Design of Bio-Based Supply Chains

Master Thesis
Abstract

The biomass feedstock is geographically dispersed and subject to variability. Therefore, strategic supply chain design can significantly influence the economic viability of bio-derived products. Preprocessing through decentralized fast pyrolysis facilities followed by centralized upgrading has recently gained a lot of attention because it limits costly biomass transportation. In this work, a modeling framework that captures the main characteristics of bio-based supply chains has been developed to investigate cost-optimal system configurations. The model provides a valuable tool to determine the optimal area that the supply chain must cover. Furthermore, optimisation results indicate that through the geographical concentration of biomass sources the profitability of the supply chain is increased enormously, and that preprocessing followed by upgrading is no longer by definition the preferred processing strategy. However, further analysis revealed that under uncertain scenarios the clustering of biomass sources has a slight adverse effect on the robustness of the supply chain.
Contents

List of Figures vii
List of Tables ix

1 Introduction 1

1.1 Key Aspects of Bio-based Supply Chains . . . . . . . . . . . . . . . . . . . . . . 3
1.1.1 Biomass Supply and Transportation . . . . . . . . . . . . . . . . . . . . . . 3
1.1.2 Biomass Storage . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
1.1.3 Conversion Technologies . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
1.2 Literature Review . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
1.3 Research Objectives and Approach . . . . . . . . . . . . . . . . . . . . . . . . . 6

2 Model Development 7

2.1 Problem Statement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
2.2 Mathematical Model Formulation . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
2.2.1 Harvesting Sites . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
2.2.2 Integrated Processing Facilities . . . . . . . . . . . . . . . . . . . . . . . . . 9
2.2.3 Preprocessing Facilities . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
2.2.4 Upgrading Facilities . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
2.2.5 Storage Facilities . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
2.2.6 Demand Sites . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
2.2.7 Objective: Minimize Annual Total Cost . . . . . . . . . . . . . . . . . . . . . 16
2.3 Discretization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
2.4 Piecewise Linearization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
2.5 GAMS Java API . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21

3 Case Studies 23

3.1 Case Independent Model Input . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
3.2 Case Study I: Geographical Topologies . . . . . . . . . . . . . . . . . . . . . . . 27
3.3 Case Study II: Feedstock Allocation Freedom . . . . . . . . . . . . . . . . . . . . 29
# List of Figures

1.1 Schematic representation of the bio-based supply chain. ........................................ 3

2.1 Bio-based supply chain superstructure. .............................................................. 7

2.2 Average distance to the center for a $3 \times 3$ grid. ........................................... 18

2.3 Average distance to the center for a varying number of grid units. ......................... 19

2.4 An illustration of the MILP regression model. .................................................... 20

3.1 Integrated processing facility capital investment scaling equation and its linear approximation. .......................................................... 26

3.2 The center-point and corner-point spatial distributions with increasing rural boundary region. .......................................................... 28

4.1 Schematics of the optimal system configurations. .............................................. 34

4.2 Cost fractions for all scenarios. ................................................................. 36

4.3 Total annual cost for both processing strategies versus the biomass feedstock allocation freedom. .................................................. 37

4.4 Optimal system configuration for two different threshold values. ....................... 38

4.5 Cost break down associated with the optimal system configuration as a function of the threshold value $\Phi$. ........................................... 38

4.6 Optimal system configurations at extreme threshold value. ................................ 39

4.7 The number of preprocessing facilities as a function of the threshold value $\Phi$. ... 40

4.8 Cost break down associated with the optimal system configuration as a function of the threshold value $\Phi$. ........................................... 40

4.9 Total annual cost for both processing strategies versus the biomass feedstock allocation freedom for a higher yielding crop. ......................... 41

4.10 The selected optimal system configurations. ..................................................... 42

4.11 The variable cost components for each simulated biomass supply scenario with increasing standard deviation. ....................................... 43

4.12 The average and standard deviation of the total variable cost components as a function of the fluctuation strength. .......................... 44

A.1 Illustrative example of a $2 \times 3$ grid. ............................................................... 62
B.1 Preprocessing facility capital investment scaling equation and its linear approximation. .......................................................... 63

B.2 Upgrading facility capital investment scaling equation and its linear approximation. 64

E.1 The variable cost components for each simulated biomass supply scenario with increasing standard deviation corresponding to the preprocessing/upgrading configuration at a threshold value of $\Phi = 0.1$. ............................................................ 70

E.2 The variable cost components for each simulated biomass supply scenario with increasing standard deviation corresponding to the preprocessing/upgrading configuration at a threshold value of $\Phi = 0.4$. ............................................................ 70

E.3 The variable cost components for each simulated biomass supply scenario with increasing standard deviation corresponding to the integrated processing configuration at a threshold value of $\Phi = 0.1$. ............................................................ 71
# List of Tables

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Biomass storage facility options and their characteristics.</td>
<td>4</td>
</tr>
<tr>
<td>3.1</td>
<td>Agricultural land coverage and population density for three region types.</td>
<td>23</td>
</tr>
<tr>
<td>3.2</td>
<td>Capacities and capital investments of three reference plants.</td>
<td>24</td>
</tr>
<tr>
<td>3.3</td>
<td>Mean absolute error and coefficient of determination for a varying number of break-points for each processing facility.</td>
<td>25</td>
</tr>
<tr>
<td>3.4</td>
<td>Operating and maintenance cost factors.</td>
<td>26</td>
</tr>
<tr>
<td>3.5</td>
<td>Storage investment cost for bio-oil and -product.</td>
<td>27</td>
</tr>
<tr>
<td>3.6</td>
<td>Distance fixed and variable transportation cost for biomass, -oil and -product.</td>
<td>27</td>
</tr>
<tr>
<td>4.1</td>
<td>Model statistics and solution report corresponding to the largest (9 × 9) model.</td>
<td>33</td>
</tr>
<tr>
<td>4.2</td>
<td>Performance measures for the two topologies and both processing strategies.</td>
<td>35</td>
</tr>
<tr>
<td>4.3</td>
<td>Performance measures for all of the system configuration under investigation.</td>
<td>44</td>
</tr>
<tr>
<td>B.1</td>
<td>Regression constants in case of one break-point.</td>
<td>63</td>
</tr>
<tr>
<td>C.1</td>
<td>Processing facility location and capacity for the 5 × 5 grid for both topologies and processing pathways.</td>
<td>65</td>
</tr>
<tr>
<td>C.2</td>
<td>Processing facility location and capacity for the 7 × 7 grid for both topologies and processing pathways.</td>
<td>65</td>
</tr>
<tr>
<td>C.3</td>
<td>Processing facility location and capacity for the 9 × 9 grid for both topologies and processing pathways.</td>
<td>66</td>
</tr>
<tr>
<td>D.1</td>
<td>Processing facility location and capacity corresponding to the system configurations in Figure ??</td>
<td>67</td>
</tr>
<tr>
<td>D.2</td>
<td>Processing facility location and capacity corresponding to the system configurations in Figure 4.6.</td>
<td>67</td>
</tr>
<tr>
<td>E.1</td>
<td>Processing facility location and capacity corresponding to the selected system configurations.</td>
<td>69</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In the past hundred years, fossil resources have enabled large-scale industrial development and have become the lifeblood of all industrialized countries. These natural resources have been generated over a period of millions of years. Biomass has been produced via photosynthesis, utilizing solar energy and capturing carbon dioxide (CO$_2$) from the atmosphere. Via subsequent geochemical processes, driven by pressure and heat, the remains of these plants and organisms have been converted into natural gas, petroleum and coal. Today, this pocket of solar energy feeds our society in terms of fuels, chemicals and materials. Due to its seemingly unlimited availability, our industrial value chains have been fully adapted towards fossil resources as raw materials for the production of carbon-based products.

However, because of the simple fact that the exploitation of these fossil resources is proceeding much faster than its regeneration through the natural carbon cycle, they will eventually be depleted. Hundreds of scientific articles have been published regarding peak oil (the time at which the maximum rate of crude oil extraction is reached) and the depletion of other fossil resources. Some studies [1] predict complete physical or economic depletion to occur before the end of this century. Other, more optimistic studies [2], do not believe that peak oil will be reached any time soon. Although any such prediction is highly uncertain and therefore remains debatable, it is widely agreed among climate experts that anthropogenic carbon conversion significantly contributes to the increase of CO$_2$ in the atmosphere and, indirectly, to the average temperature rise on earth. The only way to stop this trend is by bounding the cumulative anthropogenic CO$_2$ emissions to a trillion tonnes of carbon (3.67 trillion tonnes of CO$_2$), ultimately equilibrating the natural carbon cycle [3]. Achieving this equilibrium is not possible without a switch from fossil to renewable resources.

Since transportation fuels account for approximately 95% of the total fossil resource consumption, biomass-derived liquid transportation fuels have been proposed as part of the solution to climate change and our heavy dependence on fossil resources. Partially because the biomass feedstock is renewable, but also because it can reduce greenhouse gas (GHG) emissions, since the CO$_2$ that is captured when the feedstock crops are cultivated balances the CO$_2$ emitted when products are (eventually) incinerated.

Consequently, in the past years many countries have set targets and provide supports to accelerate the transition towards biomass-derived liquid transportation fuels. For example, the E.U. Biofuels Directive aims to replace 10% of petroleum based fuels with biofuels by 2020. Furthermore, in the United States the Renewable Fuels Standard (RFS) targets an annual biofuel production of 36 billion gallons by 2022 [4]. Currently, the most common biofuel is bioethanol, which is made primarily from corn starch and sugar cane. However, some studies criticize bioethanol production from first generation feedstock as one of the factors responsible for the 2007-2008 global food crisis [5]. In order to avoid negative effects on food prices, the
RFS further specifies that 16 out of 36 billion gallons of biofuels produced in 2022 should be made from cellulosic biomass. This second group of terrestrial biomass feedstock consists of non-starch and non-edible materials and thereby avoids the potential impact on the food market. It can be obtained from numerous sources, for example forest residues, agricultural residues, energy crops and municipal solid waste (MSW). Typically, forest residues are left over plant or wood parts after for example logging operations or forest management. Agricultural residues are plant parts that are left behind after harvest, such as corn stover. Energy crops are typically fast-growing plants that are specifically grown for energy uses (e.g. switchgrass and Miscanthus). Yet, the current annual production of cellulosic biofuel is less than 1 billion gallons worldwide. Therefore, it is foreseeable that the cellulosic bioenergy industry will be experiencing a large expansion in the coming decades.

Although the largest part of the total fossil resource consumption is used to supply us with energy, it is also one of the main sources for our chemical industry. Besides biomass, there are numerous candidates to provide us of a renewable energy supply, however fewer alternatives are at hand to serve as a feedstock for industrial organic chemicals. At the beginning of the 20th century many industrial materials, such as solvents, dyes and synthetic fibers were made from agricultural crops and trees. By the late 1960s many of these bio-based chemical products had been displaced by petroleum derivatives. Due to the aforementioned reasons, it seems inevitable that at some point in time this switch needs to be reversed. In addition to the targets set for biofuels, similar goals have been set for biomass-derived chemicals. For example, the United States Department of Energy aims to replace 25% of the industrial organic chemicals with bio-based chemicals by 2025. It is very probable that ultimately the biomass feedstock will be processed at integrated biorefineries, that in essence parallel our current petroleum refineries. First, high-value chemicals already present in biomass, such as flavoring agents and fragrances, will be extracted. Once these relatively valuable chemicals are separated, the subsequent steps will focus on processing the remaining plant polysaccharide and lignin into feedstock for bio-based materials and biofuels. Another key aspect of a biorefinery is the imbalance between the percentage of the feedstock used for energy needs and other products. Using our conventional petroleum refineries as an illustrative example, only 5% of the total crude oil consumption is used for the production of chemical products other than fuels. Most visions do not expect this ratio to change significantly for biorefineries.[6] As of today, numerous challenges regarding the large-scale commercialization of cellulosic bio-based products remain. One of the major challenges is the design of efficient supply chains linking the biomass feedstock production with end-use, which is the main subject of this work. Due to the large imbalance in the magnitude of the processing routes in a biorefinery, the principal focus will be on the production of biofuels. It is expected that the conversion of biomass into other bioproducts will only effect the overall profitability, but will not significantly alter other key aspects of the bio-based supply chain.

In the remainder of this chapter the key components of the bio-based supply chain (BSC) and existing contributions to supply chain design/optimization are elucidated. Ultimately, the problem statement for this project is presented.
1.1 Key Aspects of Bio-based Supply Chains

The bio-based supply chain links the biomass supply with the eventual bioproduct end-use. It essentially consists of biomass suppliers, storage and processing facilities, transportation and end-use locations, as schematically depicted in Figure 1.1.

![Figure 1.1: Schematic representation of the bio-based supply chain.](image)

1.1.1 Biomass Supply and Transportation

At first sight, the bio-based supply chain depicted in Figure 1.1 may seem similar to its petroleum counterpart. While the components are similar in function, they may differ greatly in property. The most significant difference is at the biomass acquisition stage. While petroleum emerges from point sources like drill shafts and oil rigs, guarantees all-year-round supply and can be moved over long distances requiring minimal transportation cost, biomass is almost the exact opposite.

Unlike crude oil, the raw biomass feedstock is distributed over a large area. Also, biomass is low-energy density solid material, while petroleum is a high-energy density liquid. It is due to the combination of these factors that transportation costs account for almost 90% of the total estimated biofuel production cost [7]. Another critical challenge in the operation of this supply chain is that, contrary to crude oil, the raw biomass feedstock is subject to seasonality and temporal variability. As mentioned, the main contributors of the second group of biomass feedstock are forest residues, agricultural residues and energy crops. These plant matters need to be planted, cultivated and harvested. Agricultural residues are usually collected after the harvest of the crops, which are usually harvested annually during a two to three month period. Forest residues however, are grown over multiple years, making them less seasonal. While the raw biomass supply may fluctuate over time, the demand for transportation fuels is usually more steady. Various solutions to deal with the seasonal nature of biomass have been proposed. For example, a more continuous resource availability may be achieved by using multiple biomass types, each with their characteristic harvesting period. Alternatively, through the cultivation of short rotation coppice crops that can be harvested all-year-round, seasonal effects can be kept to a minimum [8]. Nevertheless, since the harvested yield is influenced by many factors, e.g. unpredictable weather circumstances, temporal fluctuations are to some extent inevitable. So far, the impacts of these inherent system dynamics have been neglected.

As there are many differences between petroleum and bio-based supply chains, similarities exist as well. Besides the temporal variabilities, concerning mainly the available quantity of feedstock, the exact chemical composition (i.e. the quality) of the feedstock may vary from batch to batch as well. Petroleum crude is known to exhibit quality variations which mainly depend on its origin. Especially the density and the sulfur content are important measures regarding the quality of a crude, effecting the efficiency and eventual yield. There is no doubt that the biomass feedstock exhibits quality variations as well. Cellulosic, hemicellulosic and lignin content have been reported as being important quality measures [9].
1.1.2 Biomass Storage

Another important characteristic that distinguishes biomass from the petroleum feedstock is material deterioration, the extent of which depends largely on the type of storage facility. The most expensive option is a closed warehouse with biomass drying capability which helps to avoid quality degradation. In this case the amount of material loss can be assumed negligible. Another option is a covered storage facility consisting of a pole-frame structure with a metal roof but without any infrastructure for biomass drying. This way, the effect of biomass degradation can be reduced to 0.5% material loss per month. Alternatively, ambient storage can be employed, meaning that the biomass is covered solely with a plastic film. The downside of this option is that it leads to significant degradation of the biomass, estimated at 1% per month. The various options for biomass storage and associated cost estimations are summarized in Table 1.1 below [10]. As can be seen, there is clearly a tradeoff between storage cost and biomass degradation.

<table>
<thead>
<tr>
<th>Storage facility</th>
<th>Covered with drying</th>
<th>Covered without drying</th>
<th>Ambient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass degradation (% per month)</td>
<td>~0%</td>
<td>0.5%</td>
<td>1%</td>
</tr>
<tr>
<td>Storage investment cost ($ / m²)</td>
<td>222</td>
<td>110</td>
<td>22</td>
</tr>
<tr>
<td>Storage O&amp;M cost (% investment / year)</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1.1: Biomass storage facility options and their characteristics.

In addition to storage of raw biomass, pretreatment processes may also be adopted to improve feedstock stability. This however, is outside the scope of this work and therefore not considered here.

1.1.3 Conversion Technologies

The conversion of biomass into liquid transportation fuels is another main component of the bio-based supply chain. Conversion technologies can roughly be divided into two categories, namely thermochemical and biochemical. One example of a biochemical pathway is the fermentation of corn to form ethanol. However, considering the second group of biomass feedstock, there are two major thermochemical conversion pathways towards which most studies are dedicated: gasification followed by Fischer-Tropsch (FT) synthesis and fast pyrolysis followed bio-oil gasification and FT synthesis. A brief explanation of both conversion technologies is given below.

Biomass Gasification Followed by Fischer-Tropsch Synthesis

In the gasification process, the cellulosic biomass is fed into a gasifier, which thermally decomposes the material at temperatures ranging from 700 to 1300 °C with limited amounts or complete absence of oxygen, to yield a mixture of carbon monoxide, hydrogen and some carbon dioxide. This gaseous mixture, also referred to as syngas, may contain some contaminants and therefore a gas cleanup process is usually required. The syngas can then be utilized by Fischer-Tropsch synthesis to produce, via a series of chemical reactions, hydrocarbon liquid transportation fuels such as diesel and gasoline.
Biomass Fast Pyrolysis Followed by Bio-oil Gasification and FT Synthesis

Fast pyrolysis also thermally decomposes the cellulosic biomass in the absence of oxygen. However, unlike gasification, fast pyrolysis is typically carried out at much lower temperatures, usually between 400 and 600 °C, thereby converting the biomass into gaseous, liquid and solid products. Fast pyrolysis can produce approximately 70% (by weight) bio-oil. The other 30% is split between non-condensable gasses (mainly hydrogen and carbon monoxide) and bio-char [11]. The non-condensable gasses can be combusted to provide heat for the process, while the bio-char can be used as a soil amendment in order to prevent exhaustion of the agricultural lands. The density of bio-oil is three to five times larger compared to the density of raw biomass [12]. Nevertheless, bio-oil needs to be upgraded before it can be used as transportation fuel. This is done via bio-oil gasification and FT synthesis.

1.2 Literature Review

There is a rich literature regarding the modelling and optimization of the bio-based supply chain. General reviews concerning this subject and the remaining challenges were presented by Yue et al. [13] and Sharma et al. [14]. Below, the papers most relevant to this work addressed are reviewed.

Decisions concerning the design of a bio-based supply chain, such as the selection of biomass suppliers, storage and (pre)conversion facility location, product destinations and transportation links have profound impact on the profitability and environmental impacts of the overall supply chain. Amongst the various candidate conversion technologies, fast pyrolysis followed by hydroprocessing has recently moved to the forefront of biofuel research [15]. As stated before, due to the relatively low energy density of biomass and its spatial dispersion, transportation comprises a large slice of the cost pie. Therefore, it has been suggested that to reduce transportation costs, the raw biomass is converted into bio-oil via fast pyrolysis near harvest sites. Subsequently, the relatively high-energy density bio-oil is transported to a centralized upgrading plant that benefits from economies of scale. The benefit of decentralized fast pyrolysis plants is that it concentrates the energy in biomass into a smaller volume, the bio-oil, which is cheaper to transport [16], [17].

The majority of the models reported in scientific literature are Mixed-Integer Linear Programming Models (MILP). Kim et al. [18] proposed a MILP model for the optimal network design of a bio-based supply chain. The model is used to make decisions regarding (1) the optimal number, locations and sizes of various types of processing facilities, (2) the amounts of biomass, intermediate products, and final products to be transported between the selected locations, with the objective of maximizing the overall profit. The proposed model was tested for designing both distributed and centralized processing network systems for a fairly large region in the Southeastern part of the United States. The distributed design was identified as being the most profitable option. Zamboni et al. [19] presented a spatially explicit MILP model for the strategic design of a biofuel production system, simultaneously minimizing the overall operating costs as well as environmental issues in terms of greenhouse gas (GHG) emissions. Northern Italy was chosen as an illustrative example. To account for other main characteristics of the supply chain, such as the temporal supply variations and degradation of biomass, You et al. [20] proposed a multi-period MILP model. With this higher resolution on the temporal scale, this model can be used to make operational decisions, e.g. what level of inventory to keep in the storage units in order to fulfill the continuous demand of biofuels with a variable supply of biomass. To illustrate its application, the model was used to solve the supply chain design problem on a county level for the state of Iowa in the United States. The results suggest a largely distributed...
1.3 Research Objectives and Approach

Despite the various attempts that have been made to model the strategic design of bio-based supply chains, it is often merely a single geographical area for which the models are tested. These areas vary significantly in size and regional topology. Therefore, there is a lack of general insights concerning the impact of the geographical area of study on the overall profitability of the supply chain. In addition, biomass supply and product demand data is usually obtained through consulting statistical databases. Since these numbers are based on recently collected information, this data is only representative under current circumstances. However, with the development of dedicated energy crops this data is likely to change radically in the near future. Added to that, since transportation comprises such a large portion of the total production cost, one may wonder whether it is wise to hold on to the current spatial distribution of biomass; the alternative being the spatial concentration of feedstock sources. Given that these changes occur somewhere in the near future, it is necessary to investigate whether a potential for decentralization remains.

Furthermore, the models described above all make the crucial assumption that all parameters are known and fixed, i.e. they are completely deterministic. However, uncertainties are ubiquitous, of various types, and present at all stages in the bio-based supply chain. Although sensitivity analysis can be conducted to evaluate the effects of changes in different parameters on the outputs and objective of the model, deterministic models fail to capture the complete dynamics of the supply chain. Thus, it is necessary to apply a stochastic approach to deal with the various uncertainties and examine the robustness of the proposed deterministic solutions.

Therefore, the focus of this study is to investigate (1) the influence of the underlying geographical area on the performance of the supply chain, (2) the possible benefit that can be gained through concentration of feedstock sources and (3) the robustness of solutions obtained under deterministic conditions.

Initially, a multi-period spatially explicit mathematical programming model is developed to determine cost-efficient bio-based supply chain network structures. In order to gain general insights regarding the supply chain design and its components, this model is used to determine the optimal design given various generic topologies. Thereafter, a slight adjustment is made in the model such that it is equipped with the freedom to allocate biomass sources. In order to quantify the robustness of the obtained solutions, the optimal system configurations resulting from the deterministic baseline model are subjected to simulated quantitative biomass supply fluctuations, being one of the important sources of variability. Despite having noted the presence of qualitative variations, their possible effects are beyond the scope of this work.

The next chapter is dedicated to the development of the modeling framework. In Chapter 3 descriptions of the various case studies are given, the results of which are presented and discussed in Chapter 4. Conclusions drawn from this work and recommendations for further research are given in Chapter 5 and 6 respectively.
Chapter 2

Model Development

In this chapter the developed framework is presented. Starting with the problem statement in Section 2.1, the mathematical model formulation is presented in Section 2.2. Some model related aspects that require a more detailed explanation are postponed to the subsequent sections.

2.1 Problem Statement

The bio-based supply chain consists of five components: (i) harvesting sites, (ii) integrated processing facilities, (iii) preprocessing facilities, (iv) upgrading facilities and (v) demand sites. In Figure 2.1 the bio-based supply chain superstructure is depicted. At the beginning (left-hand side), biomass is collected at harvesting sites. From here, it is transported to an integrated or preprocessing facility, as indicated by the arrows. At an integrated processing facility the biomass can be converted into bioproducts that can be transported to demand sites. Alternatively, if the biomass is transported to a preprocessing facility it is first converted into bio-oil, an intermediate that needs further processing. Therefore, the bio-oil is subsequently transported to an upgrading facility where it is converted into bioproducts, which can then be transported to satisfy demand. Also, each processing facility is allowed to keep inventory of its input and output species, e.g. a preprocessing facility can only store biomass and bio-oil, not bioproduct.

Figure 2.1: Bio-based supply chain superstructure.
The bio-based supply chain is assessed through the development of a multi-period spatially explicit mathematical programming model. The supply chain is assumed to repeatedly go through a yearly cycle, which may be divided into multiple equal time periods. The geographical area of study is discretized into equally sized grid units. The centroid of each grid unit is a harvesting site, demand site and potential processing facility location. Biomass supply and bioproduct demand are assumed to be homogeneously distributed over the area of each grid unit. The problem is formulated as follows:

Given

- a yearly cycle consisting of a certain number of equal time periods,
- the amount of biomass supply at each location in each time period,
- the amount of bioproduct demand at each location in each time period,

determine the

- location, processing and storage capacity of each integrated processing facility,
- location, processing and storage capacity of each preprocessing facility,
- location, processing and storage capacity of each upgrading facility,
- biomass flows from harvesting sites to integrated facilities,
- biomass flows from harvesting sites to preprocessing facilities,
- bio-oil flows from preprocessing facilities to upgrading facilities,
- bioproduct flows from integrated processing facilities to demand sites,
- bioproduct flows from upgrading facilities to demand sites,

minimizing the total expected annual cost.

2.2 Mathematical Model Formulation

The model is formulated for a geographical area discretized into \( N \) equally sized grid units \( 1, 2, \ldots, N \). The yearly cycle consists of \( T \) equal time periods numbered \( 1, 2, \ldots, T \). Furthermore, the model formulation is generic and actually includes more flexibility than required for the problems that will be solved in this work. An overview of sets, parameters and variables can be found in the Nomenclature section at the end of this work. To facilitate readability, the convention is adopted to use calligraphic letters for sets, capital and Greek letters for parameters and lower case letters to indicate variables. The constraints used to model the major characteristics of biomass harvesting sites, bioproduct demand sites, processing and storage facilities are presented in Sections 2.2.1 to 2.2.6. In Section 2.2.7 the objective function is presented.

Intrinsically, this problem is a mixed integer non-linear program: The material flow and storage variables are continuous and non-negative. Binary variables are used specify the processing facility locations. Furthermore, all the constraints are linear, but the objective function (i.e. the total expected annual cost) is non-linear. To avoid non-linearities, the objective function is approximated by a piece-wise linear function. The resulting problem is therefore a mixed integer linear program (MILP).

2.2.1 Harvesting Sites

The amount of biomass type \( b \) available at harvesting site \( i \) in time period \( t \) \((A_{b,i,t})\) is equal to the total amount of agricultural land available at that particular harvesting site \((AL_i)\), multiplied
by the fraction that is used for biomass type $b$ production ($\phi_{b,i}$) and the biomass type $b$ yield per surface area in time period $t$ ($\eta_{b,t}$) in terms of dry weight:

$$A_{b,i,t} = AL_i \cdot \phi_{b,i} \cdot \eta_{b,t} \quad \forall \ b \in \mathcal{B}, i \in \mathcal{I}, t \in \mathcal{T}$$ (2.1)

where $\mathcal{B} = \{1, \ldots, B\}$ is the set of biomass types, $\mathcal{I} = \{1, \ldots, I\}$ is the set of harvesting sites and $\mathcal{T} = \{1, \ldots, T\}$ is the set of time periods. It is noted that seasonality, harvesting windows and temporal variability can be taken into account through different values of the parameter $\eta_{b,t}$.

The fraction of agricultural land at harvesting site $i$ used for the production of biomass type $b$ is constrained by a certain maximum total fraction of agricultural land that may be utilized:

$$\sum_b \phi_{b,i} \leq \Phi \quad \forall \ i \in \mathcal{I}$$ (2.2)

where $\Phi$ is always less than or equal to 1. Notice that the maximum total fraction $\Phi$ is the same for all sites $i$.

The total amount of biomass type $b$ acquired from harvesting site $i$ in time period $t$ may not exceed its available amount in terms of dry weight:

$$h_{b,i,t} \leq A_{b,i,t} \quad \forall \ b \in \mathcal{B}, i \in \mathcal{I}, t \in \mathcal{T}$$ (2.3)

where $h_{b,i,t}$ is the amount of biomass type $b$ obtained from harvesting site $i$ in time period $t$.

The mass balance of harvesting site $i$ in time period $t$ for biomass type $b$ is given by the following equation:

$$h_{b,i,t} = \sum_j fi_{b,i,j,t} + \sum_k fi_{b,i,k,t} \quad \forall \ b \in \mathcal{B}, i \in \mathcal{I}, t \in \mathcal{T}$$ (2.4)

where $fi_{b,i,j,t}$ is the flow of biomass type $b$ from harvesting site $i$ to integrated conversion facility $j$ in time period $t$. Similarly, $fi_{b,i,k,t}$ is the flow of biomass type $b$ from harvesting site $i$ to preconversion facility $k$ in time period $t$.

### 2.2.2 Integrated Processing Facilities

The mass balance of biomass type $b$ at integrated processing facility $j$ requires that the total amount of feedstock type $b$ transported from harvesting sites $i$ to integrated processing facility $j$ in time period $t$ plus the inventory level of biomass type $b$ at the end of the previous time period is equal to the amount of biomass used for the production of bioproducts plus the inventory level at the end of the current time period. This relationship is expressed through the following constraints:

$$\sum_i fi_{b,i,j,t} + \left(1 - \varepsilon_b\right) \cdot sj_{b,j,t-1} = u_{b,j,t} + sj_{b,j,t} \quad \forall \ b \in \mathcal{B}, j \in \mathcal{J}, t \geq 2$$ (2.5)

$$\sum_i fi_{b,i,j,t} + \left(1 - \varepsilon_b\right) \cdot sj_{b,j,T} = u_{b,j,t} + sj_{b,j,t} \quad \forall \ b \in \mathcal{B}, j \in \mathcal{J}, t = 1$$ (2.6)

where $\mathcal{J} = \{1, \ldots, N\}$ is the set of potential locations for the construction of an integrated processing facility, $u_{b,j,t}$ is the amount of biomass type $b$ utilized at integrated processing facility $j$ in time period $t$, $sj_{b,j,t}$ is the biomass inventory level at integrated processing facility $j$ at the end of time period $t$ and $\varepsilon_b$ is the percentage of biomass type $b$ deteriorated at this storage.
facility in time period \( t \). By introducing the deterioration factor \( \varepsilon_b \) the degradation characteristics of biomass are captured in the model. In order to investigate the annualized cost, the inventory balances are formulated in a “cyclic” manner; i.e. the inventory level at the end of a year is the same as the inventory level at the beginning of a year, after considering biomass deterioration. Equation (2.6) expresses this cyclic inventory balance.

The biomass consumption \( (u_{jb,j,t}) \) relates to the bioproduct production through a conversion factor \( \alpha_b \). Bioproducts are either stored or transported to demand sites directly. This leads to the following mass balance equations:

\[
\sum_b (u_{jb,j,t} \cdot \alpha_b) + s_{jp,j,t-1} = \sum_n f_{jn,j,n,t} + s_{jp,j,t} \quad \forall \ j \in J, t \geq 2
\]

\[
\sum_b (u_{jb,j,t} \cdot \alpha_b) + s_{jp,j,T} = \sum_n f_{jn,j,n,t} + s_{jp,j,t} \quad \forall \ j \in J, t = 1
\]

where \( f_{jn,j,n,t} \) is the bioproduct flow from integrated processing facility \( j \) to demand site \( n \) and \( s_{jp,j,t} \) is the inventory level of bioproduct at integrated processing facility \( j \) at the end of time period \( t \). Note that in comparison to equations (2.5) and (2.6) no deterioration factor is present. This is due to the fact that the significance of bio-oil degradation is minimal compared to that of biomass.

To model the selection of an integrated processing facility location and its capacity level a binary variable, \( x_{j,r} \), is introduced through the following constraints:

\[
\sum_r x_{j,r} \leq 1 \quad \forall \ j \in J
\]

\[
\sum_j \sum_r x_{j,r} \leq NJ
\]

where \( x_{j,r} \) is equal to 1 when an integrated processing facility is built at location \( j \) with capacity level \( r \) and \( NJ \) is the maximum allowable number of integrated processing facilities. Constraint (2.9) ensures that at most one capacity level can be chosen. Constraint (2.10) enforces that the cumulative number of integrated processing facilities remains below a certain maximum value \( NJ \).

The capacity level \( r \) refers to a certain predefined annual production capacity range. Following from this definition, the annual production capacity, in terms of gasoline gallon equivalent (abbr. GGE), of an integrated processing facility, \( c_{j,r} \), is given by the following constraints:

\[
C_{J_{r-1}} \cdot x_{j,r} \leq c_{j,r} \leq C_{J_r} \cdot x_{j,r} \quad \forall \ j \in J, r \in \mathcal{R}
\]

where \( \mathcal{R} = \{1, \ldots, R\} \) is the set of capacity levels, \( C_{J_r} \) is the upper boundary of an integrated processing facility with a capacity level \( r \).

To account for economies of scale regarding the total capital investment required for the construction of a processing facility, it is common practice to apply a certain scaling equation of the form:

\[
\frac{\text{cost}_{new}}{\text{cost}_{ref}} = \left( \frac{\text{capacity}_{new}}{\text{capacity}_{ref}} \right)^m
\]

where the subscript \( \text{new} \) refers to the facility that is to be constructed, \( \text{ref} \) refers to some reference facility and \( m \) is a certain scale factor. However, this would result in the model becoming nonlinear, which is preferably avoided. Therefore, to take into account economies
of scale, the total capital investment associated with the construction of integrated processing facility \( j \) is modeled through the following piece-wise linear curve:

\[
cc_{j} = \sum_{r} \left( CCJ_{r-1} \cdot x_{j,r} + (c_{j,r} - CCJ_{r-1} \cdot x_{j,r}) \frac{CCJ_{r-1} - CCJ_{r}}{CJ_{r-1} - CJ_{r}} \right) \quad \forall \ j \in \mathcal{J} \quad (2.12)
\]

where \( cc_{j} \) is the total capital investment associated with integrated facility \( j \). Furthermore, \( CCJ_{r} \) and \( CJ_{r} \) are the capital investment and the annual production capacity corresponding to the construction of an integrated processing facility with capacity level \( r \) respectively. Due to constraint (2.11) and (2.12), when \( x_{j,r} \) is 0, the total capital investment \( cc_{j} \) equals 0 as well. The values of the parameters \( CCJ_{r} \) and \( CJ_{j,r} \) are determined using a separate MILP regression model, which is discussed in Section 2.4.

The operating and maintenance (O&M) cost associated with integrated processing facility \( j \) consists of a fixed part and a variable part. The fixed part includes mainly capital charge and maintenance, which both are predominantly dependent on the size of the facility. Therefore, the annual fixed O&M cost is modeled as a percentage of the total capital investment:

\[
cf_{j} = CFJ \cdot cc_{j} \quad \forall \ j \in \mathcal{J} \quad (2.13)
\]

where \( cf_{j} \) is the total annual fixed O&M cost for integrated processing facility \( j \) and \( CFJ \) is the total annual fixed O&M cost as a percentage of the total capital investment for an integrated processing facility. The variable part includes primarily catalyst and energy costs and is modeled as scaling linearly with the total annual production quantity:

\[
cv_{j} = CVJ \cdot \sum_{b} \sum_{t} u_{j,b,t} \quad \forall \ j \in \mathcal{J} \quad (2.14)
\]

where \( cv_{j} \) is the total annual variable O&M cost for integrated processing facility \( j \) and \( CVJ \) is the cost per processed unit for any integrated processing facility.

The production quantity at integrated processing facility \( j \) in time period \( t \) may not exceed the annual production capacity of integrated processing facility \( j \) divided by the number of time periods in a year \( T \). Furthermore, a lower bound for the production quantity is introduced through a minimum capacity utilization percentage \( \vartheta' \). Together this results in the following production quantity constraints:

\[
\frac{\vartheta'}{T} \sum_{r} cj_{j,r} \leq \sum_{b} (u_{j,b,t} \cdot \alpha_{b}) \leq \frac{1}{T} \sum_{r} cj_{j,r} \quad \forall \ j \in \mathcal{J}, t \in \mathcal{T} \quad (2.15)
\]

### 2.2.3 Preprocessing Facilities

The mass balance of biomass type \( b \) at preprocessing facility \( k \) in time period \( t \) is similar to that of an integrated processing facility. It is given by the following equations:

\[
\sum_{i} f_{i,k,b,t} + (1 - \varepsilon_{b}) \cdot skm_{b,k,t-1} = uk_{b,k,t} + skm_{b,k,t} \quad \forall \ b \in \mathcal{B}, k \in \mathcal{K}, t \geq 2 \quad (2.16)
\]

\[
\sum_{i} f_{i,k,b,t} + (1 - \varepsilon_{b}) \cdot skm_{b,k,T} = uk_{b,k,t} + skm_{b,k,t} \quad \forall \ b \in \mathcal{B}, k \in \mathcal{K}, t = 1 \quad (2.17)
\]

where \( \mathcal{K} = \{1, \ldots, N\} \) is the set of potential locations for the construction of a preprocessing facility, \( uk_{b,k,t} \) is the amount of biomass type \( b \) utilized at preprocessing facility \( k \) in time period
$t$ and $skm_{b,k,t}$ is the inventory level of biomass type $b$ at preprocessing facility $k$ at the end of time period $t$.

The biomass consumption $(uk_{b,k,t})$ relates to the bio-oil production through a conversion factor $\beta_b$. This leads to the following mass balance equations:

$$\sum_b (uk_{b,k,t} \cdot \beta_b) + sko_{k,t-1} = \sum_l fkl_{k,l,t} + sko_{k,t} \quad \forall k \in \mathcal{K}, t \geq 2$$ (2.18)

$$\sum_b (uk_{b,k,t} \cdot \beta_b) + sko_{k,T} = \sum_l fkl_{k,l,t} + sko_{k,t} \quad \forall k \in \mathcal{K}, t = 1$$ (2.19)

where $fkl_{k,l,t}$ is the bio-oil flow from preprocessing facility $k$ to upgrading facility $l$ and $sko_{k,t}$ is the inventory level of bio-oil at preprocessing facility $k$ at the end of time period $t$.

To model the selection of a preprocessing facility location and its capacity level a binary variable, $y_{j,r}$, is introduced through the following constraints:

$$\sum_r y_{k,r} \leq 1 \quad \forall k \in \mathcal{K}$$ (2.20)

$$\sum_k \sum_r y_{k,r} \leq NK$$ (2.21)

where $y_{k,r}$ is equal to 1 when a preprocessing facility is built at location $k$ with capacity level $r$ and $NK$ is the maximum allowable number of preprocessing facilities.

The annual production capacity (in terms of dry weight) of preprocessing facility $k$, $ck_{k,r}$, is given by the following constraints:

$$CK_{r-1} \cdot y_{k,r} \leq ck_{k,r} \leq CK_r \cdot y_{k,r} \quad \forall k \in \mathcal{K}, r \in \mathcal{R}$$ (2.22)

where $CK_{kr}$ is the upper bound for a preprocessing facility $k$ with a capacity level $r$.

As with the integrated facilities, to avoid nonlinearities, the total capital investment associated with the construction of a preprocessing facility $k$ is modeled through a piece-wise linear curve:

$$ck_k = \sum_r \left( CCK_{r-1} \cdot y_{k,r} + (ck_{k,r} - CK_{r-1} \cdot y_{k,r}) \left( \frac{CK_r - CK_{r-1}}{CK_{r-1}} \right) \right) \quad \forall k \in \mathcal{K}$$ (2.23)

where $ck_k$ is the total capital investment associated with preprocessing facility $k$. Furthermore, $CCK_r$ and $CK_r$ are the capital investment and the annual production capacity corresponding to the construction of a preprocessing facility with capacity level $r$ respectively.

The fixed O&M cost associated with preprocessing facility $k$ is:

$$cf_{k} = CFK \cdot ck_k \quad \forall k \in \mathcal{K}$$ (2.24)

where $cf_{k}$ is the total annual fixed O&M cost and $CFK$ is the total annual fixed O&M cost as a percentage of the total capital investment for preprocessing facility $k$. As for integrated facilities, the variable part of the O&M cost is assigned on an annual throughput basis:

$$cvk_k = CVK \cdot \sum_b \sum_t uk_{b,k,t} \quad \forall k \in \mathcal{K}$$ (2.25)

where $cvk_k$ is the total annual variable O&M cost and $CVK$ is the cost per processed unit for preprocessing facility $k$. 
The production quantity at preprocessing facility $k$ in time period $t$ may not exceed the annual production capacity of preprocessing facility $k$ divided by the number of time period in a year. A lower bound for the production quantity is introduced through the minimum capacity utilization percentage $\vartheta''$, resulting in:

$$\frac{\vartheta''}{T} \sum_r c_{k,r} \sum_b u_{k,b,k,t} \leq \frac{1}{T} \sum_r c_{k,r} \quad \forall \ k \in \mathcal{K}, t \in \mathcal{T} \quad (2.26)$$

### 2.2.4 Upgrading Facilities

For bio-oil, the mass balance at upgrading facility $l$ is given by the following equations:

$$\sum_k f_{k,l,t} + s_{l,T-1} = u_{l,t} + s_{l,T} \quad \forall \ l \in \mathcal{L}, t \geq 2 \quad (2.27)$$

$$\sum_k f_{k,l,t} + s_{l,T} = u_{l,t} + s_{l,T} \quad \forall \ l \in \mathcal{L}, t = 1 \quad (2.28)$$

where $\mathcal{L} = \{1, \ldots, N\}$ is the set of potential locations for the construction of an upgrading facility, $u_{l,t}$ is the amount of bio-oil utilized at upgrading facility $l$ in time period $t$ and $s_{l,t}$ is the bio-oil inventory level at upgrading facility $l$ at the end of time period $t$.

The bio-oil consumption ($u_{l,t}$) relates to the bioproduct production through a conversion factor $\gamma$, leading to the following mass balance equations:

$$u_{l,t} \cdot \gamma + s_{l,T-1} = \sum_n f_{l,n,t} + s_{l,p,t} \quad \forall \ l \in \mathcal{L}, t \geq 2 \quad (2.29)$$

$$u_{l,t} \cdot \gamma + s_{l,T} = \sum_n f_{l,n,t} + s_{l,p,t} \quad \forall \ l \in \mathcal{L}, t = 1 \quad (2.30)$$

where $f_{l,n,t}$ is the bioproduct flow from upgrading facility $k$ to upgrading facility $l$ to demand site $n$ in time period $t$ and $s_{l,p,t}$ is the inventory level of bioproduct at upgrading facility $l$ at the end of time period $t$.

For modelling the selection of an upgrading facility location and its capacity level a binary variable, $z_{l,r}$, is introduced through the following constraints:

$$\sum_r z_{l,r} \leq 1 \quad \forall \ l \in \mathcal{L} \quad (2.31)$$

$$\sum_l \sum_r z_{l,r} \leq NL \quad (2.32)$$

where $z_{l,r}$ is equal to 1 when an upgrading facility is built at location $l$ with capacity level $r$ and $NL$ is the maximum allowable number of preprocessing facilities.

The annual production capacity (in terms of gasoline gallon equivalent) of upgrading facility $l$, $c_{l,r}$, is given by the following constraint:

$$CL_{r-1} \cdot z_{l,r} \leq c_{l,r} \leq CL_r \cdot z_{l,r} \quad \forall \ l \in \mathcal{L}, r \in \mathcal{R} \quad (2.33)$$

where $CL_r$ is the upper bound for a upgrading facility $l$ with a capacity level $r$. 

13 Design of Bio-Based Supply Chains / Final Version
In a similar manner as with the integrated and preprocessing, to avoid nonlinearities, the total capital investment associated with the construction of an upgrading facility $l$ is modeled through a piece-wise linear curve:

$$ccl_l = \sum_r \left( CCL_{r-1} \cdot z_{l,r} + (c_{l,r} - CL_{r-1} \cdot z_{l,r}) \left( \frac{CCL_{lr} - CCL_r}{CL_{r-1} - CL_r} \right) \right) \quad \forall \; l \in L \quad (2.34)$$

where $ccl_l$ is the total capital investment associated with upgrading facility $l$. $CCL_r$ and $CL_r$ are, respectively, the capital investment and the annual production capacity corresponding to the construction of an upgrading facility with capacity level $r$.

The fixed and variable O&M cost associated with upgrading facility $l$ are respectively:

$$cfl_l = CFL \cdot ccl_l \quad \forall \; l \in L \quad (2.35)$$

$$cvl_l = CVL \cdot \sum_b \sum_t u_{b,l,t} \quad \forall \; l \in L \quad (2.36)$$

where $cfl_l$ and $cvl_l$ are respectively the total annual fixed and variable O&M cost for upgrading facility $l$, $CFL$ is the total annual fixed O&M cost as a percentage of the total capital investment and $CVL$ is the cost per processed unit.

The production quantity at upgrading facility $l$ in time period $t$ may not exceed the annual production capacity of upgrading facility $l$ divided by the number of time periods in a year. Similar to the integrated processing facilities, a lower bound for the production quantity is introduced through a minimum capacity utilization percentage $\theta''$. Combined this results in the following production quantity constraints:

$$\theta'' T \sum_r c_{l,r} \leq \gamma \cdot u_{l,t} \leq \frac{1}{T} \sum_r c_{l,r} \quad \forall \; l \in L, t \in T \quad (2.37)$$

### 2.2.5 Storage Facilities

In principle there are three different species that propagate through the supply chain, i.e. biomass, -oil and -products. All species can be temporarily stored at designated storage facilities which are all located at the same location as the processing facilities.

Although the total inventory level at each of the storage facilities may be different in each time period, the capacity (in terms of volume) of a storage facility is dictated by its maximum inventory level over time. Therefore, the biomass storage capacity at integrated processing facility $j$, $csjm_j$ is given by:

$$csjm_j = \max_l \left( \sum_b \frac{s_{jm,b,j,t}}{\rho_b} \right) \quad \forall \; j \in J$$

where $\rho_b$ is the volumetric mass density of biomass type $b$. Since some mathematical programming software do not contain a function that returns the maximum (or minimum) element of a set, alternatively, the equation above can be incorporated as follows:

$$csjm_j \geq \left( \sum_b \frac{s_{jm,b,j,t}}{\rho_b} \right) \quad \forall \; j \in J, t \in T \quad (2.38)$$
Analogously, the required storage capacity for bioproduct at integrated processing facility \( j \), \( csjp_j \), is given by:

\[
    csjp_j \geq \left( \sum_b \frac{sjp_{j,t}}{U_{\text{bioproduct}}} \right) \quad \forall \; j \in \mathcal{J}, t \in \mathcal{T}
\]  

(2.39)

where \( U_{\text{bioproduct}} \) is the energy density (energy per unit volume) of bioproduct.

At a preprocessing facility \( j \) both biomass and -oil can be stored. The required capacities are respectively:

\[
    cskm_k \geq \left( \sum_b \frac{skm_{b,k,t}}{\rho_b} \right) \quad \forall \; k \in \mathcal{K}, t \in \mathcal{T}
\]  

(2.40)

\[
    csko_k \geq \left( \frac{sko_{k,t}}{U_{\text{bio-oil}}} \right) \quad \forall \; k \in \mathcal{K}, t \in \mathcal{T}
\]  

(2.41)

where \( U_{\text{bio-oil}} \) is the energy density of bio-oil.

In a similar manner, the required storage capacity for bio-oil and -product at upgrading facility \( l \) respectively are:

\[
    cslo_l \geq \left( \frac{slo_{l,t}}{U_{\text{bio-oil}}} \right) \quad \forall \; l \in \mathcal{L}, t \in \mathcal{T}
\]  

(2.42)

\[
    cslp_l \geq \left( \frac{slp_{l,t}}{U_{\text{bioproduct}}} \right) \quad \forall \; l \in \mathcal{L}, t \in \mathcal{T}
\]  

(2.43)

Contrary to the processing facilities, where economies of scale are relatively significant, the total capital investment required for the construction of a certain storage facility is assumed to scale linearly with its storage capacity. Therefore, the required capital investment needed for the construction of biomass storage capacity, \( ccsm \), is given by:

\[
    ccsm = CCSM \left( \sum_j csjm_j + \sum_k cskm_k \right)
\]  

(2.44)

where \( CCSM \) is the investment cost per volume of biomass storage capacity.

Similarly, the capital investment associated with bio-oil and -product storage capacity is given by:

\[
    ccso = CCSO \left( \sum_k csko_k + \sum_l cslo_l \right)
\]  

(2.45)

\[
    ccsp = CCSP \left( \sum_j csjp_j + \sum_l cslp_l \right)
\]  

(2.46)

where \( ccso \) and \( ccsp \) are the annual capital cost associated with bio-oil and -product capacity, and \( CCSO \) and \( CCSP \) are the required investment cost per volume storage capacity for each of the respective species. The reason that the benefits that can be gained from large scale storage facilities are minimal, is mainly due to the fact that no (automated) processing takes place at a storage facility.

Furthermore, the operating and maintenance cost associated with the various storage facilities consists only of a fixed part:

\[
    cfsm = CFSM \cdot ccsm
\]  

(2.47)
where \(CFSM\) is the O&M cost as a percentage of the total capital investment associated with the construction of biomass storage capacity.

Similarly, the operating and maintenance cost associated with bio-oil and -product storage capacity is given by:

\[
\begin{align*}
    cfso &= CFSM \cdot ccso \\
    cfsp &= CFSM \cdot ccsp
\end{align*}
\]  

(2.48)  

(2.49)

where \(CFSO\) and \(CFSP\) are the total annual O&M cost for bio-oil and -product storage, as a percentage of the total capital investment associated with the respective storage capacities. As mentioned, in case of processing facilities the variable O&M cost account for e.g. energy and catalyst cost, which both depend on the production quantity. Apart from hauling, there are no cost that largely depend on the amount of material that is stored. Therefore, the variable O&M are assumed to be negligibly small for storage facilities.

### 2.2.6 Demand Sites

The bioproduct flows from integrated processing facilities \(j \in J\) and upgrading facilities \(l \in L\) to demand site \(n\) in time period \(t\) are required to satisfy the bioproduct demand at demand site \(n\) in time period \(t\). This is expressed by the following constraint:

\[
\sum_j f_{jn,n,t} + \sum_l f_{ln,n,t} \geq PD_{n,t} \quad \forall \; n \in N, t \in T
\]  

(2.50)

where \(PD_{n,t}\) is the bioproduct demand at demand site \(n\) in time period \(t\).

### 2.2.7 Objective: Minimize Annual Total Cost

The objective is to minimize the annual total cost \(tc\), which is expressed with the following summation:

\[
\min tc = C_{acq.} + C_{processing \_capital} + C_{storage \_capital} + C_{processing \_O&M} + C_{storage \_O&M} + C_{biomass \_transport} + C_{bio-oil \_transport} + C_{bioproduct \_transport}
\]  

(2.51)

subject to (2.1)-(2.50), supplemented with non-negativity constraints for all variables. Each of the terms appearing in \(tc\) are specified below.

In the first term, the amount paid to the farmer, i.e. the cost associated with biomass acquisition, is given by:

\[
C_{acq.} = \sum_{b,i,t} AC_b \cdot h_{b,i,t}
\]  

(2.52)

where \(AC_b\) is the cost per unit of biomass type \(b\).

The annual capital cost associated with processing facilities includes the capital investments required for the construction of integrated processing, preprocessing and upgrading facilities. To equally spread the total capital investment over a certain time horizon, it is multiplied by a certain annuity factor:

\[
C_{processing \_capital} = \frac{IR(1 + IR)^{NY}}{(1 + IR)^{NY} - 1} \left( \sum_j ccj_j + \sum_k cck_k + \sum_l cc_l \right)
\]  

(2.53)

where

16 Design of Bio-Based Supply Chains / Final Version
where $IR$ is the discount rate and $NY$ is the lifetime of the project in years. A derivation of the annuity factor is presented in Appendix A.1.

Similarly, the annual capital cost associated with storage facilities consists of the capital investments required for the construction of biomass, -oil and -product storage capacity:

$$C_{storage Capital} = \frac{IR(1 + IR)^{NY}}{(1 + IR)^{NY} - 1} (ccsm + ccs0 + ccsp)$$ (2.54)

The annual operating and maintenance cost associated with processing facilities consists of the fixed and variable O&M cost associated with each integrated processing facility $j$, preprocessing facility $k$ and upgrading facility $l$:

$$C_{O&M}^{processing} = \sum_j (cf_j + cv_j) + \sum_k (cf_k + cv_k) + \sum_l (cf_l + cv_l)$$ (2.55)

The annual operating and maintenance cost associated with biomass, -oil and product storage capacity is:

$$C_{O&M}^{storage} = cfsm + cfs0 + cfsp$$ (2.56)

The annual biomass transportation cost is given by:

$$C_{transport}^{biomass} = \sum_{b,i,t} (DFCM + DVCM \cdot AD) \left( \frac{h_{b,i,t}}{1 - MC_b} \right) + \sum_{b,i,j,t} (DFCM + DVCM \cdot D_{i,j}) \left( \frac{f_{ijb,i,t}}{1 - MC_b} \right) + \sum_{b,i,k,t} (DFCM + DVCM \cdot D_{i,k}) \left( \frac{f_{ikb,i,t}}{1 - MC_b} \right)$$ (2.57)

where $DFCM$ and $DVCM$ are the distance fixed and variable cost of biomass, $AD$ denotes the average distance from all points inside a grid unit to its center, $D_{i,j}$ is the distance between harvesting site $i$ and integrated processing facility $j$ (see Appendix A.2) and $MC_b$ is the moisture content of biomass type $b$. The reasoning behind the first term of Equation (2.57) is given in Section 2.3.

The annual bio-oil transportation cost is more straightforward and is given by:

$$C_{transport}^{bio-oil} = \sum_{k,l,t} (DFCO + DVCO \cdot D_{k,l}) fkl_{k,l,t}$$ (2.58)

where $DFCO$ and $DVCO$ are the distance fixed and variable cost associated with the transportation of bio-oil.

At last, the annual bioproduct transportation cost is given by:

$$C_{transport}^{bioproduct} = \sum_{n,t} (DFCP + DVCP \cdot AD) PD_{n,t} + \sum_{j,n,t} (DFCP + DVCP \cdot D_{j,n}) fjn_{j,n,t} + \sum_{l,n,t} (DFCP + DVCP \cdot D_{l,n}) fln_{l,n,t}$$ (2.59)

where $DFCP$ and $DVCP$ are the distance fixed and variable cost of bio-oil. The idea behind the first term of Equation (2.59) is similar to that of Equation (2.57).
2.3 Discretization

Since the main trade-off between the various processing strategies is based on transportation cost, it is of great importance that this cost is modeled with sufficient accuracy. As stated before, the geographical area of study is discretized into a grid where the centroid of each grid unit is a potential processing facility location. The biomass supply and bioproduct demand are assumed to be homogeneously distributed over the area of a grid unit. Nonetheless, since only the distance between center points is taken into account, the biomass supply of a grid unit is modeled as if it originates from its centroid. Similarly, all bioproducts are transported to a centroid to satisfy a grid units’ demand. This would not pose a problem if the grid is divided into a very large number of grid units. However, increasing the number of grid units is associated with an even stronger increase of the number of binary variables (see e.g. Equation (2.9)). As a consequence, more computational effort is required, soon making the model impossible to be solved within a reasonable time span. At the other extreme, if the same area is “divided” into a single grid unit, there are entirely no logistics since all biomass is harvested and processed at the same centroid as where all bioproduct demand is located. This is also the case when a processing facility is constructed at every grid unit, which also leads to “free” transportation and thus to an unfair preference towards decentralization. Therefore, it is essential to include a correction factor in the model.

In comparison to biomass and bioproducts, bio-oil is an intermediate and therefore only transported between processing and storage facilities. As in the real world, these facilities will be constructed at certain (approximately) point locations. Therefore, this issue only presents itself in the transportation cost for biomass and bioproducts. To adjust for this, the first terms in Equations (2.57) and (2.59) are introduced. As mentioned, $AD$ is the average distance from all points inside a grid unit to its center. For a square grid unit of length $\ell$ its value is calculated through the following integral:

$$AD = \ell \cdot \int_{-\frac{\ell}{2}}^{\frac{\ell}{2}} \int_{-\frac{\ell}{2}}^{\frac{\ell}{2}} \sqrt{x^2 + y^2} \, dx \, dy = \frac{\ell}{6} (\sqrt{2} + \sinh^{-1}(1))$$  \hspace{1cm} (2.60)

where $x$ and $y$ refer to Cartesian coordinates. The complete derivation is presented in Appendix A.3. The core idea of these terms is that the biomass type $b$ acquired at a certain grid unit $i$ in a certain time period $t$, i.e $h_{b,i,t}$, is first transported to its center. From here, it may be transported onwards to a processing facility. Similarly, when bioproduct is transported to a certain grid unit to satisfy demand, it is basically “dispersed” from its centroid over the entire area of the grid unit.

![Figure 2.2: Biomass is first transported to the center point of each grid unit before it is transported to the processing facility, which is indicated by the center dot.](image)

To illustrate the significance of these terms, consider a 1-by-1 km hypothetical area with a completely homogeneous biomass supply distribution and a processing facility in the center.
If this area is divided into an infinite number of grid units the average distance over which the biomass is transported is equal to 0.3826 km (by evaluating Equation (2.60)). If the area is divided into a countable number of grid units, biomass is first transported to the center point of a grid unit before it is actually transported to the processing facility, as is illustrated for a 3 × 3 grid in Figure 2.2.

The average distance to the center in case this area is divided into an increasing number of grid units is shown in Figure 2.3.

As the number of grid units increases (i.e. smaller grid units), the average distance to the center decreases steeply to a value of 0.4130 for a 11 × 11 grid, which is relatively close to its limiting value. Although this approximation method somewhat overestimates the distance, it is crucial to include this approximate correction factor due to the aforementioned reasons.

2.4 Piecewise Linearization

As mentioned previously, to account for economies of scale regarding the total capital investment required for the construction of a certain processing facility, a scaling equation (see Equation (??)) is applied. To avoid nonlinearities on the one hand and to prevent loss of accuracy on the other, this equation is linearized via a method known as piece-wise linearization. In this section, the mathematical programming model that is used for the regression is presented. Although some of the parameters, variables and indices above are chosen to match closely with the model presented in Section 2.2, it should be noted that this method can be applied to other regression purposes as well.

MILP Regression Model

First, the function to be linearized is discretized into so-called “samples”. The main idea of the mathematical programming model presented here is that these samples are separated into different regions. To each region a different linear function is fitted. The model optimises the location of break-points and regression coefficient such that the least absolute error is achieved. This is illustrated in Figure 2.4.

Figure 2.3: Average distance to the center for a varying number of grid units.

As the number of grid units increases (i.e. smaller grid units), the average distance to the center decreases steeply to a value of 0.4130 for a 11 × 11 grid, which is relatively close to its limiting value. Although this approximation method somewhat overestimates the distance, it is crucial to include this approximate correction factor due to the aforementioned reasons.

2.4 Piecewise Linearization

As mentioned previously, to account for economies of scale regarding the total capital investment required for the construction of a certain processing facility, a scaling equation (see Equation (??)) is applied. To avoid nonlinearities on the one hand and to prevent loss of accuracy on the other, this equation is linearized via a method known as piece-wise linearization. In this section, the mathematical programming model that is used for the regression is presented. Although some of the parameters, variables and indices above are chosen to match closely with the model presented in Section 2.2, it should be noted that this method can be applied to other regression purposes as well.

MILP Regression Model

First, the function to be linearized is discretized into so-called “samples”. The main idea of the mathematical programming model presented here is that these samples are separated into different regions. To each region a different linear function is fitted. The model optimises the location of break-points and regression coefficient such that the least absolute error is achieved. This is illustrated in Figure 2.4.

Figure 2.3: Average distance to the center for a varying number of grid units.

As the number of grid units increases (i.e. smaller grid units), the average distance to the center decreases steeply to a value of 0.4130 for a 11 × 11 grid, which is relatively close to its limiting value. Although this approximation method somewhat overestimates the distance, it is crucial to include this approximate correction factor due to the aforementioned reasons.

2.4 Piecewise Linearization

As mentioned previously, to account for economies of scale regarding the total capital investment required for the construction of a certain processing facility, a scaling equation (see Equation (??)) is applied. To avoid nonlinearities on the one hand and to prevent loss of accuracy on the other, this equation is linearized via a method known as piece-wise linearization. In this section, the mathematical programming model that is used for the regression is presented. Although some of the parameters, variables and indices above are chosen to match closely with the model presented in Section 2.2, it should be noted that this method can be applied to other regression purposes as well.

MILP Regression Model

First, the function to be linearized is discretized into so-called “samples”. The main idea of the mathematical programming model presented here is that these samples are separated into different regions. To each region a different linear function is fitted. The model optimises the location of break-points and regression coefficient such that the least absolute error is achieved. This is illustrated in Figure 2.4.

Figure 2.3: Average distance to the center for a varying number of grid units.

As the number of grid units increases (i.e. smaller grid units), the average distance to the center decreases steeply to a value of 0.4130 for a 11 × 11 grid, which is relatively close to its limiting value. Although this approximation method somewhat overestimates the distance, it is crucial to include this approximate correction factor due to the aforementioned reasons.

2.4 Piecewise Linearization

As mentioned previously, to account for economies of scale regarding the total capital investment required for the construction of a certain processing facility, a scaling equation (see Equation (??)) is applied. To avoid nonlinearities on the one hand and to prevent loss of accuracy on the other, this equation is linearized via a method known as piece-wise linearization. In this section, the mathematical programming model that is used for the regression is presented. Although some of the parameters, variables and indices above are chosen to match closely with the model presented in Section 2.2, it should be noted that this method can be applied to other regression purposes as well.

MILP Regression Model

First, the function to be linearized is discretized into so-called “samples”. The main idea of the mathematical programming model presented here is that these samples are separated into different regions. To each region a different linear function is fitted. The model optimises the location of break-points and regression coefficient such that the least absolute error is achieved. This is illustrated in Figure 2.4.
When the number of regions $R$ is given, the number of break-points is equal to $R - 1$. These are ordered as follows:

$$C_{r-1} \leq C_r \quad \forall \quad r = 2, 3, \ldots, R \quad (2.61)$$

where $C_r$ is the location of break-point $r$.

Each sample $s$ can only belong to a single region $r$. This is modeled through the following constraints:

$$C_{r-1} - LN(1 - a_{s,r}) \leq c_s \quad \forall \quad s \in \mathcal{S}, r = 2, 3, \ldots, R \quad (2.62)$$

$$c_s \leq C_r + LN(1 - a_{s,r}) \quad \forall \quad s \in \mathcal{S}, r = 1, 2, \ldots, R - 1 \quad (2.63)$$

where $\mathcal{S} = \{1, \ldots, S\}$ is the set of samples, $c_s$ is the input value of sample $s$, $LN$ is an arbitrary large number and $a_{s,r}$ is a binary variable, which is equal to 1 when sample $s$ is partitioned in region $r$. In this case, sample $s$ falls into the region with lower bound $C_{r-1}$ and upper bound $C_r$. Note that when $a_{s,r} = 1$ the constraints become redundant due to, respectively, substraction and addition of the large number $LN$.

The following constraint ensures that sample $s$ can only be partitioned into one and only one region:

$$\sum_r a_{s,r} = 1 \quad \forall \quad s \in \mathcal{S} \quad (2.64)$$

The function value corresponding to sample $s$ (i.e. $cc_s$) that is partitioned in region $r$ is predicted through:

$$P_{r,s} = c_s \cdot S_r + I_r \quad \forall \quad s \in \mathcal{S}, r \in \mathcal{R} \quad (2.65)$$

where $\mathcal{R} = \{1, \ldots, R\}$ is the set of regions, $P_{r,s}$ is the predicted output value, $S_r$ is the slope and $I_r$ is the intercept of the fitted linear function in region $r$.

The absolute error between the actual and the predicted value of a sample $s$, i.e. $E_s$, that is partitioned in region $r$ is:

$$E_s \geq cc_s - P_{s,r} - LN(1 - a_{s,r}) \quad \forall \quad s \in \mathcal{S}, r \in \mathcal{R} \quad (2.66)$$

$$E_s \geq P_{s,r} - cc_s - LN(1 - a_{s,r}) \quad \forall \quad s \in \mathcal{S}, r \in \mathcal{R} \quad (2.67)$$

Similar to constraints (2.62) and (2.63), these constraints become redundant when sample $s$ does not lie in region $r$, i.e. when $a_{r,s} = 1$. The substraction of $cc_s$ and $P_{s,r}$ is reversed take into account the absolute deviation.
At last, the objective function is given by:

$$\min \sum_s E_s$$

subject to (2.61)-(2.67).

### 2.5 GAMS Java API

Earlier, it was stated that one of the research objectives of this work is to examine the robustness of the system configurations obtained by solving the deterministic model described in the previous sections. As will be discussed in more depth in Section 3.4, this is done by exposing the optimal network configuration to simulated biomass supply fluctuations. To quantify the robustness with certain confidence, this requires the model to be solved a large number of times. Therefore, it is necessary to have a programming interface that is able to (1) solve the MILP model, (2) sample random numbers according to a specified probability distribution and (3) solve a sequence of closely related models in an efficient way. The General Algebraic Modeling System (GAMS) is hooked up with the high performance linear programming solver CPLEX. Although it may be possible in some way, (2) and (3) are not easily accessible through GAMS. Therefore, the decision has been made to solely program the MILP model in GAMS, while (2) and (3) are carried out externally in Java. This is made possible through a GAMS Java application programming interface (API), which allows a convenient way to exchange model input data and model results.
Chapter 3

Case Studies

The framework presented in the previous chapter is applied in the following case studies:

I Geographical Topologies
II Feedstock Allocation Freedom
III Temporal Feedstock Variability

This chapter is organized as follows: In the first section the model input that is identical for all case studies is presented. In the subsequent sections the various case studies are described and their connection to the research objectives of this work is explained.

3.1 Case Independent Model Input

The model requires for the specification of many parameters. In this section, the parameters that are the same for each case study are given.

Spatial Distributions

Hypothetical biomass supply and demand scenarios are assigned through the specification of rural, semi-rural and urban region types for each grid unit. Each region is characterized by its agricultural land coverage and its population density. The former is derived from an assessment of the United Kingdom’s land-cover database and regressed against the population density. Therefore, these values are representative for the UK and, more generally, for Western Europe [21].

Table 3.1: Agricultural land coverage and population density for three region types.

<table>
<thead>
<tr>
<th>Region type</th>
<th>Agriculture (ha/km²)</th>
<th>Population (Capita/km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>65</td>
<td>75</td>
</tr>
<tr>
<td>Semi-rural</td>
<td>25</td>
<td>300</td>
</tr>
<tr>
<td>Urban</td>
<td>5</td>
<td>1500</td>
</tr>
</tbody>
</table>

Although the size of the hypothetical geographical area of study is varied in some case studies, it is always discretized into 25-by-25 kilometer grid units. The reasoning behind this size is as
follows: firstly, an urban region much larger than 25-by-25 kilometer is less common. Secondly, as discussed in Section 2.3, a certain number of grid units is necessary to allow detailed modeling of the transportation cost. With this grid size, the smallest grid evaluated in this work is $5 \times 5$.

**Biomass Supply**

The production of biomass typically takes place on cropland previously used for other agriculture such as growing food or feed. Since this agricultural production remains necessary, it may be displaced to former non-cropland such as forests or grasslands. This phenomenon is known as indirect land use change (ILUC). Due to the fact that these grasslands and forests typically absorb large amounts of carbon dioxide, there is a risk that the greenhouse gas savings from bio-based supply chains are negated. To reduce the risk of indirect land use change the European Union as well as the United States have set a quota limiting the amount of agricultural land that may be used for biomass sourcing. In this work a 10% availability of agricultural land for biomass sourcing is assumed, thereby approximating the EU quorum.

The feedstock itself is characterized by its harvested yield. In this work, the biomass feedstock is assumed to be a high-yielding energy crop such as *Miscanthus* or *Panicum virgatum*. The latter is more commonly known as switchgrass. The harvested yield is assumed to be 20 dry tonne ha$^{-1}$ year$^{-1}$. Furthermore, a short rotation coppice is considered that can be harvested all-year-round and thus seasonality effects are minimal. The feedstock acquisition cost is assumed to be 40 $ tonne$^{-1}$ [22]. Since only a single biomass type is considered, $B$ is equal to one.

**Bioproduct Demand**

The annual bioproduct demand is assumed to be equal to the amount that can be produced by utilizing 99.99% of the available biomass feedstock of the hypothetical area of study, independent of its size, through the least efficient processing method. The bioproduct demand of a certain grid unit is assumed to be proportional to its population density. This assumption is justified as long as the bioproduct market remains unsaturated. Furthermore, seasonality effects regarding bioproduct demand are assumed to be negligible.

**Processing Technologies**

The techno-economic analyses performed by Wright et al. [22], Tijmensen et al. [23], and Ringer et al. [24], provide data of three reference plants: an integrated processing facility, a preprocessing facility and an upgrading facility. In Table 3.2 below, the capacities and corresponding capital investments as well as their efficiencies are summarized.

<table>
<thead>
<tr>
<th>Facility</th>
<th>Technology</th>
<th>Capacity $^*$</th>
<th>Capital investment</th>
<th>Efficiency $^{**}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated</td>
<td>Biomass gasification + FT synthesis</td>
<td>35 MM GGE / year</td>
<td>$341.00 MM</td>
<td>0.46</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>Biomass fast pyrolysis</td>
<td>200 dry ktonnes / year</td>
<td>$31.64 MM</td>
<td>0.69</td>
</tr>
<tr>
<td>Upgrading</td>
<td>Bio-oil gasification + FT synthesis</td>
<td>35 MM GGE / year</td>
<td>$182.62 MM</td>
<td>0.58</td>
</tr>
</tbody>
</table>

$^*$ GGE stands for gasoline gallon equivalent. 1 GGE = 120.3 MJ.

$^{**}$ In terms of MJ output per MJ input.
It should be noted that processing biomass through the integrated conversion pathway is slightly more efficient compared to the two-step processing strategy. Although the efficiency may also depend on the size of a processing facility, it is assumed that this is a minor effect and therefore it is neglected.

For each of the processing facilities, the scaling equation (see Section ??) is linearized using the MILP regression model presented in Section 2.4. A scale factor \( m \) between 0.6 and 0.8 is common for processes with high pressure and temperature requirements, as is the case here. Therefore, in this work a scale factor of 0.6 is applied (the rule of six-tenths), as in Wright et al. [22]. The minimum capacity of the facilities are equal to zero. The maximum capacity of the integrated processing and upgrading facilities is chosen at 400 MM GGE year\(^{-1}\) and that of the preprocessing facility is equal to 4000 ktonne year\(^{-1}\). This way, either a single integrated processing facility is able to solely process all the biomass and completely satisfy bioproduct demand or, alternatively, one preprocessing facility and one upgrading facility are together able to achieve the same. The number of break-points in the piecewise linear approximation is varied between 0 and 2, meaning that the number of lines fitted to the scaling equation is varied between 1 and 3. The capacity of an integrated processing and upgrading facility is divided into 400 equidistant samples, and the capacity of an preprocessing facility is divided into 4000 samples. In Table 3.3 below, the mean absolute deviation, i.e. the objective function (in Equation (2.68)) corresponding to the optimal solution divided by the number of function samples, is given for all three cases. In addition, the coefficient of determination \( R^2 \) is determined to give information about the goodness of fit:

\[
R^2 = 1 - \frac{\sum_s E_s}{\sum_s (cc_s - \bar{cc})^2}
\]

where \( \bar{cc} \) is the average value of the function samples \( cc_s \).

Table 3.3: Mean absolute error and coefficient of determination for a varying number of break-points for each processing facility.

<table>
<thead>
<tr>
<th>Facility</th>
<th>Break-points</th>
<th>Mean absolute error</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated</td>
<td>0</td>
<td>41.49</td>
<td>0.9558</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13.42</td>
<td>0.9948</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6.62</td>
<td>0.9986</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>0</td>
<td>5.39</td>
<td>0.9690</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1.74</td>
<td>0.9964</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.86</td>
<td>0.9990</td>
</tr>
<tr>
<td>Upgrading</td>
<td>0</td>
<td>22.22</td>
<td>0.9690</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>7.91</td>
<td>0.9964</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.89</td>
<td>0.9989</td>
</tr>
</tbody>
</table>

As indicated by the mean absolute error as well as the coefficient of determination, increasing the number of break-points from zero to one results in the largest increase of the goodness of fit. The introduction of an additional break-point has no significant impact. Each additional break-point requires an additional binary variable per processing facility per grid unit. Therefore, to compromise between computational effort and accuracy, the capital investment for each type of processing facility is modeled by two linear regions, the details of which can be found in Appendix A.3. The scaling equation regarding the capital investment required for an integrated processing facility, as well as the fitted linear functions are depicted in Figure 3.1. As a final remark, it is noted that the largest proportion of the error is located at relatively low capacity. Since it is unlikely that a facility with such capacity is actually constructed, this will probably not have any impact on the outcome of the model.
Figure 3.1: Integrated processing facility capital investment scaling equation (blue) and its linear approximation (red) in case of a single break-point.

So far, only the capital investment associated with the various processing facilities have been considered. As stated before, the operating and maintenance cost associated with the facilities are split into a fixed and a variable part. The former is modeled as a fixed percentage of the total capital investment, while the latter scales linearly with the production quantity. The fixed element takes into account the cost that mainly depend on the size of the facility; e.g. the capital charge, insurance and maintenance cost. The variable part accounts for the cost that depend on the production quantity; e.g. catalyst and energy cost. As for the capital investments, the same approach as in Wright et al. [22], Tijmensen et al. [23], and Ringer et al. [24] is followed. In Table 3.4 an overview of the applied cost factors is given.

<table>
<thead>
<tr>
<th>Facility</th>
<th>Technology</th>
<th>Fixed *</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated</td>
<td>Biomass gasification + FT synthesis</td>
<td>17%</td>
<td>$0.130857 / GGE</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>Biomass fast pyrolysis</td>
<td>12.08%</td>
<td>-$1.485093 / dry tonne</td>
</tr>
<tr>
<td>Upgrading</td>
<td>Bio-oil gasification + FT synthesis</td>
<td>17%</td>
<td>$0.130857 / GGE</td>
</tr>
</tbody>
</table>

* Percentage of the total capital investment.

The fact that the variable production cost of a preprocessing facility is negative is due to its major byproduct, biochar, which has a value as a carbon sequestration agent.

**Storage**

In the developed framework, storage facilities are assumed to be attached to the processing facilities. As mentioned in Chapter 1, a key issue regarding biomass storage is material deterioration, the extent of which depends largely on the type of storage facility. In this work the intermediate option, i.e. covered storage without drying, is considered. Assuming that the maximum inventory height is 6 meters, the investment cost (refer to Table 1.1) is converted to a cubic meter basis. The density of biomass is assumed to be 100 kg m\(^3\), based on *Miscanthus* data.

Contrary to biomass, the significance of bio-oil and -product degradation is minimal. Due to a lack of available data considering the storage of bio-oil, it is assumed that the investment
required for a bio-oil storage facility is identical to that for petroleum crude. Similarly, the assumption is made that the requirements for bioproduct storage are the same as those for petroleum derived products. It is assumed that the operating and maintenance cost amount to 5% of the capital investment for both species. The cost data are summarized in Table 3.5. [25]

Table 3.5: Storage investment cost for bio-oil and -product.

<table>
<thead>
<tr>
<th>Species</th>
<th>Storage investment cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio-oil</td>
<td>404 $ / m³</td>
</tr>
<tr>
<td>Bioproduct</td>
<td>390 $ / m³</td>
</tr>
</tbody>
</table>

The energy densities of bio-oil and -product are assumed to be 19.7 MJ L⁻¹ and 36 MJ L⁻¹ respectively [20]. Furthermore, the energy content of one gasoline gallon equivalent (GGE) is 120.3 MJ/GGE.

Transportation

It is assumed that trucks are the only possibility for transportation of all three species. The distance fixed and variable transportation cost for biomass, -oil and -product are based on data from Searcy et al. [26] and Pootakham et al. [27]. The data is summarized in Table 3.6 below. The moisture content of biomass, \( MC \), is assumed to be 35 wt %. Since only one biomass type is considered the subscript \( b \) is omitted.

Table 3.6: Distance fixed and variable transportation cost for biomass, -oil and -product.

<table>
<thead>
<tr>
<th>Species</th>
<th>Distance fixed cost</th>
<th>Distance variable cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass</td>
<td>4.839 $ / tonne</td>
<td>0.456 $ / tonne / km</td>
</tr>
<tr>
<td>Bio-oil</td>
<td>0.00567 $ / L</td>
<td>0.000119 $ / L / km</td>
</tr>
<tr>
<td>Bioproduct</td>
<td>0.00328 $ / L</td>
<td>0.000425 $ / L / km</td>
</tr>
</tbody>
</table>

Annuity Factor

In this work a project lifetime \( NY \) of 20 years is considered and the interest rate \( IR \) is assumed to be 10%. Thus, the annuity factor is \( (0.1\cdot1.1^{20})/(1.1^{20}−1) \approx 0.1175 \). This factor applies to both the capital investment required for the construction of processing facilities as to that required for the construction of storage facilities.

3.2 Case Study I: Geographical Topologies

This case study is designed to investigate the influence of the underlying spatial distribution on the overall profitability of the supply chain. Two extreme regional topologies are mapped onto the grid to generate two generic spatial distribution scenarios. The center-point distribution ("Paris") has a central urban region surrounded with a peripheral semi-rural region and a rural boundary region. In the corner-point distribution ("London"), the urban region is located in the corner. Similarly, it is surrounded with a semi-rural region and a rural boundary. This way, a hard boundary is imposed on the rural region. In the real world, this may represent a coastline or national border. The size of a grid unit is held constant at 25 x 25 = 625 km². However, the
size of the entire grid is varied between 5 x 5 and 9 x 9, each time increasing rural boundary. The studied scenarios are schematically depicted in 3.2. As mentioned, seasonality effects of the biomass supply are minimal because a short rotation coppice is considered. In addition, seasonal effects of bioproduct demand are assumed to negligible. Hence the time resolution $T$ can be set to one.

Figure 3.2: The center-point (left) and corner-point (right) spatial distributions with increasing rural boundary region. The size of the disks indicates the relative amount of available agricultural land and population density.
3.3 Case Study II: Feedstock Allocation Freedom

As discussed in Section 3.1, the risks associated with indirect land use change has motivated governments to set quota on the amount of agricultural land that may be used for biomass sourcing. In this work a 10% availability of agricultural land for biomass sourcing is assumed. So far, this quota holds for every individual harvesting site \(i\), which is expressed in Equation (2.2). Since the production of biomass requires land, up to a certain extent its geographical dispersion is inherent to its nature. However, imposing this quota to every single harvesting site results in the biomass sources to be even more dispersed. Since the distribution of sources has a profound impact on the annual transportation cost, it is hypothesized that geographically concentrating the biomass sources leads to a reduction of the total annual cost. However, in doing so, the imposed quota must still be respected. In this case study, the model is slightly adjusted such that it is equipped with the freedom to utilize a larger amount of agricultural land for biomass sourcing at one harvesting site while it utilizes less at another, thereby still respecting the imposed overall quota. To achieve this, firstly Equation (2.2) is adjusted to:

\[
\sum_b \phi_{b,i} \leq \Phi \quad \forall \ i \in I
\]

where \(\Phi\) is the maximum fraction of agricultural land allowed to be used for biomass sourcing at each harvesting site. Recall that \(\phi_{b,i}\) is the fraction of agricultural land used for the cultivation of biomass type \(b\) at harvesting site \(i\). Secondly, the following constraint is added to assure that the overall quota remains satisfied:

\[
\sum_i (AL_i \sum_b \phi_{b,i}) \leq \sum_i AL_i \cdot \Phi
\]

Through this mathematical formulation, the parameter \(\Phi\) basically represents the degree of freedom the model has to deviate from the “country wide” quota \(\overline{\Phi}\).

The adjusted model is applied to determine the optimal configuration for the 7 × 7 center-point topology, depicted in Figure 3.2. The threshold value \(\Phi\) is varied between 0.1 and 1. The latter corresponding to the case where the model is allowed to use all of the agricultural land of a grid unit for biomass sourcing. In addition, the effect of another hypothetical crop, distinguished only by its 25% higher annual yield is studied as well. As for the previous case study, the time resolution \(T\) is set to one for both crops.

3.4 Case Study III: Temporal Feedstock Variability

So far, all the obtained solutions are based on the assumption that all parameters are known, i.e. they are deterministic. As mentioned in Chapter 1, uncertainties are of various types and sometimes not easily avoided. Hence it is necessary to examine the robustness of these deterministic solutions towards variabilities. Here, a method to quantify the robustness is proposed and applied.

In this case study, the influence of temporal biomass supply disruptions on a selection of the solutions obtained in the second case study is investigated. The configurations that are selected for this study are the ones obtained at threshold values of \(\Phi = 0.1\) and \(\Phi = 0.4\). At each threshold value, both the optimal and non-optimal processing strategy is examined, yielding a total of four solutions that are selected for this study.

Due to a lack of data regarding biomass supply fluctuations (or agricultural crops in general), it is necessary to make several assumptions. Firstly, it is assumed that biomass is harvested...
once a month. Therefore, in this case the time resolution $T$ is set to 12. Each monthly period
the harvested yield of biomass type $b$ at harvesting site $i$ ($\eta_{b,i}$) may deviate by a certain per-
centage from its average value. As in the deterministic studies, this average value is equal to
20 dry tonne ha$^{-1}$ year$^{-1}$ (= 20/12 dry tonne ha$^{-1}$ month$^{-1}$). The proportionate deviation is
introduced through a minor adjustment of Equation (2.1), yielding:

$$A_{b,i,t} = L_i \cdot \phi_{b,i} \cdot \eta_{b,t} \cdot (1 + \delta_{i,t}) \quad \forall \ b \in B, i \in I, t \in T$$

(3.4)

where $\delta_{i,t}$ is the proportionate deviation incurred at harvesting site $i$ in time period $t$. Several
assumptions are made regarding this deviation percentage $\delta_{i,t}$. Firstly, as already dictated by
this formulation, the deviation incurred at a certain harvesting site is always assumed to be in-
dependent of the deviation at a neighboring harvesting site. Secondly, the assumption is made
that $\delta_{i,t}$ is normally distributed with an average of zero and a certain standard deviation. The
latter basically determines the magnitude of the fluctuations. Since there is no data available
that provides insight in this magnitude, several standard deviations are evaluated in this study
($\sigma = 0.05, 0.10, 0.15$ and $0.20$). For a normal distribution, about 95% of the sampled values are
within two standard deviations ($\pm 2\sigma$).

One option to deal with these supply fluctuations is to make use of storage capacity. Never-
theless, a consequence of the assumption that $\delta_{i,t}$ is normally distributed is that, in theory, it is
possible that the biomass supply is not sufficient to fulfill the bioproduct demand requirement,
thereby making the model infeasible. To prevent the model from becoming infeasible, the model
is given another option, namely to purchase biomass externally. This option is only available
to integrated and preprocessing facilities and is introduced through adding an additional term
to the left-hand side of Equations (2.5), (2.6), (2.16) and (2.17), respectively yielding:

$$\sum_i f_{ijb,i,j,t} + (1 - \varepsilon_b) \cdot sjm_{b,j,t-1} + pp_{jb,j,t} = u_{jb,j,t} + sjm_{b,j,t} \quad \forall \ b \in B, j \in J, t \geq 2$$

(3.5)

$$\sum_i f_{ijb,i,j,t} + (1 - \varepsilon_b) \cdot sjm_{b,j,T} + pp_{jb,j,t} = u_{jb,j,t} + sjm_{b,j,t} \quad \forall \ b \in B, j \in J, t = 1$$

(3.6)

$$\sum_i f_{ikb,i,k,t} + (1 - \varepsilon_b) \cdot skm_{b,k,t-1} + pp_{kb,k,t} = u_{kb,k,t} + skm_{b,k,t} \quad \forall \ b \in B, k \in K, t \geq 2$$

(3.7)

$$\sum_i f_{ikb,i,k,t} + (1 - \varepsilon_b) \cdot skm_{b,k,T} + pp_{kb,k,t} = u_{kb,k,t} + skm_{b,k,t} \quad \forall \ b \in B, k \in K, t = 1$$

(3.8)

where $pp_{jb,j,t}$ and $pp_{kb,k,t}$ are penalty purchases of biomass type $b$ at integrated processing
facility $j$ and preprocessing facility $i$, respectively, in time period $t$. As suggested by its given
name, these purchases are subject to a penalty cost $PC$, which is introduced in the objective
function, Equation (2.51), by adding the following term:

$$C_{pen.} = PC \left( \sum_{b,j,t} pp_{jb,j,t} + \sum_{b,k,t} pp_{kb,k,t} \right)$$

(3.9)

where $C_{pen.}$ is the total annual penalty cost. The assigned penalty cost is set at $150 tonne$^{-1}$.
This way, domestic biomass is always cheaper compared to an external purchase.

To be able to examine the robustness of the selected configurations, the facility locations and
capacities are set based on their optimal values resulting from the deterministic solution. Ad-
ditionally, in the previous case study, the fraction of the available agricultural land used for
sourcing of biomass type $b$ at harvesting site $i$ ($\phi_{b,i}$) has also been optimized. Therefore, also these values are set accordingly. The values of $\delta_{i,t}$ are sampled from the predefined normal distribution. Subsequently, the model is solved. Using the GAMS Java API (recall Section 2.5), the process of random sampling and solving the model is repeated numerous times. Each time, the complete cost picture will be different, the extent of which is a measure of the robustness.
Chapter 4

Results and Discussion

In this chapter the results are presented and discussed. This is done for each of the three case studies separately.

4.1 Model Statistics

All computational studies in this work were performed with an Intel Core i7-4700MQ 2.40 GHz CPU and 8 GB RAM. The MILP was programmed in GAMS version 24.4.6 and solved with the CPLEX 12.6.2.0 solver. The optimality tolerances were always set to 0.00%. Since the size of the model is not constant throughout this work, a complete overview of each model is omitted, however, to provide some insight in the size of the problems in this work, the model statistics and solution report corresponding to the largest model solved in this work (the $9 \times 9$ grid) are listed in Table 4.1 below.

Table 4.1: Model statistics and solution report corresponding to the largest ($9 \times 9$) model.

<table>
<thead>
<tr>
<th>Model Statistics</th>
<th>Solution Report</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model type</td>
<td>MILP</td>
</tr>
<tr>
<td>Single variables</td>
<td>35000 Absolute gap 0.00 %</td>
</tr>
<tr>
<td>Discrete variables</td>
<td>486 Relative gap 0.00 %</td>
</tr>
<tr>
<td>Equations</td>
<td>3253 Solving time 3624.11 sec.</td>
</tr>
</tbody>
</table>

4.2 Case Study I: Geographical Topologies

In Figure 4.1, the optimal supply chain configuration for each scenario is schematically depicted. Refer to Figure 3.2 for the underlying bioproduct demand distribution. Several performance measures are used to analyze the optimal system configurations. Since the amount of bioproduct demand for each scenario is chosen such that 99.99% of the biomass supply is utilized, the amount of processed biomass increases with an increasing rural boundary. As a consequence, the value of the objective (i.e. the total annual cost) increases accordingly and therefore does not provide much insight in the overall profitability of the supply chain. For that reason, the total annual cost is divided by the total annual bioproduct production, yielding the cost per product unit. Furthermore, the average distance over which biomass, -oil and -product are transported ($D_B$, $D_O$ and $D_B$ resp.) is evaluated. These performance measures, including a number of other metrics, are reported in Table 4.2. Regardless of which processing strategy is
most profitable, both are evaluated here. In Figure 4.2 the various cost fractions are reported for each scenario.

Figure 4.1: The optimal system configuration for each scenario. Frames indicate the serving areas of the facilities. The size of the disks represents the relative size of the processing facilities. Refer to Figure 4.1 for the underlying bioproduct demand distribution. Numerical results can be found in Appendix C.

For all scenarios, preprocessing followed by upgrading appears to be the most profitable strategy. Firstly, it is noticed that with an increasing grid size, the product cost decreases in a non-linear fashion. This is observed in case of both processing strategies. As the rural bound-
ary expands, more biomass is processed, as is dictated by the increased bioproduct demand requirement (recall that the bioproduct demand is equal to the amount that can be produced by utilizing 99.99% of the total feedstock via the least efficient processing pathway). The obvious result is that more processing capacity is required. It can be seen that independent of the size of the rural area, only a single integrated processing or upgrading facility is employed. This means that with each expansion, the processing cost (annual capital cost plus O&M cost) per unit of bioproduct is reduced due to economies of scale. Due to the higher concentration of processes, this effect is strongest for the integrated processing pathway, as can be seen in Figure 4.2. If this were the only effect, a (roughly) linear decrease of the product cost would be observed. However, the decrease appears to weaken as the grid becomes larger. This is mainly due to the inevitable increase of the transportation cost. In case of integrated processing, the additional biomass that becomes available needs to be transported over a larger distance and therefore, a relatively larger cost is incurred. In other words, the transportation cost experiences diseconomies of scale. The alternative pathway attempts to counter this effect by employing multiple preprocessing facilities, as can be seen in Figure 4.1. This way, distant biomass is converted into bio-oil which can be transported more cost efficiently. The result is reflected in the lower product cost associated with this processing strategy. However, whether it is biomass, -oil or -product that is to be transported, an increase of the transportation cost is unavoidable. This effect is most clearly illustrated by comparing the transportation cost fractions in Figure 4.2 and can also be observed by examining the average distances over which the various species are transported (see Table 4.2). The distance over which biomass is transported remains more or less constant, while bio-oil and -product are transported over larger distances.

Table 4.2: Performance measures for the two topologies and both processing strategies.

<table>
<thead>
<tr>
<th>Grid size (km²)</th>
<th>125²</th>
<th>175²</th>
<th>225²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Typology</strong></td>
<td><strong>Processing pathway</strong></td>
<td><strong>Metric</strong></td>
<td><strong>Unit</strong></td>
</tr>
<tr>
<td><strong>Preprocessing &amp; upgrading</strong></td>
<td><strong>Total annual cost</strong></td>
<td>$MM</td>
<td>303.88</td>
</tr>
<tr>
<td><strong>Product cost</strong></td>
<td>$/GGE</td>
<td>3.14</td>
<td>2.84</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>km</td>
<td>30.21</td>
<td>38.49</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>km</td>
<td>18.69</td>
<td>26.35</td>
</tr>
<tr>
<td><strong>No. int. processing facilities</strong></td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Preprocessing &amp; upgrading</strong></td>
<td><strong>Total annual cost</strong></td>
<td>$MM</td>
<td>314.67</td>
</tr>
<tr>
<td><strong>Product cost</strong></td>
<td>$/GGE</td>
<td>3.12</td>
<td>2.81</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>km</td>
<td>27.19</td>
<td>34.80</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>km</td>
<td>29.87</td>
<td>42.45</td>
</tr>
<tr>
<td><strong>No. int. processing facilities</strong></td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Preprocessing &amp; upgrading</strong></td>
<td><strong>Total annual cost</strong></td>
<td>$MM</td>
<td>302.67</td>
</tr>
<tr>
<td><strong>Product cost</strong></td>
<td>$/GGE</td>
<td>3.00</td>
<td>2.78</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>km</td>
<td>17.91</td>
<td>18.21</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>km</td>
<td>29.87</td>
<td>36.28</td>
</tr>
<tr>
<td><strong>No. preprocessing facilities</strong></td>
<td>-</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td><strong>No. upgrading facilities</strong></td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

* The total annual cost divided by the total amount of satisfied demand.
** The average biomass transportation distance.

Apart from justifying the observed trend, it is important to note that eventually the benefit from economies of scale will be outweighed by the diseconomies of scale associated with the trans-
portation cost. At that point, multiple integrated processing or upgrading facilities need to be constructed in order to prevent the product cost from rising. Based on these observations, this point is expected to occur at grid sizes not much larger than the ones evaluated here. Basically, the preprocessing facilities deal with a similar trade-off. Especially by examining the complicated configurations in Figure 4.1 (c) and (d) it can be seen that a model is definitely required to determine the optimal bargain.

So far, no attention is paid to the influence of the location of the urban region. By solely looking at the difference in total annual cost between the two topologies (see Table 4.2), it can be seen that the difference is extremely small. In case of preprocessing followed by upgrading there is a slight preference towards the center-point design. This is due to the fact that by locating the upgrading facility in the center, the bio-oil and -product transportation distance are both at their minimum, something that can never achieved with a single facility in case of the corner-point design.

On the contrary, the integrated processing strategy shows a slight preference towards the corner-point design. As the rural boundary increases, the location of the integrated processing facility changes in a similar manner as the upgrading facility, as shown in Figure 4.1 (b), (d) and (f). By locating the facility in a more rural region, the biomass transportation distance and cost are reduced.

Figure 4.2: Capital, O&M, biomass acquisition and transportation cost fractions for (a) center-point with preprocessing and upgrading, (b) center-point with integrated processing, (c) corner-point with preprocessing and upgrading and (d) corner-point with integrated processing.
4.3 Case Study II: Feedstock Allocation Freedom

First, the results regarding the energy crop with an annual harvested yield of 20 tonne ha\(^{-1}\) year\(^{-1}\) are presented. The total annual cost corresponding to the optimal system configuration as a function of the threshold value \(\Phi\) is shown for both processing strategies in Figure 4.3 below.

![Figure 4.3: Total annual cost for both processing strategies versus the biomass feedstock allocation freedom, which is characterized by the threshold value \(\Phi\).](image)

As expected, the total annual cost clearly decreases. At low threshold values, converting biomass through preprocessing and upgrading is clearly the preferred strategy to employ; as expected from the results of the previous case study. However, at a threshold value of roughly 0.3, integrated processing becomes more profitable. As the threshold values further increases, the difference in total annual cost between both processing pathways becomes less pronounced, until eventually the lines coincide. Also, it is noted that the annual cost associated with integrated processing decreases in a smoother manner when compared to its competitor, an issue that will be examined shortly.

Without yet examining the exact system configuration, the decreasing trend already suggests that indeed advantage is taken from the given freedom to concentrate biomass sources. The optimal system configurations for \(\Phi = 0.2\), a case when preprocessing followed by upgrading is the most profitable strategy, and for \(\Phi = 0.4\), when integrated processing is preferred, are schematically depicted in Figure 4.4. Refer to Figure 3.2 (c) for the underlying bioproduct demand distribution.

As can be seen, in both cases biomass sources are concentrated at close proximity of processing facilities, leaving agricultural land located at larger distances unutilized. The motivation for this clustering of sources is of course a decrease in the average transportation distance. However, the reason that integrated processing becomes more beneficial at certain threshold values is not immediately obvious. To understand the drivers behind these observations, it is necessary to take a look at the cost break down. In Figure 4.5 the absolute capital, O&M, biomass acquisition and transportation cost associated with the optimal solution are shown as a function of the threshold value \(\Phi\).
Starting at a threshold value of 0.1, it can be seen that the previously observed decrease of the total annual cost can be designated primarily to the decrease of the transportation cost. At a threshold value of approximately 0.3 a pronounced change in the proportionality of the costs is observed, reflecting the point at which integrated processing becomes more profitable (recall Figure 4.3). The reason that the capital, O&M and biomass acquisition cost appear to reach a constant level is due to the fact that only a single integrated processing facility is employed. It appears that this change of optimal processing strategy is driven by a decrease of the O&M cost and the biomass acquisition cost, which are both (at least partially) proportional to the amount of biomass that is converted. Therefore, this change can be explained by the fact that
an integrated processing facility is more efficient, i.e. it requires less biomass to produce the same amount of product. Also, it can be seen that the transportation cost slightly increases. This suggests that the biomass is now transported to a fewer number of locations, making the average transportation distance longer. Although the transportation cost and the capital investment required for the construction of such a facility are significantly higher, apparently this does not outweigh a decrease of other cost components. More importantly, it does not outweigh the cost that would be incurred in case of the other processing pathway.

Up to a threshold value of approximately 0.85, the transportation cost is the only driver behind the further decrease of the total annual cost. The reason that the other costs remain constant is due to the fact that only a single integrated processing facility is employed. Onwards to higher threshold values, three more significant changes in the proportionality of the costs are observed. This indicates that the difference between the two competitors is more subtle, which is in accordance with the difference between their total annual cost. To illustrate this, in Figure 4.6 the optimal configurations for both strategies at the extreme value of $\Phi = 1$ are schematically depicted.

Figure 4.6: Optimal system configuration for (a) preprocessing followed by upgrading and (b) integrated processing at $\Phi = 1$. Frames indicate the serving areas of the facilities. The size of the disks represents the relative size of the processing facilities. Refer to Figure 3.2 (c) for the underlying bioproduct demand distribution. Numerical results can be found in Appendix D.

As can be seen, the designs are very similar. In both cases the biomass sources are located at the closest proximity of the processing facilities. The higher efficiency of the integrated processing facility is demonstrated by the smaller area required for biomass sourcing. On the other hand, in case of preprocessing and upgrading, transportation cost are kept minimal by constructing four preprocessing facilities to serve only five areas. Given the subtle difference between the two, a more detailed investigation is advisable.

So far, the decrease of the transportation cost, resulting from the concentration of biomass sources, is identified as the most important driver behind the decreasing total annual cost. It is important to note that the decrease of the transportation cost is non-linear. This is due to the transportation cost being proportional to the distance as well as the amount of biomass that is
transported. As the area that is used for biomass sourcing reduces, the distance decreases linearly. However, the area decreases with a power of two, thereby explaining the steep decrease at the beginning and the more constant descent at higher values of $\Phi$ (in Figures 4.3 and 4.5).

As mentioned before, the costs associated with preprocessing and upgrading fluctuate significantly compared to the costs associated with its competitor, which are either constant or decrease in a smooth manner. This is due to the fact that in case of integrated processing always a single facility is constructed. On the contrary, in case of the other conversion pathway, it is common to employ multiple preprocessing facilities and a single upgrading facility. In Figure 4.7 the number of preprocessing facilities as a function of the threshold value $\Phi$ is shown.

![Figure 4.7: The number of preprocessing facilities as a function of the threshold value $\Phi$.](image)

The number of preprocessing facilities varies between a minimum of three and a maximum of nine. Especially the steep increase in the number of processing facilities around $\Phi = 0.45$ is surprising at first sight. To gain insight in the motivation behind these fluctuations, it is unavoidable to examine the cost breakdown. Therefore, in Figure 4.8 the various cost components are depicted as a function of the threshold value $\Phi$ in case that preprocessing and upgrading is the only conversion pathway.

![Figure 4.8: Cost breakdown associated with the optimal system configuration as a function of the threshold value $\Phi$.](image)
The benefit of a more distributed design with more preprocessing facilities is lower transportation cost. On the other hand, a more centralized design with less facilities results in higher transportation cost while the capital and O&M cost are lower. As can be seen, all fluctuations are caused by this continuous trade-off. The large number of changes indicates that this trade-off is subtle and demonstrates the necessity of a model to accurately determine the cost optimal configuration. The discreteness of the fluctuations is caused by both the fact that a facility is a discrete object as the discrete nature of model.

Up to this point, only the energy crop with the lowest annual harvested yield is considered. In Figure 4.9 the total annual cost corresponding to the optimal configuration as a function of the threshold value $\Phi$ is shown for both processing strategies for a hypothetical crop with a 25% higher harvested yield. The dashed lines are identical to the previously depicted total annual cost functions for the case with the lower yielding crop in Figure 4.3.

![Figure 4.9: Total annual cost for both processing strategies versus the biomass feedstock allocation freedom, characterized by the threshold value $\Phi$, for a hypothetical energy crop with a 25% higher annual harvested yield. Dashed lines are identical to total annual cost functions depicted in Figure 4.3.](image)

The first noticeable difference is that the total annual cost starts off at a lower value. This is due to the fact that with this higher yielding crop, already at a threshold value of 0.1, not all of the agricultural land is necessarily utilized. By not using the outer most area, the transportation cost is reduced. In other words, the higher yield basically gives the system a head start. Secondly, it is noticed that on average the decrease of the total annual cost is stronger. This can be explained as follows; as the threshold value increases, more agricultural land in the close proximity of processing is allowed to be used. Driven by a reduction of the transportation cost, the model chooses to utilize this land instead of more distant areas. For a higher harvested yield, the actual amount of biomass that is moved to a closer location is larger. This translates itself to a stronger decrease of the transportation cost. A consequence of these two effects is that already at a threshold value of around 0.22, integrated processing becomes more profitable than its competitor.

### 4.4 Case Study III: Temporal Feedstock Variability

As stated in Section 3.4, the focus of this case study is on four of the deterministic solutions obtained previously. In Figure 4.10 the selected configurations are schematically depicted. The reason that these four configurations are selected is that at a threshold value of $\Phi = 0.1$ the
preference towards the preprocessing followed by upgrading strategy is at its maximum, while at a threshold value of roughly $\Phi = 0.4$ the preference for integrated processing is strongest (recall Figure 4.3 and 4.9). At least, this is true in a completely deterministic world. Here, it is investigated to what extent this remains valid when these configurations are subject to quantitative biomass supply fluctuations.

\[ \Phi = 0.1 \quad \Phi = 0.4 \]

Figure 4.10: The selection of optimal system configurations for this study. Frames indicate the serving areas of the facilities. The size of the disks represents the relative size of the processing facilities. Refer to Figure 4.1 (c) for the underlying bioproduct demand distribution. Numerical results can be found in Appendix E.

As the location of processing facilities and their capacity are set prior to solving the model, the annual capital cost and O&M cost will not experience any variations due to the imposed biomass supply fluctuations while all the other cost components do. Therefore, the costs that may vary due to the influence of supply disruptions are distinguished from those that do not. As stated in Section 3.4, the proportionate deviation $\delta_{i,t}$ is sampled from a normal distribution; a two parameter continuous probability distribution. The parameter $\mu$ (the average) is always zero. The parameter $\sigma$ (the standard deviation) is varied in order to gain insight in the effect
of the magnitude of the biomass supply deviations. For each value of $\sigma$ (0.05, 0.10, 0.15 and 0.20), the configurations are subjected to 70 randomly generated fluctuation scenarios. To provide some insight in the dynamics of the obtained results, as an example, the summation of the variable annual cost components is depicted in Figure 4.11 for each simulated year. The presented results correspond to the integrated processing configuration at a threshold value of $\Phi = 0.4$ (see Figure 4.10 (d)).

Despite the irregularities, it can be seen that as the standard deviation of the sampling distribution increases, the variable cost components appear to increase and fluctuate more intensively. In order to quantify this behaviour, the mean and standard deviation are determined for each of the graphs in Figure 4.11. The results are summarized in Figure 4.12.
As can be seen, the aforementioned observation is confirmed; both the average and standard deviation of the variable cost components increase as the probability on more intense deviations becomes higher. Due to the imposed supply fluctuations, the system is frequently forced to acquire more distant domestic biomass, make use of storage capacity or purchase additional biomass externally. Either way, a higher cost is incurred. A similar analysis is performed for the other three system configurations, the raw data of which is postponed to Appendix E. The results are listed in Table 4.3. In addition, the absolute change (compared to the deterministic solution) of each variable cost component for the case where the fluctuations are most intense is given as well.

Table 4.3: Performance measures for all of the system configuration under investigation.

<table>
<thead>
<tr>
<th>System configuration</th>
<th>ΔCost ($MM)*</th>
<th>Processing pathway</th>
<th>Φ</th>
<th>Acquisition**</th>
<th>Transport</th>
<th>Storage</th>
<th>Penalty</th>
<th>Total</th>
<th>σ ($MM)***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preprocessing &amp; upgrading</td>
<td>-1.8</td>
<td>0.1</td>
<td>-1.6</td>
<td>0.1</td>
<td>6.8</td>
<td>2.6</td>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.5</td>
<td>0.4</td>
<td>-1.0</td>
<td>1.1</td>
<td>9.6</td>
<td>7.2</td>
<td>3.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated processing</td>
<td>-0.9</td>
<td>0.1</td>
<td>-1.9</td>
<td>0.2</td>
<td>3.3</td>
<td>0.7</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.9</td>
<td>0.4</td>
<td>-0.9</td>
<td>1.0</td>
<td>2.7</td>
<td>2.1</td>
<td>2.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Difference between the solutions obtained where the width of the sampling distribution is at its maximum and the deterministic solutions.
** Acquisition of domestic biomass.
*** Standard deviation of the total variable cost components in case of the most intense fluctuations.

For all supply chain configurations the biomass acquisition cost appears to decrease, meaning that less biomass is acquired domestically. This is also reflected in the decrease of the total transportation cost. Nevertheless, due to the demand requirement, an equal quantity must be purchased externally at a penalty cost of 150 $ tonne$^{-1}$, while the acquisition cost of domestic biomass is only 40 $ tonne$^{-1}$. Because of this significant cost difference, the system attempts to avert this penalty cost. This can be achieved by employing storage facilities. However, there are several disadvantages regarding storage. Firstly, the investment required for its construction depends on the maximum inventory level, recall Equation (2.38). Therefore, given the supply disruptions, these facilities may be completely unutilized in some time periods. Secondly, since the capacities of the processing facilities are set based to the deterministic solution, the only species that can be stored is biomass, which exhibits the disadvantage of material degradation. Thus, both options to deal with the imposed supply fluctuations have negative aspects and the
given solutions depend largely on the cost associated with both strategies. Despite the decrease of the biomass acquisition and transportation cost, it is inevitable that the total annual cost increases. This cost increase goes hand in hand with an increase of the expected annual cost fluctuation (characterized by the standard deviation), meaning that the system becomes less robust. Nonetheless, the presence of random quantitative supply fluctuations does not have enough impact to reconsider the observations from the previous case study, i.e. preprocessing followed by upgrading remains the most profitable strategy at a threshold value of $\Phi = 0.1$ and integrated processing remains preferred at $\Phi = 0.4$. Although this is true for these specific threshold values, the preference towards a certain processing strategy is not always as strong, recall Figure 4.3 and 4.9. Overall, it appears that integrated processing is more robust compared to its competitor. This can be explained by the concept known as variability pooling, an everyday example of which is financial planning. Typically, financial advisers recommend investing in a diverse portfolio of financial instruments. Obviously, the reason is to hedge against risk. The probability that a large number of investments will perform extremely poor at the same time is very low. Similarly, it is very unlikely that they will all perform very well at the same time. Hence, well performing investments will compensate for low performing ones and as a consequence the relative variation decreases. The application of this concept on the supply chain configurations is best illustrated on the basis of Figure 4.10. For both values of $\Phi$, the number of harvesting sites that “serve” an integrated processing facility is much larger compared to number of sites that serve a single preprocessing facility. Therefore, the total harvested yield for an integrated processing facility is relatively much more stable compared to a single preprocessing facility.

In a similar manner, this concept also partly justifies the difference in robustness observed between the cases with more concentrated biomass sources ($\Phi = 0.4$) and those where the sources are more dispersed ($\Phi = 0.1$). However, there is another aspect that plays an important role as well. The fact that the deviation is proportionate means that it has a larger impact on grid units where a large amount of feedstock is concentrated. Therefore, already for the narrowest sampling distribution, the systems where $\Phi = 0.4$ experience more intense supply fluctuations per unit compared to the systems for which $\Phi = 0.1$. 

45 Design of Bio-Based Supply Chains / Final Version
Chapter 5

Summary and Conclusions

The developed modeling framework captures the main characteristics of bio-based supply chains. The objective of this model is to determine the optimal supply chain configuration such that total annual cost is minimized while at the same time bioproduct demand is fulfilled. A correction factor was introduced to assure accurate evaluation of the transportation cost. Furthermore, the spatial explicitness and generic formulation allow application on a wide spectrum of bio-based supply chain problems. Here, the model was applied to three hypothetical case studies to address the research objectives of this study.

Two pathways for the conversion of biomass into bio-derived products were considered: (i) preprocessing followed by upgrading and (ii) integrated processing. Hypothetical supply and demand scenarios were generated through the specification of rural, semi-rural and urban regions. To account for the risks associated with indirect land use change, a quotient of 10% was set on the amount of agricultural land that may be used for biomass sourcing.

The first case study was designed to gain insight in the effect of the underlying spatial distribution on the profitability of the supply chain. Preprocessing through multiple distributed fast pyrolysis facilities followed by centralized upgrading was identified as the most profitable strategy for all scenarios. It can be concluded that the influence of the geographical topology is small in case of integrated processing and even smaller for its competitor. Increasing the size covered by the supply chain has a significant positive effect on the product cost. Though, at a certain point, it is outweighed by the transportation cost. The model provides a valuable tool to determine the optimal area that the supply chain must cover, the exact size of which is influenced by multiple factors. In the second case study the model was adjusted to allow local deviation from the set quotient, while it remains overall satisfied. It was concluded that through the geographical concentration of biomass sources, the profitability of the supply chain can be increased enormously. Additionally, as sources are clustered, preprocessing followed by upgrading is not perforce the preferred processing strategy. However, to achieve this in the real world, there is a necessity for governmental regulations and/or incentives that provide direction regarding the agricultural lands that are dedicated for biomass cultivation. In the last case study, a simulation-based method was developed to study the system dynamics and to quantify the robustness of certain supply chain configurations towards temporal feedstock variability. The deterministic solutions seemed robust towards quantitative supply fluctuations. However, clear differences were observed: Integrated processing proved to be more robust compared to its competitor. Furthermore, scenarios with a geographically dispersed biomass supply were shown to be more robust compared to situations where biomass sources are clustered. Thus, it can be concluded that clustering of biomass sources has a positive effect on the profitability of the supply chain, however, it has a slight adverse effect on the robustness of the system.
Chapter 6

Recommendations

The recommendations for further research that follow from this work can be summarized as follows:

• The developed stochastic method can easily be extended to take into account multiple sources of variability. Two examples are:
  – Bioproduct demand fluctuations.
  – Feedstock quality deviations can be incorporated through the conversion factor.

It is anticipated that these may amplify each other.

• The possibility to exceed process capacity can be included via a penalty cost. This provides additional operational flexibility that may be beneficial in time periods with a feedstock surplus.

• In this study the deterministic solution was based on the reasonable assumption that biomass supply is (nearly) constant over time. However, the model can also be applied to a situation where the supply is deterministic but time dependent. It is anticipated that the solutions obtained under such circumstances are more robust.
Nomenclature

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FT</td>
<td>Fischer Tropsch</td>
</tr>
<tr>
<td>MILP</td>
<td>Mixed integer linear program</td>
</tr>
<tr>
<td>GAMS</td>
<td>General Algebraic Modeling System</td>
</tr>
<tr>
<td>GGE</td>
<td>Gasoline gallon equivalent</td>
</tr>
<tr>
<td>GHG</td>
<td>Greenhouse gas</td>
</tr>
<tr>
<td>O&amp;M</td>
<td>Operating and maintenance</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>USD</td>
<td>United States dollar</td>
</tr>
</tbody>
</table>

Main Model

Sets

<table>
<thead>
<tr>
<th>Set</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Biomass types</td>
</tr>
<tr>
<td>I</td>
<td>Harvesting sites</td>
</tr>
<tr>
<td>J</td>
<td>Integrated processing facilities</td>
</tr>
<tr>
<td>K</td>
<td>Preprocessing facilities</td>
</tr>
<tr>
<td>L</td>
<td>Upgrading facilities</td>
</tr>
<tr>
<td>N</td>
<td>Demand sites</td>
</tr>
<tr>
<td>R</td>
<td>Capacity levels</td>
</tr>
<tr>
<td>T</td>
<td>Time periods</td>
</tr>
</tbody>
</table>

Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$cc_j$</td>
<td>Total capital investment integrated processing facility $j$ USD</td>
</tr>
<tr>
<td>$ck_k$</td>
<td>Total capital investment preprocessing facility $k$ USD</td>
</tr>
<tr>
<td>$ccl_l$</td>
<td>Total capital investment upgrading facility $l$ USD</td>
</tr>
<tr>
<td>$ccsm$</td>
<td>Total capital investment biomass storage capacity USD</td>
</tr>
<tr>
<td>$ccso$</td>
<td>Total capital investment bio-oil storage capacity USD</td>
</tr>
<tr>
<td>$ccsp$</td>
<td>Total capital investment bio-product storage capacity USD</td>
</tr>
<tr>
<td>$cj_{j,r}$</td>
<td>Annual production capacity integrated processing facility $j$ with capacity level $r$ GGE</td>
</tr>
<tr>
<td>$ck_{k,r}$</td>
<td>Annual production capacity preprocessing facility $k$ with capacity level $r$ tonne</td>
</tr>
<tr>
<td>$cl_{l,r}$</td>
<td>Annual production capacity upgrading facility $l$ with capacity level $r$ GGE</td>
</tr>
<tr>
<td>$cfj_j$</td>
<td>Fixed annual O&amp;M cost for integrated processing facility $j$ USD</td>
</tr>
<tr>
<td>$cfk_k$</td>
<td>Fixed annual O&amp;M cost for preprocessing facility $k$ USD</td>
</tr>
<tr>
<td>$cfj_l$</td>
<td>Fixed annual O&amp;M cost for upgrading facility $l$ USD</td>
</tr>
<tr>
<td>$cfsm$</td>
<td>Annual O&amp;M cost for biomass storage capacity USD</td>
</tr>
<tr>
<td>$cfso$</td>
<td>Annual O&amp;M cost for bio-oil storage capacity USD</td>
</tr>
<tr>
<td>$cfsp$</td>
<td>Annual O&amp;M cost for bio-product storage capacity USD</td>
</tr>
<tr>
<td>$csjm_j$</td>
<td>Biomass storage capacity integrated processing facility $j$ m$^3$</td>
</tr>
<tr>
<td>$csjp_j$</td>
<td>Bio-product storage capacity integrated processing facility $j$ m$^3$</td>
</tr>
<tr>
<td>$cskm_k$</td>
<td>Biomass storage capacity preprocessing facility $k$ m$^3$</td>
</tr>
<tr>
<td>$csko_k$</td>
<td>Bio-oil storage capacity preprocessing facility $k$ m$^3$</td>
</tr>
<tr>
<td>$cslo_l$</td>
<td>Bio-oil storage capacity upgrading facility $l$ m$^3$</td>
</tr>
<tr>
<td>$cslp_l$</td>
<td>Bio-product storage capacity upgrading facility $l$ m$^3$</td>
</tr>
<tr>
<td>$cvj_j$</td>
<td>Variable annual O&amp;M cost for integrated processing facility $j$ USD</td>
</tr>
<tr>
<td>$cvk_k$</td>
<td>Variable annual O&amp;M cost for preprocessing facility $k$ USD</td>
</tr>
<tr>
<td>$cvl_l$</td>
<td>Variable annual O&amp;M cost for upgrading facility $l$ USD</td>
</tr>
</tbody>
</table>
Binary Variables

\[ x_{j,r} \] Equal to 1 if an integrated processing facility with capacity \( r \) is placed at site \( j \), 0 otherwise.
\[ y_{k,r} \] Equal to 1 if a preprocessing facility with capacity \( r \) is placed at site \( k \), 0 otherwise.
\[ z_{l,t} \] Equal to 1 if an upgrading facility with capacity \( r \) is placed at site \( l \), 0 otherwise.

Parameters

- \( \alpha_b \) Conversion factor of biomass type \( b \) at an integrated processing facility
- \( \beta_b \) Conversion factor of biomass type \( b \) at a preprocessing facility
- \( \gamma \) Conversion factor of bio-oil at an upgrading facility
- \( \delta_{i,t} \) Proportionate harvested yield deviation
- \( \varepsilon_b \) Biomass type \( b \) deterioration percentage
- \( \varphi_{i,t} \) Harvested yield of biomass type \( b \) in time period \( t \) tonne ha\(^{-1}\)
- \( \varphi'' \) Minimum utilization percentage integrated processing facility
- \( \varphi''' \) Minimum utilization percentage preprocessing facility
- \( \rho_b \) Biomass type \( b \) volumetric density tonne m\(^3\)
- \( \phi_{b,i} \) Fractional land usage for biomass type \( b \) at harvesting site \( i \)
- \( \Phi \) Maximum fractional land usage
- \( A_{b,i,t} \) Availability of biomass type \( b \) at harvesting site \( i \) in time period \( t \) tonne
- \( A_{b,h} \) Biomass type \( b \) acquisition cost USD
- \( A_{D} \) Average distance to the center of a grid unit km
- \( A_{L} \) Available land at harvesting site \( i \) ha
- \( \text{CC} J_{r} \) Upper bound total capital investment for an integrated processing facility with capacity level \( r \) USD
- \( \text{CC} K_{r} \) Upper bound total capital investment for an preprocessing facility with capacity level \( r \) USD
- \( \text{CC} L_{r} \) Upper bound total capital investment for an upgrading facility with capacity level \( r \) USD
- \( \text{CC} S \) Biomass storage capacity investment cost USD m\(^3\)
- \( \text{CC} SO \) Bio-oil storage capacity investment cost USD m\(^3\)
- \( \text{CC} S \) Bioproduct storage capacity investment cost USD m\(^3\)
- \( \text{CC} S \) Bioproduct storage capacity investment cost USD m\(^3\)
- \( \text{CC} S \) Upper bound annual production capacity for an integrated processing facility with capacity level \( r \) tonne
- \( \text{CC} S \) Upper bound annual production capacity for an preprocessing facility with capacity level \( r \) tonne
- \( \text{CC} S \) Upper bound annual production capacity for an upgrading facility with capacity level \( r \) tonne
- \( \text{CV} J \) Fixed annual O&M cost as a percentage of the capital investment for an integrated processing facility USD GGE\(^{-1}\)
- \( \text{CV} K \) Fixed annual O&M cost as a percentage of the capital investment for a preprocessing facility USD GGE\(^{-1}\)
- \( \text{CV} L \) Fixed annual O&M cost as a percentage of the capital investment for an upgrading facility USD GGE\(^{-1}\)
- \( D \) Distance between two grid unit center points km
- \( \text{DFCM} \) Distance fixed transportation cost biomass USD tonne\(^1\)
- \( \text{DFCM} \) Distance fixed transportation cost bio-oil USD GGE\(^{-1}\)
- \( \text{DFCM} \) Distance fixed transportation cost bioproduct USD GGE\(^{-1}\)
- \( \text{DFCM} \) Distance variable transportation cost biomass USD tonne\(^1\) km\(^{-1}\)
- \( \text{DFCM} \) Distance variable transportation cost bio-oil USD GGE\(^{-1}\) km\(^{-1}\)
- \( \text{DFCM} \) Distance variable transportation cost bioproduct USD GGE\(^{-1}\) km\(^{-1}\)
MILP Regression Model

Sets

\( S \) \hspace{1cm} \text{Samples} \\
\( R \) \hspace{1cm} \text{Regions}

Variables

\( C_r \) \hspace{1cm} \text{Upper bound input values region } r \\
\( E_s \) \hspace{1cm} \text{Absolute error between actual and predicted output value of sample } s \\
\( I_r \) \hspace{1cm} \text{Intercept in region } r \\
\( P_{r,s} \) \hspace{1cm} \text{Predicted output value in region } r \text{ for sample } s \\
\( S_r \) \hspace{1cm} \text{Slope in region } r

Binary Variables

\( a_{s,r} \) \hspace{1cm} \text{Equal to 1 if sample } s \text{ is partitioned in region } r, \text{ 0 otherwise}

Parameters

\( cc_s \) \hspace{1cm} \text{Actual input value of sample } s \\
\( cs_s \) \hspace{1cm} \text{Actual output value of sample } s \\
\( LN \) \hspace{1cm} \text{Arbitrary large number}
Bibliography


[15] Robert P. Anex; Andy Aden; Feroz Kabir Kazi; Joshua Fortman; Ryan M. Swanson; Mark M. Wright; Justinus A. Satrio; Robert C. Brown; Daren E. Daugaard; Alex Platon; Geetha Kothandaraman; David D. Hsu; Abhijit Dutta. Techno-economic comparison of biomass-to-transportation fuels via pyrolysis, gasification, and biochemical pathways. *Fuel*, 89:S29–S35, 2010. 5


Appendices
Appendix A

Miscellaneous Derivations

A.1 Annuity Factor

In Equations (2.53) and (2.54) the annuity factor is used to annualize the capital investments associated with processing and storage facilities respectively. There are multiple methods to derive this annuity factor and in this section one such method is briefly presented.

Considering a loan in general, its balance after \( n+1 \) time periods \( (B_{n+1}) \) is related to its balance after \( n \) time periods \( (B_n) \) through the following recurrence relation:

\[
B_{n+1} = B_n(1 + IR) - A_n \tag{A.1}
\]

where \( IR \) is an interest rate and \( A_n \) is the repayment made at the end of the \( n^{th} \) period. This nonhomogeneous recurrence relation has the following general solution:

\[
B_n = C(1 + IR)^n \tag{A.2}
\]

where \( C \) is a constant. In the special case of an amortized loan repayments are constant over time, i.e. \( A \) is constant. A particular solution of Equation (A.1) is found by choosing a suitable constant for \( B_n \). This constant is equal to the solution of the linear equation \( x = x + IR \cdot x - A \), yielding \( B_n = A/IR \). The actual solution of Equation (A.1) then becomes:

\[
B_n = C(1 + IR)^n + \frac{A}{IR} \tag{A.3}
\]

Substitution of \( n = 0 \) gives \( C = B_0 - A/IR \). Using the condition that \( B_{NY} = 0 \) and rewriting for \( A \) yields:

\[
A = B_0 \frac{IR(1 + IR)^{NY}}{(1 + IR)^{NY} - 1} \tag{A.4}
\]

where \( NY \) is the number of years in which the loan is paid off. The initial balance \( B_0 \) may correspond to the capital investments required for the construction of processing and storage facilities.
A.2 Distance Between Center Points

Consider a $n \times m$ grid of equally sized grid units that are numbered row-by-row, as illustrated in Figure A.1 below.

![Illustrative example of a 2 × 3 grid.](image)

The distance between the center points of two grid units $i$ and $j$ for an $n \times m$ grid with equally sized units is given by:

$$D_{i,j} = \ell \cdot \sqrt{(i - n \left\lfloor \frac{i}{n} \right\rfloor - j + n \left\lfloor \frac{j}{n} \right\rfloor)^2 + \left(\left\lfloor \frac{i}{m} \right\rfloor - \left\lfloor \frac{j}{m} \right\rfloor\right)^2}$$

(A.5)

where $\ell$ is the dimension of a square grid unit. Although the indices $i$ and $j$ correspond to harvesting sites and integrated processing facilities, it should be noted that Equation (A.5) holds in general.

A.3 Average Distance to the Center of a Square

In Equation (2.60) the integral that evaluates the average distance from all points inside a square to its center ($AD$) is introduced and its solution is presented. A derivation of this solution is given below:

$$AD = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \sqrt{x^2 + y^2} \, dx \, dy$$

$$= 4 \int_{0}^{\frac{1}{2}} \int_{0}^{\frac{1}{2}} \sqrt{x^2 + y^2} \, dx \, dy$$

$$= 8 \int_{0}^{\frac{1}{2}} \int_{0}^{\frac{1}{2}} \sqrt{x^2 + y^2} \, dy \, dx$$

$$= 8 \int_{0}^{\frac{1}{2}} \int_{0}^{\frac{1}{2} \sec \theta} \sec \theta \, r^2 \, dr \, d\theta$$

$$= \frac{1}{3} \int_{0}^{\frac{\pi}{2}} \sec^3 \theta \, d\theta$$

$$= \frac{1}{6} \left( \sec \theta \tan \theta + \ln(\sec \theta + \tan \theta) \right) \bigg|_{0}^{\frac{\pi}{4}}$$

$$= \frac{1}{6} (\sqrt{2} + \sinh^{-1}(1))$$

(A.6)
Appendix B

Piecewise Linearization

In Table B.1 below the regression constants for each processing facility in case of a single break-point are given. In addition, the capital investment scaling equation and the linear approximation for both the preprocessing and upgrading facility are shown in Figure B.1 and B.2 respectively.

Table B.1: Regression constants in case of one break-point.

<table>
<thead>
<tr>
<th>Facility</th>
<th>( r )</th>
<th>( C_r )</th>
<th>( S_r )</th>
<th>( I_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated</td>
<td>1</td>
<td>134</td>
<td>3.88</td>
<td>147.64</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>400</td>
<td>2.18</td>
<td>376.11</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>1</td>
<td>1360</td>
<td>0.06</td>
<td>23.23</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4000</td>
<td>0.03</td>
<td>58.48</td>
</tr>
<tr>
<td>Upgrading</td>
<td>1</td>
<td>134</td>
<td>2.48</td>
<td>94.45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>400</td>
<td>1.40</td>
<td>240.62</td>
</tr>
</tbody>
</table>

Figure B.1: Preprocessing facility capital investment scaling equation (blue) and its linear approximation (red) in case of a single break-point.
Figure B.2: Upgrading facility capital investment scaling equation (blue) and its linear approximation (red) in case of a single break-point.
Appendix C

Results Case Study I

The processing facility locations and capacities corresponding to the optimal supply chain configurations are given in Table C.1 to C.3. The location is indicated by the grid unit where the facility is located. Grid units are numbered row-by-row.

Table C.1: Processing facility location and capacity for the $5 \times 5$ grid for both topologies and processing pathways.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>kGGE/year</td>
<td>ktonne/year</td>
<td>kGGE/year</td>
</tr>
<tr>
<td>Center-Point</td>
<td>Integrated</td>
<td>13</td>
<td>100944</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Preprocessing &amp; upgrading</td>
<td>4</td>
<td>388</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>394</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>-</td>
<td>100944</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>388</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>388</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Corner-Point</td>
<td>Integrated</td>
<td>13</td>
<td>100944</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Preprocessing &amp; upgrading</td>
<td>7</td>
<td>419</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>650</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>-</td>
<td>100944</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>488</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table C.2: Processing facility location and capacity for the $7 \times 7$ grid for both topologies and processing pathways.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>kGGE/year</td>
<td>ktonne/year</td>
<td>kGGE/year</td>
</tr>
<tr>
<td>Center-Point</td>
<td>Integrated</td>
<td>25</td>
<td>227429</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Preprocessing &amp; upgrading</td>
<td>10</td>
<td>556</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>681</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>519</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>227429</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34</td>
<td>631</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>37</td>
<td>681</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>438</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Corner-Point</td>
<td>Integrated</td>
<td>25</td>
<td>227429</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Preprocessing &amp; upgrading</td>
<td>9</td>
<td>650</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>731</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>569</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27</td>
<td>563</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>33</td>
<td>-</td>
<td>-</td>
<td>227429</td>
</tr>
<tr>
<td></td>
<td></td>
<td>39</td>
<td>488</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td>506</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table C.3: Processing facility location and capacity for the $9 \times 9$ grid for both topologies and processing pathways.

<table>
<thead>
<tr>
<th>Typology</th>
<th>Processing pathway</th>
<th>Location</th>
<th>Integrated processing cap.</th>
<th>Preprocessing capacity</th>
<th>Upgrading cap.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>kGGE/year</td>
<td>ktonne/year</td>
<td>kGGE/year</td>
</tr>
<tr>
<td>Center-Point</td>
<td>Integrated</td>
<td>41</td>
<td>396075</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Preprocessing &amp; upgrading</td>
<td>11</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td></td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38</td>
<td></td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td></td>
<td>256</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td></td>
<td>-</td>
<td>396075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44</td>
<td></td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65</td>
<td></td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68</td>
<td></td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71</td>
<td></td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td>Corner-Point</td>
<td>Integrated</td>
<td>41</td>
<td>396075</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Preprocessing &amp; upgrading</td>
<td>11</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td>-</td>
<td>-</td>
<td>396075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68</td>
<td>-</td>
<td>731</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71</td>
<td>-</td>
<td>256</td>
<td>-</td>
</tr>
</tbody>
</table>
Appendix D

Results Case Study II

The processing facility locations and capacities corresponding to the system configurations shown in Figure ?? and 4.6 are given in Table ?? and ?? below.

Table D.1: Processing facility location and capacity corresponding to the selected system configurations depicted in Figure 4.4.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>kGGE/year</td>
<td>ktonne/year</td>
<td>kGGE/year</td>
<td></td>
</tr>
<tr>
<td>(a)*</td>
<td>11</td>
<td>-</td>
<td>818</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>-</td>
<td>713</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>227429</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>-</td>
<td>713</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>-</td>
<td>550</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>-</td>
<td>713</td>
<td>-</td>
</tr>
<tr>
<td>(b)</td>
<td>10</td>
<td>227429</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* As in Figure ??.

Table D.2: Processing facility location and capacity corresponding to the selected system configurations depicted in Figure 4.6.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>kGGE/year</td>
<td>ktonne/year</td>
<td>kGGE/year</td>
<td></td>
</tr>
<tr>
<td>(a)*</td>
<td>38</td>
<td>-</td>
<td>813</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>-</td>
<td>1068</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>-</td>
<td>-</td>
<td>227429</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>-</td>
<td>813</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>-</td>
<td>813</td>
<td>-</td>
</tr>
<tr>
<td>(b)</td>
<td>39</td>
<td>227429</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* As in Figure 4.6.
Appendix E

Results Case Study III

In Table E.1 the facility locations and capacities corresponding to the selected system configurations depicted in Figure 4.10 are given. Furthermore, in Figures E.1 to E.3, the raw data corresponding to the supply chain configurations (a) to (c) from Figure 4.10 are presented.

Table E.1: Processing facility location and capacity corresponding to the selected system configurations depicted in Figure 4.10.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>kGGE/year</td>
<td>ktomme/year</td>
<td>kGGE/year</td>
</tr>
<tr>
<td>(a)*</td>
<td>10</td>
<td>-</td>
<td>556</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>-</td>
<td>681</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>-</td>
<td>519</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>227429</td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>-</td>
<td>631</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>-</td>
<td>681</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>-</td>
<td>438</td>
<td>-</td>
</tr>
<tr>
<td>(b)</td>
<td>10</td>
<td>-</td>
<td>981</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>-</td>
<td>775</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>-</td>
<td>1100</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>-</td>
<td>325</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>-</td>
<td>-</td>
<td>227429</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>-</td>
<td>325</td>
<td>-</td>
</tr>
<tr>
<td>(c)</td>
<td>25</td>
<td>227429</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(d)</td>
<td>34</td>
<td>227429</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* As in Figure 4.10.
Figure E.1: The variable cost components for each simulated biomass supply scenario with increasing $\sigma$ (from top to bottom) of the sampling distribution. The results correspond to the preprocessing/upgrading configuration at a threshold value of $\Phi = 0.1$ (see Figure 4.10 (a)).

Figure E.2: The variable cost components for each simulated biomass supply scenario with increasing $\sigma$ (from top to bottom) of the sampling distribution. The results correspond to the preprocessing/upgrading configuration at a threshold value of $\Phi = 0.4$ (see Figure 4.10 (b)).
Figure E.3: The variable cost components for each simulated biomass supply scenario with increasing $\sigma$ (from top to bottom) of the sampling distribution. The results correspond to the integrated processing configuration at a threshold value of $\Phi = 0.1$ (see Figure 4.10 (c)).