30-9-2013

A sustainable bioenergy Future for Puerto Rico

Feedstock potentials, production costs, and environmental impact. A spatially explicit supply chain optimalization using GIS and a Mixed Integer Linear Programming Model

Benjamin A.J. Visser
UTRECHT UNIVERSITY & UNIVERSIDAD DE PUERTO RICO RECINTO MAYAGUEZ

This study was conducted in fulfilment of the Master of Science program in Energy Science. It enthails an integration of both thesis and internship. The internship was followed at the University of Puerto Rico in Mayagüez, and was made possible by funding from the Rinny Huizinga Stichting, the Netherlands.

Universiteit Utrecht

RINNY HUIZINGA STICHTING

Colofon

Title: **A sustainable bioenergy Future for Puerto Rico.** Feedstock potentials, production costs, and environmental impact**.** A spatially explicit supply chain optimalization using GIS and a Mixed Integer Linear Programming Model.

Author: Benjamin A.J. Visser
M.Sc. Energy Science student

Permanent address:
Marshalllaan 340
3527 TV, Utrecht
The Netherlands
Email: Benjamin.visser@hotmail.com

University: Utrecht University
Copernicus Institute
Department of Innovation, Environmental and Energy Sciences
Heidelberglaan 2
3584 CS, Utrecht
The Netherlands

Supervisors: André P.C. Faaij
Email: a.p.c.faaij@uu.nl

Gert-Jan Jonker
Email: j.g.g.jonker@uu.nl

Hosting Agency: Universidad de Puerto Rico recinto Mayagüez
Calle Post, Mayagüez Puerto Rico
PR 00681-9000

Hosting Supervisor: Luis R. Pérez-Alegría
Email: luisr.perez1@upr.edu

Date of Publication: September 2013

Table of Contents

List of tables

List of figures

1. Introduction

1.1. Problem definition

The island of Puerto Rico has historically been a major player in the sugar industry. During the golden age of sugar production, the island intensively cultivated 160.000 ha of land (more than 1/5th of its entire land cover), producing about 1 million tons of sugar in 34 mills in 1954 (Colón-Guásp, 2010). Due to volatile global sugar prices, and a gradual shift from an agricultural to an industrial and more recently a service oriented economy, the sugar industry and agricultural sector have since been declining. Colón-Guásp further mentions the small share (less than 1%) of agriculture in GDP (table 1 below) and high share of imports for food related products (export-import ratio 1:12 in financial terms; table 2 below) (Colón-Guásp, 2010).Also remarkable are the high expenses for the transportation sector, i.e. high expenses for petrol fuels.

Table 1. GDP share of the sectors, in million dollars. Commas indicate thousands.

	2004	2005	2006
Sectors/ Totals & Shares	79209	82650	86464
Manufacture	33268	34363	36556
Finance; real estate	13029	14016	14733
Commerce	9802	10260	10717
(Central) Government	7389	8151	8424
Services	7646	8023	8164
Transportation, public services	5343	5353	5508
Construction Mining	1905	1873	1821
Agriculture	414	360	333
Statistical discrepancies	415	251	209

Source: Colón-Guasp (2010).

Table 2. Exports and imports for 2006, in million dollars.

	export	import
Total	60,119	42,630
Manufacture	59,542	39,808
Foodstuffs	3,956	2,380
Agriculture, Silviculture	41	492
Beverages and tobacco	377	450

Source: Colon-Guasp (2010).

Looking at Puerto Rico's renewable energy portfolio, no plans are yet in place to develop alternatives for fossil fuels. The island is to increase its renewable portfolio from 1% in 2009 to 15% in 2020 (installed capacity) and initiates a fuel diversication program for its electricity production from oil (69 to 26%) to natural gas (15 to 30%) and coal (15% to 29%) for the same period (PREPA, 2011).

The total installed capacity for renewables in 2014 can be broken down into wind (38%), solar (21%), waste-to-energy (24%), hydro (10%) and ocean thermal (8%) (PREPA, 2011). In this scenario no attention is given to the possibility of fossil fuels substitution, i.e. of biodiesel and bioethanol production for the transport sector.

A bioenergy sector on the island is virtually absent. While Puerto Rico has been one of the mayor sugar exporting colonies in the past, its sugar producing sector is now nearly diminished. The year 2002 saw the end of sugarcane production and the abolishment of the two remaining sugar mills. Some stakeholders show renewed interest in such a sector for different reasons, such as a local supply of molasses for the rum industry, or to give a boost to agriculture and employment in the south and west of the island. Researchers from Mayaguez (UPR-M) and the 'Universidad Del Este-Turabo' (UPR-T) have ventured the topic and focused mainly on the production of ethanol from organic Municipal Solid Waste, or MSW (8000 ton/day feedstock flow) added by forest residues and sugar/energy cane feedstocks[1] (Colon-Guásp, Peréz-Alegría 2000). Particularly the MSW seems an attractive source because of its abundance, and the pressing situation of the landfills, i.e. the lack thereof (Colón-Guásp et all, 2000). Nevertheless, the research proposal by UPR-M/UPR-T urged the need for a preliminary feasibility study for the determination of production-, transportation- and feedstock costs, recommendations for plant size, plant locations, environmental considerations, seasonality and feedstock availability. Models need to be developed for optimal seasonal ethanol supply sources and final use destinations, as well as water and land transportation costs from all producing regions to all consuming regions.

It is no less than a coincidence that I came up with this subject and methodology, as I wrote it after reading this proposal. Unfortunately there is not much published or undertaken in the field of ethanol production after the year 2000 meaning a low interest or at least a low activity on the ethanol production front.

In addition to ethanol production, UPR-M started a research project with two municipalities and three companies to investigate the trans-esterification of waste greases, used cooking oil, and animal fats into fatty acid methyl esters (FAMEs). The project included a resource and market assessment, engineering designs and identification of industry partners (Colucci & Pandzardi, 2002).

The project indicated a collection potential of 5-10 million gallons per year of which 1-5 million gallons are already collected. Especially for a 10 million gallon production plant -the threshold for an economically viable plant, additional feedstocks are needed either from imports or local production. For now, there are 13 local projects experimenting with waste grease feedstocks, ranging from fuels for transportation to inputs for generators and boilers (Colucci & Pandzardi, 2002). These are some of the few examples that indicate the state of biofuel development on the island. This thesis aims to deliver a comprehensive contribution to the knowledge en policy development concerning an untapped renewable energy resource.

[1] Feestocks are the input resources for any industry or process. In this case the 'feedstocks' are primary inputs that are used to produce ethanol. These can be woody materials, organic waste in municipal solid waste. In this research the production of dedicated energy 'feedstocks' are investigated and are sugarcane, energy cane and elephant grass.

1.2. Justification of this thesis

The fact that an ethanol and sugarcane industry, and in general a biomass industry is virtually absent in Puerto Rico was, instead of an agrument against, a motivation precisely in favour of a study on the feasibility of this untapped renewable energy resource. The favourable climate and soil conditions, a long history of sugarcane cultivation, together with pressing socio-economic issues with high unemployment, a neglegted agricultural sector, and import dependency combined with volatile prices of fossil fuels were arguments in favour of this study.

Engaging in a bioenergy industy for the island begins with an assessment of the available land and resources. Land use changes have been very dynamic in the 20th century, due to agricultural decline and subsequent forest recovery and urban expantion. Due to the highly densed population on the island it is very important to start with an estimation of the available land for bioenergy production.

Table 3 below summarizes some earlier analyses on land uses in Puerto Rico. It assesses the land cover changes for the years 1991-2000 for the different land use classes (forests, urban, agriculture). Agricultural land use diminished to 51,000 hectares in 1991, and further to 26,000 hectares in 2000 (a total decline of more than 80% from 1954-2000). Forest cover on the other hand recovered significantly; from only 6% of total land area in 1940 to around 40% in 2000, while population grew from approximately 2 to 4 million. Forest recovery mainly occurred on abandoned pasture and agricultural land (Paréz-Ramos et al, 2008).

Table 3 also shows that agricultural land is lost (for 1991-2000) mainly to grasslands, whereas forests gained most from grass- and scrubland. Paréz-Ramos et al, but also Del Mar López et al (2001), mention that further loss of prime agricultural land will occur in the face of continuous urban expansion and re-urbanisation (i.e. sub-urbanization). A well developed and scattered road network, increased urban sprawl in coastal planes, and increasing land prices (expansion of industrial parks and other land completion) will all further increase the scarcity of agriculture land. Still, an USDA GAP Analysis from 2006 reported a combined area of grasslands, agriculture and active/abandoned pasture land to account for over 30% of the islands surface of which agriculture accounted for 260 km^2, dry grassland-pasture 420 km^2 and moist grassland-pasture 2200 km^2 (USDA, 2008).

Table 3. Transition matrix analysis of land-cover changes in Puerto Rico from 1991 to 2000.

	Land Cover 1991-2000	2000 Grassland	Forest	Shrubland	Urban	Agriculture	Sand& Rock	Total Area km2 in 1991	% in 1991
1991	Grassland	1736	669	469	302	65	6	3247	38
	Forest	287	1816	261	47	12	3	2425	28
	Shrubland	349	811	270	59	5	1	1495	17
	Urban	227	73	62	535	5	5	906	11
	Agriculture	227	46	31	30	177	1	512	6
	Sand & Rock	6	3	2	2	0	10	23	0
Total Area km2 in 2000		2832	3417	1095	975	263	26	8607	100
Percentage in 2000		33	40	13	11	3	0		

Source: Paréz-Ramos et al, 2008

On the other hand, according to Peréz-Alegría and Colón-Guásp (2010), the island has a readily available area of 40,000 hectares for dedicated energy crops. This amount is mainly from fallow land dedicated to the cultivation of sugarcane, and was abandoned completely after 2002. Their convervative estimate indicate an ethanol production potential of 200 million liters from this area from cane juice (5% of the islands gasoline use in 2007). Furthermore, land occupation can expand beyond this 40,000 ha; other conversion routes can be applied; the biomass feedstock can be diversified with agriculture- and forest residues and Municipal Solid Waste; and the plant's load factor can be improved by enhanced biomass supply management.

Next to this, historical data on the sugarcane industry reported relatively high commercial cane yieds for the production of sugar. Paragraph 3.1.e. 'Biomass yields' will elaborate more on this, but for now it suffices to say that yields in the range of 80-110 metric tons per year were maintained for decades in the early 20th century, on a commercial scale of about 5-10 million tons of sugarcane output per year. Experimental plots in the '80s and '90s have furthermore confirmed high experimental yields of 250 tons per hectare with precision fertilization and irrigation (see paragraph 4.1.b 'Biomass Yields').

With all these and other opportunities for the agricultural sector, employment, trade and energy diversification and -security, it is worthwhile investigating the potential and realistic configuration of such a bioenergy sector for Puerto Rico. The proposed combination of ArcGIS mapping possibilities together Microsoft excel and a CPLEX solver tool in the GAMS linear optimization tool will present a sophisticated approach to the analysis of a bio-ethanol supply chain.

1.3. Main aim of the research

This research investigates the best suitable configuration for a bio-ethanol sector that is most suited for the island of Puerto Rico. The supply chain presented here includes a series of three feedstocks, three types of ethanol conversion technologies, an array of locations for these plants, and an array of the size or capacities of these plants. This optimization procedure has a twofold result: first the ethanol supply chain will be optimized for its economical and emission performance, i.e. the highest Net Present Value (NPV) and according Greenhouse Gas (GHG) emission performance. Typically, a configuration with the best financial performance is also the one with the least favourable environmental performance. Conventional ethanol conversion technologies using 1^{st} generation feedstocks such as sugarcane, corn or starchy crops have a lower GHG emission avoidance than the more expensive but higher efficiency 2^{nd} generation technologies. Simultaneously a multi-period, spatially specific designation of biomass and biofuel production sites will result from the detailed geographical data input. The research question will be as follows:

1. "What are the production costs and GHG emissions of the proposed bio-ethanol supply chain for the period 2014-2030 for the island of Puerto Rico, following a multi-objective linear optimization of this supply chain?"

To support this main research question, three more subquestions will be answered. These are:

1. What is the sustainable biomass production potential and how is this potential geographically distributed?'

2. 'Which dedicated energy crops, ethanol conversion technologies and subsequent capacity assignment and –planning will answer best to the island's future bio-ethanol sector?'

3. How does the ethanol supply chain preference –of previous subquestion- change into the future, i.e. can turning points be detected where the supply chain switches feedstocks an technologies?

Two models have been utilized to answer this question. The first is a PC Raster Land Use Change (PLUC) model, developed at the University of Utrecht, to investigate the potential available land for bioenergy crops after the demand for land has been allocated to all other land uses up to 2030. The second model is a Multi-Objective, Mixed Integer Linear Programming (MOMILP) model to optimize for the most optimal supply chain configuration of ethanol considering the proposed feedstocks, technologies and supply chain specific parameters. The original script was developed by the University of Padova, Italy, and optimized for an ethanol supply chain of combined 1^{st} and 2^{nd} generation technologies with corn and corn stover for the north of Italy.

The reason for a combination of the proposed software is chosen is as follows. The PLUC model has a three-fold purpose; one is to aid the ongoing land use planning in Puerto Rico from a scientific point of view; the second is to circumvent the ongoing and credible arguments of land use competition between bioenergy crops and food production; while the third is to bring into awareness the possibilities of a re-valuation of agriculture from an energy point of view, where food and energy independency can be increased considerably. The MOMILP model optimizes for the various options as to choose the best feedstock, ethanol conversion technology, plant capacities and future planning structure for this untapped renewable energy resource. Together these models form an integral approach that is needed for a profound bioenergy system analysis.

In **chapter two** the methodology of this thesis is explained. How these models work, how they are integrated with each other, and what data is required for their successful performance. Two future scenarios are proposed for the PLUC model that intend to capture the possibilities for agricultural, livestock and urban expansion. It is explained how the model then allocates land on a cell by cell basis to cropland, pasture and urbanization according to their estimated future increases in output. For the MOMILP model are explained the Life Cycle Analysis (LCA) and the Supply Cycle Analysis (SCA) for resp. the GHG emission and the economic performance of the supply chain.

Chapter three presents the data input and intermediate results of the PLUC model. The results of the PLUC model will aid in answering subquestion 1 on the available potential land for bioenergy production. The findings of the PLUC model are answering to; 'where and to what extend are the three dynamic land use classes 'cropland', 'pasture' and 'urbanization' expanding'? A visual representation with ArcGIS land use maps is given of regions with concentrated and increasing cropland, pasture, urbanization. At last is given a visual representation of the available land for potential production of dedicated bioenergy crops. These maps are all representations of land uses for the year 2030. This potential land availability will be one of the inputs for the MOMILP model.

In **chapter four** is presented the remaining data input that was prepared for the MOMILP model. It follows the ethanol supply chain from 'seed-to-tank' and presents the data along the three main life cycle stages: (1) Biomass production; (2) Ethanol production; (3) Transportation of biomass and ethanol. Life cycle stage (1) 'Biomass production' includes estimations on biomass yields, biomass production costs and GHG emissions associated with the production of biomass. Life cycle stage (2) 'Ethanol production' includes an elaboration on the proposed ethanol conversion technologies, the technical parameters of these technology configurations, GHG emissions associated with ethanol production and GHG emission credits received from excess electricity production that is exported to the island's grid. The last life cycle stage (3) Transportation of Biomass and Ethanol includes the costs and emission for transporting both commodities.

Chapter five presents the results of the MOMILP model. The MOMILP model aids in answering the second and third subquestion on the total system costs and the impact of future trends. The MOMILP model results present the Cost of Ethanol (COE) and GHG emissions from 'seed-to-tank' for the preferred ethanol supply chain configuration. They also elaborate on the preference of the supply chain configuration at specific points in the future, when certain dynamics are accounted for. These dynamics are improvements in biomass yields, ethanol conversion efficiencies for each technology, and cost developments due to increases of production inputs, electricity and in general the inflation. By answering this last subquestion on future trends, it is clarified what the role of 2nd generation conversion technologies are within the energy mix, and understood if and when they become economically viable. This question also aids in the 'invest-now-or-postpone' decision-making, and can indicate how a future planning of an ethanol supply chain can be organized. Furthermore the results will be presented in a graphical manner, illustrating where ethanol plants and biomass plots are to be located in which time period within the 2014-2029 timeframe.

With these reseach questions question answered, stakeholders will have tangible economic and technical parameters to evaluate and engage in the establishment of a bioenergy sector that will provide at least a portion of the islands transportation fuels.

1.4. Scientific context

The use of Linear Programming (LP) models for optimization problems has been investigated since the second world war and the body of literature is therefore extensive. For the past decades, optimizations have been used for product/production allocations and global transportation networks. Increasingly, this knowledge is being applied in the field of supply chain management (SCM) in the bioenergy sector, as a switch in the short and medium term future towards such a bio-based economy clearly benefits from such an undertaking. Previous cases have sought to indicate site location and plant size, producing biofuels from one single 1st generation feedstock (Graham et al. 2000). Giarola et al (2011) investigate the most optimal site and size configuration and integrate both 1st and 2nd generation ethanol producing plants (using only grain and grain stove as feedstock), accounting for economies of scale, feedstock costs, and scaling factors.

Tittmann et al (2010) take previous research a step further by optimizing for different feedstocks (municipal solid waste, agricultural and forestry residues), applying an ArcGIS transportation network analysis and analyzing end product preferences at different prices for by-products. Still, there are quite some quality and in-depth improvements to be made in optimizing the supply chain, e.g. by providing more detailed geographical input, integrating more environmental constraints, optimizing the biomass and fuel production processes even further and leaving more room for end product competition.

Analyzing the supply of biomass feedstocks depends largely on its spatial distribution, and the use of satalite imagery of land uses with ArcGIS software is an adequate tool to do this. Graham (2000) uses digitalized yield and soil maps, incorporates formulas for farm gate costs, land rent and transportation costs and subsequently reproduces geographically specific maps of feedstock prices, land costs, yields, marginal costs of feedstock delivered and even plant locations.

Van der Hilst et al. (2010) represents in a similar way the costs of energy crops in comparison to the conventional crop production by incorporating algorithms in layer files and thus allowing for a geographically dispersed representation in ArcGIS. Nevertheless, the transportation costs are set at an average distance of 90 km, and its variability is somewhat ignored (it *does* represent around 10% of the total production costs of ethanol). In another article exploring the spatio-*temporal* variation of land uses in Mozambique (van der Hilst, Verstegen, 2011), the authors ingeniously apply ArcGIS combined with PC Raster software by developing socio-economic factors together with suitability maps for different land use classes and model these land use changes *per year* up to 2030.

These articles from Graham and Van der Hilst, but also other researchers using ArcGIS in feedstock analysis have in common that they explore the *upstream* biomass potentials; (1) potential arable land in the future or supply chains incorporating only one technology (van der Hilst, 2011); or (2) farm-gate and marginal prices of delivered biomass of only one feedstock (Graham, 2000).

This research expands on the possibilities of connecting spatially explicit biomass potential analyses (upstream) with supply chain analyses (downstream), incorporating a variety of feedstocks, conversion technologies, plant/plot locations and subsequent plant capacities.

Van der Hilst & Verstegen's use and methodology of ArcGIS and PC Raster software is worthwhile looking at to incorporate Land Use Change (LUC) factors that are of relevance in Puerto Rico. Most of their LUC-factors are developed especially for lesser developed countries (e.g. the use of wood for fuel, absence of centrally planned forest- and agricultural management, low productivity and slow efficiency improvements, extensive livestock cultivation, subsistence farming and the like, typical factors for less developed countries) and do not apply to Puerto Rico. Paragraph 2.2.a 'PLUC Model 1' will explore more her methodology to incorporate LUCs and to extract the sustainable biomass potentials.

1.5. Expansion of knowledge base

This research will expand on earlier research by breaking down the optimization procedure in more detail. Furthermore, a small geographical area (that of Puerto Rico) is a good test case to investigate the two proposed models on a smaller scale.

The first improvement relates to a more detailed spatially explicit input of data. This information will come from satalite imagery and is presented in Geographical Information Systems or ArcGIS datasets, and concerns land availability, a detailed road network, soil conditions, and water availability. Different energy crops are available, and are to be deployed to areas with optimal coinciding climatic and soil conditions. This must be done to optimally produce the biomass, while minimizing the environmental impact (e.g. minimizing fertilizer and water use). Elephant grass, for example is known to be more drought resistant and can be grown with less fertilizer input.

Secondly, the optimization of the plant's load factor will be investigated. Ethanol derived from sugarcane is usually restricted by the feedstock supply. Feedstock flows are restricted to varying yields and seasonal cycles, resulting in a plant operation time frame called the LOMS -Length Of Milling Season- and varies per region (e.g. 6 months/year in Louisiana, 9 months/year in KwaZulu Natal, South Africa) (Le Gal, 2009). Subsequently, the issue of optimizing the feedstock flow, i.e. an optimal or steady supply of biomass, will be addressed in this research. Optimization can be applied by diversifying the feedstock. Energy cane and sweet sorghum are two family members of sugar cane with closely related juice, cellulosic, hemi-cellulosic contents and can be harvested outside the milling season of sugar cane (Misook, 2011).

Another extension that will be addressed is the incorporation of small Combined Heat and Power (CHP) units, producing process heat and electricity (and excess to the grid). Due to very high electricity prices in

Puerto Rico it might occur that an increased CHP utilization at the expense of fuel production at the biofuel plant will be more attractive economically for the whole processing chain[2]. The technology configurations in this research are also optimized outside the LOMS by supplying a second feedstock, in this case elephant grass, to the CHP unit.

[2] Electricity prices were 0,22 $/KWh (2009) for Puerto Rico, against 0,09$/KWh (2010) in the US (eia, 2013).

2. Methodology

This chapter describes the methodologies that are used in the analysis of the different life cycle stages of the ethanol supply chain. It explaines the system boundaries, integration of the three used models, and elaborates shortly on the functionings of these models. A detailed description of these models is presented in the appendices, while the subsequent chapters 3 and 4 give detailed elaborations on the data requirements and procurements for these models. Chapter three will elaborated on the first PLUC model, and chapter four on the input for the MOMILP model.

2.1. Main structure

Figure 1 below clarifies the boundaries of the complete system investigated in this research. It is a representation of the energy system where energy and material input-output flows are captured. The overall ethanol supply chain (purple) under investigation consists of four parts, namely (1) agricultural production, (2) transport of biomass, (3) the industrial process of ethanol production (green), and (4) transport of ethanol to demand centers. The energy system under investigation includes several stages or sectors of the productive economy. Therefore it is important to start with a framework to delineate the boundaries of the reseach.

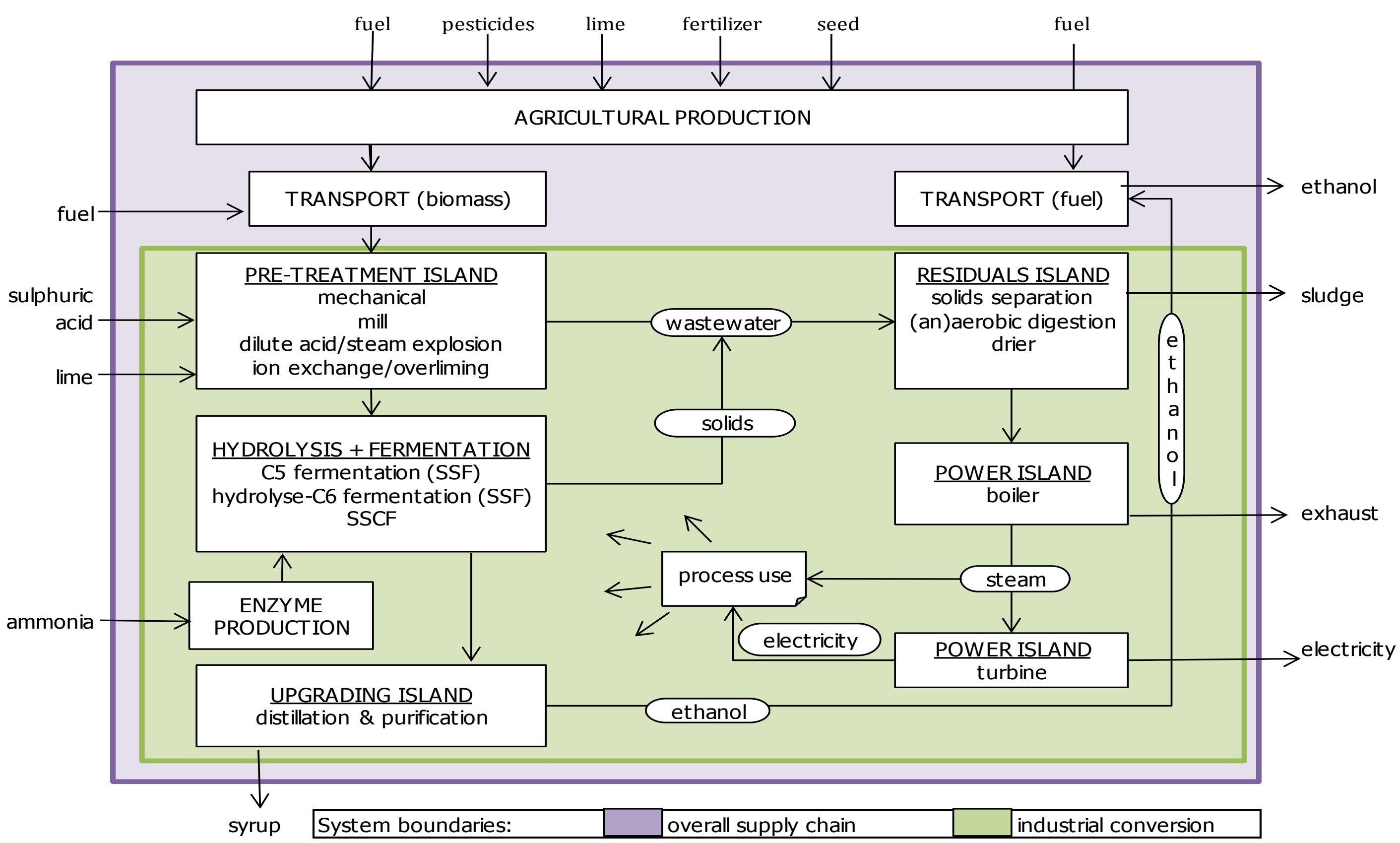


Figure 1. System boundaries of the ethanol supply chain under investigation for Puerto Rico. The life cycles included are (1) biomass production, (2) transport of biomass, (3) ethanol production, (4) transport of fuel.

The four parts mentioned above can be called 'life cycle stages', and are complementary in the 'seed-to-tank' analysis of this supply chain. The inputs for every life cycle are quantified for their costs and GHG emissions. These inputs are processed, transported and further processed, then transformed into the useful outputs ethanol, the by-product electricity, and into waste such as sludge and exhaust.

For the life cycle stage 'biomass production' will be assessed three biomass feedstocks, namely sugarcane, energy cane and elephant grass. All the inputs for this life cycle stage will be documented, and will be prepared for cost and associated emission parameters. Emissions are divided into (1) direct emissions such as exhaust from agricultural machinery and (2) indirect emissions associated with the production of inputs such as fertilizers and lime. This is done as well for the life cycle stage 'Ethanol production'.

The life cycle stage 'Ethanol production' (green in Figure 1) depicts the process flow of the conversion of biomass to ethanol that is assumed in this research. Three technologies are assessed for this supply chain for Puerto Rico, and are: (1) a 1st generation cane fermenting technology, the mainstream technology in conventional sugarcane ethanol industries; (2) a 2nd generation 'Simultanuous Saccharification then Fermentation' (SSF) technology; and (3) a 2nd generation 'Simultanuous Sacharification Combined Fermentation (SSCF) technology. Paragraph 4.2 'Life cycle stage 'Ethanol Production'' will elaborate more on the characteristics of these technologies.

At last are gathered cost and emission factors for the two transport life cycle stages: transportation of biomass to the ethanol conversion plant; and for transportation of ethanol from conversion plant to demand centers where the ethanol will be blended with gasoline.

For the analysis of these four life cycle stages of the ethanol supply chain will be used three different models. The integration of the models is illustrated in Figure 2 below. The first and foremost model is a linear optimization model (model 3 MOMILP, blue) that will optimize a range of options within the supply chain. One of these 'variables' is the amount of land that can be used for dedicated bioenergy crops. Thus, for the first life cycle stage 'biomass production', the potential land estimation will be analysed with model 1 called PLUC (pink in Figure 2). The PLUC model will be called 'model 1' as it represents the first part of analysis in a chronological order, i.e. it is used in the first life cycle stage analysis. The rest of the biomass production parameters will be gathered in the field and prepared in excel spreadsheets.

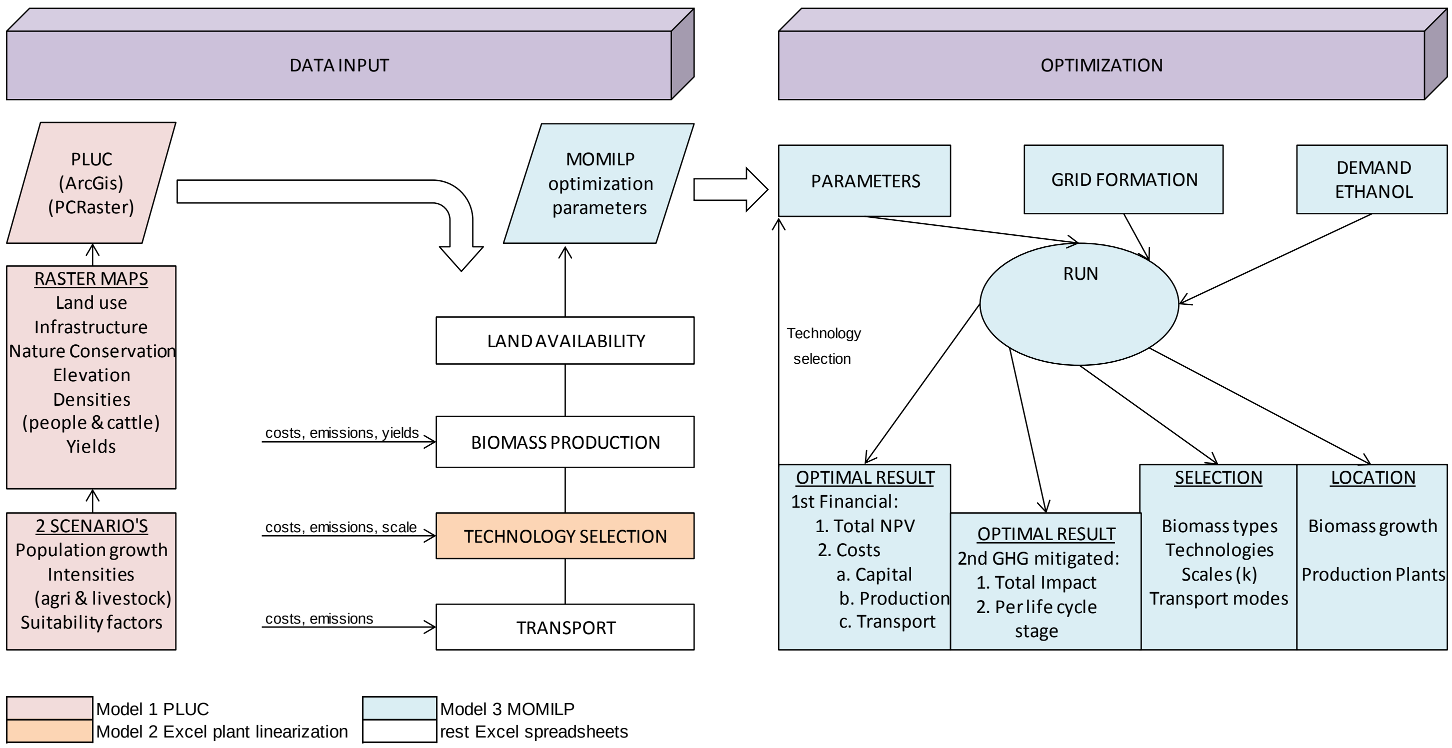
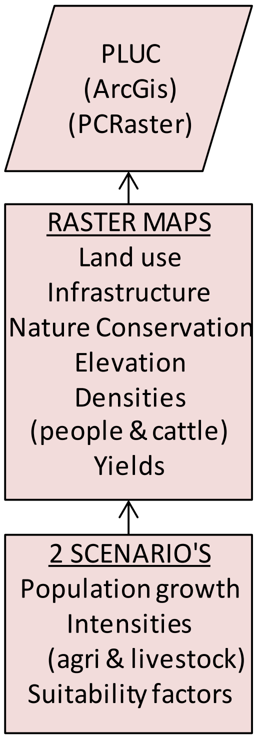


Figure 2. Integration of Model 1 (PLUC) and model 3 (MOMILP). The outputs of the second model are optimization results considering costs, emissions (GHG), selection of biomass types and optimal location of a bioenergy production plant.

Another 'variable' in the supply chain of ethanol is a range of technology options to convert biomass into ethanol. These so-called technology configurations are linearized in a second model in excel (orange in Figure 2). This excel model will analyse the third life cycle stage 'ethanol production' and will scale the plant compontents to their appropriate capacity and accompanied production cost-, ethanol conversion-, electricity- and biomass requirement parameters for the linear optimization model.

All other parameters for the remaining transport life cycle stages will be prepared in excel as well. The sum of above mentioned parameters will be used in the model 3 MOMILP.



PLUC MODEL 1 Land Use Changes. First, an estimation is given of the sustainable potential agricultural area that can be used for bioenergy crop production. Van der Hilst and Verstegen (2011) at the University of Utrecht developed a PCRaster Land Use Change model (or 'PLUC' model), and this model will be prepared in the first stage of investigation. The application of dedicated bioenergy crop production often encounters resistance because of its land competition with agriculture or other land uses. It is therefore important to include other land use classes in the analysis and account for their land use requirements in the future before allocating land to dedicated bioenergy crops. The PLUC model does this. The model requires the preparation of a Business as Usual (BAU) scenario and a Progressive Scenario (PS), that tend to capture the future agricultural, livestock and urban increases in output. The BAU scenario assumes a continuation of historical (increases in) output, whereas the PS scenario will assume a progressive agricultural policy in the future. The increases in output in the future require extra land to expand on, and this land will be allocated according to suitability factors. The suitability factors will decide which *current* land use will most likely be changed to *next year's* land use considering the satisfaction of its estimated demand described in the scenarios. Land use changes are modeled in a 1-year sequence until 2030, the timeframe of this research. Chapter 3 'Model 1: PLUC input and results' will elaborate more extensively on scenario development and other requirements of the PLUC input.

Next to data for scenario building are gathered ArcGIS geographical maps of Puerto Rico, and these are: a land use map, infrastructural, nature conservation, elevation, population and livestock densities and yield maps (see left). The PLUC model allocates all land necessary to the other land use classes up to 2029, and the remaining available land (may it be from freed-up agricultural land, or grass/scrubland) will be available for bioenergy crop production. The amount of land from the Progressive Scenario that has come available in 2029 will become one of the parameters for model 3 MOMILP.

<u>EXCEL MODEL 2</u> Ethanol Plant Linearization. The next phase of data collection is for model 2 and focuses on technical (yields, emissions) and economic input (capital & operating costs) for the technologies that are used as options for the future ethanol supply chain of Puerto Rico. This part concerns an analysis of the third life cycle stage 'ethanol production'. Nine different technology configurations are build that are combinations of the three feedstocks and three conversion technologies. From Hamelink (2004) are taken capacity-, scaling- and cost data for the specific components of a conversion plant (figure 1 in green). These components represent the (1) pretreatment, (2) Hydrolysis + Fermentation, (3) Upgrading, (4) Residuals handling, and (5) the Power island of the plant. These nine technology configurations are an energy-mass balance analysis to scale the components appropriately and will produce specific parameters such as ethanol yield, excess electricity production, biomass requirements, capital costs, operating costs.

For the life cycle stages 'transportation' and biomass production data is gathered in literature or in the field and also prepared in excel to the adequate parameters needed for MOMILP model 3.

<u>MOMILP MODEL 3</u> Supply Chain Optimization. If all input data is gathered, the optimization procedure can begin as is seen in Figure 3. For this phase of research, two other actions are needed. First, the total

amount of available land is aggregated over 52 equally squared surfaces or grid elements laid over the island. This is done to let each grid element be a *possible* option for plant and/or harvest location (after a no-go filter excluding cities, conservation, water bodies). This is the so-called 'multi-objective' (versus the 'multi-attribute) decision making approach (MODA vs. MADA) towards assignment of most optimal locations. A multi-attribute allocation procedure *pre*-determines a limited amount of locations from which to choose, whereas multi-objective allocation permits all points (52 grid elements) as options for establishment.
Second, a desired ethanol demand is set in a selected few grid elements where the ethanol is to be transported to and mixed at tank stations (these are the ten biggest population clusters). For this demand of ethanol, the MOMILP model will optimize the supply chain as to satisfy this demand, i.e. model 3 will select the technologies, their scale, location, feedstocks.

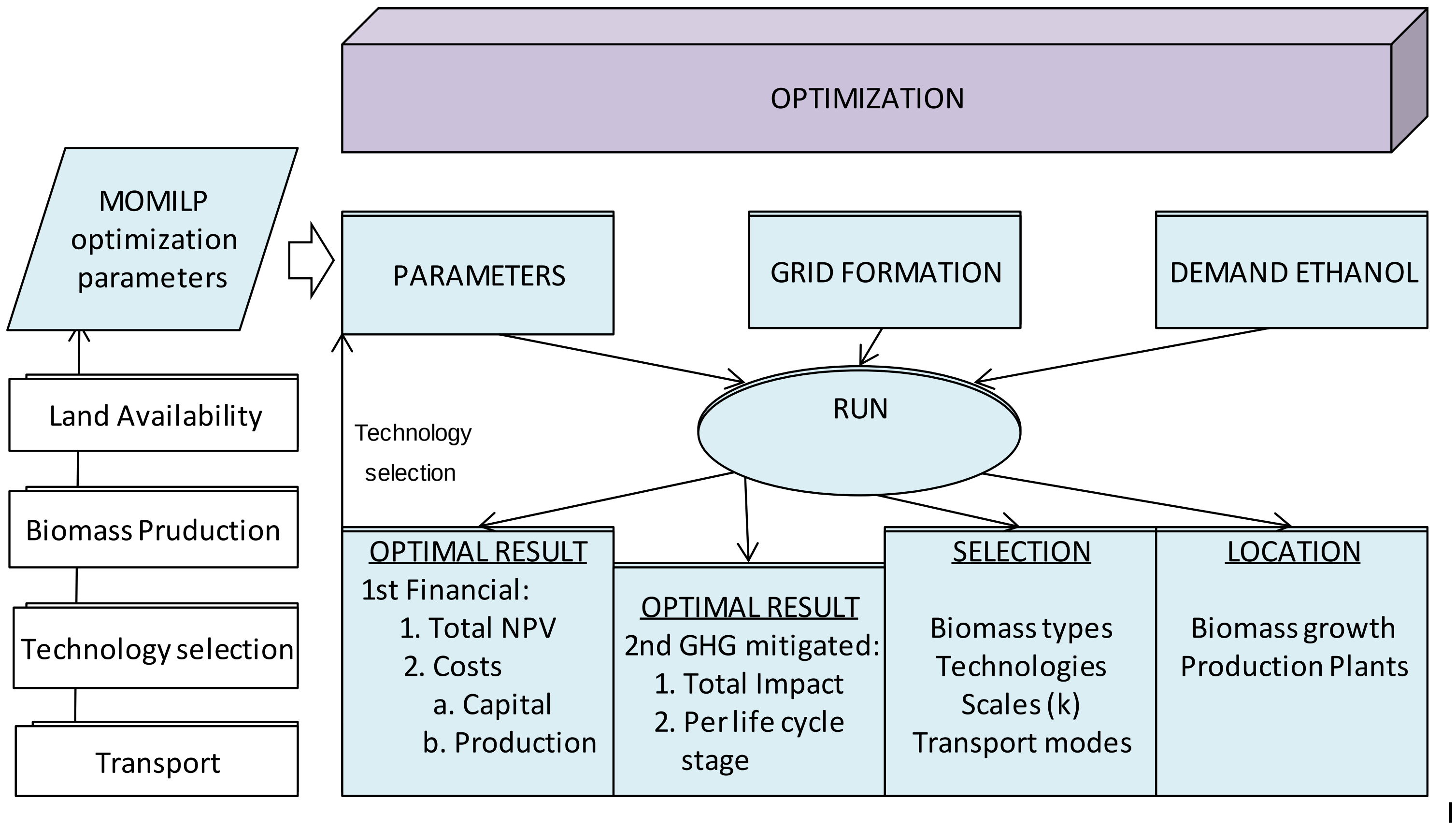


Figure 3. Model flow of the third model MOMILP: data requirements, set-up and results.

The calculation method used in this thesis can be divided into three general formulas: (1) one for ethanol conversion costs, (2) one for the Cost of Energy (COE) of the whole supply chain, and (3) one for the greenhouse gas emissions (GHG) of this supply chain representing all life cycles. These formulas are as follows:

1. Ethanol Conversion Costs:

Conversion $costs_{ethanol}$ = $CAPEX_t$ + $OPEX_t$

<u>Where</u>:

$CAPEX_t$	= Total Capital Expenditures in time t
$OPEX_t$	= Operational Expenditures in time t.

2. Overall supply chain costs:

$$EthanolCosts_{supplychain} = NPV\left(\sum_{t} BPC_t + \sum_{t} TCb_{biomass,t} + \sum_{t} EPC_t + \sum_{t} TCf_{ethanol,t} - B_{revenues,t}\right)$$

<u>Where over the timeframe 2014-2030:</u>

NPV	= Net present value of the supply chain; lifetime of 20 years; interest rate of 15%.
BPC	= Biomass production costs
$TCb_{biomass}$	= Transport cost of biomass from plot to plant
EPC	= Ethanol production costs
$TCf_{ethanol}$	= Transport costs of ethanol from plant to demand center
$B_{Revenues}$	= Benefits or revenues from selling excess electricity to the grid.

The benefits of selling ethanol are not calculated in the NPV, only the benefits from excess electricity to the grid. In this way the NPV thus represent the Net Present Cost (NPC) of ethanol production, i.e. the associated costs for the satisfaction of the ethanol demand initiated in MOMILP. This NPC is also called the 'cost of energy' ($COE_{ethanol}$). To calculate the $COE_{ethanol}$, the following formulas have been used:

$$NPV = \sum_{t} -TCI_t * dfTCI_t - \sum_{t} C_t * dfCF_t + \sum_{t} B_t * dfCF_t$$

$$dfTCI_t = \frac{1}{(1+\zeta)^{3(t-1)}} \qquad dfCF_t = \frac{3 + 3\zeta + \zeta^2}{3 * (1+\zeta)^{2t}}$$

<u>Where</u>:

NPV	= Net Present Value over time horizon ($)
TCI_t	= Total Capital Investments per time period t ($)

$dfTCI_t$ = Discount factor TCI_t[3]
C_t = Cost factors ($)
$dfCF_t$ = discount factor CF_t
B_t = Benefits from electricity per time period ($)
ζ = Interest rate (15%)
t = Timeperiod

Where the Cost of Energy ($COE_{ethanol}$) is that price for ethanol sold, where the NPV is null, thus:

$$0 = \sum_t -TCI_t * dfTCI_t - \sum_t (C_t + B_t) * dfCF_t + COE_{etoh} * \sum_t E_t * dfCF_t$$

and:

$$COE_{etoh} = \frac{\sum_t TCI_t * dfTCI_t + \sum_t (C_t - B_t) * dfCF_t}{\sum_t E_t * dfCF_t}$$

<u>Where</u>:
COE_{etoh} = Cost of $Energy_{ethanol}$ ($/GJ)
E_t = Ethanol that is marketed per time period t (in GJ)
Calculation of E_t is explained in paragraph 4.2 'Life cycle stage 'Ethanol Production'', where the energy-mass balance is elaborated on for scaling the technologies.

3. For the greenhouse gas emissions, the formula used is:

$$GHG\ emissions_{total} = GHG_{biomass} + GHG_{transport,biomass} + GHG_{ethanol} + GHG_{transport,ethanol} - GHG_{credits}$$

<u>Where for the timeframe 2014-2030</u>:
GHG_{total} = Total greenhouse gas emissions of all life cycle stages
$GHG_{biomass}$ = GHG emissions associated with the stage 'biomass production'
$GHG_{transport,\ biomass}$ = GHG emissions associated with the stage 'transport of biomass'
$GHG_{ethanol}$ = GHG emissions associated with the stage 'ethanol production'
$GHG_{transport,\ ethanol}$ = GHG emissions associated with the stage 'transport of ethanol'

---

[3] The general formula for discounting is $\frac{1}{(1+\zeta)^t}$. But for the investigated time frame of 2014-2030, five time period of three years each are used. The Total Capital investments are expenditures in the beginning of a year, therefore this dfTCI is used. The Cash flows are for every end of the year, thus resulting in given formula.

Figure 3 above illustrated the optimization process within MOMILP. When ethanol demand, grid formation, and parameters are set, the model is ready to produce results. The time frame ranges from 2014-2029, a total of fifteen years divided in five time periods of 3 years. The model optimizes either for best financial performance (NPV, see above) or best GHG emission mitigation performance. Furthermore, the MOMILP optimization will result in:

1) selection of biomass type, technology type and appropriate scale;
2) locations of biomass plots and ethanol production plants; and
3) optimal financial performance in terms of NPV and total capital-, operating, harvesting- and transportation costs per time period.

When this is finalized, a few technologies can be selected to once again use in an optimization run, but now minimizing for its GHG emission performance.

One more important exercise in excel requires the preparation of trend parameters to include future developments related to yields and costs of both biomass and ethanol production. In the life cycle stage ‘biomass production’, trends that are relevant are (1) biomass yield increases in the future due to increased land and production management and better harvesting methods. Simultanouslly increasing are the costs for producton inputs. For these are applied price indices for fertilizers, diesel and steel (machinery). Other cost components such as land and labour are corrected with a GDP deflator.

For the life cycle ‘ethanol production’ are important the conversion efficiencies of biomass to ethanol, i.e. the efficiencies in hydrolysis, saccharification, and fermentation. The range in efficiencies given by Hamelink (2004) is the starting point and linearized over the time period of investigation. If a technology would be installed in the first time period, it would have the lower bound of this efficiency range, increasing to the upper bound in the last time period. Cost factors for this life cycle stage are corrected solely by the GDP deflator.

2.2. Model Descriptions

This paragraph elaborates more extensively on the methodology of the models. In sub-paragraph 2.2.a. it is explained what the requirement of the PLUC model are. The development of two scenarios will be explained, as well as the the method for land allocation according to suitability factors. At last will be explained how the available land will be aggregated from the PLUC model to the MOMILP model. Paragraph 4.2.a 'Scaling Nine Technology Configurations' will elaborate more extensively on the second model in excel that is used to scale the nine technology configurations. In sub-paragraph 2.2.b. the MOMILP model is explained further; the parameters that are needed as input; grid formation; setting up demand centers; the dual optimization of financial and emission performance; and technology and capacity planning over the time frame.

2.2.a. PLUC Model 1

The input data for the first PLUC model requires both scenario parameters, as well as land use data in the form of ArcGIS raster maps. These geographical maps are from satalite imagery and transposed to Geographical Information System (GIS) formats. Optimizations with linear programming are possible or can be built in the ArcGIS software, albeit in much simpler form. The level of optimization complexity in this thesis requires the use of both ArcGIS software and an optimization program like the CPLEX solver in the GAMS tool. The fact that ArcGIS can read data and generate results in Microsoft Excel makes ArcGIS a powerful tool not only for displaying geographical data but also for showing optimization results of the supply chain analysis (Nardi et al, 2007).

According to Malczewski (2006), ArcGIS Multi Criteria Decision analyses are grouped into Multi Objective Decision Analysis (MODA) and Multi Attribute (MADA) approaches, where the latter starts off with a predefined set of alternatives. In this thesis this would be a number of locations where the biorefineries are to be located, and consequently the most optimal cultivation plots + plant locations are calculated. However, in this thesis a MODA approach will be used, where the plant location can be within every grid cell.

As mentioned above ArcGIS maps will be used before and after the Solver's optimization. The first round of ArcGIS data will produce land availability maps and biomass productivity maps. The optimization of the supply chain occurs in the CPLEX solver in the GAMS tools, because the amount of locations for plant construction is gridbased. To control the calculation time, the amount of grids are reduced to 52. The level of detail in the ArcGIS analysis will be unmanageable either in Excel's solver or GAMS solver, i.e. 1 km^2 cells sum up to 80,000 cells from ArcGIS in Excel and this requires too much computational power. Therefore the ArcGIS land availability data will be aggregated to larger polygons of potential plots. The conversion from the PLUC model to the MOMILP model concerning available land will be explained at the end of this paragraph.

2.2.a.i. Introduction

The PLUC model is applied to model the land use changes (LUC) into the future and extracts the potential available land for bioenergy crop production in Puerto Rico for the time frame 2014-2030. Land requirements for crop production, livestock, and urbanization are accounted for, thus avoiding competition for land between bioenergy crops on the one hand, and existing land uses on the other hand. Figure 4 below summarizes the PLUC model in a graphic way (from van der Hilst and Verstegen, 2011). The model can be broken down in three main categories;

1. 'Drivers of Demand' that will influence the future output increases from cropland/ pasture/ urbanization,
2. 'Scenario characteristics' that shape the Demand category in two scenarios, and
3. The allocation of land to the different dynamic land uses on a yearly basis.

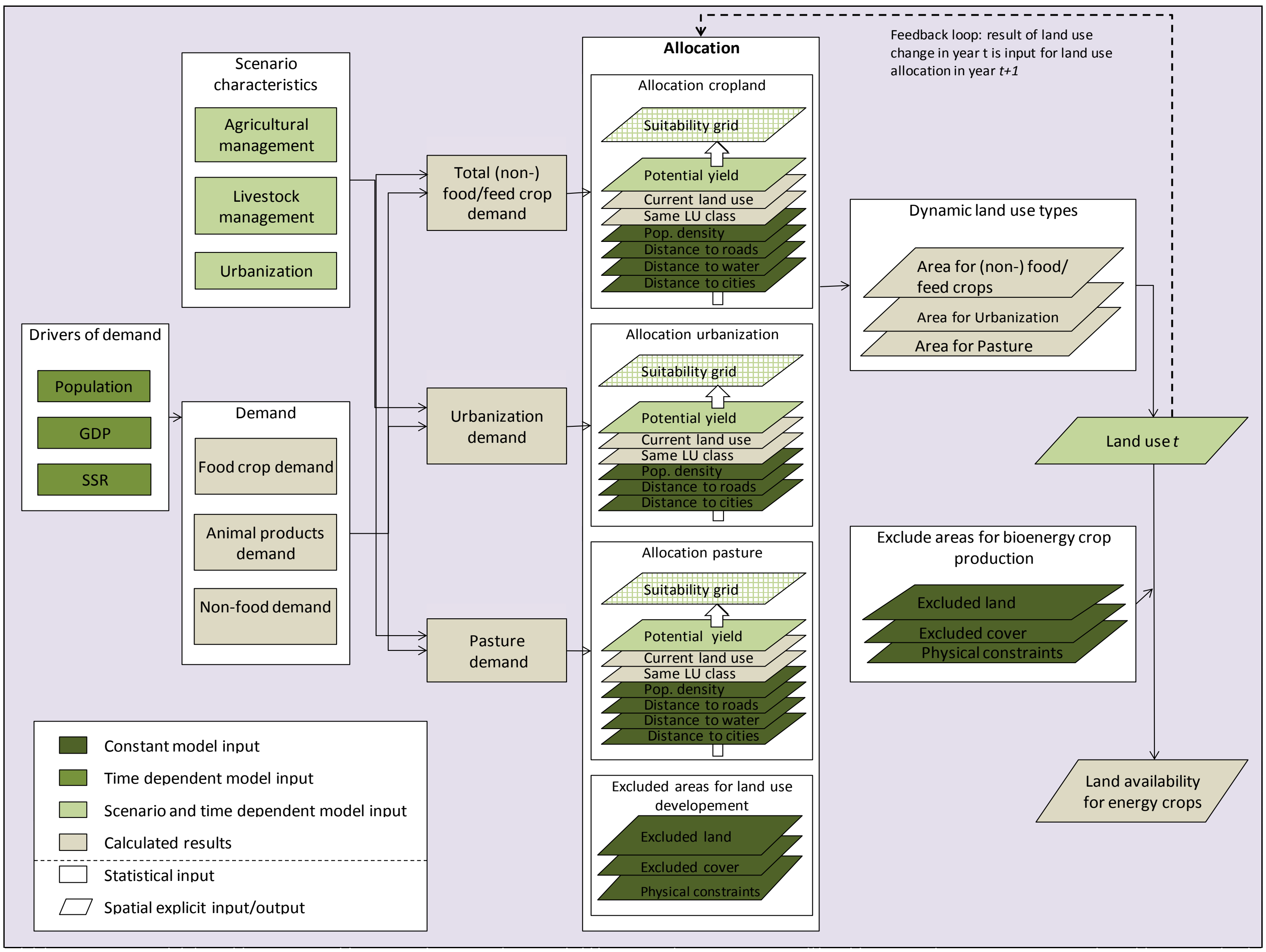


Figure 4: Scematic of the PLUC model: (left) drivers of demand and scenario characteristics formulate two scenarios; (middle) allocation of land is dependent on suitability factors, and will produce a final suitability raster grid for that land use class and year; (right) land allocation occurs with iterations of 1 year (edited from Van der Hilst and Verstegen, 2011).

The 'Drivers for Demand' influence the demand for animal products (thus pasture land) agricultural products (thus cropland) and the demand for urbanization (thus urbanized surface). The original model included deforestation for wood fuel purposes, and this wood demand was replaced with a demand for urbanization. This was done because urbanization is an important factor on the highy populated island. Demand for wood and subsequent deforestation was taken out of the group of dynamic land use classes since this practice is virtually absent in the Puerto Rican situation.

'Scenario characteristics' represent management or policies that shape the demand side of these dynamic land uses 'cropland, pasture, urbanization' and are captured in two in scenarios; the Business As Usual (BAU) and Progressive Scenario (PS). These scenarios describe a different future development for the 'demand' for agricultural land, pasture land and land for urban expansion. These 'drivers of demand' reflect developments in population, GDP and the self sufficiency ratio (SSR). Furthermore, only grass- and shrubland is considered as potential land for bioenergy crops, while forestland is excluded as a source either for fuelwood or land available for dedicated energy crops. The application of ArcGIS and this PCRaster Land Use Change model enables Puerto Rico to model future land developments in both a geographic and time dimension. This PLUC model therefore is a step forward both in the island's development of future agricultural scenarios, and in modeling land availability for bioenergy protential. Apendix 10.1 gives a full elaboration on the requirements of the PLUC model: scenario development; suitability factors; and required ArcGIS maps. the next sub-paragraphs that follow now will only give a short description of these requirements.

2.2.a.ii. Scenario development

The two scenarios that are developed are needed to structure and quantify the complex and multi-dimentional factors that are of influence on local land use dynamics. Prime agricultural land in Puerto Rico is subject to great competition between urbanization, crop- and cattle production and other land uses such as recreation and forest conservation. This competition is characterized by developments in population and GDP growth, the island's self sufficiency ratio (the ratio between local production and export/import), agricultural intensification/modernization, the population's diet and policies towards forest conservation, agricultural policy and the bioenergy sector (van der Hilst & Verstegen, 2011).

Demand for cropland and pasture land is dependent on the demand for crops and animal products, which are dependent on population growth, the local consumption and export of these products. The yields of both crops and livestock is dependent on agricultural productivity and intensification of the sectors. This productivity is dependent on the level of modernization of the sector, i.e. aplication of advanced seeds, mechanization, irrigation, feed-to-meat conversion efficiency and type of cattle industry (extensive v.s intensive cattle ranching) (van der Hilst et al, 2011).

Urbanization is dependent on developments in population growth. The mid-20th century has seen a tremendous urbanization, although the last past decades this trend was somewhat weekened. Developments in forestry are only dependent on conservation policies, as nowadays deforestation due to the harvest of wood-for-fuel is non-existent on the island. Forestland has tremendously increased in area from the early

20th century untill the late 90s, from less than 6% of total cover before 1940 to around 40% in 2000, and of that 11% between 1990-2000.

For both scenarios the current forestry conservation policies are maintained in the future, i.e. the amount of forestland will neither increase nor decrease. For the diet of the population will also be assumed that this will stay constant in the future. Table 4 below summarizes the structure of the two scenarios with parameters that are quantified for the PLUC model. These parameters are a summary of the data that is gathered and discussed in chapter 3 'Model 1: PLUC input and results'.

Table 4. Scenario parameters for the PLUC model. Two scenarios require a different agricultural and urban growth in the future, and subsequently land that is needed for expansion of these land use classes. The table is a summary of Chapter 3.

LUC drivers	2007 (historical data 1982-2007)	BAU scenario	Progressive Scenario
Population	0.97% p.a. change	Same as current, in line with historical trend	Same as current, in line with historical trend
SSR	17.65% share of local production in total consumption	Same as current; export same as current	35.30% share (doubling) of local production in total consumption. Export same as current
Farming practices	Commercial farming, widespread access to improved seeds, machinery, pesticides/ herbicides and knowledge	Same as current	Same as current
Agricultural productivity	5.4 ton/ha/yr yields for vegetables & fruits 1.58% p.a. historic increase	1.58% p.a. increase, modest improvements in yield, in line with historical trend	3.5 % p.a. increase in yields
Livestock productivity	0.53 ton/ha/yr yield 0.72% p.a. historic increase. Moderate feed conversion, mixed systems, good disease control	0.72% p.a. increase, in line with historical trend.	3,5% p.a. increase
Deforestation	Total land coverage 40%, robust forest management.	Same as current	Same as current
Urbanization	Urban yield 37 inhabitants/ha 0.41% p.a. 'urban yield' growth 12.9% of area urbanized 0.94% p.a. population growth	Same as current, in line with historical trend	Same as current, in line with historical trend

For the PLUC model are thus identified three *dynamic* land use classes which are cropland, pasture land and urbanization to which the model will allocate land to. Population growth, GDP growth and the population's Self Sufficiency Ratio (SSR) will increase the future demand of agricultural and livestock products, and subsequently its land requirement. In formula form for agriculture:

$\text{Crop intensity}_{\text{historical}}$ (tons/ha/year) = $\text{output}_{\text{historical}}$ (tons/year) / $\text{area}_{\text{cropland}}$ (ha)

First are gathered historical trends on output and associated land use (in hectares) of these three dynamic land use classes. For agriculture and livestock this is the aggregated output of the full range of produced goods. For urbanization the population size functions as output (number of inhabitants), whereas the land use is again in hectares. By dividing the output with associated area the historical *intensities* of these dynamic land use classes are retrieved.

Second are developed scenarios for a different continuation of these *intensities* into the future. The Business-As-Usual scenario (BAU) will assume a continuation of historical trends of population, GDP, agricultural-, livestock- and urbanization yields. The Progressive Scenario (PS) on the other hand, will assume a progressive increase in intensities in agriculture and livestock. For both scenarios will be assumed a progressive output in the future, i.e. a doubling of the SSR in 2030. Allocation of land to these three dynamic land use classes will happen as below formula states.

$Output_{future}$ (tons/year) / crop $intensity_{future}$ (tons/ha.yr) = Yearly $allocation_{future}$ (ha/year)

For each scenario is calculated a future output according to its specific growth trends in output and yields. The PLUC model will allocate land to each land use class *until* the required output for that year is satisfied. The allocation of land happens on a yearly basis, i.e. iterations of one year. If the demand for these products is bigger than the island's production, the model will allocate land (gridcells) to cropland *untill* this yearly demand is satisfied. It evaluates on a cell-by-cell basis the availabity and suitability of a gridcell to be converted to a specific land use class. It does this for cropland, urbanization and pasture in this sequence.

The agricultural and livestock output for both scenarios will assume a progressive increase in the timeframe 2014-2030. Both scenarios will assume in 2030 a return to output levels of 1970, the highest levels of agricultural output found in the data available. The agricultural sector has been declining since the 1950s due to a shift towards an industrial and service oriented economy. This shift has resulted in a high dependency on food imports, low quality of the imported food, high food prices, and a relatively high unemployment. But the most important reason for this progressive assumption is that food independency is at least as important as energy independency, if not more important for a sustainable future. Paragraph 3.2.a 'Self Sufficiency Ratio (SSR)' will go deeper in this discussion.

The PLUC model uses static ArcGIS maps as starting point and generates (yearly) intermediate grid rasters for each specific dynamic land use class. First excluded is land that is not to be used at all for modeling, such as nature conservation areas, roads, buildings, areas with slope higher than 16 degrees, water bodies. Additionally, grid-based yield maps are needed for crops, livestock and urbanization and also prepared in ArcGIS. These yield maps are used by the model to find the most suitable land or 'gridcells' for allocation to each land use class. In this way land of the best quality is allocated to the land use classes. Furthermore, the PLUC model will allocate land first to cropland, than to urbanization, then to pasture. This is an order of allocation specificly set for this thesis, and the order of 'who gets to choose first' can be set in the model at ones own choice. The reason for this order is to grant agriculture the best land for producing food. Urbanization is second in allocating land before pasture because of its imperative presence. Pasture is also last because cattle can graze virtually everywhere on the island due to its fertile soils and favourable climate. Land for allocation can come from future abandoned cropland or pasture (yield improvements outweight the demanded output), grassland, or scrubland that was abandoned crop/pastureland in the last 20 years.
The allocation of land to different land use classes occurs as follows. Evaluation of each gridcell to be converted to a specific land use type is governed by a set of suitability factors. Table 5 summarizes the

suitability factors and their weights for eacht land use class. Appendix 10.1 'Model 1 – PLUC description' explains in further detail the workings of the PLUC model, as well as the required maps for Puerto Rico[4].

Table 5. Suitablitiy factors and their weighting for each dynamic land use type.

	Dynamic land use type		
	Cropland	Pasture	Urban land
Suitability factor	weight		
Nr. of neigbours same class	0.2	0.2	0.2
Distance to roads	0.1	0.05	0.25
Distance to water	0.15	0.15	
Distance to cities	0.1	0.05	0.15
Yield	0.2	0.1	0.05
Population density	0.1	0.05	0.2
Cattle density		0.25	
Current land use	0.15	0.15	0.15
total	1	1	1

The weightings of the suitability factors for each dynamic land use type can be changed in the model, and these specific ones in Table 5 were assigned such as to best reflect the Puerto Rican land uses. The sum of the weight are 1, so the order is a qualitative one. Some remarkable weightings will be explained. Distance to roads for example is more important for cropland (0,1) than for pasture (0,05) because cropland needs to be reached by machinery, and cattle is ranged over a larger area by horse. Distance to water is equally important for both land uses. Population density is more important for cropland than for pasture, there the crop products are delivered faster to the city, whereas cattle is again ranged extensively. For pasture, 'current land use' was weighted highest, because rangeland was not represented as class in the land use maps. Allocation of land to pasture thus needs to happen as close to the cattle areas as possible, i.e. accurately allocated according to the cattle density map (see paragraph 3.2.e 'Map ccuracy' and 3.3 'PLUC results' for further discussion on this topic). The category urban land use was newly implemented and replaced the 'Wood demand' in the original model. This was because demand for wood for fuel was non-existent in Puerto Rico, whereas a significant urbanization certainly was.

[4] The original article is: 'Spatiotemporal land use modeling to assess land availability for energy crops – illustrated for Mozambique'. Van der Hilst, F., Verstegen, A. et all (2011).

2.2.a.iii. Linking PLUC and MOMILP

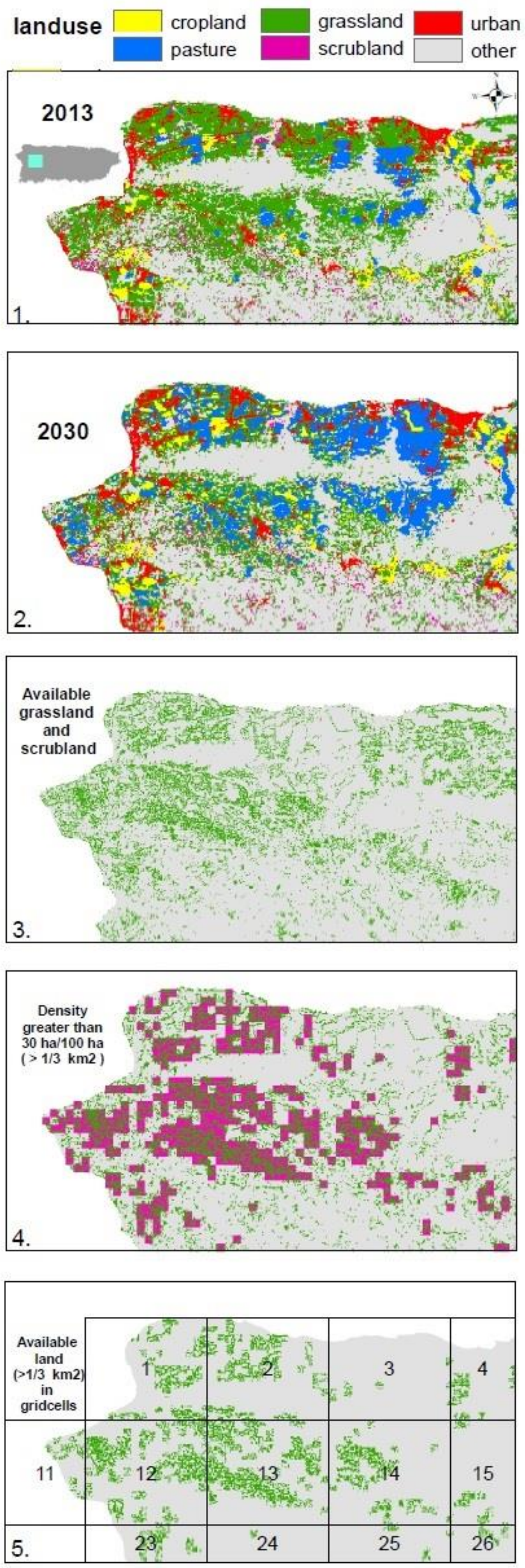


Figure 5. Process flow of grassland extraction after PLUC has run to 2029. A density layer is applied grassland is aggregated to 256 km2 gridcells.

Figure 5 left illustrates the steps that are taken to arrive at the amount of available land for bioenergy in 2030. The first land use map of 2013 contains ten land use classes, of which three dynamic ones; cropland, pasture, urbanisation. The other two important land use classes are grassland and scrubland that will be set available first for allocating to the three dynamic land use classes, and at last for potential bioenergy crop production. Figure 5 zooms in at the north-west region of Puerto Rico and shows the PLUC results and the available land extraction. The region includes (around the clock) Aguadilla, Arecibo, Utuado and at last Mayaguez in the left bottom corner.

Map two is a land use map for 2030 after a PLUC run. Land is allocated to cropland, urbanization, and pasture land. In fact, cattle is ranged extensively in open land that is available, may it be grassland, scrubland that was crop/pastureland before, or idle cropland.

The third map illustrates the available grassland/scrubland only in the year 2030. In some parts grassland is very dense, other region are less dense. For this reason another density extraction has been applied (map 4). The purple squares have 1 km^2 surfaces. If grass/scrubland amounted to 30 hectares per purple square, it is used. Thus, potential land with a density of 30% per km^2 or more will be used. Other grassland will be excluded. Map 5 then depicts the results of this density extraction. This method is still suboptimal, since some purple plots are isolated from other plots, meaning there is no economies of scale if that isolated plot is cultivated. But for now this will not be dealt with. Map 5 furthermore illustrates the gridraster that is layed over the land. The available land within every gridcell is then extracted and exported to excel and then to MOMILP for the optimization. Each grid will function as a possibility for plant location and/or biomass cultivation. Thus, there are 52 gridcells (g), each containing the MOMILP parameters GS(g) or biomass grid surface; BY(i,g) or biomass yield per biomass type i in grid g; BPC(i,g) or biomass production costs per biomass type i and grid g. For a full exploration of all the parameters see appendix 10.2.

Note that for the available land is used both grassland and scrubland. The land use maps produces by the GAP analysis from 2007 gave a very detailed overview of land in 70 different land use classes, with their descriptions about what type for vegetation, slope, climatic zone. For scrubland was also mentioned if the expantion occurred on abandoned cropland or pasture land in the last 20 years. Thus for the category 'scrubland' is only used land that falls in this category, i.e. scrupland on previous cropland and pasture land. Therefore it can be used for bioenergy production without converting forest land or destroying rich biodiversity. This scrubland is mostly populated with low density, 1-3 types of invasive trees. This does however have impact on the Co_2 balance of the ethanol conversion, but this is not included in this research.

2.2.b. MOMILP Model 2-3

The aim of applying the MOMILP model is to analize the most optimal establishment of a bioenergy sector for the island of Puerto Rico. A multi-period MOMILP (multi-objective mixed integer linear programming) model will be used for this optimization procedure. The university of Padova, Italy, developed a linear programming model for 1^{st} and 2^{nd} generation ethanol conversion from corn and stover for the north of Italy, and this model is adjusted to the Puerto Rican circumstances. This research will offer a cost-effective, quantified and optimized methodology for policymakers in the preparation of such a sector on the island or elsewhere.

Figure 6 illustrates the different links in the bioenergy supply chain, which can be divided into an upstream and downstream chain. Between both end nodes, a wide variety of deployable bioenergy supply pathways exists, all dependent on the available biomass feedstocks, conversion technologies, transportation modes, and logistical choices. The design outcomes are not unique, as they strongly depend on conversion technology and geographical context. Multiple variables are of influence, and are integrated in the MOMILP. These are from left to right (figure5) (from Giarola et. al.):

1. biomass types and sites for production
2. logistic distribution and transport characterization
3. ethanol production technology and spatially specific site location
4. capacity assignment and planning
5. demand centers

This analysis will result in a spatially specific, optimized localization of both biomass production areas and biofuels production plants. The economics of the supply chain will be assessed by means of a Supply Cycle Analysis (SCA), focusing on the biomass cultivation locations, ethanol production capacity assignment and facilities location, as well as on the transport system. The environmental impacts will be obtained by applying a Life Cycle Assessment (LCA) from 'seed-to-tank', i.e. for all three life cycle stages mentioned in the methodology (Giarola, p. 1785).

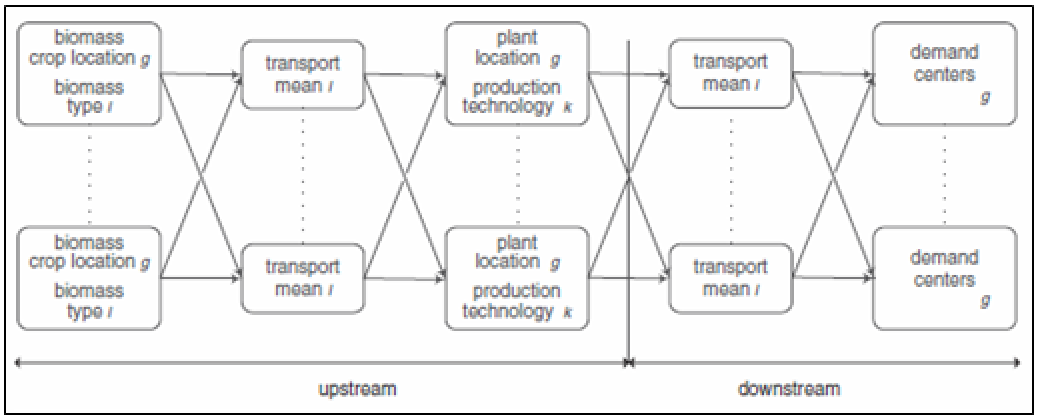


Figure 6. Biofuels network supply chain. From S. Giarola et al. (2011).

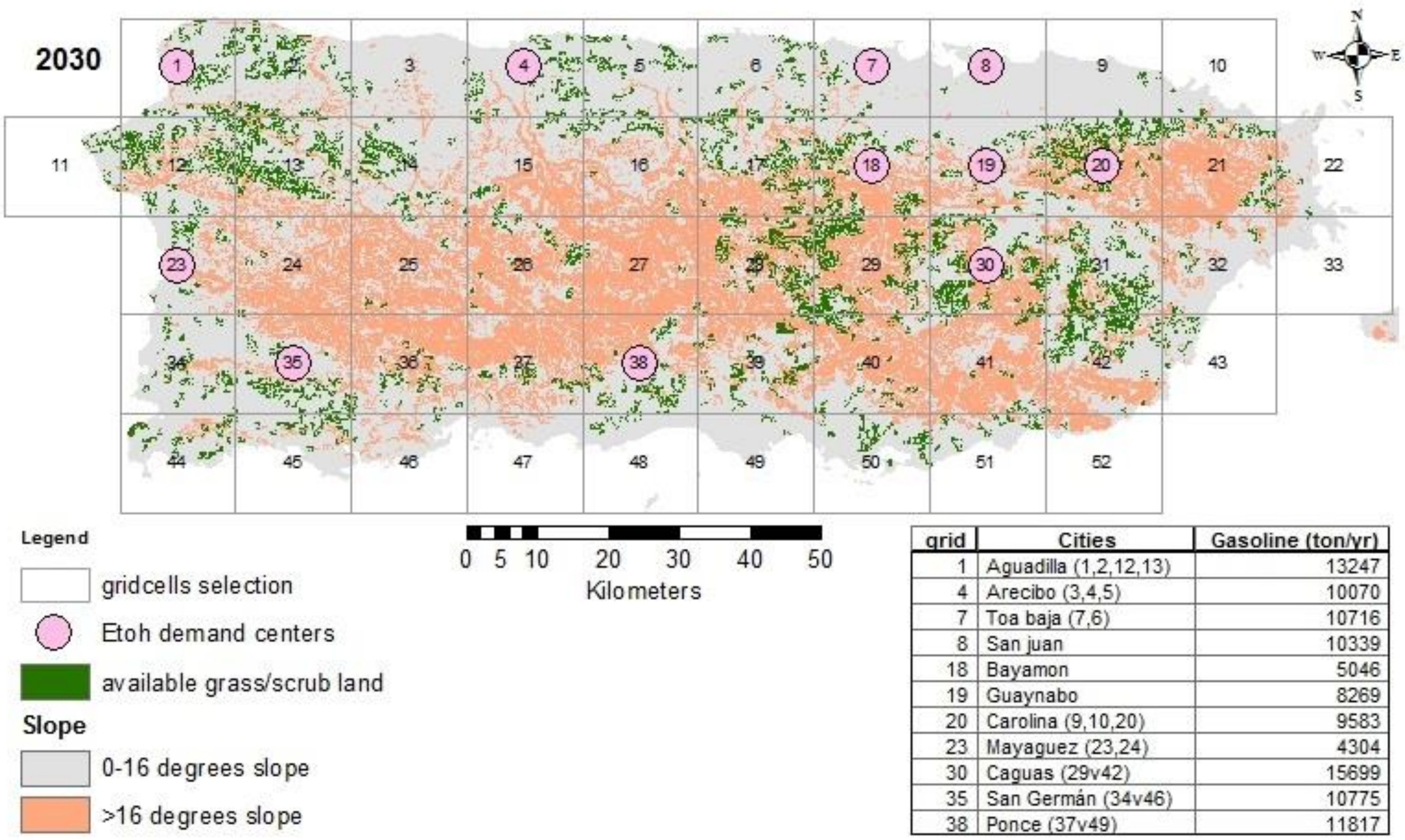
This paragraph describes shortly the input for MOMILP Model 3. In appendix 10.2. can be found a complete overview of the calculation method and the linear formulas that are used in the model. Chapter 4 'Models 2-3: MOMILP input' gives a detailed overview of the parameters that are described here.

The spatially specific, multi-objective, multi-period mixed integer Linear Programming (MOMILP) proposed in this study is derived from the methodology that Tittmann (2010) and Giarola (2011) use. 'Spatially specific' refers to the most optimal localization of ethanol and biomass production sites; 'multi-objective' relates to the dual optimization of both financial *and* environmental performance; 'multi-period' to a set of five short timeperiods of 3 years in which the 2014-2029 timeframe is divided into; whereas 'mixed integer' refers to the integration of binary, or yes/no variables into the model as to simulate the establishment of ethanol plants and/or biomass plots within each grid or not. The following inputs are necessary for the model (taken from Giarola, 2011):

1. Ethanol end users demand over the entire time horizon
2. Geographical distribution of demand centers
3. Biomass geographical availability
4. Geographical location of biomass production sites
5. Biomass production potential for each site
6. Biomass production costs as a function of geographical region
7. Technical (yields) and economic (capital and operating costs) parameters as a function of biomass type, production technology and plant scale
8. GHG emissions of biomass production
9. GHG emissions of ethanol production
10. Transport logistics (modes, capacities, distances, availability, environmental burdens and costs)

Input 1 and 2 are resp. necessary to establish the production target of the supply chain, and the locations this ethanol is blended with gasoline. The latter is partly influencing the locations where the ethanol plants will be constructed. Besides establishing the demand centers where the ethanol is blended, an realistic increase in ethanol output is to be decided on. The maximum potential ethanol production again depends on the available land, which is an output of the PLUC model. It is decided that the supply chain will produce and ethanol output equal to substitute 11,2% of the islande gasoline demand, and this amounts increases linearly to 22,4% over the timeperiod 2014-2029.

De demand centers chosen in this research are the ten biggest aggregated population centers in Puerto Rico. Figure 7 below illustrates the geography of Puerto Rico with the available land for bioenergy after a PLUC run, the ten demand centres where the ethanol will be delivered, and the slope of the island. The ten demand centres in Figure 7 represented a total annual gasoline use of 1,318,380 tons or 14 million barrels of gasoline per year in 2010 (Puerto Rico's total consumption was 25 million barrels/year of motor gasoline, eia, 2013). The demand for ethanol that will be delivered to these ten demand centers will be set at 20 to 40% of this gasoline demand, which equals a satisfaction of 11.2 to 22.4% of the islands total gasoline use over a 15-year timeframe. The demand centers mentioned in the attached table in figure 7 are not the exact location of the cities, but are the centers to which the surrounding gridcells are aggregated to. For example demand center 'Ponce' (grid 38) represents the population of grid 37, 38, 39, 47, 48, 49. Data on gasoline distribution on the island was hard to get, so the demand centers will have an ethanol demand according to their population ratio. For Ponce this is a total of 10,7% of the islands gasoline use, of which 20-40% of ethanol will be delivered by the ethanol supply chain by MOMILP.



grid	Cities	Gasoline (ton/yr)
1	Aguadilla (1,2,12,13)	13247
4	Arecibo (3,4,5)	10070
7	Toa baja (7,6)	10716
8	San juan	10339
18	Bayamon	5046
19	Guaynabo	8269
20	Carolina (9,10,20)	9583
23	Mayaguez (23,24)	4304
30	Caguas (29v42)	15699
35	San Germán (34v46)	10775
38	Ponce (37v49)	11817

Figure 7. Available grassland and scrubland in 2030, with ethanol demand centers and slope.

Inputs 3-4 on biomass geographic availability and location, are delivered by the PLUC model results. Biomass availability is the 2030 grass- and scrubland in 2030 after allocation of land to cropland, pasture land an urbanization. Biomass location is this land availability aggregated to the 52 gridcells (paragraph 2.2.a.iii. 'Linking PLUC and MOMILP').

Inputs 5-6 on biomass production potential and production costs are gathered in the field and further prepared in excel spreadsheets and presented in paragraph 4.1 ' Life cycle stage 'Biomass Production''.

Input 7 on technical (yields) and economical (capital and operating costs) parameters are the results from model 3 Excel linearizations of the technology configurations. This model will deliver the parameters (1) ethanol yield, (2) excess electricity production, (3) biomass requirement, (4) capital costs, (5) operating costs as a function of biomass type, conversion technology and plant scale. The capacities for ethanol plants will be 200 $MW_{HHV,in}$, 300, 400, 500, 600, 700, 800, 900, and 1000 $MW_{HHV,in}$. See paragraph 4.2. 'Life cycle stage 'Ethanol Production''.

Inputs 8-9 concerning GHG emissions will be gathered from literature and the field, and prepared in excel sheets.

The last input 10 on transport costs and GHG emissions is also prepared in excel (paragraph 4.3. 'Life cycle stage 'Transportation''). The key variables to be optimized over the *planning time horizon* 2014-2030 are:

1. Geographical location of biomass production sites
2. Biomass production rate and feedstock mix to the plant
3. Technology selection, location and scale
4. Transport modes
5. Financial performance (NPV) of the ethanol supply chain, including all life cycle stages
6. Cost of production per GJ and liter fuel together with a detailed cost break-down of cost factors
7. GHG impact on global warming including all life cycle stages.
8. By-product competition and credits (from excess electricity generated)

Key variables mentioned above that are optimized are crucial to prepare for a bio-ethanol industry on the island. The financial performance of the system can be specified further to deliver workable indicators for investors and policy makers.

2.2.c. Future trends

The first type of results only include an analysis of the NPV of the technology configurations over a timeframe of 15 years, where the lifetime of technologies is set at 25 years. Accompanying this NPV are the GHG emissions associated of the most optimal (and 2^{nd}, 3^{rd}, 4^{th} etc) options for a supply chain satisfying an ***increasing ethanol demand*** between 2014-2029.

Another set of results will follow this, and represent the NPV *changes* for the configurations in the future. If one wants to know what the production costs of ethanol are in 2020 (NPV of that similar supply chain from that year on, expressed in dollars per Gigajoule (GJ), a timeframe must be taken of 25 years discounting from that year on into the future, e.g. from 2020-2045. This is because technologies are discounted over a period of 25 years, and an NPV estimation for that year needs to include a timeframe with full depreciation.

One of the aspects of this thesis is to be able to also have insight in the future cost developments of ethanol. This will aid in the 'invest-now-or-postphone' investment decisions. The main trends that are identified in this research are:

(1) inflation correction for all cost calculations, i.e. a steel, fertilizer, electricity and diesel index for agricultural and ethanol production costs. For cost components that are not indexed, a general GDP deflator is applied to adjust for inflation;

(2) yield improvement of biomass feedstocks, i.e. sugarcane, energy cane and elephant grass. Yields in these crops are expected to increase between 20-60% -depending on the feedstock- in the next 20 years due to a better understanding of soil morphology, soil and crop management, better targeted water and fertilizer use and plant breeding;

(3) Ethanol conversion efficiencies, and are of influence on the financial and physical parameters of the supply chain.

This last trend is particularly difficult to incorporate in model 3 MOMILP due to the complexity of dimentions. But the following method was used to produce results in future cost estimations. So-called 'episodes' of 15 years are introduced, the duration of the MOMILP model.

The first 'episode' works with all the inputs (agricultural, transport, industrial) for the next 15 years and producing one NPV of a supply chain satisfying a ***constant ethanol demand*** for that time period. Because there is no increasing ethanol demand, neither are capital investments required in the remaining timeperiods as to satisfy this increase. Thus, all capital investments are done in the first timeperiod (2014-2017). all life cycle stage parameters are corrected by their 'trend parameter' except for the ethanol conversion efficiency parameter.

The next episode (also 15 years, or another complete MOMILP run) shifts one timeperiod of 3 years into the future, optimizing a favourable NPV while discounting for its next 25 years from 2017-2042 and so on. Now the ethanol conversion efficiencies are increased with $1/5^{th}$, that is the next 3 years in the 15-year time frame, and so forth.

The last 'episode' optimizes for a supply chain under the same ethanol demand and results in the NPV (time period 2027-2052) of such a supply system if it would be implemented in that year. For this year the ethanol conversion efficiencies are at the upperbound of the given efficiency range.

In this fashion we expect to clarify the effects of increases in biomass yields, production costs, and conversion efficiencies on the invest-now-or-delay decision making process. In other words, where and when will shifts occur from 1^{st} generation feedstocks/technologies towards 2^{nd} generation ones.

If a preferable supply chain is chosen, the optimization will be done again with constant, i.e. one average ethanol conversion efficiency and **increasing** ethanol demand. In this manner the choice for establishing an ethanol plant at any particular point in the timeframe is again included in the optimization exercise. The challenge to incorporate changes in ethanol conversion efficiencies into the future in the MOMILP model has not yet been resolved and could be done if more time and expertise is allowed.

3. Model 1: PLUC input and results

3.1. Initial estimation of available land (2013)

The first exercise in the extraction of potential land available for bioenergy crop production is to give an initial estimation of the amount of the potential land area, the location, and what the density is of this available arable land.

The following maps that are used as input data for the PLUC model are from the Geospatial Data Gateway (GDG) of the USDA Natural Resource Conservation Service center (NRCS-USDA, 2013): a Digital Elevation Model (DEM map); a TIGER 2010 transportation map containing primary and secondary roads; soils containing major land resource areas; and a TIGER 2010 census map. A land use/land cover map with 13 land use classes was retrieved from the Puerto Rico GAP analysis program (PRGAP), also conducted by the U.S. Department of Agriculture, Forest Service, International Institute of Tropical Forestry (Figure 8 below). This PRGAP land use map specified 70 different land uses, divided by the island soil, climate, vegetation, and elevation types. All maps are of high resolution, i.e. with cell size of 30-30 meters and that of th PRGAP analysis 15-15 meters. The land use classes were aggregated to the following land use classes (all 'developed' categories were later aggregated to the land use class 'urban':

1. Open Water	with less than 25% cover of vegetation or soil
2.1. Developed, Open Space	impervious surface is less than 20% of total land cover
2.2. Developed, Low Intensity	impervious surface is between 20-49%
2.3. Developed, Medium Intensity	impervious surface is between 50-79%
2.4. Developed, High Intensity	impervious surface is between 80-100%
3. Barren Land (Rock/Sand/Clay)	e.g. Sand dunes, strip mines, scarps, talus, slides
4. Evergreen Forest	trees higher than 5 meter; cover >20% of total land use
5. Shrub/Scrub	vegetation less than 5 meter; with canopy cover >20%
6. Grassland/Herbaceous	grammanoid or herbaceous vegetation; cover >80%
7. Pasture/Hay	areas of grasses or legumes for feed; cover >20%
8. Cultivated Crops	consumable crops, orchards, vineyards; cover >20%
9.1. Wetlands	soil periodically saturated or covered with water

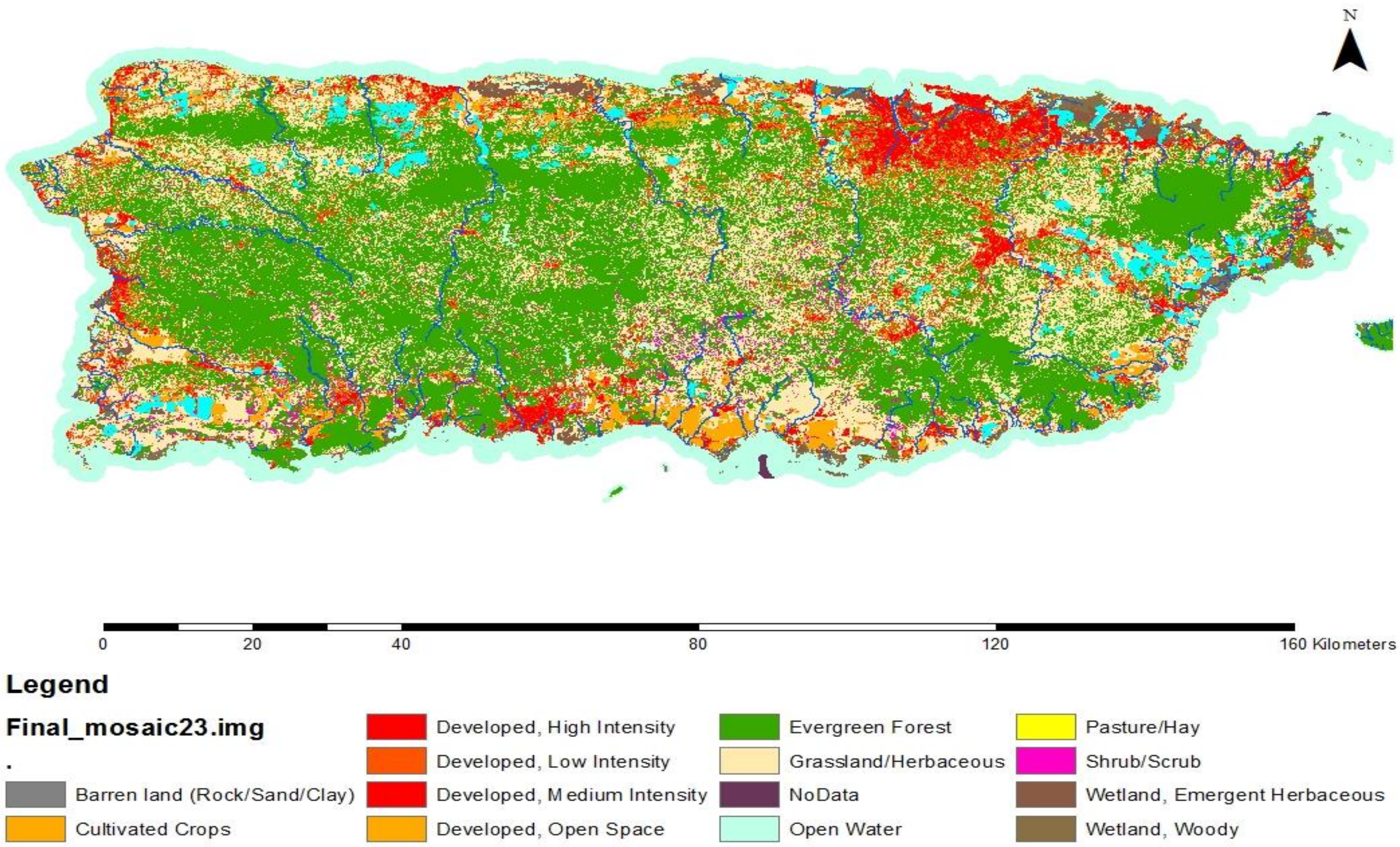


Figure 8. The land classes of Puerto Rico for the year 2006.

For a preliminary area estimation the following restrictions were applied. From the total area is extracted the area of open waters, all categories of developed spaces, barren lands, evergreen forests, wetlands, pasture/hay, and cultivated crops. These two latter categories are still important and will be analysed later after the island's future growth in agricultural and livestock intensity are accounted for. In this fashion land for energy crops can be freed-up due to more efficient cultivation processes. From these land use classes, grassland/herbaceous and scrubland is taken into account. Secondly, a slope map was combined to select areas with slopes smaller than 16 percent percent rise. As can be seen in Table 6 below, pasture/hay accounts for a total of 23.448 hectares of land; cultivated crops accounts for 19,696 hectares; grassland/herbaceous for 110,841 hectares (plots greater than 1 hectares or 41,801 ha if plots only greater than 100 ha are included); scrubland accounts for 3,987 hectares (plots between 1-60 hectares). Scrubland often represents a transitional area between forest and grass/herbaceous land, and is scattered over the island in such high degree that it can either be ignored individually, or converted to agricultural land altogether with grass-/herbaceous land.

Table 6. Amount of land for four land use classes; initial area, and net area potential.

Land use classes	Initial area hectares	Area after filter 1*	Area after filter 2**
Pasture/hay	23,448		
Cultivated crops	19,696		
Shrubland	24,182	3,987	
Grassland/Herbaceous	253,136	110,841	41,801
* plots > 1 hectare. **plots > 100 hectares			

Source: Junto de Planificacion (2007)

This land use class has seen significant growth over the past decades, due to conversion of mainly agricultural, grass/herbaceous and scrubland to forests due to industrialization as discussed in the introduction. Figure 9 illustrates the grass/scrubland availability in 2013 after all restrictions are applied. The regions with the highest concentrations of available land (numbered in the graph) are located in the:

2. Corner north-west, Aguadilla-Hatillo region
3. San Sebastian area below it,
4. region around Mayaguez,
5. large agricultural area in the corner south-west Lajas valley,
6. Center-south area around the city of Ponce,
7. Center-east region of Cajuas-Naguabo-Humacao triangle on the coast and inland, and at last in
8. the center-north left of the capital San Juan.

These areas are nearly all located along the coast in the so-called 'coastal planes', while inland rize up the forested mountains of the ‘Cordillera central’.

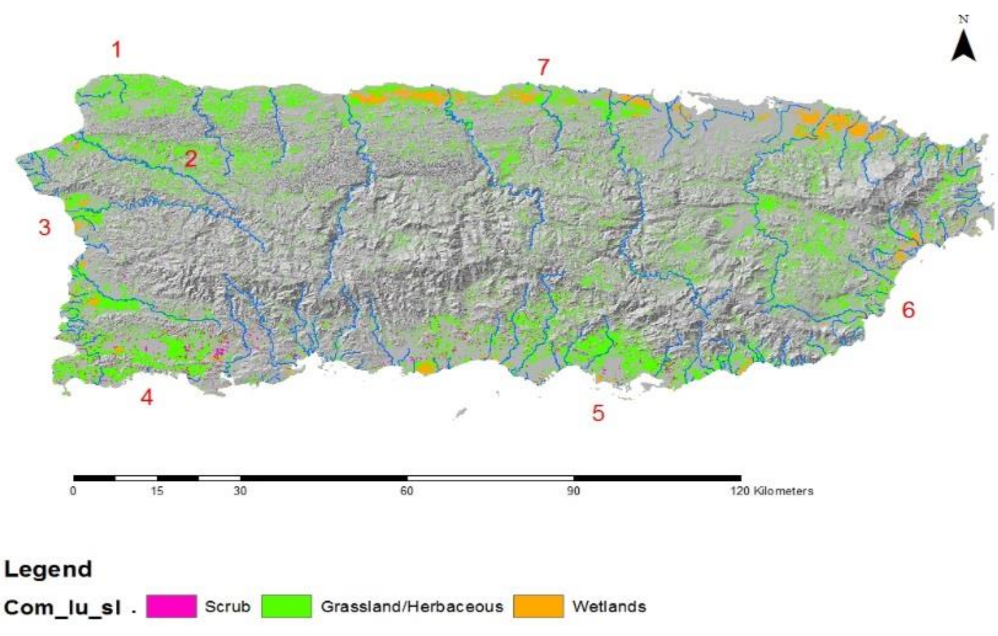


Figure 9. Available grassland and scrubland in 2006 dispersed over the island.

3.2. Data Input

One of the required parameters for the MOMILP optimization model 3 is the availability of land for dedicated bioenergy crops. As has been discussed in the methodology, the PLUC model allocates land to the three dynamic land use classes cropland, pasture land and urbanization satisfying a future demand that is calculated in the scenarios. The demand for products from these land use classes is influenced by the drivers of demand which are population growth, GDP and the *Self Sufficiency Ratio.* The future demand that is established in the scenarios is derived from the historical *output* and *yields* and its *assumed increases* in these sectors agriculture, livestock and urbanization. In this paragraph is discussed the SSR as most important driver for demand. Furthermore, agricultural and livestock output and yields, as well as population growth and urbanization are discussed as the most important scenario characteristics.

For the Business-As-Usual (BAU) scenario will be assumed a constant SSR over the timeframe, and a continuation of historical increases in output and yields. For the Progressive Scenario (PS) will be assumed a doubling of the SSR in 2030, which for agriculture equals a return to the 1970 output. This scenario assumes a progressive agricultural policy development where yields converge towards levels that resemble a modern agricultural sector. The yield parameters for this PScenario will be discussed in paragraph 3.2.b and 0 on 'agriculture and livestock intensities'. Only the available land of the Progressive Scenario will be used in the optimization process of the MOMILP model. The reason for this will be explained in the next sub-paragraph.

3.2.a. Self Sufficiency Ratio (SSR)

A strong shift from an agricultural towards an industrial and service oriented economy resulted in a neglected agricultural sector from the 50s on. The self sufficiency ratio resembles the portion of the islands local production for own consumption in its total consumption. Table 7 summarizes the sizes of local production, imports, exports and total consumption of consumables (Junta de Planificacion, 2012). From this table can be derived that the SSR is 17.65%. This is local production over total consumption, thus including local production for exports. The actual amount of consumed local production as share of total consumption (thus without local production for exports) is even lower, namely 10.84%. The PLUC model continues with a SSR of 17.65% and will thus represent the island's total agricultural production.

Table 7. Local production, imports, exports and an estimation of the Self Sufficiency Ratio of Puerto Rico.

(year 2010) Products	Local prod. (tonnes)	Imported (tonnes)	Exported (tonnes)	Net Import (tonnes)	Available for con-sumption (tonnes)	Ratio in 2010 Local production vs. Local consumtion (%)	Ratio in 2030
Milk & derivatives	320,747	622,422	30,129	592,293	913,040	35.13	70.26
Meat	50,017	440,369	2,737	437,632	487,649	10.26	20.52
Coffee/Cacao/tea	3,856	21,901	691	21,210	25,066	15.38	30.76
Fruits	60,706	682,526	34,397	648,129	708,835	8.56	17.12
Cereals	0	399,497	26,075	373,422	373,422	0.00	0.00
Vegetables	34,615	256,265	43,871	212,394	247,009	14.01	28.02
Legumes	410	21,703	389	21,314	21,724	1.89	3.78
Farinaceous/Starchy	130,757	260,244	23,704	236,539	367,297	35.60	71.20
Eggs	8,163	22,757	0	22,757	30,920	26.40	52.80
SOPAS/ESPECIES	1,719	25,239	10,304	14,935	16,654	10.32	20.64
Fish & Seafruit	1,947	43,634	927	42,707	44,654	4.36	8.72
Fats & Oils	0	85,325	13,407	71,919	71,919	0.00	0.00
Sugar	52	200,787	35,231	165,557	165,608	0.03	0.06
TOTAL	612,989	3,082,668	221,861	2,860,807	3,473,796	17.65	35.30
					Aggregated total	17.65%	35.29%

Source: Junta de Planificacion (2012). Consumo y Producion local.

Agricultural outputs have been declinging over the past four decades, as Figure 10 below illustrates. This is the island's total agricultural production in tonnes without sugarcane production, the latter accounting for nearly six million metric tons -or more than 12 times greater than the rest of the output- in 1970 (USDA, census 1970). The data in Figure 10 was derived by adding up the agricultural production of 22 commodities (excluding meat products) as reported by the USDA agricultural censuses. This source started to publicate very reliable and extensive data on the island's agriculture since 1954 (see next sub-paragraph 'Agricultural Intensity'.

With an increasing population this agricultural decline means also a decline in SSR. Modeling a negative SSR in the BAU scenario will not be challenging, since this means only a further abandonment of agricultural and pasture land up to 2030. On the other hand might it be interesting to see where the PLUC model will abandon this land first.

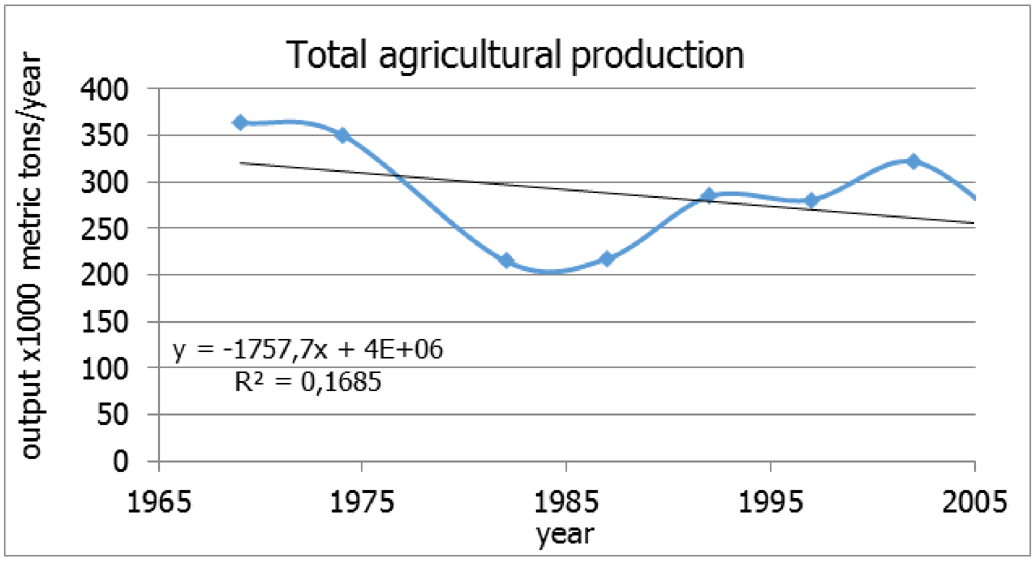


Figure 10. Agricultural output between 1970-2010. These are aggregated for all products produced on the islands in 1000 tons. Source: USDA Agricultural censuses 1970-2010.

For the progressive scenario, Prof. Luis Perez-Alegria and I decided to let a doubling of the SSR in 2030 function as a proxy demand, which equals a return to the agricultural output of 1970. We argued that this would be our –positive- contribution to the scientific and societal debates on the island concerning the role of agriculture in society in particular, and concerning land use management and -planning in general. The core reasoning behind this is the believe that within a sustainable society, an increased food independency is as important, if not more important than energy independency.

Agriculture has lost its importance and is replaced by more luctrative sectors in industry and services sectors. Its abandonment is nevertheless accompanied with higher unemployment (now 16%), lower food quality of food imports, high dependency on food imports (mostly from the U.S). A panel from the 'College of Agricultural Sciences (CCA) of the University of Mayaguez' underscores this and stresses the importance to develop a sustainable agricultural sector. The CCA estimates that the U.S. will significantly redirect exports of agricultural products towards local consumption by the year 2025 due to increased population growth, and accompanied loss of agricultural land. Furthermore, prices by that time will be 3 to 5 times higher as they are now (CCA, 2001). The panel furthermore underscores that 'Agriculture is not only about income and job generation, […] but it is also an intrinsic part of the island's culture and way of life'. We argue that the production of dedicated energy crops as a renewable energy source will induce a renewed interest in the agricultural sector. In this way the public-, stakeholders- and policymakers' attention is drawn to the possibility where bioenergy production and food production can exist along sides in Puerto Rico, even increasing food and energy independency considerably.

To reach this doubling SSR in 2030, the necessary assumption for growth in agricultural output is 5% per year between 2015-2020, an increase of 4% between 2020-2025, after which it levels of to 3.5% between

2025-2030. The same has been done for the demand for animal products where the highest output over 1982-2007 time period will be targeted in 2030. This means output increases for abovementioned 5 year time steps of 5.5%, 4% and 3.5% per year. Table 8 below summarizes these growth requirements, together with its according physical output. The table gives physical quantities of output for every five years, but since the PLUC model works with iterations of one year, these required outputs are translated to yearly future outputs.

Table 8. Calculated future output of agricultural and livestock commodities in tons at a doubling of the SSR. Percentage increases required are for 2014-2020, 2021-2025, 2026-2030.

Future estimations		year			
	Unit	2014	2020	2025	2030
Agricultural output	Metric tons/year	250,000	335.024	407.608	484.450
increase in crop output	Per annum %		5,0%	4,0%	3.5%
livestock output	Metric tons/year	75.000	97.200	120.700	150.000
increase in meat output	Per annum %		5,5%	4,0%	3,5%

Thus the self sufficiency ratio will function as the main driver for calculating future yearly outputs. In the next three sub-paragraphs are discussed the data that was gathered for agricultural, livestock and urbanization 'outputs' and 'yields', or the scenario parameters. With these findings are discussed first the current state of the sectors, and second the possibilities for future improvements. If a progressive output and yield growth is discussed, the PLUC model can allocate land to the dynamic land use classes until these outputs are satisfied.

3.2.b. Agricultural intensity

Puerto rico is considered a reasonable well developed island, with a moderate modernization of the agricultural sector. This entails mechanized crop production, medium intensity of cattle production and importation of high nutrision feed, and therefore a moderate feed conversion efficiency (FAOstat, 2013). Accesibility to inproved seeds, agro-chemicals, fertilizers, agricultural knowledge, irrigation and machinery is well established (Perez-Alegria, 2013). Furthermore, the island has an advanced agricultural university in Mayaguez with an international staff of agricultural scientists and several agricultural stations for knowledge production. In spite of the level of modernization in some agricultural subsectors, yields (especially crops) are not optimal due to a low local demand, a weak export stronghold, an uncertain agricultural framework and still a large share of smallholders in the total amount of farms. Additionally there exists a large share of land registered as farms that do not produce any crops, thus lowering yield levels even further.

For the scenarios are calculated an agricultural output (in tonnages per year), the amount of land needed for this production (in hectares), and an aggregated yield (in tonnage/hectare). This is also done for the livestock sector and urbanization. For the two scenarios will then be calculated a different continuation of both output and yields, as to investigate the effect on the availability of land for bioenergy production.

Agriculture data concerning outputs, land utilization and yields were derived from agriculture censuses of the USDA and the Junta de Planificacion (JdP, the island's planning board). The former source produces a very elaborate agricultural census every five years, and the latter co-produces and performs analyses (on) this data. Data from the USDA cencusses were analysed between 1969 until 2007. The censuses report production quantities for 20-25 types of products, ranging from vegetables, fruits, export products like tobacco and coffee, and sugarcane. Furthermore a detailed documentation is given on the amounts of land in farms, under cultivation, idle, used for grazing, with scrubland, and spefic land use per county (totaling 78), region (totaling 5), islandwide, and per farm size and crop type. The results of this literature search is found in Table 9 below.

Table 9. Agricultural land use and crop output data for the time period 1982-2007. 'Cropland for grazing' is a confusing category that can change on a yearly basis depending on the farmer's plans for his lands.

	Unit	1969	1974	1982	1987	1992	1997	2002	2007
Land in farms	ha	524.576	495.147	386.106	348.530	324.969	340.133	271.440	219.109
Total cropland, of which:									
.harvested	ha	71.634	88.177	72.395	50.964	69.122	70.789	78.295	45.666
..cropland for grazing	ha	134.507	142.119	106.000	90.295	73.971	113.887	70.215	80.417
..pasture/range other than cropland	ha	169.910	119.751	191.347	161.714	88.920	63.108	43.072	34.565
Total pasture (grazing cropland + pasture)	ha	304.418	261.870	297.347	252.008	162.891	176.994	113.286	114.982
Total crop production	ton/year	364.007	350.478	215.108	217.862	284.914	281.068	322.007	246.292
Crop yield	ton/ha/yr	5,08	3,97	2,97	4,27	4,12	3,97	4,11	5,39
Crop yield improvement	% p.a.	**0,24%**	(1969-2007)						
2014 PLUC output start	ton/ha/yr	**250.000**							

Source: USDA agricultural censuses 1982-2007.

From eight cencusses between 1969-2007, the agricultural output has been aggregated based on the weight of the products. This results in a total agricultural output in tonnages per year. For the amount of land needed for this production is used the category 'total cropland harvested' (see Table 9 for the other categories). Sugarcane production and land requirements for sugarcane are both excluded in these calculations. The reason is that sugarcane has a very high yield (averaging between 70-100 metric tons per hectare) while the rest of the foodcrops usually have yields between 5-20 tons per hectare. Especially because of its significant share in agricultural output in the mid 20th century it would give an inprecise estimate on the intensity of the foodcrop production sector. The intensity of the agricultural sector is then calculated by dividing the output by the area requirement, resulting in an islandwide tonnage/hectare intensity of the agricultural sector.

Table 9 (first line) shows a steady decline in the *total* amount of land in farms. This substantiates the general literature about agricultural land abandonment in the island (Del Már López (2001), Parez-Ramoz (2008), Colón-Guasp (2010)). Figure 11 below presents historical trends of the *total active* agricultural land, the output that is produced on that land, and the aggregated yield. The amount of land in active cultivation fluctuates over the time period 1969-2007, but the decline is also obvious. Fluctuations in this category can be described by relatively short-term decisions of farmers on how to utilize the land: to produce cashcrops; use it for cattle ranging; or let it lay idle for some time. In a broader sence this fluctuation can also be explained by the fluctuating marketing opportunities of crops for e.g. export. Figure 11 for example shows a remarkable downward trend in agricultural output and land occupation between 1982-1987 wich is caused by a downfall in plantain and banana production (from the censuses).This output is recovered again after

1987 and now supplies 100% of the island's demand. In spite of decreases in area and output, the aggregated yield still shows a steady increase.

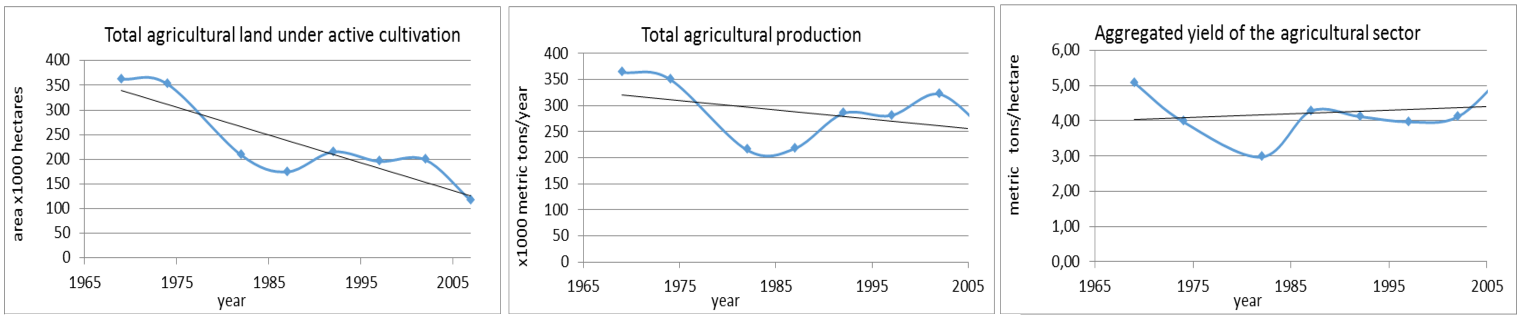


Figure 11. Agricultural land occupation, total output and aggregated yield.

As Table 9 above also shows, the *improvement* in aggregated yield is 0.24% per year. For the demand side in PLUC is furthermore needed an estimation of the 2014 agricultural output, which is 250.000 metric tons per year. For the BAU scenario is thus assumed a constant SSR (read: 'assume no further decline') and a continuation of a improvements in yield of 0.24% per year, increasing yields from 4.50 to 4.68 aggregated metric tons per year (if the *trendline* is followed) between 2014-2030.

Both yields (4-6 ton/ha) and yield improvements (0.24% p.a.) characterize a sector with poor dedicated agricultural policies. Yield improvements of 1-2% per year can be ascribed mostly to endogenous improvements such as farmers buying new machinery, production methods, seeds, and learning-by-doing. For the progressive scenario is assumed a more dedicated agricultural policy with a doubling of SSR and a yearly yield increase of 3%. The agricultural output will increase form 250,000 to 484,450 metric tons per year over the timeframe 2014-2030 (see previous paragraph 'SSR'). the aggregated yield will improve from 4.50 to 7.22 metric tons per hectare-year. These improvements are specifically propelled by an increase in crop intensity and further mechanization/modernization. In general this enthails increased land and crop management, a dedicated future agricultural vision and sthrengthening its stronghold and its outlets. In this case increasing the selling market is envisioned by increasing significantly the SSR or the share of local production in total consumption.

For the PLUC model will be inserted two data tables with yearly future (1) output and (2) yield. The model will allocate on a yearly basis the best suited available land to agriculture, pasture and urbanization. The results of this spatially explicit allocation will be elaborated on in paragraph 3.3. 'Land use changes over 2014-2030' and 3.4 'Final estimation of available land (2030)'. For now can be given a quantitative analysis of the differences in land use between the two scenarios, derived by pre-calculating the scenario effects in excell.

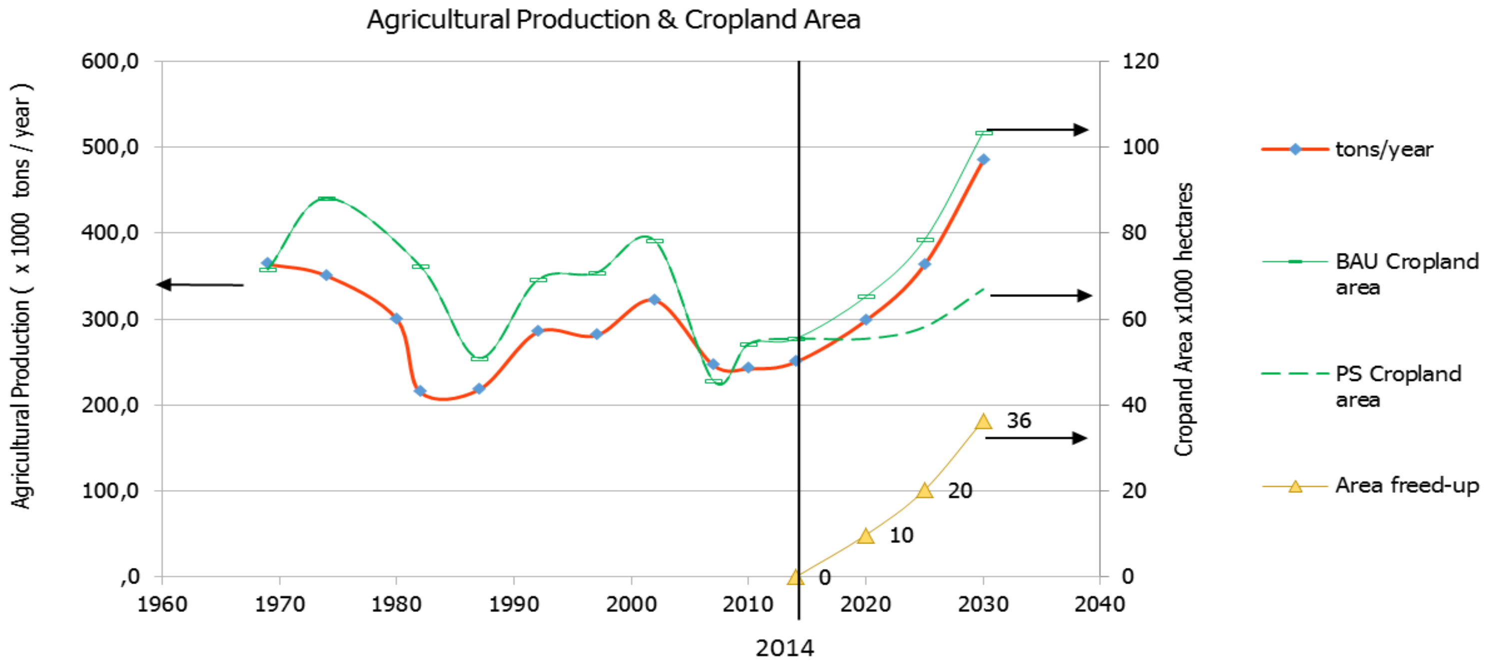


Figure 12. historical and future trends in Agricultural production and area of cropland.

Figure 12 above illustrates these findings in a graphical manner. The orange line is the agricultural output, and after the year 2014 it increases towards 484,554 metric tons per year due to a doubling of SSR in 2030 compared to 2014. Both green lines (cropland area) are identical before 2014, but afterwards they diverge due to the different assumptions in aggregated yield. It shows that the solid green line satisfies a doubling SSR with low yield increases of 0.24% per year and needs 103,444 hectares of land. The striped green line is the satisfaction of an SSR doubling attained with a progressive yield increase of 3% per year and requires a total land area of 67,000 hectares. Intensification of the sector will freeup 36,000 hectares in 2030 (yellow line) if the Progressive Scenario becomes reality. This together with the remaining grassland and scrubland can be used for dedicated bioenergy crop production.

The methods and results described in this sub-paragragh are repeated for livestock in the next sub-paragraph, and for urbanization in the one after that –although with slightly different units. For these two other *dynamic land use classes* are retrieved historical outputs and yield, whereas the two different scenarios will describe a differenct continuation of these two characteristics. For the PLUC model, these future estimated outputs and yields are calculated per year and drafted in data tables.

3.2.c. Livestock intensity

Livestock parameters are the second set of scenario characteristics that are gathered. As was done in the previous sub-paragraph, for this sector will be retrieved historical outputs, associated land use and yields also. The BAU scenario will assume a constant SSR in animal products compared to 2014 (read 'assume no further decline in output') and a continuation of historical trends in yield improvements. For the PS scenario will be assumed a doubling of SSR in animal products while yield improvements are further converged to a modern cattle sector. Data was retrieved from the same USDA agricultural cencusses between 1969-2007. Furthermore a starting livestock output for 2014 will be estimated, and depending on the yields two data tables with output and yield will be made as input for the PLUC model's scenarios.

The livestock output is calculated as the aggregated output of animal products in metric tonnages per year. the used land is the amount of land actively used for cattle. Two categories are identified and summed up: cropland used for grazing, including improved pasture; and pasture/rangeland other than cropland (see Table 10).

Livestock includes bovine animals (milk- and meatcows, bulls, and calves) and ruminants (horses, sheeps and goats; the latter two are also called small ruminants). Here it concerns livestock *yields*, this means that is accounted only for animals *sold* for meat or sold for other purpoces. The impact on grazing fields is different for bovine or (small) ruminants due to differences in size and thus in feed requirements.

To include small ruminants in the yields, a conversion factor of 0.1 is used based on feed requirements to calculate the weight of small ruminants within the livestock population. A mature cow of 500 kg consumes 10 kg of dry weight feed per day, whereas a goat or sheep of 50 kg consumes 1 kg of dry weight feed per day. Horses are assumed to have the same feed requirements as mature cows. To arrive at the outputs for livestock, an average weight of 750 kg for a mature bovine animal is assumed.

Table 10 below summarizes the findings of the data search between 1969-2007. Total pasture area is decreasing with one third over the investigated time period. Yields have been improving steadily from 1/3 of a cow per hectare in 1969 to 2/3 of a cow per hectare in 2007. Livestock output for the year 2014 has been estimated at 75,000 metric tons according to the trendline of this data (see Figure 13 below). Although the livestock sector is mostly characterized by extensive ranging, its aggregated yield is relatively high: 2.29% per year yield improvement over the time frame 1969-2007.

Land occupation	Unit	1969	1974	1982	1987	1992	1997	2002	2007
Land in farms, of which	ha	524.576	495.147	386.106	348.530	324.969	340.133	271.440	219.109
Total cropland, of which									
..harvested	ha	71.634	88.177	72.395	50.964	69.122	70.789	78.295	45.666
..cropland for grazing	ha	134.507	142.119	106.000	90.295	73.971	113.887	70.215	80.417
..pasture/range other than cropland	ha	169.910	119.751	191.347	161.714	88.920	63.108	43.072	34.565
Total pasture (grazing cropland + pasture)	**ha (sum)**	**304.418**	**261.870**	**297.347**	**252.008**	**162.891**	**176.994**	**113.286**	**114.982**
Livestock									
Total cattle and calves	amount	106.509	127.370	104.682	112.507	131.821	139.115	87.837	79.314
Horses	amount	918	999	680	1.076	1.159	2.582	1.793	1.075
Sheep	amount	1.535	809	1.752	2.045	4.089	3.319	7.742	4.394
Goats	amount	4.173	3.842	3.958	7.331	5.362	2.568	2.614	3.126
Conversion based on feed ratio (cow : other)									
Total cattle and calves	cow eq.	106.509	127.370	104.682	112.507	131.821	139.115	87.837	79.314
Horses (1:1)	cow eq.	918	999	680	1.076	1.159	2.582	1.793	1.075
Sheep (1:0.1)	cow eq.	154	81	175	205	409	332	774	439
Goats (1:0.1)	cow eq.	417	384	396	733	536	257	261	313
Milk	tons/yr								
Total livestock output (at 0.75 ton/cow.eq)	tons/yr	80.999	96.626	79.450	85.890	100.444	106.714	67.999	60.856
Livestock yields	ton/ha/yr	0,27	0,37	0,27	0,34	0,62	0,60	0,60	0,53
Output estimate 2014 for scenarios	tons/yr	**75.000**							
Livestock yield improvement	% p.a	**2,29%**	(1969-2007)						
Source: USDA agricultural census 1982-2007). http://www.agcensus.usda.gov/Puerto_Rico									

Table 10. Pasture land and livestock data (1982-2007).

Figure 13 below illustrates the livestock area, output and aggregated yield graphically. Livestock area has been decreasing steadily and remarkably. Livestock output is also declinging over the timeframe, but nevertheless can be seen a steady increase between 1982-1997 where cattle/calves sold grew from 104,000 to 139,115 pieces as export opportunities of meat products to the USA grew. Due to these fluctuation in area and output, the aggregated yield is characterized the same discontinuities. But over the whole time period a yield improvement of 2.29% per year is calculated, wich is a relatively progressive number.

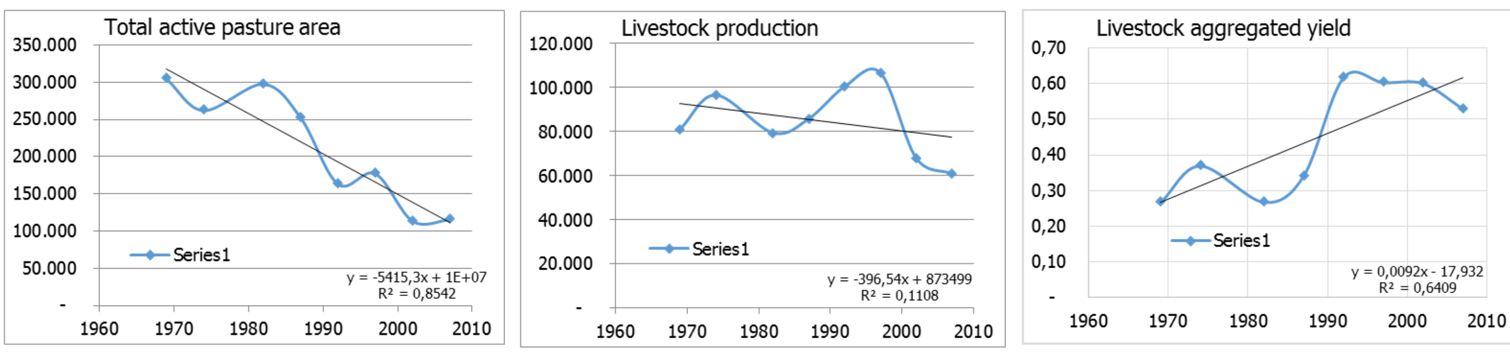


Figure 13. Agricultural developments in (1) pasture area (hectares), (2) livestock output (m.tons), and (3) livestock yields (tons/ha.yr).

For the BAU scenario will be assumed a constant SSR or output of 75,000 metric tons of meat products in 2030 compared to 2014. The intensification of the sector will continue the historical trend of 2.29% yield increase per year. This means a growth in aggregated yield from 0.70 to 1.00 metric tons per hectare in 2030.

For the PS scenario will be assumed a doubling of SSR from 75,000 to 150,000 metric tons in 2030. To reach this aim a average increase in output of 4.43% per year is required. A sectoral intensification or yield improvement of 3.5% per year is also assumed. Given the fact that livestock is currently ranged extensively with less than one cow equivalent per hectare, this assumption is reasonable. The aggregated yield will then increase from 0.70 to 1.21 metric tons per hectare.

3.2.d. Urbanization

Urbanization is a somewhat different dynamic land use class compared to the first two. It is not a 'food' commodity, and thus different parameters for this scenario character must be found. This has been established by assuming for 'output' the 'amount of population' that is added yearly. As 'yield' is then calculated the amount of 'inhabitants per hectare' added each year. This is an important land use class, as Puerto Rico has one of the highest population density in the world.

The mid-20th century has seen a tremendous urbanization, although the last past decades this trend was somewhat weekened. This urban sprawl in the '60-'80 was propelled by the shift from an agricultural, to an industrial-service orriented economy. In the '90 this shift was nearly completed, and thus a moderate urbanization will be projected into the future (Del Már-Lopez, 2001). Prime agricultural land is still subject to great competition between industrialization, urbanization, nature conservation and agricultural activities. Urbanized land accounts for around 10% of the island's surface, and is more than 1,5 times larger in size than agricultural land.

Population growth has seen a concaved increase over the last 5 decades, with an increase of around 0.97% p.a. between 1960-2007 (see Figure 14 below). Population growth was extrapolated with a concave trendline up to 2030 for both scenarios. As the graph illustrates, this also means a trend towards stabilization of population growth between 2030-2050. Urbanization, that is land with a 20-100% impervious surface, accounted for 89,500 hectares according to satellite imagery of the PRGAP analysis (PRGAP, 2008).

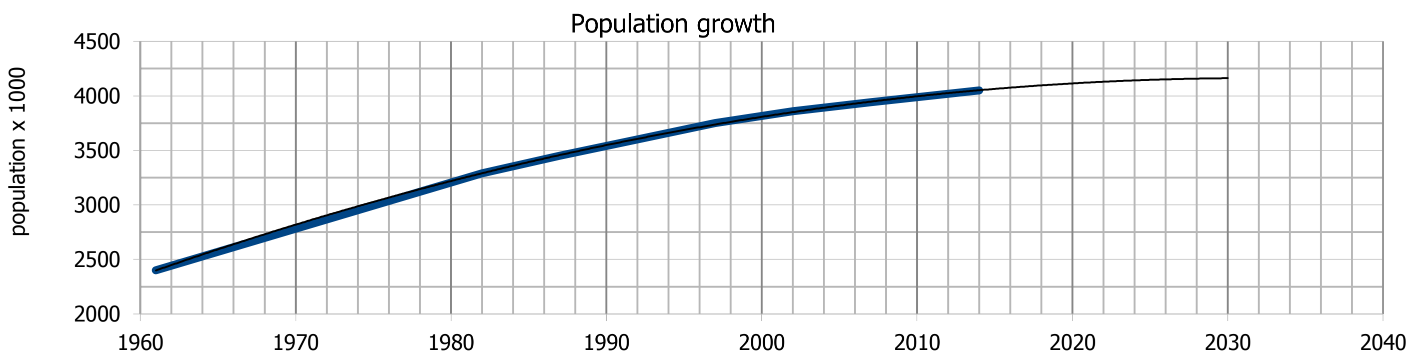


Figure 14. Historical population growth and its future trend (1960-2030).

The Urban 'yield' was derived by dividing the total population in 2007 by this area, which resulted in a yield of 46 inhabitants per hectare. This 'yield' will be held constant in the 2014-2030 timeperiod, i.e. no yield *increases* are assumed. One could think of trends in high-rise, redevelopment of abandoned urbanized areas etc. Unfortunately, historical data on urbanization statistics was deficient, and thus no *trends* in densities in the built environment could be extracted. In this research has been assumed a constant 'urbanization yield' of 46 inhabitants/hectare.

Furthermore, no yield map was developed for this land use class, which means that all (available land) can be converted to urbanization at whatever location. Nevertheless, the spatially specific allocation of urbanization is still governed by the suitability factors in the PLUC model, where (1) population density, (2) distance to roads, and (3) distance to cities receive the highest weights for this land use category.

3.2.e. Accuracy of satilite imagery

Although ArcGIS maps with very high resolution are applied in this thesis, there still exist some discrepancies in the size of (the same) land use classes between different maps. Documentation of land uses is very much dependent on the time the imagery is taken and the spectroscopy of the satellite imagery. Seasonal changes, for example, influence the density of a forest canopy which in turn reflects back different imagery to the satellite. Another difficulty can be that agricultural land planted with fruit trees will be documented as forest instead of 'agricultural land'. Yet another difficulty is the cultivation of coffee, which occurs mainly under 'evergreen forests'. Depending under which land use class this land will be categorized, either 'agriculture' or 'forest' will have a deficit in total area.

There are other issues that characterize in particular the changes in land use of Puerto Rico. One of these is the re-emergent forest cover on abandoned agricultural or pasture land. This type of tree cover is called secondary forest or scrubland/woodland and is characterized by a small selection of invasion types (mostly 2-3 varieties), fast-growing but pervious and with a very low overall biodiversity. Another particularity is the difficulty in estimating pasture land, or land exclusively dedicated to livestock. The livestock sector is characterized by extensive ranging, and cattle is ranged on a variety of land including cropland for grazing, pasture/rangeland other than cropland, grassland, and on scrubland/woodland.

One has to keep in mind that the use of spatial data is first of all a momentary, i.e. static analysis of land cover in time, and that a more comprehensive analysis of land use covers necessarily needs to be accompanied with fieldwork. Analysis of soil morphology, hydrology, elevation and taxonomy of biodiversity are all crucial to increase the precision in land cover analysis (PRGAP, 2008).

Table 11 below summarizes some sources with documentation on the size of land use classes. The first two, NRCS (2001) and GAP-Analysis (2006) are both satellite imagery produced by the USDA Forest Service, but the GAP-Analysis also included intensive groundwork concerning soil, water and biodiversity. It is a comprehensive work on biodiversity and its geographic distribution.

Table 11. Land surface comparison of the land uses between different sources.

Year of analysis	2001	2006	2000	2007
Land use types (in km2)	NRCS	(used in PLUC model) GAP-Analysis	Perez-Ramoz	USDA census
Cropland	197	342	263	457
Pasture	235			1150
Urban	1309	895	974	765
Grassland	2532	2476	2832	
Scrubland	242	1119	1095	
Forest	4049	3320	3417	
Reference:	[1]	[2]	[4]	[3]

Sources: [1] NRCS-USDA (2006) National Geospatial Management Center. [2] PRGAP Analysis-USDA (2008). [3] USDA Agriculture census 2007. [4] Perez-Ramoz (2008).

Both are very high resolution maps, with NRCS documenting in 30-30 meters resolution and 13 different land use classes, and GAP-Analysis in 15-15 meters resolution and 70 different land use classes. While of high resolution, both still document very different sizes of land use classes. The third source is from Perez-Ramoz (2008) who did a land use change analysis between 1990-2000 using Landsat satellite imagery of the Puerto Rico Gap Analysis acquired between 2001-2003. The fourth source UDSA (census 2007) is the latest agricultural census, one that has been used in extracting agricultural and pasture intensities of the previous sub-paragraphs. This source is expected to be the most accurate, since it gathers field data directly from farmers.

The main problem with a discrepancy in land cover sizes between (1) satalite data and (2) field data (that was prepared for the PLUC model, from demand and yield $\frac{Demand\ (tons/year)}{Yield\ (tons\ /\ hectare\ /\ year)} = hectares/year$) is a disproportionate allocation of land in the first year to whatever land use representing that gap in area. Solutions are to correct the data in the satalite imagery, which is a mathematical programming exercise, or to manipulate demand/yield in the PLUC model, or to adjust the amount of land in the first iteration wich is done here.

The spatial data of the GAP analysis will be used as the land use map for the PLUC model. The gap in cropland data compared to the USDA census is smallest. The reason why the census' cropland area is still larger than in the GAP analysis is due to the inclusion of coffee producing areas, which are shaded under close-canopy forest. This land use class ('coffee/evergreen forest') is documented in the GAP analysis too, and amounts to 1100 km^2 of forest, but in this thesis is kept allocated to the 'forest' land use class. Pasture land is not independently documented in the GAP Analysis at all but categorized together with grassland as 'pasture/grassland' (some 2400 km^2). The pasture are of the NRCS datasets (2001) amount to 235 km^2 and is documented as 'improved pasture' and 'hay production'. To deal with this pasture problem, the pasture area was copied from the $NRCS_{2001}$ map into the GAP_{2006} map, and is the *documented* area for pasture activities, excluding the remaining *undocumented* rangeland. This area will be the start location for pasture in the PLUC model around wich other land is added (according to suitability factor 'nr. Of neighbor same class', see paragraph 2.2.a.ii 'Scenario development' for suitability factors).

The problem concerning the gap in amount of land between satalite imagery on the one hand, and grounddata on the other, is solved as follows. The first iteration in the PLUC model (first year allocation) will occur with estimated 2014 output demand and yield. Land allocations to the tree dynamic land use classes in this first year will then be assumed to correct the satalite imagery to real-term amonts of land. One drawback of this method is that a large amount of land is allocated in the first year, which can distort the allocation logic. For example, in the first year is allocated 115 km^2 to agriculture, and an amount of 915 km^2 to pasture. Pasture allocation is significant and could occur on fertile agricultural land, that –if pasture's 'correction' allocation was phased over 5 years - could be directed to agriculture in the 2^{nd} or subsequent years.

An advantage of this great allocation of pasture land in the first year, is that a new land use class is added to the land use map, namely 'pasture/grassland'. The allocation in the first year to pasture occurred according to an accurate 'cattle density map'. Thus, where no previous spatial data sources had a precise documentation on the land use class 'pasture/grassland', i.e. where cattle was ranged, the PLUC model has now given an accurate estimate with respect to this.

This first year allocation is subobtimal, especially compared to the rest of land allocation over the 16-year timeframe (agriculture: 117 km^2. pasture: 169 km^2. These are the allocation results of the PS scenario, see next two sub-paragraphs). For agricultural land use planning purposes, especially concerning the spatial aspect, much more can and must be fine-tuned to let the PLUC model produce accurate results on allocation. But in this thesis the focus is on the allocation of best lands to the relevant land use classes over a safe period of time, before bioenergy crops are deployed. The method and assumptions described here answer to these requirements sufficiently, and the land extraction for bioenergy can now proceed to the optimization procedure of the MOMILP model.

3.3. Land use changes over 2014-2030

The most important results of the PLUC model are discussed in this paragraph. The ultimate amount of land that is available in 2030 will be input for the MOMILP model runs. First will be discussed the island-wide changes in land use classes. Then the extraction of grassland will be discussed in the next paragraph. Figure 1 below illustrates the results as the PLUC model. These are only the results for the progressive scenario land allocation. This is thus an allocation of land to the dynamic land use classes if their 2010-output is doubled in 2030, being delivered with yields that increase with 3% per year (agriculture) and 3,5% per year (livestock). Urbanization expands with a constant 'yield' of 46 inhabitants per hectare.

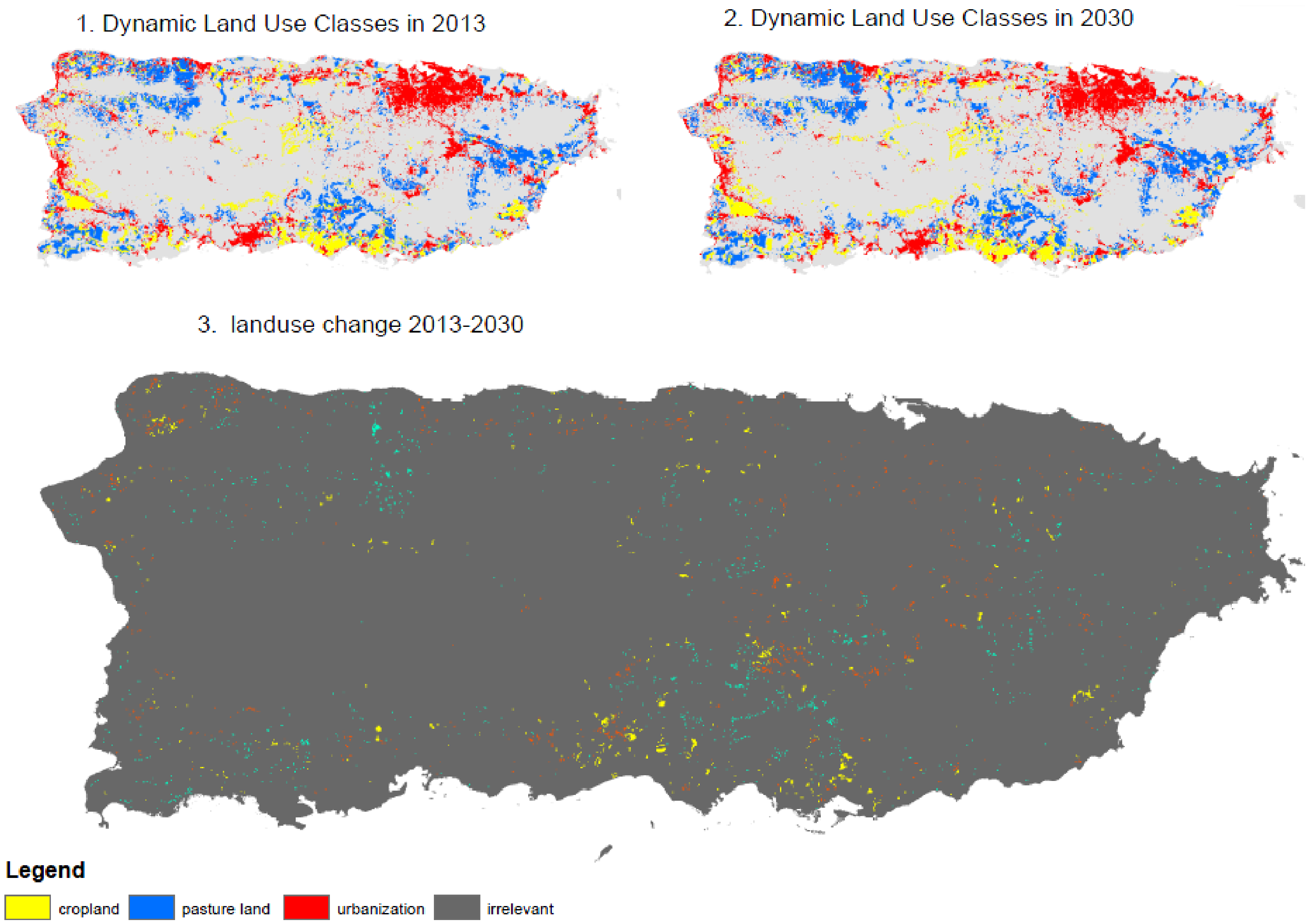


Figure 15. Land use changes between 2013-2030. Three dynamic land use classes are presented.

The first map is a land use map of 2013 of the whole island, including cropland, pasture and urbanization. All other land uses are grey for visualisation purposes. Map two is a land use map of the island for the year 2030. PLUC has been allocating land to cropland, pasture, urbanization according to their demand, as discussed in 2.2.a. 'PLUC Model 1'. For a better understaning of the *added* land, map 3 illustrates only the changes for these three dynamic land use classes. Urbanization (in red) has the smallest land use change. Expantion happens mostly near cities such as San Juan, but also in smaller conglomerates such as Cajuas, Aibonito, Cajey or Aguadilla. An amount of 28 km^2 has been urbanized over the timeframe 2014-2030.

Cropland (yellow) has also been extended, and mostly in the agricultural regions such as in Ponce (south-south-east), Lajas valley (south west) and Rincon/Arecibo (north-west). The amount of cropland increases with 116.5 km^2 over the timeperiod.

Pasture land expansion is made visible in light blue, and expands mostly in the areas with a dense cattle population: Hatillo (north-west); Ponce (south-south-east); Junco-Naguabo-Humacao triangle (east). Over the timeperiod 2013-2030 is added 169 km^2 of land for livestock purposes.

On a total island surface of 8900 km^2 these land use changes account for less then 2% change each. It is remarkable that a doubling in 2030 of the island's 2010-total production in agriculture and livestock (assuming progressive yield developments) can be accommodated with roughly the same amount of land that is already in use for these land use classes. Table 12 below summarizes the land use changes compared to 2014, and area savings between the scenarios. The netto added land in the progressive scenario compared to the current situation is 11,651 hectares for agriculture, 16,880 hectares for pasture and 2,826 hectares for urbanization. The scenario savings are based on a SSR doubling in both scenarios. Thus how much land is saved if a doubling SSR is established with progressively improved yields. In this case will be saved an amount of 36,515 hectares of agricultural land and 25,675 hectares of pasture land that can be dedicated to bioenergy crop production, next to the remaining available grass- and scrubland.

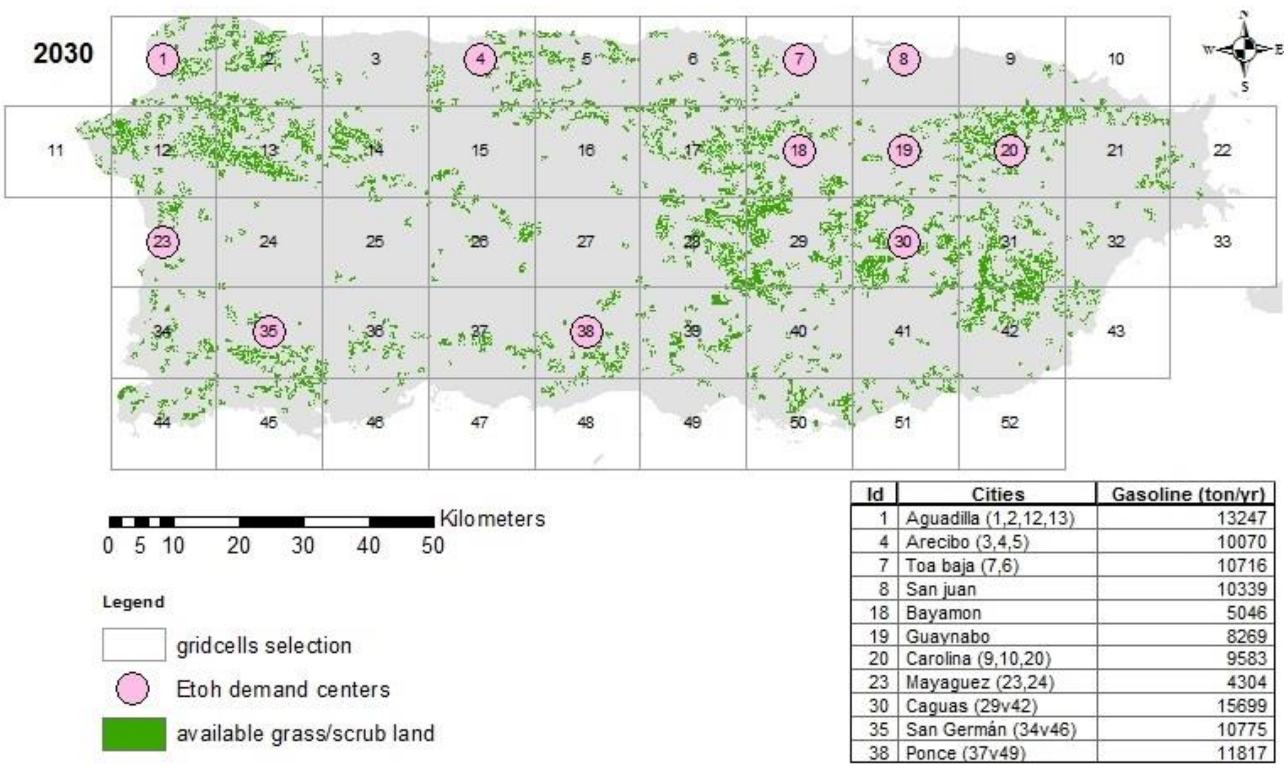
Area in year	Case	Scenarios	Description	Cropland	Pasture	Urban
2014	current		current SSR, current yields	55,432	107,143	88,043
2030	case 1	BAU	same SSR, low yield improve	53,465	74,849	90,870
2030	case 2		double SSR, low yield improve	103,599	149,698	
2030	case 3	PS	double SSR, high yield improve	67,084	124,023	90,870
2030	Netto added land		case 3 minus current	11,651	16,880	2,826
2030	Scenario land savings		case 2 minus case 3	36,515	25,675	0

Table 12. Quantitative land use changes (hectares): current land; netto added land; scenario area savings.

3.4. Final estimation of available land (2030)

The ultimate overview of available grassland and scrubland in 2030 is depicted in Figure 16 below. Also are illustrated the demand centers where ethanol will be transported to (the 10 demand centres in pink), and the grid cells to which the amount of available land will be aggregated to. The total available land for biomass production in 2030 is 82,978 hectares. There are three major regions; the San Sebastian/Rincon/Aguadilla/Arecibo north-west region (grids 1, 2, 11, 12, 13, 14); the Cajuas/Humacao/San Lorenzo eastern region (grids 30, 31, 42) and the central Morocovis/Baranquitas region (grids 17, 18, 28, 29). Smaller areas are scattered in the north and south, especially in the coastal regions. The south (grid 34-52) is the historically the agricultural region, but less available grassland is found there. This is mainly because much land in this region is allocated to crop-/pastureland in the PLUC model precicely for this region, i.e. high yields, close to already established farms.

Most of the grids have quite scattered grassland availability (see grid 34 to 40, and 44 to 52 in the south). This is partly because pasture and cropland had priority during the allocation of land over allocation to bioenergy. PLUC also allocated land in a scattered manner, instead of allocation large agricultural plots for monoculture. The grassland that is available in 2030 is then grassland that is left scattered between cropland and pasture areas. This is especially visible in the whole south region (grid 34-52). A real time planning of crop-pasture-bioenergy locations will therefore cluster this land more efficiently in commercial plots. In other words, in a specific region further planning is needed to indicate wich pasture-cropland plots can be better used for bioenergy en vise versa, for purposes of economies of scale.



Id	Cities	Gasoline (ton/yr)
1	Aguadilla (1,2,12,13)	13247
4	Arecibo (3,4,5)	10070
7	Toa baja (7,6)	10716
8	San juan	10339
18	Bayamon	5046
19	Guaynabo	8269
20	Carolina (9,10,20)	9583
23	Mayaguez (23,24)	4304
30	Caguas (29v42)	15699
35	San Germán (34v46)	10775
38	Ponce (37v49)	11817

Figure 16. Available grassland and scrubland in 2030, with ethanol demand centers and grid cells.

Figure 16 is accompanied with a table summarizing the gasoline use of the 10 largest urban centres on the island where the ethanol will be stransported to and mixed in. Data on geographic distribution of gasoline use was difficult find, so the demand centers are the population's portion in the total gasoline use. For example, if 1/10th of the population lives around demand centre 1 in Aguadilla, that demand center would account for 1/10th of the island's gasoline use. The hubs do not exactly represent the cities they depict, but represent more the urbanization around them. Demand centre 1, Aguadilla for example, represents a hub for villages and cities in gridcell 1, 2, 11, 12 and 13. Furthermore, six of the ten demand centers are located in the north-east, around the capital San Juan. Grid 8 represents the old and business centre, while grid 7, 18, 19, and 20 are highly urbanized regions around it. For all 10 demand centers is chosen to model a 20-40% gasoline substitution over the timeframe 2014-2030 which accounts for an *island-wide* substitution of 11.4 to 22.4%. This '20-40%-ethanol demand', together with the gridcells and their according land availability (and other parameters such as unit production costs of biomass per grid, biomass yields per grid) will be used in the third model MOMILP.

These geographic dependent parameters are tabularized in CSV format in excel, as Table 13 below illustrates for available land. For this parameter some other restrictions can be incorporated by omitting the grid availability. Some gridcells are occupied by cities, or have a too low land availability. In 2.2.a.iii 'Linking PLUC and MOMILP' was explained that a density filter was already applied in ArcGIS to exclude sparce land availability. In this table, gridcells with too low a density can also be omitted by simply deleting the gridcells. Omitting as much as possible gridcells increases the computation time of the MOMILP.

grid	km2	Excluding	grid	km2	Excluding	grid	km2	Excluding
1	22.2		18	31.5		35	23.7	
2	27.8		19	15.2	Guaynabo	36	11.5	
3	2.7		20	47.1		37	9.0	
4	21.6		21	18.1		38	21.1	
5	19.4		22	2.9	Density too low	39	21.3	
6	14.9		23	9.5		40	16.6	
7	6.1		24	1.6	Cordillera central	41	7.4	
8	0.4	City of San Juan	25	3.0	Cordillera central	42	30.0	
9	1.8		26	10.8	Cordillera central	43	6.5	
10	1.6	No land available	27	1.4	Cordillera central	44	12.2	
11	5.0	No land available	28	39.2		45	10.9	City of Ponce
12	43.8		29	45.5		46	1.7	Mountains
13	47.5		30	46.6		47	0.0	Mountains
14	22.7		31	46.4		48	4.0	Density too low
15	8.2	Cordillera central	32	5.2	Density too low	49	1.2	Density too low
16	12.0	Cordillera central	33	0.2	Density too low	50	12.4	
17	35.2		34	14.2		51	8.6	
						52	0.7	

Table 13. Land availability per gridcell. Some grids are excluded due to low density, or due to another land use.

Table 13 represents the most important link between model 1 PLUC and model 3 MOMILP. After each run of the PLUC model, this table is updated and inserted in MOMILP. This links is however accompanied with a loss in resolution, i.e. spatial land availability on a 30-30 meter grid resolution is aggregated to gridcells of

16-16 km. this is nevertheless a necessary process, as the data preparation requirement and computation power of the MOMILP model increases dramatically at higher resolution gridcells.

Gridbased parameters for the MOMILP model are those that vary with geography, i.e. biomass yields, biomass production costs and associated emissions. These parameters are discussed in the next chapter on the MOMILP input, and specifically in paragraph 4.1 'Life cycle stage 'Biomass Production''. The paragraph after that, 4.2 Life cycle stage 'Ethanol Production'' are presented the parameters associated with the industrial process of ethanol production.

From paragraph 4.1 'Life cycle stage 'Biomass Production'' are taken the production costs and the yields for the three biomass feedstocks, as to present a preliminary cost-supply curve of biomass cultivation (in $/GJ per PetaJoule). The yields and the production costs have been estimated for and extrapolated to 5 regions. For the energetic value of the feedstocks, and the production costs are used the following formulas:

Energy potential$_{region}$ (PJ) = Area$_{region,ha}$ * yield$_{wet,feed,\ ton/ha}$ * (1-mc$_{wet,feed,\%}$) * HHV$_{db,feed,GJ/ton}$ *10^{-6}

And:

$$\frac{\$/TON_{wet,feed}}{(1 - MC_{wet,feed}) *\ HHV_{db,feed}}$$

Where:

Area$_{region,ha}$	= Available land for biomass production (in hectares; PLUC results)
Yield$_{wet,feed,\ ton/ha}$	= Biomass yields per type and for 5 regions (in metric tons/hectare)
mc$_{wb,feed,\%}$	= moisture content of the feedstocks (60%, 54%, 35%, for sc-ec-elg, resp.)
HHV$_{db,feed}$	= Higher Heating Value of the feedstocks (17.9, 18.9, 19.2, for sc-ec-elg, resp.)

The result is shown in Figure 17 below. Note that the curves include 5 'stairs', representing the five region from which the biomass yields and production costs are derived. the supply curves are not complementary to eachother, since the available land is under competition by all three feedstock. At the given moisture contents and higher heating values, elephant grass has a total potential of 70 PJ that can be harvested between 4-4.5 $/GJ. Energy cane has a 55 PJ potential harvested at 6-6.5 $/GJ, with moisture content 54% and HHV$_{db,feed}$ 19.8 GJ/ton. Sugarcane is harvested at 7.5-8.2 $/GJ with a total of 45 PJ potential.

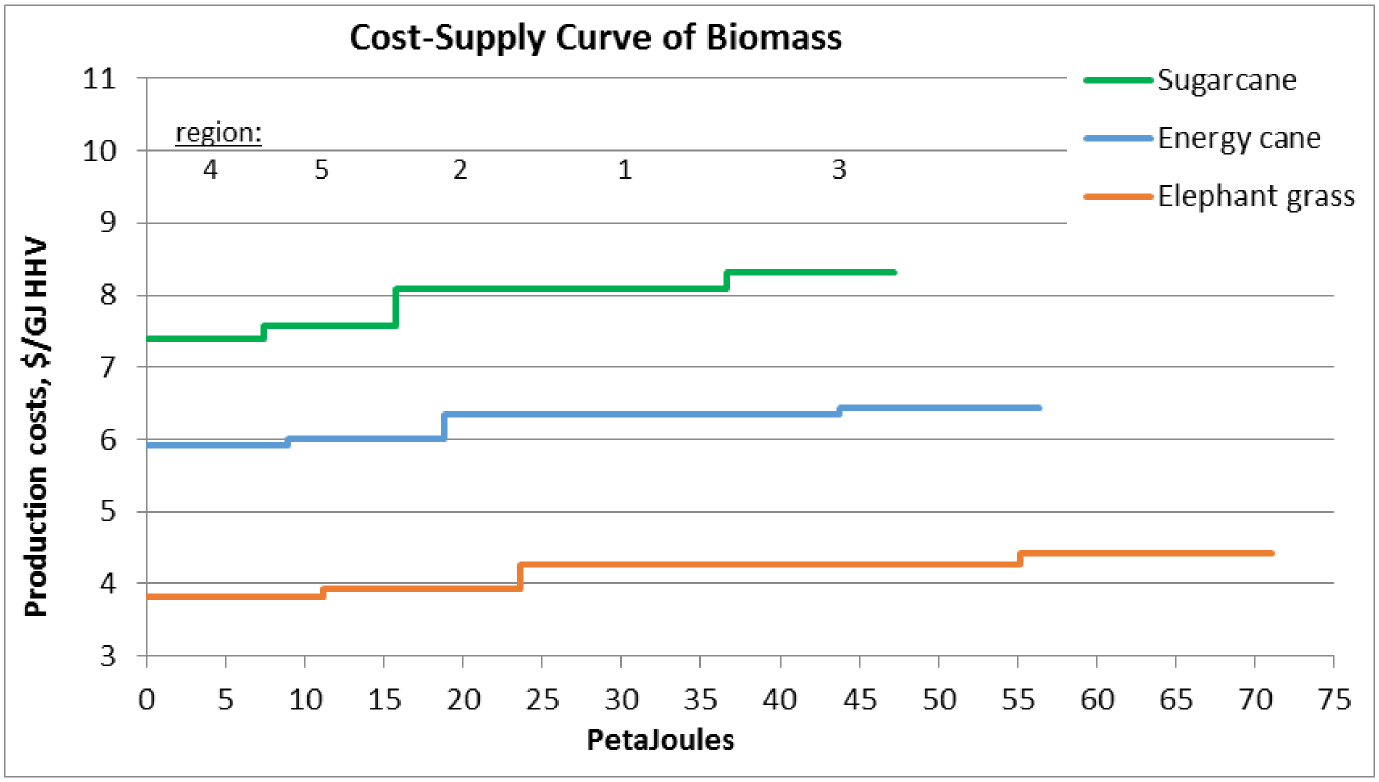


Figure 17. Cost-Supply Curve for sugarcane, energy cane and elephant grass, based on the 5 regions.

4. Models 2-3: MOMILP input

This chapter will present the remaining ethanol supply chain parameters that are needed for the second optimization model MOMILP. In the methodology it was explained that the MOMILP model is the main structure of this thesis. Complementary to this is used: (1) a PLUC model to deliver an estimation on the potential arable land for bioenergy production, discussed in the previous chapter; and (2) an Excel spreadsheet model to linearize the components of an ethanol production plant, part of the second life cycle stage 'Ethanol production'. This chapter presents the Excel linearization results of the production plant, together with the remaining parameters from all four life cycle stages.
Figure 18 below illustrates a flowchart for this chapter. For a full process flow chart of this thesis see paragraph 2.1 'Main structure'.

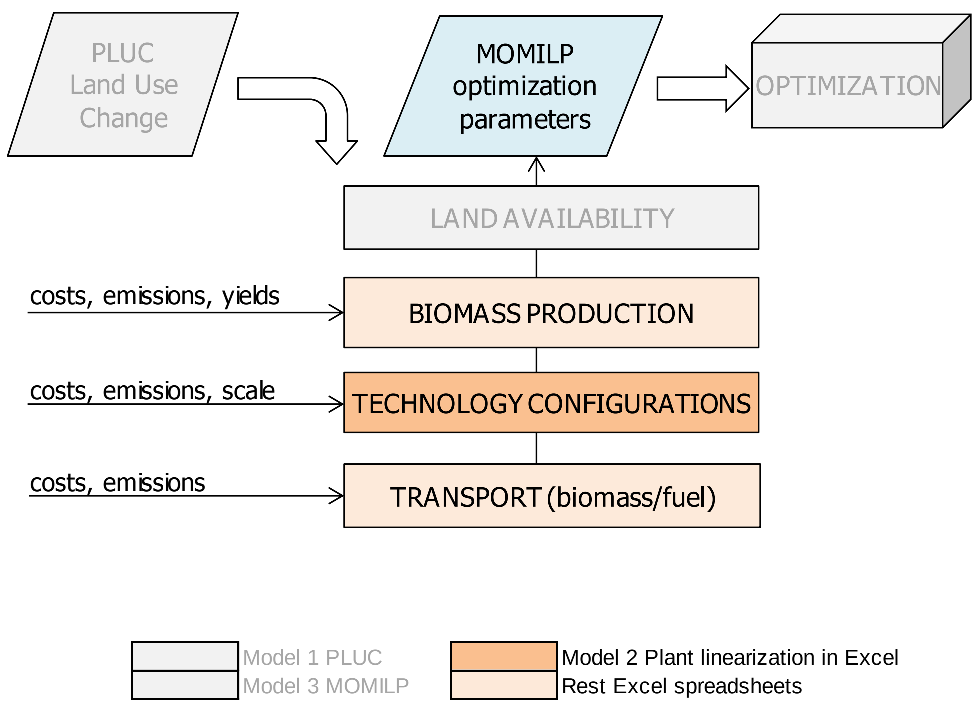


Figure 18. Process flow of this chapter: parameter gatherin from four life cycle stages.

The chapter is structured as follows. Paragraph 4.14.1. 'Life cycle stage 'Biomass Production" will present the expected yields that can be obtained in Puerto Rico, as well as the associated production costs and emissions from the cultivation of these biomass feedstocks. Three feedstocks are analysed and are sugarcane, energy cane and elephant grass. This data will be gathered from interviews or literature on the island, and prepared in spreadsheet into suitable parameters in the MOMILP model.

The second paragraph 4.2 Life cycle stage 'Ethanol Production' will present the method and results from the linearization model 2 from the excel spreadsheets. Three feedstocks are included, as well as the three ethanol conversion technologies: (1) the conventional 1^{st}-generation cane fermentation technology, (2-3) and two ligno-cellulose conversion technologies SSF and SSCF. Linearizing a production plant with a different feedstock will result in different configuration of that plant. In this way the three feedstocks and three technologies are combined and nine subsequent technology configurations are prepared. For each of these configurations are linearized the five main parts of a production plant: (1) pre-treatment, (2) Hydrolysis/fermentation, (3) Upgrading, (4) residuals handling, and (5) the power island. From these spreadsheets are produced technical ($yield_{etoh,electricity}$, GHG emissions) and economic (capital investments, operating costs) parameters as a function of biomass type, production technology, and plant scale.

The third paragraph 4.3 'Life cycle stage 'Transportation'' elaborates on cost and emissions associated with the transport of (1) biomass from plot to plant, as well as (2) ethanol from plant to demand center. For a more elaborate insight in the specific parameters and their preparation for the MOMILP model see appendix 10.2 'Model 3 - MOMILP Model description'.

4.1. Life cycle stage 'Biomass Production'

4.1.a. Biomass Options

The following feedstocks are investigated for an ethanol supply chain for Puerto Rico. It includes one 1st generation (1GEN) sugar crop and two 2nd generation (2GEN) ligno-cellulose crops:

1. Sugar cane (Saccharum oficinarum)
2. Energy cane (Saccharum spontaneum)
3. Elephant grass (Pennisetum purpureum).

Sugarcane (*Saccharum oficinarum*) for ethanol production is currently the most common feedstock in the conventional bio-ethanol production route. Sugarcane has a high sugar and low fiber content, and is currently crushed and milled and the resulting 'sugar juice' is fermented to produce ethanol. In Puerto Rico the most common cane variety is the PR-980 genotype; this variety is cultivated on the island through out the 20th century and will be used in this study. Table 14 below illustrates some crop characteristics that are relevant for technology scaling, such as the higher heating value (HHV), fiber content and total recoverable sugars (TRS). The TRS value is important for the conventional conversion technologies (1GEN), as this is the only source for ethanol production from fermentation. Bagasse is also included in the table, as this 'waste' product can be uses both for electricity production by incineration, or for ethanol production in a 2GEN conversion route.

Table 14. Feedstock parameters: fiber content, Total Recoverable Sugars, Higher Heating Value.

Parameters	Units	Feedstocks			
		Sugarcane	Energy cane	Elephant grass	Bagasse
Fiber content	% weight	13.5	26.7	78.9	87.0
..Content cellulose	% weight of fiber	41.6	43.3	49.80	41.6
..Content hemicellulose	% weight of fiber	25.1	23.8	42.20	25.1
..Content lignin	% weight of fiber	20.3	21.7	8.00	20.3
TRS-Total Recoverable sugars	% weight	13.92	5.90	-	-
HHV (dry basis)	GJ/ton	17.74	19.80	19.00	19.20 17.00 (10%mc)
References		[1] [2]	[1] [2]	[3]	[2]

Sources: [1] Alexander (1984), [2] Misook (2011), [3] Rego (2010).

Energy cane (Saccharum spontaneum) has –compared to sugarcane- a higher fiber and lower sugar content as well as a higher per hectare yield, which makes it an adequate feedstock both for conventional 1GEN and (ligno)cellulose 2GEN fermentation technologies. The type used here is the US-76-22-2 genotype investigated both in Hatillo (region 2) and Santa Isabel (region 4) of the island, both in experimental plots, and both delivering very high yields of 309 tons/ha.yr (Alexander, 1984) and 213 tons/ha.yr (Allison & Rios, 1988) total green matter, resp. Energy cane will be used both to optimize the harvesting window for sugarcane, and as feedstock for combined 1-2GEN or 2GEN technologies. As can be seen in Table 14 above, sugarcane, energy cane and its bagasse have similar contents of cellulose, hemicellulose and lignin. The overall fiber content of energy cane is roughly double that of sugarcane, and its content of total recouverable sugars roughly half of it. Due to its higher fiber content its higher heating value is 2 GJ/metric-ton higher. Even if energy cane is favoured for is fiber characteristics, its sugar content will still be used for fermentation, although the main focus will be on conversion of (hemi)cellulose to ethanol.

As a lignocellulose feedstock is chosen the perennial **elephant grass** type or *Pennisetum purpureum.* Elephant grass, native to Africa, is now grown all over the tropics as an animal forage. It is also one of the few non-legume crops that can fixate nitrogen from the air through so-called Biological Nitrogen Fixation (BNF) (Morais et all, 2009).

Morais showed in Seropédica, Rio de Janeiro, Brazil that five different genotypes of the elephant grass family were able fixate between 30-60% of their nitrogen requirements. This nutrient is required to biosynthesize basic building blocks of plants, such as amino acids for proteins and nucleotides for RNA and DNA. This ability for BNF makes this crop an adequate candidate to be harvested on infertile soils or areas where fertilizers are hard to get. Additionally, elephant grass abides relatively good in acid soils compared to the canes that need heavy liming and excess fertilization in such soils. Nevertheless, the focus in this study is on obtaining the highest attainable yield, rather than its advantage on invertile soils. Puerto Rico has abundant fertile soils and rainwater, good accessibility to fertilizers and little acid soils.

Tergas and Urrea (1985) found 53 and 72 tons/ha.yr yields with elephant grass in Colombia with intermediate and heavy fertilization, resp. They also found that elephant grass, more than the compared imperial and Guatemala grasses and sugarcane, has least variations in yield due to alternating dry-wet seasons. This relativizes somewhat the need for fertilization in dry periods.

Figure 19. Left: Sugarcane at 13 months old, Coloso, Puerto Rico. Courtesy of R. Conty (2006). Middle: ready-to-harvest energy cane at 12 months in Gaineville, Florida (Newenergyfarms, 2012)[5]. Right: elephant grass at 8 months old, Gainesville, Florida (Newenergyfarms, 2012).

4.1.b. Biomass Yields

Data for the life cycle stage 'Biomass Production' has been extracted mostly from historical data before 1998, since the sugarcane industry completely collapsed around that year. Puerto Rico was the seventh largest sugar-producing region in the world with an output of 10 million metric tons in 1950 (Fisher, 2008), and the industry has been declining up to a mere 26,000 tons in 1998. Data on the commercial sugarcane industry and its production parameters is very scarce, and the most recent source dating from 1996 documented yields of only 32 and 48 metric tons per hectare of sugarcane in resp. the west and east (Conty, 1996).

Other data sources that were found and evaluated were from Alexander (1984) in the north; Allison & Rios (1988) for the south; Perez-Alegria (2013) in the south west Lajas Agricultural valley; Conty (1996) for the west and east; Conty (2006) in the west; and from the USDA agricultural censuses (censuses 1959-2007). The first three sources were documenting sugarcane production parameters in experimental fields, where small plots were intensely cultivated with precision water and fertilizer supply, and optimal soil preparation and crop management. These sources focused not so much on cost factors, but more on optimization of yields. The only source found on the historical *commercial* sugarcane cultivation was from the US. Department of Agriculture, whom produced extensive agricultural cencusses between 1959 and 2007 (already mentioned in chapter 3 concerning agricultural and livestock production parameters).

This last data source is again used in the estimation of realistic attainable yields for sugarcane. From these yields are estimated the potential attainable yields for energy cane and elephant grass as well. The USDA documented production data for fall, spring and ratoon cane, for all sugarcane producing counties, and for five different regions the island was divided in for agricultural administration purposes (see Figure 20 below for the regional division) .

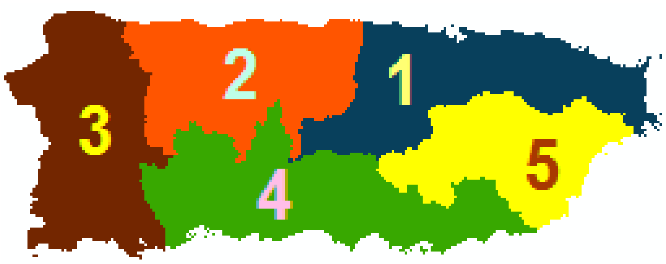


Figure 20. Regio division according to agricultural administration

This regional division will be used to extrapolate yields over the island. Regions were divided such that it represented best the climatic zones specific to the type of agriculture that is practiced there (Figure 20 above). Northern areas of the island are humid coastal and valley areas with dark loamy and clay soils, and commercial yields in the 20th century have varied between 65-110 metric tons per hectare. The south coast has a semi-arid climate, equally adequate soils but have shown average yields between 60-120 metric tons per hectare in the 20th century. The yields and biomass production costs that are discussed in this paragraph '

Life cycle stage 'Biomass Production" will be extrapolated to each zone (5 in total), then from each zone to their corresponding grids overlapping these regions (52 in total) to be used in the MOMILP, and at last rasterized to an ArgGis grid raster (cell size of 1 ha) for the use in PLUC.

Table 15 below summarizes sugarcane yield data that was extracted from the USDA censuses for the timeperiod 1959-1998, for the 5 regions in the island. The data illustrates a clear decline in yields in the last five decennia. Mid-century commercial yields were as high as 90 metric tons per hectare in region 4-5 (south), towards 80-86 tons per hectare in region 1-2 in the north. Although the climate is less favorable (semi-arid) in the south, this region has seen high yields due to a better targeted soil- and crop management combined with irrigation.

Higher yields in the south are mostly the result of a complex and a comprehensive irrigation system developed in the early 20th century consisting of artificial lakes in the central mountains that deliver water through a network of tunnels and channels to the southern agricultural regions (Perez-Alegria, 2013). Where the northern fields are usually prepared with drainage systems to alleviate the excess rain water, southern fields were managed with gravity irrigation and thus a more targeted water supply was possible.

Table 15. Average sugarcane yields for time period 1959-1998 (metric tons per hectare, dry based).

Year	Unit	Region 1	Region 2	Region 3	Region 4	Region 5	Source
1959	ton/ha	75	68	60	**89**	**82**	
1964	ton/ha	75	**75**	**67**	**92**	**91**	[1]
1969	ton/ha	**81**	74	**69**	85	80	
1974	ton/ha	72	72	67	71	72	
1978	ton/ha	**83**	**86**	62	69	73	[1]
1982	ton/ha	72	51	53	59	65	[1]
1987	ton/ha	69	45	58	60	64	[1]
1992	ton/ha			55	54	63	[1]
1998	ton/ha			32	25	48	[1]

Sources: [1] Census of Agriculture, USDA (censuse 1959-1998).

Figure 21 below illustrates the historical yield declines in Puerto Rico in a graphical manner. Yields usually decline in the subsequent years due to ratooning, i.e. yearly re-sprouting of the flora planted in year one. These yield declines can vary between 10-20% per year. Another one-time yield decline of 10-15% at the end of every 6-year-ratoon can be assigned to necessary stalks to replant sugarcane in the next crop cycle. These yield reductions have already been accounted and the presented yields in Figure 21 are the net millable cane tonnages delivered to the mill.

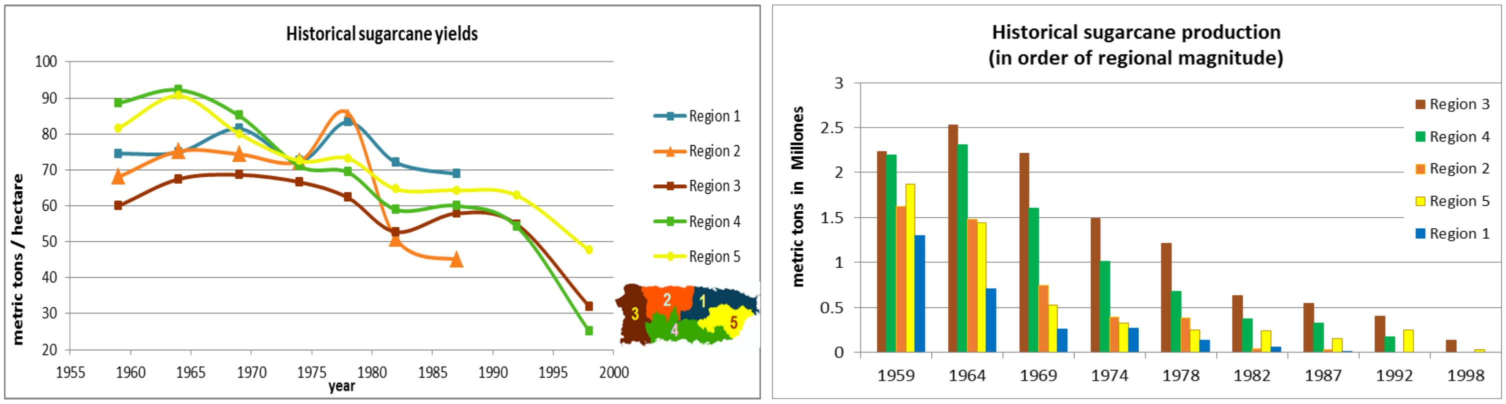


Figure 21. Sugarcane yields (metric tons/ha.yr) (left); and sugarcane output (metric tons) for the time period 1959-1998.

It is interesting to see the central role of region 3 in the sugarcane industry. It has been the largest producing area in the past (see Figure 21, right), followed by the south region 4. Dispite the region's major output, it has reported the lowest commercial yields of the island. In contrast, region 4 reports the highest yields on the island between 1949-1974 with yields somethimes up to 120 metric tons/ha.yr for some counties before 1960. Furthermore, trends in historical sugarcane output illustrate that region 1 and 5 saw the most dramatic decreases in output between 1959-1969.

To arrive at an estimate for commercially attainable yields for the timeperiod 2014-2030, it is proposed to average the two highest reported yields in each region. These yields will be called 'start-up' yields for convenience, as they are the yields that can be expected for a new, starting ethanol industry. It is reasonable to assume that yields from 60 years ago can be attained in the near future, especially concidering the higher level of technology and R&D that is available in the present. Table 16 below summarizes the 'start-up' yields that will be extrapolated to the five regions, and furthermore to the 52 grids for the MOMILP model.

Table 16. Highest and 2nd highest yields averaged for the 'start-up' yields, per region ($m.tons_{db}$/hectare).

Region [1]	Highest yield m.tons/ha.yr	2nd highest yield m.tons/ha.yr	Start-up yields (average) m.tons/ha.yr	Experimental yields [6] m.tons/ha.yr	Energy cane m.tons/ha.yr	Elephant $grass_{mc\ 50\%}$ m.tons/ha.yr
Region 1	83	81	82		102	57
Region 2	86	75	80	189 [2]	100	56
Region 3	69	67	68	180 [3]	88	48
Region 4	92	89	90	198 [4]	110	63
Region 5	91	82	86		106	60

Source: [1] Agricultural Census USDA (1959-1998). [2] Alexander (1984). [3] COnty (2006). [4] Allison & Rios (1988).

Data on energy cane concerned only experimental yields. A. Alexander (1984) compared a US-67-22-2 energy cane variety with a PR-980 sugarcane variety on a 7-hectare plot with equal fertilizer and water input (in the norgh region 2, Hatillo). After 12 months he reported yields of 215 metric tons per hectare of green

[6] These yields are total millable cane from the first-year harvest, thus not yet averaged over the ratoon period.

matter, of which 185 ton/ha millable cane. Fertilizer input amounted to 2040 tons per hectare of 20-5-10 fertilizer[7]. Allison & Rios (1988) experimented in the south region 4 (Sta. Isabel) and documented 243 metric tons per hecare green matter of which 193 tons millable cane after a 3 year ratoon. Due to lack of data on commercial cultivation of energy cane it will be assumed that this feedstock will be yielding 20 m.tons/ha.year[8] more than the 'start-up' sugarcane yields in all regions.

Unfortunately no data was found on elephant grass, although it is cultivated by some farmers as animal feed. The data for elephant grass is taken from experimental fields in Colombia (Tergas, 1985) and presents a theoretical yield of 66 m.tons/ha.yr with 1,000 tons/ha.yr of 20-9-4 fertilizer. The elephant grass yields for each region will be assumed 70% of the sugarcane 'start-up' yield at the same fertilization rate of sugarcane in that region. This results in a maximum yield of 63 m.tons/ha.yr in region 4, and 48 m.tons/ha.yr in region 3. Yield differences of sugarcane between the regions is assumed to represent soil fertility combined with water availability. This effect on yields is assumed the same for elephant grass.

<u>Trends in the future</u>. As initial yields for MOMILP's time period 1 will be assumed the highest yields, attained some 50 years ago. For future yields will be applied: a general sugarcane yield improvement growth rate of 0.96% per year; 2.39% per year for energy cane; 0.88% per year for elephant grass, similar to the Brazilian status quo.

Estimating the fertilizer inputs associated with the 'start-up' sugarcane yields seemed problematic since no quantitative data on fertilizer input was given in the USDA agricultural censuses. Fertilizer consumption was only documented as farm expense ($), and only for the censuses of 1949, 1959 and 1969. It is important to arrive at a good estimation for this, since fertilizer input accounts for a major part of the production costs. Second, a (reliable) geographic distribution of biomass yield and -production costs is the major factor in allocating biomass plots and –subsequently- ethanol plants later in the MOMILP optimization. Therefore, fertilizer inputs have been estimated by means of a 'yield-fertilization' curve that was compiled with (1) yield and input data from the experimental stations, and complemented with (2) fertilizer input *estimations* from the 1949-1969 USDA censuses.

Experimental yields reported very high sugarcane yields of 180-198 metric tons millable cane per hectare (Alexander (1984) for region 2, Rodríguez (1986) for region 3, Allison & Rios (1988) for region 4). These yields are first year harvests. To arrive at a more realistic yield these yields were corrected by accounting for a 15% ratoon loss for every subsequent year in a 7-year crop cycle, and a one-time 10% loss at the end of a ratoon for replant[9]. This experimental yield data is combined with data on soil characteristics concerning yield responses to fertilization.

---

[7] 20-5-10 is a indication of the percentage of N-P-K in the fertilizer: 20% weight content is Nitrogen, 5% weigth content is Phosphor, and 10% is Potassium. Before 1950, mainly 3-9-3, 9-6-3 and 10-6-0 was used (Capó, 1956). Between 1950-1970 more experiments with fertilization initiated a shift towards 12-4-10 or 14-6-8 use, i.e. it was realized that yields responded more to nitrogen, and less to phosphorus.

[8] Misook (2011) reports energy cane yields of 64 and 90 metric tons per hectare for resp. sugarcane and energy cane in Louissiana.

[9] Yields usually decrease with 10-20% per ratoon year. If the sugarcane is cut, and resprouted for the next year's growth, its output decreases. A crop cycle of 7 years is assumed. First 18 months: sugarcane is planted and harvested (plant cane). Then follow 5 ratoon cuts every subsequent 12 months (ratoon cane). The last half year the land is left idle to recover from the sugarcane cultivation and is often cultivated with nitrogen fixating crops like legumes or peanuts. An additional loss of 10-15% in yield includes sugarcane stalks that are needed for the next cycle's cultivation (Seabra, 2010).

The island's soils can be grouped in two categories according to sugarcane yield responses to N-P-K fertilization (Capó & Samuels, 1956). For this reason are compiled two yield-input curves: one for the south, region 4-5 and one for the north, region 1, 2, and 3 (Figure 22 below). In the humid north area (region 1, 2, north of 3, 5), yields respond significantly to nitrogen fertilization, and up to 31% yield reduction due to omission of nitrogen (Apendix 10.3). Region 5 is classified under 'humid area' but does not respond similary to fertilization as the other regions 1-2-3[10]. The semi-arid, but irrigated south area is less susceptible to fertilizer inputs and yields decrease by only 10% due to the omission of nitrogen. Sugarcane in Puerto Rico has never showed a high demand for phosphate fertilizers (Capó & Samuels, 1956). The highest response to phosphorus omission is in the humid area (8% yield decrease), and for the irrigated south it accounts for only 1% yield reduction. The same holds true for potassium, where responses are a 12% and 9% decrease, resp.

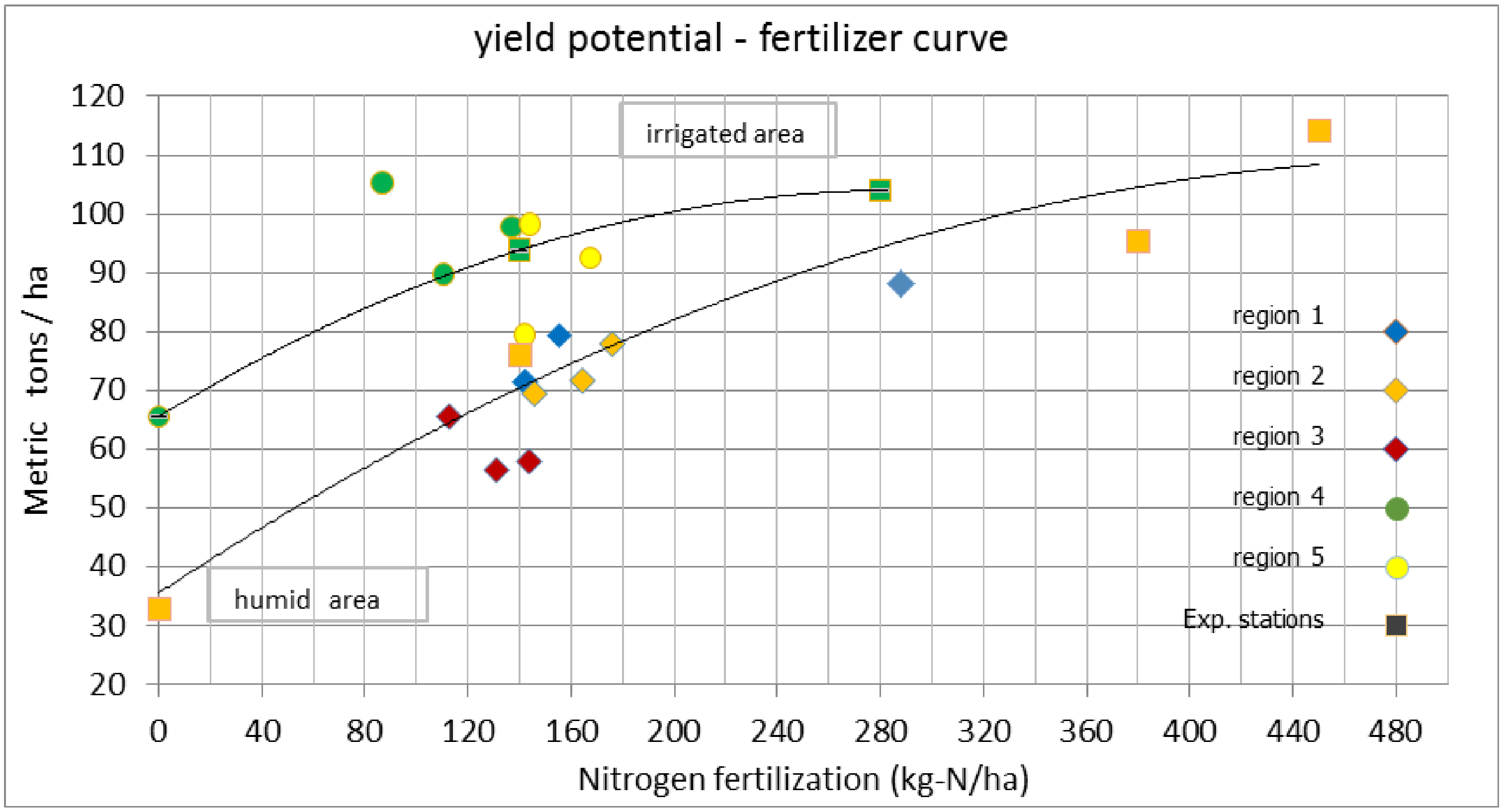


Figure 22. A potential yield-fertilizer curve for the Humid and the Semi-Arid region.

The second group of data points in the yield-fertilization curves of Figure 22 are fertilizer input estimations, extracted from the USDA censuses between 1949-1969. For 15 largest sugarcane producing counties –three for every region- are tabularized sugarcane outputs, fertilizer expenditures and area under sugarcane harvest. With a US fertilizer price index is estimated the fertilizer price for 1949, 1959 and 1969 to convert

[10] Región 5, the east of the island, is a humid region, but a major soil group 'Mabí-Rio Arriba' does not show a significant sugarcane yield response to omisssion of nitrogen (Capó, 1956). Therefore it is regarded together with region 4 as category 2, with characteristics from the 'semi-arid' soil response. The extracted USDA yield estimations also substantiate this, see Figure 22, and therefore will be plotted on the upper yield-fertilizer curve.

expenditures ($) into tonnages for each of the 15 counties[11]. Fertilizer estimates per three counties are hen averaged to represent the fertilizer intensity for that region. In Apendix 10.3 'Estimating fertilizer inputs 1949-1969' is explained in detail how this is done, what challenges were encountered and how these were overcome. Subsequently, these data points were added to the yield-fertilization curves in Figure 22 to increase its reliability. Figure 22 shows that the historical fertilizer requirements between 1949-1969 found for the 15 counties correspond reasonably with the curves. Thus, the points in the extreme right of the curves are yield-fertilizer results from experimental stations, and data points more in the center are from the agricultural cencuses. Concluding, from these two yield-fertilizer curves are read the fertilizer requirements that are needed for the 'start-up' yields. The results are shown in Table 17 below.
Since the 'start-up' yield are based on those yields documented in the 50s and 60s, and the common fertilization then was 14-6-10, the phosphorus and potash quantities will be estimated from the nitrogen estimations accordingly. For illustration the experimental yields are included in the table as well.

regions:		region 1	region 2	region 3	region 4	region 5	
Sources:		**(NA)**	**Hatillo (1986)**	**Coloso (2006)**	**Sta. Isabel (1988)**	**(NA)**	
crops:		Sugarcane	Sugarcane	Sugarcane	Sugarcane	Sugarcane	Elephant $Grass_{mc50\%}$
Experimental yields 1 (a)		(NA)	138	134	147	(NA)	
Experimental yields 2 (b) (c)			95	92	100		
Start-up' yields (d)	$m.tons/ha_{db}$	82	80	68	90	86	73.5
Material Input	**Unit**						
Nitrogen (e)	kg/ha	190	200	160	100	120	200
Phosphor (f)	kg/ha	73	77	62	38	46	88
Potassium (f)	kg/ha	132	138	111	69	83	41
Lime (g)	kg/ha	474	462	393	520	497	1000
Herbicides (h)	kg/ha	3.8	3.8	3.8	3.8	3.8	3.8
Insecticides (h)	kg/ha	0.35	0.35	0.35	0.35	0.35	0.35
Sources:			[1]	[2]	[3]	[4]	[5]

[1] A.Alexander (1984). [2] Conty (2006). [3] Allison & Rios (1988). [4] Conty (1996). [5] Tergas et all (1985).

Comments:

a. These are yields documented by the experimental stations, and are yields in the first year, excluding ratoon and replant losses.

b. Experimental yields corrected for 6-year ratoon losses of 15% p.a., and a one-time 10% replant loss. Another factor of 0.8 is applied to adjust from an experimental to a commercial nature of cultivation. Commercial cultivation does not include the precision soil- and crop management of experimental cultivation.

c. These yields are used to produce the potential yield-fertilizer curve.

d. The 'Start-up' yieds for the year 2014.

[11] Economic Research Service, USDA (2013). This US fertilizer price index is not fully applicable to Puerto Rico. Although most of the fertilizers were imported from the US, or locally produced by US companies, this index seemed too inaccurate, still. Prices derived from this index were again corrected by local data on the island's total fertilizer consumption. Apendix 10.3 'Estimating fertilizer inputs 1949-1969' elaborates in detail on the method followed.

e. Nitrogen inputs are read from the yield-input curve first, and then corrected considering the real data points, e.g. region 1 data point is (200minus10:83), whereas region 2's datapoint is (190plus10:80) because region 2 is relatively more fertilile than region 1. This is seen in the difference (=20 kg-N/ha) between the blue&orange points in figure 21.
f. For 4 out of 5 regions the 'Start-up' yields are from timeperiod 1959-1969. The most common fertilizer use in this period was 14-6-8. Therefore P and K is estimated based on this ratio.
g. Data on lime use was inconclusive, and only documented for a few counties and for the year 1959 (474 kg/ha for region 1). Seabra (2010) reports 450 kg/ha for a 86.7 m.ton/ha sugarcane yield. Therefore, 474 kg/ha lime for region 1 will be used, and for other regions the lime requirement will be scalled linearly based on N-fertilization input.
h. Neither data on herbicides nor insectisides was found in any literature, only for Hatillo (Alexander, 1986). As a default will be used data from Seabra (2011).

Table 17. Material inputs for the 'start-up' yields expected for the timeperiod 2014-2017, for 5 regions.

4.1.c. Biomass Production Costs

The previous sub-paragraph elaborated on the biomass yield potentials that can be attained in Puerto Rico. The island was divided into five regions and agricultural census data on sugarcane yields was extracted for these five regions. For the associated biomass production costs, the same extrapolation into the five regions will be applied, after wich these costs are extrapolated to the 52 gridcells in the MOMILP model.

First will be discussed the different data sources that were found on production costs of sugarcane: their deficiencies and usefull information. Then will be discussed how a realistic unit production cost is estimated using these different sources, how this is estimated for energy cane and elephant grass, and how this is extrapolated to the five regions.

For the production costs of biomass are needed a wide variety of cost factors, and these include cost estimations on (1) land preparation (land clearance/cleaning, roads, drainage); (2) cultivation costs (chemical imputs, labour, irrigation, seedings, land lease, administration, diesel consumption) and (3) new investments in machinery. For the GHG emissions are needed emission factors for fertilizers, chemicals and lime. At last are needed the previously discussed yield estimations to calculate the accurate amounts of fertilizers and chemicals, as well as to convert the production costs from \$/ha to \$/m.tons of biomass as parameter for the MOMILP model 3.

The same data sources of previous paragraph were evaluated. Experimental data focused more on yield optimization and less on cost factors (Alexander , 1984. Allison & Rios, 1988). Conty (1996) did a detailed comparative analysis between two producing areas in Coloso, region 3 and Roig, in region 5. The research included labor force characteristics, agricultural subsidies, harvest techniques and different farm practices per farm size (Conty, 1996). Also were included cost factors for machinery, fertilizers, chemicals, labor. Unfortunately this source was outdated, but more importantly reflected an industry at a point of near complete extinction. Thus the documented cost factors on fertilizer use, mechanization and labor supply were accurate, but delivering a yield of only 41 m.tons per hectare (Coloso) and 49 m.tons per hectare (Roig). The main difficulty remained finding accurate data on fertilizer and chemicals input at an optimal sugarcane production.

The fertilizer inputs that were estimated in the previous paragraph are combined with 2014 fertilizer prices to arrive at the fertilizer cost factor. Other cost factors were gathered from Conty (2006). He did a system analysis for the Hormigueros, Coloso area in region 3 (west of the island) of 11,400 hectares of sugarcane cultivation. As an agronomist at the Department of Agriculture of Puerto Rico, at the office of Economic Studies and Agricultural Planning in Coloso, west Puerto Rico, L. Conty is now preparing a restart of commercial sugar cultivation in collaboration with parties from the rum industry. From personal meetings was gathered valuable information on molasses and ethanol imports on the island, transportation costs and -capacities for sugarcane and other parameters discussed in their appropriate chapter. Furthermore, cost factors as presented in Table 18 on the next page are his estimations, complemented with the necessary data to arrive at the unit production costs (UPC) of biomass, as well as to the emission factor for the life cycle stage 'biomass growth', or $fbg_{i,g}$.

Section A presents all the relevant cost factors. These expenditures are grouped as one-time costs, ratoon-dependent or 'every.6.years' costs, or 'yearly' costs. These are grouped together and processed in a Net Present Cost calculation (20-year plantation lifetime) in section B. The cost factors are also grouped as either hectare/'area-based' 'region-based' or 'tonnage-based'. The former includes expenditures not dependent on yields such as land clearing, roads, drainage: the second concerns the fertilizer costs that are specific for each region: the latter is yield dependent, such as fertilizer/chemical input, labour supply, machinery. The 'area-based' expenditures are for all regions the same. This category cost factors was not find in any literature source (it concerns a totally new establishment of the industry) and thus is not specified spatially. The 'tonnage-based' expenditures vary per region according to a region's yield, and the initial estimation was from region 3 (Conty, 2006). New machinery (harvester & loaders) for example, was estimated by Conty as $30 million for a combined area of 11,400 hectares (or 29,000 cuerdas), at an average yield of 75 tons/ha.yr for region 3 this is 85 $/ton$_{sugarcane}$ for machinery. For the south, region 4 with average yield of 100 tons/ha.yr, this is thus an expenditure of 8480 $/ha (or 40 million for that same amount of hectares). This is done for all 'tonnage-dependent' cost factors.

This is of course a simplification of representing cost differences between regions, omitting e.g. advantages of fertile soil, specific fertilizer requirement, specific irrigation requirement of one region over the other, or differences in prices of machinery, fertilizers, labour. Nevertheless it was chosen to follow this method due to lack of cultivation data.
Section C presents the costs of fertilizers and chemicals. The costs are taken for Urea (0.45% N), Super Phosphate (0.45% P) and KCl (0.6% K) but are dollar prices for a 100% N, P, or K content. Section D presents emissions factors for these inputs and for diesel, to calculate the GHG emissions associated with biomass production. Section E presents once again the yields for sugarcane, energy cane and elephant grass to follow the 'tonnage-based' calculations. Section F presents the final results; Unit Production costs or $UPC_{i,g}$ per biomass type i and region; and the emission factor for biomass growth $fbg_{i,g}$ per biomass type i and region.

$UPC_{i,g}$ for sugarcane is on average 58-59 $/m.tons for region 1-3, and 53-54 $/m.tons for region 4-5. The latter regions do have higher overall costs, mainly due to more machinery & labour requirements (once and yearly costs, resp.), but these costs are divided by a larger biomass output. For energy cane the regional differences are flattened out somewhat, and the costs are 50-55 $/m.tons. elephant grass price ranges between 58-61 $/m.tons for regions 4-5 and between 65-68 $/m.tons for region 1-3. Emissions are between 16-25 kg Co_2 eq./m.tons of biomass for sugarcane, 13-19 and 23-25 kg Co_2 eq./m.tons for energy cane and elephant grass, resp. over the regions.

(A) Cost factor: (a	Unit:	Costs:	Cycle:	Dependent on: (b)			Source:
Clearance/cleaning	$/ha	1235.48	once	area	[1]		
Drainage	$/ha	247.10	once	area	[1]		
Roads	$/ha	247.10	once	area	[1]	[8]	
Cultivation Costs	(cost factors for region 1)					sc. (12-4-10)	
Fertilizer (c)	$/ha	483.74	yearly	region	kg/ha	1461.5	[8]
Lime	$/ha	43.49	every.6.yrs	region	kg/ha	450.00	[8]
Chemicals	$/ha	3.04	yearly	area	kg/ha	4.15	[4]
Machinery (b)	$/ha	2157.71	once	tonnage	$/ton	26.31	[1]
Labor supply	$/ha	2085.79	yearly	tonnage	$/ton	25.44	[1]
Irrigation system	$/ha	3681.32	once	area			[1]
Land preparation	$/ha	1235.48	every.6.yrs	area			[1]
Seeds & Planting	$/ha	741.29	every.6.yrs	area			[1]
Administration	$/ha	247.10	yearly	area			[1]
Land lease	$/ha	197.68	yearly	area			[1]
Diesel consumption							
..Agri operation (1st yr)		92.34	every.6.yrs	area	L/ha	102.60	[4]
(ratoon)	$/ha	8.19	yearly	area	L/ha	9.10	[4]
Harverster	$/ha	54.24	yearly	tonnage	L/tc	1.05	[4]
Loader	$/ha	8.42	yearly	tonnage	L/tc	0.16	[4]
Tractor hauler/trar	$/ha	19.42	yearly	tonnage	L/tc	0.38	[4]
..Other activities	$/ha	60.30	yearly	area	L/ha	67.00	[4]
Diesel price	$/L-2013	0.90					

(C) Fertilizer/Chemicals costs				Source:
Urea (N)	0,45 % N	$/ton 100% N	1231	[5]
Super Phosphat	0,45 % P	$/ton 100% P	1478	[5]
Kcl (K)	0,60 % K	$/ton 100% K	1078	[5]
Quicklime (2009)		$/ton	97	[6]
Herbisides (diuron)		$/ton	800	[6]

(D) Emission factos fertilizers, chemicals, diesel			[4]
Nitrogen	kg CO2eq/kg	3.97	[4]
Phosphor	kg CO2eq/kg	1.3	[4]
Potash	kg CO2eq/kg	0.71	[4]
Lime	kg CO2eq/kg	0.001	[4]
Herbicides	kg CO2eq/kg	25	[4]
GHG emission factor diesel	kg CO2eq/L	2.68	[7]

(E) Yields		sugarcane	energy cane	elephant grass	
sugarcane			(sugarc. +20)	(sugarc. * 0,7)	
regio 1	ton/ha	82	102	57	[8]
regio 2	ton/ha	80	100	56	[8]
regio 3	ton/ha	68	88	48	[8]
regio 4	ton/ha	90	110	63	[8]
regio 5	ton/ha	86	106	60	[8]

(B) total costs	$/ha	7568.70	once
	$/ha	2112.61	every.6.yrs
	$/ha	3167.92	yearly
NPC (20 years)	$/ha	3762.24	
NPC (20 years) su(	$/ton	65.54 (regio 3)	

(F)		regio 1	regio 2	regio 3	regio 4	regio 5
Unit production	sugarcane	57.78	57.69	59.37	52.86	54.07
costs	energy cane	54.22	54.08	54.88	50.46	51.35
$/metric tons	elephant grass	65.54	65.42	67.82	58.52	60.24
Emissions Biomass	sugarcane	22.89	23.97	24.53	16.27	17.98
growth	energy cane	18.40	19.18	18.95	13.31	14.59
kg CO2/ton-b	elephant grass	32.69	34.24	35.04	23.24	25.69

Sources:

[1] Conty, R. (2006). Coloso sugarcane business model.

[2] Alexander, A. (1986). Sugarcane and Energy cane in Hatillo.

[3] Tergas, L. (1985). Fertilization effects on forages in Colombia.

[4] Seabra, Macedo (2011). Life Cylce analysis of Brazil sugarcane products.

[5]National Agricultural Statistics Service, USDA (2013

[6] http://minerals.usgs.gov/minerals

[7] http://www.epa.gov

[8] USDA Agricultural censuses (1959-2007)

Comments:

a. This cost estimation is for region 1, and sugarcane.

b. Costfactors depend on area, region, or tonnage. Area-cost factors are equal for all regions. Region-cost factors are based on specific fertilizer e.a. specified in paragraph 'Biomass yield'.
tonnage-cost factors depend on the yield of a region, e.g. machinery and labour costs are estimations from Conty (2006), region 3, sugarcane. Tonnage-dependent cost factors are scaled for other regions dependent on the yield in that region. Conty (2006) estimated a required machinery investment of $30 million for the cultivation of 11,400 hectares of potential land around the Hormigueros area (region 3)

c. Fertilizer inputs are derived from paragraph 4.1. 'biomass yields' and are for sugarcane cultivation. It is assumed that the same input is used for energy cane, yielding 20 m.tons/ha extra in eve
For elephant grass is also assumed the same fertilizer input as sugarcane for every region, and the yield 0.7 of that of sugarcane.

Table 18. Cost factors for the life cycle stage 'Biomass Production' and associated GHG emissions.

4.1.d. Future trends

Trends for the future. As this thesis investigates an optimization for the next 15 years from 2014 to 2030, so are several trend developments incorporated to capture future changes in prices, biomass yields, ethanol conversion costs. As the choice for ethanol plant and biomass production plot may fall in every time period of 3 years, parameters representing that year must be used in the MOMILP model. For the timeframe 2014-2030 trends that are used in biomass yields are: 0.96% per year increase for sugarcane, or a total of 30% (similar to Brazilian estimations) (Regis et al., 2012); 2,39% per year increase for energy cane, or a total increase of 89.4% (Regis et al., 2012); 0.88% per year increase for elephant grass or a total of 26,7% (FAO, 2013).

Biomass production costs have been scaled with the relevant indexes for fertilizer (ERS-USDA 2013), steel, diesel (AEE, 2008)[12] and electricity (PREPA, 2011) prices and an overall GDP deflator (indexmundi GDP deflator, 2013) for other expenses such as labor, administration and land lease. In appendix 10.4 'Future trends in prices and GDP' are presented these historical price developments and subsequent future estimations of annual growth in prices. Technology parameters are also prepared with trends into the future, but these will be elaborated on in paragraph 3.3.1. on the technology configurations. Availability of land is also dynamic in the future, and will be discussed in a separate paragraph that follows now.

[12] Specific historical pump diesel prices could not be found. Instead historical data for distillate and residual #6 fuel oil prices have been used as guideline to predict future diesel prices. These two diesel fuels follow the same curve, and it is assumed that pump diesel follows the same curve.

4.2. Life cycle stage 'Ethanol Production'

This chapter elaborates on the conversion of ethanol within the ethanol supply chain. Different ethanol conversion routes that are chosen in this study are discussed in this paragraph. The first sub-paragraph 0 discusses the three main conversion technologies. The second paragraph 4.2.a 'Scaling Nine Technology Configurations' will discuss the technology configurations that are chosen for this optimization procedure. how installation components are scaled, and what subsequent characteristics of these technology configurations are. In excel is built a energy-mass balance spreadsheet for nine different technology parameters that will deliver the required technology parameters for the third model MOMILP. Technology Options

The following conversion technologies are investigated for this thesis:

1.Sugar juice fermentation (1st generation conversion technology)
2.Lignocellulose fermentation (2nd generation technology; SSF and SSCF)
3.Combined feedstock-to-ethanol (1st + 2nd generation; SSF and SSCF)

Since a bio-ethanol sector on the island is absent, the conversion technologies are newly built, i.e. no additional capital is installed at existing sugarcane factories with the purpose of (further) fermenting sugar juice or the bagasse. With this in mind, it could be possible that an ethanol production based on energy cane or elephant grass will be most attractive, skipping the conventional sugar juice fermentation route altogether. Nevertheless, Prof. Peréz-Alegría indicated that the rum industry on the island is very much interested in a local supply of molasses for the production of rum and consequently in setting up such a sugarcane-based supply chain (Peréz-Alegría, 2013).

Sugar juice fermentation (1st generation)
Figure 23 below illustrates the basic process flow of the conventional ethanol conversion technology which is called a 1st generation –or 1GEN- fermentation technology. It is usually combined with sugar production, i.e. the ethanol conversion route is added to existing sugarcane installations in large sugar industries such as Brazil, South Africa or Australia.

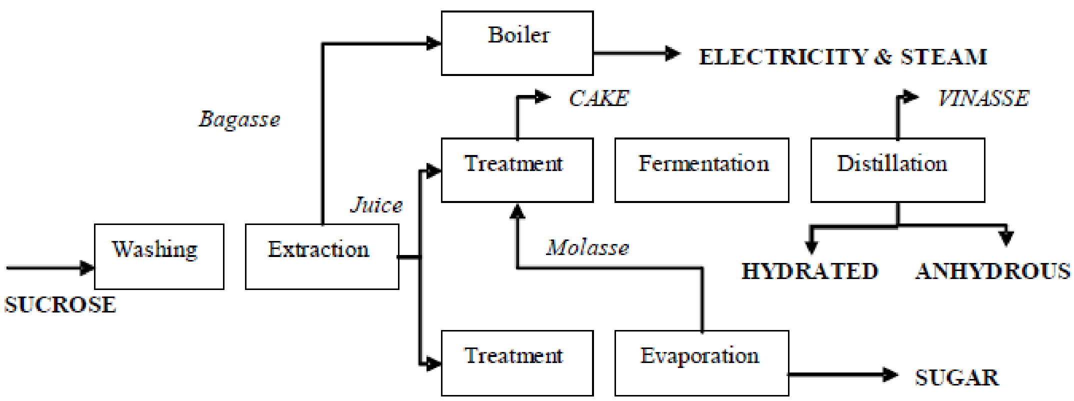


Figure 23. Process flow of a cane-based 1st generation ethanol conversion technology.

This addition entails a relatively simple incorporation of fermentation and distillation columns next to the sugar production train. Most of the rest industrial processes are shared, such as sugarcane washing, milling and further treatment. In such way a sugar installation now has an additional end-product to market, which, concerning the market demand, increases the economic optimization of the installation. Nevertheless, the sugar production is ignored, and a for this research this conversion option is focused solely on ethanol production. The production process is relatively simple, where the cane is washed and crushed, the sugar juice –or *guarapo*- fermented and distilled, and the remaining bagasse used in a boiler for process and excess steam and electricity.

(Ligno)cellulose fermentation (2nd generation)

Furthermore, two 2nd generation -2GEN- ethanol conversion technologies are included to enable a (ligno)cellulose based ethanol route. This technologies are one *Simultaneous Saccharification and Fermentation* (SSF) technology, and *one Simultaneous Saccharification and Combined Fermentation* (SSCF) technology.

Both technologies break the polysaccharides of cellulose and hemicellulose -the building blocks of plants- into smaller mono-saccharides (mono-sugars) by means of hydrolysis technologies using dilute acid, undiluted acids, steam explosion or liquid hot water (LHW). These monosaccharides, either 5 or 6 carbon ring sugars, are then fermented by yeasts. Thus four main processes are distinguished: C5 hemicellulose hydrolysis, C6 cellulose hydrolysis, C5 sugar fermentation, C6 sugar fermentation. Figure 24 below illustrates the general process flow of both technologies.

For the conventional 2GEN technology *Separate Hydrolysis then Fermentation* or SHF, C5-C6 hydrolysis, and C5-C6 fermentation occurs in separate units. Methods of hydrolysis and fermentation are subject to continuous improvements. The main difference in the two conversion methods SSF and SSCF is the optimization of the specific components within the plant, i.e. reducing the amount of hydrolysis and fermentation reactors.

SSF is an improvement on the conventional SHF technology that it combines the C6 hydrolysis and the C6 fermentation. Hydrolysis and fermentation of C5 sugars still occurs in separate units. The SSCF technology further integrates these processes by combinging C5 fermentation, C6 hydrolysis and C6 fermentation in one unit. Figure 24 below illustrates the process the difference in SSF/SSCF technologies in a graphical manner. The SSF technology is now readily available for commercial application. The SSCF technology is expected to become available around 2017-2020 (Hamelink, 2004).

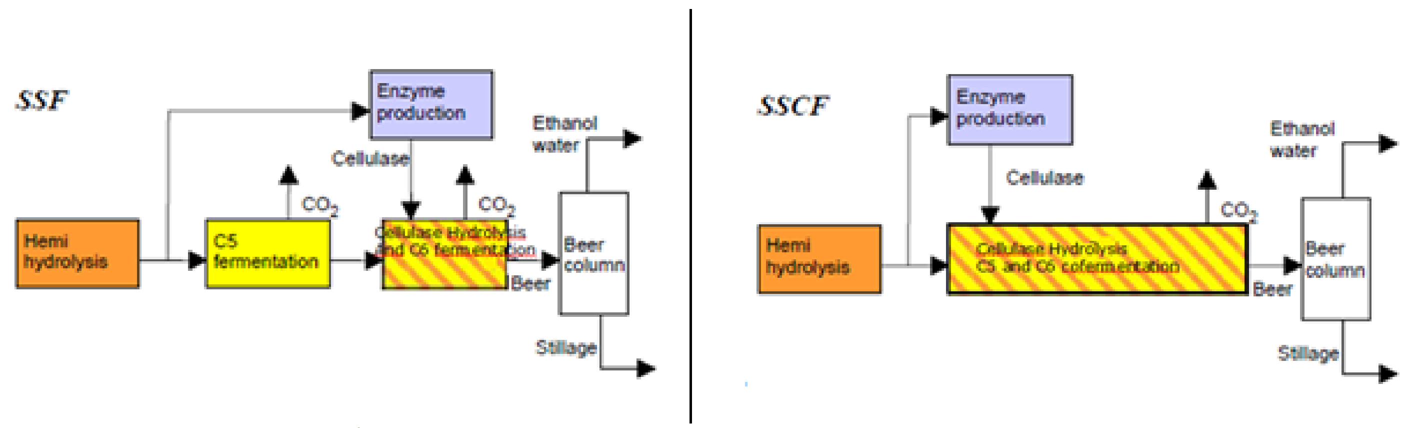


Figure 24. Process flow of two 2nd generation technologies: (1) Simultanous Saccharifiction and Fermentation (SSF) and (2) Simultanous Saccharifiction Combined Fermentation (SSCF).

The biomass feedstock is washed, cleaned, milled, chipped (~mm size) and then pretreated or hydrolyzed with dilute acid for SSF, or with steam explosion for SSCF. In both technologies the cellulose (C6) hydrolysis is enzymatic, and the enzymes are produced in a separate reactor 'on board', and not bought elsewhere (see Figure 24 above).

The SSCF technology represents -besides component integration- also other efficiency improvements which are: more efficient hydrolysis by shifting from dilute acid to steam explosion; less power use for the chipper, air compressors, auxiliaries and reactor vessels, mainly associated with electricity use in transporting the mass flows to and between the reactors; reduced use of process steam, also associated with a reduced number of units. The SSCF choice is still in the available/pilot stage and is assumed to become available in the next 3-6 years. This time dependent availability will be incorporated in the MOMILP model.

For this research will be used elephant grass as the main (ligno)cellulose feedstock. This technology will also be tested with energy cane and sugarcane as the main feedstock, and represent the 1-2GEN or combined 1st and 2nd generation conversion technology configuration.

<u>Combining juice- and lignocellulose fermentation (1st & 2nd generation)</u>

Thirdly, the integrated *guarapo*-(ligno)cellulose process is included, integrating both 1st and 2nd generation feedstocks. In the 1GEN technology the bagasse is fired in a Combined Heat & Power (CHP) unit to produce process steam and electricity. This bagasse waste product is also a (ligno)cellulose feedstock and can be treated further in a 2GEN installation. Both SSF and SSCF technologies are used, and are now fed with sugarcane or energy cane as their main feedstock. For the SSF technology, the *guarapo* will be treated separately as in the 1GEN option, whereas the bagasse is hydrolysed and fermented as in the 2GEN option. For the SSCF technology the guarapo fermentation will be combined with the C6 hydrolysis/C5-C6 fermentation unit.

For all technology configurations, specific components are also optimized by a 2nd and 3rd feedstock. Table 19 below summarizes all the technology configurations that are prepared in excel spreadsheets. In the cases where a cane is used as main feedstock (supplying 4000 hours/year cane), a second cane is supplied to increase the LOMS or lenth of milling season, and thus the operation time of the plant from 4000 to 5000 hours/year. In the 1Gen fermentation, SSF and SSCF technology configurations a third feedstock of switchgrass is supplied firstly to increase the load factor of the fermentation island from 5000 to 8000 hours/year, and secondly to optimize the power island (boiler and turbine) from 5000 to 8000 hours/year.

Table 19. Nine technology configurations with their feedstock and optimization characteristics.

	Name	k1	k2	k3	k4	k5	k6	k7	k8	k9
COMPONENTS	**Type**	1GEN	1GEN	1GEN	1-2GEN	1-2GEN	1-2GEN	1-2GEN	2GEN	2GEN
1GEN Fermentation		x	x	x	x	x	x	x		
SSF					x		x		x	
SSCF						x		x		x
CHP		x	x	x	x	x	x	x	x	x
Feedstocks										
Sugarcane		xxx	xxx		xxx	xxx				
Energy cane		x	x	xxx	x	x	xxx	xxx		
elephant grass			xx	xx	xx	xx	xx	xx	xxx	xxx
Load factor (h/yr)	Guarapo columns	4380	5000	5000	5000		5000		-	-
	Hydrolyse columns				8000	8000	8000	8000	8000	8000
	CHP	4380	8000	8000	8000	8000	8000	8000	8000	8000

The first three configurations are 1GEN. Configuration k3 is energy cane-based, and was included for its high electricity production. The succesive configurations are recurring SSF and SSCF choices, but subsequently with sugarcane, energy cane, or switchgrass as a main feedstock. It was necessary to assign a main feedstock to each technology configuration, since linearlizing the scale of the individual components is based on the specific characteristicis of a feedstock, such as fiber and sugar contents. Configuration k6 (SSF, energy cane) has, compared to k4 (SSF, sugarcane) a fairly bigger power island due to a larger fiber content and thus residuals stream. The detailed characteristics of each configuration will be discussed further in paragraph 4.2.a.

4.2.a. Scaling Nine Technology Configurations

For this second life cycle stage 'ethanol production', a second model in excel has been prepared where the plant components are scaled according to an energy-mass balance analiysis. This is done for 9 technology configurations discussed in the previous sub-paragraph. The following parameters are then derived from these spreadsheets, as a function of biomass type (i), conversion technology (k) and plant scale (p):

1. Total Capital Investments or $TCI_{k,p}$
2. Ethanol production costs $c_{'slope'}$
3. Ethanol production rate ε_p
4. Exceeding electricity production ω_k
5. Biomass requirement $\beta_{i,k}$
6. Biomass-to-ethanol conversion ratio $y_{i,k}$
7. Ethanol emission factor ffp_k
8. Emission credits from excess electricity production fec_k.

The capacity series or MW_p are 200, 300, 400, 500, 600, 700, 800, 900, and 1,000 MW of Higher Heating Value of the biomass in (HHV_{in}).

Table 20 below summarizes the production components, their base scale, base costs, scaling and installation factors. The maximum size of the components indicate the threshold capacities after which the components will not be scaled anymore, but placed in parallel. The installation factor includes additional costs such as indirect costs, start-up and working capital (buildings, piping etc). The components are scaled as follows:

$C_1 = C_0 * (\frac{P_1}{P_0})^r * x$; where

C_1 = the cost of the linearized technology
C_0 = the cost of the base scale
P_1 = the scale of the linearized technology
P_0 = the base scale
r = the scale factor
x = the installation factor.

Table 20. Costs in M$$_{2010}$ and additional scaling data for the individual components of the production train of ethanol

Components	Install. factor	Base Inv. costs	Scale factor	Base scale	Max. size	Units	Technologies 1GEN	SSF	SSCF
Pre-treatment									
Mechanical	2	5.94	0.76	83.3	83.3	Tonne-dry/h	x	x	x
Mill	1	0.50	0.7	50	50	Tonne-wet/h	x	x	
Dilute acid	2.36	18.87	0.78	83.3	83.3	Tonne-dry/h		x	
Steam explosion	2.36	1.89	0.78	83.3	83.3	Tonne-dry/h			x
Ion exchange	1.88	3.20	0.33	83.3	83.3	Tonne-dry/h		x	
Overliming	2.04	1.03	0.46	83.3	83.3	Tonne-dry/h		x	
Hydrolysis + fermentation									
Cellulase production (SSF)	2.03	1.71	0.8	50	50	kg/h cellulase	x	x	
Seed fermentors (SSF+SSCF)	2.2	0.35	0.6	3.53	3.35	tonne/h etoh	x	x	x
C5 fermentation (SSF)	1.88	0.90	0.8	1.04	1.04	tonne/h etoh		x	
Sugar fermentation (SC)	1.88	0.90	0.8	1.04	1.04	tonne/h etoh	x		
Hydrolyse – fermentation (SSF)	1.88	0.90	0.8	1.04	1.04	tonne/h etoh		x	
SSCF	1.88	0.90	0.8	1.04	1.04	tonne/h etoh			x
Upgrading									
Distillation and purification	2.75	3.96	0.7	18.47	18.47	tonne/h etoh	x	x	x
Molecular sieve	2.75	3.91	0.7	18.47	18.47	tonne/h etoh			
Residuals									
Solids separation	2.2	1.40	0.65	10.1	10.1	Tonne-dry/h	x	x	x
(An)aerobic digestion	1.95	2.06	0.6	43	43	tonne/h wst.water	x	x	x
Drier	1.86	10.68	0.8	33.5	33.5	Tonne-wet/h	x	x	x
Power island									
Boiler	2.2	36.26	0.73	173	-	MW steam raised	x	x	x
Gas turbine	1.86	22.61	0.7	26.3	-	MW-e	x	x	x
Steam system + turbine	1.86	7.17	0.7	10.3	-	MW-e	x	x	x

Source: Hamelink (2004)

This linearization data was taken from Hamelink (2004) for the technologies SSF and SSCF. The right side of the table also illustrates which technologies include which components, e.g. the SSF technology uses the dilute acid pre-treatment hydrolysis, with an ion exchange and overliming unit, while SSCF uses steam explosion as a pre-treatment option.

For SSF, we can see that the *hydrolysis and fermentation island* consists of 4 different components, and this is reduced to two for SSCF. A specific component is added for sugar fermentation in that section, based on the scaling data of *C5 fermentation (SSF)* or *hydrolysis-fermentation (SSF)* components, to process the *guarapo* or sugar-juice stream from the juice processing. This mass flow cannot be added to the hydrolysis-fermentation (SSF) unit, since the glucose presence limits the enzymatic hydrolysis of the cellulose. Adding a sucrose stream from the juice treatment to this reactor would impair the hydrolysis and will thus be fermented separately. For the SSCF option, this unit will also be separated from the SSCF units.

Since SSCF represents an increased integration of conversion components, this also means a reduction in electricity use (e.g. in the chipper and reactor vessels) and steam use (in the distillation). This will be explained below if the technology configurations are elaborated on in more detail. Furthermore, it is assumed that for all technologies the cellulase production is done on board and not bought elsewhere. Feedstocks in the SSF technologies are pre-treated by milling and dilute acid, while for SSCF this is done by steam explosion.

An energy-mass balance for an appropriate scaling of components is clarified by the following set of formulas:

The mechanical unit is based on the mass flow rate of the dry feedstock, that is sugarcane/energy cane without water, but including the sugars:

$$Ton_{dry,\ feed} = \frac{MW_{HHV,\ in} * 3.6\ GJ/MW}{HHV_{db,\ feed}\ (GJ/ton)}.$$

The rest of pre-treatment components (excluding mill) are based on the dry feedstock mass flow, where the guarapo (including sugars) is already separated. This is thus the dry weight of bagasse:

$$ton_{wet,feed} = \frac{Ton_{dry,\ feed}}{(1-mc_{feed,\ 60\%})}.$$

And,

$$Ton_{dry,\ bagasse} = ton_{wet,\ feed} * \%wt_{fiber\ content\ in\ wet\ feed}$$

The milling unit is also based on the wet feedstock weight. Units in the hydrolysis-fermentation and upgrading islands are scaled on the ethanol mass flow rate:

$$Ton_{etoh,sugars} = ton_{wet,\ feed} * TRS * \eta_{C6\ fermentation} * \eta_{C6\ chemical\ conversion}$$

$$Ton_{etoh,C5} = ton_{wet,\ feed} * \%wt_{C5} * \eta_{C5\ hydrolysis} * \eta_{C5\ fermentation} * \eta_{C5\ chemical\ conversion}$$

$$Ton_{etoh,C6} = ton_{wet,\ feed} * \%wt_{C6} * \eta_{C6\ hydrolysis} * \eta_{C6\ fermentation} * \eta_{C6\ chemical\ conversion}$$

Scaling the components in the residuals island, i.e. the waste stream for the boiler/turbine, is dependent on the content of the solids that remain in the sludge after C5 and C6 sugars are extracted. One way to estimate the residuals stream is to account for the specific hydrolyse/fermentation efficiencies of each specific sugar (mannose, arabinose, galactose, dextrose, and xylose). A simpler method is to make an overall mass balance, where $flow_{in}$ is the biomass in, and the $flow_{out}$ represents the Co_2, ethanol and residuals out:

$$Ton_{dry,\ feed\ in} = ton_{etoh,\ out} + ton_{Co2,\ out} + ton_{residuals,\ out} \quad (\%wt_{residuals} = ton_{residuals,out} / ton_{dry\ feed\ +\ co2\ +\ etoh}).$$

Calculating the mass flow rate of CO_2 - a fermentation by-product- is done with the molar mass formula:

From C6 sugars to ethanol: $C_6H_{12}O_6 \rightarrow 2\ C_2H_6O + 2\ CO_2$
180.16 g/mol = 92,14 g/mol + 88.02 g/mol, thus yielding 0.92 kg CO_2/kg_{etoh}

From C5 sugars to ethanol: $3\ C_5H_{10}O_5 \rightarrow 5\ C_2H_6O + 5\ CO_2$
450.41 g/mol = 230.35 g/mol + 220.06 g/mol, also yielding 0.92 kg CO_2/kg_{etoh}

Feedstock for the drier is:

$$Ton_{wet,residuals} = \frac{Ton_{dry,\ feed} * \%wt_{residuals}}{(1-mc_{residuals,63\%})}$$

The solids separator is based on the feed already dried:

$$Ton_{dry,\ residuals} = \frac{ton_{wet,\ residuals}}{(1-(mc_{residuals,63\%} - mc_{residuals,10\%}))}$$

This dry residual feedstock is delivered to the boiler and subsequently used to scale the power island.

Table 21 shows the most important parameters that represent both feedstock characteristics and characteristics of the technology configurations. For the technology part a global mass balance was applied to come to the capacities of the individual components. With global is meant that a general mass balance is given that include average xylose, glucose and rest-sugar saccharification and fermentation efficiencies, but exclude each individual fermentation efficiency for mannose, arabinose, galactose, dextrose, and xylose. Furthermore it excludes efficiency losses due to sugar consumption by the cellulase. The *range* of efficiencies of fermentation will, on the other hand, be incorporated. This is explained later in this paragraph when *trends* are discussed. The left-out details have a minor influence on the ultimate capital cost investments, and do not have a significant effect if compared with the revenues of the end products. The last indicator for the mass flow rate is the LHV_{wb} at which the solid waste stream is burnt. This of course indicates the energy content of the flow, and subsequently decides on the size of the boiler.

Table 21. Feedstock & technical parameters for excel model 2 (modified from Hamelink, 2004).

Feedstock parameters	**Units**				
		Sugarcane	Energy cane	Elephant $grass_{db}$	Bagasse
Fiber $content_{wb}$	% weight	14.0	26.7	78.6	87.0
..Content cellulose	% weight of fiber	41.6	43.3	49.80	41.6
..Content hemicellulose	% weight of fiber	25.1	23.8	42.20	25.1
..Content lignin	% weight of fiber	20.3	21.7	8.00	20.3
TRS-Total Recoverable sugars	% weight	14.0	10.0	-	-
Moisture content	% weight	60.0	54.4	0	0
HHV (dry basis)	GJ/ton	17.85	18.71	19.20	17.37
$LHV_{bgss,\ 10\%mc}$	GJ/ton				13.62 (a)
$LHV_{bgss,SSF,10\%mc}$	GJ/ton				17.76 (b)
$LHV_{bgss,SSCF,10\%mc}$	GJ/ton				16.83 (c)
		[1] [2]	[1] [2]	[4]	[2] [5]
Technical parameters					**References**
Moisture content from the field	%	50			[1]
Moisture content before for drier	%	63			[3]
Moisture content after drier	%	10			[3]
Turbine efficiency	%	55			
Generator efficiency	%	98			
90 bar boiler efficiency	%	90			
converse c6>oh	kg etoh/kg c6	0.51			
converse c5>oh	kg etoh/kg c5	0.61			
Density ethanol	(kg/l)	0.789			
			Technologies		

		1GEN fermentation	SSF	SSCF	
Fermentation efficiencies					
C5 saccharification efficiency	%		75-90	45-65	[3]
C5 fermenting efficiency	%		80-92	80-90	[3]
C6 saccharification efficiency	%		70-80	90-96	[3]
C6 fermenting efficiency	%	90-95	90-95	90-95	[3]
Electricity use	kWe/MW-hhv-in	_	_	_	[3]
..chipper		5	5	1	
..air compressor			31	31	
..reactor vessels		7	27	13	
..auxiliaries		29	29	29	
Steam use					[3]
..pre-treatment	ton/ton dry feed (4 bar)		0.20	0.20	
	(11 or 25 bar ssf/sscf)		0.30	0.10	
..drier	ton/twe (11 bar)	1.01	1.01	1.01	
..distillation	ton/ Kliter etoh (4 bar)	2.57	2.57	1.03	

Sources: [1] Alexander (1984), [2] Misook (2010), [3] Hamelink (2004). [4] Rego (2010). [5] sugartech (2013)

Comments:

(a) The residuals waste stream to the CHP has a HHV_{db} of 17 GJ/ton with given lignocellulose content (and HHV of 16-18 GJ/$ton_{HHV,db}$), cellulose content (also with HHV of16-18 GJ/$ton_{HHV,db}$) and lignin content (with HHV of 25 GJ/$ton_{HHV,db}$.

(b) For the SSF configurations, an average of 73% and 70% (lower bound) of the (ligno)cellulose is extracted for etoh conversion, resulting in a residual waste stream with higher lignin content, thus higher HHV/LHV of the mass flow to the CHP.

(c) For the SSCF configurations, an average of 38% and 83% (lower bound) of the (ligno)cellulose is extracted for etoh conversion, resulting in a residuals waste stream of 18.26 GJ/ton HHV,mc10%, and a 16.83 LHV,mc10%.

On scaling. The scaling sequence is as follows (and illustrated in Figure 25 below). For the 1st generation technologies, the only optimalization of the plant is for the power island. The boiler/turbine are scaled on the residuals flow rate of sugarcane or energy cane processing in the first 5000 hours of the year. for the remaining 3000 hours of the year, the power island is utilized for elephant grass combustion. For the 2nd generation technologies, the 2nd feedstock is scalled on the capacity of the C6 hydrolysis/fermentation unit of the first feedstock (Figure 25 right). If elephant gras is used as a 2nd feedstock, then the C5 fermentation and residuals flows are larger due to a higher fiber content of elephant grass. This is the case for configurations k4-7, were sugarcane/energy cane is the first feedstock. For these cases, the C5 fermentation and power units are scaled on the capacity of the 2nd feedstock. A third feedstock flow rate (also elephant grass) is delivered to be co-fired with cane bagasse in the first 5000 hours of the year to maximize the load factor of the increased turbine.

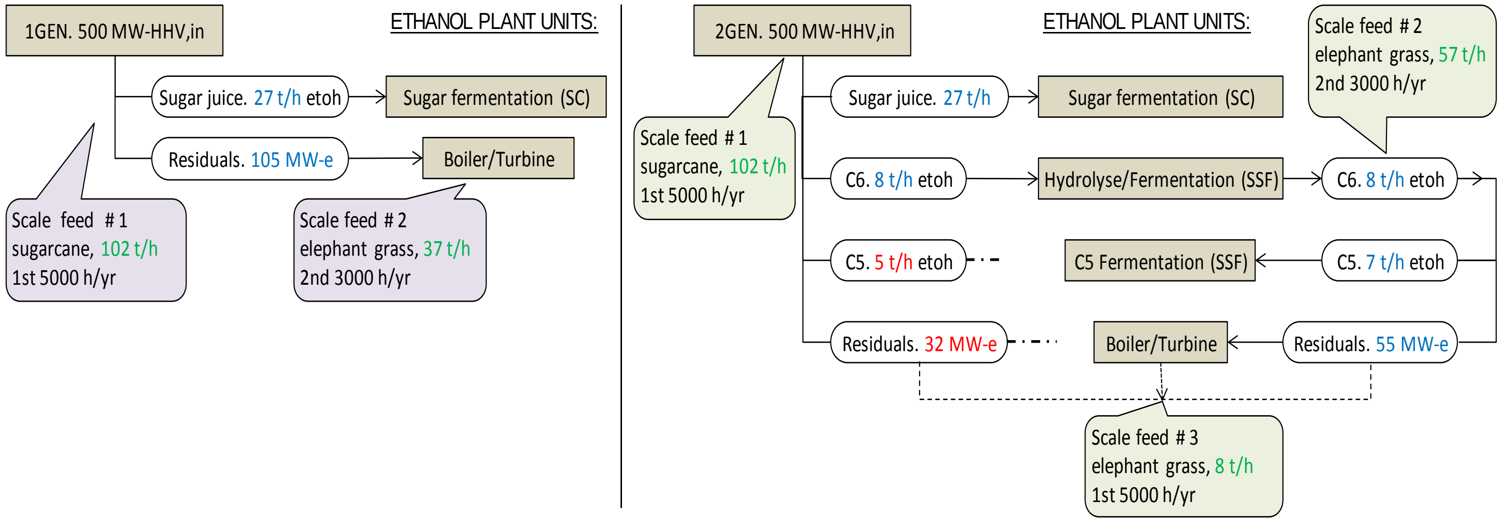


Figure 25. Method of scaling the (1) ethanol plant units; and (2) feedstock flows.

Figure 25 right illustrates this for a 500 $MW_{HHV,in}$ plant with sugarcane as its main feedstock. The hydrolysis/fermentation unit is scaled at 8 ton/hour ethanol. Sugarcane's C5 capacity is 5 tons/hour, while for elephant grass this is 7 tons/hour. Residuals from elephant grass requires a 55 MW_e turbine, while sugarcane only requires a 32 MW_e turbine. A 55 MW_e capacity will then be installed. A third feedstock of 8 tons/hour elephant grass will then supply this 55-32=23 MW_e in the first 5000 hours of the year during cane harvesting.

In this analysis elephant grass is used both for the 2nd and 3rd feedstock flow –in 1GEN configurations. For the CHP optimization, also other feedstocks can be used, such as forest/agricultural residues. These CHP units then function as regular electricity plants outside of the cane milling season.

Table 22 below summarizes the most important production parameters of the technology: the industrial yields of ethanol in liters per ton $biomass_{wet}$, and excess electricity production per ton $biomass_{wet}$. Remember that ethanol in some cases, and electricity in all cases, is produced by two different feedstocks.

Table 22. Nine technology configurations with their feedstock and optimization characteristics.

Technology	Ethanol yields			Electricity co-production		ratio feedstocks 2nd feed : 1st feed
	$L_{etoh}/ton_{wet,feed}$			$kWh/ton_{wet,feed}$		
	Sugar cane (a)	energy cane (a)	elephant grass (b)	1st feedstock	2nd feedstock optimization	wt%
1GEN Ferm (e) 2	90			351	211	8%
1GEN Ferm 3		59		613	368	18%
1-2GEN SSF (f) 4	133		247	19	379	12%
1-2GEN SSCF (c) 5	127		202	56	379	13%
1-2GEN SSF (d) 6		140	233	117	122	25%
1-2GEN SSCF 7		134	215	184	867	24%
2GEN SSF (g) 8			247	302		100% elg
2GEN SSCF (h) 9			215	571		100% elg

(a). Load factor 5000 h/yr. If sugarcane is the base feedstock, energy cane is used to expand the LOMS.

(b). Load factor 3000 h/yr.

(c). 92% of the elephant grass is first used for etoh production, optimizing the C6 unit. 8% is directly co-fired to optimize the power island.

(d). Id for k6; 91% fermentation first; 9% direct fo-firing.

(e). Seabra (2012) reports a yield of 90.12 L/ton wet sugarcane in standard Brazilian industries, The distillery gobal efficiency evolution outlook reports 85 L/tonwet,feed (Seabra, 2011). Dias (2010) reports (configuration ll) 89,3 L/ton cane, and 92.6 kWh/ton cane, but the bagasse is co-fired at 50% mc, thus 7.57 GJ/ton LHV-wb. At these bagasse parameters, k2 produces 65 kWh/ton wet cane.

(f). Dias (2011) reports a yield of 131.5 L/ton wet sugarcane, applying steam explosion as pretreatment, 70% and 80% pentose hydrolysis-fermentation efficiency, resp. Electricity production for his configuration is 72 kWh/ton cane, but includes 50% trash-from-field burning.

(g). The overall energetic efficiency (GJ based) in Hamelink (2004), for SSF is 35% (etoh) and 5% (electricity), while in this thesis, k8 has efficiencies of 34% and 8%, resp. if average fermentation efficiencies are considered. The 3% electricity default can be assigned to the simplification of steam extraction calculations in this research, where different pressures are treated equally.

(h). The overall energetic efficiency of Hamelink's SSCF configuration is 39.5% (etoh) and 10% (electricity). Configuration k9 has efficiencies of 38% and 9%, resp. This relative precision means that the 9 technology configurations are built succesfully.

K2. Configuration k1 is the status quo fermentation technology, treating sugarcane for juice fermentation and burning the bagasse in a Combined Heat and Power (CHP) island with a load factor similar to the Length of Milling Season (LOMS), i.e. between 4000-5000 hours of the year. The load factor of the CHP is optimized to 8000 h/year by providing elephant grass the remaining 3000 hours. The benefits of this optimization result in an increased electricity output to the grid.

K3. This option is included in particular to emphasize electricity production. The CHP is again optimized up to 8000 hours/year, as are the rest of the technology configurations. Energy cane is still handled first to produce ethanol, but it has a roughly 2 times higher excess electricity production than k2. Nevertheless, it has a lower ethanol output, and a Total Capital Investment roughly 20% higher). This is because energy cane has a substantial higher fiber content than sugarcane, and thus a greater power island. Due to Puerto Rico's high electricity prices (up to $\$0.27_{2013}$ per kwh) these first three conversion trains can play an important role in power production next to ethanol production.

K4. This configuration has the lowest excess electricity production of all configurations. Due to a low fiber content, which is now also utilized for (ligno)cellulose ethanol conversion, only 19 kWh/ton cane is produced.

K5-k9. The rest of the configurations are scaled and planned as discussed above with k4, but with diversifying with sugarcane, energy cane and elephant grass as first feedstock. The ratio of the 2^{nd} feedstock per 1^{st} feedstock 2^{st} : 1^{nd} ranges between 8-24% on a yearly basis depending on the configurations. The SSCF option is in all cases cheaper due to (see Table 22 above) less production components: less fermentation reactors, no mill, ion exchange or overliming components needed. The difference in ethanol output between SSF and SSCF is due to lower hydrolysis efficiencies of the latter. This difference increases for configurations with higher fiber feedstocks. This is because the efficiency of sugar fermentation for both technologies is 91%. Thus, for configurations with increasing fiber contents in their feedstock (sugarcane, then energy cane, then elephant grass) the difference in ethanol output between SSF-SSCF are increasingly due to the efficiency differences. This is illustrated in Table 23, with ethanol, electricity and overall conversion efficiencies of the 9 configurations. Difference in ethanol output between SSF and SSCF increases with increased fiber feedstocks: 2% difference between $k4\text{-}k5_{sugarcane}$, 4% between $k6\text{-}k7_{energycane}$, 6% between $k8\text{-}k9_{elephantgrass}$.
roughly equal ethanol output (higher fermentation but lower hydrolyses efficiencies)

	Sugarcane, 1GEN		Sugarcane, SSF-SSCF		Energycane, SSF,SSCF		Eleph. grass, SSF, SSCF	
	k2	k3	k4	k5	k6	k7	k8	k9
Ethanol efficiency	22%	12%	37%	35%	36%	32%	41%	35%
Electricity efficiency	37%	41%	16%	15%	15%	19%	8%	14%
Total	59%	53%	52%	50%	51%	52%	49%	49%

η_{etoh} (%) = E_{etoh} (GJ) / $E_{biomass,in}$ (GJ) η_{power} (%) = E_{power} (GJ) / $E_{biomass,in}$ (GJ) η_{total} (%) = $\eta_{etoh} + \eta_{power}$

Table 23. Conversion efficiencies of the 9 plant configurations: ethanol-, electrical-, and overall conversion efficiency.

The electricity conversion efficiency increases with higher fiber feedstocks: due to lower ethanol conversion efficiencies, a higher residual flow rate remains for the CHP. This increase in electric output is largest for k8-k9 (6% increase). From an ethanol point of view, a high fiber feedstock is preferred with SSF-SSCF technologies. From an electricity point of view, a 1GEN technology is preferred, but an optimization of the CHP is essential!

The SSCF technology will become available only in the next 3-6 years but is still applied in the optimization procedure of *fu5ture* ethanol supply (timespan 2017-2030). For SSCF the upgrading/residual and power islands is comparatively bigger than that of SSF due to a bigger residuals supply (less (ligno)cellulose utilization due to lower efficiencies in hydrolyse and C5 fermentation), and thus a higher electricity output by this bigger boiler/turbine. Moreover, the electricity output is even larger due to more efficient electricity and steam use in the production processes of SSCF configurations.

On trends. Trends for the future concern mainly the developments in hydrolyze-fermentation efficiencies of the lignocellulose, cellulose en sugars in the feedstocks. These efficiency ranges are shown again below (from Table 21 on page 84). And were integrated as follows. It was assumed that these efficiencies have matured, i.e. reached their upper bound in 15 years. Hamelink (2004) gives an indication in which time period these technologies will become available, that is in 5 years (SSF) and 10-15 years (SSCF) from the moments of his paper (2004), thus 2009 for SSF and in 2017 for SSCF (averaging 10-15 years from 2003 on).

SSF	SSCF	
%	%	
75-90	45-65	C5 hydrolysis
80-92	80-90	C5 fermentation
70-80	90-96	C6 hydrolysis
90-95	90-95	C6 fermentation

Over that 'maturity' time period of 15 years the efficiencies will increase linearly until the upper bound is reached. For every time period (of 3 years) the new efficiencies are translated to the excel spreadsheets about technology scaling, wereafter the parameters from these excel sheets are refreshed. These parameters are included into MOMILP as to optimize until 2028 for a constant ethanol demand per time period. The demand in ethanol is set *constant* because this will result in the installation of whatever technology in the first time period.

It was not possible to incorporate the efficiency increases described above in one iteration in MOMILP, i.e. if technologies would be installed in time period 3 due to an *increased* ethanol demand, they would still have the efficiencies of the first time period. The technology parameters in MOMILP are all constant, i.e. equal to those of the first year. Nevertheless, all other trend-sensitive parameters were corrected accordingly.

4.2.b. Technology Parameters

This paragraph summarizes the remaining technological parameters for each configuration: Total Capital Investments or $TCI_{k,p}$ for every technology k and capacity p; production costs '$c_{'slope'}$' including industrial inputs; excess electricity production to the grid w_k; nominal ratios $b_{k,i}$ or the ethanol share from each feedstock type; and the ethanol conversion factor from every feedstock $y_{k,i}$ in $ton_{ethanol}$ / $ton_{biomass}$.

Table 24. $TCI_{k,p}$ or Total Capital Investments for the technology configurations (M$2014 for SSF and M$2016 for SSCF)

TCI M$/ MW-in Technology k	Capacity (MW-HHV,in								
	200	300	400	500	600	700	800	900	1000
2	143	198	251	303	354	403	450	498	546
3	150	207	261	316	368	418	467	517	567
4	139	194	249	303	353	403	455	505	555
5	122	171	220	268	315	358	403	449	493
6	172	240	304	369	429	489	554	617	677
7	148	206	262	319	371	423	473	526	577
8	177	246	311	384	452	516	578	643	711
9	149	208	264	321	375	427	480	532	585

Table 24 above summarizes the total capital investments TCI for the technology configurations and for the different capacities. Note that all investments are corrected to 2014 dollars, the starting date of the optimization, except for for the SSCF configurations which are in 2016 dollars. The SSCF technologies are available from that date on.

Also remember that these figures represent the costs of the technology configurations at their lower bounds of hydrolyse-fermentation efficiencies. Into the future when these efficiencies are increasing, the TCI will decrease somewhat. While the fermentation island is scalled somewhat larger due to higher ethanol outputs, the residuals and power island is subsequently scaled smaller because a larger fraction of the biomass is utilized for ethanol production. Since the residual and power island constitute the largest share in TCI, the costs decrease somewhat with increasing efficiencies.

Production costs of ethanol represent the variable costs associated with ethanol production in dollars per ton ethanol produced. These variable costs consist of fixed variable costs such as maintenance, labor and insurance, and consumables such as dilute acid H_2SO_4, lime, ammonia and dolomite. Table 25 below summarizes the costs for the inputs necessary for ethanol production. Table 26 below summarizes the production costs $c_{'slope'}$ in $\$_{2014}$ for all technology configurations.

Table 25. Cost components in the calculation of c’slope’
or Production costs in etoh production M$2014 (Hamelink, 2004).

Parameter Etoh production	**Unit**	**Quantity (M$2014)**
Fixed Variable		
..Maintenance	% of TCI	0.0404
..Labour	% of TCI	0.0067
..Insurance	% of TCI	0.0013
Consumables		
..Dilute acid	M€/ton-dry	1.10
..Lime	M€/ton-dry	1.17
..Ammonia	M€/L etoh	0.0200
..CSL (corn steep liquor)	M€/L etoh	0.0231
..Dolomite	M€/ton-dry	20.18

Table 26. C'slope' or production costs
of ethanol ($2014/ton ethanol)

Technolog k	$/ton etoh
2	215.21
3	243.28
4	477.57
5	132.15
6	156.98
7	225.85
8	309.37
9	185.82
2	206.03

The rest of the parameters w_k for excess electricity production, $b_{k,i}$ as nominal ratios of the contribution of feedstocks to ethanol production, and $y_{k,i}$ as the conversion factors from biomass to ethanol were already presented in paragraph 4.2.a ‘Scaling Nine Technology Configurations’.

4.2.c.GHG Emissions and Emission Credits from Ethanol Production

This section summarizes both Green House Gasses (GHG) emissions that are associated with ethanol production, and emission credits received from the excess electricity produced as a byproduct. Greenhouse gas emissions associated with the production of ethanol includes the actual CO2 as a byproduct in the fermentation process of sugars to ethanol, as well as the inputs that are necessary for the industrial process. Concerning the fermentation process, 0.92 and 0,92 kg of CO2 are co-produced with 1 kg of ethanol from resp. C5 and C6 sugar fermentation. Emissions for the consumables are the emissions associated with the industrial production of these consumables, i.e. an emission factor in kg CO_2 eq/kg of consumable is applied. The consumables in the ethanol production process were already presented in Table 25 on the previous page, and subsequent emission factors are summarized in Table 28 below and are for dilute acid (H_2SO_4), lime, ammonia and dolomite. Table 27 below summarizes the emissions per ton of ethanol produces for the 9 technology configurations.

Table 28. Emission Factor for Fuel Production per technology k (in kg CO2 eq/$ton_{ethanol}$) or ffp $_k$.

Parameter	**emission factor** kg Co2 eq/kg
Consumables	
..Dilute acid H2SO4	0.053
..Lime	0.013
..Ammonia	2.059
..Dolomite	0.305

Source: all GWP100a emission factors, from ipcc, 2007.

Table 27. Emissions for fuel production or ffp(k).

Technology	kg CO2eq/t-etoh
2	71.00
3	796.07
4	352.58
5	363.34
6	430.14
7	451.46
8	471.83
9	527.54

Emission credits are also applicable in this analysis, since the excess electricity that is produced as a byproduct of the ethanol production process replaces electricity that is produced by fossil fuels. There are different allocation methods to be found in literature, but the following method will be applied here.

This thesis investigates the financial and emissions performance of a supply chain of ethanol. The end result is a total cost and emission figure for the production of a liter ethanol from 'seed-to-tank', i.e. including all life cycle stages from planting the biomass up to delivery of the ethanol at fuel stations. Therefore the emissions avoided due the by-product electricity will be allocated completely to this ethanol production.

Calculationg the emission credits that are received from excess electricity, an emission factor for the island's fossil based electricity (that is being replaced) is needed. Table 29 below includes all the fossil fuel sources of Puerto Rico, the island's total electricity production, and those fuels specifically used for electricity production are highlighted in green. Table 30 below presents the energy densities and emission factors for those fuels used for electricity production.

Table 29, Fossil fuel use in the electricity mix of Puerto Rico, electricity production from these sources, and the emission factor for PR electricity.

Fuel sources	Units	Quantity [1]	Eimissions in metric tons
Total fossil fuel			2.90E+07
Total petroleum	bbl/day	1.76E+05	2.40E+07
motor gasoline	bbl/day	4.80E+04	6.34E+06
jet fuel	bbl/day	1.00E+04	2.32E+06
kerosene	bbl/day	1.00E+03	2.32E+05
Distillate no. 2	bbl/day	2.80E+04	3.23E+06
Residual no. 6	bbl/day	6.00E+04	7.97E+06
Liq.Petr.Gas	bbl/day	4.00E+03	2.59E+04
Other Petroleum	bbl/day	2.60E+04	3.99E+06
Natural gas	cu ft	2.60E+10	1.44E+06
Bituminous coal	short tons	1.65E+06	4.00E+06
kWh Net production	kWh	2.10E+10	
kWh from petr/gas/coal	kWh	2.08E+10	
GHG from petr/gas/coal	metric tons	2.04E+07	
emission factor PR	kgCO2/kWh	**0.80**	

[1] Energy Information Administration (2010).

The total emissions from those fossil fuel sources is then divided by the sum of electricity produced by those sources, and consequently the emission factor is derived of 0.80 kgCO2eq/kWh of electricity produced. The last

below then summarizes the emissions avoided by the electricity that is produced per ton of ethanol, for all 9 technology configurations.

Table 30. Fossil fuel use in the electricity mix of Puerto Rico,electricity production from these sources, and the emission factor for PR electricity.

fuels	emission factor kgCO2eq/GJ	energy density MJ/L	emission factor kgCO2eq...	
Distillate no. 2	73.53	36.00	2.65	/L
Residual no. 6	76.83	39.71	3.05	/L
Natural gas (MJ/m3)	56.00	35.00	1.96	/m3
Bituminous coal	95.00	27.00	2.67	/kg

Source: Blok (2007)

Table 31. Emission credits from byproducts of the ethanol production, or fec(k)

Technology	kg CO2eq/t-etoh
2	6273
3	16820
4	585
5	765
6	857
7	2154
8	1154
9	2593

As was expected, configuration k3 has the highest avoidance factor of 16.8 ton of CO2eq. avoided per ton of ethanol produced. The emission factor of 0.80 kg CO2 eq./kWh is a very high number, similar to societies with coal as a dominant source for electricity production such as China (0.95 kg CO2 eq./kg), several eastern European countries (0.94 for Czech republic to 1.2 for Poland), higher than most of the direct neighbors such as Cuba (0.94), Netherlands Antilles (0.74) or Trinidad (0.77) and nearly twice as high as the US (0.54 kgCO2eq/kWh) (Brander et al, 2012).

This is explained by the high share of Distillate and Residual fuel oils in the electricity production mix of the island. It must be mentioned that the island has been preparing policies to initiate a fuel switch in its energy resource portfolio for the next 20 years.

Figure 26 below illustrates this switch, where 'long term' was intended to be in 2020, but is now pushed forward. Whether this switch is accomplished in 2025, 2030 or later, it still does not have a significant impact on the emission factor of electricity. The emissions of residual and distillate fuel oil avoided are offset by the increase in gas and bituminous coal in the electricity mix.

Table 30 shows that the emission factor for oil is between 73-76 kgCo_2eq/GJ, and the average gas-coal factor around 75 kgCo_2eq/GJ. Figure 26 shows a 50% decrease of oil consumption on the long run but a doubling of the gas-coal share (time frame 2009-'long-run'). The emission factor would drop by 0.03 kgCo_2eq./kWh from 0.8 to 0.77 kgCo_2eq./kWh, but this is due to the increased share of 15% renewable energy sources in the mix. Therefore an emission factor of 0.80 kgCO2eq/kWh will be used in the time frame of the PLUC/MOMILP models.

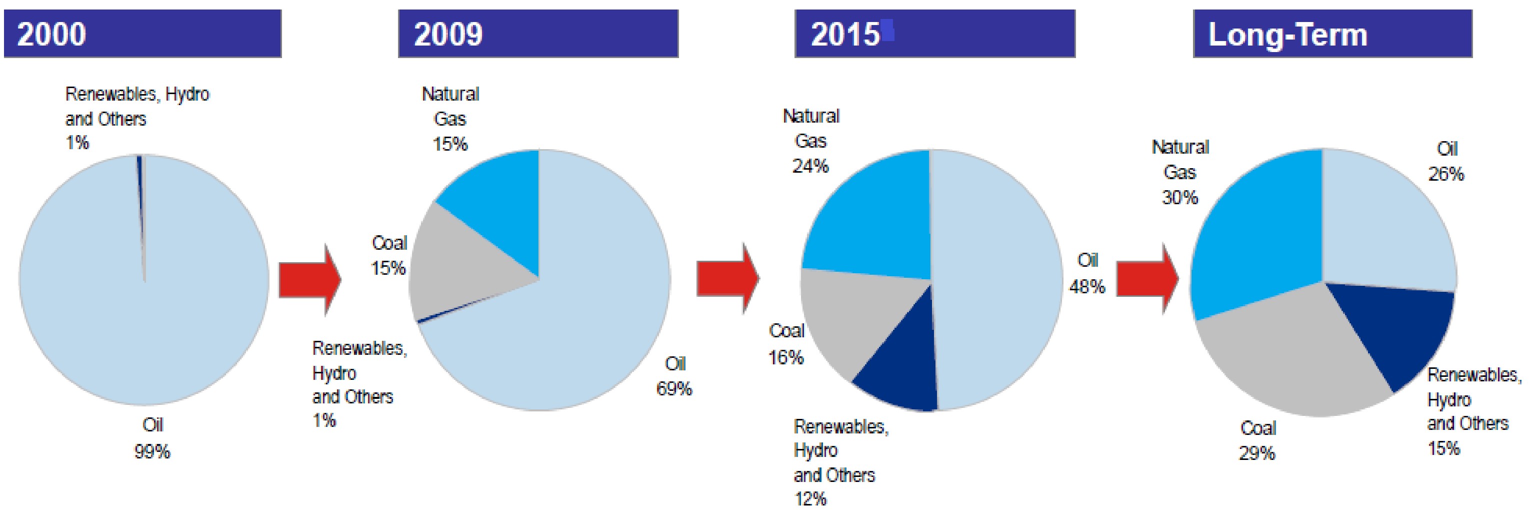


Figure 26. Fuel difersification plan for the timespan 2000-2020. The long term was initially intended to be 2020.

4.3. Life cycle stage 'Transportation'

4.3.a. Transportation of Biomass & Ethanol. Associated GHG Emissions

Transportation represents the two following life cycles and the parameters associated with 1) biomass from farm to plant, and 2) fuel to demand centers/mixing stations. These parameters are TCb_l or Transport Costs for biomass, TCf_l or Unit Transport Costs of ethanol through transport mode l, fbt_l and ffd_l or emission factors for biomass or fuel through transport mode l, resp., $Tcap_i$ and $Tcap_i$ on capacities of transporting biomass and fuel, resp.

Transport Costs for biomass in ($/ton.km) is derived from Conty (1997) and equals 0.50 $/ton.km for transport with a truck-trailer carrying 30 tons per trip, with an average delivery distance to the mill of 16 km charging 8.33 $\$_{2010}$ per trip (Conty, 1997).

$Tcap_i$ or the carrying capacity for biomass transport is 28 tons for in this research. Due to an absent sugarcane industry these parameters will change if such a supply chain is really implemented. For now these parameters will suffice.

Fbt_l or the emission factor for biomass transport is calculated with data from Macedo et all (2008) and is based on machinery in the Brazilian sugarcane industry. Table 32 below summarizes the result.

Table 32. . Emission factor for biomass transport per transport mode L (kg CO2eq/ton-b.km), or fbt(l).

Trucks energetic efficiency	t.km/L	52,40	[1]
	L/t.km	0,019	
GHG emission factor diesel	kg CO2eq/L	2,680	[2]
fbt(truck)	**kgCO2eq/t.km**	**0,051**	

[1] Macedo et all. (2008)
[2] http://www.epa.gov

Fuel distribution in Puerto Rico is done with trucks of 10,000 and 12,500 gallons (or 28.0 and 35.0 tons gasoline per trip (with 3,7854 L/gallon volume and 0.74 kg/L gasoline density). With a density of 0.789 kg/L ethanol the largest truck will carry 37.33 tons of ethanol per trip. Thus, $Tcap_{etoh}$= 37.33 tons. The parameters associated with transport are once again summarized in Table 33 below.

Table 33. Costs and GHG emissions for biomass and ethanol transport.

		TCb_l	TCf_l	Fbt_l	Ffd_l
L	Unit	$/t.km	$/ton.km	kgCo2eq/t.km	kgCo2eq./t.km
truck		0.50	0.5	0.051	0.123

5. Results

One peculiarity of combining a bottom-up techno-economic analysis of best suited technologies and according feedstocks, emissions and other parameters, with a top-down analysis that must result in the optimal choices from a wide variety of variables to satisfy a certain ethanol demand, is that one must find a balance between what to make static and what dynamic as to give a clear comparison between options.

One of the advantages of the MOMILP was that it would include so much variables to choose from, but at the same time this approach will impair a good technical comparison between technologies and their performance. For example the MOMILP can allocate an ethanol plant quite a distance outside an area where the land availability for bioenergy crop production is very large, just because it has to supply ethanol to other demand stations as well, and therefore these transport distances are included in the optimization. This was precisely one of the attractive sides of MOMILP, but so much seems to depend on the scenario that one constructs, i.e. where are the demand centres, where does one impair transport to flow, that results will differ if other scenarios are used or the same ones changed.

This chapter elaborates on the most important results from the analysis of a future supply chain of ethanol. In the first paragraph will be discussed the Costs of Energy$_{etoh}$ (COE_{etoh}) for all eight technology configurations prepared for this suppy chain for the time frame 2014-2030. The COE is broken down in their specific cost components that include all life cycle stages from biomass production, to biomass transport, ethanol production, ethanol transport and credits from electricity supplied to the grid.

The next set of results will elaborate on the developments of the COE-thanol into the future. Because multiple factors are of influence on the performance of the supply chain, such as biomass yield improvements, price increases and improvements in ethanol conversion efficiencies, it is interesting to follow how these factors influence the $COE_{ethanol}$ into the future as to provide a better understanding for the invest-now-or-postpone decision making of an ethanol supply chain.

A sensitivity analysis of the electricity price on economic performance will be given also. The second paragraph will illustrate for the best performing supply chain configuration (composed of two technologies) an overview of the geographical planning of biomass plots, ethanol plant locations and accompanying capacity scaling over the time frame 2014-2030. Along this overview will be given their subsequent GHG emissions versus the COE in a 'Paretto'-graph.

5.1. Analysis of supply chain configurations

The conditions for a comparative analysis of the technologies and according supply chain characteristics is as follows.This ethanol production demands a land availability of around 20,000-75,000 hectare for energy crops over the timeframe and depending on the supply chain configuration, i.e which conversion technology is chosen with according feedstock. The supply chain is not optimized for the full potential of 82,978 hectare - i.e. all available land is cultivated- in order to keep the allocation of plants and plots open. In other words, the playing field for the allocation of plants/plots must be kept open enough as to analyse changes in allocation due to a range in transportation costs that is discussed in the sensitivity analysis at the end.

Figure 27 below presents the Cost Of Energy$_{etoh}$ for the timeframe 2014-2030 (in \$/GJ), its breakdown in Total Capital Investments (TCI), Taxation (only when the net cash flows are positive), Ethanol Production Costs (EPC), Biomass Production Costs (BPC), transport costs for biomass (TCb) and –fuel (TCf) and the benefits of excess electricity ('Incomes power'). The electricity price is set at 0.12 \$/kWh sold to the grid, and the biomass transport costs at 250 \$/short-trip (average of 20 km^2 distance, 30 tons/haul, and thus 0.42 \$/ton.km).

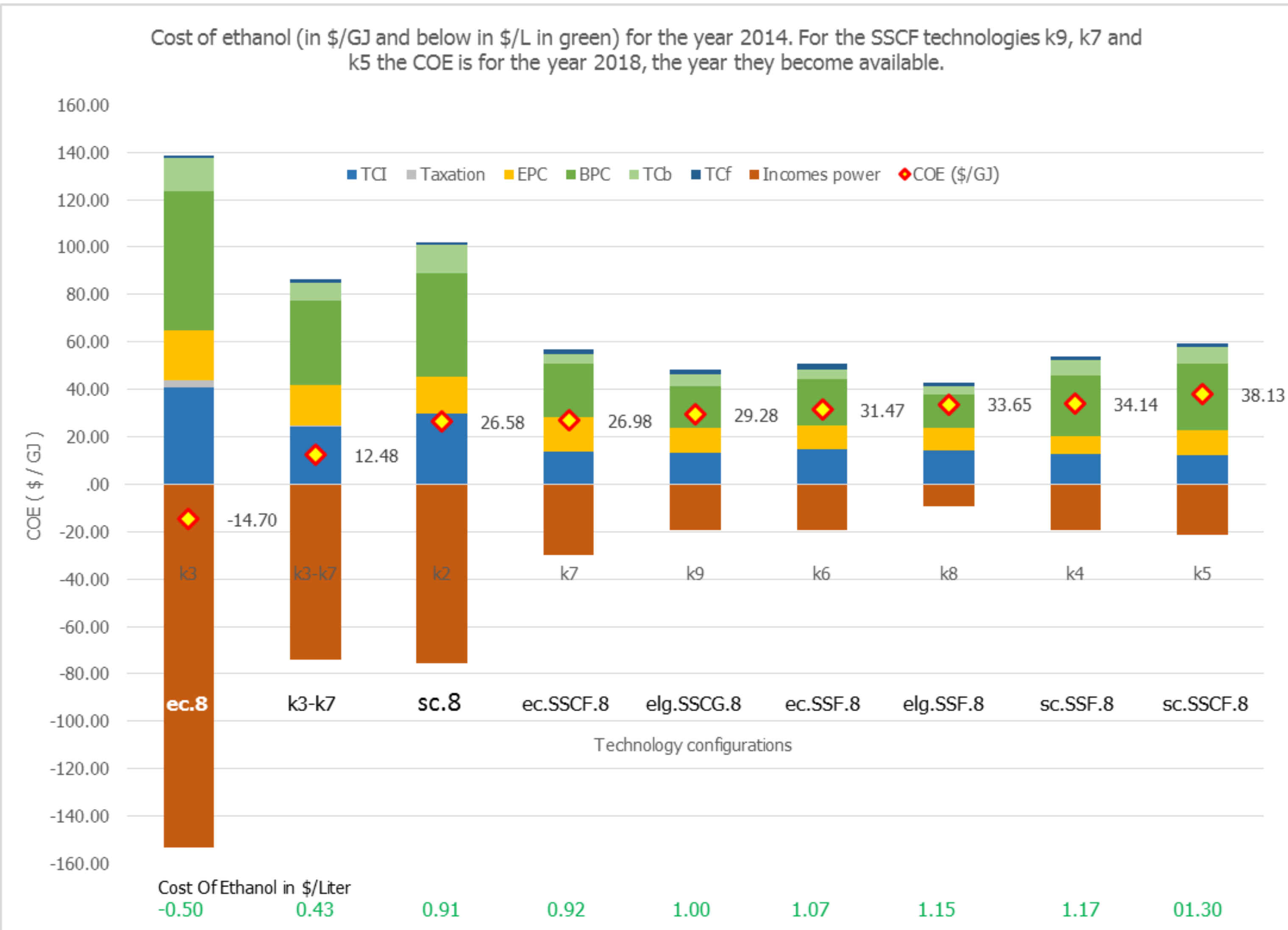


Figure 27. Cost of ethanol (in \$/GJ and below in \$/L in green) for the year 2014. For the SSCF technologies k9, k7 and k5 the COE is for the year 2017, the year they become available.

The best option is a supply chain based on energy cane with conversion technology k3=ec.8 or the normal juice extraction and 1st generation fermentation, where the power island is further optimized from 4000 to 8000 hours per year of operation supplied by elephant grass. This option has the highest total cost factors, namely 140 $/GJ, specifically in energy cane production and capital investments. Nevertheless, due to a high electricity production and revenues, the net COE is -14.70 $/GJ meaning not a cost but a benefit of 14.70 $/GJ for this conversion route. On an island scale and given timeframe this means a NPV of 1,017.4 million dollars$_{2013}$.

The third best option is an equal technology configuration but now with sugarcane as base feedstock, or configuration k2=sc.8. A high electricity revenue is also the main factor for its favorable performance. In paragraph 0 'Strategic Policy Planning' a supply chain configuration with both k3 and k7 technology configurations will be advised on: the former for its electricity revenue advantages, the latter for its favourable ethanol conversion. For this reason configuration k3-k7 is already included in Figure 27 above to compare for COEs. Its COE is second best and amounts to 12.48 $/GJ.

The influence of electricity credits plays a significant role in the economics of the supply chain. Due to the high cost –and marketing- of electricity on the island (wind and solar power is purchased by PREPA at 12-15 $ct/kWh), supply chain configurations with a 1GEN technology and a high additional power production have a very large power revenue.

The two most favorable options k3/k2 are followed by two 2nd generation technology configurations of the SSCF type, namely k7=ec.SSCF.8 based on energy cane, and k9=elg.SSCF.8 based on elephant grass.The 2nd generation configurations based on sugarcane are performing worst of all, mostly due to the low fiber content of the feedstock which is now also utilized for 2nd generation conversion thus leaving a significantly lower residual waste stream for the CHP. SSCF options are also better performing than their counterpart SSF with the same feedstock because their ethanol conversion efficiencies are lower, subsequently have greater residual waste streams for the CHP and thus higher electricity revenues. The total capital investments of the SSCF options is also lower due to component integration, and this outweighs the extra fermentation capacity that is needed at an equal ethanol demand. Later it is tried to find the turning point in SSF to SSCF on an increasing kWh price.

The intermediate conclusion here is that a 1GEN technology with either energy cane or sugarcane (with power island optimization using a 2nd feedstock), or a 2GEN technology-SSCF type with a high fiber feedstock (elephant grass or energy cane) shows the best economic performance of an ethanol supply chain.

The price of electricity that is sold to the grid plays a significant role in the technology choice. Preceding analysis of the supply chain configuration with according best-choice technologies were based on an electricity price of 0.12 $/kWh.

Figure 28 below illustrates the effect of the electricity price on the economic performance of the ethanol supply chain configurations, i.e. kWh prices versus COE. If the electricity is delivered at PREPA for less than 7 cents per kWh, a supply chain based on energy cane and technology k3 (dark blue line) will be the most expensive among all options[13]. This changes progressively towards a price of 8 cents per kWh or more, then

[13] Curves for the blue lines are cut-off at 55 $/GJ to illustrate the sensitivity for the other configurations, but they reach 70-75 $/GJ, resp. at an electricity price of 50$/MWh.

this option becomes the cheapest configuration of all. At a 11 ct/kWh price or higher, option k3=ec.8 becomes a business case with positive NPV, as is shown in previous Figure 27.

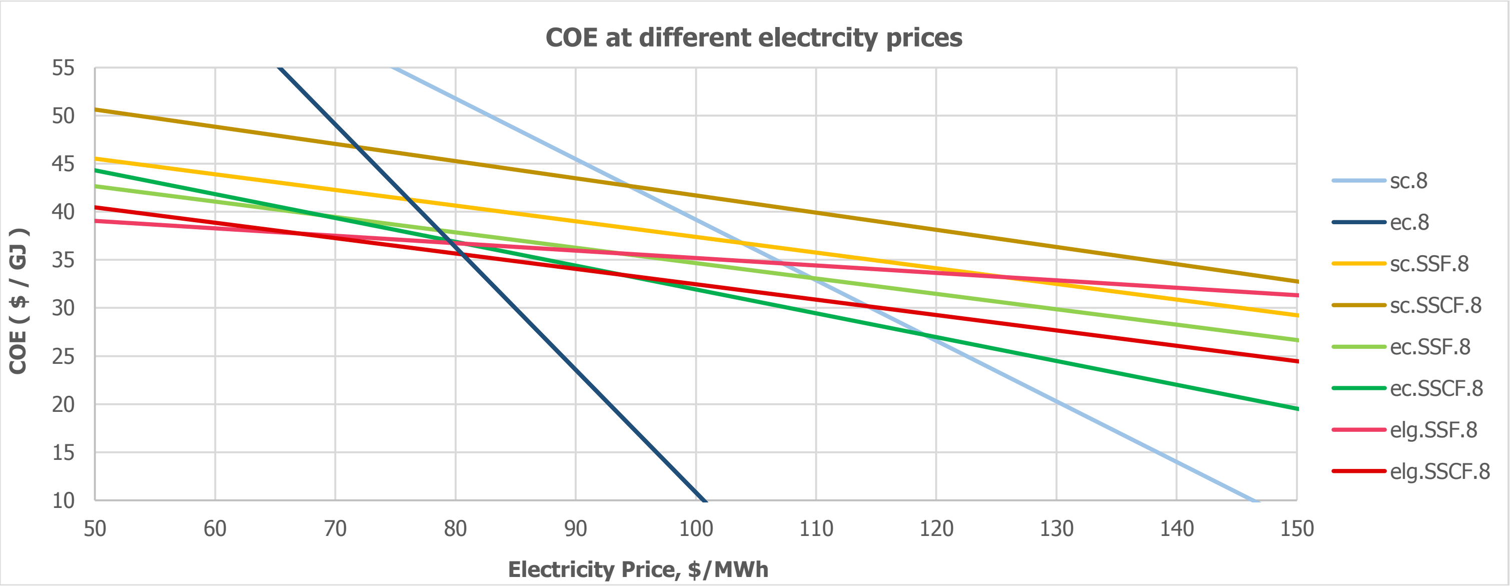


Figure 28. COE in $/GJ at different electrcity prices

Technology k2=sc.8 is characterized by the same influence that the electricity price has on its performance, albeit in less severe form (light blue line). Using sugarcane and 1st generation fermentation (k2) will become second cheapest only after an electricity price of 12 ct/kWh. The rest of the technologies do of course perform better at increased electricity prices, but this effect is not so dominant due to a fairly smaller electricity co-production. For the range in electricity price 9-12 ct/kWh, a SSCF ethanol conversion route with feedstock energy cane (dark green line) is second cheapest of all options.

Precedingly is discussed the economic performance of the supply chain configurations for the timeframe 2014-2030 at a linear increase in ethanol production of 20-40% of the demand, at an electricity price of 12 cents per kWh, and at 42 cents per ton.km for biomass transport. Trend developments in costs and biomass yields were also incorporated (see textbox 1 below). Efficiencies in ethanol conversion (saccharification and fermentation) were kept constant, i.e. the necessary extra capacity that must be added to satisfy the *increasing* 20-40% ethanol demand is all at a constant conversion efficiency at its corresponding time period. Preceding supply chain comparison in Figure 28 were thus with ethanol conversion efficiencies at the lower bound (see paragraph 4.2.a 'Scaling nine technology configurations').

To also incorporate these *conversion efficiencies* in the analysis of supply chain preference, it is chosen to keep the ethanol demand at 30% of the total gasoline demand of the ten demand centres[14]. The model is run with constant fermentation efficiencies over a 15-year period (the length of a MOMILP run) starting in 2014. Next, a new 15-year run is made over the time period 2017-2032, shifting the time frame 3 years into the future. For every new 15-year optimization run the ethanol conversion efficiencies are linearly increased, and all the time dependent and technological parameters renewed. In such way five 15-year runs are made

[14] For the previous comparison of configurations, an increasing ethanol demand of 20, 25, 30, 35, and 40% for the timeperiods was dictated. In the calculation of COE, also the E_t or energy production$_{etoh}$ is discounted (see paragraph 2.1 'Main structure' for the calculation method of COE). The discounted amount of produced ethanol for the 20-40% output is 79,023,000 GJ. A constant output of 30% has an equal discounted ethanol output of 79,023,000 GJ.

until the last timeframe 2026-2041. For all the respective years the appropriate cost and biomass yield corrections are applied (see textbox 1 above).

BOX 1 : Reminder. How are trends into the future incorporated in the MOMILP model?

Remember the range of future trends that were incorporated into MOMILP to take into account for price, yield and efficiency increases? These trends were summarized in parameters for the optimization: $trBY_{i,t}$ is the biomass yield trend for the future, for every biomass type i and time period t; $trGDP_t$ to correct prices with a GDP deflator which was used for all payments in capital investments, production costs of ethanol and transport costs; $trPCb_{i,t}$ or biomass production cost developments including fertilizer-electricity-land price indexes and increases in biomass yields; $trMP_{j,t}$ or price increases of the products ethanol and electricity according to a diesel and electricity price index; $trfbg_{i,t}$ or decreases in the emission factor for biomass growth due to yield improvements over the same agricultural input.

Fertilizer use is assumed constant while yields improve due to better land and crop management. Efficiencies with the conversion technologies also increase, but are difficult to incorporate in one 15-year run with MOMILP. To circumvent this it is chosen to set the demand for ethanol constant, so the capital investments will be done in time period one. The MOMILP model is run for every technology individually to derive the COE_{etoh}. The result is the $COE_{ethanol}$ (or NPV_{2014} in $/GJ) if one would invest in that first time period for that 15-year timeframe. The next step is to increase the fermenting efficiencies of the technologies, assuming a linear increase, insert the according technical parameters, and run for the subsequent 15-year time period to result in a new NPV_{2017} for every technology, if one chooses to invest in time step 2 (that is 2017-.2020). This is done until time period 5 with scope 2027-2042 is run with the highest fermentation efficiencies, the highest biomass yields, and all prices corrected to that time frame. The meta-result is a development over time of the most optimal supply chain options, including technology, distance, and feedstock.

The results are shown in
Figure 29 below. Technology ec.8 is drawn on the positive y-axis for visibility purposes of the other configurations. But realistically its economic performance is between -14.07 and -64.04 $/GJ, far below zero and a good business case. The electricity marketing price is kept at 12 cents per kWh. Note that the 2nd generation configurations SSCF (dark lines) are all available after 2017. What first comes to mind in
Figure 29 is that all SSCF configurations are cheaper than their SSF partners (dark versus light colours; except sc.SSCF) due to their reduced component requirement and increased electricity revenues.

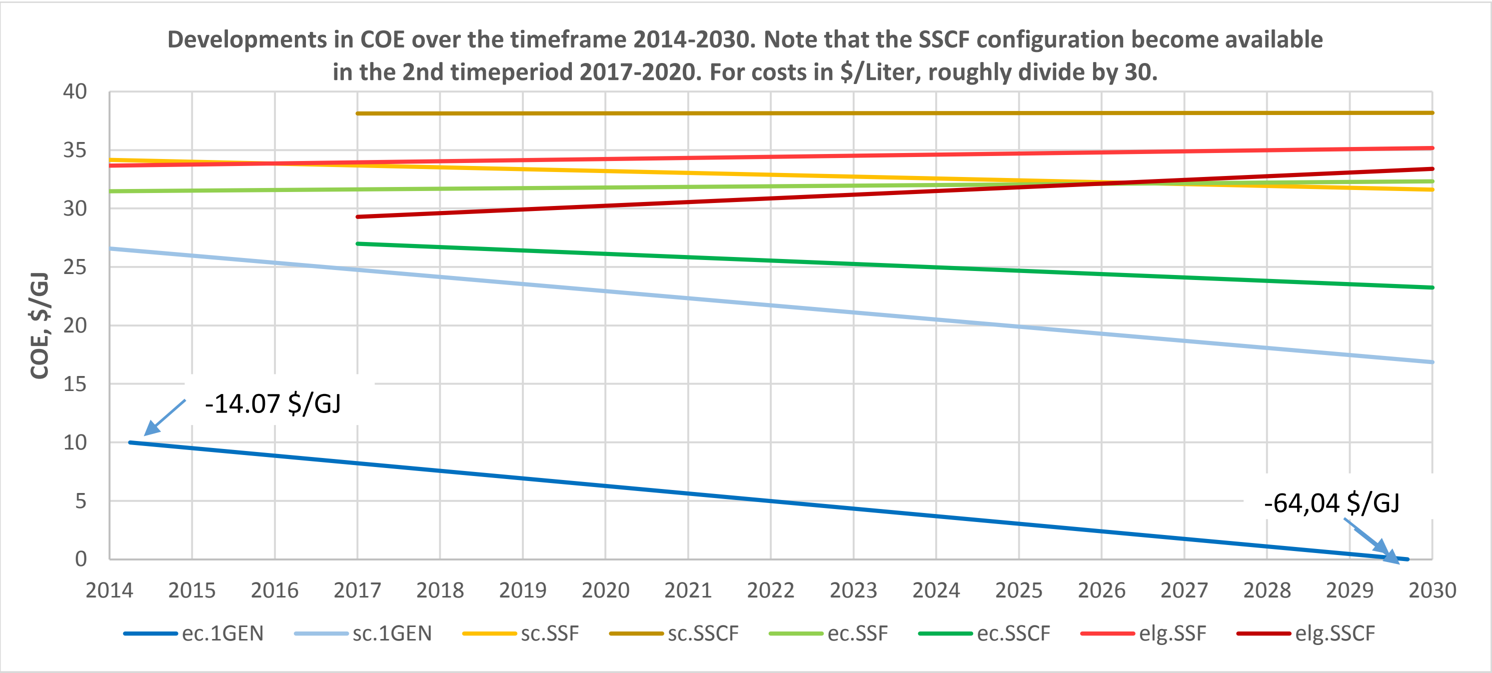
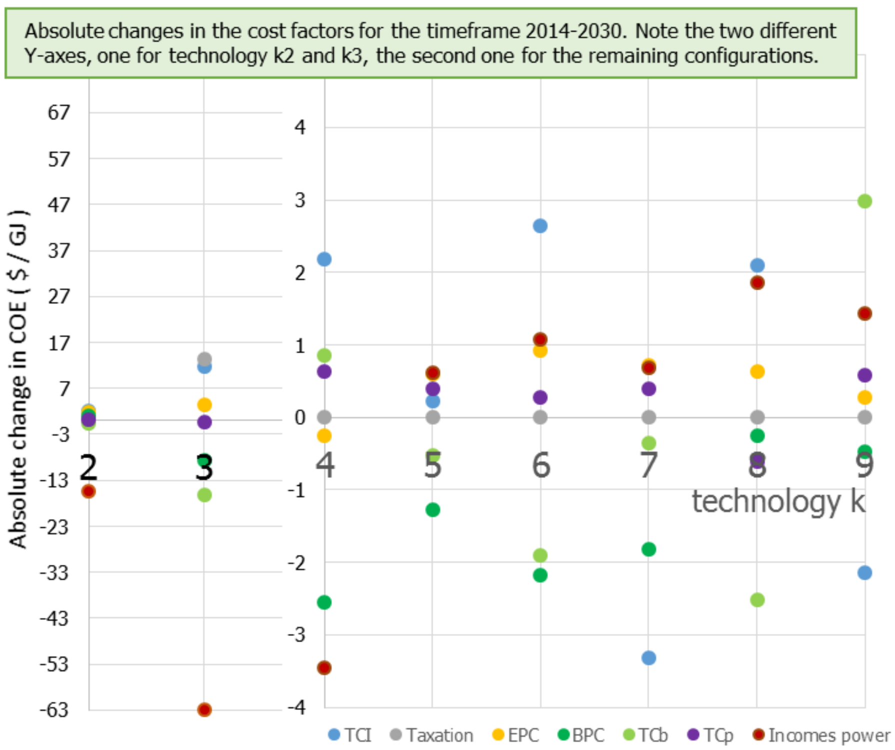


Figure 29. Developments in COE ($/GJ) over the timeframe 2014-2027. note that the SSCF options become available only in 2017. For $/L costs, roughly divide by 100 then multiply by 2,5.

If one goes from a SSF to SSCF configuration, the fermentation components are smaller (lower ethanol scaling per $MW_{HHV,in}$) and cheaper (decreased number of components), see paragraph 4.2.a 'Scaling nine technology configurations'. These decreased efficiencies result in a higher residuals stream for the CHP, and since the power island outweights the fermentation island in price one could expect an increase in total capital investments. But to counteract this, more electricity is produced and at this high electricity price, results in higher revenues.

Conversion routes with a 1GEN technology remain the best options in the future. Second best is the sc.1GEN configuration based on sugarcane. After these 1GEN configurations, the most attractive configuration continues to be the ec.SSCF conversion route (dark green line). No 2nd generation technology configuration becomes cheaper than its 1st generation feedstock counterpart at any point in the future, as one could have expected. The 1st generation sugarcane configuration remains cheaper than both SSF and SSCF sugarcane configurations, and 1st generation energy cane remains significantly cheaper than its SSF and SSCF counterparts.

Figure 30 below illustrates the absolute changes over the time period 2014-2030 for each cost component. Note the two different y-axes, one for the 1GEN supply chain configurations and one for the 2GEN technologies. Both axes are in \$/GJ.



Technology (k)	Nomenclature
2	sc.1GEN
3	ec.1GEN
4	sc.SSF
5	sc.SSCF
6	ec.SSF
7	ec.SSCF
8	elg.SSF
9	elg.SSCF

Figure 30. absolute changes of the cost fctors for the timeframe 2014-2029. Note the two different Y-axes, one for technology k2 and k3, the second for the rest-technologies, both Y-axes are in \$/GJ.

Configurations k2 and k3 show a significant increase in electricity revenues over the time frame solely due to the assumed increase in electricity price (3.02% per year increase for 2013-2042). This increases the electricity revenue with 63 \$/GJ for technology k3. This is the only configuration where a taxation is present as cost component, since taxation is applied only if the NPV is positive. The rest of the technologies have fairly smaller changes in cost factors of between -3.5 and 3 \$/GJ. The most important change is again the electricity revenue (all other configurations except k4), which is above zero in Figure 30 but is actually a decrease since 'incomes from power' is a negative *cost* factor. This means that electricity revenues –except for k4- decrease for all technologies because of increasing fermentation efficiencies.

The two most important cost decreases are assigned to biomass production costs and transport costs for biomass (light and dark green). Yield improvements reduce the unit production costs (in \$/ton biomass). Transport costs of biomass are reduced because more can be produced closer to the facility, and biomass from farther away can be omitted. Total capital investments do increase for all configurations due to general price increases for capital over time, except for configuration 7 and 9 (both SSCF). For these technologies, the ethanol production rises by resp. 13% and 19% due to efficiency improvements (highest increases of all configurations). Due to this increase the plant capacity requirement at a 20-40% ethanol demand decreases with roughly 100 $MW_{HHV,in}$ for both configurations (energy cane and elephant grass) in 15 years, and this translates in a net decrease in total capital investments. Other cost factors stay relatively within the -1 to 1 $/GJ marge, which is mostly ascribed to the application of price indexes.

Greenhouse gas (GHG) emission performance of the ethanol supply chain configurations discussed in this thesis are all very favourable for Puerto Rico. This has to do mainly with the high emission credits that are received from excess electricity production that replaces fossil based electricity with an island-wide emission factor of 0.80 kg CO2 eq./kWh (as discussed in paragraph 3.2.k 'GHG emissions from ethanol production').

In Figure 31 are outlined the GHG emissions against the COE of the supply chain configurations. Also has been added one configuration where two technologies are combined, namely k3 with k7. If both k3 and k7 were used, this configuration would have a COE of 12.48 $/GJ (exclusively energy cane-based, optimized with elephant grass) (see Figure 31 below).

All technologies have a negative GHG emission performance, with the 1GEN configurations (k3 or k2) showing the highest GHG emissions *avoided*. All SSCF technologies have better GHG performance than their SSF counterparts due to their higher electricity production.

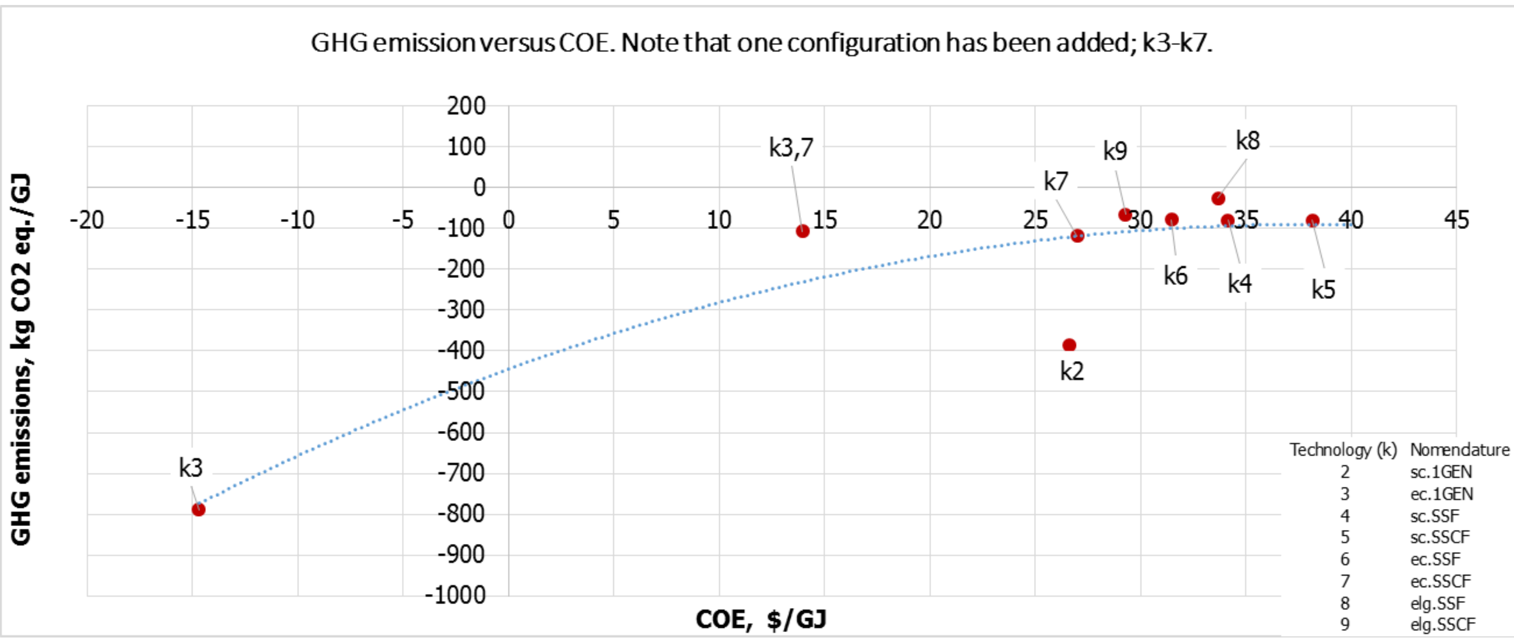


Figure 31. Total system greenhouse gas emissions versus NPC. Note that two configurations have been added; one combination with of k3-k9; and one between k3-k7.

To go deeper into a location analysis a choice must first be made of the technologies that are best fitted. The choice between a k3-, k7-, or k9 configuration is not so easy as it seems, as each has its particular advantages. An energy cane 1GEN configuration has the most favourable NPV, but is limited in its ethanol production.

5.2. Strategic policy planning for an ethanol supply chain

Table 34 summarizes some characteristics of technology k3 (energy cane, 1GEN), k9 (elephant grass, SSCF) and k7 (energy cane, SSCF) if all 82,978 hectares would be cultivated with biomass yields of 2014. It presents the ethanol production in million barrels per year and in Petajoule (x 1,000,000 Gigajoule), as well as the excess electricity in $MW_{electricity}$ installed capacity and in Petajoule.

Table 34. Maximum ethanol yields for k3, k9 and k7 at an area of 82,978 hectares with 2014 biomass yields

Technology	Ethanol	Power	Ethanol	power	Total energy
k	million BBL/year	MW-e Excess	PJ	PJ	PJ
k3	3.3	1041	8.2	30.0	38.2
k9	9.2	336	23.1	9.7	32.8
k7	8.5	481	21.3	13.9	35.2

Configuration k3 has the highest electricity potential of 30 PJ (1041 MW_e of excess electricity capacity to the grid) but a low ethanol potential of 8.2 PJ. Elephant grass configuration k9 is characterized by the reverse, with an ethanol potential of 23.1 PJ but lower excess power of 9.7 PJ. Energy configuration k7 has a roughly similar ethanol conversion (21.3 PJ) but slightly higher electricity potential (13.9 PJ). Totalling on an energetic value, option k7 has a potential of 35.2 PJ against 38.2 PJ for k3. Additionally, due to the higher electricity production, option k3 has the preferred COE. Since this thesis is focused mainly on the development of an untapped renewable energy resource that is bioethanol, a technology with a favourable ethanol conversion will be added to the k3 configuration. As second technology will be chosen configuration k7=ec.SSCF. This configuration has the best overall energetic efficiency (52%), and has the most favourable COE of the 2nd generation options. Furthermore, it preserves these advantages in the future.

One additional assumption is made before a spacially explicit allocation of plant capacities is presented. Since a 20-40% ethanol demand was dictated for the supply chain in the MOMILP model, and in timeperiod one only configuration k3 is available, the first 20% of the ethanol capacity must come from this configuration. However, due to the low ethanol conversion efficiency of this configuration (only 12%), this initial 20% ethanol output would require a total installed ethanol capacity of 2700 $MW_{HHV,in}$ and a land cultivation of 51,000 hectares of energy cane/elephant grass *immediately* in the first timeperiod 2014-2017. Given the unfamiliarity of Puerto Rico with this renewable energy resource, together with the start-up time, -investments and -research required to organize this supply chain, this scale is unrealistic. Therefore it is chosen to run the optimization again for configurations k3-k7, but now with dictated ethanol output of 10-20-25-30-35% in subsequent timeperiods. Now the MOMILP model presents a supply chain that consists of configuration k3 supplying 10% ethanol demand in timeperiod $1_{2014-2017}$ while the increased output 20-35% is supplied by configuration k7 in the remaining $timeperiods_{2017-2030}$[15]. The COE of this k3-k7 configuration was presented in Figure 27 and amounted to 12.48 $/GJ. Figure 32 below illustrates the geographical planning of plants and plots for the entire timeframe 2014-2030.

[15] In the calculation of COE, also the E_t or energy $production_{etoh}$ is discounted (see paragraph 2.1 'Main structure' for the calculation method of COE). The discounted amount of produced ethanol for the 20-40% output is 79,023,000 GJ. An ethanol output of 10-20-25-30-35% has an equal discounted ethanol output of 79,023,000 GJ. This strategy is followed in the previous paragraph as well, when a constant 30% ethanol output was dictated to investigate the effects of increased ethanol conversion efficiencies on theCOE over the timeframe 2014-2030.

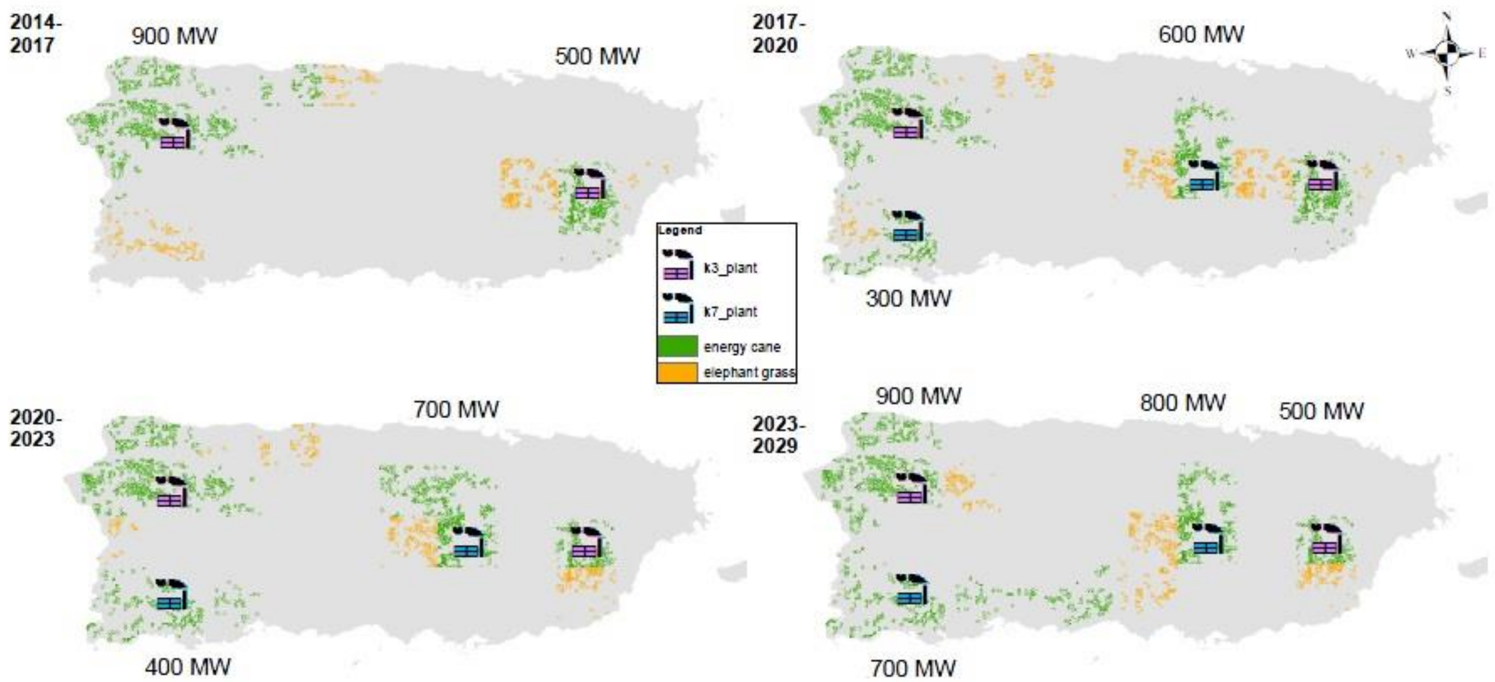


Figure 32. Planning scructure of ethanol plant and biomass plot locations. Green is enegy cane; red is elephant grass.

This ethanol supply chain will produce 1.4 to 4.9 million barrels of ethanol and will substitute 11.2% increasing to 22.4%[16] of Puerto Rico's gasoline consumption over a 2014-2030 timeframe. On four locations an ethanol plant is built. According to time period, configuration, county, region and capacity:

2014-2017;	1st generation fermentation, energy cane (k3);	San Sebastian, region 3, north;	900 MW;
2014-2017;	1st generation fermentation, energy cane (k3);	Yabucoa-Humacao, region 5;	500 MW;
2017-2020;	2nd generation SSCF, energy cane (k7);	Orocovis, region 1;	600 MW;
2017-2020;	2nd generation SSCF, energy cane (k7);	San German, region 3, south;	300 MW;

For the remaining timeperiods, the latter two plants are expanded to satisfy the increase in ethanol demand: for Orocovis, region 1 the plant is expanded from 600, 700 then to 800 MW; for San German it is expanded from 300, 400 then to 700 MW. The two expanding plants are immediately build at 800 MW (Orocovis) and 700 MW (San German), thus the capital investments (TCI) are for these capacities. The mentioned capacity increases are merely answering to the ethanol increases as dictated in the MOMILP model. It also enables a sustained expansion of biomass cultivation over the timeperiod.

The total capital investments for these plants for this timeframe are 2,010 million dollars$_{2010}$. The orange grids prepresent the cultivation of elephant grass that is needed to optimize the CHP units of the plants. Increases in demand of land is clearly visible in the south-west, where the k7 plant is expanding from 300-700 MW ethanol capacity, whereas land availability in the south is less abundant and scattered more. On the other hand, for the regions in the viscinity of the other three plants, biomass plots are diminishing over the timeframe due to improvements in biomass yields. Table 35 below summarizes some more characteristics of this supply chain configuration.

[16] Remember that for the MOMILP optimization an ethanol output was dictated of 20-40% gasoline substitution *of the 10 largest population centers*. This 20-40% output equals a gasoline substitution of 11.2% to 22.4% of the island's *total* gasoline use.

Table 35. Ethanol and power capacities and biomass requirements per grid g and timeperiod t, of the k3-k7 supply chain configuration based on energy cane.

			Time period				
Product	Technology	Grid	1	2	3	4	5
MW-HHV ethanol installed capacity							
Ethanol (region 3)	3	13	900	900	900	900	900
Ethanol (region 5)	3	31	500	500	500	500	500
Ethanol (region 1)	7	29		600	707	800	800
Ethanol (region 3)	7	35		300	357	476	689
MW-e installed capacity (at 8000 hour loadfactor)							
power	3	13	319	319	319	319	319
power	3	31	134	134	134	134	134
power	7	29		117	133	151	151
power	7	35		43	67	90	130
Total installed capacity (MW)			453	613	653	693	733
Feedstocks delivered to the sum of plants (million tons)							
energy cane			2.83	4.62	5.06	5.51	5.96
elephant grass			0.47	1.04	1.18	1.33	1.47

The second part of Table 35 above also includes the excess power capacity per grid g and timeperiod t. For this 15-year timeframe and ethanol demand, this configuration will contribute from 453 towards 733 MW_e electricity capacity with a loadfactor of 91% to the national grid, if the feedstock flows to the plant are optimized as proposed. With the island's total installed capacity of 5,365 $MW_{e,2010}$ this supply chain configuration will sustainably substitute 14% of the island's electricity production.

The last part of the table summarizes the feedstock flows to the sum of plants; this is 2.83 million tons of energy cane and 0.47 million tons of elephant grass in timeperiod $1_{2014-2017}$ produced on 27,500 hectares of energy cane and 8,200 hectares of elephant grass. For the last $\text{timeperiod}_{2026-2029}$ this is 5.96 million tons energy cane on 33,000 hectares and 1.47 million tons of elephant grass on 18,000 hectares.

Sensitivity analysis. Although a great efford has been made in this research to present a thorough analysis on the perspectives of an ethanol supply chain, there remain some uncertainties that can have a profound effect on its feasibility. In a sensitivity analysis it is tried to identify the extend of some of these uncertainties. Figure 33 below illustrates the the extend to which the COE_{etoh} will vary if a selected amount of parameters will change. The COE at the base case (100%) is the economic performance of the presented supply chain if all parameters are as presented in this research. Subsequently are varied five parameters ranging between minus and plus 16-50% from their base case values.

The price at which an excess electricity production is sold to the national grid has the most significant effect on the COE, and already discussed in previous paragraph 5.1 'Analysis of supply chain configurations'. Its range is varied from 12 $ct/kWh (100%, base case) to 10 $ct/kWh (-16.67%) and 14 $ct/kWh (+16.67%). At the former range the COE is more than doubled, whereas at the latter range the COE is zero and thus a break even point. Further debate and policy development can indicate the appropriate and feasible selling price of electricity, as this parameter can be utilized as an effective support measure.

The second most important parameter is the unit production costs of biomass production, which is varied with +/- 20% from the base case. This parameter has a significant influence on the overall performance of the supply chain, and reduces/increases the COE with roughly 50%. It is also the most uncertain parameter, where more reseach is needed to adjust for the right biomass production cost levels. Although it was tried to differentiate geographically as much as possible cost factors (fertilizer input, tonnage dependent costs etc), most investment estimations for the life cycle stage 'biomass production' (machinery, labour, administration, land clearance and preparation) were only from region 3 in the south (Conty, 2006).

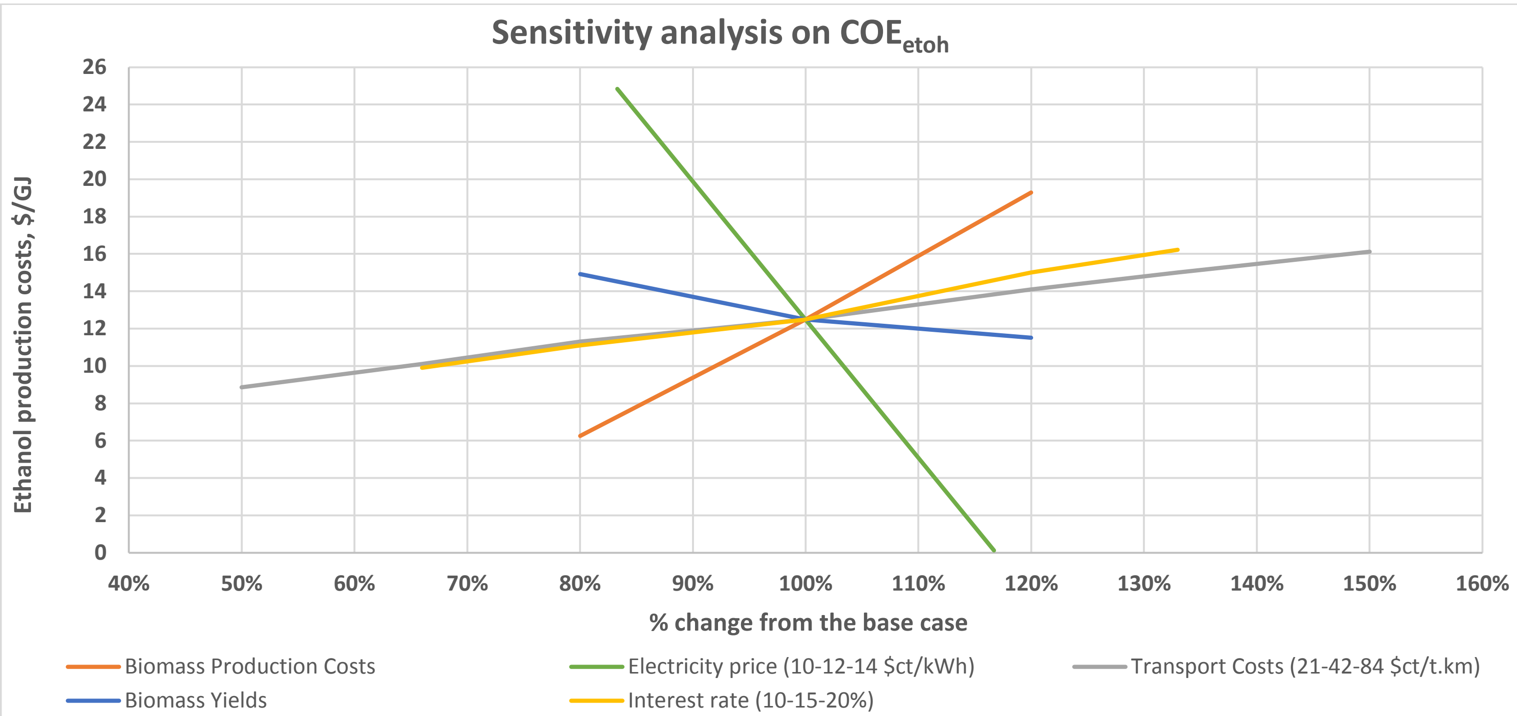


Figure 33. Sensitivity analysis on the COE_{etoh}.

The biomass yield levels were also varied with +/- 20% compared to the base case. But this parameter has less influence on the overall COE, because its influence is mainly reflected in transportation costs of biomass, i.e lower yields mean a larger biomass sourcing area. It has to be mentioned that this parameter is varied independently from its production costs. In other words, a yield increase/decrease of 20% is assumed at the base case production costs of its region. It is more realistic to assume an according increase of biomass production costs if a region is able to produce larger outputs of biomass, but this relationship is not quantified in this sensitivity analysis.

The blue line is not a straight one, and deceases less steep after 100%. The most ideal locations with this COE of 12 $/GJ were in San Sebastian (north region 3; 900 MW, timeperiod 1) and Yabucoa-Humacao (region 5, south, 500 MW, timeperiod 1) equipped with k3 plants (discussed in this paragraph). Average estimated yields in region 3 for energy cane were 90 metric tons per hectare, and 106 tons/ha in region 5. San Sebastian is still preferred for a 900 MW k3 plant due to the high density of available land in that area. However, at a minus 20% yield sensitivity analysis, this advantage is diminished and the 900-500 MW capacities switch location. In other words, the 900 MW is allocated to region 5, and only 500 MW to region 3. This is reversed again at plus 20% yields, although less than linear.

The last two parameters 'interest rate' and 'transport costs of biomass+fuel' are the least influential on the COE and follow nearly similar deviations. Transport costs are increased/decreased with 50% for both biomass and fuel transport, i.e. 0.21 and 0.63 $/t.km with 0.42 $/t.km as base case. The COE of the proposed supply chain is increased with 30% at this upper bound of the transportation range, from 12 to 16 $/GJ. With increased transportation costs, an earlier allocation of smaller plants over different regions occurs, as will be discussed in more detail in the next paragraph.

The interest rate has a similar effect on the overall COE of the supply chain as transportation costs. The rates are varied with +/- 33%, or an interest rate of 10% or 20%, with 15% as base case. An interest rate of 20% would equally increase the COE with 30%. Nevertheless, an interest rate on the lower bound can be expected, as private investment rates are usually between 8-15% and social interest rates between 3-7%. Given the fact that this analysis included a broad supply chain economic performance, where multiple facets of the productive economy must be directed through a concerted public-private effort, an interest rate between these two ranges can be expected.

5.3. Regional capacity allocation

The allocation of ethanol production plants and according capacities is closely related to the biomass sourcing area and subsequent transportation costs of biomass. Figure 34 right illustrates this general relationship. Where the plant's investment costs (per unit product, e.g. $ per liter or GigaJoule ethanol) usually decline at increasing capacity due to economies of scale, so does the transport cost of biomass increase due to a larger required sourcing area of biomass. One of the main advantages of the MOMILP model was precisely to solve for this trade-off. In this section this capacity-transport relationship is analyzed for Puerto Rico.

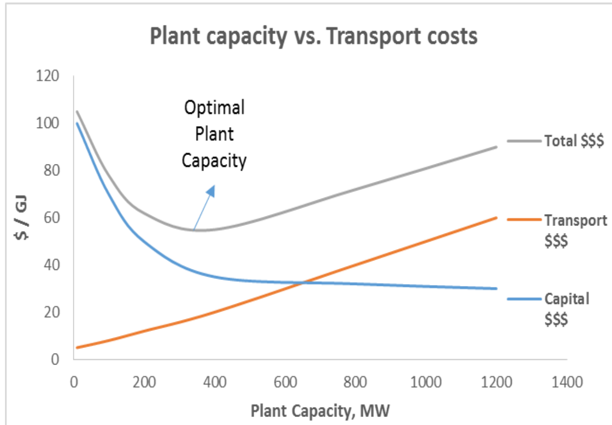


Figure 34. Capacity-transportation relationship.

The capacity-transport relationship will be presented for one technology, namely k7 or ec.SSCF.8, using energy cane as a main resource and SSCF as conversion technology. In paragraph 5.1 'Analysis of supply chain configurations' it is argued that technology k3 for its electricity generation advantage, combined with k7 for its favorable ethanol production is preferred. Due to the low biomass-to-ethanol conversion of configuration k3, it was thought most interesting to present configuration k7 in this analysis. Neither feedstocks of other configurations are included due to the relative cost differences between the three feedstocks. In paragraph 4.1 'life cycle stage 'Biomass Production' it was assumed that energy cane would harvest +20 metric tons/hectare, and elephant grass *0.7 tons/hectare, compared to sugarcane, for every individual region. Thus, due to these simplifications a switch to another feedstock/configuration will be only proportional to this k7-analysis. See paragraph 5.1 'Analysis of supply chain configurations' for the differences in cost performance of the configurations.

Also the number of locations for ethanol plants has been reduced to five, namely within the largest biomass resource areas. The Unit Production Costs of biomass (UPC_b) are yield dependent, and both the UPC_b and biomass yields (BY) are aggregated to five regions, i.e. spatial differentiation of these two parameters is simplified to five. Therefore allocation to different grids in the same region will result in roughly the same costs, and only differ on transportation costs of biomass/fuel. Additionally, all runs with all grids set available indicate these five locations as most optimal, due to their immediate vicinity of large biomass potentials, and highest biomass density within these biomass areas.

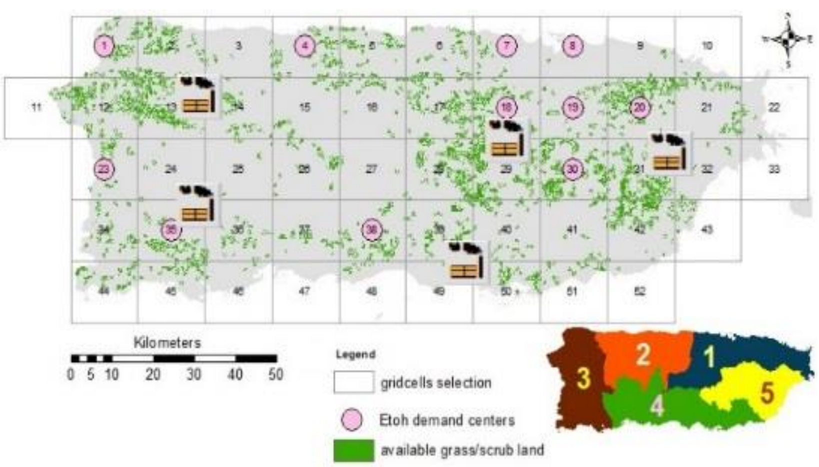


Figure 35. Potential plant locations.

The five potential plant locations are illustrated in Figure 35 above, and are –around the clock- in:
Region 3a (brown): located in the north-west region 3, in grid 13 (San Sebastian);
Region 1 (blue): located in the center-east region 1, in grid 29 (Orocovis);
Region 5 (yellow): located in the east region 5, in grid 31 (Yabucoa-Humacao);
Region 4 (green): located in the south region 4, in grid 49 (Ponce area);

Region 3b (red): located in the south-west of region 3, in grid 35 (San German).

From now on these potential plant locations will be named according to their region. The electricity price is set at 120 $/MWh, whereas transportation costs at 150% of the base case (0.63 instead of 0.42 $/ton.km). The latter is chosen such as to increase the role of transportation costs in the analysis, and to stay on the conservative side. Location 5 and 4 are furthermore restricted to capacities of 800 $MW_{HHV,in}$ due to surrounding mountains (north of 5, north of 4) and low biomass availability (4). At last, the capacity range of the plants are from 100 to 4000 $MW_{HHV,\ in}$, with intervals of 100 between 100-800 MW, of 400 between 800-2000 MW, and of 500 between 2000-4000 MW. An island-wide capacity of 4000 MW for configuration 7 means an area utilization of all 82,000 hectares available from the PLUC model. Figure 36 below illustrates the capacity-transport curves for all 5 regions.

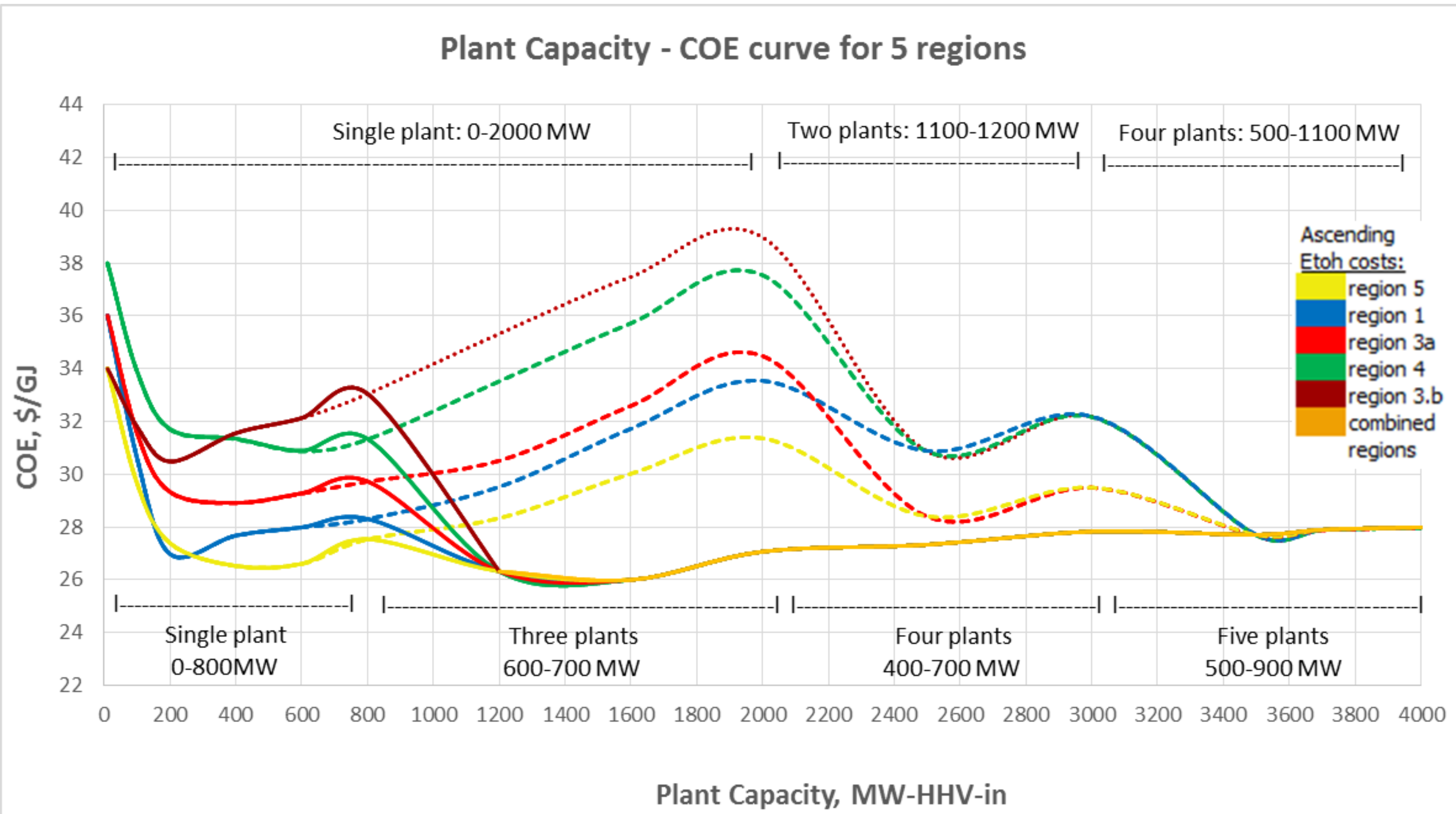


Figure 36. Capacity allocation curves for five regions.

The graph consists of three main parts of information: (1) the dashed lines, (2) two sets of peaks at 2000 and 3000 $MW_{HHV,in}$, and (3) the solid lines converging in the lower orange line.

The solid lines proceeding on the dashed lines are the total supply chain OCE_{etoh} for the regions at increasing capacity of *one single* plant. After peak #1 around 2000 $MW_{HHV,in}$ ethanol capacity, single plants in one region diverge in two plants in distinct regions. For example, point [2600:28] is the COE of 28 $/GJ, of a supply chain with two smaller plants in region 5 (yellow; 1250 MW) and region 3a. (red; 1250 MW). Remember that this is a theoretical curve, since region 5 is restricted to 800 MW due to northern mountains[17]. After peak #2

[17] This physical restriction of mountains, or more specificly the transportation over these mountain ranges, was difficult to restrict in the MOMILP model. The only factor to restrain transport between gridcells was the 'tau' unit (see page 129 in appendix 10.2.b 'MOMILP linear formulation'). But passage between gridcells cannot be closed-of permanently because the transportation of ethanol would be closed off, too. A better way to represent transportation is to include an ArcGIS transport network to produce real-life

at 3000 MW the two plants diverge to smaller plants in four distinct regions, thus converge to a COE_{etoh} of 28 $/GJ.

Above course of events occur when the plant capacity range in the MOMILP model is set maximum, i.e. from 100-4000 $MW_{HHV,in}$ with earlier mentioned intervals. The actual convergence to smaller plants in different regions at increasing plant capacity occurs after 800 MW to a COE of 26 $/GJ. This occurs when the capacity range in MOMILP is limited to 800, 900 or 1,000 MW. Interestingly, the COE now increases slowly towards 28 $/GJ at whatever capacity addition until 4000 MW, as it allocates this added capacity to either region.
The maximum plant capacity before an occurring divergence of smaller plants to other regions, seemed 700 MW. Per island-wide capacity range, region, individual plant capacity, and COE:

800 to 1200 MW,	region 5, 1, 3a,	capacity 300, 500, 400 MW, resp.	COE 26.0 $/GJ
1200 to 2000 MW,	region 5, 1, 3a,	capacity 600, 700, 700 MW, resp.	COE 27.0 $/GJ
2000 to 2500 MW,	region 5, 1, 3a, 3b/4,	capacity 500, 600, 700, 700, resp.	COE 27.5 $/GJ
2500 to 3000 MW,	region 5, 1, 3a, 3b, 4,	capacity 700, 700, 700, 600, 300, resp.	COE 27.8 $/GJ
3000 to 4000 MW,	region 5, 1, 3a, 3b, 4,	capacity 800, 900, 900, 700, 700, resp.	COE 28.0 $/GJ

North regions are allocated first, in region 5, 1 and 3.b which are the cheapest (5), or have the largest biomass potential (5, 1, 3b). Increasing to 3000 MW, a plant in region 4 or 3b can be added, i.e. there is no difference in overall COE adding either two regions. Increasing to 4000 MW, all 5 regions are equipped with 700-900 MW ethanol plants. In this island-wide capacity range, two plants of 900 MW are included (region 1, 3a), one of 800 MW (region 5), and two of 700 MW (region 3b, 4).

It is remarkable that region 4 is last in allocation. This region has the lowest production costs of biomass, and has seen the highest sugarcane yields in the island's history (see paragraph 4.1.b. 'Biomass Yields'). This is also the reason why the PLUC model allocated almost all agricultural land to the south, and leaves the region with scattered and little land available in 2030 (see chapter 3 'Model 1: PLUC inputs & results' for a detailed elaboration on the PLUC model and its results).

Furthermore, region 3a at San Sebastian, is a favorable location despite its high(est) UPC and lowest estimated BY. Solely due to the high density of biomass (thus low transport costs) this location is one of the favorable ones for plant location (after region 1 and 5). Nevertheless, the sensitivity analysis showed that at a 20% reduction in biomass yields, the allocated 900 MW capacity of k3 configuration in region 3a would switch with the 500MW capacity in region 5.

Above analysis on capacity allocation and transportation is illustrated graphically in Figure 37 below. The illustration shows the sourcing area at according COE_{etoh} for the five plant locations at increasing capacities of an *individual* plant (top for the regions 3a and 1; middle for the regions 3a, 4, and 5). Colors represent the COE, and every next color -starting from each individual plant- indicates a +400MW increase in installed capacity. The bottom picture illustrates the COE if a supply chain includes multiple smaller plants in different regions as described above. Its COE stays below 28 $/GJ.

distances between grid centroids, and use the 'tau' unit to represent e.g. inclination. But still the transportation of fuel is far easier than that of biomass, in the first place because of its fairly smaller volume.

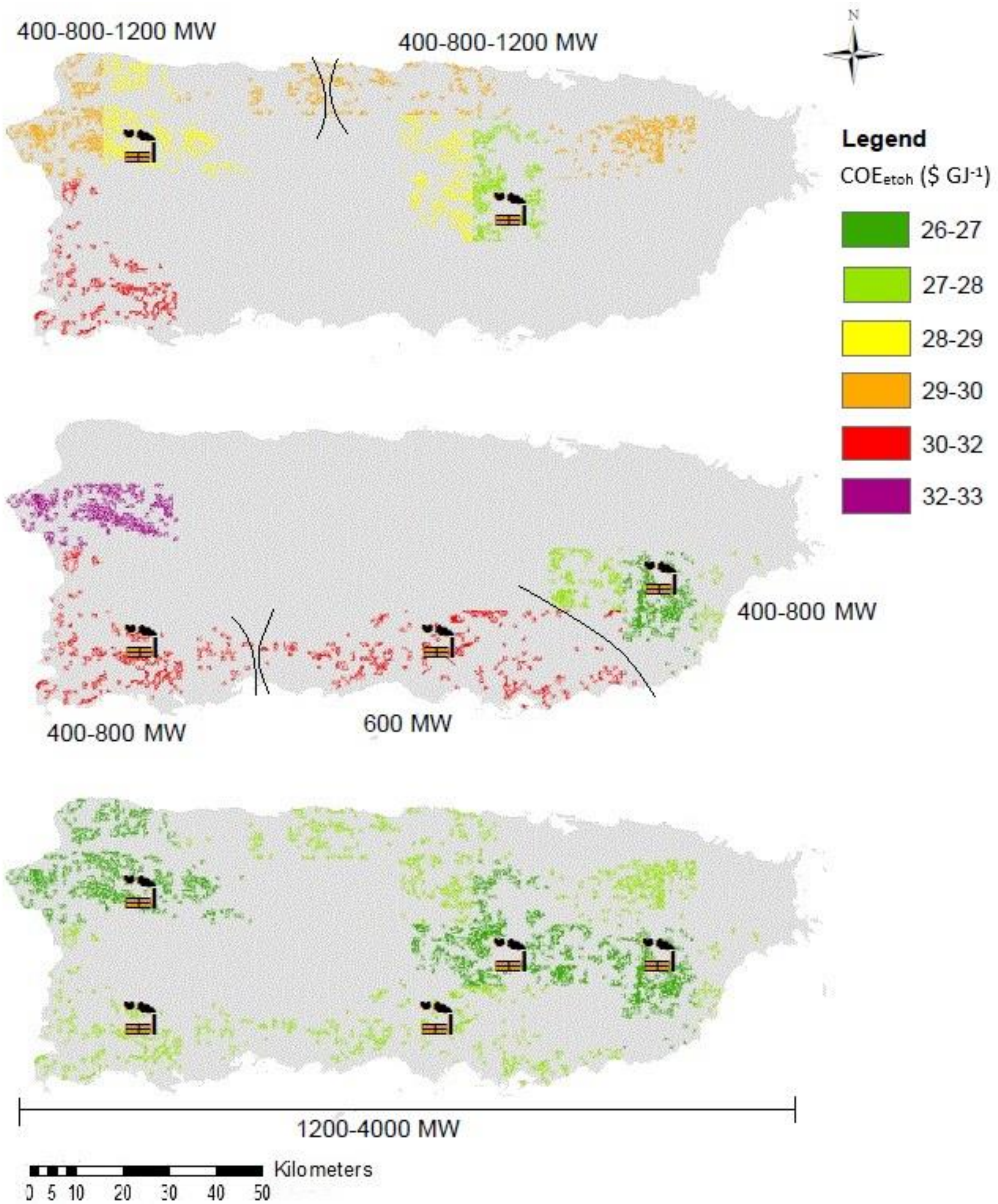


Figure 37. Biomass Sourcing areas for individual plants (upper,middle) and for an integrated supply chain (bottom).

6. Conclusion

The availability and fertility of the land, together with socio-economic circumstances of high fossil energy use, -imports, and accompanied high costs, together with an urge to switch towards local energy sources makes a biomass based energy sector an attractive option to the renewable energy mix of Puerto Rico. Bio-energy can and must be stimulated to provide a solid energy basis for the island. Nevertheless, the high population density situation makes it increasingly important to incorporate all relevant land uses if the land requirements are being assessed. The first model PLUC has successfully allocated land until 2030 to all other dynamic land use classes 'urbanization', 'cropland' and 'pasture' according to their estimated future demand.

The main aim in the preparation of the first model PLUC was a sustainable future for Puerto Rico, where sustainability is highly intertwined with self-sufficiency, first in food then in energy. The assumption that the island will give a renewed impulse to the agricultural sector by doubling its agricultural and livestock output in 2030, left enough room (read 'land') to start-up a bio-ethanol supply chain. Due to the progressive scenario with increased agricultural output and 3,5% yield improvements per year, an area of 36,000 hectare is freed-up in 2030 compared to the scenario where yield improvements would follow the historical trend. The total amount of grassland and scrubland (on former crop/pastureland) that is still available in 2030 is 82,978 hectares. This is after PLUC's allocation of land to all other land use classes.

The choice for the three feedstocks combined with three different ethanol conversion technologies (resulting in 9 technology configurations) was necessary as to whether the newly built ethanol supply chain would use its historical sugarcane, of other feedstocks with increasing fiber contents (energy cane or elephant grass). The land potential of 82,978 hectares has an energy potential of 38.2 PJ (8.2 PJ ethanol and 30 PJ electricity) if a supply chain is chosen based solely on energy cane being converted through a conventional 1st generation juice fermentation, or k3 configuration. A supply chain based on energy cane and a 2nd generation 'Simultaneous Saccharification Combined Fermentation' (SSCF) technology, or k7, has the potential of 35.2 PJ (21.3 PJ of ethanol and 13.9 PJ of electricity). This technology becomes available after 2017. The k3 option is the best performing configuration, while the k7 option is the best performing 2nd generation conversion configuration. On ethanol terms, this potential substitutes an amount of 16.5% (k3) or 41% (k7) of Puerto Rico's gasoline consumption based on energy content.

These two technology configurations resulted in the most favourable cost of ethanol (COE_{etoh}) out of all the nine configurations. The cost and emission performance of this supply chain includes biomass production, ethanol production and transport of biomass and ethanol. With a second model called MOMILP, or Linear Programming model, the most optimal ethanol supply chain configuration is chosen out of a range of options, such as 3 feedstocks, 9 technology configurations, 52 locations for ethanol plants or biomass plots and plant capacities in $MW_{HHV,in}$. The model is dictated to deliver over a time period 2014-2030 an amount of ethanol equal to 20-40% of the gasoline consumption of the 10 largest urban centres on the island, corresponding to 11.2 to 22.4% of the islands gasoline use in 2010. If MOMILP model is run, it will select the technology and feedstock that supplies this ethanol demand the cheapest way over that time period. It also indicates in which area the plants/plots will be located, what the plant's capacities will be, and what the GHG emissions are of that supply chain. Below are once more summarized the technology configurations.

An energy cane-based supply chain configuration k3 (ec.1GEN) resulted in a **benefit** of 14.07 $/GJ (0.50 $/Liter ethanol) and thus good business case for Puerto Rico. The second-best option is a supply chain based also on energy cane (ec.SSCF) with a **cost** of 26.98 $/GJ (0.92 $/Liter ethanol) and must be subsidized. The first option thus shows the best performance in economic terms, but also on emissions terms. This energy cane-based configuration has an emission factor of -790 kg CO2 eq/GJ, if emission credits from excess electricity is allocated to ethanol production. These two measures (COE and GHG emissions) are so favourable due to the high electricity by-production of configuration k3, and subsequently its substitution of fossil-produced electricity with emission factor of 0.80 kg CO2 eq/kWh.

These two configurations are attractive on different terms in that configuration k3=ec.1GEN has a favourable electricity production, while configuration k7=ec.SSCF has a favourable ethanol production. For this reason a combination of k3 and k7 configurations is the preferred ethanol supply chain for Puerto Rico, using the high electricity revenues to keep the costs down of the whole system over the timeperiod 2014-2030. The marketing price of electricity to PREPA is kept constant at 12 $cents/kWh, this is an important factor in the economic performance of the supply chains. In fact, both configurations continue to be the best performing options at an electricity price between 8 and 15 $cents per kWh. Below a price of 7 $cents configuration k3 becomes the most expensive one.

If one would wait with investing in such a supply chain, would the conversion efficiencies (and thus performance) of the technologies be more favourable in the overall costing? For this reason is incorporated for the time frame 2014-2030 cost increases due to inflation, biomass yield improvements, subsequent GHG emission reductions on farm-level, and ethanol conversion efficiencies on plant level. Technology configurations k3 and k7 remain the two best options for an ethanol supply chain in the future. First generation configurations (based on energy cane or sugarcane) continue to be cheaper than their 2^{nd} generation counterpart SSF or SSCF based on the same feedstock. This is because the SSF/SSCF configurations increase their ethanol output, but at the expense of their by-product electricity which is a significant cost (read benefit) component. The higher the fiber content of the feedstock, the lower the cost of energy due to more (ligno)cellulose utilization. For this reason technology k7 is best option after energy cane k3.

The most important cost factor changes over the timeframe 2014-2030 are the Capital Investments and electricity revenue due to general price increases (increasing prices). The most important decreasing cost factors are 'biomass production costs' and 'transport costs of biomass' due to the assumed yield increases of biomass. Concequently, distances to the production plant decrease due to more biomass availability in the vicinity of the ethanol plant, thus decreasing transport costs. In the Puerto Rican case it can be concluded that advantages in fermentation efficiencies of 2^{nd} generation technologies combined with the effect of biomass yield increases are undone by the effect of price increases (for inputs, electricity, capital). All SSF/SSCF configurations increase in price over the time frame. This in contrast to the first generation fermentation technologies, that decrease over time due to the increasing valuation of electricity into the future (…and they produce the most electricity).

The sensitivity analysis showed that the electricity price had the largest influence on the economic performance of the supply chain. At a kWh price of 14 $ct the supply chain would reach a break even point, or NPV of zero. The second most important factor of influence on the economic performance is the Unit Production Cost of biomass. An increase of 20% in production costs of biomass would increase the overall economic performance of the supply chain with 50%.

A spatially explicit analysis on plant capacity allocation favoured the establishment of multiple ethanol production plants of 700-900 $_{MWHHV,in}$, in different regions, above the establishment of one or two megaplants of +1000 MW. This was analysed for a supply chain including only configuration k7. An allocation of these capacities to three, four then five regions over the timeframe 2014-2030 would keep the overall COE between 26-28 $/GJ. At scales of between 700-900 MW, an optimum is found between economies of scale on plant level, and its according biomass sourcing area.

These findings are just a start of what I would call 'the sustainable bio-energy future' endeavour for Puerto Rico. This thesis is the first research of its kind on the island in that it provides a profound energy analysis, combining buttom-up analysis with a top-down optimization, delivering a good basis for policy planning. It furthermore accentuates the possibility for an increased energy *and* food independency. Nevertheless, the results must be interpreted with care, since the literature pool on the island was very poor, out to date, or non-existent on some subjects. Biomass production data was either too old, too experimental (extreme high yields) or absent (data for elephant grass was gained from Colombian sources). Nevertheless, historical sugarcane yield documentations by the national planning board and USDA agricultural censuses were trustworthy and reported relatively high yields compared to Brazil or South Africa, and equalled yields in Colombia. How fast these yields will be reached after a restart of the industry must be seen. Much more research is required to fine-tune the input parameters that were used for these results.

7. Discussion & Recommendations

This chapter will discuss some difficulties that were encountered while working with the two models, and some recommendations for further improvement may the two models be used for futher planning of the agricultural and bioenergy sector. I will also discuss some fallacies that I encountered and could not resolve due to a lack in linear programming skills. Furthermore I will discuss the success of integrating the two models.

Concering the PLUC model. The most important obstackle in using the PLUC model, or better, using satalite imagery as input for PLUC, is the discrepancy between documented statistics on the ground on the one hand, and the magnitude of land use classes of the ArgGIS maps on the other hand. As I discussed earlier in paragraph 3.1.d 'Biomass Land Availability', and in paragraph 4.1. 'PLUC results', the amount of land used for agriculture and pasture as documented by the USDA agricultural cencuses was much larger than the amount of land of these classes in the ArcGIS maps. Two cases were particularly remarkable. For agriculture the documentation of coffee plantations were included in the censuses but were difficult to extract with satalite imagery. This is still a significant sector for Puerto Rico, and is mostly shade-grown. Thus coffee plantations are at the same time classed as 'forest'. Of course these kind of difficulties arise when one aggregates agricultural output over an aggregated land use area, as has been done in this thesis.

The other case is the documentation of pasture area in satalite imagery. This is difficult if cattle is ranged on cropland, scrubland, or grassland. The GAP Analysis, from where the ArcGIS maps are from, documented all grassland as 'pasture/grassland'. In this particular case I could have made a particular 'pasture/grassland' class, but that would include all grassland and/or scrubland. I decided to enable pasture land to expand to the documented amount of land (USDA censuses) in the first year 2014-2015 according to a relatively accurate cattle density map. Nevertheless, the amount of pasture land that expanded in the first year was around 660 km^2, while for the rest of the timeframe pasture expansion according to the scenarios was only 2 km^2 (assuming a progressive yield improvement). Thus, the former land use increase can be better approached as a documentation of a new land use class 'pasture/grassland'. Within this area there is much room for intensification of the sector. The most important objective, that enough land is allocated to the competing land use classes cropland, pasture and urbanization for the foreseeable future, is upheld.

Another particularity concerns the scale of the geographical region the PLUC model is applied to. This is the first time the PLUC model has been applied to a very small (island) region. It was only tested on the very large countries Mozambique, Poland, and in preparation for Brazil. In these cases, land allocation is visibly allocated to distinct areas, i.e. the allocation of land is more visible due to a relative concentration within a larger area.

Applying PLUC to Puerto Rico may be a simplified representation of land use dynamics, especially due to the high population density and subsequent fragmented landuse. Land was allocated in a quite scattered manner, especially for agriculture. Allocation occured mostly in agricultural regions, as the model was instructed, but in small plots scattered between pasture land, forest, urbanization. This for one, does not favour large scale commercial cultivation. On the other hand, this may be the exact scale of allocation that represents the island's situation and size. If a progressive agricultural policy is aspired, small(er) farmers will again be stimulated to engage, invest, expand, thus stimulating small-scale farming. Figure 38 below illustrates the Land Use Changes 2013-2030 in the south region of Coamo, Santa Isabel, Salinas. Added land in 2030 is marked with darker colors.

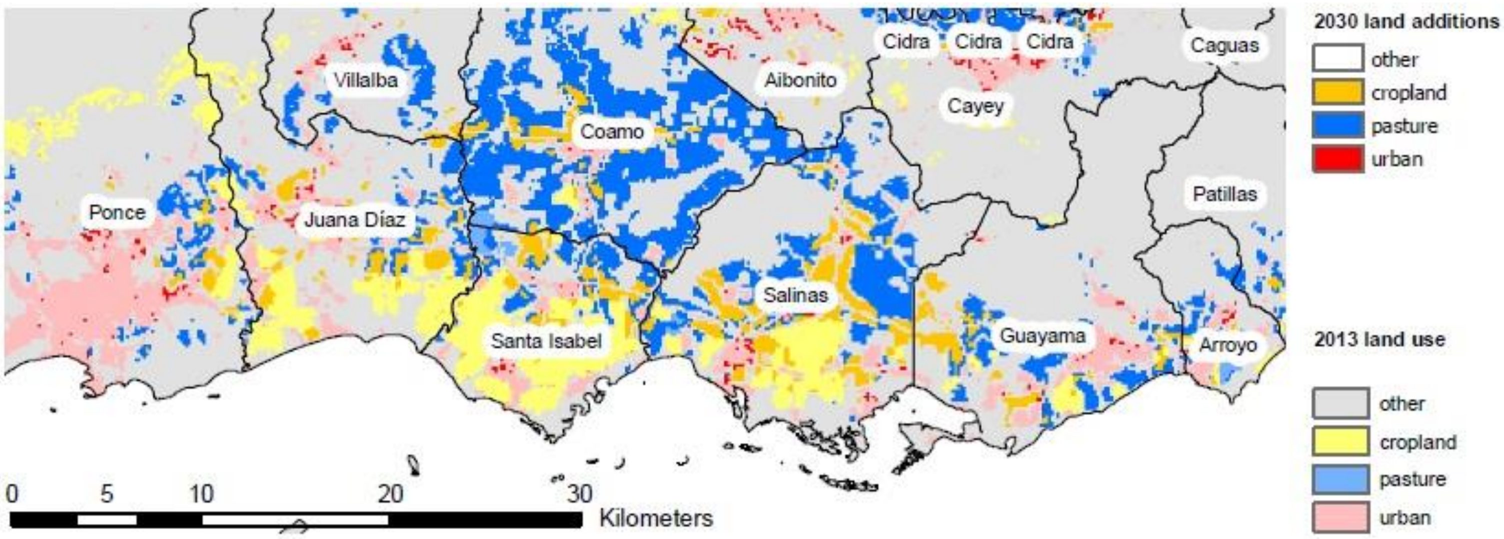


Figure 38. Land allocation in 2030 in the South triangle region of Ponce-Aibonito-Guayama.

Concerning the MOMILP model. First of all, I enjoyed working with the MOMILP model. It was a tough challenge with a lot of data input, linear programming skills, and exiting supply chain prospects. My linear programming skills were not that advanced, and this translated in some challenges –particularly in applying some specific Puerto Rican characteristics- that I could not resolve or that required some very time-consuming manual modifications. Other adjustmenst that were envisioned were just not feasible or manageble. Some of these challenges I will try to elaborate on.

It must be mentioned that the MOMILP model was developed for a specific geographical location –northen Italy- with scenario specific characteristics, especially concerning specific feedstocks and technology configurations. Applying this model to another region with other characteristics proved problematic in some cases. One example is the range of technology capacities. In the original script the available range in plant capacities was much smaller: between 160-460 $MW_{HHV,in}$ for corn, and 220-630 MW for stover, with five small intervals between 20-60 MW. In this thesis a larger range was intended (200-1000 MW; for ec.1GEN up to 1500 MW) with 9 intervals of 100 MW. This was chosen on the one hand to validate which capacity would be most optimal according to the island's specific input parameters. On the other hand this extended range was needed because technology configurations with divergent biomass-to-ethanol conversion factors were applied. Configuration k3=ec.1GEN for example, had a very low factor of 8.12 $ton_{etoh}/MW_{HHV,in}$, and was chosen for its electricity generation advantage. Whereas configuration k7=ec.SSCF had the reverse characteristics with a 25.82 ton_{etoh}/MW conversion factor. The problem arose when this 1^{st} generation technology was optimized together with a 2^{nd} generation technology. The model would choose a very high k3 plant capacity in time period 1, would introduce an SSF/SSCF technology in time period 2 and would let this technology produce nearly all the ethanol. In other words, a large part of the k3 capacity would remain unutilized. Installing a large k3 capacity and utilizing only a fraction of it in later time periods turned out to be the most optimal solution (i.e. most economical) but is of cource not efficient. The defined ethanol demand (20-40%) is also influencing this. This scaling up and down as soon as new plants were added in sussecive time periods occurred between all technology combinations, not only between k3 and others. An intermediate solution was a tiresome manual manipulation of scales, ethanol demand, and technology combinations.

Another particularity concened the decision-making for plant location. In the methodology I argued that the MOMILP model was attractive in that it had the opportunity for a larger range of possibilities for plant location (discussion 'MODA versus MADA decision making characteristic of MOMILP', paragraph 2.2 'modeling with PLUC'). MOMILP would be able to use a large range of 'gridcells' as location for a plant. In reality it very much favoured the speed and accuracy of the model-runs if the window of plant locations were indeed limited to a select few. MOMILP would find lower costs every time a few other grids were turned off. Fortunately some tireless days of model-running provided some experience in handling the model, and resulted in a selected amount of gridcells that were most favourable.

Two other comments concern the data availability and the delimitation of the investigation. Firstly, such a heavy analysis -as this combined PLUC-MOMILP methodology - necessitates very accurate input data from the field. Unfortunately, data availability on feedstocks, production costs, production inputs was very limited or outdated. I encountered only five sources on biomass cultivation, either from the '80-'90 or from experimental plots. My first attempts to work with these data sources, for example, repeatedly allocated ethanol plants in one specific region where the biomass production costs were cheapest, but the data from this region was based on one outdated source. I only later discovered this, but on the scale of this investigation this difference would represent 80 million dollars in production costs. Fortunately I was able to fine-tune the production costs of biomass after some drafts (see appendix 10.3 'Estimating fertilizer inputs 1949-1969') as to enable a relatively trustworthy spatially explicit allocation with the MOMILP model.

A second and last comment concerns the delimitation of the research. Much depend on the assumptions one makes in scetching the system under investigation. In this case some important aspects were the location of the demand centres for ethanol. One can imagine that another arrangement for delivery stations, or fewer stations, could play a role in the allocation of plants. Certainly, transportation of fuel accounts for +/- 2 $/GJ in the overall COE. I decided to point out ten centres for ethanol delivery, but I could have pointed out the three existing harbors where the gasoline is imported to, as well. This would make ethanol blending with gasoline much easier. Other important 'meta-assumptions' as I would call them, are the available land for bioenergy production, wich is again dependent on the scenario development in the PLUC model. Much of these assumptions and variables are better delineated when a region is already advancing on a bioenergy sector, on a serious agricultural sector, or even just in general scenario building. This research may have been too heavy for this location, and a simpler case study would have been better.

On the other hand, this research could be exactly the right starting point for policymakers, farmer cooperatives and other stakeholders alike, in a further consolidation of a sustainable island-state future. Applying two heavy models requiring a significant amount of data input that is not yet conclusive, may seem over the top, complex or not feasible. But the preliminary results of this thesis show the feasibility of utilizing the island's bioenergy potential. The broad scope of this system analysis also touches on other aspects of societal organisation; 'how do we plan our future more sustainably and independently'; 'how do we increase our food and energy independency'; 'can agriculture play a renewed and revalued role in this respect'; 'do we see and take advantage of this opportunity that this renewable energy resource presents us'; and 'how do we mobilize our societal resources for this opportunity'. A large-scale employment of this renewable energy source could provide ample opportunities in fighting unemployment, strengthening existing farmers, developing new businesses, providing a local supply for motor fuels, and containing the effects of fluctuating supplies and costs of fossil fuels on the general economy.

8. Internship documentation

This research included an internship in Puerto Rico at the University of Puerto Rico at Mayaguez from 6th of February to 6th of June 2013. The workstation was at the Department of Agricultural sciences under supervision of Prof. Luis Peréz-Alegria, professor in water and soil management, and waste management. The internship enthailed gathering input data for the PC Raster Land Use Change model (chapter 3 'Model 1: PLUC input and results'), as well as production parameters for the life cycle stage 'Biomass Production' (chapter 4 'Models 2-3: MOMILP input'). Furthermore, linearization of nine technology configurations were prepared in an excel model for the technical parameters of the life cycle stage 'Ethanol production'.

Model preparation for PLUC. Preparing data for the PC Raster Land Use Change (PLUC) model consisted of two parts: map preparation in ArcGIS software and the development of the scenarios for land use change between 2014 and 2030. Map availability for Puerto Rico is very detailed since the US Geological Survey (USGS), and the National Resource Conservation Service centre (NRCS) of US. Department of Agriculture include Puerto Rico in the production of geospatial data. Spatial data that was retrieved from these sources included land use data, soil, infrastructure, nature conservation data, digital population censuses, and waterbodies (PRGAP, 2008; NRCS, 2006; USGA, 2001).

Analysis of these maps included ArgGIS map preparation to the specific PLUC format. Apendix 10.1.b 'Suitability factors' gives an oversight of the required maps for the PLUC model, and include conversion to Boolean, scalar or nominal formats of cities-, water-, road-, NoGo-, digital elevation model-, forest- and land use maps. Population density, cattle density, crop yield and sugarcane yield maps were not available and were produced from the available data. A population density layer was extracted from the category 'urbanization', one of the land use classes in the land use map. A cattle density layer was produced by gathering livestock data per county from the USDA agricultural census of 2007, tabularized in excel and converted to ArcGIS format. Crop and sugarcane yield data was constructed by combining soil maps with maps of water availability, crop occurrence and climatic zone maps.

A special preparation concerned the production of an up to date land use map. A land use map from the PRGAP Analysis project (PRGAP Analysis-USGS, 2008) was used because of its high 15-15 meter resolution, and elaborate list of land use classes (70 in total). Due to its high precission it was able to identify reliable land use classes for forest, urbanization, cropland, pasture, scrubland. Scrupland was categorized in 13 classes, and therefore could be accurately aggregated to either (1) scrubland from recently abandoned crop- or pastureland (last 15 years), (2) mature scrubland/young forest (older than 15 years;aggregated to 'forest'), and (3) scrubland on (semi)arid or rocky soil (aggregated to 'barren'). Only the first category was enabled in the PLUC model as an option for land use change, and together with 'grassland' made available for crop/pasture/urbanization land to expand on.

Another improvement was made for the land use class 'pasture', which was not categorized as such in either sources (USGS, NRCS). Pasture was categorized with grassland as 'grassland/pasture' and included all grassland in Puerto Rico. From the NRCS (2006) source was copied into the PRGAP Analysis land use map the 'pasture' land use class, which enthailed 'cropland for hay and cultivated pasture'. This class was nevertheless 5 times smaller than the documented pasture area by the USDA agricultural sensus. To update this land use class, a preliminary run with the PLUC model was made, as to enable 'pasture' land expand to a realistic size according to the cattle density map mentioned earlier.

The last preparation in ArcGIS concerned preparation of geographic data for the MOMILP model. First, a grid raster of 16-16 km^2 surface was made to aggregate the available grass/scrubland (PLUC results) to suitable destination-locations for MOMILP (52 gridcells in total, see paragraph 2.2.a.iii 'Linking PLUC and MOMILP'). Second, 10 demand centers for ethanol delivery were indicated for 10 gridcells based on the population density of 10 aggregated urbanization centers.

The second part in the preparation of PLUC model inputs concerned the scenario development for agriculture, pasture, Self Sufficiency Ratio, and Urbanization, parameters that influence the future demand for land of other land use classes. For agriculture and livestock were needed sectoral intensities derived from historical trends. This was mostly done thought literature and statistics research. Most of these intermediate results were derived from statistics analysis of USDA agricultural censuses. Intermediate results were enhanced by interviews with professors within the agricultural department of the Mayaguez University. During my stay in Puerto Rico I had conversations with numerous agricultural students in the field of soil sciences, agronomy, livestock sciences. This was a valuable source of information on the status quo of the agricultural sector, that placed in context some general trends that were discovered in the data acquisition.

At the end of februari Prof. Peréz-Alegria and I organized a small conference were the outline of the research, together with the intermediate results of the PLUC model were presented to an audience of students and professors. This also enabled me to formally introduce the topic and objectives to professors whith whom I had consultations later on. Preparation of all the inputs of the PLUC model accounted for 25% of the time of the internship, or around 20 days.

Model preparation for MOMILP. Preparations for the MOMILP model represented three distinctive group of activities: (1) preparation of the MOMILP model itself in the CPLEX solver of GAMS; (2) data collection for the life cycle stage 'Biomass production' concerning yields and biomass production costs; (3) the preparation of technical parameters for ethanol conversion, linearized in nine configuration excel spreadsheets.

Preparation of the MOMILP model in the CPLEX solver of GAMS consisted of translating the North-Italy oriented model to the Puerto Rican case. The most important adjustment was the integration of 2nd generation conversion technologies that would become available only in a later timeperiod. This was successfully done by integrating a binary variable '$tech.av_{k,t}$' or 'technology availability' for each technology k and time period t, where 1 = 'available' and 0 = 'unavailable'.

Another adjustment included the integration of trend parameters, i.e. incorporating future increases of biomass yields, costs, and ethanol conversion efficiencies. These were '$trBY_{i,t}$' or trend in biomass yield for every biomass type i and timeperiod t; $trUPC_{g,t}$ or cost trend for Unit Production Costs of biomass for every grid g and timeperiod t; $trGDPdef_{k,t}$, or a GDP deflator trend for each technology k and timeperiod t; $trTrnsp_{l,t}$ or cost parameter for increases in transportation costs of biomass and ethanol.

Incorporating changing ethanol conversion efficiencies seemed particularly challenging, since the complete set of technical parameters would change if the conversion efficiencies would change. Therefore were build 5 different optimization scripts, each with the technical parameters associated with the ethanol conversion efficiencies of a particular time period (2014-2030 included 5 different time periods of 3 years). Conversion efficiencies would increase linearly from the lower to the upper bound in this timeframe. MOMILP model adjustments represented around 15% of the time of the internship, or around 100 hours of work. This included the adjustments mentioned above, while getting accuainted with the modeling in Linear Programming represented much more preparation in Holland.

A second set of activities represented data gathering for the life cycle stage 'biomass production', or investigating the biomass yield potential, and its accociated production costs and emissions. Data acquisition included both reviewing existing literature on sugarcane and energy cane production, as well as interviews with experts. Sugarcane cultivation data was both found in historical statistics of the USDA agricultural censuses, as well as in literature from experimental stations. This data was discussed and enhanced through interviews with Luis Raul Conty, agronomist at the Department of Agriculture of Puerto Rico, at the office of Economic Studies and Agricultural Planning in Coloso, west Puerto Rico. He concluded some sugarcane experiments between 2002-2006, and investigated a renewed start of a sugarcane industry in collaboration with parties of the rum industry, aiming at a local supply of molasses for rum production. Most of the production costs included data from his consultations, and were especially valuable since it considered a renewed start of the industry, thus including start-up costs such as new machinery, land clearance, new roads, administration. Interviews with Prof. David Sotomayor, professor in soil management, were conducted to investigate soil properties of the different regions in Puerto Rico. With this data and data from literature was then constructed a biomass yield potential-fertilizer analysis for the island.

At last were conducted some interviews on general topics relevant on the island concerning the state of the agricultural sector, prospects for bio-energy development, regional preferences for biomass cultivation, water availability. Water availability in particular was an important subject. Water sources are subject to competition between urbanization, agriculture and consumption, or are heavily polluted. Thus, water consumption as well as pollution risks from fertilizer use are subjects of great importance that need further attention if a bioenergy sector is organized.

The last part of the data preparation for MOMLP concerned the linearization excercise for the 9 technology configurations. Together this part of data preparation for the MOMILP model represented around 60% of the internship, or around two months. Unfortunately no conclusive MOMILP results were ready for presentation before my return to the Netherlands. Nevertheless, a follow-up visit may be realistic to give some presentations to stakeholders in San Juan, Turabo University at the Puerto Rico Energy Center (PREC), and at Mayaguez University.

9. Sources

Alexander, A., Samuels, G., Rios, C., Garcia, M. (1984). *The Production of Energy Cane in Puerto Rico: The Hatillo project*. American Society of Sugarcane Technologists ASSCT, Volume 3, Florida and Louisiana Divisions.

Albanese, L. (2007). Induniv: *A new alternative energy project on the boards for Puerto Rico*. Caribbean Business; Vol. 35 Issue 36, Special section p15.

Allison & Rios (1988). *Experiments with sugarcane and Energy cane in Sta. Isabel*. University of Puerto Rico, Mayaguez.

Blok, K. (2007*). Introduction to Energy Analysis*. Techne Press, Amsterdam, The Netherlands.

College of Agricultural Sciences (CCA, 2001). *Strategic Plan to Foster Sustainable Agriculture in Puerto Rico*. University of Puerto Rico, Mayaguez Campus.

Colón-Guasp, W. (2010). *Cambio de Paradigma, Producción de energía o de alimentos: el caso de Puerto Rico*. La Universidad del Este, Carolina, Puerto Rico.

Colón-Guasp, W., Perez-Alegría, L.R. (2000). *Establishment of an Environmental Scheme to Contribute to the Economic Development of Puerto Rico: Ethanol Production and Use*. Caribbean Food Crops Society, Carolina, Puerto Rico.

Conty Pérez, L. R. (1996). *Estudio Socioeconómico de la produccion de cana de azúcar en Puerto Rico y su comparación con las áreas productoras de cana de Louisiana, Hawaii, Florida, Texas, Colombia y Australia.* Universidad de Puerto Rico Recinto Universitario de Mayaguez.

Conty Pérez, L. R., Santos, F. (2006). *Plan estratégico para desarollar un industria canjera in Hormigueros.* Autoridad de Tierras de Puerto Rico, San Juan.

Capó, B., Samuels, G. (1956). *Research with Sugarcane Fertilizers in Puerto Rico, 1910-1954*. Agricultural Experiment Station, University of Puerto Rico, Rio Piedras, Puerto Rico.

Da Rosa, A. (2005). *Fundamentals of Renewable Energy Processes*. Elsevier. ISBN 978-0-12-088510-7, pp. 501–502.

Del Már López, T., Et al (2001*). Urban Expansion and the Loss of Prime Agricultural Lands in Puerto Rico*, Puerto rico University at Turabo.

Dias, M, et al (2011). *Second generation ethanol in Brazil: Can it compete with electricity production*? CTBE and University of Campinas, Brazil.

Fisher, G., Teixeira, E., Hiznyik, E., Velthuizen, H. van. (2008). *Land use dynamics and sugarcane production*, Chapter 2 of '*Sugarcane Ethanol-Contributions to Climate Change'*. Wageningen Academic Publishers, The Netherlands.

Giarola et al. (2011). *Spatially explicit multi-objective optimization for design and planning of hybrid first and second generation biorefineries*. Universitá di Padova, Italy.

Graham, R.L., English, B.C., Noon, C.E., (2000). *A geographic information system-based modeling system for evaluating the cost of delivered energy crop feedstock*. Biomass and Bioenergy 18 (4), 309–329.

Hamelinck, C., Hooijdonk, G.v., Faaij, A.P.C. (2004). *Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term.* Copernicus Institute for Sustainable Development and Innovation, Utrecht University, The Netherlands.

Junta de Planificacion (2012). *Resumen Grupos Alimenticios Total 2010 Preliminar.* Departamento de Agricultura, External Trade Statistics 2010.

Leite et al. (2011). *The Choice of Biofuels to Mitigate Greenhouse Gas Emissions*. University of Campinas, Brazil.

Le Gal, P. (2009). *Coupled modeling of sugarcane supply planning and logistics as a management t*ool. CIRAD, UMR Innovation, Montpellier, France; University of KwaZulu-Natal, School of Bioresources Engineering and Environmental Hydrology, South Africa.

Malczewski, J. (2006). *GIS-based multi-criteria decision analysis: a survey of the literature*. Department of Geography, University of Western Ontario, London, Ontario, N6A 5C2, Canada.

Misook, K., Donal, F.D. (2011). *Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills.* Society for Industrial Microbiology, Louisiana State University Agricultural Center, USA.

Morais, R., de Souza, B., Leite, J., Soares, L. et all (2009). *Elephant grass genotypes for bioenergy production by direct biomass combustion.* Universidade Federal Rural de Rio de Janeiro, Departamento de Agronomia e Ciencia do Solo, Seropédica, RJ, Brazil.

Nadathur, G., Sen, A. (2009). *Combining Agriculture with Microbial Genomics to make fuels.* Microbe magazine, Volume 4, Number 6, 2009. 1752 N Street, NW, Washington DC, USA.

Nardi, M., Sperry, S., Davis, T. (2007*). Grain Supply Chain Management Using ArcGIS in Argentina*. Clemson University, applied Economics and Statistics, Clemson, SC, US.

Perez-Alegria, LR. (2013). *Personal conversations during Internship at University of Puerto Rico at Mayaguez*.

Parés-Ramos, I., Gould, W., Aide, T. (2008). *Agricultural Abandonment, Suburban Growth, and Forest Expansion in Puerto Rico in 1991 and 2000.* Ecology and Society 13 (20):1.
Available on the world wide web: http://www.ecologyandsociety.org/vol13/iss2/art1/

Regis, M., Seabra J., Walter, A. (2012). *Sugarcane as an Energy Source*. Biomass Conversion and Biorefinery. Brazilian Bioethanol Science and Technology Laboratory, Campinas, SP, Brazil.

Rego, M., Neiva, J., do Rego, A. et all (2010). C*hemical and bromatological characteristics of elephant grass silages containing a mango by-product.* Revista Brasileira de Zootecnia, v.39, n.1, p.81-87, 2010.

Rodriguez (1986). *Sugarcane production experiments in Lajas Valley*. Agricultural Experimental station, Lajas, University of Puerto Rico, Mayaguez, Puerto Rico.

Seabra et al (2010). *A techno-economic evaluation of the effects of centralized cellulosic ethanol and co-products refinery options with sugarcane mill clustering*. NREL, USA and NIPE Campinas, Brazil.

Macedo, I., Seabra, J., Chum, H., Sarto, C. (2011*). Life Cycle Assessment of Brazilian sugarcane products: GHG emissisons and energy use.* UNICAMP and Brazilian Bioethanol Science and Technology Laboratory (CTBE), Campinas, SP, Brazil.

Tergas, L., Urrea, G. (1985). *Effect of Fertilization on the yield and nutritive value of Tropical Forages on an Ultisol in Colombia.* Tropical Pastures Programme, CIAT, Apartado Aéreo 6713, Cali, Colombia.

Tittmann et all (2010). *A spatially explicit techno-economic model of bioenergy and biofuels production in California*. University of California, USA.

USDA (1949-2007). *Cencus of Agriculture, Puerto Rico Island and Municipio Data*. Volume 1, Geographic Area Series. USDA, National Agricultural Statistics Service (consulted censuses: 1949, 1959, 1964, 1969, 1974, 1978, 1982, 1987, 1992, 1998, 2007). Available at: http://www.agcensus.usda.gov/Puerto_Rico .

PRGAP Analysis-USDA (2008). *The Puerto Rico GAP Analysis Project: Land cover, Vertebrate Species Distributions, and Land Stewardship*. International Institute of Tropical Forestry. 1201 Calle Ceiba, Río Piedras, Puerto Rico. http://gapanalysis.usgs.gov/data/

Van der Hilst, F. et all (2010). *Potential, spatial distribution and economic performance of regional biomass chains: The North of the Netherlands as example*. Copernicus Institute for Sustainable Development and Innovation, Utrecht University, The Netherlands.

Van der Hilst, F., Verstegen, J., KarssenbergD., Faaij, A. (2011). *Spatiotemporal land use modeling to assess land availability for energy crops – illustrated for Mozambique*. Copernicus Institute for Sustainable Development, Utrecht University, The Netherlands.

World Wide Web

AEE - Autoridad de Energia Electrica (2008).
http://aceer.uprm.edu/pdfs/pres_aee_rum_feb08.pdf.

Brander, M. et all (2012). *Electricity-specific emission factors for grid electricity*.
www.emissionfactors.com

BRC (2012). Biotechnology Research Consortium: Partners, goals and projects.
http://www.uprm.edu/cima/biotech.pdf

Colucci, J., Pandzardi, P. (2002). Biodiesel for Puerto Rico. Congreso Interamericano de Ingeniería Sanitaria y Ambiental Cancún, México. http://www.bvsde.paho.org/bvsaidis/mexico26/iv-013.pdf

Economic Research Service, USDA (2013). *Fertilizer Price index.*
http://www.ers.usda.gov/data-products/fertilizer-use-and-price.aspx

EIA (2010). *Energy Profile Puerto Rico*.
http://www.eia.gov/state/print.cfm?sid=RQ (2010)

EIA (2012). *Energy prices of USA and Puerto Rico*.
http://www.eia.gov/energyexplained/index.cfm

FAOSTAT (2013). *Country Profile Puerto Rico*.
http://faostat.fao.org/country_profile/Puerto_Rico

FAO database (2013). Crop profile '*Pennisetum purpureum Schumach.*
http://www.fao.org/ag/agp/AGPC/doc/gbase/data/Pf000301.HTM.

Geoexplorers (2012). Consortium for MOAA, GIS, environmental science and business related interests.
http://research.ut.pr/centers&institutes/G.I.S.Research.Center

Indexmundi GDP deflator for Puerto Rico.
http://www.indexmundi.com/facts/puerto-rico/gdp-deflator.

IPCC (2007) GWP100a emission factors for etanol production inputs.
http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter2.pdf

National Resource Conservation Service Center, USDA (2013). Geospatial data on Puerto Rico: DEM, TIGER 2010 transportation map, TIGER 2919 census map.
http://datagateway.nrcs.usda.gov/. Cited 8 october 2012.
http://landcover.usgs.gov/landcoverdata.php (2001).

Newenergyfarms (2012). Pictures from http://www.newenergyfarms.com/site/Energy.html. Cited 3 Augustus 2013.

PREC (2012). Puerto Rico Energy Center.
http://www.suagm.edu/utdoctoral/prec/index.html.

PREPA fuel diversification plan (2011).
http://www.aeepr.com/INVESTORS/DOCS/Oerational%20Profile/Fuel%20Diversification%20Plan2.pdf

Sugartech (2013) HHV of bagasse extraction.
http://www.sugartech.co.za/extraction/bagasseCV/index.php

10. Apendixes

10.1. Model 1 – PLUC description

The PLUC model 1 is applied to model land use changes (LUC) and used to extract the potential available land for bioenergy crop production in Puerto Rico. Land requirements for crop production, livestock, and urbanization are accounted for, thus avoiding competition for land between bioenergy crops on the one hand, and existing land uses on the other hand.

The timeframe for this model run is chosen between 2014-2029 with ArcGIS land use data from 2006 (2006 is the date of the most recent ArcGis land use maps available). Figure 4 below summarizes the PLUC model in a graphic way (from van der Hilst and Verstegen, 2011). The model can be broken down in three main categories;

1. 'Drivers of Demand' that will influence the future output increases from cropland/pasture/urbanization (left-below);
2. scenario characteristics that shape the demand category in two scenarios (middle) and
3. the allocation of land to the different dynamic land uses on yearly basis (right).

The 'Drivers for Demand' influence the demand for animal products (thus pasture) agricultural products (thus cropland) and the demand for urbanization (thus urbanized surface). 'Scenario characteristics' on the other hand, represent management or policies that shape the demand side and is attempted to be captured in two in scenarios; the Business As Usual (BAU) and Progressive Scenario (PS).

These scenarios describe a different development of the same drivers that influence the demand for agricultural land use, pasture land and land for urban expansion. These drivers reflect developments in population, diet and the self sufficiency ratio (SSR). Furthermore, only grass- and shrubland is considered as potential land for bioenergy crops, while forestland is excluded as a source either for fuelwood or land available for dedicated energy crops. The application of ArcGIS and PCRaster enables Puerto Rico to model future land developments in both a geographic and time dimension. This PLUC model therefore is a step forward both in the island's development of future agricultural scenarios, and in modeling land availability for bioenergy protential.

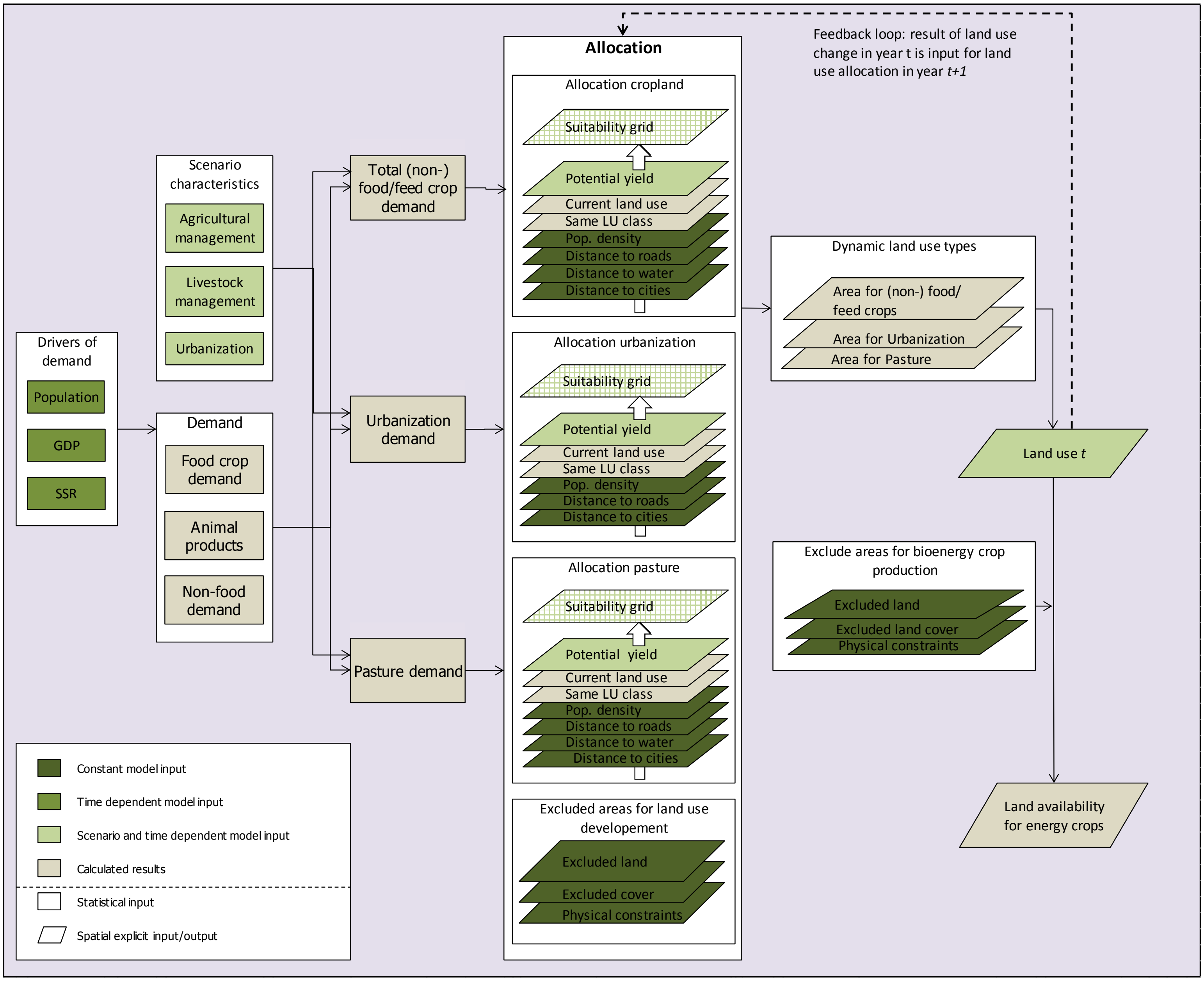


Figure 39: Complete overview of the PLUC model: (left) drivers of demand determine two scenarios;(middle) allocation of land is dependent on suitability factors, and will produce a final suitability raster grid for that land use class for every year; (right) the system has time steps of one year, after the last model run in 2030, the available land can then be assigned to bioenergy crops. Note that the 'Wood Demand' is replaced by 'Urban Demand' (van der Hilst and Verstegen, 2012).

10.1.a. Scenario development

The scenarios that are developed need to structured and quantify in some way or another the often complex and multi-dimentional factors that are of influence on local land use dynamics.

Prime land in Puerto Rico is subject to great competition between urbanization, crop and cattle production - the dynamic land uses- and other land uses such as recreation and forest conservation. This competition is again dependent on developments in population and GDP growth, the island's self sufficiency ratio (the ratio between local production and export/import), agricultural intensification/modernization, the population's diet and policies towards forest conservation, agricultural development and bioenergy sector (van der Hilst & Verstegen, 2012).

Demand for cropland and pasture is dependent on the demand for crops and animal products, which are again dependent on population growth, the local consumption of these products and the diet of the population. The yields of both crops and livestock is dependent on agricultural productivity and intensification. This productivity is dependent on the level of modernization of the sector, i.e. aplication of advanced seeds, mechanization, irrigation, feed-to-meat conversion efficiency and type of cattle industry (extensive v.s intensive cattle ranching).

Puerto rico is considered a reasonable well developed island, with a moderate modernization of the agricultural sector; this entails mechanized crop production, intermediate intensity of cattle production and importation of high nutrision feed, and therefore a moderate feed conversion efficiency. Accesibility to inproved seeds, agro-chemicals, fertilizers, agricultural knowledge, irrigation and machinery is well established. In spite of the level of modernization in some agricultural subsectors, yields (especially crops) are not optimal due to a low local demand, a weak export stronghold, an uncertain agricultural framework and still a large share of smallholders in the total amount of farms. Additionally there exists a large share of land registered as farms that do not produce any crops, thus lowering yield levels even further.

Urbanization is dependent on developments in population and GDP growth. The mid-20th century has seen a tremendous urbanization, although the last past decades this trend was somewhat weekened. This urban sprawl in the '60-'80 was propelled by the shift from an agricultural, to an industrial-service orriented economy. In the '90 this shift was nearly completed, and thus a moderate urbanization will be projected into the future.

Developments in forestry are only dependent on conservation policies, as nowadays deforestation due to the harvest of wood-for-fuel is non-existent on the island. Forestland has tremendously increased in area from the early 20th century untill the late 90s, from less than 6% of total cover before 1940 to around 40% in 2000, and of that 11% between 1990-2000. For both scenarios the current forestry conservation policies are maintained in the future, i.e. the amount of forestland will neither increase nor decrease.

Table 36. Scenario parameters for the PLUC model. This data is used to predict and produce yearly demands for food/feed/meat up to 2030. Available land will be allocated to land uses that supply this demand, while all surplus land will be used for bioenergy crop. The table is a summary of field work that was performed in Puerto Rico. In the next paragraph 3.2. 'Data Input' will elaborate more on the findings in this table.

LUC drivers	**2007 (historical data 1982-2007)**	**BAU scenario**	**Progressive Scenario**
Population	0.97% p.a. change	Same as current, in line with historical trend	Same as current, in line with historical trend
Diet		Same as current	Same as current
SSR	17.65% share of local production in total consumption	Same as current; export same as current	35.30% share (doubling) of local production in total consumption. Export same as current
Farming practices	Commercial farming, widespread access to improved seeds, machinery, pesticides/ herbicides and knowledge	Same as current	Same as current
Agricultural productivity	5.4 ton/ha yields for vegetables & fruits 1.58% p.a. historic increase	1.58% p.a. increase, modest improvements in yield, in line with historical trend	3.5 % p.a. increase in yields
Livestock sector	yields 0.53 ton/ha/yr with 0.72% p.a. increase Moderate feed conversion, mixed systems, good disease control	0.72% p.a. increase, in line with historical trend.	3,5% p.a. increase
Deforestation	Total land coverage 40%, robust forest management.	Same as current	Same as current
Urbanization	Urban yield 37 inhabitants/ha at 0.41% p.a. 'urban yield' growth. 12.9% of area urbanized; 0.94% p.a. population growth	Same as current, in line with historical trend	Same as current, in line with historical trend

Table 4 above summarizes the structure of the two scenarios with the parameters for the used land use change drivers in the PLUC model. These parameters are a summary of the data that is gathered and discussed in paragraph 3.1.d. 'Biomass land availability'.

Population growth, GDP growth and the population's Self Sufficiency Ratio (SSR) will increase the yearly demand for agricultural and livestock products. This will increase the demand for agricultural land, and this increase is again somewhat lessened by increases in agricultural/livestock yields and intensity improvements. Out of all this data, a yearly agricultural and livestock demand is calculated for which the PLUC model will allocate land to.

The allocation of land happens on a yearly basis, i.e. iterations of one year. The PLUC model uses static ArcGIS maps as starting point and generates (yearly) intermediate maps for each specific dynamic land use type. First excluded is land that is not to be used at all for modeling, such as nature conservation areas, roads, buildings, areas with slope higher than 16 degrees, water bodies. Secondly PLUC allocates land on a yearly basis to each dynamic land use type until their yearly demand is met, i.e. available land after a year's allocation will be used as available land in the next year's allocation. This is done until 2029, after which that land can be categorized as potential land for dedicated bioenergy crops.

10.1.b. Suitability factors

From the drivers of land use change mentioned above, a growth rate in demand for food, and meat is identified for the period 2014-2030, and this allows for a yearly demand estimation for these products. This is also done for urbanization. Additionally, yield maps are needed for crops, livestock and urbanization. On a yearly basis, the PLUC model calculates the island's production of crops, meat and urbanization by multiplying the ArcGis cell area times the yield (in tons/area) of that cell.

If the demand for these products -determined in the scenarios- is bigger than the island's production, the model will allocate land (gridcells) to cropland *untill* this yearly demand is satisfied. The PLUC model will first allocate land to cropland, than to urbanization, then to pasture. This is an order of allocation specificly set for this thesis, and the order of 'who gets to choose first' can be set at ones own choise.

The allocation of land to different land use classes occurs as follows. Evaluation of each gridcell to be converted to a specific land use type is governed by a set of suitability factors. Table 2 below summarizes the suitability factors and their criteria. Within the model it is to be specified:

the order of 'who gets to choose first' for each dynamic land use type, i.e. allocate first to crops, then to urbanization, then to pasture;

which suitability factors are to be applied to which dynamic land use types, e.g. a gridcell to be evaluated as future cropland is evaluated by its suitability factors 'distance to water/city/road, yield etc';

the parameters of each suitability factor (see Table 37 below);

the weight of each suitability factor within its land use class (between 0-1, and need to sum up to 1);

Table 37. Applied suitability factors and their accompanying parameters.

nr	description	parameter 1	parameter 2	parameter 3	parameter 4
1	nr of neighbours same class	window length	-	-	-
2	distance to roads	direction	max distance efect	friction	relation type
3	distance to water	direction	max distance efect	friction	relation type
4	distance to cities	direction	max distance efect	friction	relation type
5	yield	friction	-	-	-
6	population density	direction	-	-	-
7	cattle density	direction	-	-	-
8	current land use	suitability current lu	-	-	-

1. window length (in m.) in which neighbours are counted; e.g. 3000 for 3x3 window when cell length is 1000 m
2. direction of the distance function; 1 = positive; -1 = negative
3. max. distance of effect (in m.) of the distance function; e.g. 100,000 for effect op to 100 cells away when cell length is 1000 m.
4. friction in the distance function; used in e^***friction*** * *distance* only for an exponential distance function; use 1 when unknown or when the relation is not exponential
5. type of distacne function; 0 = linear; 1= exponential; 2 = inversely proportional
6. Phthon dictionary with suitability of current land use for placing the new land use; e.g. 3:0.7 means that land use type 3 has a suitability of 0.7 for becomming the land use type that holds this suitability factor (types not specified will have no additional suitability due to factor 9); especially useful to give bandoned areas a higher suitability.

Source: Van der Hilst, Verstegen (2011)

Table 37 above summarizes these specifications in suitability factors for Puerto Rico's dynamic land use classes. If for example a grassland gridcell is to be evaluated for future cropand purposes, the following suitability factors apply:

1. ‘number of neighbouring cells of the same class’, i.e. if more cropland-cells are already clustered there, a grassland-cell will likely be converted to cropland if it is near this cluster;
2. ‘distance to roads/water/cities’, i.e. the closer a grassland-cell is located to roads/water/cities, the higher the probability of conversion;
3. ‘yield’, i.e. areas with higher yields on the yieldmap have higher probability of converison at that location;
4. ‘current land use’, i.e likelyness of the other land uses to be converted to cropland, i.e. grassland{0.8} & pasture{0.5} means that grassland has a 0.8 weight of being converted to cropland, whereas pasture land only a weight of 0.5. the weighting of the suitability factors in Table 38 below indicates the preference of each factor for each land use type. For cropland is most important ‘nr. Of neigbours same class’ (0.2) and ‘yield’ (0.2), followed by ‘distance to water’ (0.15) et cetera.

By applying these suitability factors, a range of cells will be selected that are most suitable to be converted to cropland. They will be converted to cropland up to the point that the total demand for crop products for that year is satisfied by cultivating this land.

The weightings of the suitability factors for each dynamic land use type can be changed at will in the model, and these specific ones in table 6 were assigned such as to best reflect the Puerto Rican land uses. The sum of the weight are 1, so the values are qualitative, and based on priority of importance. Some remarkable weightings will be explained. Distance to roads for example is more important for cropland (0,1) than for pasture (0,05) because cropland needs to be reached by machinery, and cattle is ranged over a larger area by horse. Distance to water is equally important for both land uses. Population density is more important for cropland than for pasture, there the crop products are delivered faster to the city, whereas cattle is again ranged extensively. For pasture, ‘current land use’ was weighted highest, because rangeland was not represented in the land use maps, and thus allocation of land to pasture needed to happen as close to the cattle areas as possible, i.e. accurately allocated according to the cattle density map (see paragraph 3.2.g. ‘Map ccuracy’ and 5.1. ‘PLUC results’ for further discussion on this topic). The category urban land use was newly implemented and replaced the ‘Wood demand’ in the original model. This was because demand for wood for fuel was non-existent in Puerto Rico, whereas a significant urbanization certainly was.

Table 38. Suitablitiy factors and their weighting within each dynamic land use type

	Dynamic land use type		
	Cropland	Pasture	Urban land
Suitability factor		weight	
Nr. of neigbours same class	0.2	0.2	0.2
Distance to roads	0.1	0.05	0.25
Distance to water	0.15	0.15	
Distance to cities	0.1	0.05	0.15
Yield	0.2	0.1	0.05
Population density	0.1	0.05	0.2
Cattle density		0.25	
Current land use	0.15	0.15	0.15
total	1	1	1

Table 39 summarizes the needed ArcGIS maps, and include yield maps, density, infrastructural, elevation and spatial maps.

Table 39. Required ArcGIS maps to run the PLUC model. The maps contain yield, infrastructural, nature conservation and density maps.

Map	contents	data type	unit
noGo.map	all areas that cannot be changed and do NOT have a specific class in the land use map (protected areas, roads, water, nature areas, slope > 16 degrees (cannot be changed = true)	Boolean	-
cities.map	whether or not a cell contains a city (city = true)	Boolean	-
roads.map	whether or not a main road is present in a cell (road = true)	Boolean	-
water.map	whether or not a river or water body is present in a cell (water = true)	Boolean	-
bioNoGo.map	all areas that cannot be used by bioenergy crops in addition to noGo.map (cannot be changed = true)	Boolean	-
dem.map	Digital Elevation Model of the study area (in m)	scalar	-
nullMask.map	value 0 for cells included in the study area and No Data for cells outside the study area	scalar	-
popDensity.map	number of people per area unit	scalar	people/area
cattleDensity.map	number of cows and goats per area unit	scalar	animals/area
scYield.map	fraction of the maximum yield a cell can reach for sugar cane	scalar	-
yield.map	fraction of the maximum yield a cell can reach for food crops and pasture	scalar	-
biomass.map	fraction of the maximum biomass a forest cell produces	scalar	-
landuse.map	land use classes; all dynamic land use types must exhibit at least one cell in the initial land use map	nominal	-

10.2.Model 3 - MOMILP Model description

10.2.a. MOMILP parameters and variables

The model consists of four 'chapters', each nominating and declaring static or variable data. The model starts with the chapter 'Sets', where all the indices are introduced over which formulas will govern. Below are the sets summarized that are used in the model. The most important ones will be explained briefly; time period (t) are 3 year time lapses, over which the model will be run. In each time laps technologies can be build and biomass plots developed. Grid elements g are equally squared surfaces, a total of 52 to cover the island. Each element has an amount of potential arable land, and can be a location for technology development. The following sets are used: life cycle stages (s) are biomass growth, -transport, fuel production, -transport, and emission credits; biomass types are sugarcane, energy cane, and elephant grass; transport modes (l) are only trucks; products (j) are ethanol and power; technology scales (p) are nine in this case (see paragraph 4.2.b. 'Life cycle stage 'Ethanol Production'); objective functions (o) are the financial and environmental ones; cost linearization (c) provide production costs per ton ethanol produced; grid elements g' are different than g, and allow transportation from grid g to grid g'.

<u>Sets</u>

t	time periods (á 3 years)	/ 1*6 /
g	grid squares	/ 1*52 /
s	life cycle stages	/ bg, bt, fp, fd, ec /
i	biomass types	/ sc, ec, elg/
k	production technologies	/ 1*9 /
l	transport modes	/truck /
j	product types	/ ethanol, power /
p	discretisation intervals for plant size linearization	/ 1*9/
o	objective functions	/ eco, env /
c	coefficients for costs linearization	/ slope, intercept /
G	grid squares different than g	/ 1*52/

The next 'chapter' within the MOMILP modelin interface contains the parameters and these are static data. With static data means that the data is prepared in excel spreadsheets or gathered from literature or in the field in Puerto Rico.

The parameter arable land (AD) is derived from the PLUC model, and represents the total available land in 2029 for potential bioenergy production. Here follows an overview of the parameters needed for model 3 MOMILP, and are given in their abbreviation, description, unit, and the subsequent paragraph where it is discussed in more detail.

Parameter	Description	Units	Paragraph
Economic parameters			
CI(p,k)	Capital investments at each interval p & technology k	M€	[2][18] [6]
dfCI(t)	Discount factor for Capital Investment in time t		
dfCF(t)	Discount factor for Cash Flows in time t		
UPC(i,g)	Unit Production Costs per biomass type i per grid g	€/$t_{biomass}$	[3]
UTCb	Unit Transport Costs for biomass	€/$t_{biomass}$	[3]
UTCf(l)	Unit Transport Costs for fuel	€/t_{fuel}	[3]
Technical parameters			
AD(g)	Arable land density km2 arable/ km2 grid surface		[4]
BY(i,g)	Biomass Yield of crop i per grid element g	$ton_{biomass}$/period.km2	[4]
BA(i,g)	Biomass Availability per type ii in grid g (= AD(g) * BY(i,g))	t/time period	[4]
burn(i,k)	Fraction of biomass i burned in every technology k		[2], [6]
y(i,k)	Conversion factor for biomass i and technology k to etoh	tetoh/$t_{biomass}$	[2], [6]
ER(p)	Ethanol production rate for each plant size p	tetoh/year	[2]
w(k)	electricity conversion factor for every technology k	kwh/l_{etoh}	[2], [6]
Emission parameters			
fbg(i,g)	Emission factor for biomass ii **growth** per grid element g,	kg CO2 eq./$t_{biomass}$	[3]
fbpt(i)	E.f for biomass **pre-treatment** for biomass type ii	kg CO2 eq./$t_{biomass}$	GJJ, [3]
fbt(l)	E.f for biomass **transport** per transport mode l	kg CO2 eq./$t_{biomass}$	[3], [5]
ffp(i)	E.f for **fuel production** per biomass type ii	kg CO2 eq./t_{etoh}	[2]
fec(k)	Emission **credits** for each technology k	kg CO2 eq./t_{etoh}	[2]
Contextual parameters			
phi	% over incomes: 15% to calculate FixC(t)		[3], [6]
xi	Interest rate: 10% (to calculate the discount factor)		
Tr	Taxation rate: 41% (PR average)		
$Pcap_{min}$	Minimum ethanol production capacity: 100.000	t_{etoh}/time period	
etperc(t)	ethanol blending percentage at time t		

The third chapter in this model represents the continuous variables. These continuous variables function also as intermediate and end results presentable to stakeholders. Examples are total biomass & ethanol production costs and emissions, cash flows, capital costs, production rates, biomass and fuel flows, transport costs and locations of production sites.

Continuous Variables (intermediate and end results)

BPC(t)	biomass production cost in time t	€/time period
CCF	cumulative cash flow discounted	€
CF(t)	cash flow	€/time period
D(t)	depreciation	€/time period
Db(i,g,t)	biomass demand in region g at time	$t_{biomass}$/time period

[18] [2] Hamelink, 2004. [3] UPR-M. [4] PLUC model. [5] CFCS. [6] PREC.

EPC(t)	ethanol production costs	€/time period
FFC	discounted facilities capital costs	€
FixC(t)	fixed cost	€/time period
Imp(s,t)	impact for life cycle stage s at time t	kg CO2 equiv./time period
Inc(t)	gross earnings	€/time period
Pb(i,g,t)	production rate of biomass i in region g in time t	t/time period
Pf(i,k,g,t)	etoh production rate from i through k g t	t/time period
P(j,k,g,t)	total production rate for product j throught k, g	t t/time period or MWh/time period
PBT(t)	profit before taxes	€/time period
Qb(i,g,l,g',t)	flow rate of biomass i between g and g' via transport l	t/time period
Qf(g,l,g',t)	etoh flow rate between g and g' via transport l	t/time period
TAX(t)	tax amount in time t	€/time period
tcb(t)	transport costs biomass	€/time period
tcf(t)	transport costs fuel	€/time period
TCI(t)	total capital investment at time t	€
TGHG	total GHG impact	kg of CO2 equiv
TI(t)	total impact at time t	kg CO2 equiv./time period
TPot(i,t)	total potential production of biomass i at time t	t/time period
VarC(t)	variable costs at time t	€/time period

10.2.b. MOMILP linear formulation

The calculation methods for costs, emissions and transport are relatively straightforward and are linearized to fit into the model. This section gives a detailed overview of these model's formula's on capital expenditures (CAPEX), operation and maintenance (O&M), i.e. production costs, emissions and other system parameters and variables. The method followed here is from Giarola et al. (2011), but applied to fit the Puerto Rican case.

The first objective function is the economic performance, while the second objective function is the emission performance of the supply chain. The NPV is calculated as:

$$NPV = CCF - FCC \qquad CCF = \sum_t CF_t * dfCF_t \qquad FCC = \sum_t TCI_t * \ dfTCI_t$$

$$dfCF_t = \frac{3 + 3\xi + \xi^2}{3 * (1 + \ \xi)^{2t}} \qquad dfTCI_t = \frac{1}{(1 + \ \xi)^{3(t-1)}}$$

Where
CCF = Discounted Cumulative Cash Costs over the whole time frame of 15 years
FCC = Discounted Facility Capital Costs over a time frame of 15 years
CF = Cash Flows
TCI = Total Capital Investment
dfCF = discount factor for the Cash Flows
dfCI = discount factor for the Capital investment
ξ = Future interest rate, assumed to be constant.

Cash flows **CF** over the time frame are calculated with

$$CF_t = PBT_t + D_t - TAX_t \quad PBT_t = Inc_t - VarC_t - FixC_t \qquad TAX_t \geq Tr * PBT_t$$

$$Inc_t = \sum_j \sum_k \sum_g P^T_{j.k.g.t} * MP_j$$

$$P^T_{ethanol.k.g.t.} \leq \sum_p \lambda_{p.k.g.t.} * 3 * MW_p * ecf_k$$

$$VarC_t = EPC_t + BPC_t + TCb_t + TCf_t$$

$$FixC_t = \emptyset * Inc_t$$

Where

PBT = Profit Before Tax

D = Depreciation charge

Tax = Taxes

Tr = Tax ratio

Inc = Income generated over product marketing

VarC = Variable Costs

$P^T_{j.k.g.t}$ = Total Production of product j from technology k in grid element g and time period t

MP = Market Price for product j

EPC = Ethanol Production Costs

BPC = Biomass Production Costs

TCb = Transport Costs for biomass

TCf = Transport Costs for Fuel

$\lambda_{p.k.g.t.}$ = continuous recursive variable which will assume a non-zero value since the moment an investment decision is taken

MW_p = Nominal production rate of ethanol for each plant size p. The conversion factor 3 accounts for each time period length.

ecf_k = ethanol conversion factor per technology k

TAX_t is derived by applying a 'normal-tax net income' tax rate Tr of 20% to the Profit-Before-Tax or PBT_t, which is a general approximation of the taxation for corporations in Puerto Rico (Praxity, 2012)[19].

Fixed costs that account for fixed variable costs are not calculated here but already accounted for in the estimation of the capital investment costs, in table (Hamelink) represented by the 'installation factor'.

Incomes per time period or Inc_t constitutes of marketing the end products ethanol at 0.80 $/$L_{etoh}$ and electricity at 0.10 $/kWh. Variable costs or $VarC_t$ constitute the different production and transport costs of both biomass and ethanol, or EPC_t, BPC_t, TCb_t and TCp_t and will be discussed in chapter 3 'Data Input' as this data was gathered in Puerto Rico. The production rate of ethanol $P^T_{ethanol.k.g.t.}$ in tons per technology k in grid

[19] Praxity (2012). Business and Taxation Guide to Puerto Rico. Praxity ™ Global Alliance of Independent Firms, San Juan, Puerto Rico.

g and time period t, is calculated by multiplying; with the capacity of the plant MW_p with an ethanol conversion factor ecf_k that represents the amount of ethanol produced per MW of capacity installed; with 3 for the amount of years in each time period; and with $\lambda_{p.k.g.t.}$ that represents a continuous recursive variable and that has a non-zero value since the moment an investment decision is taken. This planning variable $\lambda_{p.k.g.t.}$ will be explained later at the capital linearizations.

The variable costs are constitute the following cost components:

$$VarC_t = EPC_t + BPC_t + TCb_t + TCp_t$$

$$BPC_t = \sum_i \sum_g Pb_{i.g.t} * UPC_{i.g}$$

$$EPC_t = \sum_k c_{k.slope} * \sum_g P^T_{ethanol.k.g.t}$$

$$TCb_t = \sum_{i.l.} UTCb_l * \left(\sum_{g.g'} Qb_{i.g.l.g'.t} * LD_{g.g'} * \mathrm{tau}_{\mathrm{g.l.g'}} \right) + \sum_{i.g.} UTCI^* * Pb_{i.g.t} * LD_{g.g'}$$

$$TCf_t = \sum_{i.l.} UTCf_l * \left(\sum_{g.g'} Qf_{i.t.l.t'.t} * LD_{g.g'} * \mathrm{'tau'}_{\mathrm{g.l.g'}} \right)$$

Where:

$Pb_{i,g,t}$	= production rate of biomass i in region g at time t	t/time period
$UPC_{i,g}$	= unit production costs for biomass type i in grid g	\$/tbiomass
$Y_{k,g,t}$	= 1 if a production facility k is already established in region g at time t, 0 otherwise	
UTCb	= unit transport costs for biomass type i	\$/tbiomass
$Qb_{i,g,l,g',t}$	= flow rate of biomass i between g and g' via transport l	t/time period
$Qf_{g,l,g',t}$	= ethanol flow rate between g and g' via transport l	t/time period
$LD_{g,g'}$	= local delivery distance between grids g and g'	km
$Tau_{g,l,g'}$	= tortuousity factor in each grid	
$UTCb_L$	= unit transport costs biomass per transport type l between grid g and g’	\$
UTCl*	= unit transport costs per transport type l within grid g	\$
$UTCf_L$	= unit transport costs for fuel, per transport type	\$

Biomass production costs are calculated by using a Unit production cost (\$/ton.km) over i and g for the biomass Production rate $Pb_{i,g,t}$ (ton/tp) and summing up for all grids g and biomass types i. Ethanol production costs or EPC_t are derived by applying a production cost factor c_{slope} to the production rate of ethanol $P^T_{ethanol.k.g.t}$.

Transportation costs are divided in two parts; the costs of transport between grids and the costs of transport within the grid. The unit transportation costs UTCb for biomass is multiplied by the *product* of biomass flow rate $Qb_{i.g.l.g'.t}$ and the average length between the centroids of 2 grid cells g and g' or $LD_{g,g'}$ *and* a tortuosity factor 'tau' that represents the curvature of the road between the grids. Tau will be higher in the center of the island that is populated by a mountain belt. The same accounts for transport costs of fuel TCf_L.

Some production constraints are needed. These are

$$P^{T}_{ethanol.k.g.t.} \geq PCap^{min} * Y_{k.g.t.}$$

$$P^{T}_{'power'.k.g.t.} = P^{T}_{'ethanol'.k.g.t.} * \frac{\omega_k}{\rho}$$

$$P^{T}_{ethanol.k.g.t.} = \sum_{i} Pf_{i.k.g.t.}$$

$$Pf_{i.k.g.t.} = P^{T}_{ethanol.k.g.t.} * \beta_{i.k.}$$

$$Db^{T}_{i.g.t.} = \sum_{k} \frac{Pf_{i.k.g.t.}}{\Upsilon_i}$$

Where:

$Pcap_{min}$	= Minimum ethanol production capacity: 100.000	t_{etoh}/time period
$Y_{k,g,t}$	= 1 if a production facility k is already established in region g at time t, 0 otherwise	
ω_k	= electricity conversion factor for every technology k	kwh/l_{etoh}
$\rho_{i,k}$	= ethanol density, 0.7891	kg/l
$Pf_{i,k,g,t}$	= ethanol production rate from i through k g t	t/time period
$\beta_{i.k.}$	= Fraction of ethanol rate from biomass type i and technology k	
$Db^{T}_{i.g.t.}$	= biomass demand in region g at time	$t_{biomass}$/time period
Υ_i	= Conversion factor for biomass i and technology k to etoh	$t_{etoh}/t_{biomass}$

The total production of ethanol $P_{ethanol}$ must be greater than the minimum plant capacity $PCap^{min}$ set at 32400 ton ethanol for each time period. Power produced is derived by multiplying total ethanol production with ω_k electricity conversion factor (kWh/L_{etoh}) divided by the density of ethanol (0.7891 kg/L_{etoh}). At last, specific fuel production $Pf_{i.k.g.t.}$ per biomass type i, per k,g and t is derived by multiplying the fuel production per technology k with the nominal ratio $\beta_{i.k.}$ which represent the ratio of each specific feedstock to the plant.

Biomass demand $Db^{T}_{i.g.t.}$ in grid g is equal to the specific ethanol production rate $Pf_{i.k.g.t.}$ divided by the conversion factor biomass-ethanol (Υ_i) for each specific biomass type i.

The model has to comply with some mass balance constraints, too. These are

$$\sum_{k} P^{T}_{ethanol.k.g.t.} = Df^{T}_{g.t.} + \sum_{l}\sum_{g'}(Qf_{g.l.g'.t.} - Qf_{g'.l.g.t.})$$

$$Pb^{T}_{i.g.t.} = Db^{T}_{i.g.t.} + \sum_{l}\sum_{g'}(Qb_{i.g.l.g'.t.} - Qb_{i.g'.l.g.t.})$$

$$Pb_{i.g.t.} \leq BA_{g.i.}$$

$BA_{g.i.} = BY_{i.g.} * GS_g$

$TPot_{i.t.} = \sum_g (BA_{i.g.} * IBF_g)$; where

$Qf_{g.l.g'.t}$ = Fuel flow rate from grid g through transport mode l to grid g' in time t
$Qf_{g'.l.g.t}$ = Fuel flow rate from grid g' through transport mode l to grid g in time t
$Qb_{i.g.l.g'.t}$ = Biomass flow rate from grid g through transport mode l to grid g' in time t
$Qb_{i.g'.l.g.t}$ = Biomass flow rate from grid g' through transport mode l to grid g in time t
$BA_{g.i}$ = Biomass availability in grid g
$BY_{i.g}$ = Biomass yield in grid g (discussed in previous paragraph)
GS_g = Grid Surface biomass availability in grid g (discussed in previous paragraph)
$TPot_{i.t}$ = Total potential biomass i production in time period t
IBF_g = Binary variable for biomass growth possibility in grid g (given in previous paragraph)

[...] meaning that the specific ethanol production $P^{T}_{ethanol.k.g.t.}$ per technology k in that grid g must be equal to the total ethanol demand $Df^{T}_{g.t.}$ in every grid where the demand is set (10 demand centres indicated), *minus* the fuel brought in from other grids. In other words, if de total fuel demand $Df^{T}_{g.t.}$ is brought to the left-hand side of the equation, logically the fuel demand is the local fuel production $P^{T}_{ethanol.k.g.t.}$ *plus* the ethanol brought into that grid $Qf_{g'.l.g.t.} - Qf_{g.l.g'.t.}$ This also applies for the total biomass production $Db^{T}_{i.g.t.}$ in grid g (if brought to the left-hand side of the equation), that must be equal to de local biomass demand in that grid $Pb^{T}_{i.g.t.}$ *plus* the biomass brought in from other grids $Qb_{i.g'.l.g.t.} - Qb_{i.g.l.g'.t.}$ The biomass demand must be upper strained by the biomass availability $BA_{g.i}$ in every grid.

Total Capital Investments **TCI** are calculated with:

$$TCI_t = \sum_p \sum_k \sum_g \lambda^{plan}_{p.k.g.t.} * CI_{p.k.}$$

$$CI_{p,k} = a_k * ER_p^{\,R_k}$$

$$CI_{p,k} = CI_{p,k,t\_ref} * \frac{M\&S_t}{M\&S_{t\_ref}}$$

$$PC_{p,k} = c_{k,'slope'} * PCap_p + c_{k,'intercept'}$$

Where
$\lambda^{plan}_{p.k.g.t}$ = Continuous planning variable that is only a non-zero value by the time that investment has actually occurred.
CI = Capital investment for technology k and plant size p
ER = Nominal Production Rate of ethanol
CI_{p,k,t_ref} = Base cost for CI
$M\&S_{t\_ref}$ = The Marshall & Swift Equipment Costs in a base time t_ref
$M\&S_t$ = idem, but then for the desired time t ('present time' when that actually is)
PC = Production Costs for scale p and technology k

PCap = Plant Capacity for every scale p.

Total Capital Investments TCI is calculated by multiplying the Capital Investments CI for scale p and technology k with $\lambda_{p.k.g.t}^{plan}$ and summing up all relevant grids g, technologies k and scales p. these parameters will be discussed in paragraph 3.2.i. 'Technology configurations'. $\lambda_{p.k.g.t}^{plan}$ is a continuous planning variable that is only a non-zero value by the time that investment has actually occurred. The Capital Investments $CI_{p,k}$ for each scale p and technology k is calculated by $CI_{p,k} = a_k * ER_p^{R_k}$ where the parameters a_k and ra_k are derived from literature. This formula is again not linear, and is made so by dividing the ER nominal ratio into 6 intervals p.

Giarola et all suggested that the Capital Investments be updated by the Marshall & Swift Equipment Cost index by above formula. Production costs $PC_{p,k}$ for every technology k and scale p are supposed to be a linear function of the plant's Production Capacity PCap over three year time iteration length. The technology specific parameters $c_{k,'slope'}$ and $c_{k,'intercept'}$ are the slope and intercept of the linear equation derived by regressing the production costs specific to each scale p of each technology.

Emissions. The second objective function is the greenhouse gas balance, and is formulated as beneath. The total GHG balance (TGHG) is the sum of all the impacts (TI) of all life cycle stages. The most important knowledge here is that in all life cycle stages are used emission factors for that stage (fbg, fbt, ffd, ffp, fec, i.e. for biomass growth, biomass transport, ethanol production and end-product credits, resp.) These emission factors are discussed in the paragraphs that elaborate on their specific life cycle stage (chapter 3).

$$Obj_{env} = TGHG$$

$$TGHG = \Sigma \mathrm{TI}_{t}$$

<u>Biomass growth:</u>

$$Imp_{bg.t} = \sum_{i} \sum_{g} fbg_{i.g} * Pb_{i.g.t}$$

<u>Transport system:</u>

$$Imp_{bt.t} = \sum_{i.l} fbt_l * \left(\sum_{g.g'} Qb_{i.g.l.g'.t} * LD_{g.g'} * \tau_{g.l.g'} \right) + \sum_{i.g} fbtl^* * Pb_{i.g.t} * LD_{g.g}$$

$$Imp_{ft.t} = \sum_{l} ffd_l * \left(\sum_{g.g'} Qf_{i.g.l.g'.t} * LD_{g.g'} * \tau_{g.l.g'} \right)$$

<u>Fuel Production:</u>

$$Imp_{\mathrm{fp}.t} = \sum_{i} \sum_{g} \sum_{k} ffp_i * Pf_{i.k.g.t}$$

<u>Emission credits:</u>

$$Imp_{ec.t} = -\sum_{k}\sum_{g} fec_k * \mathrm{P}^{\mathrm{T}}_{\mathrm{ethanol.k.g.t.}}$$

Where
TGHG = Total GHG emissions
$Imp_{s,t}$ = (GHG) Impact of life cycle stage g in time t
$Imp_{bg,\ t}$ = Impact of biomass growth in time t
$fbg_{\ i,g}$ = (emission) factor for biomass growth, for biomass type i in grid g
$Pb_{\ i,g,t}$ = Production rate of Biomass, for type i in grid g for time t
$Imp_{fp,t}$ = (GHG) Impact of fuel production, in time t
$ffp_{\ i}$ = (emission) factor for ethanol production of biomass type i
$Pf_{\ i,k,g,t}$ = Production rate of ethanol from biomass type i, technology k, grid g in time t
$Imp_{ec,\ t}$ = Impact of emission credits in time t
fec_k = (emission) factor for emission credits for each specific technology k (summing up credits from ethanol and power exported)
$\mathrm{P}^{\mathrm{T}}_{\mathrm{ethanol.k.g.t.}}$ = Total Production of ethanol at each technology k, g and t

The only emission data that is needed for the models, are thus five emission factors for the five life cycle stages. Carbon penalties due to land use changes from agricultural or grassland to bioenergy crop lands are beyond the scope of this research. Emissions per life cycle stage are caused by the following sources:

Life cycle stage	Emissions
Biomass growth:	Fuel use for machinery, fertilizers, pesticides, insecticides
Biomass transport:	Fuel use of trucks
Ethanol production:	Industrial inputs (dilute acid, dolomite, ammonia) and CO_2 emissions as a by-product from the fermentation process.
Fuel transport:	Fuel use

10.3.Estimating fertilizer inputs 1949-1969

Estimating the fertilizer inputs for the 'target' yields is done by making a yield-fertilization curve, that resembles as realistic as possible the yield-fertilization relationship of a *commercial* cultivation. This curve represents the two climatic zones of the island, distinctive for two categories of sugarcane yield responses to fertilization. Table 40 below summarizes for 12 soils, how much the yields vary due to omission of either N, P or K fertilizers. In the yellow part below this is refined by phasing the fertilizer dosis: for nitrogen this is either 0, 125 or 250 pounds per acre applied, while P and K are equal; for phosphorus the dosis is either 0, 150 or 300 pounds/acre, while N and K remain constant; etc for Potassium (Capó & Samuels, 1956).

The weighted mean of sugarcane tonnage reduction for the 'humid' area is 31%, and 10% reduction for the 'irrigated' area. Sugarcane yields from Allison & Rios' experiments (1988) in Sta. Isabel, near Ponce (region 4; 'irrigated area') were used to plot the yield reductions as a function of fertilizer units applied (kg/ha N). Allison & Rios yielded 147 m.tons/ha millable cane at 134 kg-N/ha, i.e. 100.6 metric tons/ha if ratoon/replant losses are accounted for. The yield-fertilization curve was based only on N fertilization. A similar curve was made for the humid area using Alexander's yield data in Hatillo, region 2.

Yields at the three fertilization inputs (0, 140, 280 kg-N/ha) were also corrected with a factor that represents the translation from the experimental to the commercial nature of cultivation (below-left in yellow table). Take the full omission of nitrogen as example (0 kg-N/ha). Yields would reduce with 31% in the 'humid area', but from a starting point of 100 tons/ha sugarcane, this would mean a harvest of 70 tons/ha at zero nitrogen addition. This is not realistic on the long term, but applying an EtC factor of 0.5 and this becomes 35 tons/ha sugarcane for region 4, which is quite realistic for an average lower bound of sugarcane yield. Furthermore, this EtC factor increases at decreasing fertilizer application. The rationelle: at increasing fertilization (i.e. when fertilization is not a limiting factor) the experimental results can be expected in real life, at least on a short timeperiod. But at decreasing fertilization towards zero, experimental data does not resemble real life, since depletion of the soil occurs only later.

Table 40. Yield responces to fertilizer application in Puerto Rican soils.

		Percentage decrease in cane tonnage due to the omission of … (a)		
Soil	Regions:	N	P	K
	Humid area			
Coloso clay	1-2-3.	29	5	11
Coloso silty clay	1-2-3.	33	2	8
Lares Clay	2	27	2	8
Vega Alta clay loam	1-2.	24	-	-
Vega Baja silty clay	1-2.	36	13	10
Weighted mean		31	8	12
	Irrigated area			
Aguirre clay	4	-	2	0
Altura silt loam	4	1	1	2
Fraternidad clay	3-4.	16	1	11
Coamo clay	4	16	-	-
machete clay loam	4	12	0	2
Mercedita clay	4	-	7	2
Santa Isabel silty clay	4	18	0	4
Weighted mean		10	1	5
Weighted mean of all soils		25	7	9

Exper>commercial (EtC) Correction factor		Fertilizer units applied: (a)	Relative cane yields (%) were fertilizer varied was: N	P	K
0.5	Humid area	0	69	92	88
0.8		1	91	100	97
1.0		2	100	100	100
0.7	Irrigated area	0	90	99	95
0.9		1	97	99	97
1.0		2	100	100	100

(a). Amount of fertilizers used in kg per ha.

(orignal P,K in pounds/acre: 0, 125, 300)
(original N in pounds/acre: 0, 125, 250)

Fertilizer units:	N	P	K	N	P	K	N	P	K
0	0	336	336	280	0	336	280	336	0
1	140	336	336	280	170	336	280	336	170
2	280	336	336	280	336	336	280	336	336

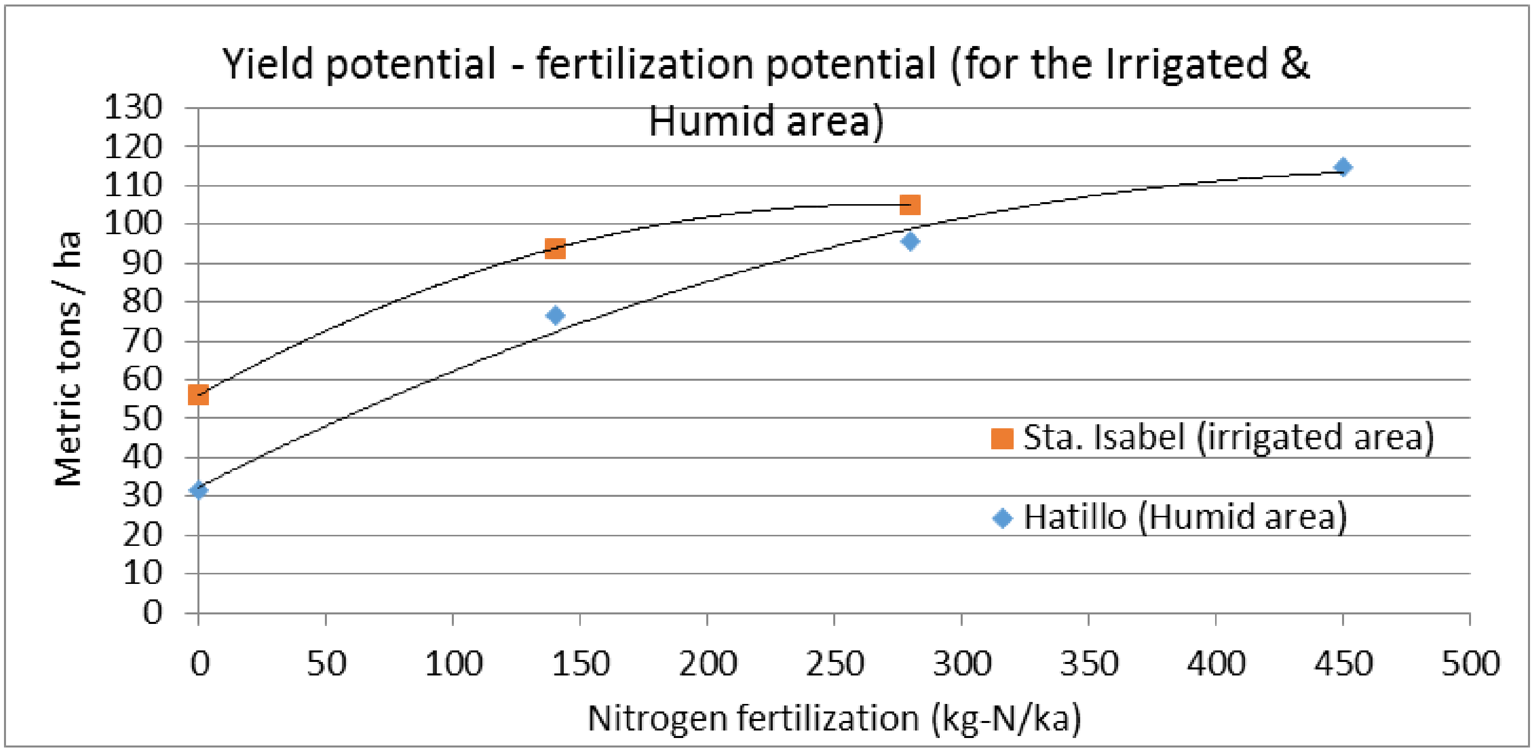


Figure 40. Initial yield-fertilization curve for the humid and irrigated area

Figure 40 illustrates the two curves for Hatillo and Sta. Isabel representing the . To substantiate these curves are needed actual fertilizer inputs at yields documented between 1959-1998. Sugarcane cultivation data on output, required area and fertilizer expenditures were found for 1949, 1959, and 1969 and were documented for all 78 counties. If fertilizer prices ($/m.ton) from these years are found, the fertilizer expenditures ($) can be converted to tonnages (m.tons), and with subsequent sugarcane outputs (m.tons) and area (ha), als converted to intentsities (tons/ha fertilizer with associated tons/ha sugarcane yield). Below are presented the sugarcane cultivation parameters in Table 41 (1949), Table 42 (1959), and Table 43 (1969). Here the fertilizer prices have already been estimated to complete the tabularization. Later is explained where this is derived from.

1949		Sugarcane			Fertilizers			
Region	County	Area planted (ha)	Output (m.tons)	Yields (m.tons/ha)	Expenditure ($-1949)	Input (m.tons)	Intensity (kg/ha)	Average
	Carolina	2,763	199,882	74	282,050	3,318	1,201	
1	Loiza	2,591	183,552	72	300,571	3,536	1,365	1,286
	Veja baja	2,910	196,337	69	319,764	3,762	1,293	
	Arecibo	5,556	366,319	67	512,461	6,029	1,085	
2	Camuj	3,194	198,034	63	320,328	3,769	1,180	1,216
	Manati	2,474	190,391	78	290,497	3,418	1,382	
	Cabo Rojo	4,968	293,183	60	303,960	3,576	720	
3	Lajas	3,044	147,323	49	197,517	2,324	763	977
	San sebastian	4,625	302,224	66	569,672	6,702	1,449	
	Juana Diaz	3,408	288,378	86	258,341	3,039	892	
4	Ponce	4,209	366,705	89	370,851	4,363	1,036	970
	Sta. Isabel	3,252	302,224	94	271,102	3,189	981	
	Humacao	3,197	322,027	**102**	390,529	4,594	1,437	
5	Naguabo	3,605	322,669	**91**	395,681	4,655	1,291	1,391
	yabucoa	3,392	337,814	**101**	416,342	4,898	1,444	

Table 41. 1949 sugarcane cultivation parameters for fertilization intensities.

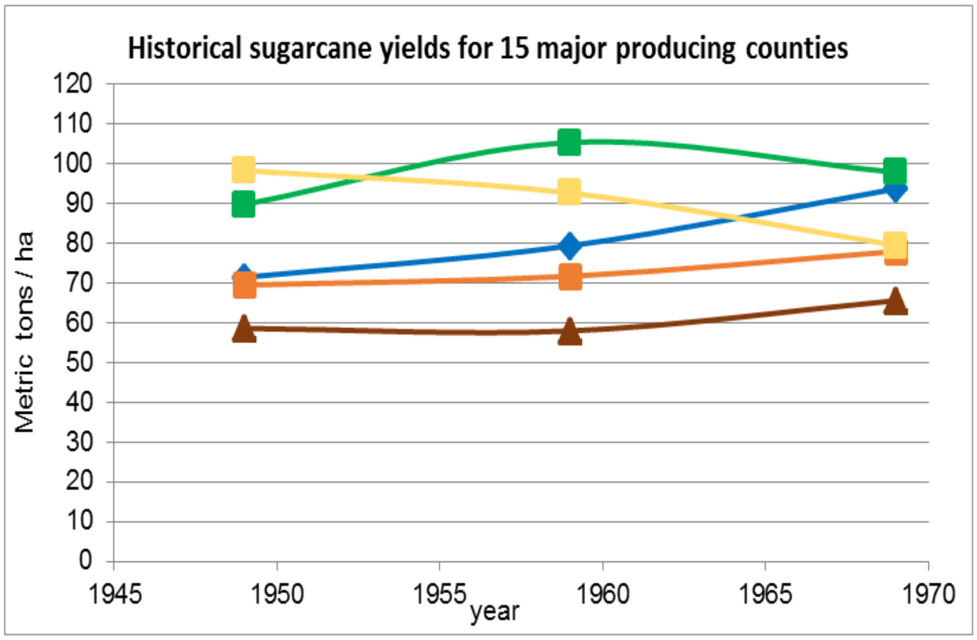
1959		Sugarcane			Fertilizers			
Region	County	Area planted (ha)	Output (m.tons)	Yields (m.tons/ha)	Expenditure ($-1949)	Input (m.tons)	Intensity (kg/ha)	Average
	Carolina	1,936	154,753	80	171,709	2,453	1,267	
1	Loiza	2,014	170,672	85	150,411	2,149	1,067	1,195
	Veja baja	3,004	220,278	73	263,425	3,763	1,253	
	Arecibo	6,578	478,542	73	526,271	7,518	1,143	
2	Camuj	3,418	207,908	61	295,397	4,220	1,235	1,472
	Manati	2,028	165,062	81	289,384	4,134	2,039	
	Cabo Rojo	4,186	265,469	63	303,729	4,339	1,036	
3	Lajas	3,233	187,571	58	205,426	2,935	908	1,105
	San sebastian	5,445	284,964	52	522,878	7,470	1,372	
	Juana Diaz	3,888	398,263	**102**	146,096	2,087	537	
4	Ponce	4,099	384,876	**94**	272,387	3,891	949	666
	Sta. Isabel	3,299	394,338	**120**	118,493	1,693	513	
	Humacao	3,556	388,641	109	389,903	5,570	1,566	
5	Naguabo	3,602	298,545	83	218,696	3,124	867	1,289
	yabucoa	3,028	259,328	86	303,889	4,341	1,433	

Table 42. 1959 sugarcane cultivation parameters for fertilization intensities.

1969		Sugarcane			Fertilizers			
Region	County	Area planted (ha)	Output (m.tons)	Yields (m.tons/ha)	Expenditure ($-1949)	Input (m.tons)	Intensity (kg/ha)	Average
	Carolina	87	9,076	104	81,905	780	8,941	
1	Loiza	487	41,399	85	131,934	1,257	2,583	2,321
	Veja baja	1,076	98,741	92	232,601	2,215	2,059	
	Arecibo	3,683	263,451	72	519,827	4,951	1,344	
2	Camuj	2,077	142,491	69	217,487	2,071	997	1,418
	Manati	986	92,238	94	198,062	1,886	1,914	

	Cabo Rojo	5,326	359,886	68	399,355	3,803	714	
3	Lajas	4,460	267,605	60	425,611	4,053	909	908
	San sebastian	3,798	262,185	69	439,071	4,182	1,101	
	Juana Diaz	1,226	113,699	**93**	292,569	2,786	2,273	
4	Ponce	5,361	479,851	**90**	332,114	3,163	590	1,128
	Sta. Isabel	2,676	297,314	**111**	146,083	1,391	520	
	Humacao	347	29,454	85	71,635	682	1,966	
5	Naguabo	656	48,493	74	217,766	2,074	3,162	2,121
	yabucoa	2,027	161,025	79	263,087	2,506	1,236	

Table 43. 1969 sugarcane cultivation parameters for fertilization intensities.



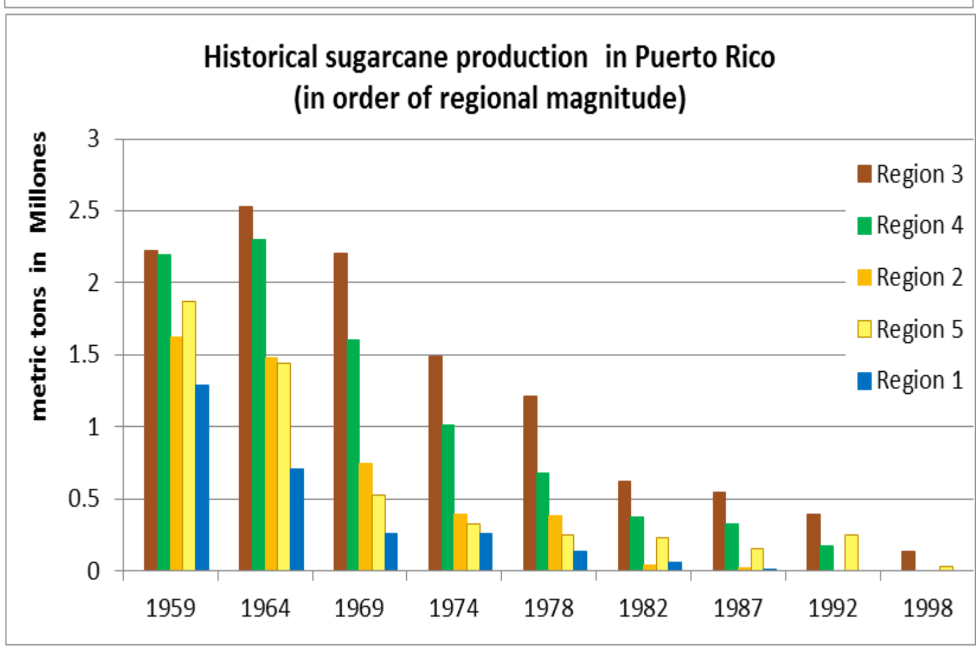


Figure 41. 1949-1969 yields for 18 counties (above) and the island's total cane output for 1959-1998.

For each region is picked the three largest sugarcane cultivating counties. For each county the fertilizer intensity is calculated, and at last averaged again with the other two regional counties, as to represent the fertilization intensity of that region.

Although this method is a reasonable option to extract fertilizer intensities, the results must be seen as a preliminary analysis. For example, the fertilizer expenditures from the USDA censuses were overall expenditures per county, not distincting for specific crop use. It could be assumed to ignore this aggregation, since on an island level (e.g. 1959) the rest agricultural output against sugarcane output was only 6%, or 364,000 tons 'other output' against 10,115,000 tons of sugarcane. For some major sugar regions the output of 'other' crops were even less than 3% of total output. Furthermore, the island's fertilizer consumption according to Capó (1956) accounted for 80% of total fertilizer consumption.

In such a detailed analysis it is very interesting to get aquiainted with success stories of some cultivation regions, to encounter some inefficiencies, and some falacies. Counties Juana Diaz, Ponce and Sta Isabel (region 4) report consistently higher yields (86-120 m.tons/ha) and

2nd highest output of all regions for the timeperiod 1949-1969, and also upto 1998. Figure 41 left&up confirms again the yield distribution between the regions, where region 3 report consistently the lowest yields of +/- 60 tons/ha. This is remarkable since this region is the largest sugarcane producer of the island for the whole timeperiod 1949-1998. This in contrast with region 4, whom is the 2nd largest producer but reports the highest yields.

Yields in region 1 (blue) seem to increase from 70 to 95 m.tons/ha, but this is desceptive. Carolina reports a yield of 104 m.tons/ha in 1969, but this is only from 9,000 m.tons and 87 hectares. Regions 1 and 5 illustrate the sharpest decrease in area & output, and this is illuding the parameters on intensity. Carolina also reports an extremely high fertilizer expenditure for these 87 hectares, which will illude the fertilization intensity even more. These are the effects of sugarcane producing regions in decline: fertilizer rates are desperately increased in the hope to obtain high output, or fertilizer expenditures are reported but not used anymore, or an oversupply of fertilizers being applied to a reduced area. Because of lack of insight into these causalities, results from these regions are excluded in further calculations.

The fertilizer inputs were were derived at a fertilizer price of 100, 70, and 110 $/m.ton (for resp. 1949, 1959, 1969). Application of a USA fertilizer producer price index was not conclusive (see Figure 42 below) since, an index of 40 would give prices between 45-55 $/m.ton of fertilizer, which would again result in fertilizer intensities of 2,000-3,000 m.tons/ha. This has been resolved by averaging the 15 regions to one island-wide intensity, and plotting this with intensity data from Cabó (1956). This is illustrated in Figure 42 below-right: prices are adjusted until the datapoints fit closest the trendline of Cabó's data, but still resemble fertilizer increases in most regions. The result is shown in Figure 43 below.

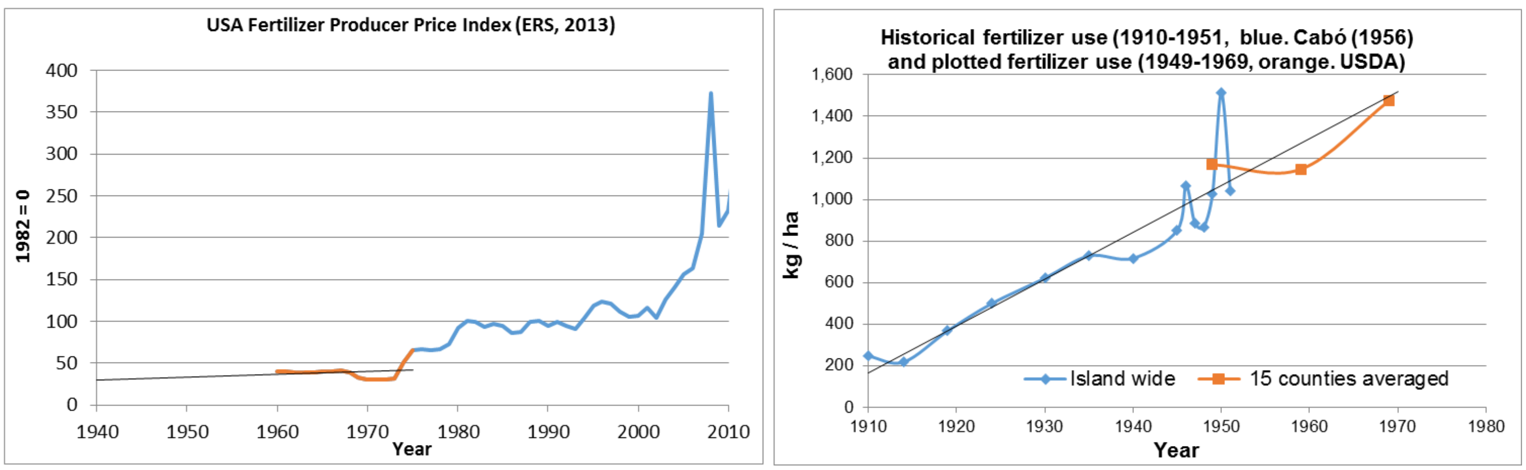


Figure 42. USA Fertilizer Producer Price Index (ERS-USDA, 2013) (left); Historical fertilizer intensity in Puerto Rico (right).

The indicated fertilizer prices of 100, 70, and 110 $/m.ton $fertilizer_{12-4-10}$ result in the fertilizer intensities of 1,000-1,300 kg/ha for regions 1, 2, and 5, while for the regions 3 and 4 these are lower, between 750 and 1,000 kg/ha. These results are plotted in the yield-fertilizer curves in Figure 43-right. Regions 1 and 2 (blue & orange) fit best to the 'humid' area curve, and yield/fertilizer data from region 3 to a lesser extent. The 'irrigated' area (upper curve) was plotted to represent region 4 and 5, since most of their soils correspond to the qualities of that category. Exept for one data point of region 4 (87:105) and one data point of region 5 (142:79), all remaining data points fit closesly to the graph. Both curves are thus valuable tools to estimate input requirements for the 'target' yields as a start for the 2014-2030 optimization in MOMILP.

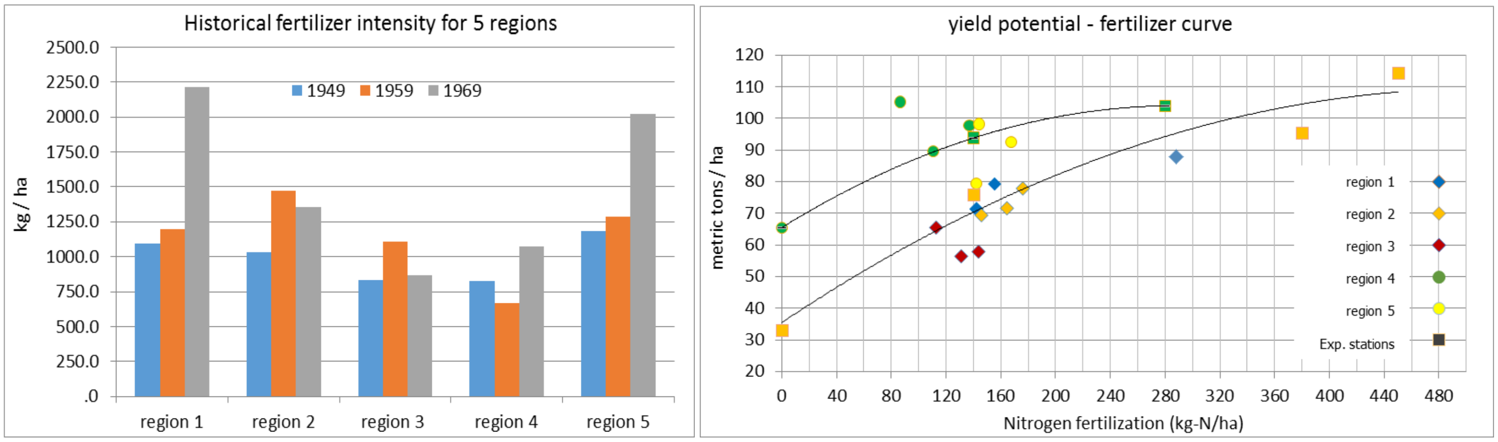


Figure 43. Historical fertilizer intensity for 5 regions (left); and the final yield-fertiziser curve (right).

10.4. Future trends in prices and GDP

Biomass production costs have been scaled with the relevant indexes for fertilizer (ERS-USDA 2013), steel, diesel (AEE, 2008) and electricity (PREPA, 2011) prices and an overall GDP deflator (indexmundi GDP deflator, 2013) for other expenses such as labor, administration and land lease. These trends in prices and the GDP deflator are illustrated in Figure 44. The fertilizer price index is from the US and constitutes the producer price index for the category 'all fertilizers'. The future price developments for diesel (used for cultivation machinery) is derived from historical cost trends of distillate and residual fuel oil. Both graphs follow a similar curve, since they are products of the same industrial process and relatively similar in type. These curves are for distillate/residual fuel that is being used in electricity production on the island. Price developments for normal vehicle diesel was not found. Therefore it is assumed that normal pump diesel will follow the same trend as the distillate/residual trends. The annual growth percentage is an average of these two curves, relative to their volume of consumption.

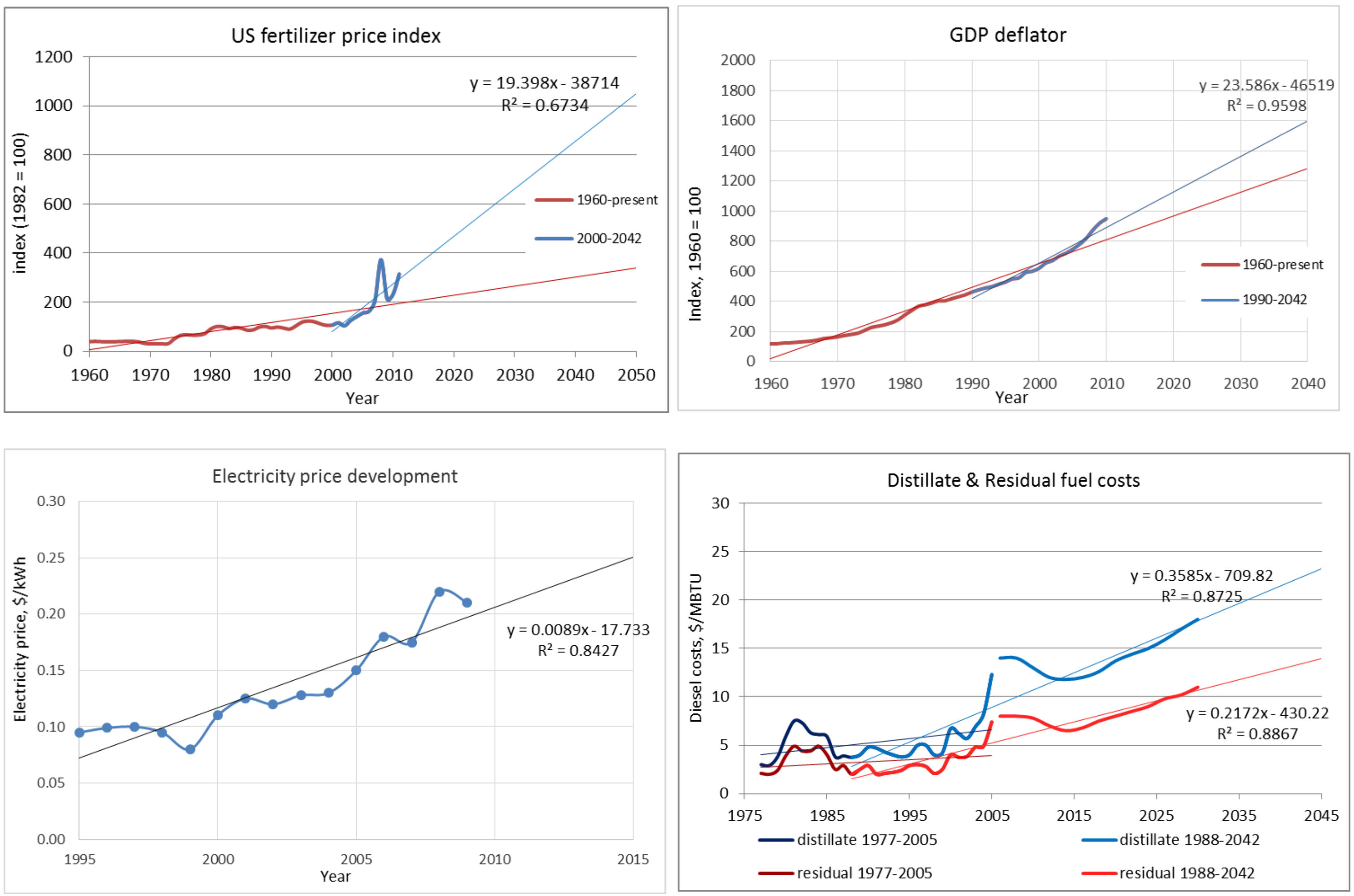


Figure 44. Future developments in fertilizer, electricity and diesel prices, and a GDP deflator.