1, n_2, \ldots, n_n$ number of scenarios. Each set T_l contains $\{S_l^1, S_l^2, \ldots, S_l^{n_l}\}$ scenarios. The transition from T_l to T_{l+1} (denoted as $T_l \rightarrow T_{l+1}$) is carried out by finding two scenarios; one from T_l and another from T_{l+1}. The scenarios are selected from the corresponding set based on the minimum distance between them in terms

of decision variables. All the scenarios from the T_l set are compared with all the scenarios from T_{l+1} set based on the distances. The scenarios that have minimum distance among all the pairs, are selected as transient scenarios. Mathematically $T_l \rightarrow T_{l+1}$ is calculated as follows (by considering $n_l, n_{l+1} > 0$):

$$\begin{aligned} A &:= \text{argmin}_{0 \leq i_l < i_{l+1} < \max\{n_l, n_{l+1}\}} \{d(S_l^{i_l}, S_{l+1}^{i_{l+1}})\} \\ &= \{(i_l, i_{l+1}) : d(S_l^{i_l}, S_{l+1}^{i_{l+1}}) \leq d(S_l^{j_l}, S_{l+1}^{j_{l+1}}), 0 \leq i_l, j_l < n_i, 0 \leq i_{l+1}, j_{l+1} < n_{l+1}\} \\ B &:= \min_{0 \leq i_{l+1} < n_{l+1}} \min_{0 \leq i_l < n_l} \{(i_l, i_{l+1}) : (i_l, i_{l+1}) \in A\} \end{aligned} \tag{8.1}$$

Where $d(S_l^{i_l}, S_{l+1}^{i_{l+1}})$ is the Euclidean distance between two scenarios (i_l^{th} scenario from T_l and i_{l+1}^{th} scenario from T_{l+1} set) with respect to decision variables/technologies and the distances are symmetrical. Set A contains all the indice pairs that have minimum distances [5]. Finally, set B contains only one pair of indices. The indices contained in B (i.e., i_l^{th} and i_{l+1}^{th} scenario from T_l and T_{l+1}, respectively) are considered as transient scenarios.

The previous formulation is only applicable for a single-stage transition (e.g, $2020 \rightarrow 2030$), however, it is also possible to extend the process by considering multi-stages transitions (i.e., n-stages). To design a multi-stage transition ($T_1 \rightarrow T_2 \rightarrow \cdots \rightarrow T_n$), it is required to minimize the summation of the all a 1-stage distances. This is essentially a comparison among all possible pairs contained within the sets. As an example, to find $T_1 \rightarrow T_2 \rightarrow T_3$, all pairs of $T_2 \rightarrow T_3$ are considered for each pair of $T_1 \rightarrow T_2$. Mathematically $T_1 \rightarrow T_2 \rightarrow \cdots \rightarrow T_n$ is calculated as follows (by considering, $n_1, n_2, \ldots, n_n > 0, n \geq 2$):

$$\begin{aligned} A &:= \text{argmin}_{\substack{0 \leq i_l < i_{l+1} < \max\{n_{i_l}, n_{i_{l+1}}\} \\ 1 \leq l \leq n-1}} \left\{ \sum_{j=1}^{n-1} d(S_j^{i_j}, S_{j+1}^{i_{j+1}}) \right\} \\ B &:= \min_{0 \leq i_n < N_n} \min_{0 \leq i_{n-1} < N_{n-1}} \cdots \min_{0 \leq i_1 < n_1} \{(i_1, i_2, \ldots, i_n) : (i_1, i_2, \ldots, i_n) \in A_k\} \end{aligned} \tag{8.2}$$

Where $i_1, i_2, \ldots, i_n$ are the indices of the transient scenarios from $T_1, T_2, \ldots, T_n$ sets, respectively.

8.8.2 Results

In the current section, we will identify transition scenarios for VdN. In the previous few sections (section: 8.6.1, 8.6.2 and 8.6.3), we have selected sets of target scenarios for different time periods. We would like to identify scenarios for transition from 2020 to 2030 ($2020 \rightarrow 2030$) and 2030 to 2050 ($2030 \rightarrow 2050$).

For each 2020 scenario (in Table 8.7), we identify a 2030 scenario using equation 8.1. Please note that all the values of decision variables are normalized to calculate the distance between scenarios. Each two rows of the Table 8.10 shows the

[5] This is unlikely that there will be more than one set of pairs that have same minimum distances, however, in theory it may be happen. Therefore, the next formulation is provided that will return only one set of pairs.

Table 8.10: Energy transition 2020 → 2030 for all 2020 scenarios.

Transition from 2020 to 2030	Capacities							Number of Cars		Normalized distance
	PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kWth)	NGas Boiler (kWth)	Biomass Boiler (kWth)	Solar Thermal (kWth)	ICE cars	EVs	
TS1 to TS8	1209	2096	8635	1	7	19263	2239	18726	654	0.572
	1071	53	7466	3	5	21389	86	15509	3871	
TS2 to TS6	1253	174	8639	6	174	20959	2195	18654	726	1.074
	1726	43	7467	5	15	21376	169	14865	4515	
TS3 to TS7	4275	148	8736	7	21	20613	2477	18910	470	0.607
	1082	318	7494	9	2	20952	731	15252	4128	
TS4 to TS1	958	173	8463	12	118	22016	3424	18868	512	0.771
	1164	258	7531	15	27	20748	903	13742	5638	
TS5 to TS1	989	21	8473	12	96	22137	2978	18911	469	0.696
	1164	258	7531	15	27	20748	903	13742	5638	
TS6 to TS8	3060	515	8719	2	24	20344	1220	19149	231	0.602
	1071	53	7466	3	5	21389	86	15509	3871	
TS7 to TS7	1652	1167	8679	11	11	19923	2978	19230	150	0.589
	1082	318	7494	9	2	20952	731	15252	4128	
TS8 to TS7	1262	2431	8680	9	2	18670	3746	19328	52	0.600
	1082	318	7494	9	2	20952	731	15252	4128	
TS9 to TS7	1279	9	8681	10	1	21082	3933	19328	52	0.580
	1082	318	7494	9	2	20952	731	15252	4128	
TS10 to TS7	1247	697	8594	10	7	20886	802	19373	7	0.551
	1082	318	7494	9	2	20952	731	15252	4128	

details of the transient scenarios. First column shows the names of the transition scenarios, for example, *TS1 to TS8* refers that the scenario named *'TS8'* of 2030 (see Table 8.8) is the best match for the 2020 scenario named *'TS1'* (see Table 8.7). Moreover, the last column shows the normalized distance between the scenarios in terms of decision variables. Please note that different 2030 scenarios are selected for different 2020 scenarios. The reason is that different 2020 scenarios have different values for decision variables; based on the values, a 2030 scenario is chosen that has minimum distance with the particular 2020 scenario. Table 8.11 shows the details of the transition 2030 → 2050.

In last few paragraphs we have showed the results for 1-stage transition, however, a multi-stage transition is required to do a long-term planning. In the case of VdN, a 2020 → 2030 → 2050 transition planning can be done based on the available targeted optimized scenarios. Using equation 8.2, a double-stage transition planning is performed. The results are reported in Table 8.12. For each 2020 scenario, the best 2030 and 2050 scenarios are presented (each row presents transient scenarios for each 2020 scenario). Please note that the results presented in Table 8.12 are not combined results from Table 8.10 and Table 8.11. The later results minimize the summation of 2-steps distances (distances from 2020 to 2030 scenarios and distances from 2030 to 2050 scenarios) while the other results minimize only 1-step distances. The results shows that there are mainly two scenarios (i.e., TS3 and TS7) in 2030 which are more compatible with scenarios of 2020; while single 2050 scenario (i.e., TS7) is compatible in this case. The last column of the table reports the normalized total distance for each transition. Finally, the blue shaded row presents best possible transition scenarios where less differences are required among different technologies/decision variables.

Figure 8.10 illustrates all the target scenarios and best transition scenarios in objective space. All the boxes present optimized scenarios for different time periods

Table 8.11: Energy transition 2030 → 2050 for all 2030 scenarios.

Transition from 2030 to 2050	Capacities							Number of Cars		Normalized distances
	PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kWth)	NGas Boiler (kWth)	Biomass Boiler (kWth)	Solar Thermal (kWth)	ICE cars	EVs	
TS1 to TS7	1164	258	7531	15	27	20748	903	13742	5638	1.189
	2091	239	7470	3	20	18845	583	46	19333	
TS2 to TS9	1164	39	7531	20	21	20964	822	13742	5638	1.300
	1019	6	8603	8	24	11695	236	1306	18074	
TS3 to TS7	1063	2	7443	3	17	21567	444	13884	5495	1.022
	2091	239	7470	3	20	18845	583	46	19333	
TS4 to TS6	1077	64	7431	1	4	21599	367	14271	5108	1.056
	1218	288	7109	3	0	21170	1557	1	19379	
TS5 to TS7	1100	40	7449	3	17	21491	985	14509	4871	1.067
	2091	239	7470	3	20	18845	583	46	19333	
TS6 to TS7	1726	43	7467	5	15	21376	169	14865	4515	1.096
	2091	239	7470	3	20	18845	583	46	19333	
TS7 to TS7	1082	318	7494	9	2	20952	731	15252	4128	1.165
	2091	239	7470	3	20	18845	583	46	19333	
TS8 to TS7	1071	53	7466	3	5	21389	86	15509	3871	1.141
	2091	239	7470	3	20	18845	583	46	19333	
TS9 to TS9	1066	32	7467	18	4	21388	1005	15986	3393	1.398
	1019	6	8603	8	24	11695	236	1306	18074	
TS10 to TS6	1074	78	7418	4	0	21662	298	16221	3159	1.193
	1218	288	7109	3	0	21170	1557	1	19379	

Table 8.12: Transition 2020 → 2030 → 2050 for all the 2020 scenarios.

Scenario	Transition		Normalized
2020	2030	2050	total distance
TS1	TS3	TS7	1.665
TS2	TS3	TS7	2.122
TS3	TS3	TS7	1.714
TS4	TS3	TS7	1.958
TS5	TS6	TS7	1.882
TS6	TS3	TS7	1.686
TS7	TS7	TS7	1.753
TS8	TS7	TS7	1.765
TS9	TS7	TS7	1.745
TS10	TS7	TS7	1.716

(different colors represent different time periods). Down-pointing triangles present the identified target scenarios (different colors represent different time periods). Red colored down-pointing triangles represent the best scenarios for 2020 → 2030 → 2050 transition. Finally, the two arrows show the transitions.

8.9 Conclusion

In this chapter, we have shown how our proposed framework can be applied when performing long-term energy system planning. As most of the local European communities are signing the Covenant of Mayors, the communities are bound to reduce progressively their CO_2 emissions, following reduction targets at 2020, 2030 and

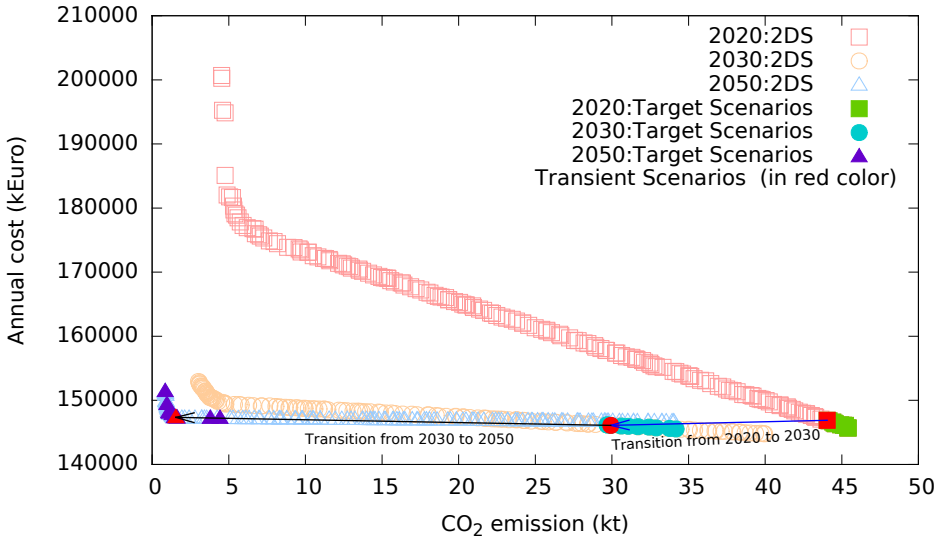


Figure 8.10: Transition 2020 → 2030 → 2050 shown in objective space

2050. Therefore, it is necessary to provide the long term energy planing for corresponding regions. We take an example of Val di Non and applied our proposed framework to find out optimized scenarios for different time periods. Afterwards, we have identified some target scenarios (in term of CO_2 emissions) from the optimized scenarios, since the community has to reach different emissions goals in different time periods.

Decision makers may choose any scenario within the target sets, however, the scenarios of consecutive time periods should be consistence to each other (i.e., should not being too different from each other). In this regard, we propose a technique that can identify best possible scenarios to have a smooth transition from one time period to another. The technique can not only be applied for single-stage transition (i.e., 2020 → 2030), but also be applied for multi-stage transition (i.e., 2020 → 2030 → 2050). We have applied the proposed technique on Val di Non target scenarios and different transient scenarios are identified.

In the next chapter we will propose modified MOEAs that can deal with given preferred (targeted) regions. In the precious two chapters, we have manually identified some target scenarios from the optimized scenarios (after finishing the optimization phase). However, with the help of the proposed technique, the target regions can be set before the optimization phase and the modified MOEA will identify only the scenarios within the regions.

In both this and the previous chapter, an approach is considered where we have manually selected some target scenarios from identified optimized scenarios (after the identification of optimized scenarios). In the next chapter we will propose modified MOEAs that can deal with the problem of identification of solutions in given preferred regions. With the help of the proposed technique, it would be possible that the target regions are set before the optimization phase and the modified MOEA will identify the scenarios only within the specified regions.

Chapter 9

Multi-objective optimization with multiple preferred regions

We are trying to prove ourselves wrong as quickly as possible, because only in that way can we find progress.

— Richard P. Feynman

9.1 Introduction

For the last two decades, multi-objective evolutionary algorithms (MOEA) have been successfully used to solve multi-objective optimization (MOO) problems. Typically, the goal of a MOEA is to find a set of trade-off solutions that are well-distributed over the objective space. Sometimes, however, the decision makers/users have little interest in exploring the entire objective space. It may be more interesting to them to explore some preferred regions of a front due to market demands, due to financial pressure, or simply due to curiosity.

Let us consider an energy system optimization problem, where the goal is to minimize CO_2 emission and annual cost. In this case, a reference system is analyzed (generally the current system), optimized systems are identified, and then compared with the reference system. Decision makers are often interested in exploring several regions defined by either CO_2 emission or annual cost. For example, if a reference scenario has x amount of CO_2 emissions, the interesting regions could be 10%–20% and 35%–40% reduction of x (a manual approach is considered in Chapter 7 and 8).

There are a few algorithmic advantages in exploring preferred regions over exploring the entire objective space. These include faster convergence speed and better approximation of Pareto-front. Based on the idea of incorporating user preference, a wide range of different concepts and algorithms has been proposed (see [172, 27] for comprehensive surveys, as it is impossible to list all relevant work here), such as

M. S. Mahbub, M. Wagner, and L. Crema, "Multi-objective optimisation with multiple preferred regions," in proceedings of *Artificial Life and Computational Intelligence: Third Australasian Conference, ACALCI 2017*. Springer, pp. 241–253 [online] `http://dx.doi.org/10.1007/978-3-319-51691-2_21`.

The text presented in chapter 9 is revised versions of the articles. Most of the sections of the chapter are paraphrased and extended for better understanding to the readers.

(i) defining reference-point(s) [144, 55, 148] and specifying weights in the objective space [123, 74]. However, the problem with these approaches typically is that it is difficult to set the corresponding parameters without knowing the shape of the true Pareto-front. Therefore, we propose modifications of generic algorithms that require only very intuitive preference encoding in the form of intervals, as outlined above in the energy system example above. The modified algorithms are tested on seven benchmark problems. The results show that the algorithms perform than generic ones. Finally, one of the algorithms is applied on energy system optimization problem with satisfactory results.

The structure of this chapter is as follows. First, we introduce the idea of preferences for MOEAs in Section 9.2. In Section 9.3 we show how we integrate preference information in two algorithms. Lastly, we present and discuss the results of our experimental studies in Section 9.4 and Section 9.5.

9.2 Basic principles

There is a fundamental difference between how reference points and weights are defined, and how our preferred regions are defined. A reference point is defined by specifying values for all the dimensions for a point in the objective space. In contrast to this, a preferred region is defined by specifying an upper and lower bound of one particular dimension. For example, if a two-objective problem (objectives are plotted along x and y-axes) is considered, a user can define three preferred regions by setting three upper and lower bounds for intervals along either the x or y axis. Figure 9.1 illustrates an example of three preferred regions (bounded by three different color vertical lines). In this figure, we also show three reference points (gray crosses) which might have been set by a user. As the user lacks knowledge about the shape of the front, these points are not on the true Pareto-front. Consequently, it is left up to the MOEA to follow its own interpretation of "closeness" in order to distribute the solutions around the reference. One outcome is that the solution density is high near the reference point and the density decreases with increasing distance (see Figure 9.1 for the outcome of one run by r-NSGAII [55] using the shown reference points). As mentioned before, it is also possible to use weights in the objective space in order to encode preferences. In Figure 9.1 we indicate this using a color gradient along the x-axis, where a preference for smaller x-values is encoded. The preference formulation for a single objective using weights is relatively simple, however, the formulation becomes tricky when multiple objectives are preferred, and it becomes very complicated when reference points or preferred regions are to be encoded (using weights). The concept of our interval-based regions, on the other hand, is straightforward to use even for laypeople. To the best of our knowledge, even though there are lots of similar approaches, this is the first time this rather simple concept of intervals along axes is used in the context of MOO.

9.3 Preferred regions for different MOEAs

In the following sections, we present the ideas related to preferred regions and the adaptation of the ideas into different MOEAs. In the first section, preferred region based NSGAII (pNSGAII) is presented. In the second section, two variants of AGE

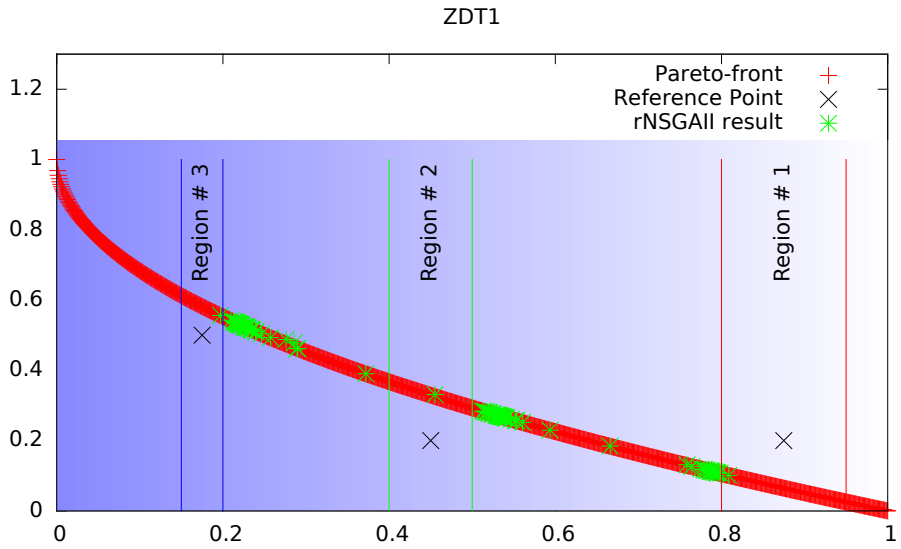


Figure 9.1: Reference points, weights in the objective space, and preferred regions.

(pAGEonline and pAGEoffline) are presented.

9.3.1 Ideas adopted in pNSGAII

We adopt several ideas in NSGAII [53] resulting in pNSGAII. Algorithm 6 presents the main loop of the proposed pNSGAII. There are couple of modifications with respect to the original NSGAII. Firstly, each solution is associated with a particular region (step # 4). Secondly, a modified parent selection procedure is used (step # 8); thirdly, the individuals of a merged population (containing solutions of the previous generation and the offspring) are associated with nR regions (step # 13) such that for each region, $2 * \alpha_i$ individuals are associated. Based on the association, the merged population is divided into nR sub-populations (step # 14). Sub-populations are used to increase the likelihood of achieving the targeted α_i well-distributed solutions per preferred region. Lastly, a modified ranking procedure (step # 16) is applied to rank the solutions of the sub-populations.

The first important addition to pNSGAII is the association of solutions to regions. Before the optimization, the user provides the preferred number of solutions associated with each region. During the optimization, the solutions are assigned greedily to the regions based on the distance between solutions and regions (see Algorithm 7). The distance from a solution (that is outside of the region) to a region is calculated as $min(|f_i(A) - R_u|, |f_i(A) - R_l|)$, where $f_i(A)$ is the objective value (the objective dimension on which a user specifies the ranges) of a solution. The distance is 0 if a solution is inside a preferred region. Finally, Figure 9.2 illustrates an example of Algorithm 7 in step # 13 of Algorithm 6. In the example, $\alpha 1$ and $\alpha 2$ are set to 6; which refers that 12 solutions will be associated with each region. In the figure, two regions are bounded by different colors vertical lines and different colored markers show the solutions associated with different regions.

In the proposed parent selection procedure, there are two ways the parents can be selected. Firstly, a parent can be selected from the same region that the procedure is currently working on (z^{th} region, step # 6 of Algorithm 6). Secondly, the parent

Algorithm 6 Main loop of pNSGAII

1: nR ▷ Number of given regions
2: α ▷ A set containing user-defined preferred number of solutions for each region; α_i is the number of preferred solutions for i^{th} region
3: Initialize population P with $\sum_{i=1}^{nR} \alpha_i$ random individuals and $O \leftarrow \emptyset$
4: Associate α number of solutions with nR regions
5: **while** Stopping criteria not met **do**
6: **for** $z \leftarrow 1$ to nR **do**
7: **for** $j \leftarrow 1$ to $\alpha_z/2$ **do**
8: Select two parents using modified parent selection procedure
9: Generate offspring and add to O
10: **end for**
11: **end for**
12: $P \leftarrow P \cup O$
13: Associate $2 * \alpha$ number of solutions with nR regions
14: Divide P into nR sub-populations ($SP_i; i = 1, \ldots, nR$)
15: **for** $i \leftarrow 1$ to nR **do**
16: Rank SP_i and select α_i solutions based on ranking and crowding distance
17: Add these solutions to SP_i
18: **end for**
19: $P \leftarrow \bigcup_{i=1}^{nR} SP_i$
20: **end while**

can be selected from other regions. A user defined probability (P_{ps}) determines which way the parent will be selected. The details of the algorithm are presented in Algorithm 8. This approach enables us to prevent an algorithm from getting trapped on a local multi-dimensional front (which we have observed in preliminary experiments) by considering outliers [1].

However, when a parent is selected from the currently working region (z^{th} region of step # 6 of Algorithm 6), a tournament selection based approach is adopted. To select a parent, a given number of tournaments are played between randomly selected associated individuals (associated with z^{th} region). The winner of the tournaments is chosen as a parent. The winner (between two randomly selected individuals) of a tournament is decided based on the following criteria:

- Distance from a given region
- Overall constraints violation [49]
- Dominance Relation

The order of the criteria is strictly followed. Therefore, if a solution is closer than another solution with respect to a given region, then the following two criteria are not considered (even 2^{nd} solution violates constrains more than 1^{st} or 2^{nd} solution dominates 1^{st} solution). The overall constraint and dominance relations come into play when two solutions are within the given region.

[1] We will refer to solutions as being *outliers* if they are not within preferred regions.

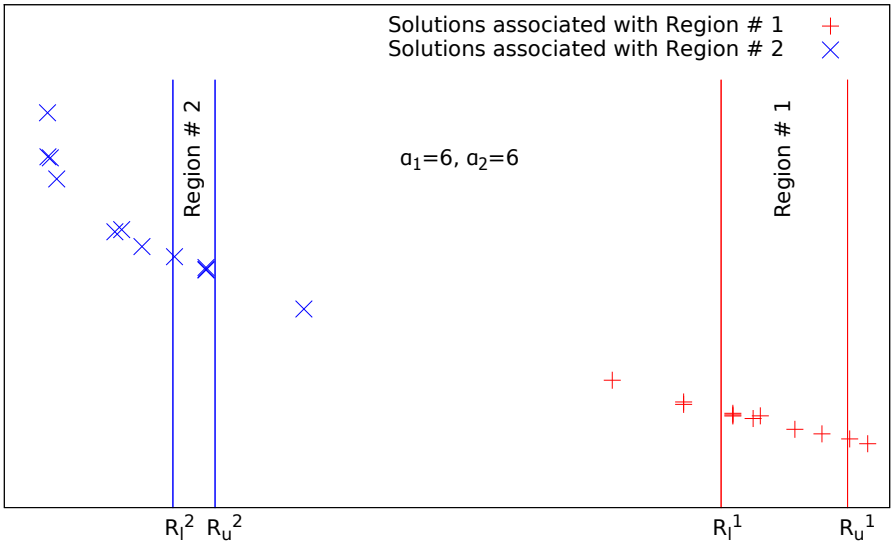


Figure 9.2: Association of solutions with regions.

Ranking in the original NSGAII is performed based on dominance. Individuals are divided into different fronts depending on the relative dominance relationship among the individuals. Therefore, the individuals which are not dominated by any other individuals are placed in the first front. In addition, the second front contains the individuals that are dominated by minimum number of individuals and so on. However, for pNSGAII we propose a ranking procedure based on dominance relations and closeness of an individual to the preferred regions (same ordered criteria as in parent selection). We do not apply the proposed ranking procedure in all generations, as narrowing down the search to some particular regions from the beginning may be problematic. Therefore, for a particular generation (Step # 16, Algorithm 6), only one of the two ranking procedures (i.e., default NSGAII ranking procedure and proposed ranking procedure) is applied with probability $p_{rk} = \left(\frac{usedBudget}{totalBudget}\right)^n$. The shape of this schedule can be controlled through the exponent n. For $n = 1$, the probability of applying the proposed ranking procedure is increased linearly over time. For larger values of n, the probability increases sharply in later stages of the algorithm's run.

In Step # 14, we propose to sub-divided the population into nR number of population (we call it local sub-population). Solutions in a local front (front in local sub-population) have to be non-dominated in term of other local fronts as we want all non-dominated solutions in final output. A solution may be non-dominated in local sub-population, however, may be dominated by a solution of other sub-populations, has to be eliminated. Therefore, in order to rank a sub-population, we not only compare a solution within the local sub-population, but also compare the solution against other solutions of other sub-populations (we call it global population). Therefore, after ranking the local population as usual, each individual of local population is compared (i.e., according to dominance) against all the global individuals. If some local individuals are in the same local front and the individuals are dominated by the same number of global individuals (or not being dominated), then the local individuals will be in the same front. However, if the local individuals are dominated by different number of global individuals, they will be in different

Algorithm 7 Associating regions to solutions

```
nR                                          ▷ Number of given regions
R_u, R_l                                    ▷ Upper and lower bounds of given regions
S                                           ▷ A set containing all the solutions
cα ← 0                                      ▷ Current number of associated solutions with regions
for s ∈ S do
    for i ← 1 to nR do
        if s is within R_u^i and R_l^i And cα_i < α_i then
            Associate s with i^th region; cα_i ← cα_i + 1 and exit the loop
        end if
    end for
    if s is not yet associated then
        Depending on the status of cα_j, associate s with the region j that has
    minimum distance to s (if cα_j = α_j, associate s with the region that has second
    least distance)
        cα_j ← cα_j + 1
    end if
end for
```

Algorithm 8 Parents selection procedure

```
P_ps                                        ▷ Probability of selecting a parent from other region
r                                           ▷ A random number drawn from uniform distribution
if r > P_ps then
    Return a parent based on tournament selection and ordered criteria
else
    Return a random parent from a different region
end if
```

fronts depending on number of global dominators. Therefore, the local individual with smaller numbers of global dominators will be placed in a lower ranked front, and with higher numbers of global dominators will be placed in an higher ranked front. With this approach, the local individuals that are dominated within the global population will have smaller chances to stay in the next generation.

9.3.2 pAGE

The algorithm Approximation-Guided Evolution (AGE) [29] in its original formulation uses an archive A in which it maintains a list of all non-dominated solutions seen. This archive can grow and thus slow down the algorithm. In its newer version, AGE maintains an archive A^ϵ that is an ϵ-approximation of all non-dominated solutions encountered [170]. In the following, we present two straightforward uses of the archive to guide the optimisation towards preferred regions (see Algorithm 9):

- pAGEonline largely corresponds to any of the above-mentioned AGE variants. After the generation of the offspring set O based on the population P, AGE would normally proceed to consider the union $P \cup O$ and then reduce this set greedily to approximate the archive. At this point, we insert one action (step # 11): from the union $P \cup O$ we remove each of the solutions that are outside the preferred regions with probability p_r.

Algorithm 9 $(\mu + \lambda)$-Approximation Guided Evolution with preferences

1: Initialize population P with μ random individuals
2: Set archive $A \leftarrow P$
3: **for** each generation **do**
4: Initialize offspring population $O \leftarrow \emptyset$
5: **for** $j \leftarrow 1$ to λ **do**
6: Select two random individuals from P, and apply crossover and mutation
7: Add new individual to O, if it is not dominated by any individual from P
8: **end for**
9: Insert each offspring in the archive A
10: Add offspring to population, i.e., $P \leftarrow P \cup O$
11: *[pAGEonline]* remove each outlier from P with p_r
12: **while** $|P| > \mu$ **do**
13: Remove p from P for which the approximation of A by P is the smallest when p is left away
14: **end while**
15: **end for**
16: *[pAGEoffline]*
17: 17.1: Remove all outliers from archive A.
18: 17.2: $P \leftarrow A$.
19: 17.3: if $|P| > \mu$ then apply lines 14–16 to reduce the P

- pAGEoffline uses the original AGE with the non-approximated archive for the entire optimisation process. The preference selection only happens once in the very end as a post-processing step using already existing methods. First, pAGEoffline removes from the archive A^ϵ all outliers, and it assigns a copy of this reduced archive to the population P. Since P might be larger than the desired population size, the original AGE's internal mechanism from steps # 12–14 is used so that P approximates A^ϵ well.

For pAGEonline, the probability p_r can be static for the entire optimisation or it can change dynamically according to some predefined schedule. In preliminary experiments we observed that a static choice of $p_r = 1$ can be problematic, as this always removes all outliers. As an alternative to this we decided to increase p_r with the number of evaluations performed, which is similar to our approach in pNSGAII. In additional preliminary experiments, we noticed that even a linear schedule such as $p_r = \frac{\text{usedBudget}}{\text{totalBudget}}$ can prevent pAGEonline from finding all regions. As a consequence, we decided to lower the initial "pressure" by using $p_r = \left(\frac{\text{usedBudget}}{\text{totalBudget}}\right)^n$. If $n = 10$ is considered, the pressure remains low for a long time which allows pAGEonline to find the front. For example, when 90% of the runtime has been reached the probability of removing an outlier is just $p_r = 0.35$, and at 95% it is $p_r = 0.60$. Then, in the last couple of generations, the pressure increases quickly and pAGEonline can focus on spreading out the solutions within the preferred regions. Note that this schedule is by no means optimal, however, it performs well enough to demonstrate the feasibility of the proposed approach.

9.4 Experimental study

We conduct a range of experiments to analyze the performance of our proposed algorithms pNSGAII, pAGEonline, and pAGEoffline. The benchmark problems include five two-dimensional benchmark problems from the ZDT family [179] and two three-dimensional problems from the DTLZ family [54]. To the best of our knowledge, no directly comparable algorithms for multiple preferred regions are available from the literature; algorithms that consider reference point(s) have a different goal, which puts them at a disadvantage by definition (see Section 9.2). Therefore, we compare our approaches with their original algorithms, and we vary the evaluation budgets and population sizes to investigate the effectiveness.

9.4.1 Evaluation metrics

Over the years, a number of evaluation metrics have been proposed. We use widely used metrics available in the jMetal framework, such as, the covered hypervolume (HV) [177], inverted generational distance (IGD) [37] and additive ϵ-approximation (EPS) [181] to measure the performance of the MOEAs. We use them with a simple modification, i.e., separately for each preferred region. Therefore, indicators' values are measured for each preferred region. As the true Pareto-front is required for the calculation of IGD and EPS values, we use those provided by the jMetal framework. From the true fronts, we extract the regional fronts from the original ones by considering only the solutions that are inside the given regions. To calculate HV values, we define the reference point for each region to be based on the extreme values in the preferred region. For example, in the introductory Figure 9.1, these reference points are (0.2, 0.62), (0.5, 0.38), and (0.95, 0.1).

It is important to note that performance indicators for preference-incorporating algorithms exist (e.g. [120, 79]), however, these are for reference point-based approaches and thus not applicable.

9.4.2 Experimental setup

We developed all algorithms in the jMetal framework [58]. Initially, each pMOEA variant is tested on the ZDT family with two configurations, based on population size μ and maximum function evaluations (FE). Table 9.1 presents the different configurations used in the experiments. The configurations are chosen in this way to demonstrate the efficiency of pMOEA in terms of convergence speed.

Problem	Algorithm	μ	EF
ZDT	pMOEA	30	12000
	MOEA	30	12000
		100	12000
		100	24000
DTLZ	pMOEA	30	50000
	MOEA	150	49950

Table 9.1: Configurations in terms of population size μ and evaluation budget EF to test the efficiency of the interval-based preferences.

Table 9.2: Parameter settings

Parameter	Value	Used on
$[R_l, R_u]$	[0.80, 0.95], [0.40, 0.50], [0.15, 0.20]	all algorithms
α	[10, 10, 10]	pNSGAII
Crossover	SBX [53]	all algorithms
Crossover probability	0.90	all algorithms
Mutation	Polynomial mutation [53]	all algorithms
Mutation probability	$1/ndv$ [a]	all algorithms
Distribution Index	20	all algorithms
Parent selection	Binary tournament	NSGAII
Number of tournaments	5	pNSGAII
ϵ_{grid}	0.01	pAGE online, AGE
n	10	all pMOEA
P_{ps}	0.20	pNSGAII

[a] number of decision variables

We consider short and long runs with an evaluation budget of $FE = 12000$ and $FE = 24000$, respectively. On the ZDT[2] family, we conduct only short runs of the pMOEAs, and we compare these with short and long runs of the original MOEAs to investigate the efficiency. For the DTLZ[3] functions, we only use a single configuration of the original algorithm since the solutions with an increased population size, are otherwise the solutions would be very thinly spread out over the three-dimensional front.[4]

Table 9.2 presents the other parameters used in the investigations. In this study, the regions are defined in terms of R_l and R_u along the first objective. Moreover, we set $\alpha = [10, 10, 10]$ for pNSGAII so that 10 solutions will be associated with each region - this was chosen arbitrarily, however, it is possible to have different numbers of solutions associated with different regions. The remaining parameters are mostly the default values from the jMetal framework. In addition, $n = 10$ is used for p_{rk} and p_r for pNSGAII and pAGEonline, respectively. $p_{ps} = 0.20$ is used within pNSGAII's parent selection procedure. In pAGEonline, $\epsilon_{grid} = 0.01$ is used for the approximating archive. We run each algorithm independently 100 times and report the averages in the following.

9.4.3 Results and discussion

Firstly, we present in Figure 9.3 an example of Pareto-fronts obtained by different pMOEAs. We can observe that the solutions are concentrated in the user-defined

[2]The ZDT functions are used as provided by the jMetal framework. The number of decision variables is 30 for ZDT1/2/3 and 10 for ZDT4/6.

[3]Number of decision variables is 12 for DTLZ2/3, as set in the jMetal framework.

[4]If we use $\mu = 30$ for typical MOEA (please see Table 9.1) then it is less probable to find adequate number of solutions in preferred regions, that makes it difficult to compare with pMOEA. In addition, compared in terms of FE, MOEA uses 50 less function evaluations than pMOEA only because the number is compatible with μ (no extra function evaluations after completing last generation).

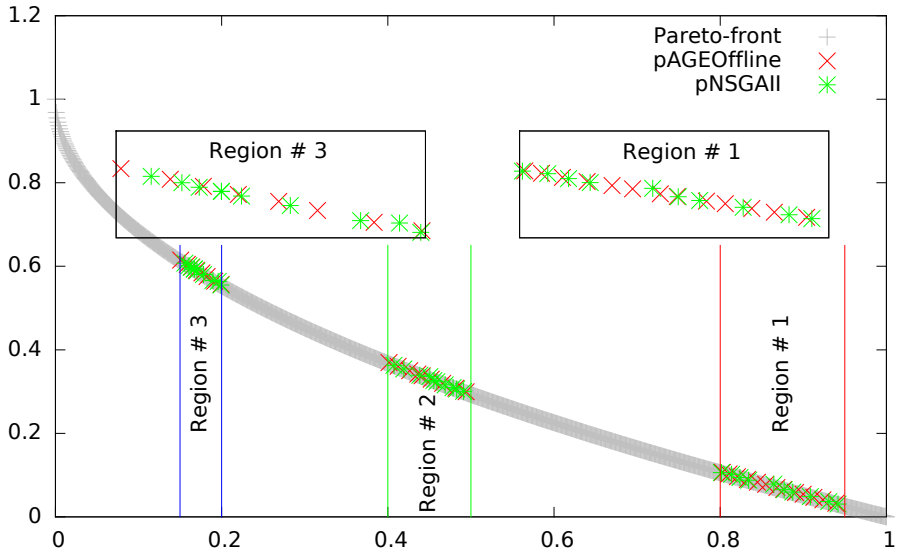


Figure 9.3: Pareto-fronts obtained by pAGEoffline and pNSGAII on ZDT1 problem.

regions. For more details, zoomed views of the obtained fronts of regions #1 and #3 are shown that allow us to visually assess the distribution of the solutions.

Next, we report the results in terms of mean values and the corresponding standard deviations of EPS, IGD and HV for each region for each pMOEAs. Figure 9.4 and 9.5 show the result obtained for NSGAII and AGE, respectively. By using markers of different shapes, we show how our pMOEAs perform compared to their original variants. In the following, we summarize the results.[5]

Most of the time, pNSGAII outperforms NSGAII with the same evaluation budget ($EF = 12000$) regardless of μ (Figure 9.4). pNSGAII performs similarly to NSGAII ($EF = 24000$) a number of times, i.e., on all regions for ZDT1/ZDT2 and regions #1/#2 for ZDT6. pNSGAII fails to converge on ZDT4 due to many local optima.

The Figure 9.5 demonstrates the comparison of pAGE variants (i.e., online and offline) and AGE with different configurations. pAGEonline and offline perform consistently better in comparison to the generic AGE (with $EF = 12000$ regardless of μ) for almost all the regions and all the problems. When comparing with AGE ($EF = 24000$), most of the time it perform similarly (all the ZDT problems expect ZDT). In the case of ZDT4, generic AGE performs better than pAGE variants. Please not that in this comparison, AGE uses two times function evaluations than pAGE variants.

It can be concluded from the results that our pMOEAs perform very well in short runs ($EF = 12000$); however, in compare with long run ($EF = 24000$) the performance is not so consistent.

To briefly demonstrate that the approach also works on three-dimensional problems, we show a few results in Figure 9.6; all other configurations performed worse and were left away. It is clear from Figure 9.6 that pAGE achieves better approximations than pNSGAII. The plot also provides a comparison with the generic AGE with two different configurations. The original AGE is not able to find the front given the computational budget (being 1–10 units away), whereas our pMOEAs achieve

[5]all code and results are available to `https://github.com/shaikatcse/pMOEAs`.

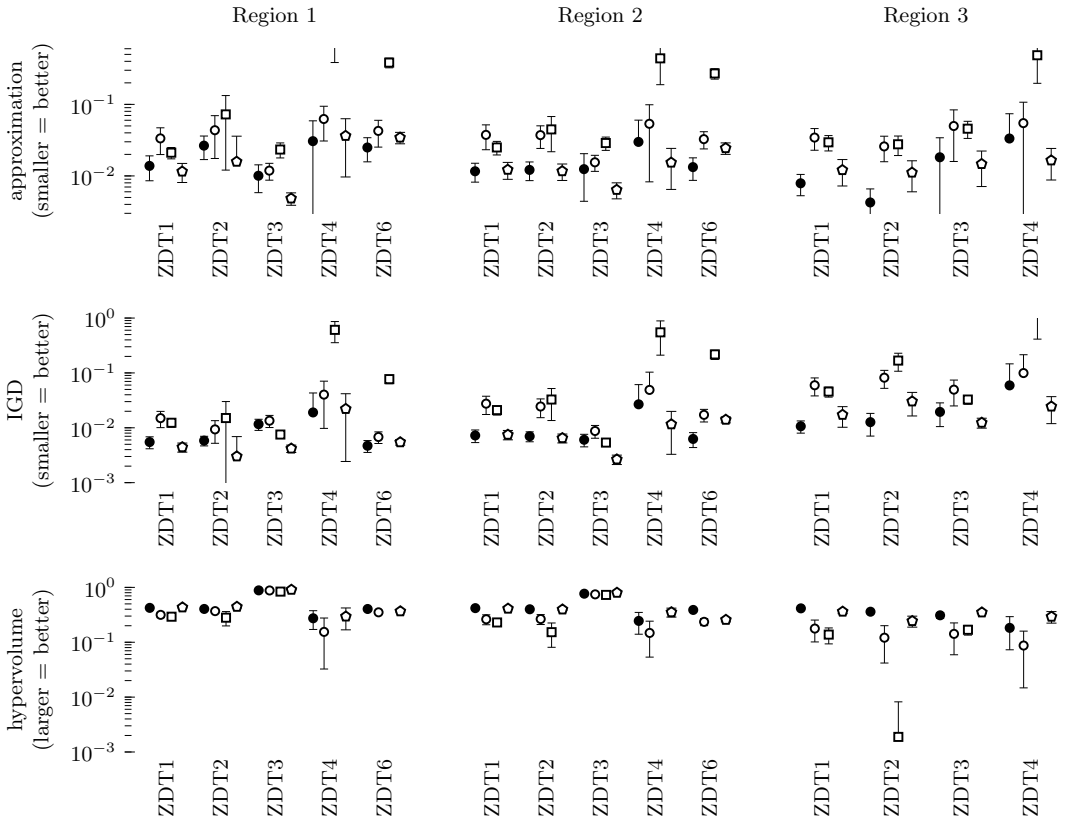


Figure 9.4: Comparison of our pNSGAII with their original variants on the ZDT functions with $m = 2$. The Regions 1–3 are defined in Section 9.4.2. Shown are the means and standard deviations of 100 independent runs. From top to bottom, we show the results of epsilon, IGD and HV. Within each block of four markers, we first show our pNSGAII with ●, then with ○ and □ the original algorithm with two population sizes ($\mu = 30$ and $\mu = 100$, $EF = 12,000$ each), and then with ⬠ the original algorithm with twice the evaluation budget ($\mu = 100$, $EF = 24,000$).

good approximations of the fronts with the same budget (being only 0.05–0.5 unit away). Between pAGEonline and pAGEoffline, there is no clear winner.

9.5 pNSGAII on energy system optimization problem

To investigate the performance of our approach on a real-world problem, we have applied pNSGAII on an energy system optimization problem. The general goal of the problem is to identify multiple optimal systems in order to minimize CO_2 emissions and annual cost. Here, we want to identify multiple optimal systems for three specific regions of interest (i.e., 10 solutions for each region) for the Aalborg energy system. The three regions are defined in terms of CO_2 emissions (i.e., [0.40, 0.5], [0.0, 0.15] and [-0.40, -0.50] Mtons of emissions). For example, we are interested in identifying 10 optimal solutions in a region within 0.40 to 0.5 million tons of CO_2 emissions.

The result is illustrated in Figure 9.7; x-axis presents emissions in million tons and y-axis presents annual cost in million Danish krone. The gray points represent

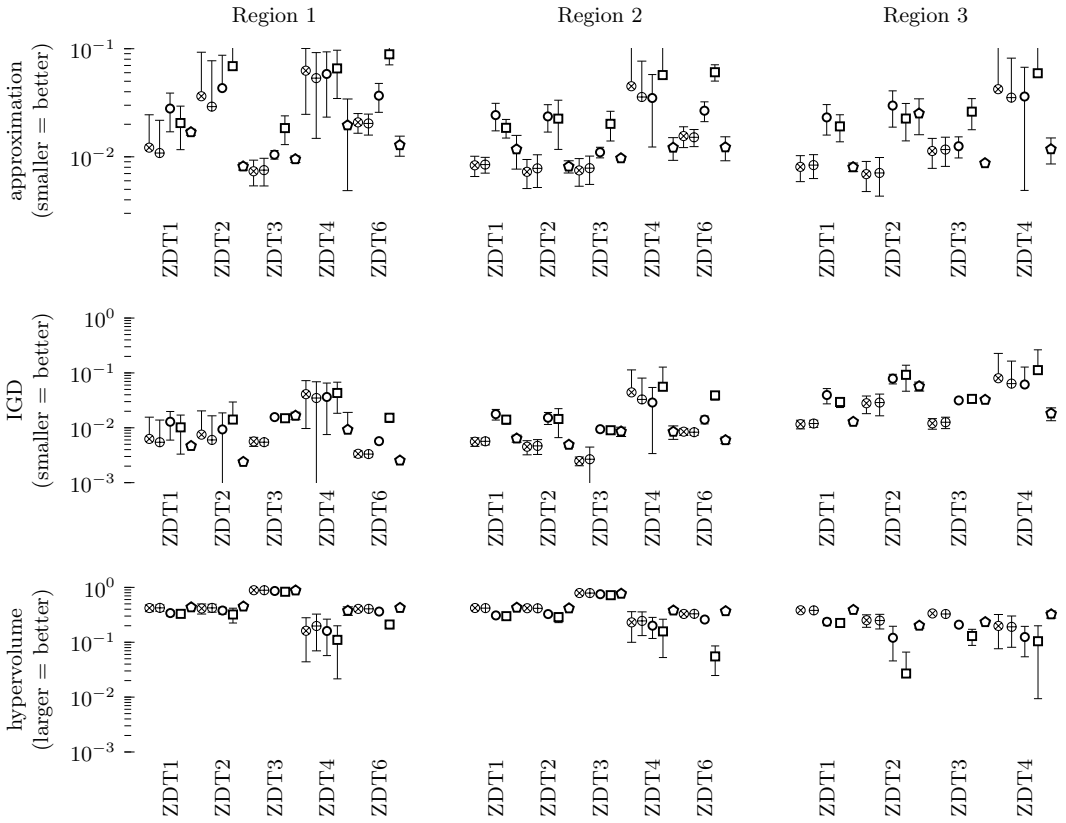


Figure 9.5: Comparison of our pAGE with their original variants on the ZDT functions with $m = 2$. The Regions 1–3 are defined in Section 9.4.2. Shown are the means and standard deviations of 100 independent runs. From top to bottom, we show the results of epsilon, IGD and HV. Within each block of five markers, we first show our pAGE variants with ○ (pAGEoffline with × and pAGEonline with +), then with ○ and □ the original algorithm with two population sizes ($\mu = 30$ and $\mu = 100$, $EF = 12,000$ each), and then with ⌂ the original algorithm with twice the evaluation budget ($\mu = 100$, $EF = 24,000$).

the true Pareto-front, which is approximated by considering the outcomes of 240 independent runs of NSGAII and SPEA2 with $\mu = 100, \xi = 7000$ (in Chapter 4). The red markers show the solutions found by our pNSGAII ($\mu = 30, \xi = 6000$, problem-specific constraint handling was added): 10 solutions per region are found, and they are very close to optimal solutions. As the experiment achieved these set goals, we conclude that our proposed approach can not only be successfully applied to test functions, but also to real-world optimization problems.

9.6 Conclusions

We proposed the concept of incorporating multiple user preferences into MOEAs via the use of intervals. The concept was designed with laypeople in mind who might not have detailed knowledge about the objective space.

We presented modifications for two MOEAs to handle multiple preferences, and we demonstrated the resulting capability on two- and three-dimensional test problems. On two-dimensional problems, our pMOEAs typically achieve the same hy-

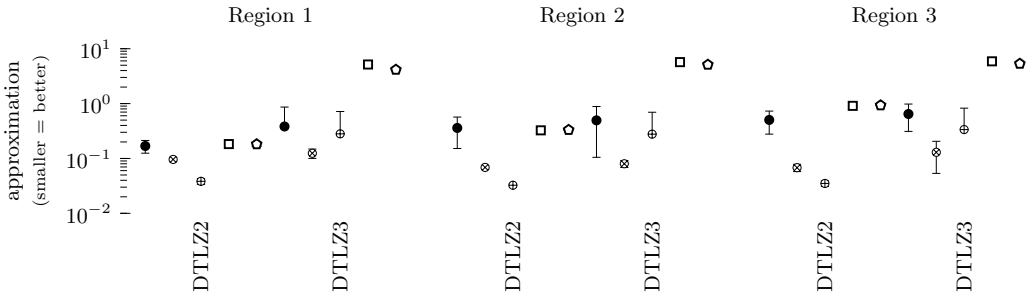


Figure 9.6: Comparison of our pMOEAs ($\mu = 30$, $EF = 50,000$) on a subset of the DTLZ functions with $d = 3$. ● denotes pNSGAII and ○ denotes our pAGEoffline ($\times$) and pAGEonline ($+$) variants, and □ and ⬠ denote the original AGE algorithm with $\mu = 100$, $EF = 50,000$ and $\mu = 150$, $EF = 49,950$ respectively. Typically, our pMOEAs achieve significantly better approximations of the preferred regions.

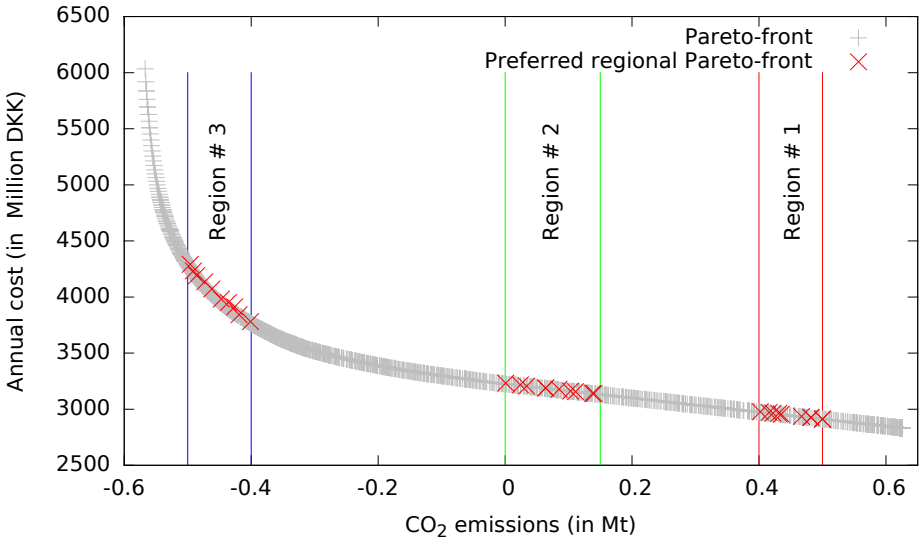


Figure 9.7: Result of the energy system optimization. Shown is the final solution set computed by pNSGAII.

pervolume, inverted generational and additive approximation values as the original algorithms, where the latter had twice the evaluation budget. On three-dimensional problems, our online and offline variants of AGE with preferences perform best. Finally, the effectiveness of the algorithm is investigated on a real-world problem and satisfactory performance is achieved.

As solutions can be spread too diversely over the objective space for higher dimension problems (having more than 3 objectives), we think that preferences in general can be an interesting option for decision makers.

Chapter 10

Conclusion

If you can't explain it simply, you don't understand it well enough.

— Albert Einstein

10.1 Thesis summary

The aim of the thesis is to develop a generalized, efficient, and user-friendly energy system optimization framework that can be used by decision makers to identify optimized energy scenarios. Since energy system optimization problem is a multi-objective optimization problem (the problem requires to optimize multiple objectives simultaneously), in this regards, a framework is proposed based on multi-objective evolutionary algorithms and EnergyPLAN. A multi-objective evolutionary algorithm is chosen as the problem on hand is a combinatorial optimization problem, at the same time, the problem is discontinuous in nature. In addition, EnergyPLAN is mainly selected for the ability to incorporate all the energy sectors together. Therefore, the proposed model is a generalized one as it can handle the synergies among all the major energy sectors which is essential to integrate large scale renewable energies. Moreover, simple domain knowledge regarding energy systems is incorporated in several stages of the model and a robust stopping criterion is proposed to improve the efficiency of the model.

It has been found from the results that the framework is able to find out multiple optimized scenarios while minimizing multiple conflicting objectives simultaneously. Moreover, the investigation shows that the domain knowledge exploited model performs better than default approach. The satisfactory results are found when the proposed stopping criterion is tested on benchmark problems. We have also found that the robust stopping criterion stops the energy system optimization process early to save expensive computational time while achieving good results.

Finally, we have suggested a user-friendly way to incorporating user preferences into the optimization process. The approach helps the user to explore specific regions of a Pareto-front. The proposed approach is very useful in the energy domain as decision makers generally want to identify optimized scenarios that fulfill some specific targets (such as scenarios within certain ranges of emissions reduction). The perform investigation shows that the modified algorithms explore the user-defined regions more efficiently than the generic algorithms on both benchmark and practical problems.

Afterwards, the framework is put forward to solve two energy systems optimization problems (energy systems for two Italian Alpine valley). Both energy systems evolve all three major sectors. The first system is optimized based on four objectives (handle economical, environmental, political and technical aspects) by considering recent demands. The optimization results show that a reasonable number of scenarios has identified, among them 13 scenarios are better than the current one (reference) in all four objectives. Moreover, all the identified scenarios are less emitting and less dependent than the reference one. For the second system, a long-term planning approach is considered. In this regards, optimization is performed to determine future scenarios; more specifically optimizations are carried out in three phases (for three different time periods) by considering future demands. The results show that it is possible to find some particular scenarios (i.e., target scenarios) from the optimized scenarios that fulfill the decision makers' goals, as decision makers want to reach distinct goals in term of emissions reductions. The optimized scenarios also reveal that it is required to transform thermal sector in near future and a radical transformation of transport sector should be carried out after the year of 2030. Finally, a technique is proposed to choose two scenarios from the targeted ones (of two different time frames) for a smooth transition from one time period to another. The technique is successfully employed to determine transient scenarios for the valley.

We conclude from all the results that proposed framework is a general and complete tool that can be applied to design virtually all kind of energy systems. The tool can be applied for not only short-term designing but also for mid-to-long term energy planning. We hope that the present contributions will stimulate energy planners to explore the potential of all these options.

10.2 Future work

There are some unanswered questions found in the thesis that can be considered as potential directions. The considered directions are two folded: i) possible next steps to energy domain, and ii) interesting extensions regarding the improvements of different components of multi-objective evolutionary algorithms. Following two paragraphs are dedicated to energy domain and the last two paragraph tries to provide some lights towards evolutionary algorithms.

As stated earlier that the approach presented here is used to identify optimal scenarios of an energy system, however, the thesis does not investigate the problem *"how to reach there?"*(a transition path [34] to optimized scenario). Even the framework can be used to identify optimal scenarios for different time frames, still we do not investigate which transition path is suitable for reaching one optimized scenario to another. Transition path is guided by different policies such as emission trading scheme [34], carbon taxation [34], feed-in-tariffs (FIT) [93]. Chappin investigates the impact of different policies on energy transitions; results from [34] show the strong dependence of a final state (scenario) on the adopted policies i.e., policy parameters. On the other hand, it is not only important to reach the target scenario, but also to approach it in a *smooth* way, i.e., with a transition compatible with economic and stability concerns. From an economic point of view, a sudden/abrupt change within an energy system is not expected to be sustainable. Therefore, it is important and strategic to have instruments to monitor the evolution of a transition and adapt/change the policy framework accordingly. Hence, potential research direction

could be the *optimization of policy parameters* taking into account both the desired final scenario and the transition path smoothness.

It may be possible that in future an energy system would be so complex that optimizing two objectives would not be enough to evaluate it properly. A number of evaluation criteria (objectives) have been proposed in [135] and [159]. The framework proposed here may not scalable to deal with many-objective problems [48]. Therefore, an interesting extension could be the incorporation of a many-objectives evolutionary algorithm such as NSGA-III [48], MOEA/D [176] into the framework.

More investigation can be carried out about the proposed stopping criterion in Chapter 5. As a future work, an interesting step forward would be the investigation of the performance of different metrics in the context of stopping criteria. In this thesis, average Hausdorff distance and diversity are used for monitoring objective and decision space, respectively. However, other objective space metric such as hypervolume with a combination with diversity metric may improve the performance of the stopping criteria. Only in-depth comparison of different combinations may reveal the best approach for the criteria. Moreover, probably the most difficult task in the field is to benchmark/testing different stop criteria as we do not know true point to stop. To our best knowledge, no study has been found in literature. Therefore, it remains an interesting topic for future.

Regarding preferred regional evolutionary algorithm proposed in Chapter 9, the future work will include the adoption of the techniques to the higher dimensional problems. Technically, the extension is straightforward, as the intervals just have to be added to an internal array. Whether the approaches are effective in higher-dimensional objective spaces remains to be seen.

Appendices

Appendix A

A.1 Local PV productivity and maximum reasonable PV capacity

The PV productivity (PV_{LP}) can be calculated by dividing total yearly PV production by peak power of PV capacity.

$$PV_{LP} = \frac{6,111,381kWh}{4,998kW} = 1,223kWh/kW \tag{A.1}$$

The maximum reasonable PV capacity is decided by dividing the annual maximum electrical demand (el_d^{con}) (assuming that the energy system is completely electrified: existing electrical demand, existing thermal demand covered by GSHP, existing transport demand covered by electric car) by the local PV productivity PV_{LP} :

$$PV_{max} = \frac{el_d^{con}}{PV_{LP}} = \frac{el_d + Q_{sh}/COP + (d_{pc} + d_{dc}) * elC_{eff}}{PV_{LP}} \tag{A.2}$$

$$= \frac{26,177,984kWh + 55,826,546kWh/3.2 + 48,084,100km * 0.168kWh/km}{1,223kWh/kW} \tag{A.3}$$

$$= \frac{51,701,908kWh}{1,223kWh/kW} \tag{A.4}$$

$$= 42.28MW \tag{A.5}$$

A.2 Assessment of the local sustainable wood resource

In order to verify the potential of a sustainable use of the wood resources in the studied area, the following calculation is performed. It considers the parameters (i.e., surface area and potentiality of sustainable use of biomass resources) provided by a recent study performed by province of Trento (i.e., "Energetic and Environmental Plan of the province of Trento" [71]):

$$SWR = S_{for} * P_{SWR} \tag{A.6}$$

Therefore, it is possible to use nearly 57 GWh/year (which is an increase by 30% over the reference scenario) of wood in a sustainable way for the studied area. Table A.1 shows the corresponding calculation.

S_{for} (ha)	$P_{SWR}(MWh/(ha*year))$	SWR (GWh)
13,671.35	4.16	56.87

Table A.1: Parameters for assessing sustainable wood resource.

A.3 Cost and efficiency related data

Table A.2: Investment cost, lifetime, fixed O&M cost

Technology	Unit	Investment cost (KEuro)	Lifetime (year)	Fixed O&M (%)	Reference
Hydro	kW_e	1.9	50	2.7	[47]
PV	kW_e	2.6	30	0.77	[47]
Biogas	kW_e	4	20	3.8	[7]
Individual wood, LPG and oil boilers	kW_{th}	0.588	15	2.1	[47]
Individual GSHP	kW_e	1.188	15	0.6	[47]
Borehole for GSHP	kW_e	3.2	100	0	[3]
Wood ORC CHP	kW_e	6.7	15	1.45	[60]
Fossil fuel based car	1 unit	9.45	15	0	[10]
Electric car	1 unit	18.69	15	5.5	[16]

Table A.3: Variable O&M cost

Technology	Unit	Variable O&M	Reference
Hydro	€/MWh	1.19	[47]
Wood ORC mCHP	€/MWh	2.7	[60]

Table A.4: Generation efficiency

Technology	Efficiency	Reference
Individual wood boiler	Th = 0.75	[47]
Individual gas boiler	Th = 0.9	[47]
Individual oil boiler	Th = 0.8	[47]
Individual GSHP	COP = 3.2	[3]
Wood ORC mCHP	El = 0.18, Th = 0.8	[166]
Electric car	0.168 kWh/km	[16]

Table A.5: Fuels' prices and additional cost

Fuel	Price (€/MWh)	Reference
Oil & Diesel	145	[126]
Petrol	181.29	[126]
LPG	162	[126]
Wood	35	[2]
Electricity import	Hourly price (average 61.58)	[11]
Electricity export	Hourly price (average 61.58)	[11]
Electricity internal use additional cost (Supply and sale + general system charges + grid and metering cost + taxes – CEIS discount (15%))	106.27	[32]

Appendix B

B.1 Parameters regarding the energy system optimization framework for VdN

Table B.1: Parameters for framework used for finding optimized scenarios for VdN

	Parameters	Value
Initialization	θ	6.0
related	k	3
parameters	β	[0,1,2]
Stopping criterion	nGenLT	20
related	nGenUnCh	5
parameters	α	0.05

Table B.2: Domain-knowledge related to decision variable and objectives

	PV	OB	NGB	BioB	BioCHP	HP	ICEC	EC	ST
2020 & 2030									
Emission	True	False	Null	True	True	Null	False	True	True
AC	False	Null	Null	True	False	Null	True	False	False
2050									
Emission	True	False	Null	True	True	Null	False	True	True
AC	Null	Null	Null	True	False	Null	Null	Null	null

Where PV: photovoltaics, OB: oil boiler, NGB: natural gas boiler, BioB: biomass boiler, BioCHP: biomass CHP, ICE: Internal combation engine cars, EC: electric car and ST: solar thermal.

B.2 Results for 4DS case for VdN

Table B.3: Comparison of target scenarios with reference scenarios for 2020 and 4DS case

Scn.	CO2 emission (ktons)	Red wrt RS2008(%)	Red wrt RS2020 (%)	AC (kEuro)	Inc wrt RS2008 (%)	Red wrt RS2020 (%)	Capacities							Number of Cars	
							PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kW_{th})	NGas Boiler (kW_{th})	Biomass Boiler (Kth)	Solar thermal (kW_{th})	ICE cars	EVs
RS2008	98.09	-	-	140863	-	-	936	0	0	27463	21275	15983	13310	19379	0
RS2020	91.84	6.37	-	160363	13.84	-	936	0	0	26832	20880	15522	13286	19379	0
TS1	44.14	55.00	51.94	147370	4.62	8.10	1763	184	8508	33	11	21858	4108	18733	647
TS2	44.40	54.73	51.65	147100	4.43	8.27	1410	91	8625	10	13	21308	985	18843	537
TS3	44.81	54.31	51.20	146883	4.27	8.41	2804	944	8541	15	15	20924	2477	19125	254
TS4	44.95	54.18	51.06	146688	4.14	8.53	2993	30	8571	1	20	21672	1169	19169	210
TS5	45.32	53.79	50.65	146514	4.01	8.64	1063	277	8645	6	8	21017	3337	19303	77
TS6	45.42	53.69	50.54	146306	3.86	8.77	969	61	8625	12	5	21344	235	19344	36
TS7	45.42	53.69	50.54	146305	3.86	8.77	969	61	8625	12	4	21344	235	19344	36

Table B.4: Comparison of target scenarios with reference scenarios for 2030 and 4DS case

Scn.	CO2 emission (ktons)	Red wrt RS2008(%)	Red wrt RS2020 (%)	AC (kEuro)	Inc wrt RS2008 (%)	Red wrt RS2020 (%)	Capacities							Number of Cars	
							PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kW_{th})	NGas Boiler (kW_{th})	Biomass Boiler (Kth)	Solar thermal (kW_{th})	ICE cars	EVs
RS2008	98.09	-	-	140863	-	-	936	0	0	27463	21275	15983	13310	19379	0
RS2030	86.05	12.27	-	167656	19.02	-	936	0	0	25794	20138	14911	13286	19379	0
TS1	29.82	69.60	65.35	147875	4.98	11.80	1879	152	7397	16	11	21697	725	13845	5535
TS2	30.13	69.29	64.99	147836	4.95	11.82	988	5	7404	17	26	21781	384	13983	5397
TS3	30.74	68.66	64.28	147828	4.94	11.83	959	1	7518	6	5	21115	106	14316	5063
TS4	31.04	68.36	63.93	147803	4.93	11.84	1911	8	7404	0	6	21819	78	14521	4859
TS5	31.25	68.14	63.69	147807	4.93	11.84	1886	126	7409	2	15	21660	121	14632	4747
TS6	32.26	67.12	62.51	147765	4.90	11.86	1409	110	7402	3	23	21707	311	15170	4210
TS7	32.63	66.73	62.08	147758	4.89	11.87	1017	29	7517	1	6	21099	180	15359	4021
TS8	32.88	66.48	61.79	147746	4.89	11.88	1911	69	7395	3	13	21800	333	15526	3854
TS9	33.11	66.25	61.52	147718	4.87	11.89	1020	69	7396	3	3	21804	423	15638	3742
TS10	34.09	65.25	60.38	147692	4.85	11.91	1091	62	7397	1	43	21763	433	16158	3222

Table B.5: Comparison of target scenarios with reference scenarios for 2050 and 4DS case

Scn.	CO2 emission (ktons)	Red wrt RS2008(%)	Red wrt RS2020 (%)	AC (kEuro)	Inc wrt RS2008 (%)	Red wrt RS2020 (%)	Capacities							Number of Cars	
							PV (kWe)	Biomass CHP (kWe)	GSHP (kWe)	Oil Boiler (kW_{th})	NGas Boiler (kW_{th})	Biomass Boiler (Kth)	Solar thermal (kW_{th})	ICE cars	EVs
RS2008	98.09	-	-	140863	-	-	936	0	0	27463	21275	15983	13310	19379	0
RS2050	82.78	15.61	-	181145	28.60	-	936	0	0	25880	20227	14923	13286	19379	0
TS1	0.78	99.20	99.05	150444	6.80	16.95	39985	24033	6863	1	2	153	185489	1	19379
TS2	0.87	99.12	98.95	147299	4.57	18.68	39952	18944	7291	0	1	2221	24945	1	19379
TS3	1.03	98.95	98.75	145813	3.51	19.50	39984	4056	6865	0	2	19171	19192	1	19379
TS4	0.86	99.12	98.96	147344	4.60	18.66	39739	20043	7441	1	6	194	18298	0	19380
TS5	1.09	98.89	98.68	145408	3.23	19.73	39461	45	6828	1	6	23221	13459	0	19380
TS6	1.08	98.90	98.70	145505	3.30	19.67	39954	934	6829	1	8	22372	16240	0	19379
TS7	1.09	98.88	98.68	145364	3.20	19.75	39790	208	7025	0	0	21798	2499	0	19380
TS8	0.97	99.01	98.83	146371	3.91	19.20	39802	9712	7039	3	2	12650	21171	0	19379
TS9	0.88	99.10	98.93	147157	4.47	18.76	39993	18284	7467	2	0	1703	15078	1	19379
TS10	0.87	99.12	98.96	147318	4.58	18.67	39463	19496	7421	0	1	849	20439	0	19379

B.3 Projected investment, operational and maintenance costs, lifetimes and efficiencies

Table B.6: projected investment cost for different technologies for different targeted years.

Technology	Unit	2008 (MEuro/Unit)	2020 (MEuro/Unit)	2030 (MEuro/Unit)	2050 (MEuro/Unit)	Reference
Hydro Power	MW_e	3.297	3.297	3.297	3.297	[88]
PV (Residential)	MW_e	6.2	1.3	1.1	0.9	[90, 46]
Solar Thermal for HSW+SH (+ heat infrastructure)	MW_{th}	0.958	0.697	0.479	0.383	[87]
Solar Thermal for HSW+SH (+ heat infrastructure)	TWh/year	2,281	1,660	1,140	571	[46]
Individual boilers (+ heat infrastructure)	MW_{th}	0.528	0,565	0,595	0,667	[46]
Individual GSHP (+ heat infrastructure) + Borehole		2.328	2.183	2.062	1.93	[87, 3]
Biomass cogeneration (<10MW)	MW_e	5.988	5.326	4.775	4.775	[89]
District heating	1000Euro/TJ		145	145	145	[46]
Alkaline electrolyser	MW_e		0.834	0.631	0.508	[91]
H2 station 500 kg/day (incl H2 compr& storage)	1000 car		1.444	1.111	1.111	[61]
Residential BEV charging	1000 car		1.4	1.4	1.4	[61]
ICE (Diesel) car	1000 car	18.763	20.357	21.048	21.177	[124]
BEV car	1000 car	39.258	27.78	24.321	22.57	[124]
FCEV car	1000 car	122.672	31.332	24.884	22.385	[124]

Table B.7: Projected operational and maintenance cost for different technologies for targeted years.

Technology	Unit	2008	2020	2030	2050	Reference
Hydro Power	% of Investment cost	2.65	2.65	2.65	2.65	[88]
PV		2.94	2.09	1.38	1.15	[46]
Solar Thermal for HSW+SH (+ heat infrastructure)		0.11	0.13	0.15	0.15	[46]
Individual oil boilers (+ heat infrastructure)		4	4	4	4	[46]
Individual GSHP (+ heat infrastructure)		1.1	1.1	1.1	1.1	[46]
Individual GSHP (+ heat infrastructure) + Borehole		-	0,6	0,57	0,53	[46]
Biomass cogeneration (<10MW) + district heating		-	4,06	3,92	3,92	[46, 89]
Alkaline electrolyser		-	5	5	5	[91]
ICE Car (all fuels)		4.09	4.09	4.09	4.09	[46]
BEV car + Residential BEV charging		-	6,65	4,10	4,09	[46]
FCEV car + H2 station 500 kg/day (incl H2 compr & storage)		-	6,88	4,24	4,24	[46, 61]

Table B.8: Projected lifetime for different technologies for targeted years.

Technology	Unit	2008	2020	2030	2050	Reference
Hydro Power	years	50	50	50	50	[46]
PV	years	20	25	30	30	[46]
Solar Thermal (+ heat infrastructure)	years	20	25	30	30	[46]
Individual boilers (+ heat infrastructure)	years	20	20	20	20	[46]
GSHP + Borehole	years	34	34	34	34	[46, 3]
Biomass cogeneration <10MW) + DH	years		24	25	25	[46, 89]
Alkaline electrolyser	years		10	13	13	[91]
ICE car (all fuels)	years	15	15	15	15	[46]
BEV car + Residential BEV charging	years	13	13	13	13	[91, 61]
FCEV car + H2 station 500 kg/day (incl H2 compr & storage)	years	13	13	13	13	[91]

Table B.9: Efficiencies for different technologies for targeted yeas.

Technology	Efficiency 2008	Efficiency 2020	Efficiency 2030	Efficiency 2050	Reference
Solar Thermal for HSW+ SH (+ heat infrastructure)	0.533 MWh/kW	0.533 MWh/kW	0.533 MWh/kW	0.533 MWh/kW	[87]
Individual oil boilers (+ heat infrastructure)	Th = 0.8	Th = 0.85	Th = 0.87	Th = 0.89	[46]
Individual biomass boilers (+ heat infrastructure)	Th = 0.75	Th = 0.80	Th = 0.82	Th = 0.84	[46]
Individual gas boiler	Th = 0.90	Th = 0.95	Th = 0.97	Th = 0.99	[46]
Individual GSHP (+ heat infrastructure)	COP = 3.90	COP = 4.54	COP = 5.07	COP = 5.46	[87]
Biomass cogeneration (<10MW)	El = 0.16, Th = 0.80	El = 0.17, Th = 0.80	El = 0.18, Th = 0.80	El = 0.18, Th = 0.80	[166]
District heating: Net loss (%)	-	14.5%	14.5%	14.5%	[46]
Alkaline electrolyser	-	74	75	78	[91]
H2 station 500 kg/day (incl H2 compr & storage)	-	80	80	80	[91]
ICE car (diesel and petrol)	0.684 kWh/km	0.607 kWh/km	0.549 kWh/km	0.497 kWh/km	[61]
BEV car	-	0.169 kWh/km	0.145 kWh/km	0.117 kWh/km	[61]
FCEV car	-	0.334 kWh/km	0.267 kWh/km	0.200 kWh/km	[91]

Bibliography

[1] The National Institute for Statistics, italy. [Online]. Available: https://www.istat.it/en/

[2] AIEL- Italian agriforenergy association. [Online]. Available: http://www.aiel.cia.it/en/

[3] ANIG HP - Associazione nazionale impianti geotermia heat pump (national association for geothermal heat pump). [Online]. Available: http://www.anighp.it/

[4] Comunità della Val di Non - Community of Val di Non. [Online]. Available: http://www.comunitavaldinon.tn.it/

[5] COVENANT OF MAYORS. [Online]. Available: http://www.eumayors.eu/IMG/pdf/covenantofmayors_text_en.pdf

[6] Domoliti network. [Online]. Available: http://www.novareti.eu/

[7] ENAMA-Ente nazionale meccanizzazione agricola, macchine agricole, agricoltura, meccanizzazione agricola (ENAMA, italy). [Online]. Available: http://www.enama.it/en/index.php

[8] "Energy technology perspectives 2015 (ETP 2015)," Tech. Rep. [Online]. Available: http://www.iea.org/etp/etp2015/

[9] Home - Eurostat. [Online]. Available: http://ec.europa.eu/eurostat

[10] fiat.it - Punto 2012 - Discover. [Online]. Available: http://www.fiat.it/punto

[11] GME - Gestore dei Mercati Energetici SpA (Energy Market Operator). [Online]. Available: http://www.mercatoelettrico.org/it/

[12] H2RES model. [Online]. Available: http://h2res.fsb.hr/

[13] HOMER - Hybrid Renewable and Distributed Generation System Design Software. [Online]. Available: http://www.homerenergy.com/

[14] LEAP-Long-range Energy Alternatives Planning System. [Online]. Available: http://www.energycommunity.org/

[15] Meteotrentino. [Online]. Available: http://hydstraweb.provincia.tn.it/web.htm?ppbm=T0414&rs&1&df

[16] Nissan LEAF - auto 100% elettriche - veicolo elettrico nissan. [Online]. Available: http://www.nissan.it/IT/it/vehicle/electric-vehicles/leaf.html

[17] Sistema informativo degli indicatori statistici (database for statistical indicators), province of trento. [Online]. Available: http://www.statweb.provincia.tn.it/indicatoristrutturalisubpro/?q=riscaldamento

[18] TIMES. [Online]. Available: http://www.iea-etsap.org/web/Times.asp

[19] A. Alarcon-Rodriguez, G. Ault, and S. Galloway, "Multi-objective planning of distributed energy resources: A review of the state-of-the-art," vol. 14, no. 5, pp. 1353–1366. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032110000146

[20] P. Alberg Østergaard, B. V. Mathiesen, B. Möller, and H. Lund, "A renewable energy scenario for aalborg municipality based on low-temperature geothermal heat, wind power and biomass," vol. 35, no. 12, pp. 4892–4901. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544210004779

[21] P. Alberg Østergaard, "Transmission-grid requirements with scattered and fluctuating renewable electricity-sources," vol. 76, no. 1, pp. 247–255. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261903000655

[22] S. Amicabile, J.-I. Lee, and D. Kum, "A comprehensive design methodology of organic rankine cycles for the waste heat recovery of automotive heavy-duty diesel engines," vol. 87, pp. 574–585. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1359431115003774

[23] Automobile Club d'Italia. ACI studi e ricerche - open data (ACI studies and research - open data). [Online]. Available: http://www.aci.it/laci/studi-e-ricerche/dati-e-statistiche/open-data.html

[24] J. L. Bernal-Agustín and R. Dufo-López, "Efficient design of hybrid renewable energy systems using evolutionary algorithms," *Energy Conversion and Management*, vol. 50, no. 3, pp. 479–489, Mar. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0196890408004524

[25] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed. Princeton University Press, 7 2009.

[26] P. P. Bonissone, R. Subbu, N. Eklund, and T. R. Kiehl, "Evolutionary algorithms + domain knowledge = real-world evolutionary computation," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 256–280, June 2006.

[27] J. Branke, *Multiobjective Optimization: Interactive and Evolutionary Approaches*. Springer, 2008, ch. Consideration of Partial User Preferences in Evolutionary Multiobjective Optimization, pp. 157–178.

[28] K. Bringmann, T. Friedrich, and P. Klitzke, "Generic postprocessing via subset selection for hypervolume and epsilon-indicator," in *Parallel Problem Solving from Nature – PPSN XIII*, ser. Lecture Notes in Computer Science, T. Bartz-Beielstein, J. Branke, B. Filipič, and J. Smith, Eds. Springer International Publishing, no. 8672, pp. 518–527. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-319-10762-2_51

[29] K. Bringmann, T. Friedrich, F. Neumann, and M. Wagner, "Approximation-guided evolutionary multi-objective optimization," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume*

Two, ser. IJCAI'11. AAAI Press, 2011, pp. 1198–1203. [Online]. Available: http://dx.doi.org/10.5591/978-1-57735-516-8/IJCAI11-204

[30] L. T. Bui, S. Wesolkowski, A. Bender, H. A. Abbass, and M. Barlow, "A dominance-based stability measure for multi-objective evolutionary algorithms," in *2009 IEEE Congress on Evolutionary Computation*, May 2009, pp. 749–756.

[31] P. Capros, A. De Vita, N. Tasios, D. Papadopoulos, P. Siskos, E. Apostolaki, M. Zampara, L. Paroussos, K. Fragiadakis, N. Kouvaritakis, L. Hoglund-Isaksson, W. Winiwarter, P. Purohit, H. Bottcher, S. Frank, P. Havlik, M. Gusti, and H. P. Witzke, *EU Energy, Transport and GHG Emissions: Trends to 2050, Reference Scenario 2013*, Dec. 2013. [Online]. Available: http://trid.trb.org/view.aspx?id=1285618

[32] CEIS, "Bilancio sociale 2013 - CEIS (company balance 2013 - CEIS)." [Online]. Available: http://www.ceis-stenico.it/clientfiles/upload/175_NewsFile.pdf

[33] T. Cerovac, B. Ćosić, T. Pukšec, and N. Duić, "Wind energy integration into future energy systems based on conventional plants – the case study of croatia." [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261914006370

[34] È. J. L. CHAPPIN, "Simulating energy transitions," Ph.D. dissertation, Delft University of Technology, 2011.

[35] Z. Chen and F. Blaabjerg, "Wind farm-a power source in future power systems," vol. 13, no. 6, pp. 1288–1300. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032108001433

[36] C. C. Coello, G. B. Lamont, and D. A. v. Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd ed. Springer.

[37] C. A. Coello Coello and M. Reyes Sierra, "A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm," in *MICAI 2004: Advances in Artificial Intelligence: Third Mexican International Conference on Artificial Intelligence, Mexico City, Mexico, April 26-30, 2004. Proceedings*, R. Monroy, G. Arroyo-Figueroa, L. E. Sucar, and H. Sossa, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 688–697. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-24694-7_71

[38] R.-G. Cong, "An optimization model for renewable energy generation and its application in china: A perspective of maximum utilization," vol. 17, pp. 94–103. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032112005023

[39] R.-G. Cong and S. Shen, "How to develop renewable power in china? a cost-effective perspective," vol. 2014. [Online]. Available: http://www.hindawi.com/journals/tswj/2014/946932/abs/,http://www.hindawi.com/journals/tswj/2014/946932/abs/

[40] D. Connolly, H. Lund, B. V. Mathiesen, and M. Leahy, "Modelling the existing irish energy-system to identify future energy costs and the maximum wind penetration feasible," vol. 35, no. 5, pp. 2164–2173. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544210000484

[41] D. Connolly, H. Lund, B. V. Mathiesen, S. Werner, B. Móller, U. Persson, T. Boermans, D. Trier, P. A. Østergaard, and S. Nielsen, "Heat roadmap europe: Combining district heating with heat savings to decarbonise the EU energy system," vol. 65, pp. 475–489. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0301421513010574

[42] D. Connolly, H. Lund, B. Mathiesen, and M. Leahy, "A review of computer tools for analysing the integration of renewable energy into various energy systems," vol. 87, no. 4, pp. 1059–1082. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261909004188

[43] D. Connolly and B. V. Mathiesen, "A technical and economic analysis of one potential pathway to a 100% renewable energy system," vol. 1, no. 0, pp. 7–28. [Online]. Available: http://journals.aau.dk/index.php/sepm/article/view/497

[44] D. Connolly, H. Lund, B. V. Mathiesen, and M. Leahy, "A review of computer tools for analysing the integration of renewable energy into various energy systems," *Applied Energy*, vol. 87, no. 4, pp. 1059–1082, 2010.

[45] C. Cormio, M. Dicorato, A. Minoia, and M. Trovato, "A regional energy planning methodology including renewable energy sources and environmental constraints," vol. 7, no. 2, pp. 99–130. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032103000042

[46] C. David, "EnergyPLAN cost database, version 3.0," p. Denmark. [Online]. Available: http://www.energyplan.eu/costdatabase

[47] C. David, "EnergyPLAN cost database, version 1.0," p. Denmark. [Online]. Available: http://www.energyplan.eu/

[48] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, Aug 2014.

[49] K. Deb, "An efficient constraint handling method for genetic algorithms," *Computer Methods in Applied Mechanics and Engineering*, vol. 186, no. 2-4, pp. 311–338, 2000.

[50] K. Deb, "An efficient constraint handling method for genetic algorithms," *Computer Methods in Applied Mechanics and Engineering*, vol. 186, no. 2–4, pp. 311–338, Jun. 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0045782599003898

[51] K. Deb, *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons.

[52] K. Deb and S. Jain, "Running performance metrics for evolutionary multi-objective optimizations," in *Proceedings of the Fourth Asia-Pacific Conference on Simulated Evolution and Learning (SEAL'02)*, 2002, pp. 13–20.

[53] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[54] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization*, ser. Advanced Information and Knowledge Processing. Springer London, 2005, pp. 105–145. [Online]. Available: http://dx.doi.org/10.1007/1-84628-137-7_6

[55] K. Deb, J. Sundar, U. B. R. N, and S. Chaudhuri, "Reference point based multi-objective optimization using evolutionary algorithms," in *International Journal of Computational Intelligence Research*. Springer-Verlag, 2006, pp. 635–642.

[56] L. Devroye, *Non-Uniform Random Variate Generation*. Springer-Verlag, 1986.

[57] C. Dong, G. H. Huang, Y. P. Cai, and Y. Liu, "An inexact optimization modeling approach for supporting energy systems planning and air pollution mitigation in beijing city," vol. 37, no. 1, pp. 673–688. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S036054421100689X

[58] J. Durillo, A. Nebro, and E. Alba, "The jMetal framework for multi-objective optimization: Design and architecture," in *2010 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–8.

[59] J. J. Durillo and A. J. Nebro, "jMetal: A Java framework for multi-objective optimization," *Advances in Engineering Software*, vol. 42, pp. 760–771, 2011.

[60] A. Duvia, A. Guercio, and C. Rossi, "Technical and economic aspects of biomass fuelled CHP plants based on ORC turbogenerators feeding existing district heating networks," in *Proceedings of the 17th European Biomass Conference*.

[61] C. Econometrics, "En route pour un transport durable," Cambridge Econometrics, UK, Tech. Rep., Nov. 2015. [Online]. Available: http://www.camecon.com/Libraries/Downloadable_Files/En_route_pour_un_transport_durable_-_technical_report.sflb.ashx

[62] O. Ellabban, H. Abu-Rub, and F. Blaabjerg, "Renewable energy resources: Current status, future prospects and their enabling technology," vol. 39, pp. 748–764. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032114005656

[63] European Commission, "Energy and climate change," BELGIUM, Tech. Rep., 2008. [Online]. Available: http://www.consilium.europa.eu/uedocs/cmsUpload/st17215.en08.pdf

[64] European Commission, "Energy roadmap 2050."

[65] European Commission, "EUROPE 2020: A strategy for smart, sustainable and inclusive growth."

[66] European Commission, "The European Union explained: Climate action," BELGIUM, Tech. Rep., Nov. 2014.

[67] European Commission, "A policy framework for climate and energy in the period from 2020 to 2030."

[68] M. Fadaee and M. Radzi, "Multi-objective optimization of a stand-alone hybrid renewable energy system by using evolutionary algorithms: A review," vol. 16, no. 5, pp. 3364–3369. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032112001669

[69] M. Falchetta, "Fonti rinnovabili e rete elettrica in italia. considerazioni di base e scenari di evoluzione delle fonti rinnovabili elettriche in italia (renewable energy and electricity grid in italy. basic considerations and evolution scenarios of electric renewable sources in italy)," ENEA; Italian national agency for new technologies, energy and sustainable economic development, Tech. Rep. ENEA-RT-2014-08, 2014.

[70] A. M. Fathoni, N. A. Utama, and M. A. Kristianto, "A Technical and Economic Potential of Solar Energy Application with Feed-in Tariff Policy in Indonesia," *Procedia Environmental Sciences*, vol. 20, pp. 89–96, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1878029614000140

[71] D. T. A. Forest, "Piano energetico-ambientale provinciale 2013-2020 (provincial environmental and energtic plan 2013-2020)." [Online]. Available: http://enerweb.casaccia.enea.it/enearegioni/UserFiles/Trento/PEAP_Tremto_2013_2020.pdf

[72] A. Franco and P. Salza, "Strategies for optimal penetration of intermittent renewables in complex energy systems based on techno-operational objectives," vol. 36, no. 2, pp. 743–753. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0960148110003460

[73] T. Friedrich and M. Wagner, "Seeding the initial population of multi-objective evolutionary algorithms: A computational study," *Applied Soft Computing*, vol. 33, no. 0, pp. 223 – 230, 2015.

[74] T. Friedrich, T. Kroeger, and F. Neumann, "Weighted preferences in evolutionary multi-objective optimization," in *Advances in Artificial Intelligence (AI)*. Springer, 2011, pp. 291–300.

[75] J. E. Gentle, *Random number generation and Monte Carlo methods*. Springer Science & Business Media, 2006.

[76] T. Goel and N. Stander, "A non-dominance-based online stopping criterion for multi-objective evolutionary algorithms," *International Journal for Numerical Methods in Engineering*, vol. 84, no. 6, pp. 661–684, 2010.

[77] D. Gómez-Lorente, I. Triguero, C. Gil, and O. Rabaza, "Multi-objective evolutionary algorithms for the design of grid-connected solar tracking systems," *International Journal of Electrical Power & Energy Systems*, vol. 61, pp. 371–379, Oct. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0142061514001768

[78] N. R. C. Government of Canada, "RETScreen International Home." [Online]. Available: http://www.retscreen.net/ang/home.php

[79] Y. Guo, J. Zheng, and X. Li, "An improved performance metric for multiobjective evolutionary algorithms with user preferences," in *Congress on Evolutionary Computation (CEC)*, 2015, pp. 908–915.

[80] D. A. Hagos, A. Gebremedhin, and B. Zethraeus, "Towards a flexible energy system – a case study for inland norway," vol. 130, pp. 41–50. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261914005170

[81] Y. HAIMES YV, L. LASDON LS, and D. WISMER DA, "On a bicriterion formation of the problems of integrated system identification and system optimization," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-1, no. 3, pp. 296–297, 7 1971.

[82] H. Hao, Z. Liu, F. Zhao, W. Li, and W. Hang, "Scenario analysis of energy consumption and greenhouse gas emissions from China's passenger vehicles," *Energy*, vol. 91, pp. 151–159, Nov. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544215011299

[83] L. D. D. Harvey, "Global climate-oriented transportation scenarios," *Energy Policy*, vol. 54, pp. 87–103, Mar. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S030142151200938X

[84] K. Hedegaard, B. V. Mathiesen, H. Lund, and P. Heiselberg, "Wind power integration using individual heat pumps – analysis of different heat storage options," vol. 47, no. 1, pp. 284–293. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544212007086

[85] L. Hong, H. Lund, B. V. Mathiesen, and B. Möller, "2050 pathway to an active renewable energy scenario for jiangsu province," vol. 53, pp. 267–278. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0301421512009408

[86] L. Hong, H. Lund, and B. Möller, "The importance of flexible power plant operation for jiangsu's wind integration," vol. 41, no. 1, pp. 499–507. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544212001338

[87] International energy agency, "Technology Roadmap Energy-efficient Buildings: Heating and Cooling Equipment," INTERNATIONAL ENERGY AGENCY, France, Tech. Rep., 2011.

[88] International energy agency, "Technology Roadmap Hydropower," INTERNATIONAL ENERGY AGENCY, France, Tech. Rep., 2012. [Online]. Available: https://www.iea.org/publications/freepublications/publication/2012_Hydropower_Roadmap.pdf

[89] International energy agency, "Technology Roadmap Bioenergy for Heat and Power," INTERNATIONAL ENERGY AGENCY, France, Tech. Rep., 2012.

[90] International energy agency, "Technology Roadmap Solar Photovoltaic Energy," International energy agency, France, Tech. Rep., 2014. [Online]. Available: https://www.iea.org/publications/freepublications/publication/TechnologyRoadmapSolarPhotovoltaicEnergy_2014edition.pdf

[91] International energy agency, "Technology Roadmap Hydrogen and Fuel Cells," Internation energy agency, France, Tech. Rep., 2015.

[92] S. Jiménez-Fernández, S. Salcedo-Sanz, D. Gallo-Marazuela, G. Gómez-Prada, J. Maellas, and A. Portilla-Figueras, "Sizing and maintenance visits optimization of a hybrid photovoltaic-hydrogen stand-alone facility using evolutionary algorithms," *Renewable Energy*, vol. 66, pp. 402–413, Jun. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0960148113007052

[93] K.-K. Kim and C.-G. Lee, "Evaluation and optimization of feed-in tariffs," *Energy Policy*, vol. 49, pp. 192–203, Oct 2012.

[94] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," vol. 91, no. 9, pp. 992–1007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832005002012

[95] C. Koroneos, M. Michailidis, and N. Moussiopoulos, "Multi-objective optimization in energy systems: the case study of lesvos island, greece," vol. 8, no. 1, pp. 91–100. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032103000911

[96] A. Kusiak, F. Tang, and G. Xu, "Multi-objective optimization of HVAC system with an evolutionary computation algorithm," vol. 36, no. 5, pp. 2440–2449. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544211000314

[97] P. S. Kwon and P. Østergaard, "Assessment and evaluation of flexible demand in a danish future energy scenario," vol. 134, pp. 309–320. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261914008472

[98] P. S. Kwon and P. A. Østergaard, "Comparison of future energy scenarios for denmark: IDA 2050, CEESA (coherent energy and environmental system analysis), and climate commission 2050," vol. 46, no. 1, pp. 275–282. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544212006536

[99] P. S. Kwon and P. A. Østergaard, "Priority order in using biomass resources – energy systems analyses of future scenarios for denmark," vol. 63, pp. 86–94. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544213008384

[100] T. Lambert, P. Gilman, and P. Lilienthal, "Micropower system modeling with homer," in *Integration of Alternative Sources of Energy*. John Wiley & Sons, Inc., pp. 379–418. [Online]. Available: http://dx.doi.org/10.1002/0471755621.ch15

[101] H. Lund, "EnergyPLAN: Advanced energy system analysis computer model." [Online]. Available: http://www.energyplan.eu/wp-content/uploads/2013/06/EnergyPLAN-Documentation-V11.4-2014.pdf

[102] H. Lund and B. V. Mathiesen, "Energy system analysis of 100% renewable energy systems—the case of denmark in years 2030 and 2050," vol. 34, no. 5, pp. 524–531. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544208000959

[103] H. Lund, "EnergyPLAN: Advanced energy system analysis computer model," Aalborg University, Denmark, Tech. Rep., 2014.

[104] H. Lund, *Renewable Energy Systems: A Smart Energy Systems Approach to the Choice and Modeling of 100% Renewable Solutions*, 2nd ed. Academic Press.

[105] H. Lund and W. Kempton, "Chapter 5 - analysis: Large-scale integration of renewable energy," in *Renewable Energy Systems (Second Edition)*, H. Lund, Ed. Academic Press, pp. 79–129. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780124104235000055

[106] H. Lund and W. Kempton, "Integration of renewable energy into the transport and electricity sectors through v2g," vol. 36, no. 9, pp. 3578–3587. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0301421508002838

[107] H. Lund, A. N. Andersen, P. A. Østergaard, B. V. Mathiesen, and D. Connolly, "From electricity smart grids to smart energy systems – a market operation based approach and understanding," vol. 42, no. 1, pp. 96–102. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544212002836

[108] H. Lund, N. Duiċ, G. Krajačiċ, and M. d. Graça Carvalho, "Two energy system analysis models: A comparison of methodologies and results," vol. 32, no. 6, pp. 948–954. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S036054420600291X

[109] T. Ma, P. A. Østergaard, H. Lund, H. Yang, and L. Lu, "An energy system model for hong kong in 2020," vol. 68, pp. 301–310. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544214002436

[110] M. S. Mahbub, "A domain knowledge-based multi-objective evolutionary algorithm for optimizing energy systems," in *MENDEL 2014: 20th International Conference on Soft Computing*, vol. 20. [Online]. Available: mendel-conference.org

[111] M. S. Mahbub, M. Cozzini, P. A. Østergaard, and F. Alberti, "Combining multi-objective evolutionary algorithms and descriptive analytical modelling in energy scenario design," *Applied Energy*, vol. 164, pp. 140 – 151, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261915014920

[112] M. S. Mahbub, D. Viesi, and L. Crema, "Designing optimized energy scenarios for an italian alpine valley: the case of giudicarie esteriori," *Energy*, vol. 116, Part 1, pp. 236 – 249, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544216313317

[113] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The annals of mathematical statistics*, pp. 50–60, 1947.

[114] R. Marler and J. Arora, "Survey of multi-objective optimization methods for engineering," *Structural and Multidisciplinary Optimization*, vol. 26, no. 6, pp. 369–395, 2004. [Online]. Available: http://dx.doi.org/10.1007/s00158-003-0368-6

[115] R. Marler and J. Arora, "Survey of multi-objective optimization methods for engineering," *Structural and Multidisciplinary Optimization*, vol. 26, no. 6, pp. 369–395, Apr. 2004. [Online]. Available: http://link.springer.com/10.1007/s00158-003-0368-6

[116] L. Martí, J. García, A. Berlanga, and J. M. Molina, "An approach to stopping criteria for multi-objective optimization evolutionary algorithms: the MGBM criterion," in *IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 1263–1270.

[117] L. Masisi, V. Nelwamondo, and T. Marwala, "The use of entropy to measure structural diversity," in *Computational Cybernetics, 2008. ICCC 2008. IEEE International Conference on*. IEEE, 2008, pp. 41–45.

[118] B. V. Mathiesen, H. Lund, and D. Connolly, "Limiting biomass consumption for heating in 100% renewable energy systems," vol. 48, no. 1, pp. 160–168. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544212006123

[119] B. Mathiesen, H. Lund, D. Connolly, H. Wenzel, P. Østergaard, B. Möller, S. Nielsen, I. Ridjan, P. Karnøe, K. Sperling, and F. Hvelplund, "Smart energy systems for coherent 100*Applied Energy*, vol. 145, pp. 139 – 154, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261915001117

[120] A. Mohammadi, M. N. Omidvar, and X. Li, "A new performance metric for user-preference based multi-objective evolutionary algorithms," in *2013 IEEE Congress on Evolutionary Computation*, June 2013, pp. 2825–2832.

[121] F. G. Montoya, F. Manzano-Agugliaro, S. López-Márquez, Q. Hernández-Escobedo, and C. Gil, "Wind turbine selection for wind farm layout using multi-objective evolutionary algorithms," *Expert Systems with Applications*, vol. 41, no. 15, pp. 6585–6595, Nov. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S095741741400267X

[122] A. J. Nebro, J. J. Durillo, C. A. C. Coello, F. Luna, and E. Alba, "A study of convergence speed in multi-objective metaheuristics," in *Parallel Problem Solving from Nature–PPSN X*. Springer, 2008, pp. 763–772.

[123] A. Nguyen, M. Wagner, and F. Neumann, "User preferences for approximation-guided multi-objective evolution," in *Simulated Evolution and Learning (SEAL)*. Springer, 2014, pp. 251–262.

[124] Nikolas Hill, Sujith Kollamthodi, Adarsh Varma, Stephanie Cesbron, Peter Wells, Shane Slater, Celine Cluzel, Phil Summerton, Hector Pollitt, Sophie Billington, and Terry Ward,, "Fuelling Europe's Future: How auto innovation leads to EU jobs," Cambridge Econometrics, Tech. Rep., 2013. [Online]. Available: http://www.camecon.com/EnergyEnvironment/EnergyEnvironmentEurope/FuellingEuropesFuture.aspx

[125] T. Novosel, B. Ćosić, G. Krajačić, N. Duić, T. Pukšec, M. S. Mohsen, M. S. Ashhab, and A. K. Ababneh, "The influence of reverse osmosis desalination in a combination with pump storage on the penetration of wind and PV energy: A case study for jordan." [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544214003521

[126] M. of Economic Development. Statistiche dell'energia - ministero dello sviluppo economico (energy statistics- ministry of economic development). [Online]. Available: http://dgsaie.mise.gov.it/dgerm/

[127] I. Oropeza-Perez and P. A. Østergaard, "The influence of an estimated energy saving due to natural ventilation on the mexican energy system," vol. 64, pp. 1080–1091. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544213009717

[128] P. A. Østergaard, "Reviewing energyplan simulations and performance indicator applications in energyplan simulations," *Applied Energy*, vol. 154, pp. 921 – 933, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261915007199

[129] P. A. Østergaard, *Baggrundsrapport for Energivision for Aalborg Kommune 2050*. Institut for Samfundsudvikling og Planlægning, Aalborg Universitet. [Online]. Available: http://www.energyplanning.aau.dk/Publications/AalborgKommune-EnergivisionBaggrundsrapport.pdf

[130] P. A. Østergaard, "Geographic aggregation and wind power output variance in denmark," vol. 33, no. 9, pp. 1453–1460. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544208001187

[131] P. A. Østergaard, "Modelling grid losses and the geographic distribution of electricity generation," vol. 30, no. 7, pp. 977–987. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S096014810400374X

[132] P. A. Østergaard, H. Lund, F. Hvelplund, B. V. Mathiesen, A. Remmen, L. M. Odgaard, and B. Möller, *Energivision for Aalborg Kommune 2050.* Institut for Samfundsudvikling og Planlægning, Aalborg Universitet. [Online]. Available: http://www.energyplanning.aau.dk/Publications/AalborgKommune-Energivision.pdf

[133] P. A. Østergaard, "Ancillary services and the integration of substantial quantities of wind power," vol. 83, no. 5, pp. 451–463. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0306261905000619

[134] P. A. Østergaard, "Comparing electricity, heat and biogas storages' impacts on renewable energy integration," vol. 37, no. 1, pp. 255–262. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544211007705

[135] P. A. Østergaard, "Reviewing optimisation criteria for energy systems analyses of renewable energy integration," vol. 34, no. 9, pp. 1236–1245. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544209001777

[136] P. A. Østergaard and H. Lund, "A renewable energy system in frederikshavn using low-temperature geothermal energy for district heating," vol. 88, no. 2, pp. 479–487. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261910000826

[137] P. A. Østergaard, H. Lund, and B. V. Mathiesen, "Energy system impacts of desalination in jordan," vol. 1, no. 0, pp. 29–40. [Online]. Available: http://journals.aau.dk/index.php/sepm/article/view/515

[138] PAES Srl., "Piano di azione per l'energia sostenibile comunità della valle di non," 2015.

[139] J. R. Pillai, K. Heussen, and P. A. Østergaard, "Comparative analysis of hourly and dynamic power balancing models for validating future energy scenarios," vol. 36, no. 5, pp. 3233–3243. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544211001708

[140] A. Pina, C. A. Silva, and P. Ferrão, "High-resolution modeling framework for planning electricity systems with high penetration of renewables," *Applied Energy*, vol. 112, pp. 215 – 223, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S030626191300487X

[141] A. Pina, C. A. Silva, and P. Ferrão, "High-resolution modeling framework for planning electricity systems with high penetration of renewables," *Applied Energy*, vol. 112, pp. 215–223, 2013.

[142] R. D. Prasad, R. C. Bansal, and A. Raturi, "Multi-faceted energy planning: A review," vol. 38, pp. 686–699. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032114004730

[143] P. Prebeg, G. Gasparovic, G. Krajacic, and N. Duic, "Long-term energy planning of Croatian power system using multi-objective optimization with focus on renewable energy and integration of electric vehicles," *Applied Energy*. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261916304123

[144] R. C. Purshouse, K. Deb, M. M. Mansor, S. Mostaghim, and R. Wang, "A review of hybrid evolutionary multiple criteria decision making methods," in *Evolutionary Computation (CEC), 2014 IEEE Congress on*. IEEE, 2014, pp. 1147–1154.

[145] I. Ridjan, B. V. Mathiesen, and D. Connolly, "Synthetic fuel production costs by means of solid oxide electrolysis cells." [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544214004149

[146] D. Roche, D. Gil, and J. Giraldo, "Detecting loss of diversity for an efficient termination of EAs," in *15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, 2013, pp. 561–566.

[147] O. Rudenko and M. Schoenauer, "A steady performance stopping criterion for Pareto-based evolutionary algorithms," in *6th International Multi-Objective Prpgramming and Goal Programming Conference*, 2004.

[148] G. Rudolph, O. Schütze, C. Grimme, and H. Trautmann, "A multiobjective evolutionary algorithm guided by averaged hausdorff distance to aspiration sets." Springer, 2014, pp. 261–273.

[149] B. Saavedra-Moreno, S. Salcedo-Sanz, A. Paniagua-Tineo, L. Prieto, and A. Portilla-Figueras, "Seeding evolutionary algorithms with heuristics for optimal wind turbines positioning in wind farms," *Renewable Energy*, vol. 36, no. 11, pp. 2838–2844, Nov. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S096014811100190X

[150] A. Sadri, M. M. Ardehali, and K. Amirnekooei, "General procedure for long-term energy-environmental planning for transportation sector of developing countries with limited data based on LEAP (long-range energy alternative planning) and EnergyPLAN," *Energy*, vol. 77, pp. 831–843, Dec. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544214011256

[151] F. Sáfián, "Modelling the hungarian energy system – the first step towards sustainable energy planning," vol. 69, pp. 58–66. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544214002011

[152] S. Salcedo-Sanz, D. Gallo-Marazuela, A. Pastor-Sánchez, L. Carro-Calvo, A. Portilla-Figueras, and L. Prieto, "Evolutionary computation approaches for real offshore wind farm layout: A case study in northern Europe," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6292–6297, Nov. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417413003527

[153] J. D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms," in *Proceedings of the 1st International Conference on Genetic Algorithms*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1985, pp. 93–100. [Online]. Available: http://dl.acm.org/citation.cfm?id=645511.657079

[154] O. Schutze, X. Esquivel, A. Lara, and C. A. Coello Coello, "Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 4, pp. 504–522, 2012.

[155] O. M. Shir, M. Preuss, B. Naujoks, and M. Emmerich, "Enhancing decision space diversity in evolutionary multiobjective algorithms," in *Evolutionary Multi-Criterion Optimization*. Springer, 2009, pp. 95–109.

[156] A. R. Solow and S. Polasky, "Measuring biological diversity," *Environmental and Ecological Statistics*, vol. 1, no. 2, pp. 95–103, 1994.

[157] N. Srinivas and K. Deb, "Muiltiobjective optimization using nondominated sorting in genetic algorithms," *Evol. Comput.*, vol. 2, no. 3, pp. 221–248, Sep. 1994. [Online]. Available: http://dx.doi.org/10.1162/evco.1994.2.3.221

[158] P. A. Østergaard, "Heat savings in energy systems with substantial distributed generation," vol. 23, no. 4, pp. 169–176. [Online]. Available: http://vbn.aau.dk/en/publications/heat-savings-in-energy-systems-with-substantial-distributed-generation(09759160-cd71-11db-9e19-000ea68e967b).html

[159] P. A. Østergaard, "Reviewing energyplan simulations and performance indicator applications in energyplan simulations," *Applied Energy*, vol. 154, pp. 921 – 933, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261915007199

[160] P. A. Østergaard, "Regulation strategies of cogeneration of heat and power (CHP) plants and electricity transit in denmark," vol. 35, no. 5, pp. 2194–2202. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544210000575

[161] P. A. Østergaard, "Wind power integration in aalborg municipality using compression heat pumps and geothermal absorption heat pumps," vol. 49, pp. 502–508. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544212008912

[162] P. A. Østergaard and K. Sperling, "Towards sustainable energy planning and management," vol. 1, no. 0, pp. 1–5. [Online]. Available: http://journals.aau.dk/index.php/sepm/article/view/559

[163] R. Tran, J. Wu, C. Denison, T. Ackling, M. Wagner, and F. Neumann, "Fast and effective multi-objective optimisation of wind turbine placement," in *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '13. ACM, pp. 1381–1388. [Online]. Available: http://doi.acm.org/10.1145/2463372.2463541

[164] H. Trautmann, U. Ligges, J. Mehnen, and M. Preuss, "A convergence criterion for multiobjective evolutionary algorithms based on systematic statistical testing," in *Parallel Problem Solving from Nature–PPSN X*. Springer, 2008, pp. 825–836.

[165] H. Trautmann, T. Wagner, B. Naujoks, M. Preuss, and J. Mehnen, "Statistical methods for convergence detection of multi-objective evolutionary algorithms," *Evolutionary computation*, vol. 17, no. 4, pp. 493–509, 2009.

[166] Turboden, "Turboden Biomass solutions," 2015. [Online]. Available: http://www.turboden.eu/en/public/downloads/12-COM.P-18-rev.17.pdf

[167] T. Ulrich, "Exploring structural diversity in evolutionary algorithms," Ph.D. dissertation, ETH Zurich, 2012.

[168] Unione Petrolifera, "Relazione annuale 2014 (annual report 2014)." [Online]. Available: http://www.unionepetrolifera.it/?page_id=38

[169] K. Vaillancourt, Y. Alcocer, O. Bahn, C. Fertel, E. Frenette, H. Garbouj, A. Kanudia, M. Labriet, R. Loulou, M. Marcy, Y. Neji, and J.-P. Waaub, "A Canadian 2050 energy outlook: Analysis with the multi-regional model TIMES-Canada," *Applied Energy*, vol. 132, pp. 56–65, Nov. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261914006540

[170] M. Wagner, K. Bringmann, T. Friedrich, and F. Neumann, "Efficient optimization of many objectives by approximation-guided evolution," *European Journal of Operational Research*, vol. 243, no. 2, pp. 465 – 479, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0377221714009552

[171] T. Wagner, H. Trautmann, and B. Naujoks, "OCD: Online convergence detection for evolutionary multi-objective algorithms based on statistical testing," in *Evolutionary Multi-Criterion Optimization*, ser. Lecture Notes in Computer Science. Springer, 2009, vol. 5467, pp. 198–215. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-01020-0_19

[172] A. P. Wierzbicki, *Reference Point Approaches*. Boston, MA: Springer US, 1999, pp. 237–275. [Online]. Available: http://dx.doi.org/10.1007/978-1-4615-5025-9_9

[173] M. Wineberg and F. Oppacher, "The underlying similarity of diversity measures used in evolutionary computation," in *Genetic and Evolutionary Computation-GECCO 2003*. Springer, 2003, pp. 1493–1504.

[174] R. Wurbs, "Reservoir-system simulation and optimization models," vol. 119, no. 4, pp. 455–472. [Online]. Available: http://dx.doi.org/10.1061/(ASCE)0733-9496(1993)119:4(455)

[175] X. Yu, *Introduction to evolutionary algorithms*, ser. Decision engineering. Springer.

[176] Q. Zhang and H. Li, "MOEA/d: A multiobjective evolutionary algorithm based on decomposition," vol. 11, no. 6, pp. 712–731.

[177] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.

[178] E. Zitzler and L. Thiele, "An evolutionary algorithm for multi-objective optimization: The strength pareto approach," Computer Engineering and Communication Networks Lab TIK, Swiss Federal Institute of Technology (ETH), Gloriastrasse 35, CH-8092, Zurich, Switzerland, Tech. Rep. TIK-Report 43, May 1998.

[179] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evolutionary Computation*, vol. 8, no. 2, pp. 173–195, 2000. [Online]. Available: http://dx.doi.org/10.1162/106365600568202

[180] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength pareto evolutionary algorithm," Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH), Zurich, Tech. Rep. TIK-Report No. 103, 2001.

[181] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. Da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, 2003.