EVALUATING THE EFFECTS OF UNCERTAINTY ON PROJECTIONS OF GREENHOUSE GAS EMISSIONS:

A BIOFUEL CASE STUDY IN BRAZIL

Student:

Renan Maron Barroso

Dissertation supervised by:

Dr. Judith Verstegen

Institute for Geoinformatics, University of Münster

Co-supervised by:

Dr. Floor van der Hilst

Copernicus Institute, Utrecht University

Dr. Carlos Granell Canut

Institute of New Imaging Technologies, Jaime I University

February 2019

DECLARATION

I, Renan Maron Barroso, aware of my responsibilities of the penal law, declare that the thesis entitled ***Evaluating the effects of uncertainty on projections of greenhouse gas emissions: a biofuel case* study in Brazil** is the result of my own research to obtain the degree of Master of Science in Geospatial Technologies.

I confirm that this document is not copied from any other person's work (published or unpublished) and has not previously submitted for assessment anywhere.

I understand that literal citing without using quotation marks, citing the contexts of any work without making the references, or using the thoughts of someone whose work was published as of my own thoughts are counted as plagiarism.

I declare that I understood the concept of plagiarism and that my research must be accurately referenced. Therefore, I declare that all the sources used by me are adequately cited and listed.

I acknowledge that my thesis will be rejected in case of plagiarism.

Münster, 25th February 2019

____

Renan Maron Barroso

EVALUATING THE EFFECTS OF UNCERTAINTY ON PROJECTIONS OF GREENHOUSE GAS EMISSIONS: A BIOFUEL CASE STUDY IN BRAZIL

ABSTRACT

The use of projections of greenhouse gas emissions (GHG) estimates are fundamental to design appropriate policies to combat climate change, but the inherent complex nature of the climate system results in projections with a significant degree of uncertainty. An important source of uncertainty in GHG emissions estimates refers to land use changes (LUC) due to the complexity of the land system. As the land domain plays a relevant role in climate change mitigation, understanding the effects of uncertainty on projections of LUC-related GHG emissions estimates is crucial to better support the process of decision making. Based on a case study conducted by van der Hilst *et al.* (2018), this thesis evaluates the effects of uncertainty on the projections of LUC-related GHG emissions in Brazil towards 2030, given an expected increase in the global biofuel demand and distinct scenarios of LUC mitigation measures. With the use of Monte Carlo simulation technique, we developed a spatially explicit, stochastic model in Python programming language to perform the uncertainty analysis. As uncertainty can be derived from many sources, we focused on adding uncertainty in the model input data to assess its effects on the LUC-related GHG emissions estimates resulting from an increase in the global biofuel demand. As van der Hilst *et al.* (2018) performed an analysis of the same case study, but without uncertainty analysis, this thesis compares the stochastic results of the deterministic results. The comparison of the results obtained between the deterministic and the stochastic approach provides valuable insights about the effects of uncertainty in the final estimates of emissions. We run the model for six distinct LUC scenarios and computed the LUC-related GHG emission estimates given the changes in soil organic carbon (SOC) and biomass stocks, resulting in estimates with an associated uncertainty. We performed a statistical test to verify the existence of significant differences in the emission estimates between the scenarios and we run a sensitivity analysis to evaluate the contribution of the model components in the overall uncertainty of the emission estimates. The outcomes allows saying that adding uncertainty in the input data results in estimates with great uncertainty, specially in the emissions resulting from the changes in SOC stocks. The emission estimates obtained in this thesis have similar values when comparing to results of the deterministic approach of van der Hilst *et al.* (2018). The statistical test allows saying that the LUC-related GHG emission estimates resulting from an additional ethanol demand are significantly different between all scenarios, therefore the emission estimates could be used to support decision making e.g. to define or prioritize the implementation of a new LUC mitigation measure in Brazil.

Keywords: *greenhouse gas emissions, land use changes, land use change projections, mitigation measures, Brazil, carbon stocks, biofuel, uncertainty, stochastic modelling, Monte Carlo simulation*

ACKNOWLEDGMENTS

It is with great joy that I express my deepest gratitude to my supervisor Prof. Dr. Judith Verstegen for continuous encouragement, comprehension, critical thinking and precise advice. Her support was essential to keep my motivation up, especially in the moments of difficulty. This thesis could not have been realized without her assistance, and I will always be thankful for that. I also address my sincere gratitude to the co-supervisor Dr. Floor van der Hilst for allowing me to use her work as the basis of this thesis and to the co-supervisor Dr. Carlos Granell Canut for the positive feedbacks and motivational words.

An immense thanks to all my Erasmus Mundus classmates. Because of them, this Master program has become unique and unforgettable. I thanks all the professors I had during this master program and all the staff of NOVA IMS and IFGI. My special acknowledgment is addressed to Prof. Dr. Marcos Painho who accepted me to join this master and whose classes were inspiring, to Prof. Max Pfeiffer for all his will to help during this last semester, and to Mr. Christoph Brox and Mr. Karsten Höwelhans for their exceptional administrative support.

A profound thanks to my friends Cris, Matheus for allowing me to share this journey throughout this master program with you. Also, a special thanks to Prof. Jonas Masetti and Prof. Edilberto Moura, who influenced me to be here in this masters.

Finally, I want to express my eternal love and gratitude to my parents Nelson and Juçara, and my sister Camila. Thank you for your unconditional love and for being so supportive in every step of my life. You make me stronger.

ACRONYMS

AFOLU	Agriculture, Forestry and Other Land Use
AGB	Above ground biomass
BGB	Below ground biomass
CGE	Computable General Equilibrium model
CO_2	Carbon dioxide
FAO	Food and Agriculture Organization
GHG	Greenhouse gas
GIS	Geographic Information System
ICONE	Instituto de Estudos do Comércio e Negociações Internacionais
IPCC	Intergovernmental Panel on Climate Change
LUC	Land use change
OECD	Organisation for Economic Co-operation and Development
PDF	Probability density function
PLUC	PCRaster land use change model
SOC	Soil organic carbon

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

One of the major challenges faced today by global society is the climate change resulting from anthropogenic activities. A significant effort has been made to establish international agreements and strengthen global policies to cope with climate issues, with the Paris Agreement being the current long-term vehicle addressing mitigation goals worldwide. Most of the actions determined in this pact aim to reduce the build-up of greenhouse gases (GHG) in the atmosphere.

Global actions have historically concentrated on the reduction of fossil fuels used in the energy sector to decrease GHG emissions (Bryngelsson, 2015). Such focus has mainly occurred to mitigate carbon dioxide (CO_2) emitted from fossil fuel combustion and industrial processes, which represented 78 % of the increase of total GHG emissions from 1970 to 2010 (Smith *et al.*, 2014).

New evidence calls to action mitigation efforts in the land domain mainly because land use changes (LUC) lead to both GHG emissions and removals. For example, converting forests to crop fields releases CO_2 in the atmosphere due to the removal of biomass and soil, while afforestation increases carbon stocks thus contributing to carbon sequestration. While changes in the land account for about 9-11% of total anthropogenic emissions, LUC-related mitigation actions can contribute from 20 to 60% of the total cumulative emissions abatement up to 2030 (IPCC, 2014). The actions are mostly related to the promotion of carbon sequestration, conservation of carbon pools, and replacement of fossil fuels by biological products (Smith *et al.*, 2014).

Although the importance of the land system on GHG emissions is today recognized by science, there is no consensus in the scientific community about the amount and rate at which CO_2 flux occurs between the land and the atmosphere (Ross *et al.*, 2016). Additionally, many factors related to land use dynamics contribute to this lack of consensus. The land domain is a complex system in which LUC are influenced by an extensive range of socio-economic and environmental drivers interacting through space and time (van der Hilst et al., 2018). Such complexity might hinder any prediction in this domain, resulting in projections of LUC-related GHG emission estimates with a significant degree of uncertainty.

As policymakers consider the outcomes of projections for decision making, e.g. to design appropriate policies for climate change mitigation, identifying sources of uncertainty and understanding its effects on GHG emissions estimates is essential. Compared to estimates with no uncertainty analysis, quantifying uncertainty in scenario projections allows a more realistic interpretation of estimates, and the results are more justifiable from a scientific perspective (Puig, 2015). Therefore, ignoring uncertainty hinders the evaluation of possible ranges of GHG emissions estimates which might lead to wrong decisions with regards to the development of new policies.

A sound manner to cope with uncertainty in projections of LUC-related GHG emissions estimates is with models. Although this modelling approach can be used for uncertainty analyses, Warner *et al.* (2014) revealed a gamut of studies in which models have neglected uncertainty in the estimates, i.e., the models were set through a deterministic approach.

Deterministic models are built in such a way that they do not account for uncertainty analysis. This approach has a limitation when modelling complex systems because the nature of the drivers lying behind the system is intrinsically heterogeneous, and this is not considered in the model. On the other hand, stochastic models acknowledge for heterogeneity thus allowing the analysis of the inherent uncertainty of the system being modelled (Renard, Alcolea and Ginsbourger, 2013).

Many researchers declare that uncertainty in projections of LUC-related GHG emissions estimates must be investigated more rigorously to better support decision making (e.g. Wicke *et al.*, 2012; Warner *et al.*, 2014; Verstegen *et al.*, 2016). Motivated by this claim, in this thesis we intend to contribute with additional research in this domain. Specifically, we produce stochastic results of GHG emissions estimates from a case study conducted by van der Hilst *et al.* (2018, hereinafter referred to as reference study), then we compare with their deterministic results.

The referenced study developed a modelling framework consisting of a macro-economic model, a spatially explicit LUC model, and a GIS-based carbon module. By running this framework deterministically, they projected LUC-related GHG emissions estimates in Brazil up to 2030, taking into account distinct scenarios of LUC mitigation measures in Brazil combined with an increase in global biofuel demand.

Their approach of van der Hilst *et al.* (2018) was the first in integrating macro-economic drivers, spatially explicit socio-economic and biophysical drivers together with the spatial heterogeneity in carbon stocks to estimate the GHG emissions. This study aims to keep on with their innovation by adding uncertainty information in a component of the framework that does not support stochastic runs, namely the GIS-based carbon model.

The comparison of the results obtained between the deterministic and the stochastic approach can provide valuable insights about the effects of uncertainty in the final estimates of emissions. The reference study has shown that mitigation measures could reduce LUC-related GHG emissions derived from the increase in ethanol production in Brazil up to 2030.

the GIS-based model of the reference study is replaced by a stochastic model implemented in Python programming language (Python Software Foundation, 2014) to account for uncertaint. The developed model is built to perform a Monte Carlo simulation, which is a conventional technique to deal with uncertainty analysis related to LUC and GHG emissions (e.g. Ogle *et al.*, 2003; Kim and Sohngen, 2009; Verstegen *et al.*, 2012; Mustafa *et al.*, 2018). Since uncertainty can be derived from many sources that are both extrinsic and intrinsic to models (Deser *et al.*, 2012), hereto we choose to focus on the uncertainty related to the model inputs.

Both in the reference study and hereto, the LUC-related GHG emissions are estimated by spatially explicit calculations given the changes in soil organic carbon (SOC) and biomass stocks resulting from LUC. The computation of carbon stocks accounts for the spatial heterogeneity in land use, soil and climate conditions. With the use of the spatially explicit approach, the uncertainty can be quantified and geographically allocated (Prestele *et al.*, 2016).

The case study of this research is related to an expected increase in biofuel demand worldwide. This is an important issue because the production of biofuels is one of the climate change mitigation actions that has been extensively promoted in the last decades, as they have been considered an essential alternative to replace fossil fuels and reduce GHG emissions (Chum *et al.*, 2011; Smith *et al.*, 2014).

The increase in biofuels demand leads to additional pressure in the land domain, as more area of land for planted biomass is required. The allocation of new crops for bioenergy production results in direct and indirect LUC that might cancel out the climatic benefits of replacing fossil fuels by biofuels (Fargione *et al.*, 2008; Searchinger *et al.*, 2008). The fact is that the extent to what LUC changes induced by biofuel production affect the GHG emissions is still unclear by researchers.

In this context, Brazil is a subject of much study because its production of ethanol from sugar cane places the country as the second largest ethanol producer in the world, with production expected to increase substantially (Macedo, Seabra and Silva, 2008; FIESP and ICONE, 2012). Besides, LUC dynamics are significantly complex in Brazil. Its land heterogeneity, geographical extension, favourable climate conditions, the richness of natural resources, together with other socioeconomic drivers stimulate land use competition. This complexity contributes to mask the influence of LUC-related GHG emission estimates derived from biofuel production. Hence, performing uncertainty analyses is essential to enhance the understandings of the LUC dynamics resulting from biomass feedstock production and their influence on projections of GHG emissions estimates in Brazil.

To sum up, in this thesis we add uncertainty in the input data of a stochastic model to assess the effects of uncertainty on the projections of LUC-related GHG emissions in Brazil towards 2030, given an increase in the global biofuel demand and distinct scenarios of LUC mitigation measures.

1.1 AIM AND RESEARCH QUESTIONS

This research aims to evaluate the effects of uncertainty in the input data of a stochastic, spatially explicit model developed to calculate LUC-related GHG emissions derived from scenarios of increased biofuel production and LUC mitigation measures in Brazil towards 2030.

1.2 RESEARCH QUESTIONS

a) What are the input data uncertainties of the model developed to calculate LUC-related GHG emissions?
b) What is the impact of uncertainty in the input data on LUC-related GHG emission estimates derived from scenarios of increased biofuel production and LUC mitigation measures in Brazil towards 2030?

1.3 THESIS STRUCTURE

The document is structured in four chapters, starting with the introduction. Chapter two presents the methodology and includes the description of the carbon model, input data, scenario approach, runs of the model, expression of uncertainty and uncertainty analysis. The third chapter includes the results and discussion. The conclusion is presented in the last chapter.

2 METHODS

2.1 OVERVIEW

Given an initial land system state and a projected scenario of land use dynamics, the LUC-related GHG emissions estimates caused by an increase in global biofuel demand in Brazil towards 2030 are performed using a spatially explicit model. Taking into account the spatial heterogeneity in land use, soil and climate conditions together with uncertainty information in the input data, the model calculates the emissions based on the changes in carbon stocks in the time frame 2012-2030.

The model makes use of two types of input data: spatial data, represented by land use, climate and soil data; and IPCC data. The former is the input data in which uncertainty information is added. The IPCC data refers to parameters used to calculate SOC and biomass carbon stocks. The parameter values are extracted from the IPCC Guidelines for National Greenhouse Gas Inventories from the Intergovernmental Panel on Climate Change (IPCC, 2006). The inclusion of uncertainty herein allows for stochastic runs of the model with the use of the Monte Carlo simulation technique.

Using a scenario approach provided by van der Hilst *et al.* (2018), six distinct scenarios of LUC mitigation measures are analyzed, plus a reference scenario with no mitigation strategies. The model is set to run each scenario twice: with and without an increment in the global demand in the biofuel production. By doing that, it is possible to investigate the effects of an increase in the biofuel demand in the GHG emission estimates in Brazil when different LUC mitigation measures are taken into account.

Also, the inclusion of uncertainty in the model input data allows assessing the effects of uncertainty on the LUC-related GHG emissions estimates derived from such an increase in ethanol demand. As the reference study provided deterministic results of the case study used hereto, in the uncertainty analysis we compare their model outputs with the stochastic outputs.

2.2 MODELLING FRAMEWORK

In this research, we adapt the framework of the reference study (Figure 1). The original framework integrates a macro-economic model, a spatially explicit LUC model, and a GIS-based module. The former is the component that computes the LUC-related GHG emissions given the changes in carbon stocks, but it is designed to run deterministically. Therefore, hereto we replace the GIS-based module by a stochastic model to account for uncertainty (from now on, referred to as carbon model).

It is important to mention that the macro-economic model and the spatially explicit LUC model are not re-run hereto. Even though, understanding the process in which the emissions are estimated is important because both the reference study and this research use the same case study. Also, the outputs of the LUC model are used as input in the carbon model.

The process of estimating the LUC-related GHG emissions estimates in the reference study started with the simulation of distinct scenarios of demand and supply of commodities for Brazil in the time frame 2012-2030. This was done by running the macro-economic model MAGNET, a global Computable General Equilibrium model (Woltjer *et al.*, 2014). Based on local projections of LUC mitigation measures and expected global developments (e.g. gross domestic product, population growth, and agricultural demands), the MAGNET output provided information regarding the amount of land required in 2030 to meet the demand for crop and livestock production, including bioethanol production from crops.

Next, for each simulated scenario, van der Hilst *et al.* (2018) used PLUC model (PCRaster Land Use Change Model; Verstegen *et al.*, 2012) to allocate the land use requirements of MAGNET spatiotemporally. The allocation process is based on the spatial variability of the suitability for each land use type. The results of PLUC are the LUC dynamics per scenario, i.e. land use spatial data per scenario projected for Brazil in 2030. Lastly, they developed the spatially explicit GIS-based module to quantify the LUC-related GHG emissions, given the changes in carbon stocks in Brazil.

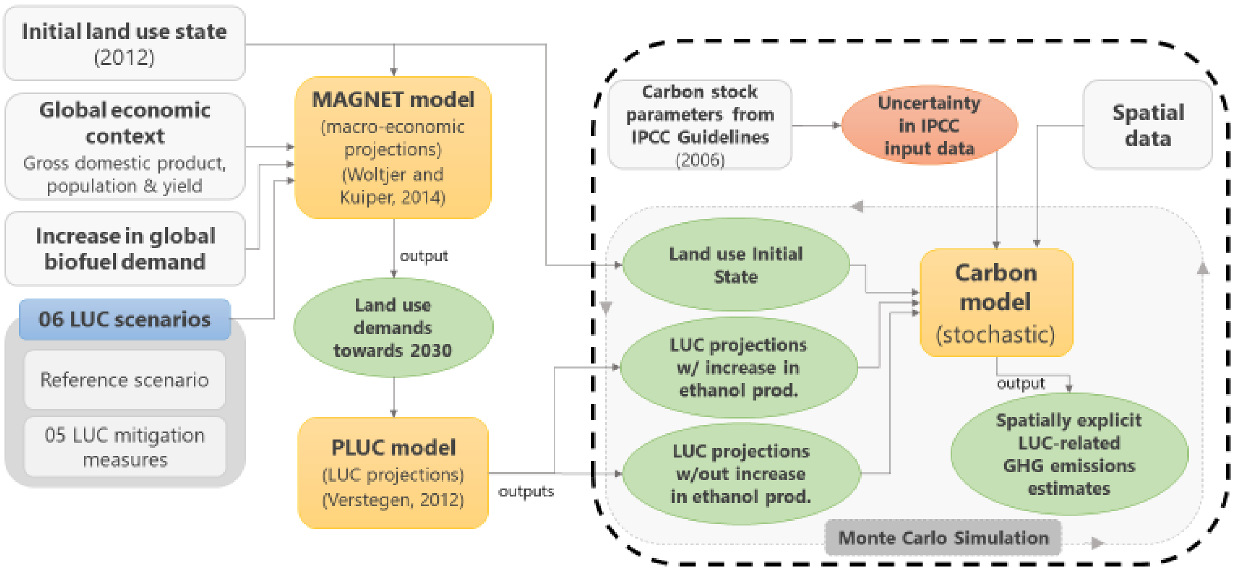

Figure 1 – Framework of van der Hilst *et al.* (2018) adapted for this research. The traced rectangle with sharp corners refers to the work developed hereto, i.e., the part in which the original framework is adapted. The steps outside the rectangle are not performed in this research, namely the runs of MAGNET and PLUC models.

Figure 1 illustrates the adapted framework of van der Hilst *et al.* (2018). The traced rectangle with sharp corners represents the component modified in this research. It is shown that the carbon module developed hereto uses three sources of input data: spatial data; the PLUC outputs and IPCC input data in which uncertainty information is added. Also, Monte Carlo simulation is added to provide a stochastic approach to determine the LUC-related GHG emission estimates given the changes in carbon stocks.

More detailed information of the carbon model, including the process of computing the carbon stocks, is described in the next section and Figure 2.

2.3 THE CARBON MODEL

The carbon model is based on the stock difference approach in line with the Tier 1 method for estimating emissions of the IPCC Guidelines for National Greenhouse Gas Inventories (IPCC, 2006). The LUC-related GHG emissions are spatially explicitly calculated in terms of carbon stock changes in biomass and SOC, given two points in time. The model initially computes the total carbon stocks for each point in time, and the results are subtracted to obtain the changes. Given the uncertainty in the input data and by running the model in Monte Carlo simulation, the final output consists of GHG emissions estimates with uncertainty ranges.

Figure 2 presents a scheme of the process to calculate carbon stocks. It is shown two types of input data necessary to run the model: spatial data and IPCC data. The spatial data consists of climate, soil and land use data. The IPCC data represent a set of parameters necessary to calculate biomass carbon stocks and SOC. Their values and the calculation method are given by the Tier 1 method of estimating GHG emissions of the IPCC guidelines (IPCC, 2006).

The calculation of biomass stocks (BCS) involves four parameters. BCS is calculated by the sum of above-ground biomass (AGB) and below-ground biomass (BGB) in terms of dry matter multiplied by a carbon fraction (CF) (Equation 1). A root-to-shoot ratio (r) between AGB and BGB is used to calculate BGB (Equation 2). The CF parameter is then used to convert the dry matter of AGB and BGB to biomass carbon stocks. IPCC provide Tier 1 default values for AGB, r and C.

The calculation of SOC stocks also involves four parameters in which IPCC provide Tier 1 default values for all. SOC is calculated according to the amount of SOC in mineral soils in the top 30 cm of the soil profile (SOC_R, hereinafter referred to as SOC reference value) multiplied by three factors (Equation 3), namely agricultural inputs (I_F), land management (M_F) and land use type (L_F).

$$BCS = AGB + BGB * C \quad \text{(Equation 1)}$$

$$BGB = AGB * r \quad \text{(Equation 2)}$$

$$SOC = SOC_R * I_F * M_F * L_F \quad \text{(Equation 3)}$$

It is important to mention that the parameters are assumed to be spatially dependent of at least two of three spatially heterogeneous factors (climate region, soil condition, and land use type). The SOC factor and the parameters used to calculate BCS are dependent on the land use type and climate region. The SOC reference values are spatially dependent on the land use type and soil condition. For instance, the AGB value for forests in the South of Brazil and forests in the North of Brazil cannot be the same since the climate condition in those regions is not equal. This means that IPCC provides for a single parameter many values that can be used. Therefore, the parameter value depends on those spatial attributes (for better understanding, see tables 2, 3 and 4).

In that sense, before doing the calculations and running the carbon model, we prepared the IPCC input data by identifying all the values that are applicable to Brazil, based on the spatial data we had. To account for uncertainty in those values, we used the uncertainty ranges given by IPCC (2006). It is important to mention that some uncertainty ranges are not provided by IPCC. In cases like that, no additional uncertainty from other sources was added.

Given the parameters with uncertainty, we build the probability density functions (PDF) describing the range and relative likelihood of possible values for each of the IPCC input data. When the run starts, the PDF are provided to the model for the selection of random values. After the random values are chosen, three text files (.txt) are created representing the BCS calculation, SOC_R and SOC_F random values, respectively. The calculation and allocation of the carbon stocks in the study area performed given those files and the spatial data of land use, climate, and soil.

The output of this process shown in Figure 2 represents the total carbon stocks for a given point in time according to the land use data that is used. To account for the GHG emissions estimates, the model must run with a different land use data representing a different point in time. The difference between the carbon stocks in this time frame allows the estimation of emissions.

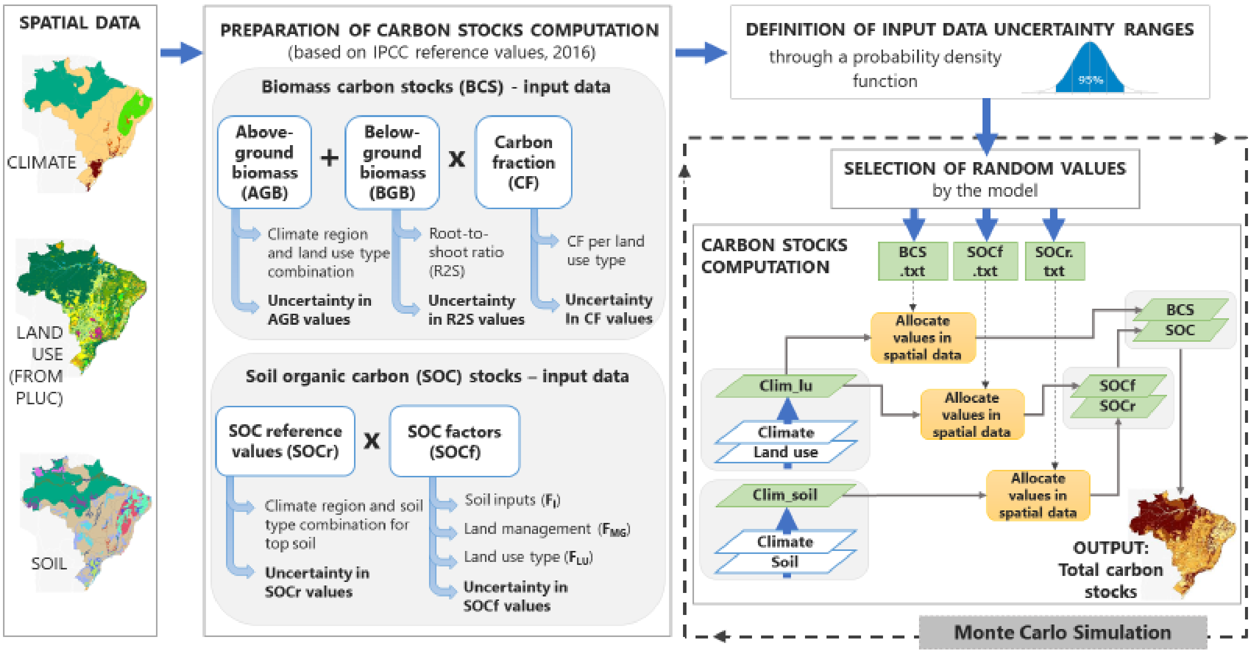

Figure 2 – Process of calculating carbon stocks with regards to the carbon model. The parallelograms represent spatial data

This process of selecting random values until the allocation of total carbon stocks represent a single run of the model. The model is set to perform this process 10,000 To run stochastically, which represents the Monte Carlo simulation. What the simulation does is making use of the probability density functions given to the model to generate random values in each run that is processed (i.e., a Monte Carlos realization).

By running the model 10,000 times, it is possible to evaluate uncertainty. After all the realizations are finish, one final distribution of the output values per run can be built to define the central estimate of the variable of interest and its related uncertainty. In the case of this thesis, the Monte Carlo input variables are the IPCC default parameter values related to SOC and biomass, while the variable of interest is represented by the total carbon stocks.

2.3.1 INPUT DATA

2.3.1.1 Spatial data

The spatial data used hereto (Table 1) regards to raster data at a cell size of 25 km^2 in the same spatial reference (WGS 1984 geographic coordinate system and Albers equal-area conic projection), with the same number of cells representing the total area in Brazil.

The climate spatial data is provided by van der Hilst *et al.* (2018). It includes five climate regions distinguished for Brazil that were obtained by combining elevation data (NASA and NGA, 2001) with temperature and precipitation data (Hijmans *et al.*, 2005). The soil spatial data (EMBRAPA and IBGE, 2001) includes five soil types. Both the spatial data are in line with classes shown in IPCC guidelines (IPCC, 2006; for climate classes, see vol. 4, chpt. 2, pg. 31; for soil types, see vol. 4, chpt. 4, pg. 46).

The land use spatial data are the land use outputs from PLUC obtained by van der Hilst *et al.* (2018). They represent the LUC mitigation scenarios projected for Brazil towards 2030, plus a land use data representing the initial state of the system in 2012. The land use types of the PLUC outputs are represented by 11 classes: natural forest, grass and shrubs, planted forest, rangeland, sugar cane, (other) cropland, planted pasture, abandoned land, urban, water, and bare soil.

Table 1 – Spatial data source used in the carbon model

Spatial Data	Format	Description	Data source
Climate in Brazil	Raster	Data provided by van der Hilst *et al.*, 2018. It distinguishes five climate regions as a result of the combination of temperature, precipitation and elevation data	Hijmans *et al.*, 2005 (temperature and precipitation); NASA and NGA, 2001) (Elevation)
Soil types in Brazil	Raster	Data provided by van der Hilst *et al.*, 2018. It has five classes of soil types for Brazil	EMBRAPA and IBGE, 2001
Land use	Raster	12 land use data representing the LUC scenarios for Brazil in 2030 and one land use data in 2012 (used to compute net changes)	It is the output of the PLUC model run by van der Hilst *et al.*, 2018

2.3.1.2 IPCC data with uncertainty

Tables 2, 3 and 4 show the parameter values used to calculate SOC reference values in top soil, SOC factors and biomass stocks, respectively. In all tables, only the land use classes, soil types and climate regions occurring in Brazil's spatial dataset are shown. Also, the land use classes Urban, Water and Bare Soil are not presented because they are assumed to have no carbon stocks.

The tables also show the values' uncertainty ranges given by IPCC, representing the 95% confidence interval expressed as a percentage of the central estimate of the values. As those values express a central estimate of a parameter that is uncertain, i.e., its true value is unknown, the "lack of knowledge" of the real value can be represented by a PDF to indicate the range and likelihood of possible values. Because the implementation of this model follows the use of default values from the Tier 1 method of IPCC guidelines, we assume a symmetrical PDF for all the input data (IPCC, 2006, vol. 1, chpt. 3., pg. 18). This means that the 95% confidence interval is expressed as plus or minus half the confidence interval width, divided by the estimated value of the variable (precisely as it is shown in the tables).

The same input data from IPCC was also used in the research of van der Hilst et al. (2018). Special cases in which the values differ are because hereto we try to use to the maximum the uncertainty information given by IPCC. Also, if the uncertainty is not provided by IPCC, no additional uncertainty from other sources is added.

Table 2 – Reference values with uncertainty for SOC in topsoil (30 cm depth), derived from IPCC guidelines (2006, vol. 4, chpt. 2, pg. 31)

Climate region	**Soil type (tonne C ha^{-1}) [1,2]**				
	Sandy	**Wetland**	**High Activity Clay**	**Low Activity Clay**	**Spodic**
Warm temperate moist	34 ± 90%	n/a	88 ± 90%	63 ± 90%	n/a
Tropical dry	31 ± 90%	n/a	38 ± 90%	35 ± 90%	n/a
Tropical moist	39 ± 90%	86 ± 90%	65 ± 90%	47 ± 90%	115 ± 90%
Tropical wet	66 ± 90%	86 ± 90%	44 ± 90%	60 ± 90%	115 ± 90%
Tropical montane	34 ± 90%	n/a	88 ± 90%	63 ± 90%	115 ± 90%

1 – Uncertainty values are expressed as a percentage of the central estimate. If a percentage does not follow the value, no uncertainty is given by IPCC; 2 – N/a means that this climate-soil combination does not occur in Brazil. Therefore they are not accounted in this thesis.

Table 3 – Values with uncertainty of SOC factors, derived from IPCC guidelines (2006, vol. 4, chapters 4, 5 and 6)

Land use type	**Climate**	**SOC factor[1]**		
		Land use factor (FLU)	**Management factor (FMG)**	**Input factor (FI)**
Natural Forest	Warm temperate moist	1	1	1
	Tropical dry	1	1	1
	Tropical moist	1	1	1
	Tropical wet	1	1	1
	Tropical montane	1	1	1
Rangeland[2]	Warm temperate moist	1	0.95 ± 13%	1
	Tropical dry	1	0.97 ± 11%	1
	Tropical moist	1	0.97 ± 11%	1
	Tropical wet	1	0.97 ± 11%	1
	Tropical montane	1	0.96 ± 40%	1
Planted Forest	Warm temperate moist	1	1	1
	Tropical dry	1	1	1
	Tropical moist	1	1	1
	Tropical wet	1	1	1
	Tropical montane	1	1	1
Crops [3]	Warm temperate moist	0.69 ± 12%	1	0.92 ± 14%
	Tropical dry	0.58 ± 61%	1	0.95 ± 13%
	Tropical moist	0.48 ± 46%	1	0.92 ± 14%
	Tropical wet	0.48 ± 46%	1	0.92 ± 14%
	Tropical montane	0.64 ± 50%	1	0.94 ± 50%
Grass and Shrubs [4]	Warm temperate moist	1	1	1
	Tropical dry	1	1	1
	Tropical moist	1	1	1
	Tropical wet	1	1	1
	Tropical montane	1	1	1
Sugar Cane [5]	Warm temperate moist	0.69 ± 12%	1.08 ± 5%	1.11 ± 10%
	Tropical dry	0.58 ± 61%	1.09 ± 9%	1.04 ± 13%

Land use type	Climate	SOC factor[1]		
		Land use factor (FLU)	Management factor (FMG)	Input factor (FI)
	Tropical moist	0.48 ± 46%	1.15 ± 8%	1.11 ± 10%
	Tropical wet	0.48 ± 46%	1.15 ± 8%	1.11 ± 10%
	Tropical montane	0.64 ± 50%	1.09 ± 50%	1.08 ± 50%
Planted Pasture[6]	Warm temperate moist	1	1.14 ± 11%	1
	Tropical dry	1	1.17 ± 9%	1
	Tropical moist	1	1.17 ± 9%	1
	Tropical wet	1	1.17 ± 9%	1
	Tropical montane	1	1.16 ± 40%	1
Abandoned[7]	Warm temperate moist	0.82 ± 17%	1.15 ± 4%	0.92 ± 14%
	Tropical dry	0.93 ± 11%	1.17 ± 8%	0.95 ± 13%
	Tropical moist	0.82 ± 17%	1.22 ± 7%	0.92 ± 14%
	Tropical wet	0.82 ± 17%	1.22 ± 7%	0.92 ± 14%
	Tropical montane	0.88 ± 50%	1.16 ± 50%	0.94 ± 50%

1 – Uncertainty values are expressed as a percentage of the central estimate. If the value is not followed by a percentage, no uncertainty is given by IPCC; 2 – The management factor of rangeland is assumed to be 'Moderately degraded grassland'; 3 – Cropland is assumed to be 'long term cultivated', with full tillage and low fertilizer input. Land use, management, and input factors are set accordingly; 4 – Grass and shrubs is assumed to be 'unmanaged land' so no factors are applied; 5 – Sugar cane is assumed to be 'long term cultivated', with reduced tillage and high fertilizer inputs (without manure); 6 – Planted pasture is assumed to be 'improved grassland', with medium input. Land use, management, and input factors are set accordingly; 7 – The values for 'set aside land' are assumed for abandoned land, with no tillage and no inputs.

Table 4 – Parameters with uncertainty values used to estimate biomass carbon stock, derived from IPCC guidelines (2006, vol. 4, chapters 4, 5 and 6)

Land use type	Climate	Parameter[1,2]			
		Above ground biomass (tonne dry matter ha^{-1})	Root to shoot	Carbon Fraction	Biomass Carbon (tonne C ha^{-1})
Natural Forest	Warm temperate moist	245 ± 14.3%	0.275 ± 20%	0.465 ± 5.38%	n/a
	Tropical dry	305 ± 34.4%	0.275 ± 1.8%	0.465 ± 5.38%	n/a
	Tropical moist	245 ± 14.3%	0.275 ± 20%	0.465 ± 5.38%	n/a
	Tropical wet	260 ± 53.8%	0.37	0.465 ± 5.38%	n/a
	Tropical montane	145 ± 58.6%	0.275 ± 1.8%	0.465 ± 5.38%	n/a
Rangeland[4]	Warm temperate moist	2.7 ± 75%	4 ± 150%	0.5 ±	n/a
	Tropical dry	2.3 ± 75%	2.8 ± 95%	0.5 ±	n/a
	Tropical moist	6.2 ± 75%	1.6 ± 130%	0.5 ±	n/a
	Tropical wet	6.2 ± 75%	1.6 ± 130%	0.5 ±	n/a
	Tropical montane	2.3 ± 75%	1.6 ± 130%	0.5 ±	n/a
Planted Forest[4]	Warm temperate moist	170.42	0.275 ± 20%	0.465 ± 5.38%	n/a
	Tropical dry	94.68	0.275 ± 1.8%	0.465 ± 5.38%	n/a
	Tropical moist	132.12	0.17 ± 47.1%	0.465 ± 5.38%	n/a
	Tropical wet	223.40	0.370	0.465 ± 5.38%	n/a
	Tropical montane	100 ± 70%	0.275 ± 1.8%	0.465 ± 5.38%	n/a
Crops[5]	Warm temperate moist	n/a	n/a	n/a	5 ± 75%
	Tropical dry	n/a	n/a	n/a	5 ± 75%
	Tropical moist	n/a	n/a	n/a	5 ± 75%
	Tropical wet	n/a	n/a	n/a	5 ± 75%
	Tropical montane	n/a	n/a	n/a	5 ± 75%
Grass and Shrubs	Warm temperate moist	2.7 ± 75%	2.8 ± 144%	0.47	n/a
	Tropical dry	2.3 ± 75%	2.8 ± 144%	0.47	n/a
	Tropical moist	6.2 ± 75%	2.8 ± 144%	0.47	n/a
	Tropical wet	6.2 ± 75%	2.8 ± 144%	0.47	n/a

Land use type	Climate	Parameter[1,2]			
		Above ground biomass (tonne dry matter ha^{-1})	**Root to shoot**	**Carbon Fraction**	**Biomass Carbon** (tonne C ha^{-1})
	Tropical montane	2.3 ± 75%	2.8 ± 144%	0.47	n/a
Sugar Cane	Warm temperate moist	19.69	0.20	0.47	n/a
	Tropical dry	19.69	0.20	0.47	n/a
	Tropical moist	19.69	0.20	0.47	n/a
	Tropical wet	19.69	0.20	0.47	n/a
	Tropical montane	19.69	0.20	0.47	n/a
Planted Pasture	Warm temperate moist	2.7 ± 75%	4 ± 150%	0.47	n/a
	Tropical dry	2.3 ± 75%	2.8 ± 95%	0.47	n/a
	Tropical moist	6.2 ± 75%	1.6 ± 130%	0.47	n/a
	Tropical wet	6.2 ± 75%	1.6 ± 130%	0.47	n/a
	Tropical montane	2.3 ± 75%	1.6 ± 130%	0.47	n/a
Abandoned[6]	Warm temperate moist	n/a	n/a	n/a	2.5 ± 75%
	Tropical dry	n/a	n/a	n/a	2.5 ± 75%
	Tropical moist	n/a	n/a	n/a	2.5 ± 75%
	Tropical wet	n/a	n/a	n/a	2.5 ± 75%
	Tropical montane	n/a	n/a	n/a	2.5 ± 75%

1 – Uncertainty values are expressed as a percentage of the central estimate. If the value is not followed by a percentage, no uncertainty is given by IPCC; 2 – N/a means not given by IPCC; 3 – IPCC does not differentiate between grassland and rangeland. It is assumed that the figures provided by IPCC for grassland are representative for rangeland; 4 – The figures are based on the ratio 77 % eucalyptus and 23 % of pine based on the current composition of planted forest (ABRAF, 2013); 5 – For cropland, IPCC (2006) does not provide numbers for above and below ground biomass, just for the total amount of biomass; 6 – No information is available for abandoned land. Therefore it is assumed that half of the available biomass of cropland is available in abandoned land.

2.3.2 IMPLEMENTATION

The model is implemented in Python programming language. It consists of a single script of which the main package used during the implementation is ‘NumPy’. This package is mainly used to work with raster files in the format of multidimensional arrays thus providing a fast and powerful way to process spatial data. The conversion from raster to array or vice-versa is performed with the use of ‘GDAL’ and ‘OSR’ libraries. The ‘random’ module and ‘SciPy’ library are used to generate random values necessary to run the Monte Carlo simulation. ‘Pandas’ library is used to convert arrays in structured data and analyse them. ‘Matplotlib’ library is used for plotting. The modules ‘glob’, ‘os’, and ‘time’ are also used for other tasks.

2.4 SCENARIO APPROACH AND GLOBAL ETHANOL DEMAND

Both the scenario approach and the global ethanol demand used hereto are provided by van der Hilst *et al.* (2018). Six potential LUC mitigation scenarios were used in their evaluation, plus a reference scenario with no measures as it is shown in Table 5. The strategies include an increase in agricultural productivity, shifting towards second-generation of ethanol production using sugar cane or eucalyptus, and implementing land conservation policies.

Table 5 – Brief description of the LUC mitigation scenarios for Brazil up to 2030

Code	Scenario[1]	Description/assumptions
Ref	Reference scenario	Brazil will develop towards 2030 according to historical trends that are in line with the SSP2 scenario for global development of the Shared Socio-economic reference Pathways (O'Neill *et al.*, 2014, 2017). No additional measures are considered to improve either agricultural productivity or strict conservation policies. Incremental improvements are assumed to occur in the first-generation ethanol production chain. No shifting towards the second-generation of ethanol is considered. Land use changes do not occur in military, indigenous, federal and state conservation areas.
HP	Improved agricultural productivity	The annual yield increase is twice as high compared to the reference scenario.
2nd Gen. SC	A shift towards the 2nd generation of ethanol (sugar cane)	Brazil will combine improvements in the first-generation ethanol production chain with a shift towards second-generation ethanol from bagasse and sugar cane straw.
2nd Gen. EU	A shift towards the 2nd generation of ethanol (eucalyptus)	Brazil will combine improvements in the first-generation ethanol production chain with a shift towards second-generation ethanol from bagasse and sugar cane straw until 2020. From 2020 onwards, a full shift towards second-generation ethanol from eucalyptus is considered.
CP	Strict land conservation policies	Together with military, indigenous, federal and state conservation areas, natural forests cannot be converted to any other land use from 2015 onwards.
All	All mitigation measures	Represent a scenario in which the LUC mitigation measures are combined, namely: high agricultural productivity, shift to second-generation ethanol from sugar cane, and strict conservation policies.

1 – The scenarios are provided by van der Hilst et al. (2018). For a full description, consult the study.

The projections of global ethanol demand concern exclusively to ethanol, no other biofuel. They were based on the OECD-FAO Agricultural outlook (OECD and FAO, 2014) and ICONE, the Brazilian Institute for International Trade Negotiations (FIESP and ICONE, 2012). As the scenarios are evaluated twice in this thesis (with and without additional global demand for biofuels), in the evaluation without the demand, it is assumed that the global demand remains at the level of 2013.

The results of van der Hilst *et al.* (2018) obtained by the model MAGNET show that ethanol production in Brazil is projected to more than double if the additional demand is considered.

2.5 MODEL RUNS

The carbon model is run to analyse the carbon stocks and LUC-related GHG emissions in the period between 2012 and 2030. In total, the Monte Carlo simulation is performed 13 times (Table 6), which characterizes a total of 130,000 realizations. The first Monte Carlo simulation is run for 2012, and it represents the initial state of the land use system. Its results are compared with the results of all other scenarios to assess the net changes in carbon stocks and GHG emissions. The other Monte Carlo simulations are run for the LUC mitigation scenarios, representing the LUC mitigation scenarios that are evaluated twice (with and without the increment in the global ethanol demand).

Table 6 – Description of the Monte Carlo simulations performed by the model runs

MC simulation	Code	Scenario description	Abbrev.	Ethanol demand (mln liters)[1,]
1	SC0	Initial state of the land use system (2012)	2012	23901
2	SC1	Reference scenario – without additional ethanol demand	REF	28246
3	SC2	Reference scenario – with additional ethanol demand	REF +Eth	54234
4	SC3	High agricultural productivity – without additional ethanol demand	HP	28768
5	SC4	High agricultural productivity – with additional ethanol demand	HP +Eth	55072
6	SC5	Shift towards 2nd generation ethanol (sugar cane) – without additional ethanol demand	2nd SC	30946
7	SC6	Shift towards the 2nd generation of ethanol (sugar cane) – with additional ethanol demand	2nd SC +Eth	58583
8	SC7	Shift towards the 2nd generation of ethanol (eucalyptus) – without additional ethanol demand	2nd EU	27787
9	SC8	Shift towards the 2nd generation of ethanol (eucalyptus) – with additional ethanol demand	2nd EU +Eth	53471
10	SC9	Strict conservation policies – without additional ethanol demand	CP	28151
11	SC10	Strict conservation policies – with additional ethanol demand	CP +Eth	54234
12	SC11	All LUC mitigation measures – without additional ethanol demand	ALL	30871
13	SC12	All LUC mitigation measures – with additional ethanol demand	ALL +Eth	58503

1 – Ethanol production for Brazil in 2030, based on projections of global ethanol demand (van der Hilst et al. 2018),

Once each of the Monte Carlo simulations is finished, the model gathers the outputs of the realizations and produce final distributions of the simulation, thus allowing the computation of mean carbon stock values in Brazil. The quantification of uncertainty as a result of the Monte Carlo simulation is obtained by identifying the 95% confidence interval of the distributions (see description in subsection 2.6).

Although 130,000 realizations are performed, the random values that are selected from the PDFs of the input data are only obtained 10,000 times. This is because the Monte Carlo realizations share the same random values between the scenarios to avoid uncertainty that we do not want. For instance, we do not know the SOC value of the initial state of the system, but we know that the initial soc (2012) does not depend on the land use dynamics in the future (2030). So, we do take into account uncertainty, but not between runs, as they start from the same state for sure.

The changes in carbon stocks are computed per Monte Carlo realization, accounting for the difference between stocks in 2012 and 2030, for all scenarios. Also, the difference of stocks between the scenarios with and without an increase in ethanol production is computed per Monte Carlo realization. As a result, the carbon model runs produce distributions to assess:

a) Carbon stocks estimates and associated uncertainty for 2012 and for each scenario;

b) Changes in carbon stocks and associated uncertainty between 2012 and 2030, for each scenario;

c) Changes in carbon stocks and associated uncertainty between the scenarios with and without addition ethanol production.

A factor of 44/12 representing the ratio of the molecular weights of CO_2 (44) and carbon (12) is used to carbon stock changes in LUC-related GHG emissions..

Regarding the difference of carbon stocks between the scenarios with and without increase in ethanol production, once the conversion to emissions is realized, the LUC-related GHG emissions specifically allocated to ethanol production are calculated as the total LUC-related GHG emissions resulting from the additional ethanol demand divided by the total ethanol production that can be obtained in 20 years. The amortization period of 20 years is in line with the IPCC Guidelines for National Greenhouse Gas Inventories (2006).

2.6 QUANTIFYING AND EXPRESSING UNCERTAINTY

Each distribution derived from the Monte Carlo simulation allows the quantification of uncertainty. This is performed through the identification of the 95% confidence interval of the distribution, which is represented by the 2.5th and 97.5th percentiles of the distribution.

The mean and median are used as statistical measures to identify the symmetry of one distribution to define how to compute the confidence interval. If the mean and median are equal, we assume the distribution is normal, and the 2.5th and 97.5th percentiles are obtained from both mean and median that the confidence interval will be the same. If the mean and median differ, then it is assumed a non-normal distribution and the confidence interval is given by calculating the percentiles concerning the median.

The expression of uncertainty in the model outputs follows the format of the uncertainty shown in the IPCC input data (see subsection 2.3.1.2) i.e. the 95 confidence interval is expressed as a percentage of the mean value of the distribution. For example, considering a normal distribution, if the mean value is calculated as 100 tonnes of carbon, the 2.5th and 97.5th percentiles are 70 and 130 tons/C, respectively. The mean value would be expressed as 100 tons/C ±30%. If the distribution is non-normal, then the uncertainty range is asymmetric. Taking the same 100 tons/C as an example, but now with the 2.5th percentile equals to 50 ton/C and the 97.5th percentile equals to 200, the mean value with uncertainty would be expressed as 100 tons/C -50% to +100%.

2.7 SENSITIVITY ANALYSIS AND STATISTICAL TEST

The global sensitivity analysis is realized by using the Sobol’ method (Sobol’, 1993 in: Convertino *et al.*, 2014). The analysis is done in each scenario to compute the contributions of the carbon stock main components (hereto, SOC and biomass stocks) to the total uncertainty obtained in the LUC-related GHG emission estimates resulting from the addition ethanol demand. The method represents the contributions of the components as a fraction of the total variance in the model output (i.e., the LUC-related GHG emission estimates resulting from the addition ethanol demand). The fractions are calculated by running the model two more times: by setting the SOC input data to run deterministically and the biomass input data to run stochastically; and vice-versa.

The Kruskal–Wallis test (Kruskal and Wallis, 1952) is used to perform the statistical test, which analyses if there is a significant difference in the GHG emission estimates resulting from the addition ethanol demand between the simulated scenarios of LUC mitigation strategies. If significant, the test indicates that at least one scenario is significantly different from the others. Next the post-hoc tests after Nemenyi are applied (Nemenyi, 1963) to identify the differences between the scenarios. The test is realized in R (R core Team, 2018) with the ‘stats’ and ‘PMCMR’ (Pohlert, 2014) packages. The ‘multcompView’ package is used to plot the test results, based on the compact letter display method (Piepho, 2004). The letters in the plot are assigned considering a significance level of 0.01 (i.e., p-value).

3 RESULTS AND DISCUSSION

3.1 TOTAL CARBON STOCKS

The results of the carbon model show a reduction in total carbon stocks for all scenarios in 2030, with and without an additional ethanol demand. The 120.32 x 10^9 tonnes of carbon in 2012 are reduced to 117.78 x 10^9 tonnes of carbon in the worst-case scenario, represented by the reference scenario with additional ethanol production. The scenario with the lowest reduction is when all the mitigation measures are implemented, with the carbon stocks estimates of 119.78 x 10^9 tonnes of carbon.

Brazil has considerably more carbon stocks of biomass than stocks of SOC (Figure 3). For all scenarios and 2012, it is verified that biomass carbon stocks represent 62 to 63% of the total stocks, while 37 to 38% are related to SOC. This is mainly because the North region strongly influences biomass stocks in Brazil, where Amazonia is located, with huge stocks of carbon in comparison to other regions. (for evidence, see mean carbon stocks for the reference scenario in Figure 6). When the additional global ethanol production is taken into account, there is a reduction in the amount of carbon stocks estimates for all scenarios in comparison to the scenarios without additional ethanol demand.

The uncertainty in SOC stock estimates is higher than in biomass stocks estimates for all scenarios (Figure 3). This is verified for all scenarios, including 2012. For instance, the SOC stocks estimates projected for the reference scenario are 45.06 x10^9 tons C, with an associated uncertainty of -39.5% to +42.3%, while the biomass stocks estimates are 75.27 x10^9 tons C +-32%.

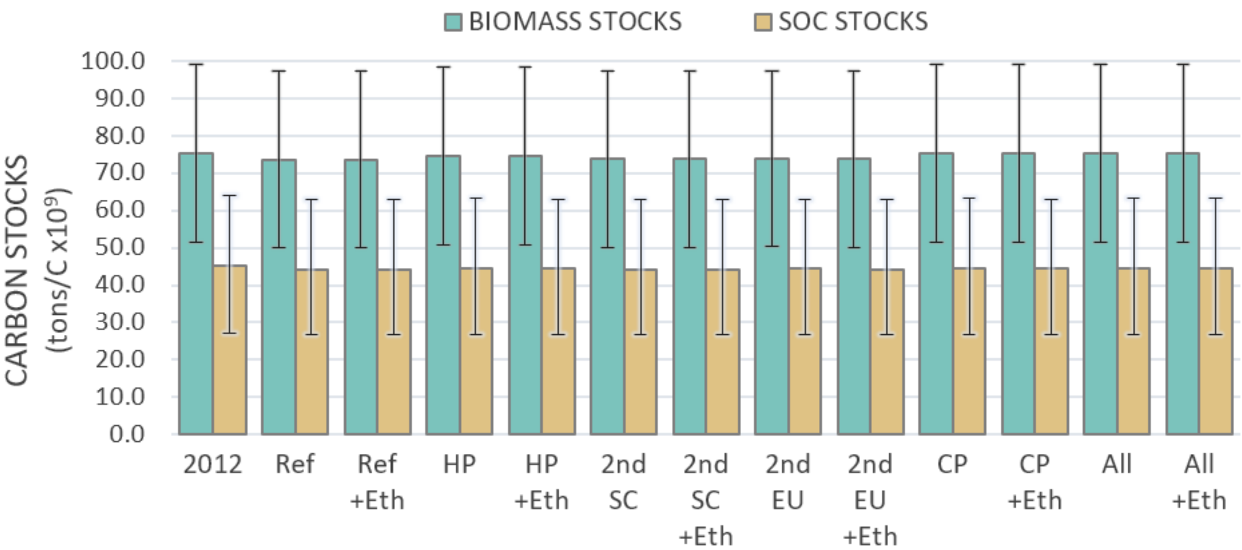

Figure 3 – Total carbon stocks estimates in Brazil for 2030, given the LUC mitigation scenarios with and without an increase in ethanol production.

Figure 3 also illustrates that the carbon stocks estimates and their associated uncertainty are very similar among the scenarios. From an implementation perspective, this can be explained because the scenarios share the same random values between the Monte Carlo runs. The final distributions of carbon stocks estimates resulting from the Monte Carlo simulation demonstrate such similarity. This is illustrated by the examples in Figure 4 in which the distributions for the initial state of the system (2012) and the reference scenario without an increase in ethanol are shown.

Furthermore, the model uncertainty was only added in the input data from IPCC, but not to spatial data e.g. in the land use data (PLUC outputs). As shown by Verstegen et al. (2016), projections of LUC are highly uncertain. Therefore we expect that the inclusion of uncertainty is derived from the land use dynamics in the model would lead to different estimates and to more variance in the final uncertainty ranges between the scenarios.

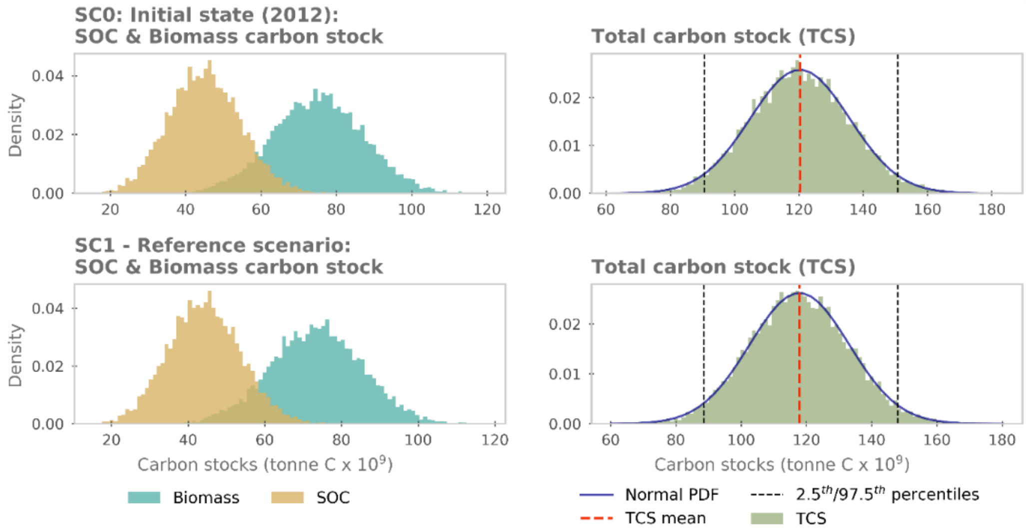

Figure 4 – Final distributions of carbon stocks estimates resulting from the Monte Carlo simulation, demonstrated for the initial state of the system (2012) and the reference scenario

The higher uncertainty of SOC stocks estimates in comparison with biomass stocks estimates suggests that the main source of uncertainty from the input data lies on the parameters used to compute SOC stocks. A considerable source of uncertainty that is likely influencing the uncertainty of SOC stocks estimates is the input parameter SOC reference (SOC_R), which has an uncertainty of 90%. Some inputs related to biomass stocks also have high uncertainties, even higher than 90%, as the cases of the root-to-shoot ratio parameter (up to 150% in a warm moist climate for planted pasture), but the contribution to the overall uncertainty might not be so significative when compared to SOC_R values.

The SOC_R affect all the land use types that are assumed to have carbon stocks, while the root-to-shoot ratio only has high uncertainty values for specific land use types, namely planted pasture, shrublands, and rangelands. Hence, when the model runs the uncertainty of SOC_R is propagated to every raster cell, while the uncertainty of root-to-shoot ratio is only propagated to the mentioned land use types. This is illustrated in Figure 5, where the reference scenario without additional ethanol demand is considered. It is possible to verify the location of the land use types of which the root to shoot parameter with high uncertainty is propagated in comparison to the location of which SOC_R is propagated.

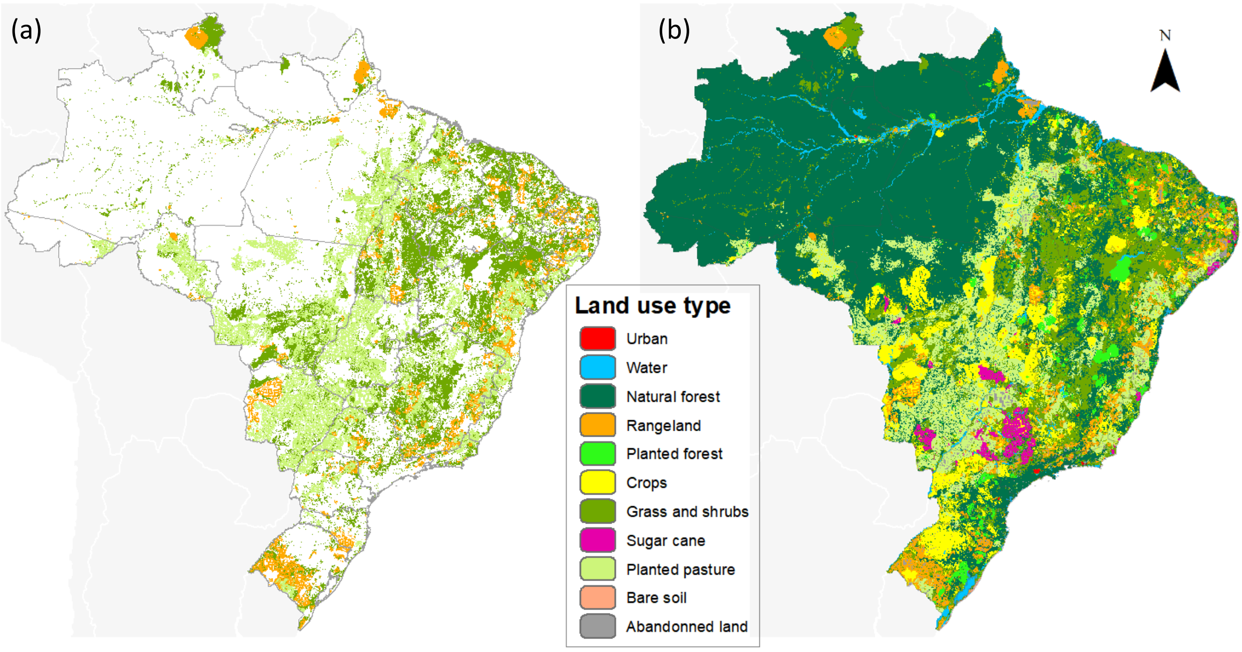

Figure 5 – (a) Location of the land use types of which the root to shoot with high uncertainty is propagated; (b) Location of the land use types of which SOC reference with high uncertainty is propagated

Furthermore, there is a large spatial variability in the uncertainty because IPCC parameters are spatially dependent on land use, soil type, and climate conditions. This influences the way that uncertainty in the input data is propagated. This is specifically evident if e.g. we compare the SOC stock and biomass stock estimates in the South of Brazil (see enlargements of Figure 6). This region is within a climate transition area involving two climate regions with different soil types and many land use types, i.e., the area involves all those particularities that are taken into account when the uncertainty is propagated. Therefore, we see high spatial variability in the allocation of SOC stocks, as exemplified in the Monte Carlo realizations shown in Figure 6. For biomass stocks, the spatial variability occurs, but it is not so expressive as it is for SOC stocks. This is because biomass stocks do not account for soil factors.

On the other hand, if we analyse the Amazon region, we see less spatial variability for both SOC and biomass stocks because this region is mostly represented by one land use type, i.e. forests. Consequently, the spatial variability in the uncertainty associated with this region is lower than the uncertainty associated with the South region.

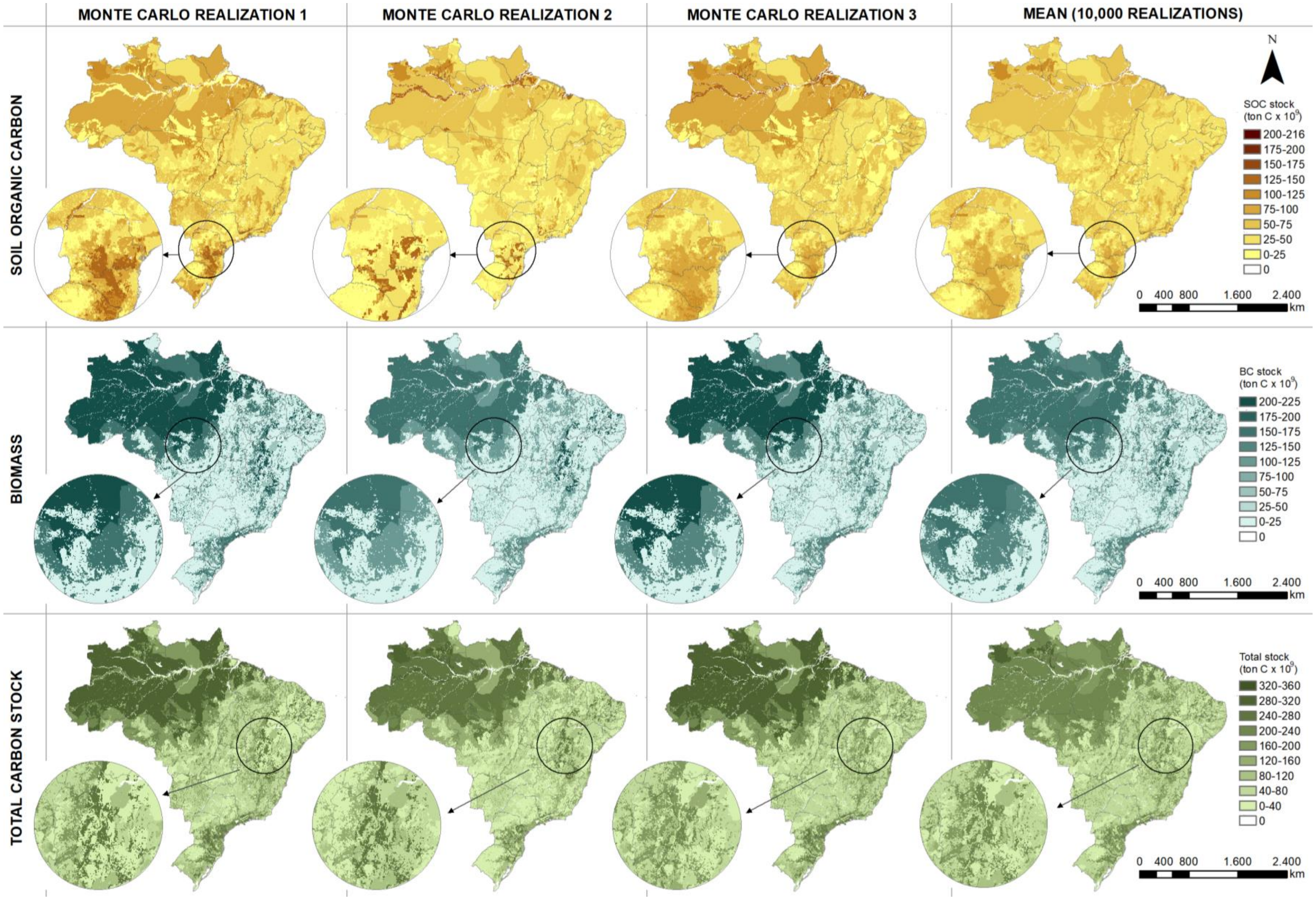

Figure 6 – Example of different Monte Carlo realizations and the mean carbon stocks obtained for the reference scenario without an increase in ethanol production.

3.2 LUC-RELATED GHG EMISSIONS RESULTING FROM AN INCREASE IN BIOFUEL DEMAND

The direct effect in the land use caused by an additional demand for biofuels is the increase in the land requirements for ethanol production from sugar cane in Brazil. This occurs for all the scenarios, resulting in more GHG emissions when compared to the scenarios without the additional demand. The increase in ethanol production mostly affect SOC stocks. Therefore, the LUC-related GHG emissions estimates resulting from the addition ethanol production are mainly caused by the net changes in SOC stocks. However, this does not occur when the shift towards the 2nd generation of ethanol from eucalyptus is considered. In this scenario, the main source of emissions derives from the changes in biomass.

In general, the emissions estimates resulting from the additional ethanol production are similar to the results of the deterministic approach performed by van der Hilst *et al.* (2018) (Figure 7). A substantial difference accounts for the emissions in the scenario related to the shift towards the 2nd generation of ethanol from eucalyptus. They computed 5.4 g CO_2-eq/MJ of emissions from biomass, while hereto the estimate is 7.4 g CO_2-eq/MJ -43% to +44%. However, given the associated uncertainty, we can state that their value is within the 95% confidence interval of our estimates.

Although the estimates are similar between the studies in the other scenarios, we consider that the estimates of this stochastic approach come with great uncertainty (Figure 7). This is especially evident in the emissions derived from the net changes in SOC stocks, where the uncertainty represent at least 75% of the emission estimates e.g. in the reference scenario with 20 g CO_2-eq/MJ -77% to +110%. In the case of net changes related to biomass, uncertainty is higher in the emission estimates of the reference scenario and strict conservation policies (6.0 CO_2-eq/MJ -87% to +107% and 4.4 CO_2-eq/MJ -92% to +119%), but this does not mean that uncertainty is low in the other scenarios e.g., the emission estimates in the scenario of a shift towards the 2nd generation of ethanol from eucalyptus are 7.6 CO_2-eq/MJ -50% to +65%.

Because of the high uncertainty, the results depict that some of the GHG emissions estimates can even represent GHG savings, as the 95% confidence interval reaches values below zero (see e.g. the confidence interval shown in the boxplot of the reference scenario in Figure 7). However, what it most concerns is that the emissions can be much higher than the estimates.

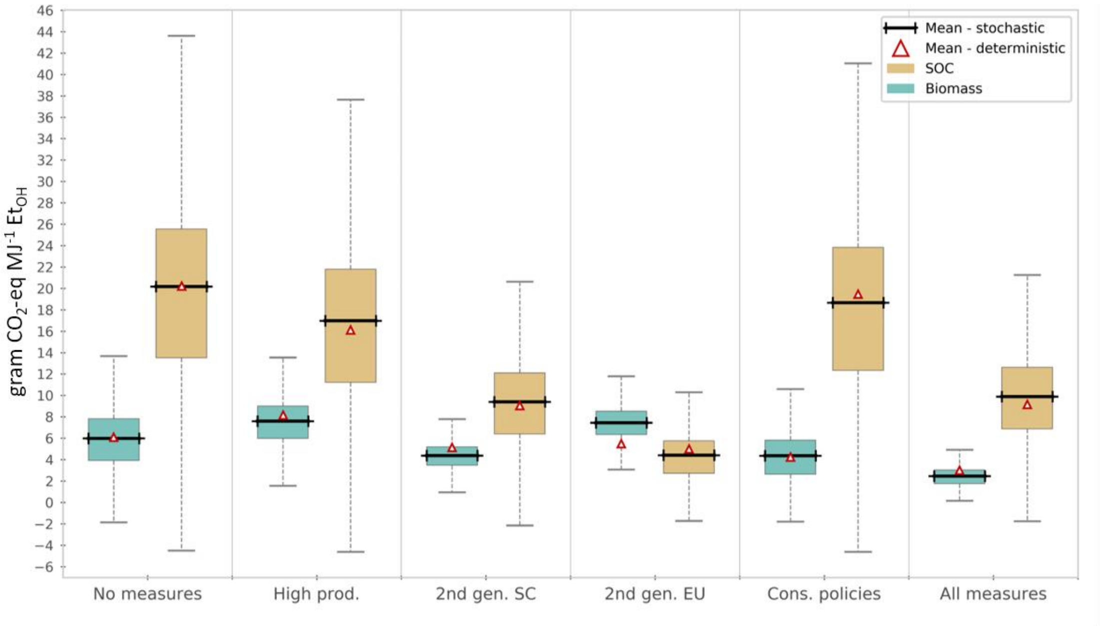

Figure 7 – Boxplots of LUC-related GHG emissions resulting from an increase in ethanol production for Brazil up to 2030.

In addition, the results indicate that the scenario related to the shift towards the 2nd generation of ethanol from eucalyptus is the most dissimilar in comparison with the others. This is because the production of ethanol from eucalyptus results in the increase of planted forests at the expense of natural forests. Although GHG savings are promoted by the increase in eucalyptus plantations, they are not enough to compensate the emissions resulting from loss of forests, even if we account for uncertainty (Figure 8).

Apart from the LUC mitigation scenario of ethanol production from eucalyptus, Figure 8 shows a clear pattern among the scenarios with regards the influence of the additional ethanol production in the LUC-related GHG emissions per land use type. The increase in areas for sugar cane predominantly results in emissions in rangelands. Also, no emissions occurs directly from crops. However, crops influence the emissions in other land use types. This can be explained by the cascading pattern explained by van der Hilst *et al.* (2018): sugar cane expands predominantly at the expense of cropland, which in turns expands at the expense of mostly rangeland and planted forest, which successively results in the conversion of other land use types.

The GHG savings associated to sugar cane shown in Figure 8 might be explained by the fact that sugarcane sequesters more carbon from SOC and biomass when it expands at the expense of croplands. In comparison to cropland, sugar cane has more biomass and higher factors of SOC (see tables 2 and 3 regarding the IPCC input data).

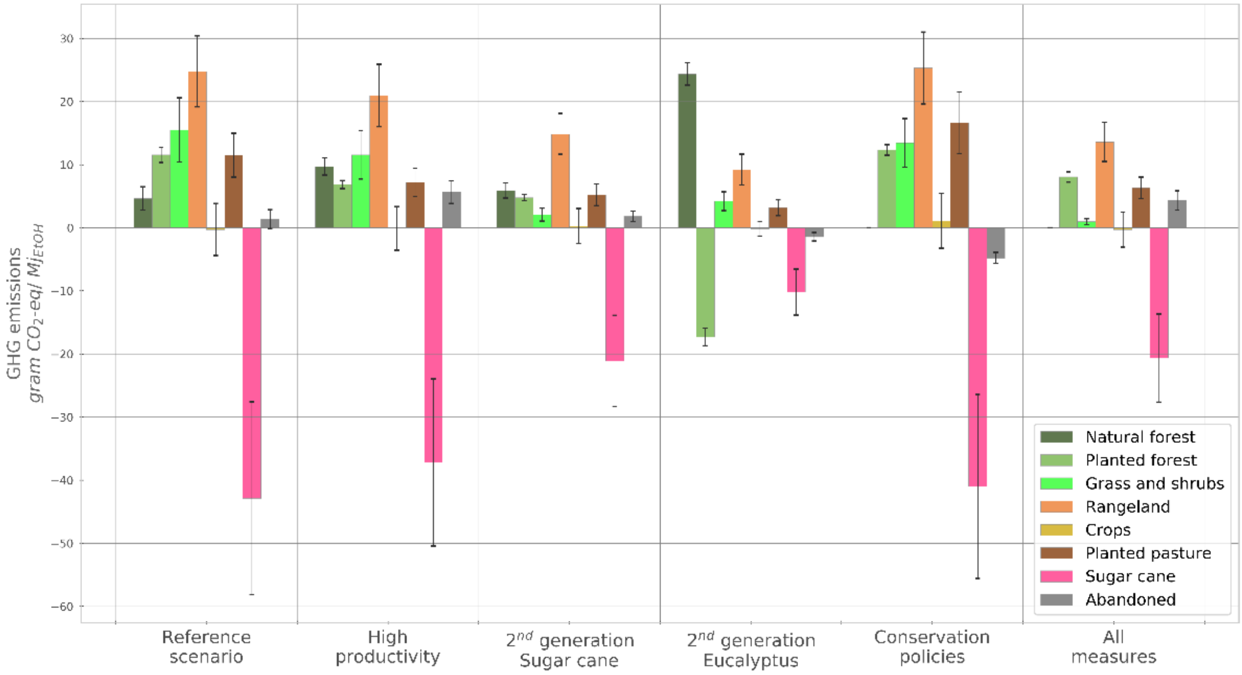

Figure 8 – LUC-related GHG emissions per land use type resulting from an increase in ethanol production

3.3 Sensitivity analysis and statistical test

Considering the two components analysed (SOC stock and biomass stock), the results of the global sensitivity analysis shows that the main contributor of the uncertainty in the LUC-related GHG emission estimates resulting from the additional ethanol demand refers to SOC stock. This is verified for all the scenarios, with a subtle difference in the scenario related to the shift towards the 2nd generation of ethanol from eucalyptus, where biomass contributes more than in other scenarios. As mentioned in subsection 3.2, we assume that the LUC dynamics are very particular for this scenario because the areas of planted forest increase to product eucalyptus and this occurs at the expanse of natural forests. Therefore, the LUC dynamics occurring in this scenario predominantly affect the land use types that most have biomass stocks.

The SOC stock component contribution represented as a fraction of the total variance in the overall uncertainty of the GHG emissions estimates, varies between 57% in the scenario related to the shift towards the 2nd generation of ethanol from eucalyptus and 83.2% in the scenario where all mitigation measures are considered. The biomass component contribution varies between 3.9% in the scenario where all mitigation measures are considered and 3.1.0% in the scenario related to the shift towards the 2nd generation of ethanol from eucalyptus. In all scenarios, about 12~13% of the total variance represent a contribution related to the model interactions (Figure 9).

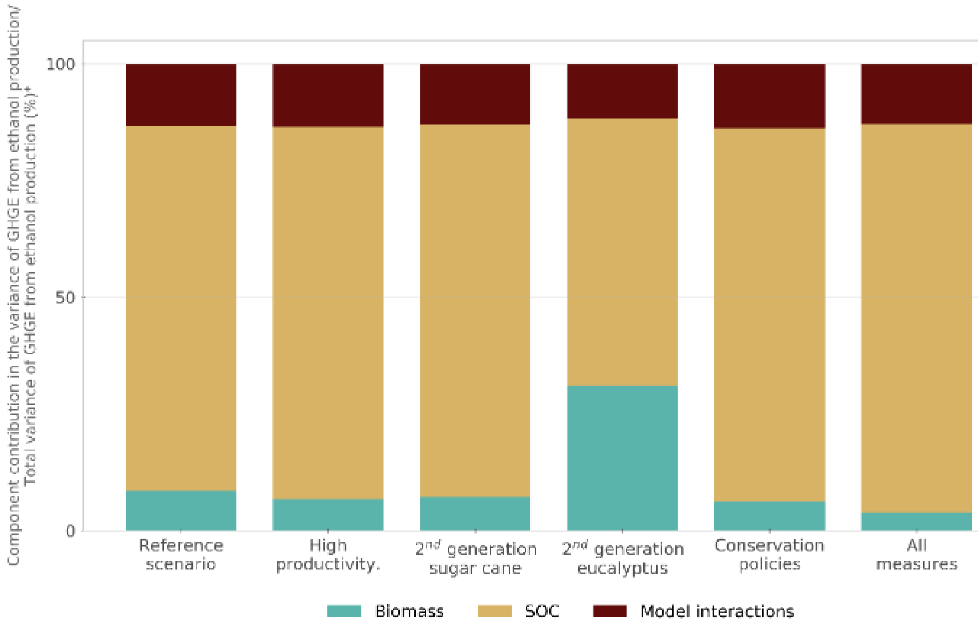

Figure 9 – Global sensitivity analysis results showing the contribution of SOC and biomass stocks to the total variance in the LUC-related GHG emissions estimates resulting from an increase in ethanol production

The results of the statistical test applied hereto given a p-value of 0.01 allows saying that the LUC-related GHG emission estimates resulting from the additional ethanol demand are significantly different between all scenarios (see Figure 10 with the plot of the compact letter display). This indicates that the hypothesis of similarity of mean among the scenarios is rejected. In other words, the emission estimates could be used to support decision making e.g. to define or prioritize the implementation of a new LUC mitigation measure in Brazil.

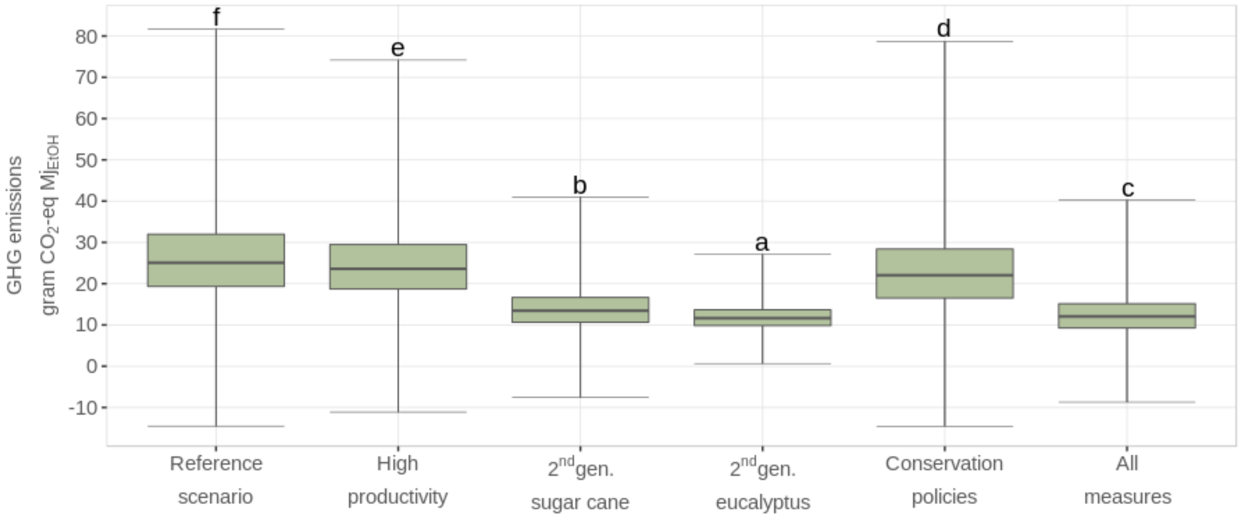

Figure 10 - Boxplots of LUC-related GHG emission estimates with the compact letter display: if two boxplots have the same letter, the hypothesis that they come from the same population cannot be rejected under p-value equals to 0.01

4 CONCLUSION

In this study, we developed a spatially explicit, stochastic model that accounts for uncertainty information in the input data used to calculate GHG emissions, given the net changes in SOC and carbon stocks. We applied the model in a case study based on the work of van der Hilst *et al.* (2018) in Brazil. They stochastically calculate the LUC-related GHG emissions resulting from an increase in ethanol production up to 2030, given six distinct scenarios of LUC mitigation measures.

To run stochastically, we added uncertainty information in the model input data from IPCC, which is the data representing the parameters used to calculate carbon stocks. The model also uses spatial data, namely: climate, soil type and land use. Although there is an inherent uncertainty in those data, we had not considered them in the model because there was no uncertainty information available for it. In the case of Brazil, as the country has a large amount of carbon stocks and many processes occurring in land, the addition of uncertainty in the spatial data would improve the uncertainty analysis.

The results of the model runs show that the addition of uncertainty in the IPCC input data results on GHG emissions estimates with great uncertainty for all scenarios. For example, the highest uncertainty was found in the GHG emission estimates resulting from changes in SOC stocks in the scenario related to the shift towards the 2nd generation of ethanol from sugarcane (20.2 g CO_2-eq/MJ -77% to +109%), while the lowest uncertainty was found in the GHG emission estimates resulting from changes in biomass stocks in the scenario related to the shift towards the 2nd generation of ethanol from eucalyptus (7.4 g CO_2-eq/MJ -43% to +44%).

The emission estimates obtained in this thesis have similar values when comparing to the results of the deterministic approach of van der Hilst *et al.* (2018), but a substantial difference accounted for the emissions in the scenario related to the shift towards the 2nd generation of ethanol from eucalyptus. While they computed 5.4 g CO_2-eq/MJ of emissions from biomass stocks, we estimated 7.4 g CO_2-eq/MJ -43% to +44%.

Considering the two components analysed in the global sensitivity analysis (SOC stock and biomass stock), we verified that the main contributor of the uncertainty in the LUC-related GHG emission estimates resulting from the addition ethanol demand refers to SOC stock.

The results of the statistical test applied in this thesis allows saying that the LUC-related GHG emission estimates resulting from the additional ethanol demand are significantly different between all scenarios. This means that the emission estimates could be used to support decision making.

We believe that GHG emission estimates with uncertainty ranges provide crucial information to decision makers and allow for more realistic interpretations in comparison to deterministic estimates. Based on that, they could make wiser decisions e.g. to define or prioritize the implementation of a new LUC or climate change mitigation measure. In that sense, ignoring uncertainty in scenario projections is not recommended since the information of possible ranges of GHG emissions estimates is not taken into account. By increasing the knowledge about uncertainty , we reduce the chance of policy makers to make wrong decisions.

Finally, as the model hereto was developed to replace part of the modelling framework proposed by van der Hilst et.al (2018) in order to account for uncertainty, we believe that the work shown in this thesis represents an additional step for a fully stochastic run of their modelling framework.

5 REFERENCES

ABRAF (2013) *Statistical yearbook 2013, base year 2012*. Brasilia: Associaçao Brasileira de Produtores de Florestas Plantadas.

Bryngelsson, D. (2015) *Land-use competition and agricultural greenhouse gas emissions in a climate change mitigation perspective (Doctoral Thesis)*. Chalmers University of Technology.

Chum, H. *et al.* (2011) *IPCC special report on renewable energy sources and climate change mitigation*, *Energy Sources*. Edited by O. Edenhofer et al. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139151153.

Convertino, M. *et al.* (2014) 'Untangling drivers of species distributions: Global sensitivity and uncertainty analyses of MaxEnt', *Environmental Modelling and Software*. doi: 10.1016/j.envsoft.2013.10.001.

Deser, C. *et al.* (2012) 'Uncertainty in climate change projections: The role of internal variability', *Climate Dynamics*, 38(3–4), pp. 527–546. doi: 10.1007/s00382-010-0977-x.

EMBRAPA and IBGE (2001) 'Brazil's soil map, from compilations (Embrapa, IBGE), conversion to digital and soil surveys produced by RADAMBRASIL. 1:5.000.000'. Laboratório de Processamento de imagens e Geoprocessamento (LAPIG) do Instituto de Estudo Sócio-ambientais (IESA) da Universidadefederal de Goiás (UFG).

Fargione, J. *et al.* (2008) 'Land Clearing and the Biofuel Carbon Debt', *Science*, 319(5867), pp. 1235–1238. doi: 10.1126/science.1152747.

FIESP and ICONE (2012) *Outlook Brazil 2022. Agribusiness projections.* Sao Paulo, Brazill.

Hijmans, R. J. *et al.* (2005) 'Very high resolution interpolated climate surfaces for global land areas', *International Journal of Climatology*, 25(15), pp. 1965–1978. doi: 10.1002/joc.1276.

van der Hilst, F. *et al.* (2018) 'Mapping land use changes resulting from biofuel production and the effect of mitigation measures', *GCB Bioenergy*, (December 2017). doi: 10.1111/gcbb.12534.

IPCC (2006) *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds).*, *IGES*. Japan. Available at: http://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/0_Overview/V0_0_Cover.pdf.

Kim, Y. H. and Sohngen, B. (2009) 'Assessing the uncertainty of land based carbon sequestration: A parameter uncertainty analysis with a global land use model', *2009 Annual Meeting, July 26-28, 2009, Milwaukee, Wisconsin*, (49416). Available at: http://ideas.repec.org/p/ags/aaea09/49416.html%5Cnpapers2://publication/uuid/84D09086-D9F4-434B-98D3-5D048354CE56.

Kruskal, W. H. and Wallis, W. A. (1952) 'Use of Ranks in One-Criterion Variance Analysis', *Journal of the American Statistical Association*. doi: 10.1080/01621459.1952.10483441.

Macedo, I. C., Seabra, J. E. A. and Silva, J. E. A. R. (2008) 'Green house gases emissions in the production and use of ethanol from sugarcane in Brazil: The 2005/2006 averages and a prediction for 2020', *Biomass and Bioenergy*, 32(7), pp. 582–595. doi: 10.1016/j.biombioe.2007.12.006.

Mustafa, A. *et al.* (2018) 'A Time Monte Carlo method for addressing uncertainty in land-use change models', *International Journal of Geographical Information Science*. Taylor & Francis, 32(11), pp. 2317–2333. doi: 10.1080/13658816.2018.1503275.

NASA and NGA (2001) ‘Shuttle Radar Topography Mission (SRTM) 30m Digital Elevation Data’. Washington: CGIAR Consortium for Spatial Information (CGIAR-CSI).

Nemenyi, P. . (1963) *Distribution-free Multiple Comparisons*.

O’Neill, B. C. *et al.* (2014) ‘A new scenario framework for climate change research: The concept of shared socioeconomic pathways’, *Climatic Change*. doi: 10.1007/s10584-013-0905-2.

O’Neill, B. C. *et al.* (2017) ‘The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century’, *Global Environmental Change*. doi: 10.1016/j.gloenvcha.2015.01.004.

OECD and FAO (2014) *OECD-FAO agricultural outlook 2014-2023*. Paris: OECD Publishing.

Ogle, S. M. *et al.* (2003) ‘Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agricultural lands between 1982 and 1997’, *Global Change Biology*, 9(11), pp. 1521–1542. doi: 10.1046/j.1365-2486.2003.00683.x.

Piepho, H. P. (2004) ‘An algorithm for a letter-based representation of all-pairwise comparisons’, *Journal of Computational and Graphical Statistics*. doi: 10.1198/1061860043515.

Pohlert, T. (2014) ‘The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR)’, *R package*.

Prestele, R. *et al.* (2016) ‘Hotspots of uncertainty in land-use and land-cover change projections: a global-scale model comparison’, *Global Change Biology*, 22(12), pp. 3967–3983. doi: 10.1111/gcb.13337.

Puig, D. (2015) *Uncertainty in greenhouse-gas emission scenario projections: Experiences from Mexico and South Africa*. Copenhagen, Denmark: UNEP DTU Partnership.

Python Software Foundation (2014) ‘Python language reference, version 2.7’. Python software foundation.

R core Team (2018) *R: A Language and Environment for Statistical Computing*, *Vienna, Austria*.

Renard, P., Alcolea, A. and Ginsbourger, D. (2013) ‘Stochastic versus Deterministic Approaches’, in *Environmental Modelling*. Chichester, UK: John Wiley & Sons, Ltd, pp. 133–149. doi: 10.1002/9781118351475.ch8.

Ross, C. W. *et al.* (2016) ‘Land use, land use change and soil carbon sequestration in the St. Johns River Basin, Florida, USA’, *Geoderma Regional*. Elsevier B.V., 7(1), pp. 19–28. doi: 10.1016/j.geodrs.2015.12.001.

Searchinger, T. *et al.* (2008) ‘Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change’, *Science*, 319(5867), pp. 1238–1240. doi: 10.1126/science.1151861.

Smith, P. *et al.* (2014) ‘Agriculture, Forestry and Other Land Use (AFOLU). In: Climate Change 2014: Mitigation of Climate Change’, *Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, p. 83. Available at: http://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/4_Volume4/V4_04_Ch4_Forest_Land.pdf.

Sobol’, I. M. (1993) ‘Sensitivity Estimates for Nonlinear Mathematical Models’, *Mathematical Modeling and Computational Experiment*. doi: 1061-7590/93/04407-008.

Verstegen, J. A. *et al.* (2012) ‘Spatio-temporal uncertainty in Spatial Decision Support Systems: A case study of changing land availability for bioenergy crops in Mozambique’, *Computers,* *Environment and Urban Systems*. Elsevier Ltd, 36(1), pp. 30–42. doi: 10.1016/j.compenvurbsys.2011.08.003.

Verstegen, J. A. *et al.* (2016) 'What can and can't we say about indirect land-use change in Brazil using an integrated economic - land-use change model?', *GCB Bioenergy*, 8(3), pp. 561–578. doi: 10.1111/gcbb.12270.

Warner, E. *et al.* (2014) 'Challenges in the estimation of greenhouse gas emissions from biofuel-induced global land-use change', *Biofuels, Bioproducts and Biorefining*, 8(1), pp. 114–125. doi: 10.1002/bbb.1434.

Wicke, B. *et al.* (2012) 'Indirect land use change: review of existing models and strategies for mitigation', *Biofuels*, 3(1), pp. 87–100. doi: 10.4155/bfs.11.154.

Woltjer, G. *et al.* (2014) 'The MAGNET model', (August), p. 148. Available at: http://edepot.wur.nl/310764.