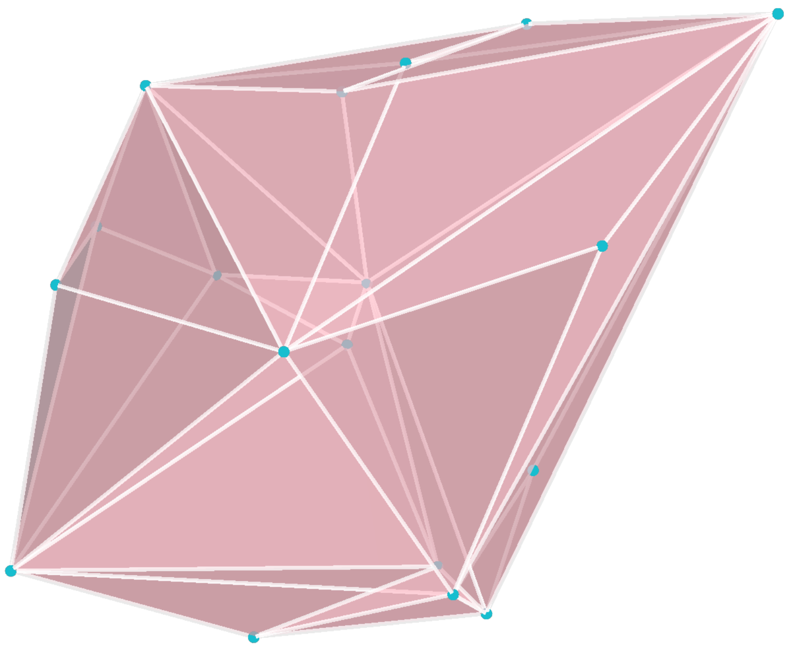
A novel modeling to generate alternatives approach: Determining the convex hull containing all near optimal solutions

Tim T. Pedersen

January 2020



Title page

Title:
A novel modeling to generate alternatives approach: Determining the convex hull containing all near optimal solutions

Project:
Mester's Thesis, Department of Engineering , Aarhus University.

Date finished:
3rd January 2020

Author:
Tim Tørnes Pedersen - 201403848 - au518895

Project supervisors:
Gorm B. Andresen, Associate professor, Department of Engineering, Aarhus University
Marta Victoria, Assistant Professor, Department of Engineering, Aarhus University

This project is lawfully written by the following author:

Abstract

To limit the extent of irreversible climate change and accepting public opinion expressed in Greta Thunberg's speech at the UN Climate Action summit, drastic measures are needed to reduce the emission of greenhouse gases. The majority of CO_2 emitted by humans are results of energy production to cover the ever-rising energy demand including transportation, heating, and electricity. To assist scientists and policymakers in their stride to reach ambitious goals in the reduction of CO2 emissions, analysis tools must be developed. An important tool, when it comes to the planning of global and local energy networks are numeric techno-economic energy models. These models are capable of providing great insight into complex systems such as the European electricity grid and allow the user to make predictions about future needs and design strategies.

Numeric energy-economic models do however suffer from great uncertainties arising from flaws in the mathematical formulation and construction of the energy-economic model referred to as structural uncertainty. An example of flaws in the mathematical formulation could be unmodeled constraints such as public acceptance issues. If these uncertainties are not addressed, the results become untrustworthy and end up providing little to no insight. Until recently, no methods for addressing structural uncertainty of the techno-economic models existed. This changed in 2010 when J. DeCarolis published a paper proposing a technique called "Modeling to generate alternatives (MGA)" doing just so. The root cause of structural uncertainty cannot be addressed, as the origin of structural uncertainty is hard to define. Instead, one must investigate all solutions near the one found to be optimal, and estimate the likelihood of these near-optimal solutions being the true optimal solution.

The proposed technique by J. DeCarolis does, however, suffer from a range of flaws, arising from lacking structure in the manner near-optimal solutions are found. To obtain a complete picture of all near-optimal solutions, a structured method of finding these is needed. The objective of this thesis is to explore the characteristics of all near-optimal solutions contained within the near-optimal feasible decision space and to develop a new technique that in a structured manner can explore all solutions located within this space.

Analysis of the common mathematical formulation of the numeric techno-economic model reveals that the model consists of linear constraints and therefore, the near-optimal feasible decision space, containing all near-optimal solutions to the model, must be convex. Knowing these properties, a technique has been developed capable of searching the entire near-optimal feasible decision space. The technique iteratively converges towards the full solution and provides statistical information about all near-optimal solutions. Furthermore, a method reducing the complexity of the mathematical problem, by a grouping of variables is proposed. Grouping the variables in the model to form a new set of variables does however reduce the amount of information obtained by solving this simplified problem. The effects of grouping the model variables are explored, and the effect is found to be significant, but predictable. The developed method is applied to a model of the European electricity grid. The usefulness of the technique is proven as it provides information about the distribution of technology capacities in all near-optimal solutions to the used model of the European electricity grid.

Part I

Main Report

1 Introduction

As climate change and the effects thereof have become more and more evident in recent years, the effort to reduce the emission of greenhouse gases has increased. This comes to show in ambitious goals of reduced emissions on a national and international scale. The majority of greenhouse gas emissions arise from the energy sector including transportation, heating/cooling, and electricity demand [1], therefore efforts must be put into reducing emissions from the energy sector.

To achieve reductions in greenhouse gas emissions from the energy sector, large amounts of variable renewable energy sources, such as wind and solar energy is being implemented in the energy grid. The volatile nature of these variable renewable energy sources drives the need for collaboration between countries and energy sectors to handle periods of scarce resources and fluctuations in energy demand. This complicates the already complex task of energy system synthesis even further, hereby requiring decision-makers to have greater in-depth knowledge. Therefore, the need for analysis tools providing insights into the constraints and possibilities of the complex energy systems has never been more present.

A frequently used tool to gain insight in future energy grid compositions is numeric techno-economic energy models on either regional, national or international scale. These models can be used to study the behavior and composition of existing and future energy networks, together with the impact of new technologies or structural changes in the networks [2]. The models often approximate the behavior of the physical systems, through a series of mathematical constraints based on energy laws. By inputting time series for wind and solar availability and energy demand, the behavior of the energy network can be studied with alternative configurations of energy generating and storage technologies. Using an optimization algorithm, it is possible to find the cheapest configuration of the energy network that satisfies all user prescribed constraints.

Techno-economic energy models do, however, suffer from large uncertainties and the lack of validation possibilities, resulting in unreliable and therefore less informative results. Model uncertainty can either arise from uncertainty in input parameters and data, such as uncertainties in future technology prices and energy demand. This type of uncertainty is referred to as parametric uncertainty. Parametric uncertainty is well understood and often studied with sensitivity analysis or Monte Carlo simulations. A different type of uncertainty introduced by an incomplete or faulty mathematical description of the problem has been found to have just as large effects on the results produced by these techno-economic models [3]. As this type of uncertainty relates to the mathematical foundation or structure of the model, they are referred to as structural uncertainty. Structural uncertainty is inevitable as it is impossible to create completely accurate mathematical models representing physical systems. A common source of structural uncertainty is unmodeled objectives such as public acceptance issues or political ambitions. Furthermore, the economy of energy systems, which is often the parameter to be optimized, is highly influenced by politics and public opinion. It is easy to imagine that a scenario with large amounts of aerial electricity transmission lines would be met with public resistance, and therefore a scenario with less transmission at the same cost might be favorable. This does in term mean that the model objective is not only to reduce system cost but also to satisfy as many of the involved stakeholders as possible. It is however impossible to model the satisfaction of all involved stakeholders, let alone the challenge of identifying all future parties involved.

Recently an approach addressing the structural uncertainty of the techno-economic models has been proposed by J. DeCarolis. In an article [3] a technique called Modeling to Generate Alternatives (MGA), from the field of management research/planning science [4], is applied to the field of energy planning. The root cause of structural uncertainty cannot be addressed as it can with parametric uncertainty, as the origin of structural uncertainty is hard to define. Instead, one must investigate all solutions near the one found to be optimal, and estimate the likelihood of these near-optimal solutions being the true optimal solution. In the technique proposed by J. DeCarolis, a finite set of maximally different near-optimal solutions are found. The difference in the found solutions can then be used as a measure of structural uncertainty and provides a variety of alternatives to the optimal configuration of the energy system. The concept of using MGA algorithms on energy planning problems have been further studied and the result presented in a range of articles and papers; [3], [5], [6], [7], [9] and [10].

The MGA technique introduced by Brill et al. [4] and implemented on a techno-economic energy model by J. DeCarolis [3], is referred to as the Hop Skip Jump (HSJ) MGA algorithm. It will produce a small number of alternative near-optimal solutions, by altering the objective function of the optimization algorithm. Instead of minimizing cost, the objective is changed such that the model seeks to implement previously unused technologies. In order to ensure that the found alternative solutions are near the optimum, a new constraint is added to the model. This constraint allows for the new MGA solutions to deviate in cost by a certain amount from the optimal solution. This constraint is referred to as the MGA constraint.

The alternative solutions found with the HSJ MGA algorithm, do provide some insights into the characteristics of all near-optimal solutions, but it is not possible to determine if all possible solutions have been found. Furthermore, the HSJ method of finding alternative solutions is somewhat random and the found alternative solutions are highly dependent on the starting point. In order to fully understand the characteristics of all near-optimal solutions, a more structured method is needed.

In this project the model presented in [11] of the European electricity grid, will serve as a base model, to use for testing and validation. The model is built in the open-source framework PyPSA [12], and formulated as a techno-economic linear optimization problem. The objective in the model is to minimize total annual system cost while satisfying a range of constraints ensuring feasible operation. The model groups the European electricity network into 30 nodes, each one representing a single country. Countries are linked with power lines approximating the current layout of the European transmission grid. Each node in the network, will in this project, only be granted access to three electricity-generating technologies and no storage technologies, simplifying the network drastically compared to the configuration used in [11]. The energy-generating technologies chosen are open cycle gas turbines (OCGT), wind turbines and solar photovoltaic (solar PV).

The current use of techno-economic models combined with optimization tools, not using MGA algorithms, provides a very rigid solution to the future configuration of energy networks. Usually, a single optimal solution is found and used as an end goal to strive towards. This method of doing energy planning provides little guidance, in the case of unforeseen events and changing objectives. Using the MGA approach presented by J. DeCarolis, it is possible to have a set of alternative near-optimal solutions providing guidance if the initial optimal solution, should become unachievable or unfeasible. Having the characteristics of all near-optimal solutions available, it would be possible to define a tolerance on the optimum, and to provide details on minimum required capacities and must-have technologies. Furthermore, it would be possible to provide a cost estimate for deviating from the optimal design, guiding decision-makers in the case of unforeseen events and changing objectives. Having this information would add value to techno-economic optimization studies, as the result would be relevant for longer periods of time, requiring less frequent reevaluations of the studies.

In this project the MGA concept will be further explored, in an attempt to map the entire set of near-optimal solutions, providing a detailed description of all feasible near-optimal configurations of a given techno-economic model. The objective of this project is to develop a structured method of investigating the characteristics of all feasible near-optimal solutions to a techno-economic optimization problem. The developed method should be verified on the techno-economic model from [11] and the quality of the insights provided should be discussed.

The structure of this project will be as follows. Initially, a detailed analysis of the mathematical formulation of a techno-economic model will be presented. The properties of the mathematical constraints used in the model are analyzed together with the objective function. Having established the formulation of the model, the characteristics of the set containing all near-optimal solutions will be presented. The use of optimization algorithms used to find optimal solutions in this set will briefly be discussed. With a good understanding of the techno-economic optimization problem, the working principles of existing MGA algorithms will be introduced followed by a presentation of the MGA method developed in this project. In chapter 3, the techno-economic model used in this project will be presented, explaining the implemented technologies and accounting for input data. The results of the computational experiments performed in this project, using the reference model and the presented MGA algorithm, will be presented in chapter 4. Several experiments have been performed highlighting the flexibility and usability of the proposed MGA algorithm. The found results will be discussed in chapter 5, analyzing the characteristics of the techno-economic model of Europe used in this project, together with an evaluation of the MGA algorithm itself. Finally, the project will be concluded upon, highlighting relevant findings and achieved objectives.

2 Theory

The purpose of this chapter is to explain the mathematical concepts behind the numerical models used and to provide useful insight into the working concepts of the algorithms developed. Initially, the constraints defining a techno-economic energy model is presented to provide a good understanding of the optimization problem. Next, the working principles of existing MGA algorithms are presented, as they serve as the foundation for the MGA method developed in this project.

In this project, the following mathematical formulations will be used. Scalar values will be non-bold characters such as d. Vectors will be represented by bold characters $\mathbf{x}$. Single values in a vector are indexed using a subscript $\mathbf{x}_i$, as in this example where the i'th element of $\mathbf{x}$ is represented. Vectors of higher dimensions will be indexed with as many variables as dimensions. For example, a single value from a three-dimensional vector may be accessed as $\mathbf{g}_{n,s,t}$. A specific point given as a vector is indexed with superscript such that $\mathbf{x}^1$, refers to a specific configuration of $\mathbf{x}$. Sets will be assigned non-bold capital letters such as W.

2.1 Techno-Economic model mathematical formulation

To fully understand the techno-economic model, the mathematical formulation of the model must first be understood. In this section, the objective function and the constraints defining the numeric techno-economic model will be presented and explained.

The goal is to formulate the techno-economic model in the form of a classic optimization problem as explained in [13], where an objective function is defined along with a set of constraints on the form presented in Equation 2.1. All constraints are collected in the vector functions $\mathbf{f}_i$ and $\mathbf{h}_i$, and are rewritten to be either less than or equal to 0. The vector $\mathbf{x} \subseteq \mathbb{R}^d$ contain the optimization variables to be optimized.

$$\begin{aligned} \text{minimize} \quad & \mathbf{f}_0(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{f}_i(\mathbf{x}) \leq 0 \quad i = 1..m \\ & \mathbf{h}_i(\mathbf{x}) = 0 \quad i = 1..p \end{aligned} \tag{2.1}$$

It is possible to formulate the techno-economic optimization problem in this form by studying the individual components of the model and defining their individual relations with constraints. But first of all the components in the techno-economic model must be identified.

The techno-economic energy model is built as a network where each country is represented as a node connected to the surrounding countries through a link. Each country/node has three energy-producing technologies available, gas turbines (OCGT), wind turbines, solar PV and an energy load that must be satisfied. In the model a range of snapshots is defined, each representing an hour of the year. For each hour all energy demands must be satisfied and to do so the installed capacities of the available technologies can be increased if necessary, at the expense of added system cost.

Following the naming convention from [11], indexing the nodes in the network with the variable n, the power generating technologies by s, the hours in the year by t and the possible connecting power lines by l, all variables to be optimized can be defined as:

- $\mathbf{g}_{n,s,t}$: Hourly dispatch of energy from the given plants in the given countries with the marginal cost $\mathbf{o}_{n,s}$.
- $\mathbf{G}_{n,s}$: Total installed capacity of the given technologies in the given countries with the capital cost $\mathbf{c}_{n,s}$.
- $\mathbf{F}_l$: Total installed transmission capacity for all lines with the fixed annualized capacity cost $\mathbf{c}_l$.

Thus, the optimization variables $\mathbf{x}$ becomes:

$$\mathbf{x} = \{\mathbf{g}_{n,s,t}, \mathbf{G}_{n,s}, \mathbf{F}_l\} \tag{2.2}$$

The objective for the optimization of the techno-economic model, is to reduce total system cost, leading to the following formulation of the objective function:

$$\text{minimize } f_0(\mathbf{x}) = \sum_{n,s} \mathbf{c}_{n,s}\mathbf{G}_{n,s} + \sum_{l} \mathbf{c}_l\mathbf{F}_l + \sum_{n,s,t} \mathbf{o}_{n,s}\mathbf{g}_{n,s,t} \tag{2.3}$$

This objective function is subject to a range of constraints ensuring realistic behavior of the system. As described in [11] a power balance constraint is issued to ensure stable operation of the network, by requiring that the sum of produced and consumed energy sum to zero for each time step. The hourly electricity demand at each node is described by $\mathbf{d}_{n,t}$, the incidence matrix describing the line connections is given by $\mathbf{K}_{n,l}$ and the hourly transmission in each line is described as $\mathbf{f}_{l,t}$. Then the power balance constraint becomes:

$$\sum_{s} \mathbf{g}_{n,s,t} - \mathbf{d}_{n,t} - \sum_{l} \mathbf{K}_{n,l}\mathbf{f}_{l,t} = 0 \;\; \forall n, t \tag{2.4}$$

Where the first term represents all generated energy, the second term describes energy demand and the last term provides the transmission between the individual nodes. It is important to note that the transmission lines are modeled without transmission loss. All transmission lines in the model are modeled with a controllable dispatch constrained by the fact that there must be energy conservation at each node the line is connected to.

For all conventional generators, the maximum hourly dispatch of energy is limited by the installed capacity. It is important to note that for all simulations performed in this project the installed capacity is a variable.

$$0 \leq \mathbf{g}_{n,s,t} \leq \mathbf{G}_{n,s} \; \forall n, s, t \tag{2.5}$$

Rewriting this equation to be of the form presented in 2.1 it becomes two constraints.

$$\begin{aligned} -\mathbf{g}_{n,s,t} &\leq 0 \; \forall n, s, t \\ \mathbf{g}_{n,s,t} - \mathbf{G}_{n,s} &\leq 0 \; \forall n, s, t \end{aligned} \tag{2.6}$$

The dispatch of variable renewable energy sources (wind and solar) is not only limited by the installed capacity, as availability, hence the name, is variable. Therefore, the constraint for dispatch of variable renewable energy sources become:

$$0 \leq \mathbf{g}_{n,s,t} \leq \overline{\mathbf{g}}_{n,s,t} \mathbf{G}_{n,s} \; \forall n, s, t \tag{2.7}$$

This constraint must also be rewritten to the correct form, and thereby becomes:

$$\begin{aligned} -\mathbf{g}_{n,s,t} &\leq 0 \; \forall n, s, t \\ \mathbf{g}_{n,s,t} - \mathbf{G}_{n,s} \overline{\mathbf{g}}_{n,s,t} &\leq 0 \; \forall n, s, t \end{aligned} \tag{2.8}$$

Here $\overline{\mathbf{g}}_{\mathbf{n,s,t}}$ represents the normalized availability per unit capacity.

The installed capacity is constrained by the geographical potential ensuring that unrealistic capacities are not installed in favorable countries.

$$0 \leq \mathbf{G}_{n,s} \leq \mathbf{G}_{n,s}^{max} \; \forall n, s \tag{2.9}$$

Rewriting this constraint to the desired form it becomes:

$$\begin{aligned} -\mathbf{G}_{n,s} &\leq 0 \; \forall n, s \\ \mathbf{G}_{n,s} - \mathbf{G}_{n,s}^{max} &\leq 0 \; \forall n, s \end{aligned} \tag{2.10}$$

As energy dispatch from the energy generating technologies is limited by the installed capacity, so is the transmission in the individual lines limited by the installed capacity.

$$|\mathbf{f}_{l,t}| \leq \mathbf{F}_l \; \forall l, t \tag{2.11}$$

$$|\mathbf{f}_{l,t}| - \mathbf{F}_l \leq 0 \; \forall \; l, t \tag{2.12}$$

There is no limit on the maximum allowable transmission capacity installed.

In the model it is possible to activate a CO_2 constraint, limiting the allowed CO_2 emissions for the entire energy network. As in [11] the constraint is implemented using the specific emissions $\mathbf{e}_s$ in CO_2-tonne-per-MWh of the fuel for each generator type s, with the efficiency η_s and the CO_2 limit CAP_{CO_2}.

$$\sum_{n,s,t} \frac{1}{\eta_s} \mathbf{g}_{n,s,t} \mathbf{e}_s - CAP_{CO_2} \leq 0 \; \forall \; n, s, t \tag{2.13}$$

The full set of constraints needed to define a techno-economic model is now defined together with the objective function. It is seen that the objective function and all constraints are linear. The optimization of this problem hereby falls under a category of optimization problems, called linear problems. These problems allow the use of special optimization solvers, optimized to solve linear problems and have several characteristics that will be used later in this project.

Using this set of constraints and optimizing the total system cost for an entire year of energy production, means that this model assumes perfect foresight, as weather and demand for the entire year are known to the optimization algorithm. The effect of assuming perfect foresight is however not that critical when no storage technologies are implemented. As all capital and marginal prices used in this model are constants, this model also assumes perfect competition and long-term market equilibrium. Meaning that over the entire simulation period, the technologies recover their total cost by their hourly market revenues (capital and marginal).

2.2 Properties of the near-optimal feasible space

In this section, the properties of the near-optimal feasible space of the optimization problem will be discussed. If the number of decision variables in $\mathbf{x}$ is d, then all configurations of $\mathbf{x}$ is contained in $\mathbf{x} \in \mathbb{R}^d$. The set containing all $\mathbf{x}$ is referred to as the decision space and is formulated as:

$$W \subseteq \mathbb{R}^d \tag{2.14}$$

All points $\mathbf{x}$ in W that satisfies the constraints $\mathbf{f}_i(\mathbf{x}) \leq 0$ and $\mathbf{h}_i(\mathbf{x}) = 0$, is said to be feasible. The set containing all feasible solutions can be defined as:

$$X = \{\mathbf{x} \mid \mathbf{f}_i(\mathbf{x}) \leq 0 | \mathbf{h}_i(\mathbf{x}) = 0\} \tag{2.15}$$

This set X is what is referred to as the feasible set or feasible space. The feasible set X is a subset of W and the dimension of X might be lower than d, as the equality constraints from Equation 2.4, force the solutions to be located on a hyperplane in W.

Analyzing the original optimization problem one can deduct that the set including all feasible solutions X, given by Equation 2.15, must be convex. This is due to the fact that all constraints $\mathbf{f}_i$, $\mathbf{h}_i$ and the objective function f_0, are linear and therefore satisfy Equation 2.16, thus ensuring convexity [13].

$$\mathbf{f}_i(\alpha\mathbf{x} + \beta\mathbf{y}) \leq \alpha\mathbf{f}_i(\mathbf{x}) + \beta\mathbf{f}_i(\mathbf{y}) \;\; \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^d \text{ and } \alpha, \beta \in \mathbb{R} \tag{2.16}$$

Linear problems further have the characteristic that any optimal solution will be located on the boundary of the decision space as shown in theorem 8.2 in [14].

Because all constraints are linear, the shape of the decision space will be given by a polyhedral, as all boundaries are linear. This means that the entire volume of the decision space can be described with a finite set of vertices. When performing MGA optimization, to describe the entire decision space, this feature is very important.

It is important to note that the variables $\mathbf{x}$ that defines the decision space variables in the original problem include all hourly technology dispatches $\mathbf{g}$, all installed technology capacities $\mathbf{G}$ and all installed line capacities $\mathbf{F}$ as shown in Equation 2.2.

Therefore, the dimensionality of the decision space must be given by the number of individual dispatch decisions given by $n \cdot s \cdot t$ plus the number of capacities to optimize, given by $n \cdot s$ plus the number of line capacities to optimize l. The dimensionality of the full problem is therefore given by Equation 2.17.

$$d = n \cdot s \cdot t + n \cdot s + l \tag{2.17}$$

In the case of the reference model used in this project where 30 countries, 3 technologies, 52 transmission lines and all 8765 hours of the year are included, the dimensionality becomes:

$$30 \cdot 3 \cdot 8765 + 30 \cdot 3 + 52 = 788992$$

Performing optimization on a problem with this many variables with the goal of finding a single optimal solution is feasible, but exploring the near-optimal feasible space in this high dimension will be difficult. Therefore, a method of reducing the number of decision variables is needed.

Dimensionality reduction

For the applications of this project, a decision space of lower dimension is desired. Therefore, a method of reducing the dimensionality must be found. One option is to ignore the hourly dispatch of energy from the individual generators, thereby reducing the dimensionality by a substantial amount. Ignoring all information about hourly plant operation will reduce the dimension of the decision space by a significant amount, and as the focus of this report is to investigate the composition of technologies, rather than operational tactics, losing the information on hourly energy dispatches is of little concern. It is however important that all solutions still satisfy the original constraints. Choosing to ignore the variable concerning hourly dispatch of energy, a new vector containing all variables can be formed as:

$$\mathbf{x}^* = \{\mathbf{G}_{n,s}, \mathbf{F}_l\} \tag{2.18}$$

This new vector $\mathbf{x}^*$ is containing only some of the elements of the original $\mathbf{x}^*$, and is thereby a sub-vector $\mathbf{x}^* \in \mathbf{x}$.

Requiring that all solutions satisfy all the constraints of the original problem, the new reduced decision space can be defined as:

$$W^* = \{\mathbf{x}^*(\mathbf{x}) | \mathbf{x} \in W\} \tag{2.19}$$

As $\mathbf{x}^*$ is a sub-vector of $\mathbf{x}$, W^* is also a subspace of W, $W^* \subseteq W$. The length of the vector $\mathbf{x}^*$ and thereby the dimension of the new reduced decision space W^* then becomes:

$$d^* = d \cdot s + l \tag{2.20}$$

Applying this to the model used in this project, the dimensionality would only be $d^* = 30 \cdot 3 + 52 = 142$. This is a very significant reduction compared to the dimension of the true solutions with more than half a million decision variables.

Ignoring all hourly dispatch, but requiring that all constraints of the initial problem are still satisfied, means that the model and optimization problem itself has not changed. But instead of analyzing the entire decision space with MGA algorithms, only a subspace W^*, is considered.

Further reduction by a grouping of variables

If desired it is possible to further reduce dimensionality. This is done by grouping variables by summation, into a new set of variables. An example would be to group all individual technologies in groups containing all installed capacity of that given technology across the entire network. Three new variables would then be formed, with one containing the entire sum of installed wind capacity, one with solar capacity and one with OCGT capacity. The vector containing all variables then becomes:

$$\mathbf{x}^{**} = \{\sum_n G_{n,s} \; \forall \; s\} \tag{2.21}$$

As it is still required that all solutions satisfy the constraints of the original problem, the set containing all summed variables then becomes:

$$W^{**} = \{\mathbf{x}^{**}(\mathbf{x}) | \mathbf{x} \in W\} \tag{2.22}$$

The dimension of the set containing all summed variables therefore becomes:

$$d^{**} = s \tag{2.23}$$

Which in this project is only 3. This is now a very manageable dimension size, ideal to perform MGA analysis on. If desired other variables could be included in the set W^{**}, such as summed transmission capacity.

2.3 Numeric optimization

The core of this project is a numeric optimization problem, and therefore the relevant optimization theory should be understood. The aim of this section is not to provide a full walkthrough of the optimization algorithms used in the project, instead, an introduction to relevant concepts and methods will be provided, giving the necessary insights.

In this project, the optimization tool Gurobi [15], have been used in combination with the open-source energy modeling tool PyPSA [12]. Gurobi offers a wide range of optimization solvers, with their individual advantages and disadvantages. Because the optimization problem addressed in this project is linear, it is possible to use the simplex method, which utilizes some of the characteristics of linear problems to elegantly find optimal solutions.

A linear problem has two important characteristics. Namely that the feasible space is convex, and the fact that the optimal solution will be located on the boundary, as stated in theorem 8.1 and 8.2 in [14]. Using this information the simplex method is capable of finding the optimal solution using the Gauss-Jordan elimination method as explained in [14]. The benefit of using the simplex method is that it requires no calculation of the objective function gradient, but instead relies on simple matrix row operations.

Alternatively, one could use the barrier method to find the optimal solution. The barrier method transforms the constrained problem into an unconstrained one by adding a penalty term to the objective function. The penalty term ensures that the constraints are satisfied, as the penalty term drastically increases as the constraints are violated. Once the problem has been converted to an unconstrained one, gradient-based methods such as gradient descent or Newton's method [14], can be used to find the optimum. Although the barrier method requires fewer iterations to converge than the simplex method, the added computational cost of calculating the objective function gradient makes this method unfavorable.

2.4 Modeling to Generate Alternatives

In this section, the basic principles of modeling to generate alternatives (MGA) will be explained together with the benefits and challenges this technique introduces. MGA was first introduced by E. D. Brill in 1982 [4], for use in the field of resource planning. This new method was designed to address the large amounts of uncertainty in resource planning problems, by proposing alternative near-optimal solutions, to the optimization problem. The field of resource planning and techno-economic optimizations both share the problem of having a lot of stakeholders with unclear objectives, which makes the use of MGA very useful in both fields.

Motivation for using MGA

In the field of mathematical modeling, the scientist aims to produce models representing physical systems as realistically as possible. However, some degree of uncertainty in the models is inevitable as model fidelity is limited by a range of factors including numeric precision, the uncertainty of data, model resolution, etc. Modeling of energy systems is a field especially prone to large model uncertainties, deriving not only from lack of fidelity but from factors such as unmodeled objectives and structural uncertainty [3]. The unmodeled objectives derive from desires and opinions from all involved stakeholders such as the public, private companies, and policy-makers. As for the structural uncertainty it originates from the fact that predicting the exact layout of the future energy grid is an almost impossible task.

The core problem can be summarized as follows: Since structural uncertainty will always exist as it is impossible to produce a complete mathematical model of this type of problems, the ideal solutions are much more likely to lie within the models' inferior region, rather than at a single optimal point, [4]. The purpose of MGA is therefore to search the inferior region of the decision space surrounding the optimal solution, to gain insight into the solutions located in this region.

The concept of MGA covers a wide range of techniques to search the near-optimal feasible decision space, approaching the problem in different manners. All solutions presented in literature so far does, however, share the fact that, instead of providing general insight on the near-optimal decision space, they seek to provide a finite set of maximally alternative solutions from this region. Although it is very easy to relate to a few alternative solutions, these methods fail to provide any general information about the characteristics of the near-optimal solutions.

Technical explanation of the Hop Skip Jump algorithm

The term "modeling to generate alternatives" (MGA), was first introduced in 1982 by Brill et. al in the article [4] and later rediscovered by DeCarolis in [3] for use in energy system optimization. The presented method named "Hop Skip Jump" or HSJ iteratively produces maximally different solutions located within the near-optimal feasible space, by altering the objective function of the original problem. The technique lets the user search the near-optimal feasible decision space of an optimization problem such as the one addressed in this project described in Section 2.1, and presents a limited set of alternative solutions to the original problem.

To define the near-optimal feasible space, the extent of the full decision space must be constrained, to only include the near-optimal solutions. In all work performed on MGA, the term "near-optimal" refers to all solutions that have an objective value near that of the optimal solution. This means that to define the near-optimal space, the optimal solution must be found first.

By solving the optimization problem defined in Equation 2.1 the optimal solution $\mathbf{x}^0$ is found with the objective value $f_0(\mathbf{x}^0)$. The near optimal criteria can then be defined as all solutions with an objective value near that of the optimal solution. Written as a constraint this becomes:

$$f_0(\mathbf{x}) \leq f_0(\mathbf{x}^0) \cdot (1+\epsilon) \;\; \forall \, \mathbf{x} \in W \tag{2.24}$$

The constant ϵ is the MGA slack that determines the range of the near-optimal feasible space. ϵ is often chosen to lie within the range of 0 to 0.1, equivalent to 10% of slack on the objective value. A simple sketch illustrating this new constraint is shown in Figure 2.1.

The introduction of this new constraint often referred to as the MGA constraint is universal for all MGA approaches, as it is essential to limit the decision space. The different MGA methods do however use different techniques to explore this new, near-optimal feasible space, and in this section, the HSJ method presented in [4] will be presented.

The goal of the HSJ MGA technique is to explore a finite set of maximally different alternative solutions located within the near-optimal feasible space. In the original article by Brill et. al [4] the HSJ MGA technique is described with the following steps:

1. Obtain an initial optimal solution for the problem at hand $f_0(\mathbf{x}^0)$.
2. Define target value, Equation 2.27, for the objective function by adding a user-specified amount of slack ϵ to the value of the objective function in the initial solution.

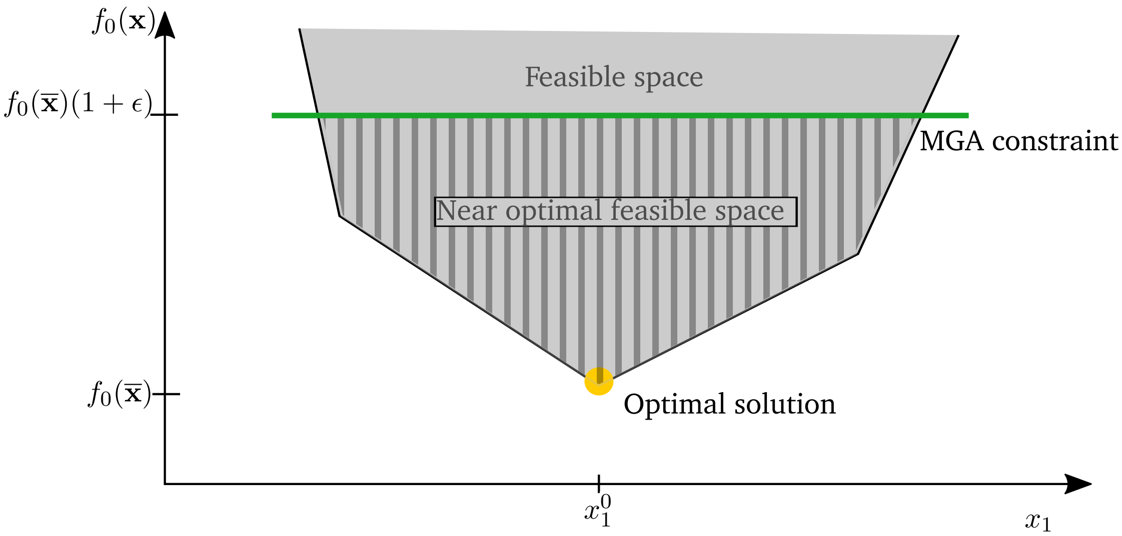


Figure 2.1: A sketch showing the near optimal feasible space constrained by the MGA constraint.

3. Introduce the constraint limiting the objective function to surpass this target value, to the model (Equation 2.26).
4. Formulate a new objective function that seeks to minimize the sum of decision variables that had non zero values in the previous solution of the problem (Equation 2.25).
5. Iterate the reformulated problem, updating the objective function every time.
6. Terminate the optimization when the new solution is similar to or close to any previously found solution.

The MGA objective function defined in step 4 is given by:

$$\text{Minimize } p = \sum_{k \in K} x_k \tag{2.25}$$

$$\text{Subject to } f_0(\mathbf{x}) - T \leq 0 \quad \forall\, \mathbf{x} \in W \tag{2.26}$$

$$T = f_0(\mathbf{x}^0) \cdot (1+\epsilon) \tag{2.27}$$

In this formulation, k represents the variable indices for the variables with nonzero values in the previous solution, $f_0(\mathbf{x})$ is the evaluation of the objective function and T is the target value specified. In the formulation of the constraint $\mathbf{x} \in W$ specifies that all previously defined constraints still apply as all-new solutions $\mathbf{x}$ must be a part of the set of feasible solution vectors from the original formulation W.

How the new objective function precisely is formulated and which variables to include is discussed in [5], where two alternative approaches to defining the new objective function are presented. One approach suggests giving all nonzero variables from the last iteration a weight of 1 in the new objective function. This approach does not consider weight from previous iterations. However, the second approach suggests adding on to the coefficient with a factor of +1 for every time one variable has appeared with nonzero in a row, thereby further increasing the insentience to reduce the use of that specific technology.

The HSJ approach introduces the MGA constraint and then seeks to find alternative solutions within the near-optimal feasible space by changing the objective function. The new objective function in the HSJ approach becomes to minimize the capacity of non-zero technologies from the previous solutions. This means that the HSJ method only will produce new alternative solutions as long as there is at least one technology with a capacity of zero in the previous solution. In practice, this leads to a maximum number of alternative solutions equal to the number of technologies with a capacity of zero in the optimal solution. The HSJ method is therefore not ideal if information about the entire decision space is desired, and therefore other alternative approaches should be explored.

2.5 Other MGA approaches

The primary step in the HSJ method that requires improvement is the selection of objective function for the iterative solutions to the problem. Analyzing the problem of selecting objective function it becomes clear that, one essentially needs to choose a weighting of the decision variables included in the objective function. This weighing could be represented with the vector $\mathbf{n}$ with the same length as the vector $\mathbf{x}$ containing the decision variables. The objective function then becomes:

$$\text{Minimize } p = \sum_i \mathbf{n}_i \mathbf{x}_i \tag{2.28}$$

The vector $\mathbf{n}$ can essentially be thought of as a search direction in the decision space, forcing the optimization to find solutions located on the edge of the near-optimal decision space in that direction.

If infinitely many MGA iterations were performed with a new direction vector $\mathbf{n}$ every time, all solutions on the border of the decision space would be found. Therefore, a natural alternative MGA approach would be to select a random vector $\mathbf{n}$ for every iteration with the individual components of $\mathbf{n}$ being between 0 and 1. The only thing needed to make this a full MGA algorithm is a termination criterion. Here one could choose to set a fixed number of iterations to perform.

An alternative approach for selecting the MGA objective function is presented in [16], where an objective function seeking to maximize the difference between alternative solutions is proposed. Using the Euclidean distance between two points $\mathbf{x}$ as a measure of the difference between two points the objective function for the first MGA iteration becomes:

$$p = \sum_i |\mathbf{x}_i - \mathbf{x}_i^0| \tag{2.29}$$

Where $\mathbf{x}_i^0$ represents the optimal solution. For all following MGA iterations the MGA objective should then be changed to:

$$p = \sum_{k=0}^{A-1} \sum_{i} |\mathbf{x}_i - \mathbf{x}_i^k| \tag{2.30}$$

Hereby seeking to maximize the distance from the current solution to all previously found solutions. Using this set of objective functions provides a more structured method of finding alternative solutions in the "corners" of the near-optimal decision space. Using Euclidean distance as a measure of the difference between solutions does however come with the cost of the objective function becoming quadratic. This vastly complicates the optimization problem, compared to the regular linear objective function, increasing computation time by a large factor.

In the work presented in [10], an additional MGA approach is presented. Instead of seeking to find maximally different solutions, the goal is to determine the feasible range of the variables making the near-optimal feasible space. The decision variables can either be all variables included in the model, or variables representing groups of the true variables, as it was suggested in Section 2.2. Using the objective function formulation from Equation 2.28, then the search directions used in [10] would correspond to setting a single value in $\mathbf{n}$ as either 1 or -1 and all other values as 0. The search directions would then become:

$$\begin{aligned}
\mathbf{n}^1 &= \{1, 0, 0, ..., 0\} \\
\mathbf{n}^2 &= \{-1, 0, 0, ..., 0\} \\
\mathbf{n}^3 &= \{0, 1, 0, ..., 0\} \\
&\vdots \\
\mathbf{n}^d &= \{0, 0, 0, ..., -1\}
\end{aligned}$$

Using this MGA method would result in $d \cdot 2$ number of alternative solutions with d being the number of decision variables included. Understanding such a large number of alternative solutions is infeasible as the complexity of the individual solutions is too great. This MGA method does, however, provide a bounding box for technology capacities, as maximum and minimum values for technology capacities are found.

A completely different approach for addressing MGA is to use genetic algorithms as proposed in [16]. Here a population-based niching algorithm is proposed to generate a finite set of alternative solutions that are maximally different. Where in the algorithm presented in Equation 2.29 and 2.30, the MGA solutions are generated iteratively, the genetic niching algorithm proposes to create a population of solutions initiated at random locations in the decision space:

$$P = \{\mathbf{x}^1, \mathbf{x}^2, .., \mathbf{x}^n\} \tag{2.31}$$

The objective then becomes to maximize the total distance between all individuals in the population.

$$\text{Maximize } p = \sum_{k=1}^{n} \sum_{j=1}^{n} |\mathbf{x}^j - \mathbf{x}^k| \tag{2.32}$$

This reduces the number of iterations needed to generate the set of alternative solutions to just one. Creating an initial population further allows the modeler to select the exact number of alternative solutions desired. Although genetic algorithms might be computationally efficient, they still only produce a limited insight into the general characteristics of the near-optimal feasible space.

2.6 Novel MGA approach

The novel approach towards MGA optimization of energy networks, developed in this project, will be presented in this section. Building on the concepts presented in Section 2.4 and 2.5 this method seeks to explore not only a few alternative solutions from the decision space, but instead seeks to define the entire near-optimal feasible space, and hereby provide useful statistical data through a strategic sampling of this space.

The method developed can be divided into two phases. In the first phase, the shape of the feasible near-optimal decision space is found, and in the second phase, relevant data is extracted from the found space through sampling.

Feasible space mapping

As explained in Section 2.2, the near-optimal decision space is convex and can either be closed or not before the MGA constraint is applied. The space will however always be closed when the MGA constraint from Equation 2.26 is introduced.

Knowing that all constraints used, including the MGA constraint, are linear the convex set defining the near-optimal feasible space must be a polyhedron and therefore it is possible to define the shape of this set with a finite number of vertexes. The goal of the first phase of this MGA approach is to find enough of these vertices to approximate the shape of the near-optimal feasible space.

Because the optimization problem is closed, any choice of objective function will provide a solution lying within the feasible region. Using this, it is possible to search in any desired direction in the decision space by altering the objective function of the numerical optimization problem. Using a unit vector pointing in the desired direction $\mathbf{n}$, multiplied with the variables to be optimized $\mathbf{x}$, as objective function (Equation 2.28), provides full control over direction of search.

To use this tool to map the feasible region, a framework is needed to select search directions $\mathbf{n}$ that will lead to the discovery of all vertices defining the feasible region. There are several ways to solve this problem, and the simplest choice would be to search in random directions until no new solutions are found. This is however not very efficient, and as dimensionality and model evaluation time increases, this method becomes unfeasible.

Instead, the method proposed here suggests an approach that reduces the number of model evaluations by a clever choice of search direction. Initially, the optimization problem is solved using the original objective function to define the MGA constraint (Equation 2.26). This provides a single point $\mathbf{x}^0$ located within the feasible region. By selecting search directions that seek to maximize and minimize every single variable in $\mathbf{x}$ one by one, as the MGA method presented in [10], an additional $2 \cdot d$ solutions are found. Knowing that the feasible region is defined by a polyhedron, it makes sense to imitate this shape by computing the hull containing all points found so far.

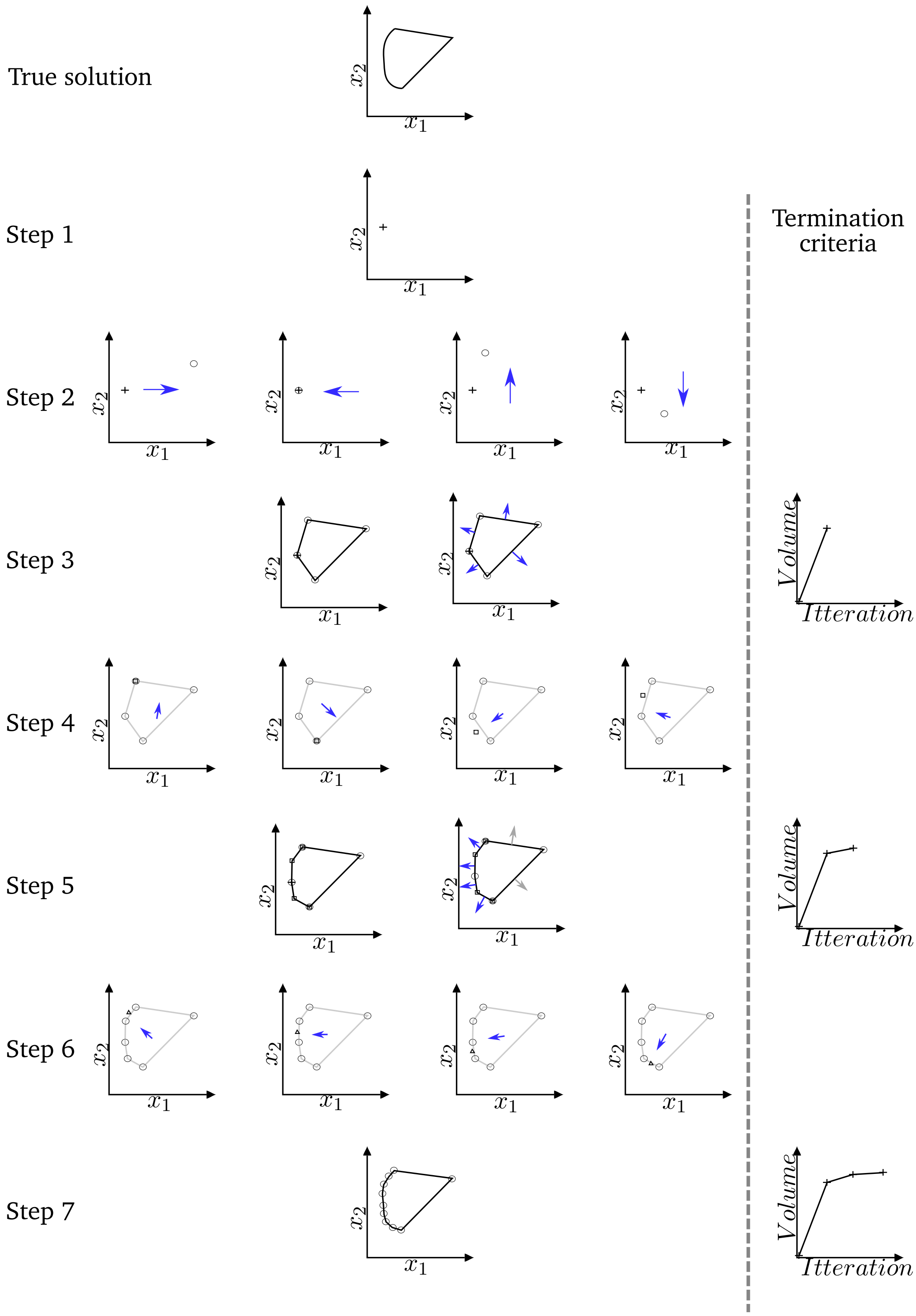


Figure 2.2: The figure shows a step by step schematic of the working principles of the proposed novel MGA method capable of determining the convex hull of the near optimal feasible space. In Step 1 the optimal point is found. Step 2 maximizes and minimizes all variables and in Step 3 the convex hull containing all found points is found. In Step 4 the MGA algorithm searches in all face normal directions. Step 5 and 6 repeats the process of computing the hull and searching in the face normal directions.

Using the face normal vectors of this hull to define the next set of search directions ensures that if one of the faces in the hull is not part of the polyhedron defining the feasible region, then a new point will be found when searching in the normal direction of that particular face. Using the newly found points combined with all previously found points, to repeat the process of defining a hull and searching in the face normal directions, will as long as the hull computed isn't the full solution to the problem, continue finding new points within the feasible region, until all points defining the feasible region are found. In other words, this method ensures that the solution converges towards the full solution if enough iterations are performed.

If the feasible region was to have a very complex shape being defined by a high number of vertexes, or if a non-linear constraint was introduced, and thereby preventing that the entire region from being represented by a finite number of vertices, it would be necessary to have a termination criteria that does not require that the complete solution is found. The volume of the hull estimating the feasible region will converge towards the size of the feasible region and implementing convergence criteria on the hull volume provides a good termination criterion. The entire process of searching the feasible region is listed in bullet form and visualized in Figure 2.2.

1. Find the initial solution and add MGA constraint
2. Maximize and minimize all variables
3. Based on these points define a convex hull
4. Compute all face normals of the hull
5. Iterate over each face normal, discard any previously searched direction, and change the objective function to search in that direction 2.28
6. Add the newly found points to list of points and define a new hull
7. Check if the hull volume satisfies the convergence tolerance. If yes then termite, else return to step 4

The result of following these steps is a set of points, located on the border of the near-optimal feasible decision space, defining the hull containing the estimated near-optimal feasible decision space.

The method presented here can, in theory, be used to approximate any n-dimensional near-optimal decision space. But as the dimension of the near-optimal feasible space increases, the complexity of the hull containing all near-optimal solutions, and thereby the number of facets on the hull also increases. This means that more search directions are needed and thereby computing time increases. The algorithm used for computing the convex hull containing all near-optimal points increases in computing time by a power of three, $O^*(n^3)$. The allowable dimension size of the decision space is therefore determined by the available computing time.

Sampling of decision space

At this point, this novel MGA approach doesn't differ much from the MGA solutions previously proposed in literature, as the result of the first phase is a set of alternative solutions located on the perimeter of the near-optimal feasible space. Analyzing the solutions found in the first phase would provide a wrong picture of the complete set of solutions contained in the near-optimal feasible space as they represent only the extreme points of the decision space. Instead, a method of extracting information about all solutions located within the hull spanned by the solution from the first phase should be defined.

To extract meaningful statistical data about the model from the found feasible region, a method for sampling the feasible space is needed. Assuming that all solutions to this particular problem are evenly distributed across the entire feasible region, sampling randomly evenly across the region will provide a good dataset representing the true data if enough points are sampled. The assumption of even distribution of solutions does however only hold true for the full-dimensional original decision space W. This will be further discussed in Section 2.7.

Drawing random samples evenly distributed inside the feasible space is however no trivial task. More generally speaking the task is to draw samples evenly inside a polyhedron, which can be further reduced to simply drawing samples evenly spaced inside a simplex, as the polyhedron can be split into a range of simplexes as seen on Figure 2.3a and 2.3b. Then by drawing a number of samples equivalent to the volume fraction of the simplex multiplied with the number of total sample points desired, from each simplex, it is possible to sample the entire decision space evenly.

The simplex is given by a list containing all its vertices $P = \{\mathbf{p}^1, \mathbf{p}^2, ..., \mathbf{p}^m\}$. The number of vertices needed to describe a simplex will always be $m = d + 1$, where d describes the dimension of the space the simplex is located within. A new point inside this simplex can be found by summing the points describing the simplex scaled with a vector s if this scaling vector has the property of summing to one $\sum_i \mathbf{s}_i = 1$.

$$\mathbf{p}_{\text{new}} = \sum_{i=1}^{m} \mathbf{p}^i s_i \tag{2.33}$$

Where m is the number of points used to describe the simplex. The challenge is then to select s, such that the space inside the simplex is sampled evenly. There are several ways of doing this. The method chosen in this project is called the Bayesian Bootstrap and is further explained in [17] together with the proof that this method will generate evenly distributed points.

Using the Bayesian Bootstrap method, an initial $m - 1$ random numbers are drawn, from an even distribution with a range from 0 to 1.

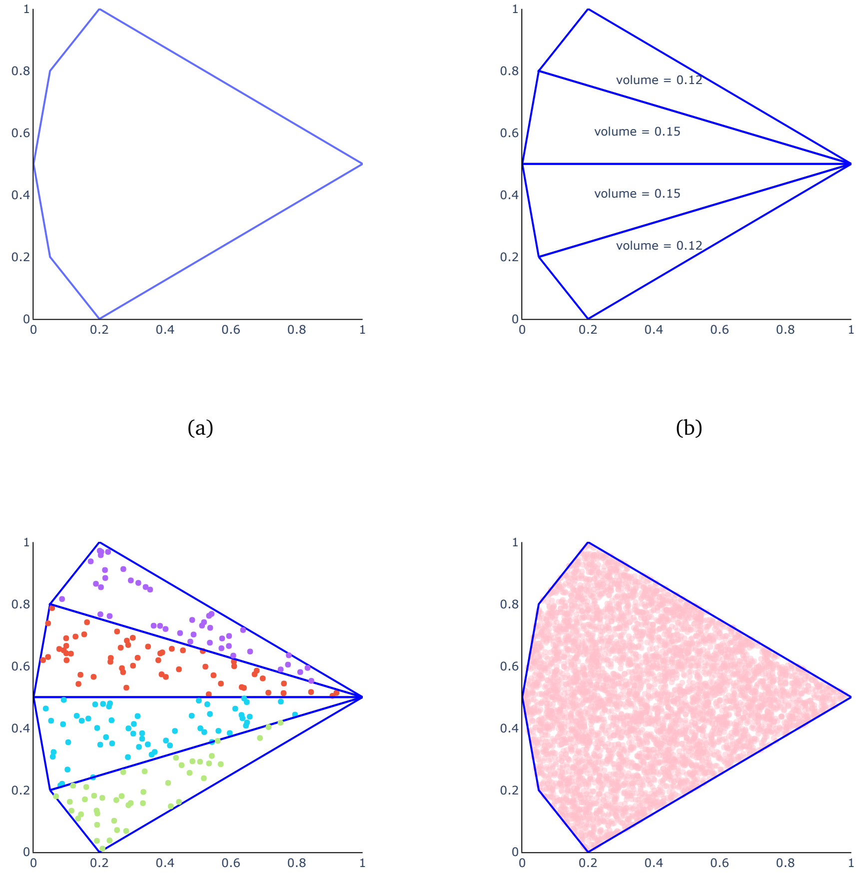
$$\mathbf{r} = \{r | r \in \mathbb{R}, 0 \leq r \leq 1\} \tag{2.34}$$

Then sorting the components of the vector $\mathbf{r}$ by increasing value, and adding 0 as the first entry and 1 as the last, this new $\mathbf{r}$ vector can be used to define a scaling vector. The length of this new $\mathbf{r}$ vector is now $m + 1$ as 0 and 1 has been added. Using the difference between the components in $\mathbf{r}$ to define a new vector:

$$\mathbf{s} = \{r_{i+1} - r_i\} \; \forall \; i = 1, 2, .., n \tag{2.35}$$

The vector s has the property that the sum of the components will always be equal to 1, by definition. Using the s vector to scale the points in P it is possible to draw point randomly located within the simplex.

Following this procedure for all simplices provides an even sampling of the convex hull as it can be seen in Figure 2.3c and 2.3d.



(c) (d)

Figure 2.3: An illustration of the sampling method. Figure (a) shows the initial convex hull, figure (b) presents the hull separated into simplexes and their individual area. Figure (c) shows how 200 sample points are distributed between the individual simplexes, and figure (d) shows the final sampling with 5000 sample points evenly distributed over the convex hull.

Having sampled evenly across the entire convex hull approximating the near-optimal feasible space, a dataset containing all sample points is now available. These points can now be used to provide information about the distribution of variables and their correlations within this near-optimal feasible space. From the dataset, it is possible to find common features shared amongst all near-optimal solutions such as minimum capacity requirements of the individual technologies, etc.

2.7 Multiplicity

When grouping variables to reduce dimensionality as described in Section 2.2, a lot of information is lost. Especially one very feature of the distribution of solutions in the near-optimal feasible space is lost. In the fully dimensional decision space, where every variable/dimension correspondents to the capacity of one specific technology in a specific country, all solutions are evenly distributed. Essentially a point in this space describes a specific configuration of the European energy grid. When the variables are grouped by summation and a new decision space of lower dimension is formed, this property is lost, as a point in the reduced decision, space covers over a range of possible configurations of the energy grid. Instead of describing a specific configuration of the energy grid, a point in the reduced decision space describes a set of equality constraints that the energy grid must satisfy.

If the decision space of the original problem W has d dimensions, corresponding to the d variables in $\mathbf{x}$. When no equality constraints are introduced, the solution space X is also of d dimensions. For every equality constraint introduced the dimensionality of the solution space is reduced by one dimension. If o equality constraints are introduced the dimension of the solution space becomes $d - o$, meaning that the solution space is a $d - o$ dimensional subspace of the decision space.

In the solution space of the original problem, every element/point is unique and corresponds to a unique configuration of the European energy grid. When a new decision space is formed Y, by creating new variables $\mathbf{y}(\mathbf{x})$ consisting of sums of decision variables, this feature is lost. Instead of representing a single configuration of the European energy grid, a point $\mathbf{y}$ in the reduced decision space, represents a range of configurations. A point $\mathbf{y}$ in Y, covers over a subset of the solution space in X of dimension $n - o - k$, where k is the dimension of Y.

An example would be a three dimensional problem with the near-optimal feasible space X, variables $\mathbf{x}$ and constraints $\mathbf{f}_i(\mathbf{x})$. There are only inequality constraints, thus the dimension of the solution space X is 3.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{2.36}$$

$$X = \{\mathbf{x} | \mathbf{f}_i(\mathbf{x}) \leq 0\} \tag{2.37}$$

Where, a reduced decision space of two dimensions Y with the variables $\mathbf{y}$ is introduced to simplify the problem.

$$\mathbf{y} = \begin{bmatrix} y_1 = x_1 + x_2 \\ y_2 = x_3 \end{bmatrix} \tag{2.38}$$

$$Y = \{\mathbf{y}(\mathbf{x}) | \mathbf{f}_i(\mathbf{x}) \leq 0\} \tag{2.39}$$

In this case, where the full dimension solution space X is three dimensional and the reduced decision space Y is of two dimensions, a solution y in the reduced decision space, corresponds to a one-dimensional subspace of the original solution space. In other words a solution in Y represent all points in X that satisfies the original constraints $f(x) \leq 0$ and the two equality constraints $x_1 + x_2 = y_1$ and $x_3 = y_2$. There might only be a single point in X that satisfies these constraints, or there might be infinitely many, but as all solutions in Y must satisfy the same constraints as the solutions in X, there will be at least one solution.

In broader terms, a point in a reduced decision space Y of dimension k represents a set of equality constraints, increasing the total number of equality constraints in the original problem. Therefore a point $\mathbf{y}$ in Y represents a subspace of the solution space X of dimension $n - o - k$.

Thus by using the novel MGA approach presented in Section 2.6 on the reduced problem to determine the shape of Y, and then sampling Y evenly, would generate misleading results, as some of the sampled points represents a single or a few configurations of the energy network, and some sample points represent a vast amount of configurations of the energy network. In order to achieve correct results, the points sampled in Y, should be assigned a weight representing the size of the subspace of X that the given solution in Y spans. In the case of a three-dimensional solution space and a two-dimensional reduced decision space, the weight assigned to each sample point should be the length of the line in X given by the constraints that specific point in Y represents.

A method for determining the size of the subspace of X that a point in Y spans must be established.

The subspace of X given by a point in Y can be described through a parametric representation. Following the example from above, where the two equality constraints $x_1 + x_2 = y_1$ and $x_3 = y_2$ is to be satisfied, the parametric representation of the 1-facet in the three-dimensional space X can be found by selecting two points that satisfy the constraints. These two points could be selected as:

$$\mathbf{x}^1 = \begin{Bmatrix} y_1 \\ 0 \\ y_2 \end{Bmatrix} \ , \ \mathbf{x}^2 = \begin{Bmatrix} 0 \\ y_1 \\ y_2 \end{Bmatrix}$$

Using these two points the parametric representation of the 1-facet can be written as:

$$\mathbf{x} = \mathbf{x}^1 + t\left(\mathbf{x}^1 - \mathbf{x}^2\right) = \begin{Bmatrix} y_1 \\ 0 \\ y_2 \end{Bmatrix} + t \begin{Bmatrix} y_1 \\ -y_1 \\ 0 \end{Bmatrix} \tag{2.40}$$

The size of the 1-facet could now be found by integrating over t. But as the boundaries of t are unknown this is not an option. Instead, we can use MGA to find the span of the 1-facet combined with the knowledge of the direction of the 1-facet given by the vector $\mathbf{x}^1 - \mathbf{x}^2$.

Performing a MGA study where the two equality constraints $x_1 + x_2 = y_1$ and $x_3 = y_2$, are included and the search directions is chosen as the positive direction of the 1-facet $\mathbf{n} = \mathbf{x}^1 - \mathbf{x}^2$ and the negative direction $\mathbf{n} = -(\mathbf{x}^1 - \mathbf{x}^2)$. This provides two points that defines the ends of the 1-facet, allowing the length to be calculated as the difference between these two points projected on to the direction vector $\mathbf{x}^1 - \mathbf{x}^2$.

This method can be expanded to any dimension of solution space and subspace. The parametric representation of the subspace of X then becomes:

$$\mathbf{x} = \mathbf{x}_1 + \sum_{i=1}^{(n-k)} t_i(\mathbf{x}_1 - \mathbf{x}_{i+1}) \tag{2.41}$$

The points $\mathbf{x}_i$ should be selected such that no more than two points are co-linear. Using the vectors $(\mathbf{x}_1 - \mathbf{x}_{i+1})$ $for\ i = 1..(n-k)$ as search directions in an MGA study, and computing the size of the found facet, it is possible to determine the multiplicity factor.

This method of determining multiplicity was implemented on a small scale experiment, showing the profound effects of neglecting multiplicity. It was however not implemented on the full-scale experiments performed in this project, due to its computational cost and complexity.

Multiplicity experiment

To understand the effect of multiplicity, an experiment has been performed using a three-dimensional problem with near-optimal feasible space X and the two-dimensional reduced decision space Y. The techno-economic model used is a simplified version of the full-scale model used in this project, including only two nodes (Denmark and Sweden) connected by a single transmission line. Sweden has two technologies available, wind turbines and gas turbines (OCGT). Denmark is only allowed to install wind power. A CO_2 constraint is enforced requiring a 60% reduction in emissions compared to an unrestricted scenario. The decision variables in $\mathbf{x}$ are:

$$\mathbf{x} = \begin{Bmatrix} x_1 : \text{DK wind} \\ x_2 : \text{SE wind} \\ x_3 : \text{SE OCGT} \end{Bmatrix} \tag{2.42}$$

The variables of $\mathbf{y}$, when grouped by technology type then becomes:

$$\mathbf{y} = \begin{Bmatrix} y_1 : \text{DK wind} + \text{SE wind} \\ y_2 : \text{SE OCGT} \end{Bmatrix} \tag{2.43}$$

Finding the shape of Y using the MGA approach presented in Section 2.6 and sampling the space uniformly, yields the results presented in Figure 2.4. The shape of the reduced decision space Y presented on Figure 2.4a, is quite simple and indeed linear. The distributions of the sample points on Figure 2.4a mimics the shape of Y, by being almost triangular.

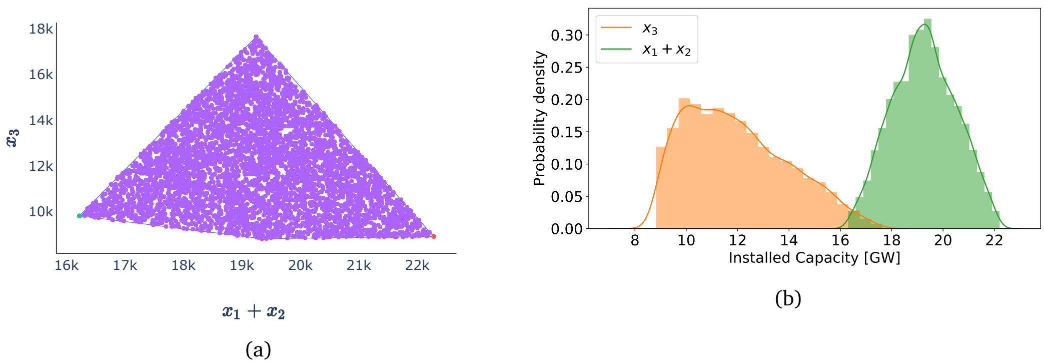
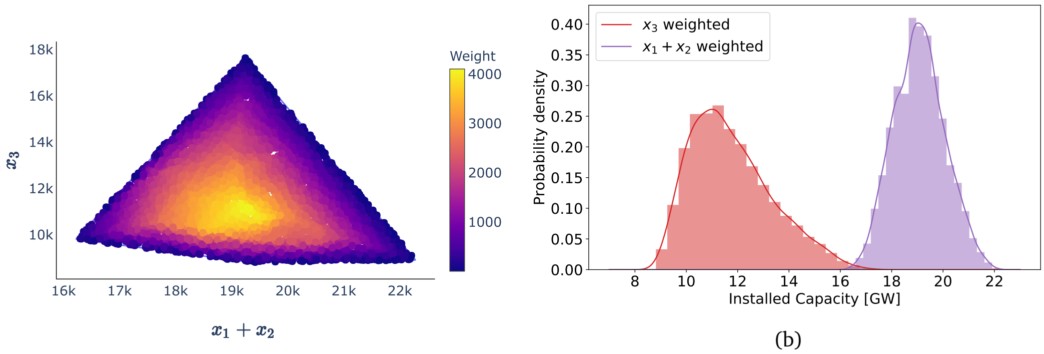


Figure 2.4: The figure shows the results of a MGA study with grouped variables. Figure (a) shows the two dimensional subspace Y and (b) presents a histogram of the sampled points from (a).

To generate a more realistic result, and to study the effect of multiplicity, the multiplicity for all sample points is evaluated. For every sample point, an MGA study of the full three-dimensional problem is performed with the two equality constraints from Equation 2.44 introduced.

$$
\begin{aligned}
x_1 + x_2 &= y_1 \\
x_3 &= y_2
\end{aligned}
\tag{2.44}
$$

The search directions for the objective function is chosen as $\mathbf{n}_1 = \{1, -1, 0\}^T$ and $\mathbf{n}_2 = -\{1, -1, 0\}^T$. These two directions are chosen, by following the method proposed in Section 2.7, where a parametric representation of the subspace of X given by a point in Y is found, and the direction vector used as search direction. The weight of the individual points is then given by the area of the subspace of X given by that specific point in Y. In this case, it is equivalent to the euclidean distance between the two points found by searching in the $\mathbf{n}_1$ and $\mathbf{n}_2$ direction. The weighted points are shown in Figure 2.5a, and a histogram using the weighted points are shown in Figure 2.5b.



(a)

(b)

Figure 2.5: MGA results of a study using grouped variables and a weighting of the sample points based on their multiplicity. Figure (a) shows the weighted sample points, and figure (b) shows a histogram of weighted sample points.

It is seen how the points closer to the center of the reduced decision space in Figure 2.5a, is assigned a much higher weight than the points on the perimeter of the reduced decision space. This leads to the distributions being narrower and taller as seen in Figure 2.5b. The distributions have now become more bell-curved, and the lower deviation suggests that scenarios located in the center of the reduced decision space are much more likely to occur than scenarios on the perimeter.

To validate the results of the weighting method, an MGA study on the full problem is also performed. This study explorers the entire three-dimensional decision space and samples it evenly. The results of this study is presented on Figure 2.6a and 2.6b, showing the three dimensional space X on Figure 2.6a and the distributions of the sampled points on Figure 2.6b.

The distributions of the sampled points from the approach without calculating the weights, and the one weighting the sample points are compared to the full solution on Figure 2.6. In Figure 2.6d it is seen that the result using weighted points matches almost perfectly with the true solution. In practice, this means that, without the weighting, solutions too close to the frontier of the near-optimal solution space seem to be "similarly probable" as the solutions in the inner part of the near-optimal space when this is not true in reality. Consequently, when including the weighting, the probability density graphs squeeze, reducing the probability for extreme values.

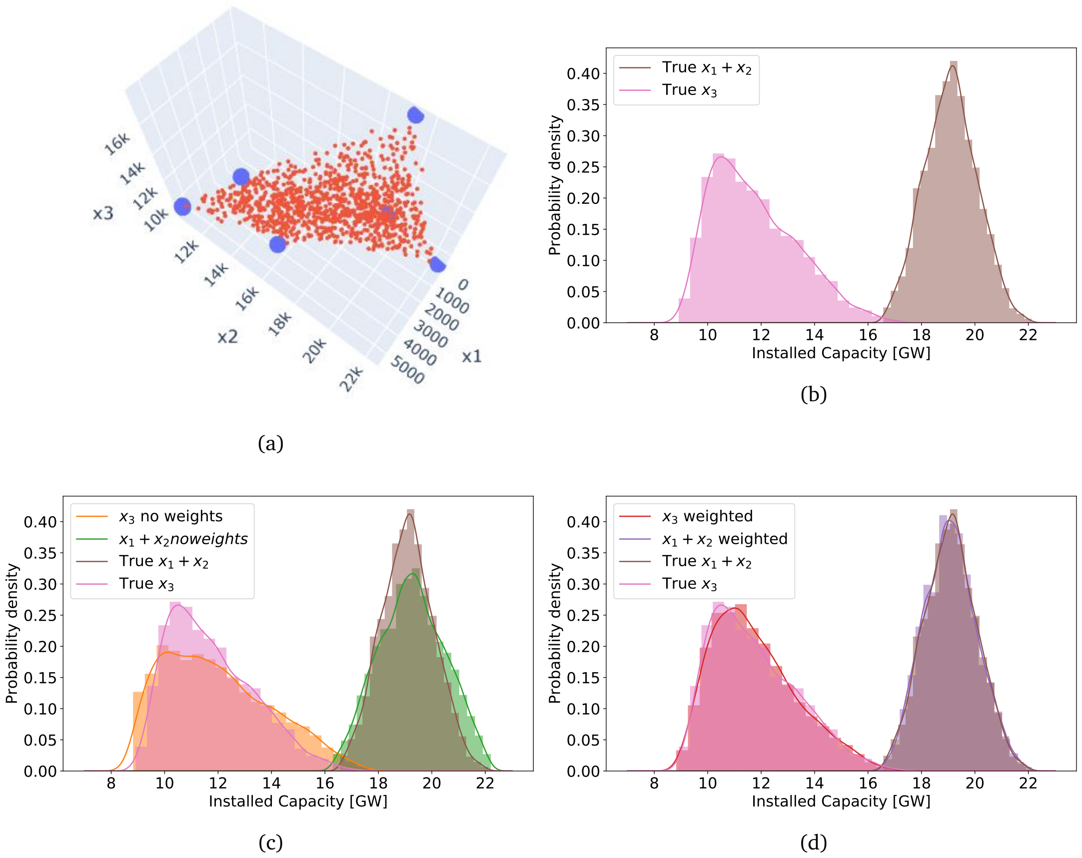


Figure 2.6: Figure (a) and (b) show the result of an MGA study using three decision variables. In Figure (a) the three-dimensional near-optimal feasible space is presented and a histogram of the sampled point is seen on (b). Figure (c) and (d) compares the results found by performing MGA on the full three-dimensional decision space with the results of two MGA studies using grouped variables. In Figure (c), results from the three-dimensional study are compared to the MGA study of grouped variables, not weighting the sample points. On figure (d) the results from the three-dimensional study are compared to the MGA study of grouped variables weighting the sample points.

This insight is very important, as previously conducted MGA studies from literature [5], [6], [7], [8], [9] and [10], have had a tendency to focus on solutions located on the very perimeter of the decision space. When analyzing only the extreme points, it is important to remember that these solutions are unique, and no other configurations of the model can generate the same results. Wheres, when the entire decision space is sampled, it is possible to consider where the flexibility of the model lies.

Multiplicity in higher dimensions

It is suspected that multiplicity become a larger factor as the dimensionality of the full problem increases. Therefore the multiplicity experiment have been repeated, with two new models with respectively 4 and 6 variables. The decision space of reduced dimension have been kept in two dimensions. The variables x in the two new experiments then becomes:

$$\mathbf{x} = \left\{ \begin{array}{ll} x_1: & \text{DK wind} \\ x_2: & \text{DK OCGT} \\ x_3: & \text{SE wind} \\ x_4: & \text{SE OCGT} \end{array} \right\} \text{ and } \mathbf{x} = \left\{ \begin{array}{ll} x_1: & \text{DK wind} \\ x_2: & \text{DK OCGT} \\ x_3: & \text{SE wind} \\ x_4: & \text{SE OCGT} \\ x_5: & \text{NO wind} \\ x_6: & \text{NO OCGT} \end{array} \right\}$$

Grouping the variables by generator type, the variables of the reduced decision spaces become:

$$\mathbf{y} = \left\{ \begin{array}{l} y_1 : x_1 + x_3 \\ y_2 : x_2 + x_4 \end{array} \right\} \text{ and } \mathbf{y} = \left\{ \begin{array}{l} y_1 : x_1 + x_3 + x_5 \\ y_2 : x_2 + x_4 + x_6 \end{array} \right\}$$

Following the same method as used in the previous multiplicity study, where the MGA method is first used on the reduced decision space, without calculating the weights of the points, and then weighting all the sample points. Both results are compared to the true solution, found by performing an MGA study on the full-dimensional problem.

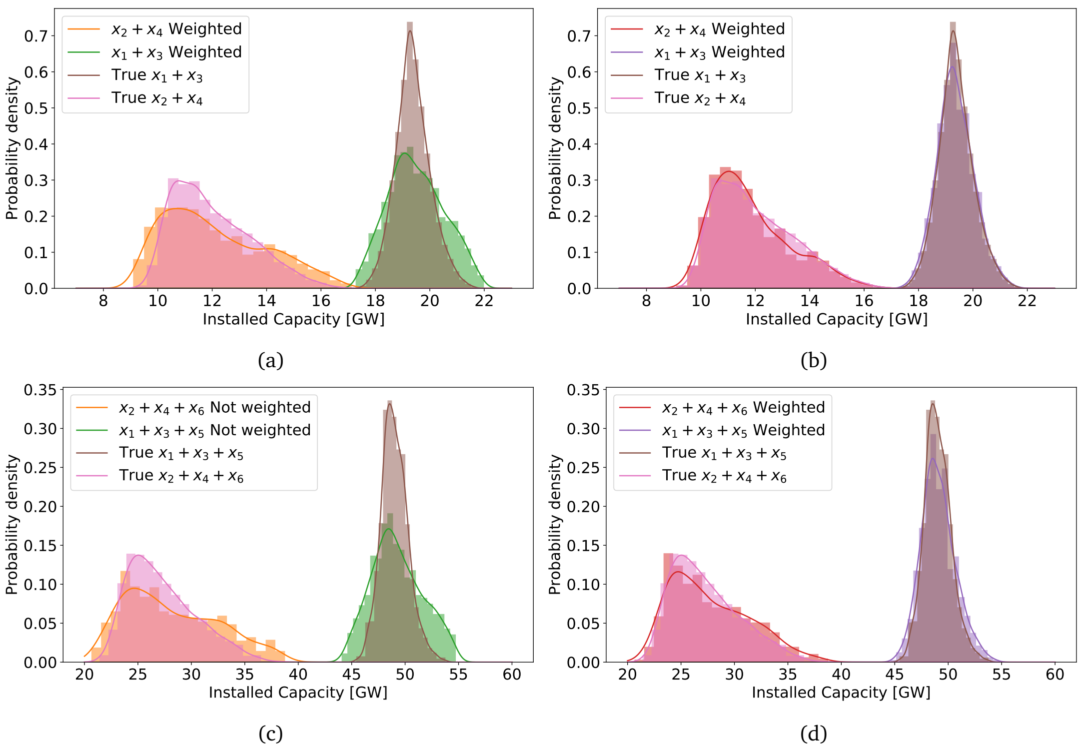


Figure 2.7: Figure (a) and (b) presents the results from a four dimensional study with a reduced decision space of two dimensions. Figure (c) and (d) shows the results a six dimensional study with a reduced decision space of two dimensions

In Figure 2.7, the results from the four and six-dimensional problems are presented. Comparing the not weighted, solution from the four-dimensional problem seen on Figure 2.7a, to the not weighted solution for the three-dimensional study on Figure 2.6, it is seen that a larger mismatch between the true solution and the estimated one, occurs as the dimension increases. Extending this analysis to the six-dimensional problem seen in Figure 2.7c, it is seen that this trend continues. However, comparing the weighted approaches in all three studies seen in Figure 2.6b and Figure 2.7b and 2.7d, a much better fit between the distributions is found. These results suggest that the proposed method of estimating multiplicity is working correctly.

Estimating the effect of multiplicity in this manner is however not feasible for larger problems. Every estimation of a sample point weight essentially requires an MGA study on its own, and as more than 1000 sample points are needed to generate well-representing results, the computational effort needed becomes too high. Instead of using the approach for estimating multiplicity presented in this section, as a tool to use in larger simulations, this study should serve as a proof and reminder, that the true distribution of solutions is much narrower than what is found when using MGA algorithms, on a reduced dimensional decision space.

2.8 Gini coefficient

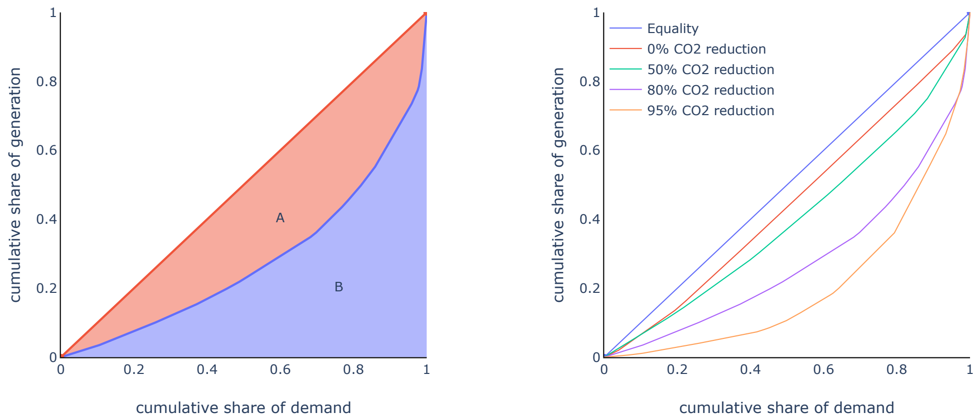
The goal of this project is to develop a method to deal with structural uncertainty in techno-economic energy models, arising from unmodeled objectives. One very important factor that isn't included in the objective function is the equality in average yearly production versus consumption. In a model including multiple countries, like the one used in this project, this could be very important as countries in general desire to be somewhat self-sufficient with energy.

To create a measure for the equality of production versus consumption, the Gini coefficient can be used. Calculating the cumulative share of demand per country and plotting it against the cumulative share of generation per country one gets the Lorentz curve for that specific scenario, as shown for four scenarios in Figure 2.8b. As inequality increases the Lorentz curve lies further and further away from the equality line.

The Gini coefficient is calculated as the relationship between the area between the Lorentz curve and the equality line (Area A on Figure 2.8a) relative to the total area under the equality line (Area A+B on Figure 2.8a). Thus the Gini coefficient becomes $G = \frac{A}{A+B}$.

A scenario where every country over the duration of an entire year, produces as much energy as it consumes, would have a Gini coefficient of 0, and represent the equality line on Figure 2.8. A scenario where one country is producing all energy and consuming none, would, on the other hand, have a Gini coefficient of 1, and represent total inequality.

The Gini coefficient calculated in this manner is a very good measure of the distribution of energy production versus consumption. As a very important unmodeled objective in a techno-economic energy model including multiple countries, is the desire for the countries to be self-sufficient, this Gini coefficient provides a very important insight.



(a) (b)

Figure 2.8: On figure (a), a schematic of the areas considered when calculating the Gini coefficient is presented. Figure (b), shows the Lorentz curves of four studies using different CO2 constraints.

3 Model

The purpose of this section is to present the techno-economic model used in this project. In the previous chapter, the mathematical framework of a techno-economic model was presented, and in this chapter, the physical interpretations of the equations making the model will be discussed. Furthermore, all the time-series and cost data used in the model will be presented. The model used in this project is heavily inspired by the work performed in [11].

3.1 Topology

The model used in this project is based on the work presented in [11], where a model spanning the electricity grid of 30 European countries is formulated as a techno-economic linear optimization problem. Countries included in the model are the EU-28 countries not including Cyprus and Malta, instead of including Norway, Switzerland, Serbia, and Bosnia and Herzegovina.

The topology of the network presented in Figure 3.1, is such that each node represents a country and the links represent international HVDC or HVAC links. The links included are based on currently installed international transmission lines.

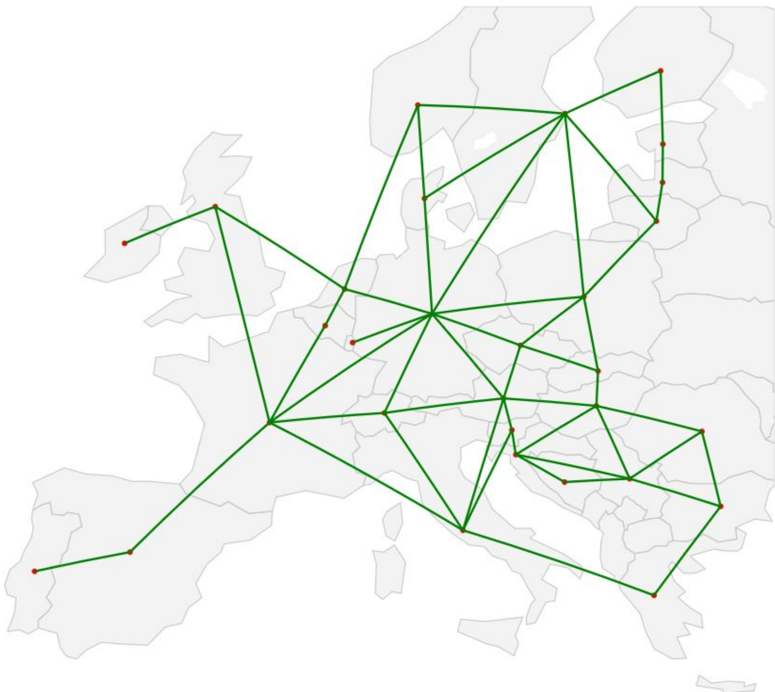


Figure 3.1: This figure shows the topology of the techno-economic model used in this project. Green lines represent transmission lines and red dots present the location of centroid of all countries included in the model.

All model input parameters are based on 2011 values as this is the earliest year with all data available. The temporal resolution of the model is hourly, with all simulations spanning a full year. Technology costs are all valued in 2011 Euros.

3.2 Energy production

Each node in the network, has energy-producing technologies available, with initial capacities being zero. The available energy-producing technologies used in this project: Onshore wind, offshore wind, solar PV and open-cycle gas turbines (OCGT). In the model, all technology capacities are expandable limited only be the geographical potential.

The geographical potentials used are calculated following the work of [11]. In the calculation of geographical potential, the potential available area suited for either onshore wind, offshore wind, and solar PV, must first be defined. These areas were found by allowing certain technologies to be installed only in areas with certain land-use types. Hereby restricting onshore wind farms from being installed in cities and solar PV plants from being installed in forests etc. The placement of offshore wind farms was restricted to areas with a water depth of less than 50m. Furthermore, all nature reserves were excluded from potential areas. As competing land use and likely public acceptance issues will occur, the found potential areas are set to bee only 20% of the found area for onshore and offshore wind and only 1% for solar PV. Assuming a maximum nominal installation density of 10 MW/km^2 for offshore and onshore wind power, and 145 MW/km^2 for solar PV [11], it is possible to calculate the geographical potential for the three technologies all across Europe.

The hourly energy production of all variable renewable energy sources is limited by the production potential given by the weather. Following [11], the availability was calculated using historical weather data for 2011 from [18] with a spatial resolution of 40x40 km and hourly temporal resolution. The weather data is first converted to generation potentials for each 40x40 km cell using the REatlas software [19], and then the national hourly means are found. The mean capacity factor for wind and solar power is presented in Figure 3.2.

The dispatchable energy sources available in all countries are chosen to be open-cycle gas turbines (OCGT), as they have high flexibility and good load following capabilities, therefore making them suitable as a backup generator in a highly decarbonized scenario. They do however not necessarily produce realistic results when used in scenarios with low decarbonization, as other plant types such as conventional coal-powered combined heat and power perform better in these cases. The capacities and energy generation of the gas turbines are contrary to the variable renewable energy sources, not limited by geographical or generation potentials. They are, however, limited by the maximum allowable CO_2 emission. The CO_2 emission intensity of the open cycle gas turbine is 0.19 t/MW [11]. In this project, the CO_2 constraint will always be calculated as a percentage reduction in emission compared to a scenario run on the same model with the same parameters without any constraint on CO_2 emission.

In countries located on the coast both onshore and offshore wind turbines are available. The capacity of these two types of wind power is however treated as one single variable in all simulations performed in this project.

In all simulations, the capacities of all energy generators are initially set to be zero, with the capability to be expanded until geographical potentials or CO_2 constraints,

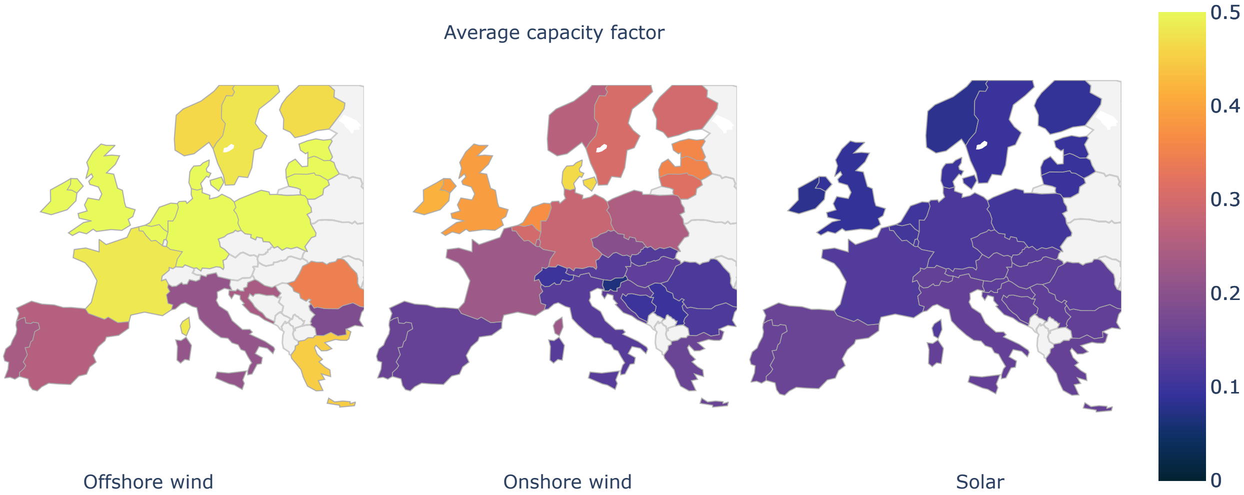


Figure 3.2: The figure shows the average generation potential of wind turbines and solar PV plants, in the individual countries included in the model.

limits further expansion. The cost of expanding capacities is calculated as annualized cost, given as the annualized investment cost plus fixed annual operations and maintenance cost. The annualized investment cost is calculated by multiplying the annuity factor (Equation 3.1) by the investment cost.

$$a = \frac{r}{1 - \frac{1}{(1+r)^n}} \tag{3.1}$$

Where r is the discount rate, and n is the expected lifetime of the given technology. In this project a discount rate of 7% [11] is used. The lifetime of the individual technologies are listed in Table 3.1. All cost data are based on the 2030 values presented in [20].

Technology	Investement [€/MW]	Fixed O&M [€/kW/year]	Marginal cost [€/MWh]	lifetime [years]
Onshore Wind	1182	35	0.015	25
Offshore Wind	2506	80	0.02	25
Solar PV	600	25	0.01	25
OCGT	400	15	58.4	30
Transmission	400 €/MW km +150000€ pr line	2%	0	40

Table 3.1: Generator parameters are based on the values from [20], and transmission parameters are based on the work presented in [21].

3.3 Energy demand

The data for the hourly electricity demand found in the European Network of Transmission System Operators data portal is used as energy demand [22]. The data has a resolution of one hour and is provided for all countries included in the model. In Figure 3.3, the summarized demand for the entire year for the individual countries is shown. A total of 3152TWh of energy was consumed by the countries combined in 2011.

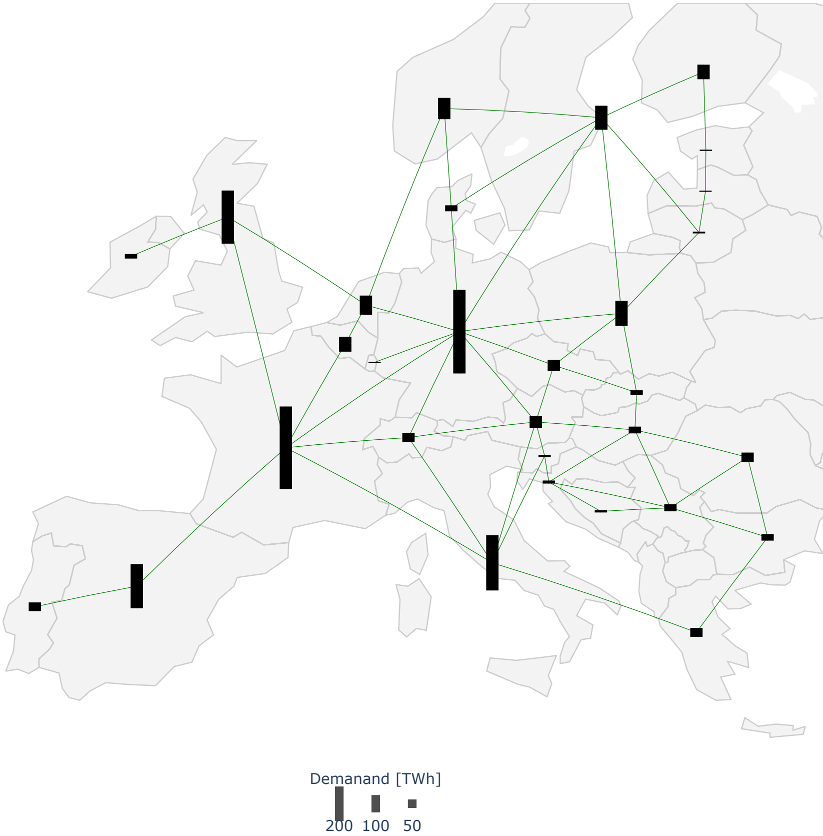


Figure 3.3: The figure shows the total electricity demand of the individual countries during an entire year.

3.4 Energy transmission

In the model used in this project, all transmission lines are treated as transport models with a coupled source and sink, only constrained by energy conservation at each connecting node. Transmission loss is thereby not considered. This approximation is assumed to be acceptable as most international transmission lines already are, or probably will be in the near future, controllable point-to-point high voltage direct current (HVDC) lines.

Line capacities initially start as zero, and can then be expanded if found feasible in the optimization, with no constraint on the maximum allowable capacity. The investment cost of line capacity is calculated as a cost pr MWkm plus an additional cost for a high voltage AC to DC converter pair. The price of a high voltage AC to DC converter pair is set to be 150000€ regardless of line capacity [21].

The length of each line is set as the distance between the centroids of each connecting country plus an additional 25%. The extra 25% is added to the line length as competitive land use and public acceptance issues will prohibit lines from being placed in optimal positions.

Furthermore, to satisfy n-1 security the price is adjusted with a factor of 1.5, to account for the extra installed capacity needed, as shown in [11].

$$c_l = (L \cdot I_s \cdot 1.25 + 150000)\,1.5 \cdot 1.02 \cdot a \tag{3.2}$$

3.5 Utilization of parallel programming and cluster computing

As the MGA approach described in Section 2.6 requires a high number of similar optimizations to be performed only with slightly changed objective functions, it is possible to achieve a great performance boost, by utilizing parallel programming and cluster computing.

The model used in this project is implemented in the open-source tool PyPSA [12], build for the programming language Python. The individual optimizations of the model are done with the optimization tool Gurobi [15]. The Gurobi solver is capable of using several computing cores to speed up the process of optimizing the model. Allowing Gurobi to use two cores, it is possible to find an optimal solution to the techno-economic model used in this project in approximately 20 minutes. As the MGA method presented in this project requires several optimizations of the problem, with a changing objective function, it is possible to reduce the computation time drastically by utilizing parallel programming.

Using the Python framework Multiprocessing [23], the MGA method presented in this project was implemented, capable of performing several optimizations at once. Using the PRIME compute cluster [24], it was possible to perform 16 parallel optimizations on the 32 core compute nodes, reducing the time needed per MGA study drastically. The Prime compute cluster [24], has 17 compute nodes with 24-36 cores available each operating at roughly 3 GHz, giving a total of 538 cores. Performing an MGA study requires anything in the range of 100-2000 optimizations of the model, each one requiring two cores for twenty minutes. Using a single compute node with 32 cores a single MGA study can be performed in a few hours up to two days, depending on the complexity of the problem.

4 Results

Several computational experiments have been performed to analyze the performance of the novel MGA approach developed in this project, and to study the techno-economic model of Europe presented. Initially, the results of a study performing regular optimization of the techno-economic model of Europe is presented, providing a basic insight into the techno-economic model used. Knowing the optimal solutions to the optimization problem, the results of a study implementing the novel MGA method presented in this project, are presented. This study seeks to investigate how the installed capacities of the different technologies are distributed across all near-optimal feasible solutions to the problem. Yet another MGA study is performed, analyzing the interplay between energy production in South and North Europe. This study also seeks to investigate the limitations of the MGA algorithm, in terms of the maximum allowable variables included in the decision space considered by the MGA algorithm. To compare the usefulness of the novel MGA approach developed in this project, a study comparing the novel MGA approach with previously presented MGA approaches have been performed. The results of this study will be presented, highlighting the benefits and disadvantages of a total of four different MGA approaches including the one presented in this project.

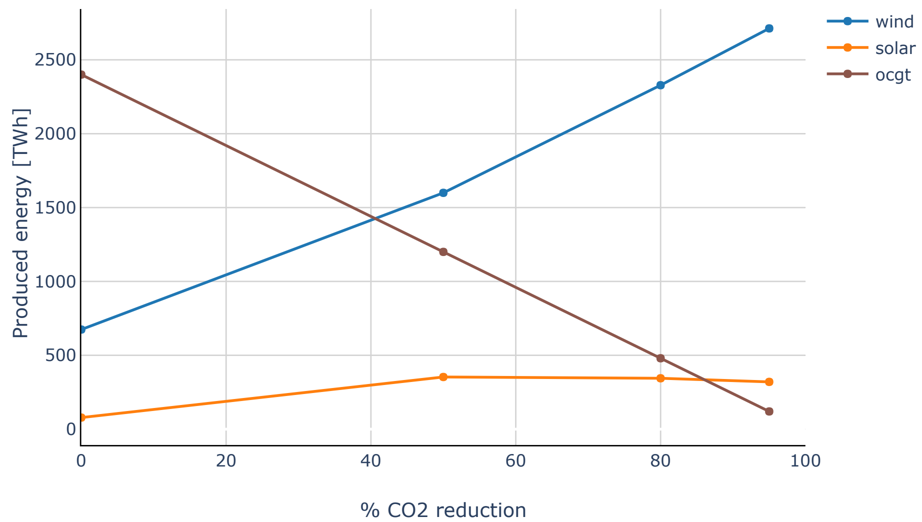
4.1 Optimal solutions

Before any MGA studies are performed, the baseline performance of the techno-economic model must be established. This is done by performing classic optimization of the models to find the optimal solutions.

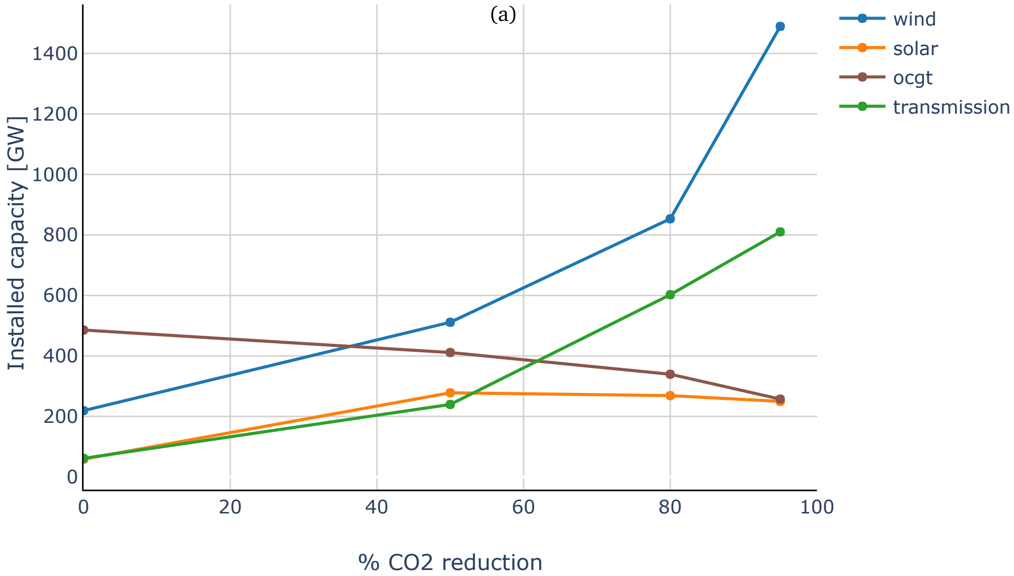
Initially, the model is optimized without introducing any CO_2 constraints, in order to establish a baseline for CO_2 emissions. The baseline CO_2 emissions measured in tonnes per year acts as the reference for CO_2 reductions in other scenarios. It is important to perform this study as real-world figures on CO_2 emissions doesn't necessarily compare to the figures found using this specific model. This is because the model used in this project simplifies the complex energy grid of Europe by a great deal, thereby introducing a lot of uncertainty. Therefore, it is relevant to compare the performance of the model used in this project to the performance of the actual energy system.

When the optimal solution for the model with no CO_2 constraints is found, a range of optimizations is performed with altering CO_2 constraints. The allowable CO_2 emissions are calculated as a percentage of the emission of the base scenario. A total of three CO_2 reduction levels was chosen to investigate being 50%, 80%, and 95% reduction compared to the base scenario. The objective of performing these optimizations is to investigate the effect of the CO_2 constraint on the model, before performing any MGA studies.

On Figure 4.1b, the summarized technology capacities are presented, showing how the two CO_2 neutral energy sources increase as the CO_2 constraint is tightened. It is important to note how the installed wind capacity increases rapidly compared to



(a)



(b)

Figure 4.1: Figure (a) shows the total amount of produced energy by the individual technologies during an entire year, for the four optimal solutions with different reductions in CO2 emissions. On Figure (b) a summary of the total installed technology capacity versus % CO_2 reduction for the four optimal solutions is presented.

solar PV that levels out, as the CO_2 reduction increases. This could indicate that any further solar PV capacity would primarily generate surplus energy, as long as no storage technologies are implemented. Studies have previously shown that solar PV is highly dependent on short term storage if it is to be utilized on a larger scale [25] [26]. This is due to the high daily fluctuations in energy production from solar PV and therefore solar PV has a great synergy with short term storage technologies.

Figure 4.1b further shows that if +90% CO_2 reduction should be achieved, a wind-solar mix of approximately 80% wind and 20% solar PV is desirable, when no storage solutions are implemented. This complies very well with the results presented in [25], where a study on the optimal wind and solar mix for Europe is performed.

Technology		Business as usual	50% CO_2	80% CO_2	95% CO_2
wind	[GW]	219.1	511.2	853.3	1489.5
Solar	[GW]	58.6	278.0	268.78	249.6
OCGT	[GW]	485.5	411.2	339.4	257.2
Transmission	[GW]	61.4	239.5	602.3	810.1
Gini coefficient		0.11	0.20	0.44	0.59
CO_2 emission	[MT]	1151.9	576.0	230.4	57.6
Objective value	[1e9€]	200.7	212.8	256.5	358.1

Table 4.1: Key data from the four optimal solutions are presented.

Comparing the produced energy in the four scenarios presented in Figure 4.1a with the installed capacities from Figure 4.1b, it is seen how the energy produced by gas turbines drops drastically compared to the gas turbine capacity as CO_2 emissions are reduced. This means that every GW of installed OCGT capacity produces less energy in the scenarios with a high CO_2 reduction. Instead of acting as a primary energy source the gas turbines change role and instead acts as backup capacity. The amount of produced energy from wind turbines compared to the installed capacity also seems to decrees when the CO_2 reduction increases above 80%. This could indicate that after this point, additional installed wind capacity generates more surplus than previously installed capacity. The decreasing capacity factors for both wind turbines and OCGT are what makes the scenario cost increase, as the same amount of energy is produced in all scenarios.

The Gini coefficients for the four scenarios is presented in Table 4.1 and on Figure 2.8. The Gini coefficient expresses the equality in production versus consumption for the four scenarios. A low Gini coefficient indicates that energy is being produced locally and opposite, a high Gini coefficient indicates that energy is being produced away from where it is needed. The Gini coefficients of the four optimal solutions appear to increase, together with the capacity of wind and solar power, as the CO_2 constraint is tightened. This could suggest that it is beneficial to install wind and solar power in countries with good wind and solar profiles and then transmit the energy to countries with less favorable wind and solar profiles. Analyzing the average capacity factors for wind power in Figure 3.2, it is seen that higher capacity factors are common in northern countries. This corresponds very well with where the large capacities of wind are placed in the optimal solutions subject to high CO_2 constraints presented in Figure 4.2c and 4.2d. Comparing the installed capacities presented in Figure 4.2 with the total demand of the individual countries presented in Figure 3.3, it is seen that as CO_2 emissions are reduced, the installed capacities move further and further away from the countries with the largest demands.

Business as usual

In the business as usual scenario, seen in Figure 4.2a, where no constraint on the CO_2 emission is implemented, energy is primarily supplied by gas turbines as expected. Any significant capacities of variable renewable energy sources are only implemented in countries where the price of energy produced from such technologies can compete with the price of energy from gas turbines. Analyzing Figure 3.2 it is found that wind energy is favorable in the northern countries and solar energy only becomes favorable in most southern countries, in this case, Spain and Portugal. Furthermore, the energy generation is spread, fairly even across the network, thereby requiring

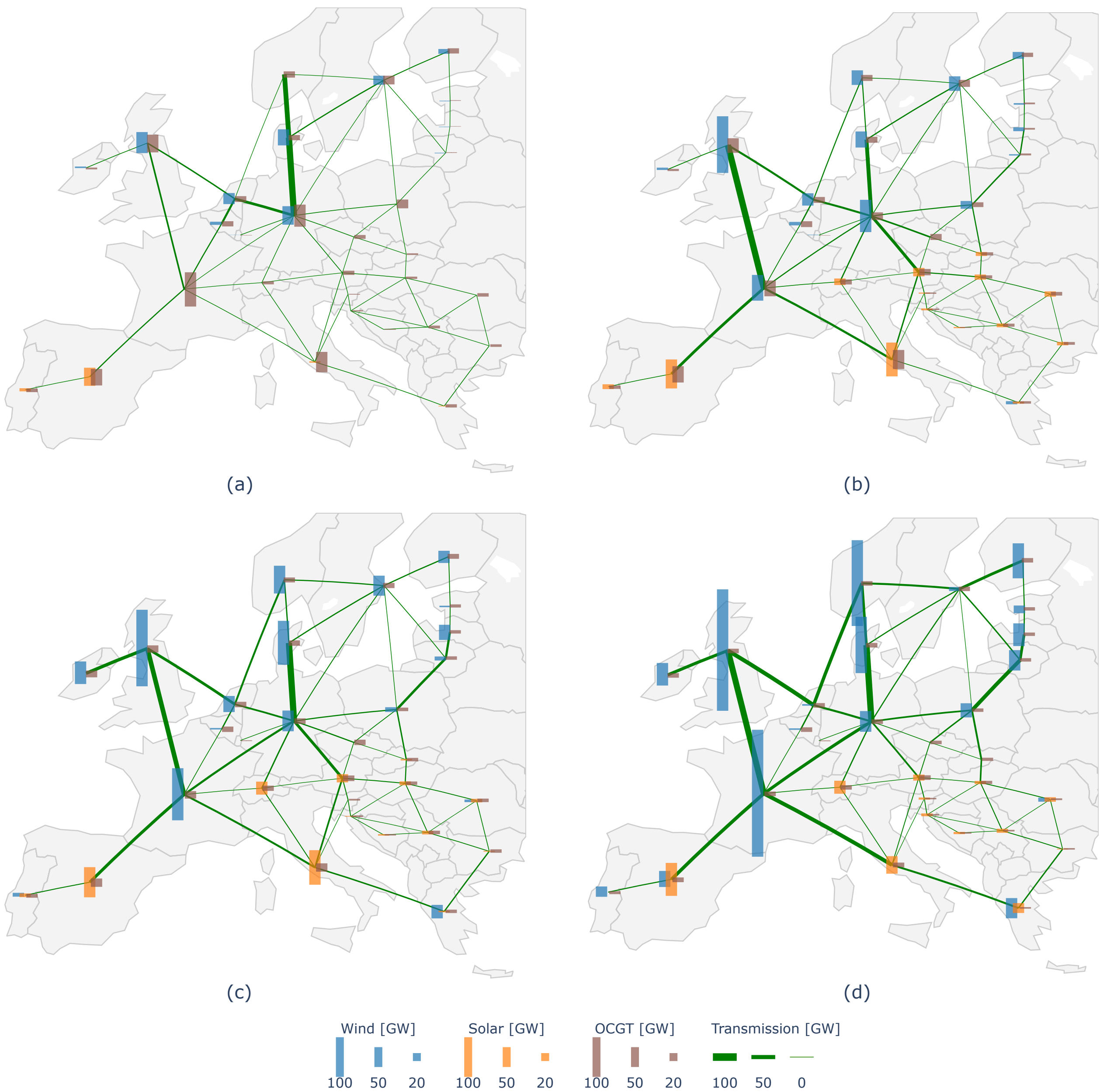


Figure 4.2: The figure presents the layout of technology capacities of all four optimal solutions. On figure (a) the business as usual scenario is presented, (b) shows data for the scenario with 50% CO2 reduction, (c) 80% reduction and (d) 95% reduction in CO2 emissions.

less transmission capacities, and thereby also resulting in a fairly low Gini coefficient of 0.11. In this scenario, transmission is purely installed between countries with significant shares of variable renewable energy sources.

The CO_2 emission in the base scenario without CO_2 constraints was found to be 1151.9 MT CO_2/year, which complies reasonably well with the 2011 CO_2 emission for the EU-28 countries energy sector, found by the European Environment Agency (EEA) to be 1517.3 MT CO_2/year [1]. Although the numbers are off by some hundred MT CO_2/year, and the numbers from EEA only represent the EU-28 countries, this comparison can conclude that the model used in this project, despite its coarse spatial resolution and a small number of included technologies, is capable of producing results with acceptable accuracy.

Reduced CO_2 emission scenarios

The distribution of installed capacities of the scenarios with reduced CO_2 emissions are presented on Figure 4.2b, c and d, with respectively 50, 80 and 95% CO_2 reduction. All three scenarios implement wind energy in northern Europe and implement large amounts of transmission capacity between all countries with high shares of wind power. OCGT capacity appears to be evenly spread across the countries. Significant shares of solar power are only implemented in southern Europe, and even with a CO_2 reduction of 95%, the most northern country to install solar power is Austria. As the CO_2 constraint is tightened, the Gini coefficient increases, as the countries with favorable conditions for solar and wind power, simply increases their capacity of these technologies. Whereas, technology capacities in less favorable countries stay unaffected.

Looking at the scenario costs in Table 4.1, it is seen that as the CO_2 constraint is tightened, the cost increases rapidly. From the business as usual scenario to the scenario with a 50% reduction, there is only an increase in the price of 6%. When the CO_2 emissions are to be reduced beyond this point, price increases rapidly, and to achieve a reduction of 95% the cost almost double, having increased with 78%, compared to the business as usual scenario. This also complies with the data shown in Figure 4.1, where it can be seen that total installed production capacity increases, even though the energy demand stays the same. These results are very much in line with the results found in [27], where surplus electricity is investigated in scenarios with large shares of variable renewable energy sources.

4.2 MGA study of four-dimensional decision space

The objective of this experiment is to document the performance of the techniques developed in this project, by studying the techno-economic model presented in Chapter 3. Focus is placed on the relevance/usability of the data extracted from the model using the newly developed technique, as well as shear performance measured in computation time. The experiment will explore the interplay between variable renewable energy sources and transmission on an international level spanning all countries in the model.

In this experiment, the decision space is reduced to just four dimensions by grouping the decision variables as explained in Section 2.2. The four variables in the new decision space are the total amount of installed gas turbine (OCGT) capacity, wind turbine capacity, solar PV capacity and the total installed transmission capacity. Using a low dimensional space allows for a thorough exploration of the feasible space as a lower-dimensional decision space, needs fewer computations per study, thereby reducing computation time.

As a thorough exploration of the near-optimal feasible space is desired, to evaluate the performance of the technique, combined with a desire to learn more about the features of the near-optimal feasible space, a range of MGA studies is performed with varying parameters. For every single MGA exploration, two parameters can be altered. These are the amount of reduction in CO_2 emission compared to the base model, and the amount of MGA slack used. For this experiment, it was chosen to iterate over both of these variables exploring CO_2 reductions of 0, 50, 80 and 95%, and MGA slacks of 1, 2, 5, and 10%. This means that a total of 16 MGA studies is to be performed.

Even though it was established that multiplicity of sampled points has a large effect on the results, when a high dimensional space is reduced to one of much lower dimension, it has not been accounted for in this study, due to the added computing time needed.

The distribution of capacities found in the four MGA studies is presented in Figure 4.3. As was seen with the optimal solutions, the figure clearly shows how decreasing CO_2 emissions leads to an increase in wind and transmission capacity and a reduction in OCGT capacity. The results presented in Figure 4.3, furthermore provide information about the distribution of capacities. When analyzing the capacity distributions presented in Figure 4.3 it is important to consider that there hasn't been accounted for multiplicity, meaning that the true distributions would be narrower, as discussed in Section 2.7.

It is particularly interesting to see how the OCGT capacity distribution has a sharp lower bound, indicating that a rather fixed minimum amount of OCGT capacity is needed in all scenarios. The minimum OCGT capacity needed will be given by the remaining energy demand in the hour with the combined lowest availability of wind and solar power. To remove the need for backup generators such as gas turbines, would require unrealistically large capacities of wind and solar power.

Figure 4.3 also shows how the distribution of capacities become wider as CO_2 emission is decreased. This essentially means that larger configuration flexibility is available as CO_2 emissions are reduced. It could, however also be a result of how the MGA slack is defined. The MGA slack is defined as a percentage of the optimal solution scenario cost. As the cost of the optimal solutions varies a lot from the scenario with no CO_2 constraint to the scenario with tight CO_2 constraint, the allowable extra cost of scenarios also varies a lot, if measured in shear size and not as a percentage of the optimal solution objective value.

The distributions of solar power capacity in all four scenarios presented in Figure 4.3, all reach a value of 0 GW, meaning that solar power can be omitted if desired. In the scenarios with larger CO_2 reduction, this would, however, be a very extreme scenario, leading to a very specific design of the energy network. It is however not possible to configure a solution in a way that omits wind power in any of the four scenarios.

In Figure 4.4 correlations of the four variables of the simplified decision space are shown. Data from all MGA studies performed have been used to generate this figure. The figure shows a strong correlation between wind power and transmission with a correlation of 0.50. This corresponds well with the results presented in Figure 4.2, where reinforcements on to the transmission grid are made as more wind is introduced to the model. Similar correlations between large shares of wind power and transmission capacity have been found in an article investigating this subject [28]. On the other hand, when analyzing the correlation between transmission and solar power a small negative correlation of -0.07 is found. This could suggest that the solar capacities installed, only produce enough energy to supply the country wherein they are installed, and therefore a need to transmit electricity generated by solar power doesn't exist. The correlation between energy production from solar PV in different countries has been studied in [29], where a strong correlation between production periods was found. As sunrise and sunset are similar in all European countries, the production of energy from solar PV happens simultaneously across entire Europe.

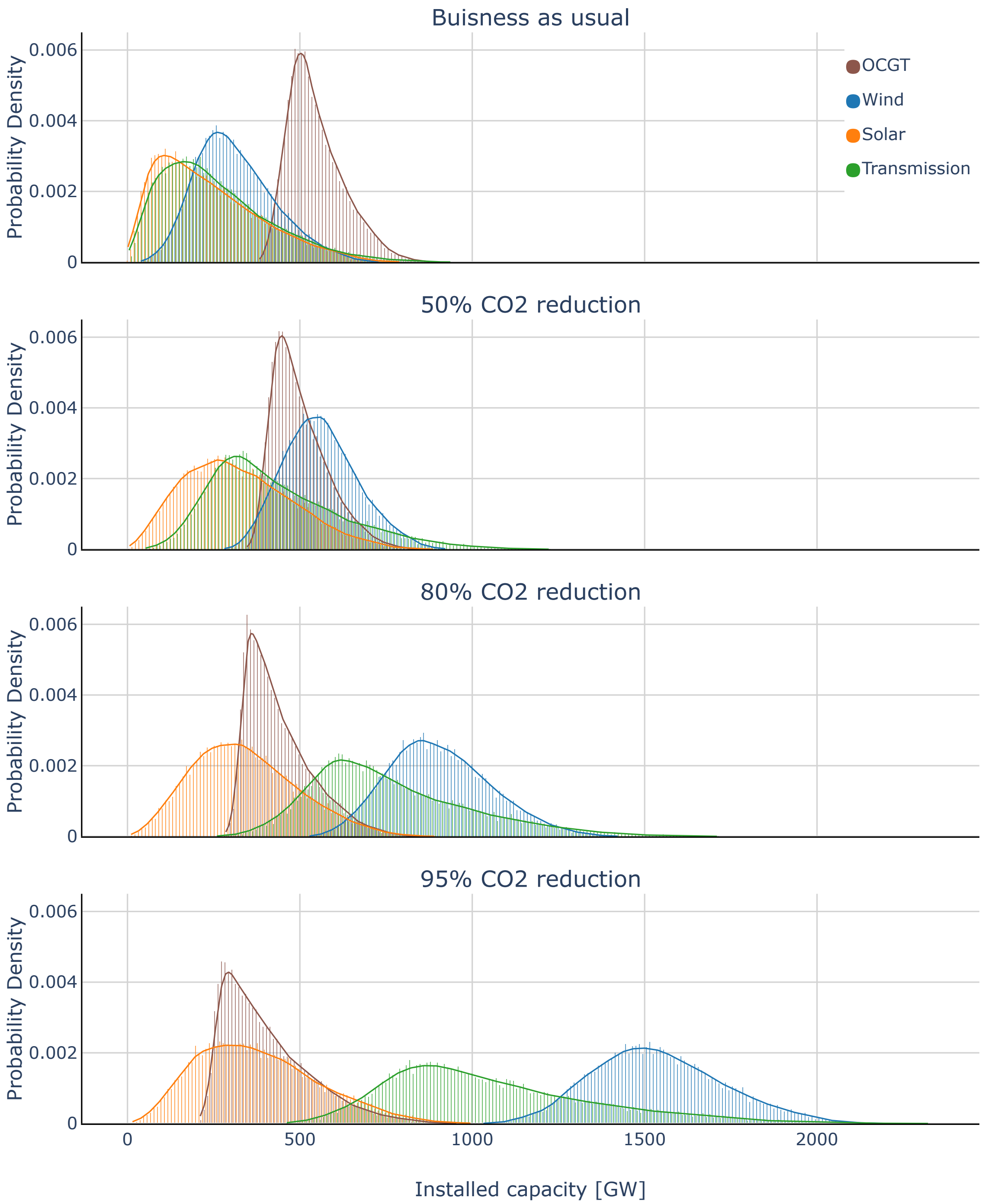


Figure 4.3: The figure shows the distribution of technology capacities for four MGA studies, where a MGA slack of 10% have been used.

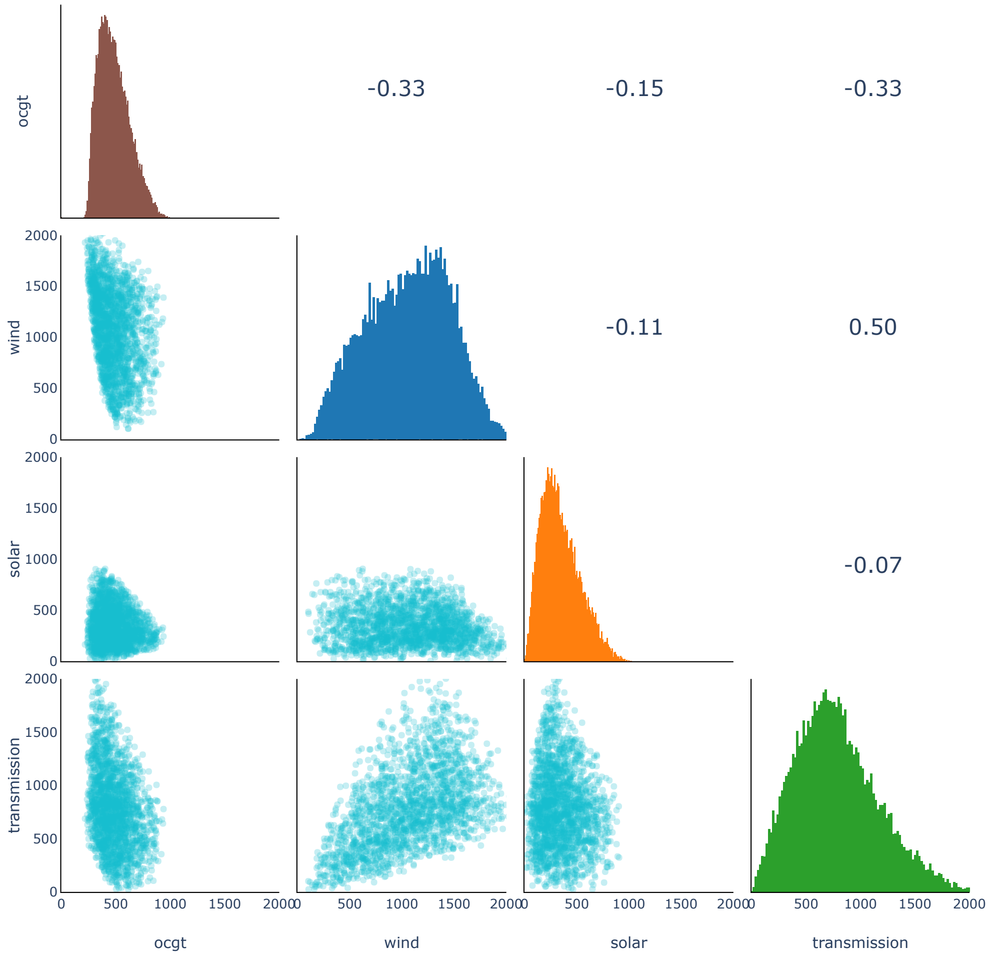


Figure 4.4: This figure presents all variable correlations of the four MGA studies performed with 10% MGA slack. Distributions of the four individual technologies are plotted on the diagonal, scatter plots of the data are shown in the lower left half and correlation values are presented in the upper right half of the plot matrix.

Therefore an increase in transmission capacity would not allow more solar PV capacity to be installed. These results from [29], corresponds very well with what's seen in this project. A very strong negative correlation between OCGT and transmission is also presented in Figure 4.4 with a correlation of -0.33. The capacity factor of OCGT does not depend on geographically determined factors and therefore OCGT capacity is installed where it is needed, leading to a negative correlation with transmission capacity. Another strong negative correlation is between wind power and OCGT with a value of -0.33. These two technologies appear to compete somewhat equally as energy sources.

Analyzing the correlations of a single MGA study, namely the one with a CO_2 constraint of 95% presented in Figure 4.5, it is a completely different story. Here the most significant correlation is between wind and solar power, which has a negative correlation of -0.72. This indicates that when a specific CO_2 reduction is desired, wind and solar power competes evenly as energy-generating technologies. It is interesting to see how wind power and transmission has a small negative correlation when the correlations are calculated for just a single MGA study, compared to the large positive correlation, when data from all scenarios were considered.

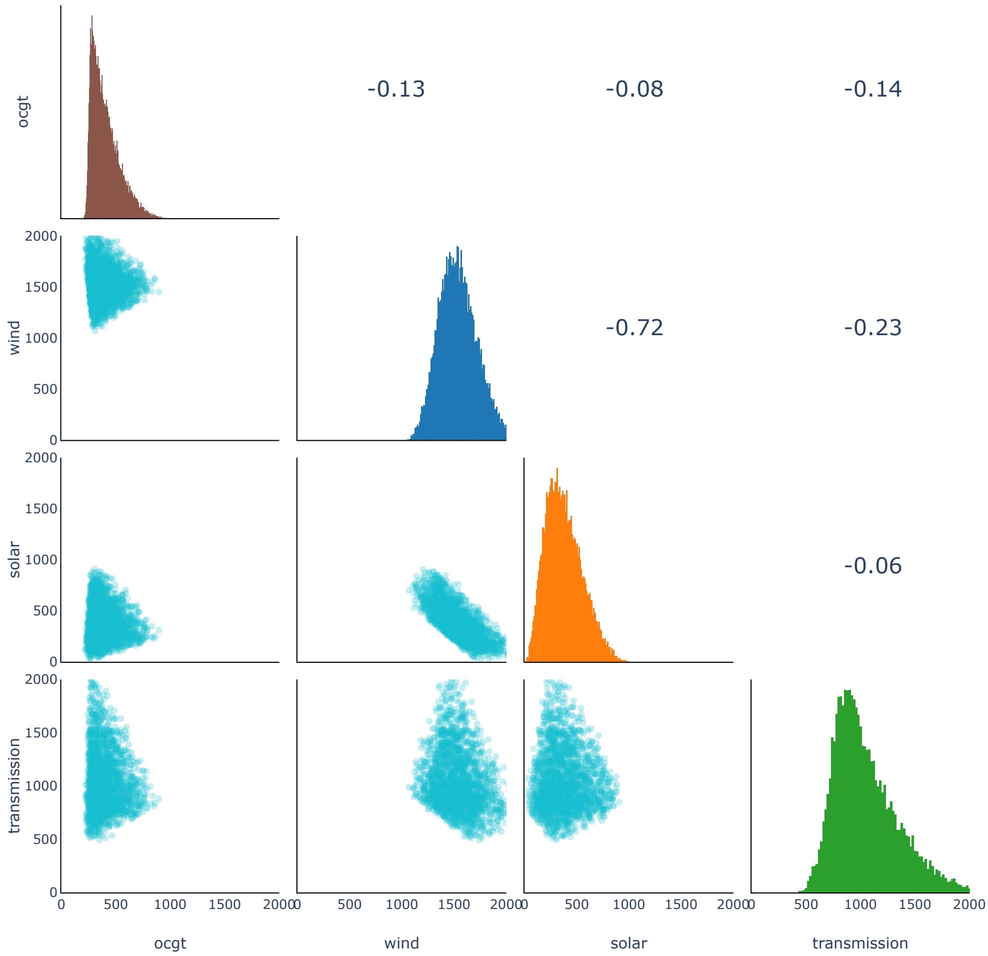
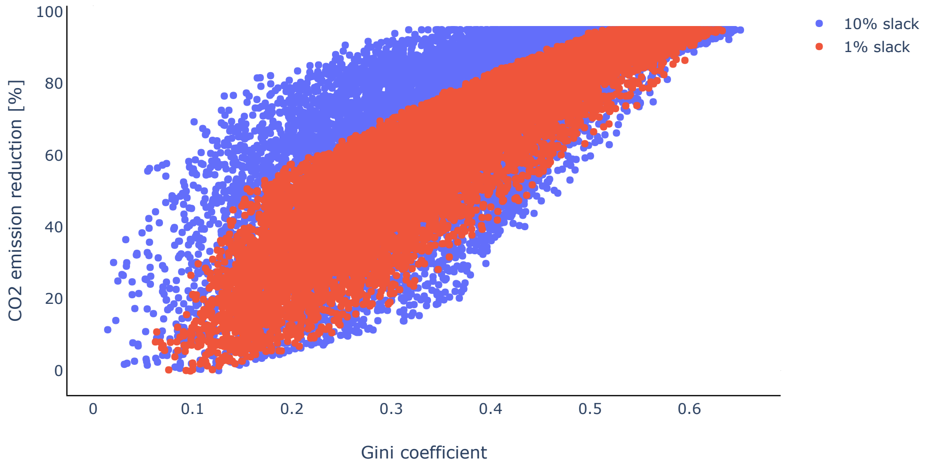


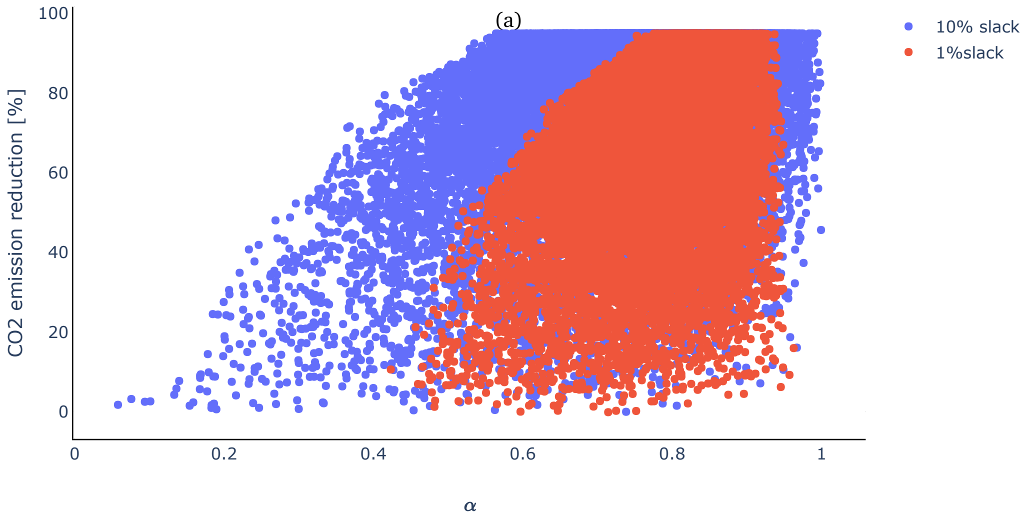
Figure 4.5: The figure shows all variable correlations (installed capacity in GW) of the MGA study with a CO_2 emission reduction of 95%. Distributions of the four individual technologies are plotted on the diagonal, scatter plots of the data are shown in the lower left half and correlation values are presented in the upper right half of the plot matrix.

In this project, the cost has been treated as one combined cost for the entire European energy network. This is, however, a very large simplification of the problem. In reality, the energy system cost is treated on a national level, and therefore it is desired to distribute the energy system cost as evenly across all countries in the model. In this project, the Gini coefficient is used as a measure for the equality of energy production versus consumption. This Gini coefficient can also be used as a measure of equality in the distribution of system costs.

In Figure 4.6a the Gini coefficient for all studies performed in this experiment is plotted against the CO_2 reduction. On the figure, it is seen that there is a very strong positive correlation between CO_2 reduction and the Gini coefficient. This means that as requirements to reductions in CO_2 emissions increase, more energy needs to be



(a)



(b)

Figure 4.6: Figure (a) plots the Gini coefficient versus the reduction in CO_2 emissions. Figure (b) plots the wind/solar mix α versus the reduction in CO_2 emissions.

produced outside the countries where it is needed. A larger Gini coefficient also means that some countries will have to invest more money in modernizing the energy grid by installing larger capacities of renewable energy sources, and other countries will be depending on the import of energy from countries with large capacities of renewable energy. As a country, it is desired to be self-sufficient with energy as energy is one of the most critical resources. The desire to reduce the Gini coefficient might, therefore, introduce a larger financial willingness. On Figure 4.6a, it is seen that a 1% slack on total system cost requires a Gini coefficient of at less 0.5 for a CO_2 reduction of 95%, but increasing the slack to 10% allows for a Gini coefficient as low as 0.3 for the same reduction in CO_2 emissions.

In this study, the only variable renewable energy sources included are the two major variable renewable energy technologies, wind, and solar power. From the correlation plot on Figure 4.4, a slight negative correlation with a value of -0.11 was seen. Such a small correlation value indicates that these two technologies interfere very little with each other and that the share of wind and solar power remain somewhat constant. On Figure 4.6b, the share of wind and solar power is plotted against CO_2 emissions. The share of wind and solar power α is calculated as the wind capacity relative to the total capacity of variable renewable energy.

$$\alpha = \frac{\text{wind capacity}}{\text{wind capacity} + \text{solar capacity}} \tag{4.1}$$

In Figure 4.6b it is seen that the wind-solar mix has an average of around 0.8, complying very well with the results from [25], where the mix of wind and solar power was studied. The penetration of wind power increases as CO_2 emissions decrease, which complies well with the results seen in the optimal solutions in Figure 4.1a, where the implementation of solar PV capacity stagnates at CO_2 reduction levels higher than 50%. Figure 4.6b, further shows that with a slack of 10% of total system cost a wide span is available in the wind-solar mix. At 95% CO_2 reduction, it is possible to have a scenario that is 100% wind dominated or a scenario where 50/50 mix between wind and solar is used.

Having analyzed the results from this experiment, it can be concluded that the developed MGA algorithm is capable of providing useful insights regarding the techno-economic model investigated. Information that would not have been made available with other methods was extracted allowing for great insights regarding the possibilities available within the near-optimal feasible space.

4.3 MGA study using seven decision variables

In the previous experiment, variables were grouped by technology type. It is, however, possible to group the variables in any way desired. Therefore a study investigating the interplay between energy production in North and South Europe has been performed, grouping the variables not only by technology type but also by spatial location. A total of 7 grouped variables is formed, including OCGT, wind and solar power from both North and South Europa, and the total amount of transmission capacity. The grouped variables then become:

$$\mathbf{x} = \begin{Bmatrix} x_1: & \text{North OCGT} \\ x_2: & \text{North wind} \\ x_3: & \text{North solar} \\ x_4: & \text{South OCGT} \\ x_5: & \text{South wind} \\ x_6: & \text{South solar} \\ x_7: & \text{Transmission} \end{Bmatrix}$$

A parting line between North and South Europe was drawn at latitude 49.2421, which is equivalent to the median of the latitude position of all country centroids included in the model. All countries with a centroid north of latitude 49.2421 are considered as North Europe, and all other countries are considered as South Europe. A total of 15 countries are included in each category. The parting line is presented on Figure 4.7.

Using the novel MGA method presented in this project, the capacities of all near-optimal solutions were found, using an MGA slack of 10% on four scenarios with respectively 0, 50, 80 and 95% reduction in CO_2 emissions.

In Figure 4.8, a histogram presenting all seven grouped variables for all four scenarios is shown. Much like the study performed using only four grouped variables seen in Figure 4.3, the wind, and solar capacities increase as CO_2 emissions are reduced.

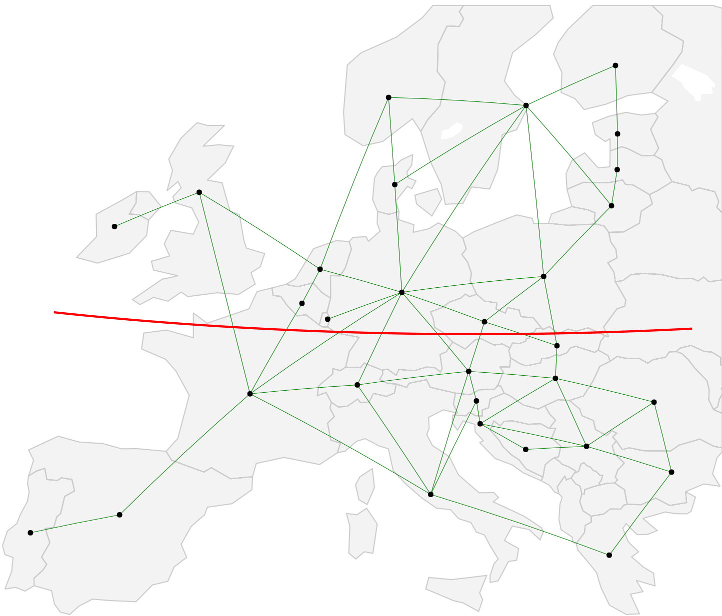


Figure 4.7: The figure shows the topology of the techno-economic model used in the study. The red line indicates the parting line between North and South Europe.

Analyzing the data presented in Figure 4.8, it is seen that wind power in northern Europe is preferred over any other energy generating technologies, as CO_2 emissions are lowered. Comparing solar power in North and South Europe, it is seen that South Europe in all scenarios has a larger amount of installed capacity. Scenarios, where large shares of solar power are installed in North Europe, are however also feasible, and in a scenario where CO_2 emissions are reduced by 95%, a scenario with 600GW of installed solar capacity in North Europe would be feasible.

Analyzing the distributions of wind and solar power in Figure 4.8, it is seen that there is an overlap between respectively wind in North and South Europe and solar power in North and South Europe. This shows, that a scenario with more solar capacity in northern Europe compared to South Europe is feasible with a 10% slack on cost or a scenario with more wind in southern Europe than northern Europe.

Analyzing the correlations presented in Figure 4.9, results show a significant correlation between wind and transmission. Especially wind in northern Europe has a strong correlation with transmission, with a correlation above 0.4. Wind power in South Europe correlates positively with transmission too, only with a correlation factor of 0.12. The results presented in Figure 4.9, further shows strong negative correlations between OCGT in North and South Europe, indicating that OCGT in North Europe competes very directly with OCGT power in South Europe. A similar tendency is seen between solar power in North and South Europe. Interestingly, no correlation is seen between wind in North and South Europe, indicating that the installed capacity of wind at either end of Europe, does not affect each other.

Each of the four studies performed in this experiment required 48 hours of computing time on the PRIME computing cluster [24] using a single 32 core computing node. Including more variables in the decision space considered by the MGA algorithm is possible but the practical maximum is estimated to be roughly 10 decision variables.

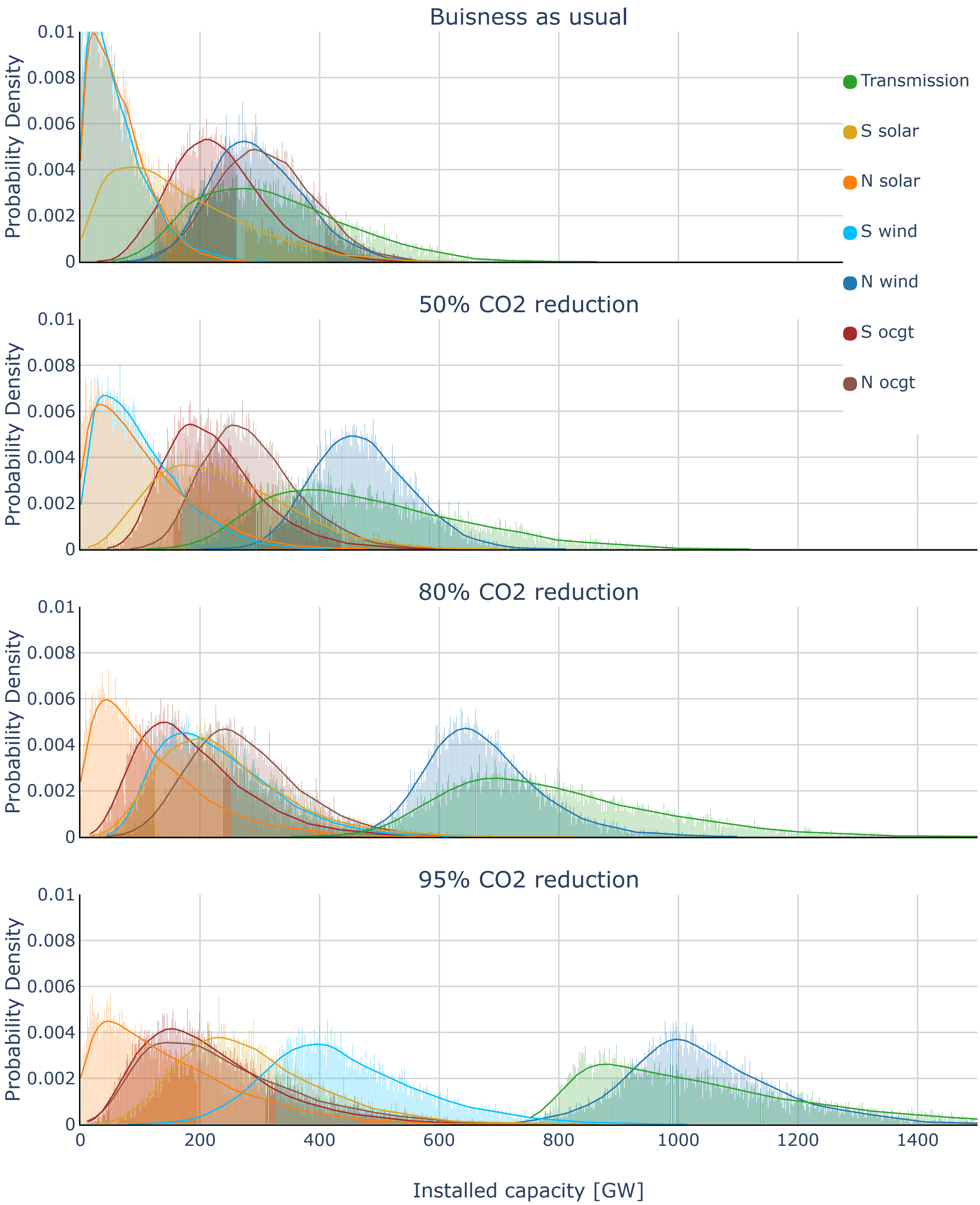


Figure 4.8: The figure shows the distribution of technology capacities for four MGA studies, where a MGA slack of 10% have been used, and seven variables has been included in the MGA study.

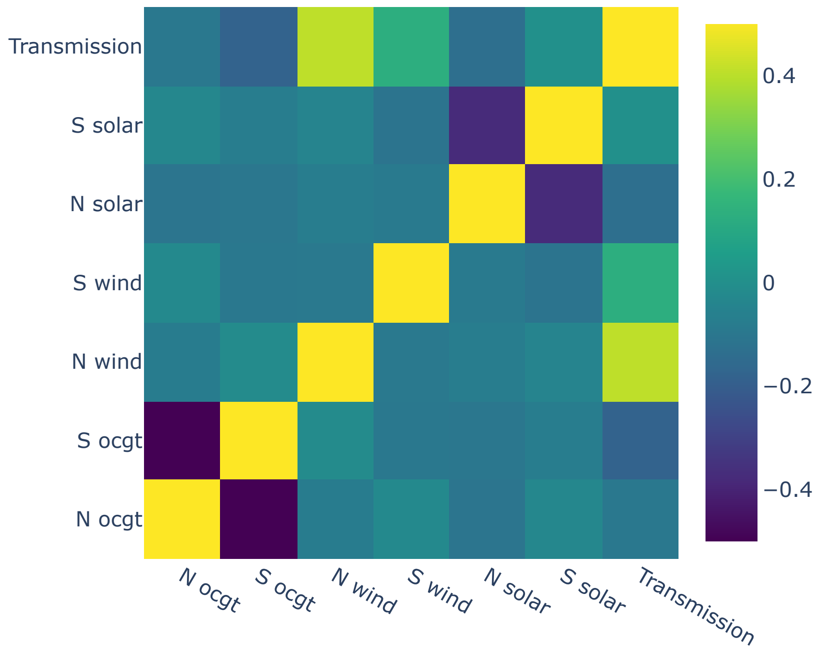


Figure 4.9: The figure shows a heatmap representing the correlation matrix of the variables in the four MGA studies performed with seven variables included in the study.

4.4 Comparison of MGA algorithms

The goal of this experiment is to highlight the benefits and weaknesses of existing MGA methods compared to the novel MGA approach presented in this project. The following four different MGA techniques will be explored: The HSJ approach presented in [3], an approach where groups of variables are maximized and minimized presented in [10] together with an approach where all decision variables are maximized and minimized and the novel approach presented in this project.

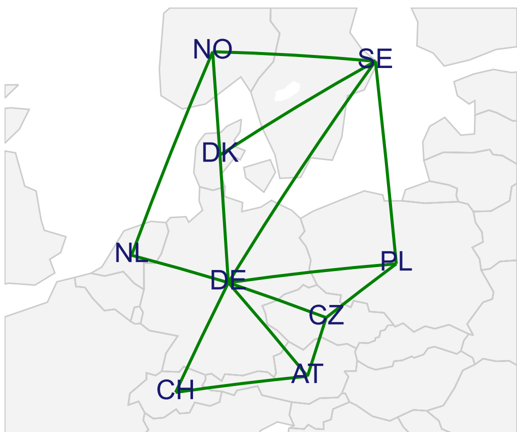


Figure 4.10: The figure shows the topology of the simplified network used in study comparing performance of MGA algorithms.

When comparing MGA approaches in this section, it is very important to keep in mind what the goal of performing an MGA analysis is, and what measure characterizes a good MGA technique. The overall objective is to explorer the possibilities for alternative solutions to the optimization problem within a certain range of economic slack. Solutions to a techno-economic problem as the one considered in this project can, however, be different in a wide range of manners, as the decision space is high dimensional. This makes it hard to determine the coverage of the decision space for a given MGA method, as the extent of the decision space is unknown. Instead, the results found with the different MGA methods will be analyzed, compared and discussed in an attempt to determine strengths, and weaknesses for the MGA methods.

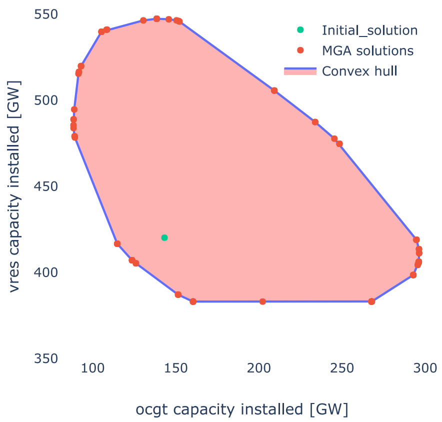
To generate comparable results, all four MGA methods are implemented on the same simplified network. As the focus of this experiment is to compare methods, and not to analyze a techno-economic model, a simplified model is used to reduce the computation time needed, and the complexity of the results. The techno-economic model used in this experiment includes only nine of the thirty countries from the full model as shown in Figure 4.10. Furthermore, only a single 24 hour period is simulated. The result is a model with 27 variables (9 countries with 3 technologies each) when hourly dispatch is not considered as a decision variable. Furthermore, a CO_2 constraint is employed forcing the model to reduce CO_2 emissions with 80% compared to an unrestricted scenario.

Using the approach presented in Section 2.2 to reduce the dimensionality, a new two-dimensional decision space is formed, by grouping the variables in a group representing all installed gas turbine capacity and the other group representing all variable renewable energy source capacity (wind and solar power). This allows for easy visualization and understanding of the results.

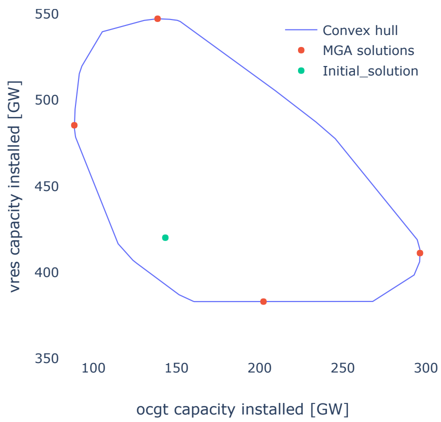
Initially, the novel MGA approach presented in this project was deployed to find the convex hull containing all solutions in the two-dimensional decision space. As this method converges towards the full solution, the result can be considered as the full solution to the two-dimensional problem. The found convex hull is presented in Figure 4.11a. The shaded area of the convex hull indicates that the novel MGA approach extracts information about the entire hull volume in contrast to all other MGA approaches that terminates when a set of different solutions is found. Despite using a simplified techno-economic model and reducing dimensionality to just two dimensions, the shape of the convex hull is still rather complex.

In Figure 4.11b, the results of the method from [10] where the grouped decision variables are maximized and minimized, are presented. This method effectively performs the same initial MGA iteration as the novel MGA approach, and therefore it finds solutions also found be the novel MGA approach. In the work presented in [10], these maximized and minimized solutions are used to create upper and lower bounds on the summarized capacities. Four individual solutions found using this technique are presented in Figure 4.12.

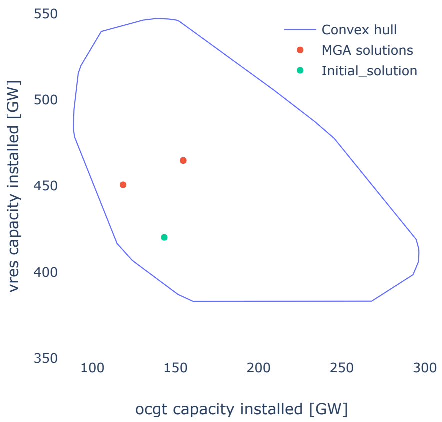
Analyzing the result of the HSJ MGA method in Figure 4.11c, it is clear that it does not search towards the edge of the convex hull as the first two methods. The reason hereof is that the HSJ method doesn't consider the grouped variables. Instead, it seeks to create maximally different solutions by implementing technologies not included in the previous solution. Here similar technologies from different countries are treated as two individual technologies. Analyzing the individual HSJ solutions presented in Figure 4.13 they are very different. Although all the solutions implement similar capacities of renewable energy sources and OCGT when summarized, the placement of the capacity implemented varies a lot from solution to solution. Looking at the two HSJ solutions from Figure 4.13b and 4.13c, where solution (b) implements all VRES in



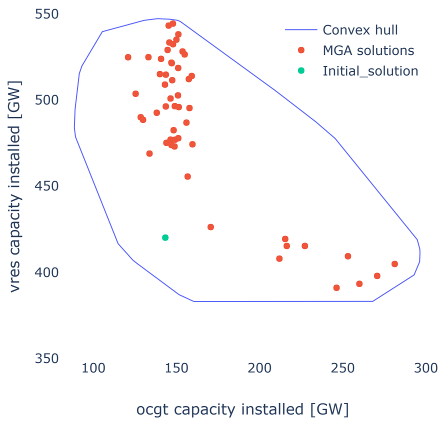
(a)



(b)



(c)



(d)

Figure 4.11: On the figure the found MGA solutions within the reduced decision space of four MGA algorithms is presented. Figure (a) presents the results from the novel MGA method. Figure (b) shows the results from the maximization and minimization of grouped variables approach, as presented in [10]. Figure (c) presents results from the HSJ MGA method from [3], and Figure (d) shows the results from the maximization and minimization of all decision variables as presented in [10].

Denmark and Germany, and solution (c) implements no VRES in Germany but spreads it to the surrounding countries, it is easy to see that the solutions are topologically different but when analyzing the summarized capacities they are very similar.

The overall conclusion from this study must be that the novel MGA approach developed in this project shows to be very effective in its ability to determine the shape of the near-optimal feasible space. Although it performs more optimizations than the other MGA methods, it uses these optimization runs wisely to investigate all regions of the decision space.

Figure 4.12: On the figure a presentation of capacity distributions from four solutions using the novel MGA results is presented. The four studies have used the following objective functions: (a) Minimize OCGT, (b) Minimize VRES, (c) Maximize OCGT, (d) Maximize VRES

Figure 4.13: On the figure a presentation of capacity distributions from three solutions using the HSJ MGA method is presented. Figure (a) shows the optimal solution, and figure (b) and (c) are HSJ solutions.

5 Discussion

This chapter will critically examine the results found in this project and the assumptions made to achieve these results. The objective of this project has been to explore all near-optimal solutions of the techno-economic model of Europe presented, to determine common characteristics of these solutions. A method capable of mapping all near-optimal solutions in a structured manner has been developed, and the usefulness of this method and its performance will be critically examined in this chapter.

The method developed to search the near-optimal feasible space has been applied to the problem of techno-economic energy system optimization on an international scale. There is however nothing preventing this method from being used in other fields of numeric modeling and optimization, as the method itself is formulated in general terms. The method developed does, however, require that the investigated problem is linear and convex.

Initially, the formulation of the techno-economic model was studied, and based on key findings, a method capable of searching the near-optimal feasible space was proposed. Several studies have been performed with the proposed MGA method analyzing the presented techno-economic model of Europe. Key results of the studies show that the proposed method is capable of providing useful insights, revealing must-have technology capacities and simultaneously indicating where solution flexibility is available.

Analyzing the results achieved in both studies using the presented MGA algorithm in Section 4.2 and 4.3, a vast complexity amongst the near-optimal solutions is found. The complexity of the near-optimal feasible spaces found using relatively small MGA slacks shows the importance of addressing structural uncertainty with MGA algorithms, as small flaws in the model formulation can result in solutions representing very different configurations of the energy network. Accepting an uncertainty in total system cost of a very large public planning project, such as modernizing the entire European electricity grid, to be at least 10% is not unrealistic. This in term means that all solutions found in the MGA studies using a 10% slack on total system cost are equally realistic. Analyzing the presented results from both studies using the MGA method on the full Europe model in Section 4.2 and 4.3, very large spans in technology capacities are found. In Figure 4.3, it is seen that in a scenario where Europe decreases CO_2 reductions with 95%, an energy network with 1000GW of installed wind power across Europe is just as feasible as a scenario with 2000GW of installed wind power. These are huge spans, and that is why a single optimal solution is far from enough when modeling future scenarios of energy systems.

The results achieved through the use of MGA provides the decision-maker with a framework exhibiting the possibilities available with a given system. This allows for an agile process when synthesizing future energy systems as a continuous range of possible energy systems are made available. Instead of striving towards a single optimal solution the decision-maker is capable of assessing the desires of all stakeholders and designing an energy solution within the frames presented in the MGA results, satisfying as many individual objectives as possible. Results generated with previously presented MGA methods such as the work presented by J. DeCarolis in [3], where a small set of alternative solutions to a techno-economic optimization problem is found by using the HSJ MGA algorithm, can be rather complex to understand. Grasping over a large number of alternative solutions at once is a very complex task, and by generating only a small number of alternative solution the algorithm fails to present all the possibilities within the near-optimal feasible space. With the MGA algorithm presented in this project a continuous span of alternative solutions is found, making it simpler for the decision-maker to alter the desired solution such that it satisfies as many stakeholders as possible. A critical flaw in previously presented MGA algorithms such as the HSJ algorithm used in [3], and the maximize/minimization of all variables used in [10], is that there is no guarantee that all possible alternative solutions have been exposed. The nature of the proposed MGA algorithm from this project ensures that it converges towards the full solution as it iterates. This provides a guarantee that all possible configurations of the energy network are well investigated, removing any bias that the other MGA methods might have.

The benefits of using MGA are clearly shown in this report, but using increasingly complex algorithms does however come at a cost. The MGA algorithm presented in this project introduces a significant increase in the computational time needed for a single study. In the MGA study searching a decision space in 7 dimensions more than 2000 optimizations of the techno-economic problem were performed. This requires much more computing time than classic optimization where a single or perhaps a few optimizations are performed. Previously presented MGA algorithms such as the HSJ [3] algorithm rely on a handful of optimizations usually less than 10, to generate its results. Therefore, one must consider if the gained insights are worth the added computational time. But with cloud computing becoming more and more widespread and prices for computing time as low as 0.64$ per hour for a 64 core compute note with 240GB of memory [30] on Google's cloud computing platform, a very large increase in computational time can be accepted for gained insights.

When increasing the number of variables included in the reduced decision space the computation time needed to perform an MGA study increases rapidly. There are mainly two reasons why the computational time needed increases. First of all, using the convex hull face normal's as searching directions, introduce a lot more search directions as the dimension of the convex hull increases. A higher-dimensional geometry simply has a more complex shape, and thereby more faces than a lower-dimensional one, thus more optimizations are needed. A different factor is that it is simply harder to compute the convex hull of a higher dimensional set of points. The computation time for the algorithm used to compute the convex hull of a finite set of points in this project [31] increases at roughly $O^*(n^3)$. The result of this is that it is unfeasible to use the presented MGA algorithm on a decision space of very large dimension. The largest number of dimensions used for a successful MGA study in this project is seven, but it is estimated that a study using ten dimensions is feasible.

Due to the increasing computing time needed as the number of decision variables increases, the developed MGA algorithm is not capable of handling the entire decision space. Therefore it is necessary to limit the decision space considered by the MGA algorithm to include only a few variables. Selecting variables to include in the reduced decision space, allows the modeler to focus the MGA study on certain areas of interest. In this project, two studies with different focus have been performed providing information on two different subjects. In future MGA studies, one could imagine the variables included in the reduced decision space being technology capacities installed in a single country or one might focus on the capacities of individual transmission lines. When selecting variables to explorer, the modeler has the opportunity to focus the study on certain elements of the model, hereby introducing the opportunity to investigate the interplay between a single technology and the entire European electricity grid.

When reducing the dimensions of the considered decision space by grouping variables together, the solutions located within this reduced decision space are no longer evenly distributed. In the reduced decision space, a single solution covers a range of possible solutions. This multiplicity of solutions in the reduced decision space was investigated in Section 2.7, where a method capable of estimating the effect of this feature on small problems was proposed and the implications of not considering multiplicity in larger problems were discussed. The effect of multiplicity is significant and to improve the proposed MGA method, future works should investigate methods capable of determining the multiplicity of larger problems.

6 Conclusion

The near-optimal feasible space of techno-economic models, such as the one presented in this project, contains information about all feasible solutions unobtainable through classic optimization. A structured method, capable of mapping the entire near-optimal feasible space of techno-economic energy models have been developed in this project, building on the principles of MGA approaches. The developed method uses a two-step approach, where the hull containing all near-optimal solutions are initially found through strategic alterations of the objective function, combined with the introduction of a constraint confining the near-optimal feasible space. In the second step of the developed method, the hull containing all near-optimal solutions is sampled, thereby extracting a dataset representing all near-optimal solutions.

Initially, the mathematical formulation of the techno-economic energy model was investigated, revealing that the near-optimal decision space is convex. As the number of decision variables in the presented model is more than half a million a method of reducing the number of variables was investigated. Neglecting hourly dispatch of energy from the individual plants, reduce the number of decision variables included drastically. Neglecting hourly dispatch introduces no significant loss in information as this project focuses on the distribution of technology capacities rather than plant operation. Further reduction of the number of decision variables can be achieved by grouping the individual variables into new variables through summation. The selection of grouped variables to include in the studies allows the modeler to focus the studies on areas of interest. An effect of grouping variables is that the distribution of solutions across the decision space is no longer even. This effect was investigated and results revealed that sample points located close to the center of the decision space of grouped variables, cover over a much larger number of solutions than sample points located on the perimeter of the decision space. The result of this is that the distributions of technology capacity squeeze together, indicating that solutions located near the center of the decision space are much more likely than extreme solutions.

Using the developed MGA method on a techno-economic model of the European electricity grid several experiments were conducted. The experiments successfully extracted data about the techno-economic models' near-optimal feasible space, revealing large solution flexibility within the near-optimal feasible space at a small variance in total system cost. These results emphasize the importance of addressing structural uncertainty in the model through the use of MGA algorithms. From the results, information about technology capacity ranges was extracted, revealing must-have technologies as CO_2 emissions are decreased. Using the Gini coefficient as a measure of equality in the location of energy-producing technology compared to energy demand, revealed that as CO_2 emissions are decreased larger inequality arises.

The computing time required by the developed method was estimated, and practical limits identified. Especially the number of variables included in the decision space investigated by the method was found to have a large effect on the required computing time. A successful study including seven variables in the decision space was conducted however, the practical limit to the number of decision variables has been estimated to be ten. Within these limitations, the model has proved very useful, and through intelligent selection of decision variables a wide range of studies can be performed. Future studies could include the investigation of models implementing storage technologies or models implementing larger selections of renewable energy technologies. Overall, the developed method has been found to provide much greater insights about the near-optimal solution space of techno-economic models, compared to previous MGA methods, and the use of the developed MGA method in future work will contribute to a greater understanding of energy systems and the available possibilities when conducting energy system synthesis.

Bibliography

[1] European Environment Agency. Annual european union greenhouse gas inventory 1990 – 2015 and inventory report 2017. 2017.

[2] K. Zhu, M. Victoria, T. Brown, G.B. Andresen, and M. Greiner. Impact of co2 prices on the design of a highly decarbonised coupled electricity and heating system in europe. *Applied Energy*, 236:622 – 634, 2019.

[3] Joseph F. DeCarolis. Using modeling to generate alternatives (mga) to expand our thinking on energy futures. *MEnergy Economics*, 33, 2010.

[4] Lewis D. Hopkins E. Downey Brill Jr., Shoou-Yuh Changt. Modeling to generate alternatives: The hsj approach and an illustration using aproblem in land use planning. *Management Science*, 28, 3 1982.

[5] J.F. DeCarolis, S. Babaee, B. Li, and S. Kanungo. Modelling to generate alternatives with an energy system optimization model. *Environmental Modelling and Software*, 79, 2016.

[6] James Price and Ilkka Keppo. Modelling to generate alternatives: A technique to explore uncertainty in energy-environment-economy models. *Applied Energy*, 195:356 – 369, 2017.

[7] Philip B. Berntsen and Evelina Trutnevyte. Ensuring diversity of national energy scenarios: Bottom-up energy system model with modeling to generate alternatives. *Energy*, 126:886 – 898, 2017.

[8] Yavuz Gunalay and Julian Scott Yeomans. Simulation-optimization techniques for modelling to generate alternatives in waste management planning. *Applied Operational Research*, 2011.

[9] Philip Voll, Mark Jennings, Maike Hennen, Nilay Shah, and André Bardow. The optimum is not enough: A near-optimal solution paradigm for energy systems synthesis. *Energy*, 82:446 – 456, 2015.

[10] Fabian Neumann and Tom Brown. The near-optimal feasible space of a renewable power system model, 2019.

[11] D.P. Schlachtberger, T. Brown, S. Schramm, and M. Greiner. The benefits of cooperation in a highly renewable european electricity network. *Energy*, 134:469 – 481, 2017.

[12] Tom Brown, Jonas Horsch, and David Schlachtberger. Python for power system analysis (pypsa) version 0.14.1. `https://doi.org/10.5281/zenodo.3233682`, may 2019.

[13] Lieven Vandenberghe Stephen Boyd. *Convex Optimization*. Cambridge University Press, 2004.

[14] J. S. Arora. *Introduction to Optimum Design*. Cambridge University Press, 4th edition, 2004.

[15] Gurobi. Optimization software. `https://www.gurobi.com`.

[16] Daniel H. Loughlin, S. Ranji Ranjithan, E. Downey Brill JR., and John W. Baugh JR. Genetic algorithm approaches for addressing unmodeled objectives in optimization problems. *Engineering Optimization*, 2001.

[17] Donald B. Rubin. The bayesian bootstrap. *The Annals of Statistics*, 9(1):130–134, 1981.

[18] Suranjana Saha, Shrinivas Moorthi, Xingren Wu, Jiande Wang, Sudhir Nadiga, Patrick Tripp, David Behringer, Yu-Tai Hou, Hui-ya Chuang, Mark Iredell, Michael Ek, Jesse Meng, Rongqian Yang, Malaquias Pena Mendez, Huug van den Dool, Qin Zhang, Wanqiu Wang, Mingyue Chen, and Emily Becker. The ncep climate forecast system version 2. *Journal of Climate*, 27(6):2185–2208, 2014.

[19] Gorm B. Andresen, Anders A. Sondergaard, and Martin Greiner. Validation of danish wind time series from a new global renewable energy atlas for energy system analysis. *Energy*, 93:1074 – 1088, 2015.

[20] Andreas Schröder, Friedrich Kunz, Jan Meiss, Roman Mendelevitch, and Christian von Hirschhausen. Current and prospective costs of electricity generation until 2050. DIW Data Documentation 68, Berlin, 2013.

[21] S. Hagspiel, C. Jagemann, D. Lindenberger, T. Brown, S. Cherevatskiy, and E. Troster. Cost-optimal power system extension under flow-based market coupling. *Energy*, 66:654 – 666, 2014.

[22] European Network of Transmission System Operators. Country-specific hourly load data. `https://www.entsoe.eu/data/data-portal/`, October 2019.

[23] https://docs.python.org/3.4/library/multiprocessing.html?highlight=process. Python multiprocessing. `https://docs.python.org/3.4/library/multiprocessing.html?highlight=process`.

[24] PRIME – Programming Rig for Modern Engineering. `https://eng.au.dk/forskning/laboratorier-og-faciliteter/prime/`. Accessed: 2019-12-18.

[25] Morten Grud Rasmussen, Gorm Bruun Andresen, Dominik Heide, and Martin Greiner. Optimal combination of storage and balancing in a 100% renewable european power system. *Proceedings of the 10th International Workshop on Large-scale Integration of Wind Power Into Power Systems As Well As on Transmission Networks for Offshore Wind Power Plants*, pages 682–684, 2011.

[26] Marta Victoria, Kun Zhu, Tom Brown, Gorm B. Andresen, and Martin Greiner. The role of storage technologies throughout the decarbonisation of the sector-coupled european energy system. *Energy Conversion and Management*, 201:111977, 2019.

[27] Gorm B. Andresen, Rolando A. Rodriguez, Sarah Becker, and Martin Greiner. The potential for arbitrage of wind and solar surplus power in denmark. *Energy*, 76:49 – 58, 2014.

[28] Arturs Purvins, Alyona Zubaryeva, Maria Llorente, Evangelos Tzimas, and Arnaud Mercier. Challenges and options for a large wind power uptake by the european electricity system. *Applied Energy*, 88(5):1461 – 1469, 2011.

[29] Marta Victoria, Kun Zhu, Tom Brown, Gorm B. Andresen, and Martin Greiner. The role of photovoltaics in a sustainable european energy system under variable co2 emissions targets, transmission capacities, and costs assumptions. *Progress in Photovoltaics: Research and Applications*.

[30] Google Cloud - cloud computing platform. `https://cloud.google.com/compute/all-pricing`. Accessed: 2019-12-16.

[31] QHull - python module. `http://www.qhull.org/`. Accessed: 2019-12-16.

Part II

Appendix

A Code

All written code for this project can be found on Github at:

`https://github.com/TimToernes/MGA-PyPSA.`

The script capable of determining the convex hull containing all near optimal solutions to a techno-economic model is located in "./PyPSA_project/hull_generation_parllel.py". The script can be controlled through setup files located in "./PyPSA_project/setup_files/".

Pseudo code

A brief pseudo code illustrating the working principles of the "./PyPSA_project/hull_generation_parllel.py" script:

```
Solve network subject to regular constraints and with original objective function
Add MGA constraint
while ε > tol
    If first loop
        directions = max and min all variables
    Else
        directions = normals to hull faces
    for direction in directions
        objective function = direction[i] * variable[i]
        point on convex hull += solve problem subject to objective function
    hull = ConvexHull ( points on convex hull)
    epsilon = new hull volume - old hull volume / hull volume
Evenly distribute points in hull
Plot histogram using evenly distributed points.
```

B Plots for section 4.2

This section contains the remaining correlation plots from section 4.2. The plots show the correlations of the used decision variables, for all four studies performed. Furthermore, a plot of the individual decision variable distributions is included.

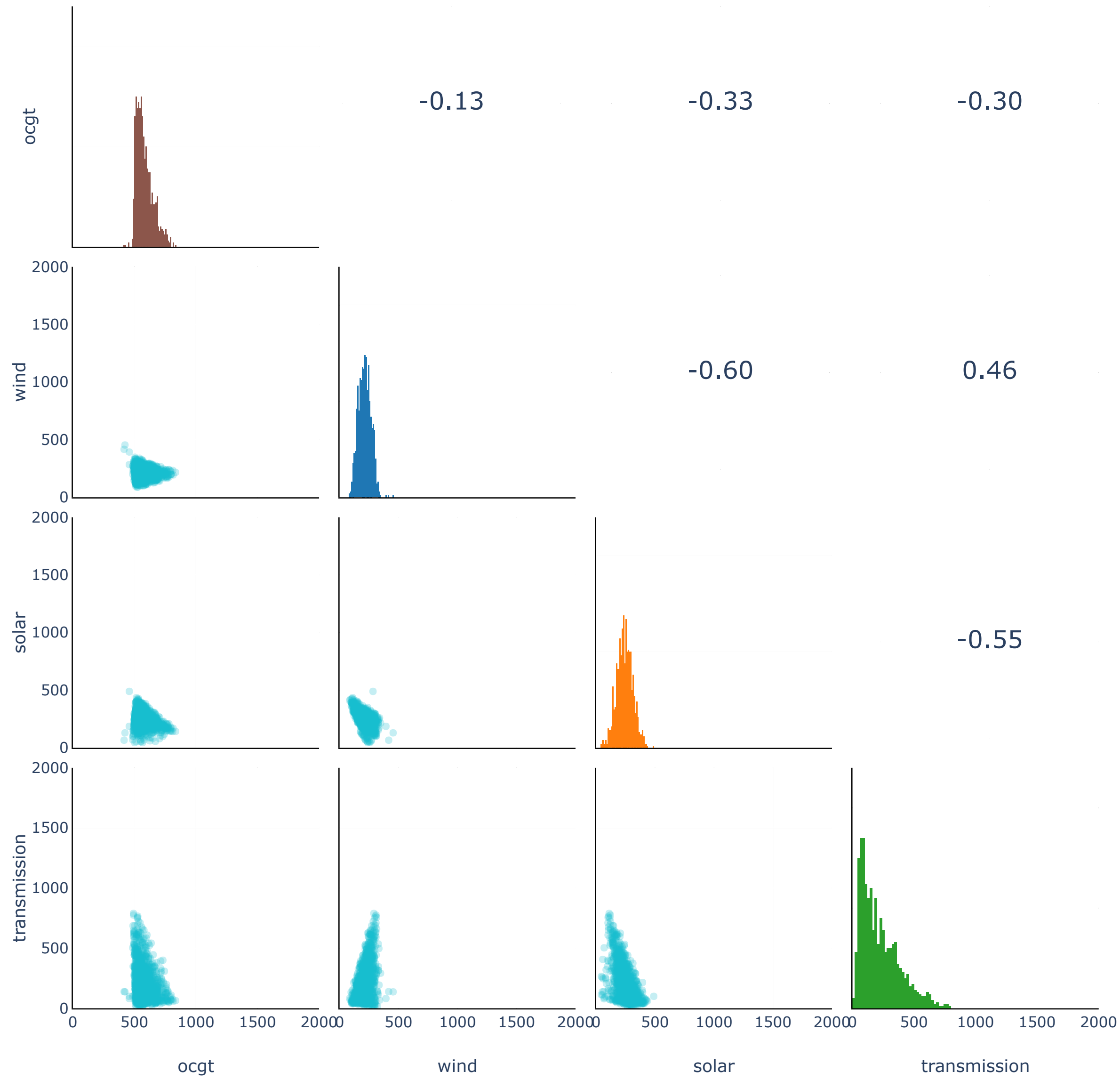


Figure B.1: Variable correlations of the MGA study with a CO2 emission reduction of 0%. Distributions of the four individual technologies are plotted on the diagonal.

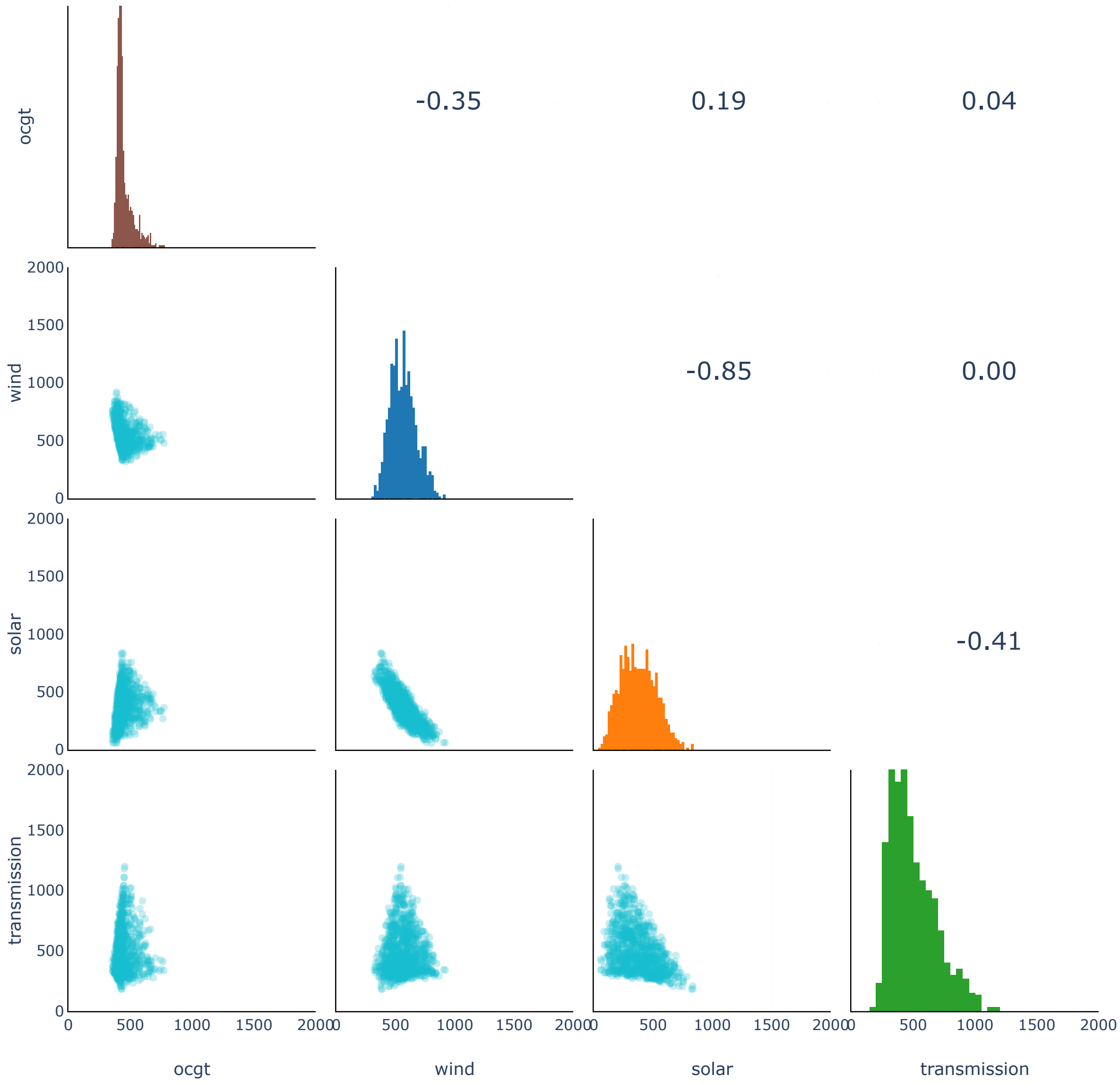


Figure B.2: Variable correlations of the MGA study with a CO2 emission reduction of 50%. Distributions of the four individual technologies are plotted on the diagonal.

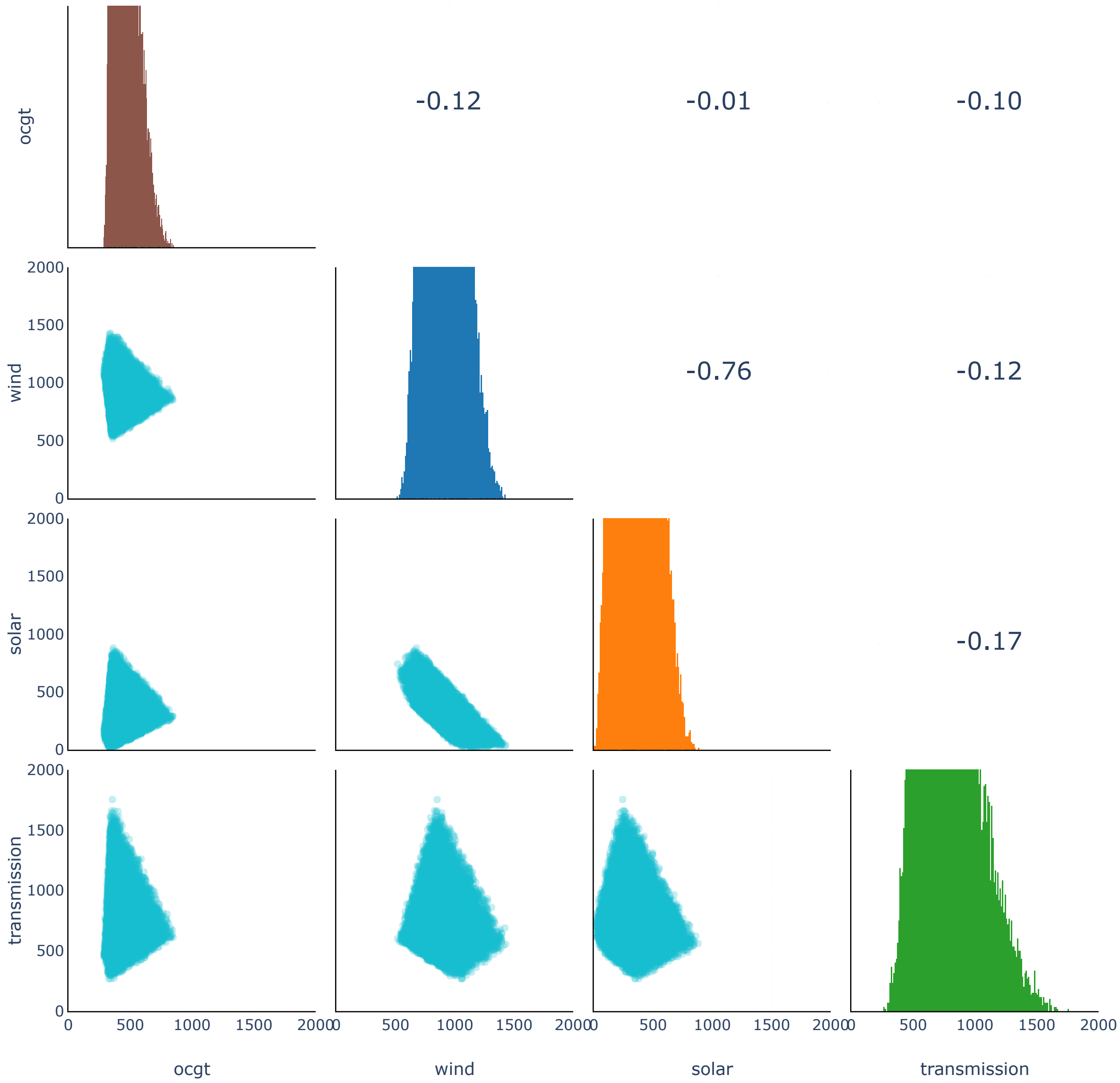


Figure B.3: Variable correlations of the MGA study with a CO2 emission reduction of 80%. Distributions of the four individual technologies are plotted on the diagonal.

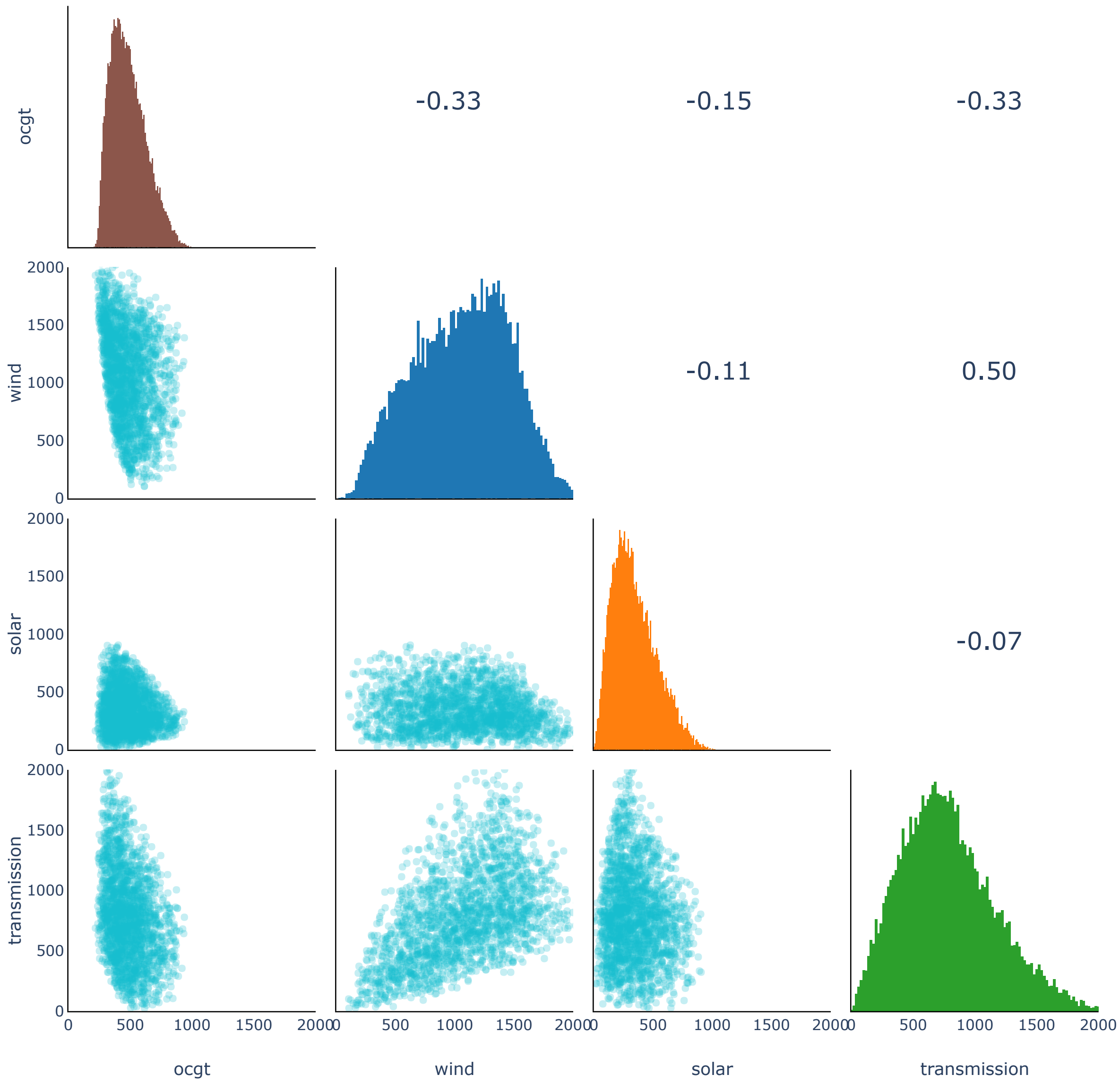


Figure B.4: Variable correlations of the MGA study with a CO2 emission reduction of 95%. Distributions of the four individual technologies are plotted on the diagonal.

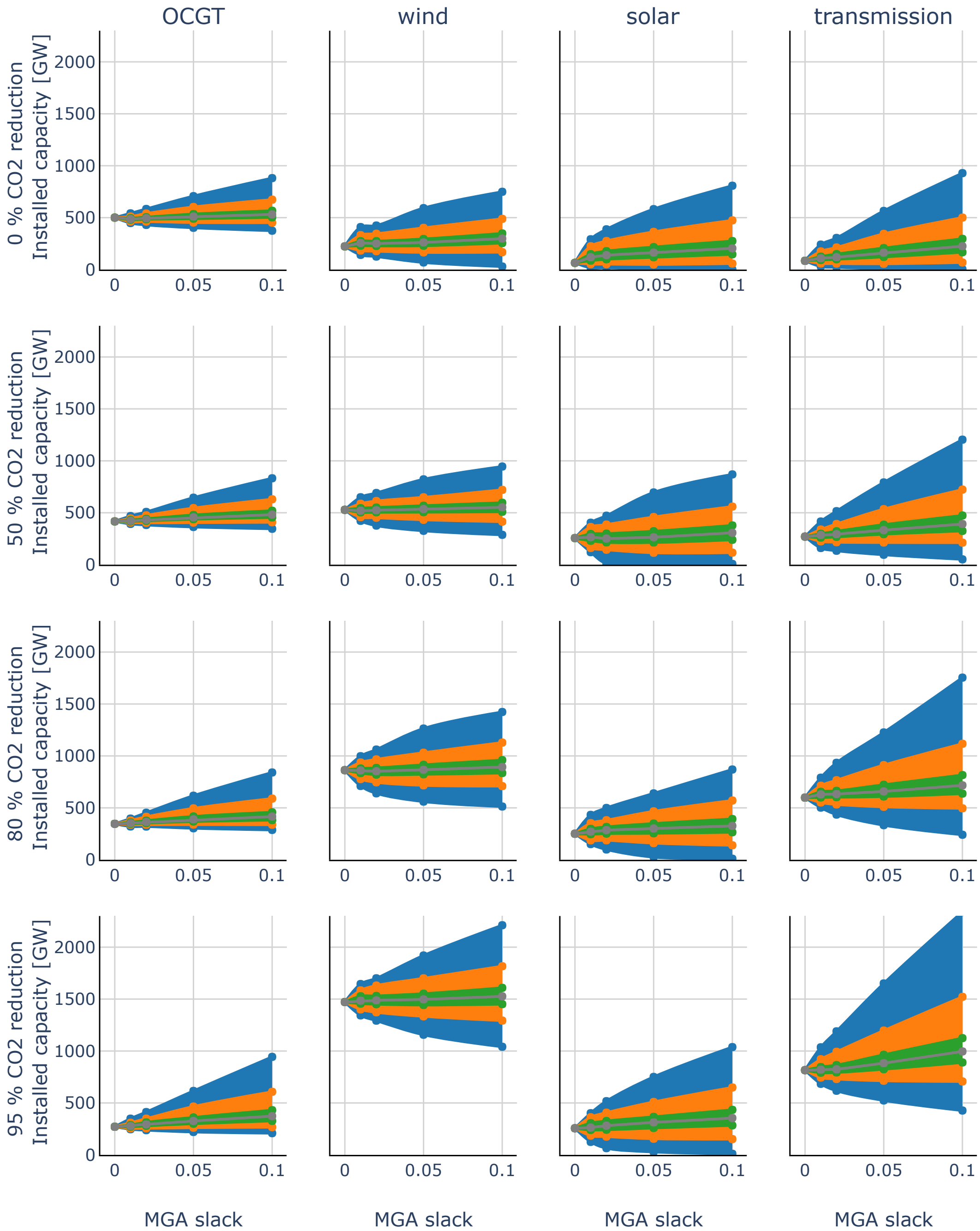


Figure B.5: Plot matrix showing the distribution of capacities as the MGA slack is changed. The columns of plots each represents their individual technology and the rows corresponds to a CO2 emission constraint. All data used is from the MGA study using four decision variables.