Environmental Modelling & Software 97 (2017) 287–302

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

How a Pareto frontier complements scenario projections in land use change impact assessment

Judith A. Verstegen [a, b, *], Jan Gerrit Geurt Jonker [b], Derek Karssenberg [c], Floor van der Hilst [b], Oliver Schmitz [c], Steven M. de Jong [c], André P.C. Faaij [d]

[a] *Institute for Geoinformatics, University of Münster, Germany*
[b] *Copernicus Institute for Sustainable Development, Utrecht University, The Netherlands*
[c] *Department of Physical Geography, Utrecht University, The Netherlands*
[d] *Energy and Sustainability Research Institute Groningen, University of Groningen, The Netherlands*

ARTICLE INFO

Article history:
Received 5 May 2017
Received in revised form
1 August 2017
Accepted 14 August 2017
Available online 14 September 2017

Keywords:
Pareto frontier
Ethanol
Optimization
Brazil
Genetic algorithm
GHG emissions

ABSTRACT

To evaluate the sustainability of potential agricultural land developments, scenario projections with land use change models are often combined with environmental impact assessments. Although this allows inter-scenario comparison of impacts, it does not permit interpretation of scenarios in the light of theoretically optimal impacts. A Pareto frontier provides this information. We demonstrate this for ethanol production in Goiás, Brazil, in 2030. For a Business-as-Usual scenario projection, the spatial configuration, production costs, and GHG emissions of the production chain are compared with those obtained from spatial optimization and summarized by the Pareto frontier. Projected production costs are 729 $/m^3 ethanol, with GHG emissions of 40 kg CO_2-eq/m^3 ethanol. The Pareto frontier indicates an improvement potential of ~50 $/m^3 ethanol when keeping emissions fixed, or ~250 kg CO_2-eq/m^3 ethanol when keeping costs fixed. Robust locations having low costs and emissions show where and how improvements are reached, offering instruments for policy (re)design.

1. Introduction

Intensifying pressure on land, by e.g., requirements for producing food, feed, fibre and bioenergy, has stimulated debates about sustainable land use (e.g. Lambin and Meyfroidt, 2011; Seppelt et al., 2014). The spatial expansion of multiple land use types can be projected by using land use change models, given expected future demands for commodities (e.g. Fargione et al., 2010; Verstegen et al., 2016b). In such spatially explicit projections, it is common to use scenarios that allow for divergent future story lines, where each line is represented by a particular set of future trends in system drivers. We refer to this modelling approach as 'scenario projection'. Scenario projections are combined with environmental impact assessments of the projected land use changes to quantify the effects of the different story lines on the indicators associated with the impacts of interest. Examples of such indicators are greenhouse gas (GHG) emissions for climate impact, employment for socio-economic impact, and mean species abundance for ecologic impact.

Although scenario projection allows for comparing impacts among a set of scenarios, it has a distinct limitation: it gives no information on the overall optimality of the projection. In other words, a scenario projection does not indicate its position in the total indicator solution space (Seppelt et al., 2013). Thereby, it remains unclear if it is possible to attain lower impacts than those evaluated by the scenario(s), and, if so, how much lower (Fig. 1a). Our aim here is to show how adding spatial optimization to a spatial scenario projection allows for the assessment of how much a scenario can potentially be improved for a given set of impact indicators.

Spatial optimization is a contrasting method to assess the impact of land use change: it involves designing an optimal land use configuration with respect to one or more impact indicators of interest (e.g., GHG emissions, employment, and/or mean species abundance) given a range of boundary conditions (e.g. Almeida et al., 2016). Thus, this approach does not actually apply any land use change model. When optimizing multiple impact indicators (objectives) simultaneously, there is typically a very large number

* Corresponding author. Institute for Geoinformatics, University of Münster, Germany.
E-mail address: j.a.verstegen@uni-muenster.de (J.A. Verstegen).

http://dx.doi.org/10.1016/j.envsoft.2017.08.006
1364-8152/

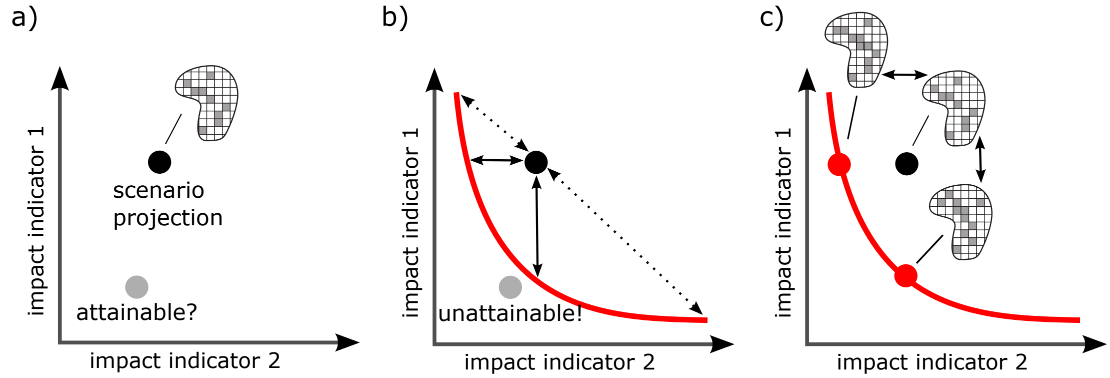

Fig. 1. A hypothetical land use change scenario projection for which two impact indicators are assessed, 1 and 2. For both impact indicators, low values are desired. (a) Scenario projection with spatial configurations projected by the land use change model (maps), the impact indicator values (black dot) derived from these, and the question of if (and by how much) lower impact indicator values are attainable (grey dot). (b) Scenario projection and calculation of Pareto frontier (red line) with an assessment of the maximum impact reduction (dotted arrows) and constrained impact reduction (solid arrows). (c) Comparison of spatial land use configurations associated with points on the Pareto frontier and calculated by the scenario projection, to identify required spatial reorganization of scenario projections to minimize impacts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of optimal land use configurations (solutions); together, these optimal solutions form a Pareto frontier (Seppelt et al., 2013). Because for all these solutions it is impossible to improve one objective without impairing another, the best solution from the Pareto frontier depends on the weighting of the impact indicators. As such, the Pareto frontier shows trade-offs between the different impact indicators.

It is hypothesized that combining spatial scenario projection and spatial optimization would provide deeper insights in the solution space of future land use, because it enables the scenario projections to be interpreted in the context of the optimal solutions given by the Pareto frontier. In particular, this approach is expected to provide information on:

1. The performance of a scenario is in terms of impact indicators, and how much each impact indicator value of the scenario projection can theoretically be improved. The maximum and constrained improvement can be calculated, where 'constrained improvement' is calculated by keeping the other impact indicator(s) at the same value as in the scenario projection (Fig. 1b) and;
2. Where and how the land use projected by the scenario should be reorganized to reach these improvements (and where it should not be reorganized) (Fig. 1c).

These two aspects are beneficial for policy making, because they assist in developing scenarios with an increased performance by quantifying improvement opportunity and identifying land use characteristics that lead to these improvements. On the other hand, a Pareto frontier alone, without scenario projection results, does not provide information about the feasibility of reaching these improvements given the current land use system dynamics and current policy instruments captured in the scenario. For example, Cotter et al. (2014), used spatial optimization to design a sustainable land use scenario, but they did not have a Business-as-Usual (BAU) scenario of land use change, thereby missing information on which part of the optimal future configuration is likely to be attained by current dynamics and policies. If it is known to what extent the BAU is sustainable, then one can assess in what way policies should be *re*designed.

Some recent studies have used spatial optimization to optimize certain parameters of scenarios (e.g. Arancibia et al., 2016; Law et al., 2017), whereby this integration of optimization and scenario projection allowed for selection of the optimal scenario. However, such integration still cannot ascertain if the retrieved impacts are the lowest attainable, as the lowest impacts might not be attainable using any of the options defined as the scenario parameters. Seppelt et al. (2013) agree with this sentiment: they argue that it is not the integration but the combination of scenario projection and optimization that can strengthen efficient decision making for sustainable land use. Yet, as a very limited number of case studies exists with which to investigate this approach (e.g. Gaddis et al., 2014), and none exists in land use change, our paper aims to provide such a case study.

At this time, biofuels are, rightly or unrightly so, at the centre of the debate around sustainable land use (Tempels and Van den Belt, 2016). Because of this, we performed an impact assessment of ethanol production from sugar cane for 2030, in Goiás, Brazil, a region with a large expected increase in ethanol production (e.g. Lapola et al., 2010). In line with other biofuel impact assessments (e.g. Akgul et al., 2012; Arancibia et al., 2016), we consider two impact indicators: ethanol production costs, as an indicator of economic competitiveness, and greenhouse gas (GHG) emissions, as an indicator of the potential to mitigate climate change. Our research questions for this case study are: What is the performance of and improvement potential for a BAU scenario projection in terms of impact indicator values? How can an assessment of the spatial differences between the projected and optimized land use configurations explain the performance of the BAU scenario projection? How can these spatial differences be used to (re)design land use policies?

We evaluated the way in which the configuration of the ethanol production chain impacts ethanol production costs and GHG emissions for the sum of the ethanol production chain's four main components: acquisition and preparation of land for sugar cane production, sugar cane cultivation and harvest, sugar cane transport to the production facilities (mills), and conversion from sugar cane to ethanol. First, we performed an impact assessment on a BAU scenario projection of the expansion of sugar cane fields and mills for 2030, using an existing land use change model (Verstegen et al., 2016b; Jonker et al., 2016); the impact indicators were calculated through a post-analysis on the configuration of sugar cane fields and mills for 2030. Next, we performed a separate impact assessment via the Pareto frontier for the two impact indicators, calculated through spatial optimization, also for 2030. In this assessment, we optimized the locations of the sugar cane fields and the locations and processing capacities of the mills using a genetic algorithm (Li and Yeh, 2005). Finally, we compared the first and second impact assessment results in terms of impact indicator values, as in Fig. 1b, and in terms of spatial land use configurations, as in Fig. 1c. Policy redesign recommendations were derived from these comparisons.

2. Methods

2.1. Case study description

Brazil started producing ethanol from sugar cane at the beginning of the 20th century, motivated by a gasoline import burden and a sugar production surplus (Walter et al., 2014). In the harvest season 2013/2014, Brazilian ethanol production reached 27.5 million m^3 (UNICA, 2015). The region in Brazil with the second highest production, with 3.9 million m^3 (UNICA, 2015), was the state Goiás, which recently experienced a fast growth of sugar cane area (Adami et al., 2012). Goiás is 340 000 km^2, roughly the size of Germany. It mainly consists of a plateau with a tropical climate, characterized by a vast area of planted pastures, especially in the west. Furthermore, there is a sizeable patch of cropland in the southwest, with mainly soy, corn and sugar cane (IBGE, 2013). Most forest areas are found in the centre north part, which has a more mountainous character.

In our study, we assume a total ethanol supply increase for Goiás of 10.2 million m^3 for 2030. This is derived from the total production of sugar cane in 2030 projected by the land use change model (see next section) minus the total production in the initial land use map from 2006 (Fig. 2), and an assumed conversion efficiency of 0.09 m^3 ethanol/tonne cane (Jonker et al., 2016). In both the scenario projection and the optimization, new sugar cane fields cannot be allocated on raster cells that are urban, water or sugar cane in the initial land use map (placing 'new' sugar cane fields over existing ones would generate no additional ethanol compared to the initial situation). Furthermore, it is assumed that sugar cane present in the initial land use map goes to existing mills; this is not remodelled.

2.2. Scenario projection of land use change for ethanol production

For the scenario projection, we build upon an existing land use projection for Brazil from an integrated economic–land use change model (Verstegen et al., 2016b). The economic model determines the total demands for all commodities, including ethanol, in 15 world regions, one of them being Brazil. The land use change model projects the expansion and contraction of a set of land use types in Brazil based on these demands and suitability factors that vary per land use type. The land use change model is calibrated between 2006 and 2012 and provides a scenario projection, with input from the economic model, between 2012 and 2030. The suitability factors for sugar cane are, with their calibrated median weight in parentheses, the area of sugarcane fields in an extended Moore neighbourhood (0.29), travel time to existing mills (0.28), potential yield (0.22) and the conversion elasticity from other land use types to sugarcane cultivation (0.21) (Verstegen et al., 2016b). The land use change model has a 5 × 5 km^2 resolution. Therefore, we employ this resolution in our Goiás case study as well (Fig. 2). We use the projected positions of the new sugar cane fields for ethanol production in Goiás for 2030 from the BAU scenario that includes an increased ethanol supply due to current and planned ethanol mandates worldwide. The word 'new', in this context, means allocated at any point in time between the initial year of the simulation (2006, Fig. 2) and 2030.

Land use change models, including our model, typically do not project the processing facilities of the cultivated crops, in our case the sugar cane mills. Therefore, we append results from Jonker et al. (2016) who have determined the number, scales and locations of ethanol production mills for 2030, based on the minimization of production cost. This is done with a mixed-integer linear programming model, while fixing the locations of the sugar cane fields as projected by the BAU scenario used here. As such, part of our scenario projection is optimized, because we considered this the best way to append the mills to our land use results. Although the optimization involved only the mills and was based on production costs only, the BAU results may be too optimistic. The effect is expected to be small because, compared to the locations of the mills, the locations of the sugar cane fields have a much larger impact on the two impact indicators used in this study (Jonker et al., 2016).

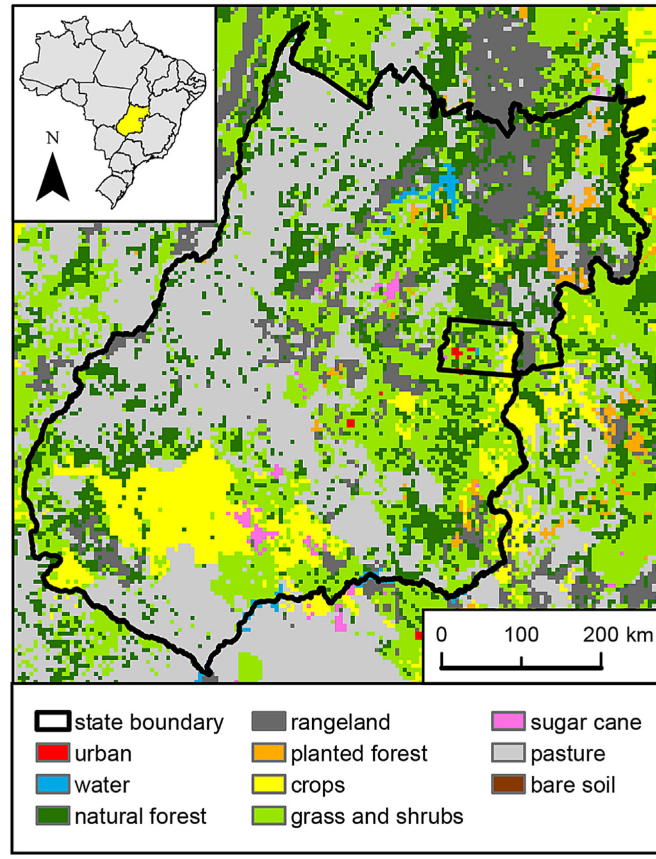

Fig. 2. Land use map of Goiás, Brazil for 2006 at a 5 × 5 km^2 resolution (adapted from Verstegen et al., 2016b). The inset shows all states in Brazil with Goiás indicated in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.3. Impact assessment

Both impact indicators were calculated for the four main components in the ethanol production chain (see Introduction). Note that we calculate costs and emissions at the 'factory gate', meaning that the revenues from selling the ethanol and the avoided emissions that come from ethanol replacing fossil fuels are not included. Total production costs, c ($/$m^3$ ethanol),[1] are:

$$c = c_l + c_c + c_t + c_p \quad (1)$$

[1] Throughout the manuscript $ means US$_{2014}$, unless explicitly stated otherwise.

In equation (1), c_l (\$/m^3 ethanol) denotes the land costs, involving the acquisition costs and the costs to turn the initial land use into sugar cane production; c_c (\$/m^3 ethanol) are the cultivation costs, consisting mainly of fertilizer, labour and machinery costs; c_t (\$/m^3 ethanol) are the transport costs, being the diesel costs of the truck, subject to the distance, speed and road type(s) between the fields and the processing mills; and c_p (\$/m^3 ethanol) are the processing costs, depending on the processing capacity (scale) of the mill. Correspondingly, total emissions e (tonne CO_2-eq/m^3 ethanol) are:

$$e = e_l + e_c + e_t + e_p \tag{2}$$

In equation (2), e_l (tonne CO_2-eq/m^3 ethanol) are the land use and land management change emissions, depending mainly on the carbon stocks of the replaced land use type (method IPCC, 2006); e_c (tonne CO_2-eq/m^3 ethanol) are the cultivation emissions, being fertilizer and machinery emissions; e_t (tonne CO_2-eq/m^3 ethanol) are the transport emissions, being the diesel emissions from the truck travelling between the fields and the processing mills; and e_p (tonne CO_2-eq/m^3 ethanol) are the processing emissions, a fixed amount per m^3 ethanol produced.

The land, cultivation, transport, and processing costs and emissions are averages over the whole study area. The first three components are calculated by averaging the spatially explicit costs/emissions (varying over space because of former land use type, fertilizer application, distance to the mill etc.) over all sugar cane fields. The processing costs are calculated by averaging the spatially explicit costs/emissions (varying over space because of the scale of the mill) over all ethanol mills. The methods to calculate the cost and emission components are based on Jonker et al. (2015, 2016). The details of these calculations, parameter values, and sources are provided in Appendix A.

For the scenario projection, the impact indicators were calculated through a post-analysis on the configuration of sugar cane fields and mills for 2030. For the optimization approach, the impact indicators were minimized for 2030 by optimizing the configuration of sugar cane fields and mills for 2030, as explained in the next section.

2.4. Optimization of impacts and calculation of the Pareto frontier

The two impact indicators of ethanol production needed to be minimized to find the Pareto frontier that shows the trade-offs between the two. One way to stimulate reduction of GHG emissions for agricultural products is to charge the producer for these emissions using a carbon price (Smith et al., 2008; Chen et al., 2012). When a carbon pricing system is established, the two objectives of minimizing production costs and GHG emissions can be combined into a single objective:

$$x = c + e \cdot p \tag{3}$$

In equation (3), x (\$/m^3 ethanol) are the aggregate costs (production costs plus GHG costs) that we aim to minimize and p (\$/tonne CO_2-eq) is the carbon price. So, x is the performance criterion, i.e. the objective value, for the optimization. The Pareto frontier between the production costs (c) and GHG emissions (e) of ethanol is found by minimizing the aggregate costs (x) for a set of different carbon prices (p). The Pareto frontier allows policy makers to evaluate the effect of carbon price on the trade-offs between costs and emissions.

For a given ethanol supply increase d (m^3 ethanol) and carbon price p, the aggregate costs x are controlled by 1) the locations of the sugar cane fields, 2) the number of processing mills and their scale (processing capacity), and 3) the locations of these mills (see previous section and e.g. de Meyer et al., 2014). The relationship between the aggregate costs and each of these three control variables is non-linear and the set of potential solutions is too large for exhaustive search. For such cases, it is advantageous to use metaheuristics, which is a set of methods designed to find near-optimal solutions in an acceptable computation time using a smart search through the solution space (Blum and Roli, 2003). We use a genetic algorithm (GA), because this metaheuristic has been proven to generate good results for optimization problems like ours (e.g., Li and Yeh, 2005; Stewart et al., 2004). A GA mimics the process of natural selection in a population of solutions, called individuals, similar to the set of samples in a Monte Carlo approach (Bennett et al., 1998). The genome of the individuals represents the setting of the control variables of the optimization task. This population evolves through a given number of iterations, called generations, towards better solutions in terms of the performance criterion. The best performing individual of the final evolved population is the optimal solution. Appendix C provides a more elaborate explanation of how the GA works. It is not feasible to simultaneously optimize the locations of all fields, the number of mills, and all locations and scales of these mills. Such a large number of variables for a large case study area becomes too computationally intensive for a GA (Haupt and Haupt, 2004). Therefore, we took a sequential approach.

The genome of the GA represents the coordinates of a fixed number of mills M (Fig. 3). For each individual in each generation, the following method was applied to calculate the objective value x. Once the locations of the mills are set by the GA, sugar cane fields are placed at the locations with the lowest $(c_l + c_c + c_t) + p \cdot (e_l + e_c + e_t)$ until the total supply 10.2 million m^3 is met.[2] The total area of sugar cane fields can differ per individual in the GA, i.e. per solution, depending on whether sugar cane is allocated on high yielding (small area) or low yielding (large area) locations. When the fields are allocated, it is known which field delivers sugar cane to which mill (lowest $c_t + p \cdot e_t$), and thus the scale of each mill (tonne cane/year) can be calculated by summing the supply of all its fields. In this process, it is ensured that the scale of each mill cannot exceed the maximum attainable scale S (tonne cane/year). If no fields are assigned to a mill, it is not considered in the calculation of the objective value. Now, because all control variables are known, the objective value x can be calculated. Next, the GA adapts the locations of the mills of a selected fraction of the population and the process is repeated for the next generation (Fig. 3).

The optimization was implemented in the Python programming language (Python software foundation, 2014) using the AMORI software (AMORI, 2009) for the GA and the PCRaster Python framework (Karssenberg et al., 2010) for the placement of mills and fields and the calculation of the objective value. The number of mills placed, M, is 30 and each mill has a maximum scale, S, of 5.5 million tonne cane/year (Jonker et al., 2016). The GA was run for five different carbon prices p of 0, 10, 100, 200 and 400 \$/tonne CO_2-eq.[3] In addition, it was run to optimize emissions only, to get the minimum attainable emissions (minimum attainable costs are reached at a carbon price of 0 \$/tonne CO_2-eq).

We used a population of 1000 individuals and stopped the GA at generation 24. The population fraction to reproduce was set to 0.1 and the mutation rate to 0.3. During reproduction, individuals were split at two locations in the bit-string of the genome. The maximum number of bits to mutate in a single individual was two. These GA

[2] Processing costs cannot be calculated yet, because they depend on the scale of the mill, which can only be determined once the fields are allocated.

[3] No carbon pricing system is currently installed in Brazil and it is unknown if and when it will be installed (Dahan et al., 2015).

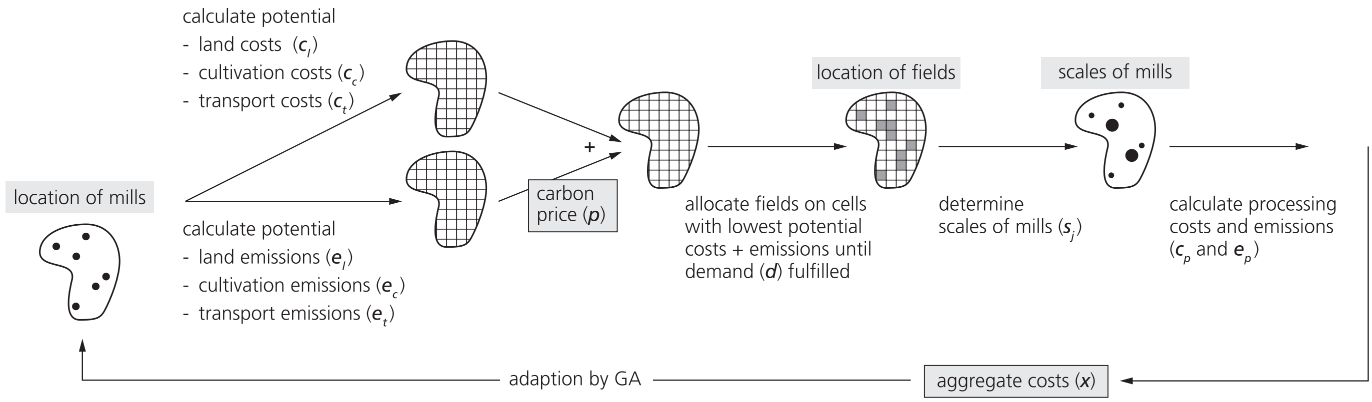

Fig. 3. Conceptual model of control variables and calculation of the objective value (aggregate costs).

settings were determined by performance tests as shown in Appendix C.

3. Results and discussion

3.1. What is the performance of and improvement potential for a BAU scenario projection in terms of impact indicator values?

For a supply increase of 10.2 million m^3 ethanol in Goiás in 2030, the BAU scenario projection has production costs of 729 $/$m^3$ ethanol and GHG emissions of 40 kg CO_2-eq/m^3 ethanol (Fig. 4). The scenario projection indicates that the region of Goiás will meet the emission savings required by the Renewable Energy Directive (RED)[4] (European Parliament and Council of the European Union, 2009).

The minimum attainable emissions, calculated from the optimization, are −399 kg CO2-eq/m3 ethanol, i.e. carbon sequestration of 399 kg CO2-eq/m3 ethanol, and the minimum attainable production costs (excluding GHG costs) are 656 $/m3 ethanol (Fig. 4). These minimum costs are similar to the 520 US$2010/m3 (approximately 650 US$2014/m3 ethanol) calculated for Brazil by Jonker et al. (2015).

The BAU scenario projection does not lie on the Pareto frontier, meaning that under the current land use system dynamics, the projected land use for 2030 is suboptimal in terms of production costs and GHG emissions. Looking at the minimum attainable costs, the projected production costs can be reduced at most by 73 $/$m^3$ ethanol (dotted arrow in Fig. 4); this is a cost improvement of 11%, but the trade-off is that emissions will become 20 times as high. The maximum possible amount by which emissions of the BAU scenario can be reduced is 439 kg CO_2-eq/m^3 ethanol. As percent improvement is a non-informative measure for GHG emissions, because the measurement scale of GHG emissions is of the 'balance' type (unbounded, positive and negative values). Therefore, we can report the percentage of shift in the total range of the y-axis of the Pareto frontier, which is about 37% for the case of maximum reduction in emissions (for maximum reduction in production costs, this value is 41%). This means that for the two impact indicators, the potential shift is almost equally large in view of the total solution space.

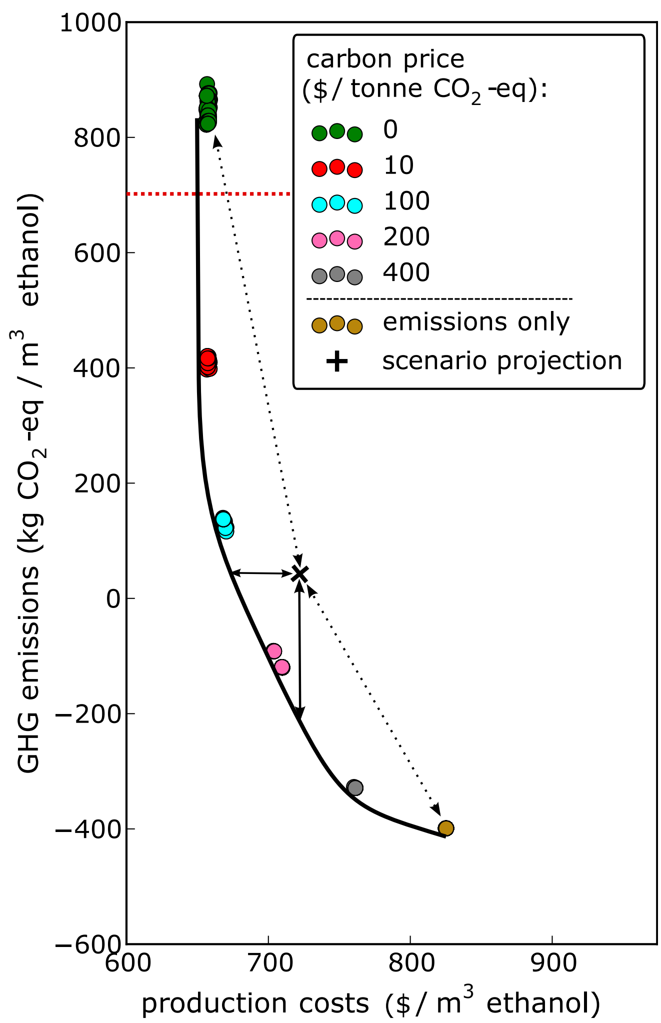

Fig. 4. Scenario projection and optimization results, with the Pareto frontier (black line) between production costs and GHG emissions for a supply increase of 10.2 million m^3 ethanol in Goiás in 2030. For each carbon price (0, 10, 100, 200 and 400 $/tonne CO_2-eq) and the optimization on emissions only, the best 10% of the final GA population is plotted, where each individual is indicated by a coloured circle. For both impact indicators, the maximum impact reduction (dotted arrows) and constrained impact reduction (solid arrows) of the scenario are visualized. The red dotted line is the maximum GHG emission value allowed for biofuels produced in installations in which production started on or after 1 January 2017 according to the Renewable Energy Directive (RED) (European Parliament and Council of the European Union, 2009). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[4] Note that RED also requires inclusion of GHG emissions from the distribution of ethanol (transport from the mills to the customers). This is not included in our study, but we expect its contribution to be small because the emissions of sugar cane transport are already small and the ethanol has a much higher energy density than sugar cane, and, therefore, lower GHG emissions per m3 ethanol (e.g. Hamelinck et al., 2005a; Wang et al., 2014).

The constrained improvement (Fig. 4, solid arrows), i.e. what can be gained in one indicator, without causing an increase in the projected value of the other, is a reduction of about 50 $/m^3 ethanol for production costs. This is a 7% improvement compared to the BAU scenario and a 28% shift in the total range of the x-axis of the Pareto frontier. The constrained improvement for GHG emissions is a reduction of about 250 kg CO_2-eq/m^3 ethanol compared to the BAU scenario, which is a 21% shift in the total range of the y-axis of the Pareto frontier. So, even the constrained ('no-loss') improvements are still significantly better than the BAU scenario.

3.2. How can an assessment of the spatial differences between the projected and optimized land use configurations explain the performance of the BAU scenario projection?

Of all optimization results, the BAU scenario projection is closest to those for carbon prices of 100 and 200 $/tonne CO_2-eq in terms of ethanol production costs and GHG emissions (Fig. 4). Yet, the spatial land use configurations of the scenario projection and these two optimization results are very different (Fig. 5, panels b, c and d). To quantify this, comparing these two optimizations with the BAU scenario regarding the locations of the sugar cane fields results in a Kappa statistic (Cohen, 1960) of 0.11 for the results of 200 $/tonne CO_2-eq and 0.07 for the results of 100 $/tonne CO_2-eq, where 1 means full agreement with the scenario projection and 0 is the agreement obtained by random allocation. Because the locations of the fields in the optimizations and the BAU scenario are so dissimilar, the mills are in very different locations as well.

Besides comparing the actual locations of the optimized and projected sugar cane fields and ethanol, we can categorize the locations in terms of their performance for the impact indicators and compare the optimized and projected results in terms of these categories. Locations can be categorized as being in use for ethanol production at (Figs. 5 and 6).

1. low production cost *and* low GHG emission points on the Pareto frontier, *and* for the scenario projection. Examples is the municipality Chapadão do Céu.[5]
2. low production cost *and* low GHG emission points on the Pareto frontier *but not* for the scenario projection. An example is Cristalina.
3. low production cost *or* low GHG emission points on the Pareto frontier *and* for the scenario projection. Example are Rio Verde, for low GHG emission points and São Luiz do Norte for low production cost points.
4. low production cost *or* low GHG emission points on the Pareto frontier *but not* for the scenario projection. An example is Silvânia, for low GHG emission points.
5. the scenario projection, *but not* for any of the points on the Pareto frontier. An example is Goiatuba.
6. *neither* the scenario projection, *nor* any of the points on the Pareto frontier.

Type 1 and type 2 locations are 'robust'. Ethanol produced here has economic competitiveness (low production costs) and climate change mitigation potential (low GHG emissions), a win-win situation. These locations are optimal under both high and low future carbon prices.

Robust locations all have high yields (Figs. 5 and 6). Both production costs and GHG emissions are reduced if sugar cane is produced in high-yielding areas, because this requires smaller areas for the production of the same amount of ethanol. A smaller area reduces costs, because some of the land costs and cultivation costs are area dependent (Appendix A). An example of area-dependent costs are machinery costs. The harvest machines have to drive over the total area of land, independent of how much is harvested per hectare. A high yield indirectly reduces the machinery costs per liter of ethanol. A high yield also reduces GHG emissions because a smaller sugar cane area means less land use change, and thus less land emissions (the largest component of the total emissions). In addition, cultivation emissions are also area dependent because the machinery has constant emissions per hectare.

The second characteristic of robust locations are large areas of cropland and/or grass and shrub land in 2006 (Figs. 5 and 6). Both production costs and GHG emissions are reduced when former cropland is used for sugar cane production. Production costs are reduced because the costs to prepare the land are low, as the land is already in use for agriculture (no trees need to be cut etc.). GHG emissions mainly depend on the type of land use replaced by sugar cane (land emissions). When the replaced land use is cropland, carbon is sequestered, because cropland consists mostly of annual crops, having a low permanent carbon stock as they are harvested each year. When the replaced land use is grass and shrubs, emissions are low, because sugar cane is similar to grass in terms of carbon stock.

Although production costs are reduced when sugar cane is cultivated on former cropland and/or grass and shrub land, the relative effects are the largest for the GHG emissions. This is shown by the selected locations of type 3 and 4, Rio Verde and Silvânia. These municipalities had large areas of cropland and/or grass and shrub land in 2006, but relatively low yields. Therefore, the optimization suggests producing ethanol here under carbon prices of 200 $/m^3 and above, but not under lower carbon prices, because the low yield increases the production costs too much. For yield, on the other hand, the relative effects are the largest for the production costs. This is shown by the selected location of type 3 São Luiz do Norte. This municipality has high yields, but had no cropland in 2006. Therefore, the optimization suggests producing ethanol here under carbon prices up to 200 $/m^3, but not under higher carbon prices, because not enough carbon can be sequestered.

The GHG emissions of the BAU scenario fall in between emissions reached at carbon prices of 100 and 200 $/tonne CO_2-eq as calculated by the optimization approach (Fig. 4). We expected that, as there is currently no carbon pricing system installed, the BAU scenario projection would have emissions comparable with the optimization results for a zero carbon price. But the projected GHG emissions are 720 kg CO_2-eq/m^3 ethanol lower than that. A possible explanation is that, although no carbon pricing system is currently installed in Brazil, ethanol producers in the current land use system take into account GHG emissions through other sustainability regulations. The RED is not likely to have an effect, because it is a European regulation and almost all Brazilian ethanol is currently used domestically. But there are some Brazilian sustainability regulations, such as the federal ecological zoning for sugar cane and certification of sustainably produced ethanol (e.g. Aguiar et al., 2011). Although such regulations were not explicitly included in the land use change model, they might have been implicitly captured in the model structure through calibration using historic data. This is a major difference between optimization, in which an explicit goal function is formulated, and scenario projection, in which past behaviour of both known and unknown processes is captured in the model structure through calibration.

[5] The official name of the municipality is Chapadão do Céu. The short name Chapadão is used in Fig. 5.

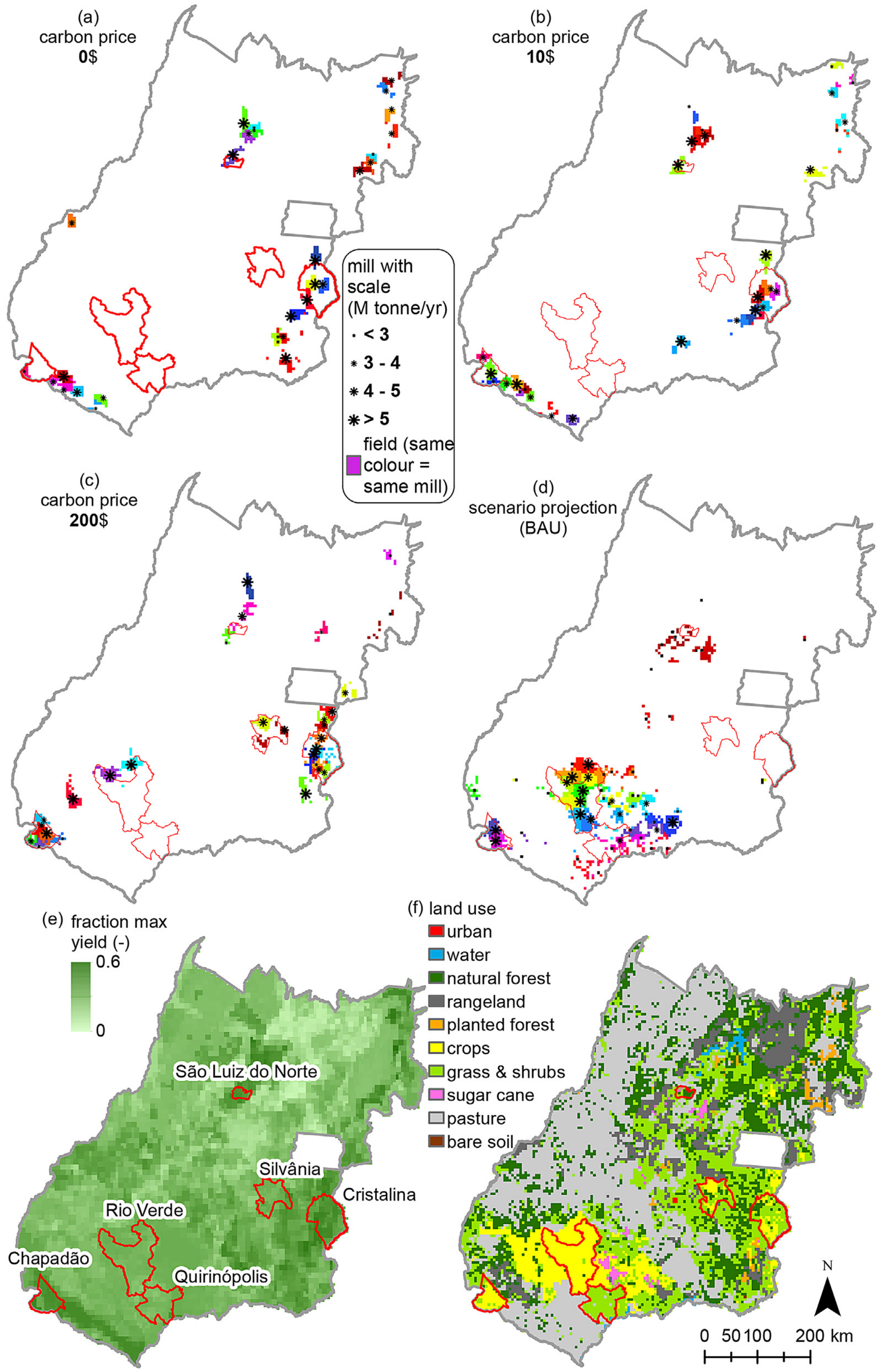

Fig. 5. Spatial patterns of sugar cane fields and mills belonging to the best-performing individual for different carbon prices of (a) 0, (b) 10, and (c) 200 \$/tonne CO_2-eq, and for (d) the scenario projection. Fields with the same colour deliver to the same mill. At the bottom the input maps (e) fraction of the maximum attainable yield (Tóth et al., 2012) and (f) initial land use (2006) (Verstegen et al., 2016b). The municipalities with red boundaries are evaluated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

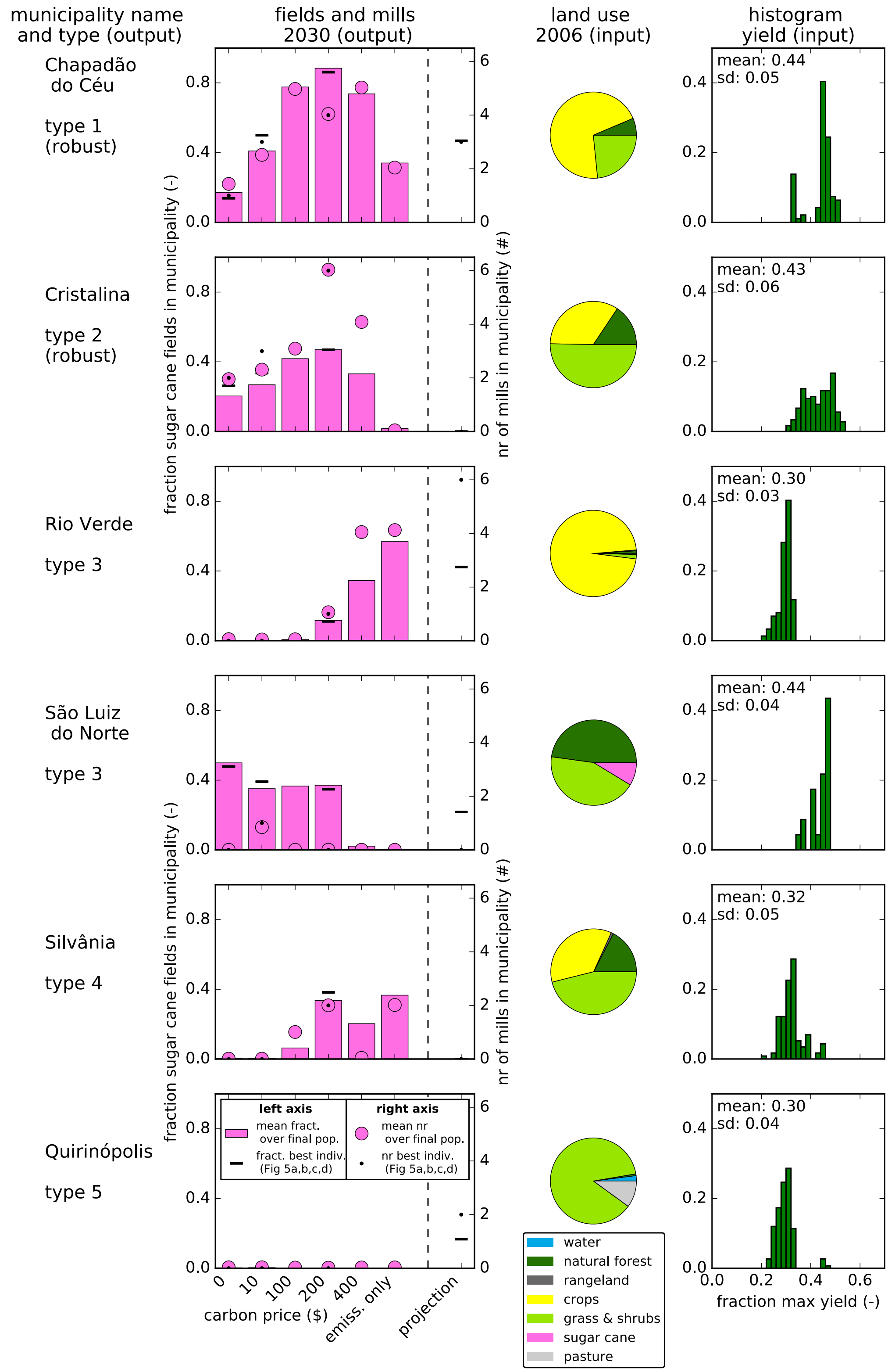

Fig. 6. Overview of the categorization and characteristics of the spatial outputs (of the optimization and scenario projection, separated by a dashed vertical line) and inputs for six example municipalities in Goiás. The locations of the municipalities can be found in Fig. 5. Left: filled bars (fraction sugarcane fields) and filled circles (number of mills) show the mean results over the total final population of the GA, for each carbon price; black horizontal lines (fraction of sugarcane fields) and black dots (number of mills) are the results of the best-performing individual (Fig. 5a,b,c) and the results of the projection (Fig. 5d). Centre: land use in 2006 (Fig. 5f) for each municipality. Right: distribution of the fraction of the maximum attainable yield (Fig. 5e) of all cells in each municipality.

3.3. How can these spatial differences be used to (re)design land use policies?

Policy redesign should be focused on directing ethanol production away from the bad-performing locations in the projected situation (Fig. 5d) towards locations in the optimized solutions (Fig. 5a, b, or c), where the choice between the optimized solutions depends on Brazil's aims regarding the minimization of production costs and/or GHG emissions.

Type 1 locations require, in theory, no policy redesign, as the BAU scenario projects ethanol to be produced here in 2030 with current policies. At type 2 locations, the BAU scenario does not project sugar cane fields and mills. It is advantageous to design policies to stimulate ethanol production here independent of the current political preference for economic competitiveness or climate change mitigation potential as there will be a profit regarding both impacts.

The case study results can be used to design policies that stimulate ethanol production in robust locations with roughly two different types of instruments. The first instrument is to give farmers credits if they certify the product came from a specific location. For example, this is currently done for soybean farmers: Banco do Brasil requires soybean farmers who apply for credit to certify the origin of their soybeans in order to assure that no deforestation took place to grow the soybeans (The Nature Conservancy, 2012). A similar location check could also be done for sugar cane ethanol, using location types 1 and 2, and potentially 3 and 4, when there is a preference for minimizing one of the two impact indicators. The second instrument is to analyse the characteristics of robust locations and then incentivise producers to build mills and farmers to start producing sugar cane for ethanol in locations with these characteristics. This is currently done in Brazil for a single impact indicator: the Low-Carbon Agriculture (ABC) programme, launched in 2010, provides low-interest rural credit for the implementation of agricultural practices, such as the restoration of degraded grasslands, to reduce GHG emissions (Newton et al., 2016).

Besides being used to create policies to stimulate sustainable ethanol production, the spatial configuration results can be used to (re)design policies for preventing *un*sustainable ethanol production. The same two instruments can be applied, but now focusing on locations of type 5, where ethanol production is sub-optimal for both impact indicators. Preventing ethanol production there can be targeted by location through the installation of national parks or agro-ecological zones. Or it can be targeted by fining or banning the purchase of commodities produced at high-emission locations (or at high-cost locations, but it is unlikely that farmers want to produce there anyway), as is done via the soy moratorium (Gibbs et al., 2015). If one of the two impact indicators is more important than the other, locations of type 3 and 4 have to be taken into account in the policy design as well. Locations of type 6 are of little concern, because sugar cane expansion is not projected to occur there, so little effort has to be made to prevent expansion there. Still, one should assess (before implementation, using scenarios including the new policies and the land use change model) and monitor (after implementation) that the redesigned policies do not cause expansion here (Gibbs et al., 2015).

3.4. Limitations and uncertainties of the case study

Three main sources of uncertainty exist in our study: uncertainty in the inputs, uncertainty in the process representations, and the capacity of the GA to find the optimum. Regarding input uncertainty, we have seen in the results that the improvement potentials and the robust locations are sensitive to the two crucial input maps: potential yield and initial land use. Especially for yield, the uncertainty is high, because we use a global yield map (Tóth et al., 2012), including little local knowledge. It is likely that the results are also sensitive to other input parameters in the impact assessment, such as the time horizon over which the land emissions are divided (see Appendix A, equation (A.6)). In theory, one could use error propagation analysis, using Monte Carlo, to assess these uncertainties. But since one optimization run, i.e. a single Monte Carlo sample, costs 28 h, this is hardly doable. Input uncertainties in the land use change model have been assessed and reported in Verstegen et al. (2016b).

Regarding process representation uncertainties, the GHG emission calculations in the impact assessment, especially land emissions, are relatively crude. But the advantage of using this crude method (IPCC, 2006) is that it is used in most emission impact assessments, which makes our results at least comparable. Another limitation regarding the process representation of our case study is that indirect land use change (iLUC) is not taken into account. ILUC is the cascading effect of a land use change: for instance, when sugar cane is allocated on land previously used for crop cultivation, the crop production has to be moved elsewhere in order to sustain the supply for this crop (Wicke et al., 2012). ILUC is likely to cause GHG emissions, additional to the direct GHG emissions. Since these emissions are (indirectly) caused by the land allocation to sugar cane, it could be argued that they should be included in the commodity's Pareto frontier and scenario projection. Incorporating iLUC requires one to determine where the displaced land use type reappears. This entails the use of either a spatial land use change model determining the dynamics of all displaceable land use types (e.g. Verstegen et al., 2016b) or a combined optimization model for ethanol production and all other commodities (e.g. Lautenbach et al., 2013). Both options are data intensive and therefore require massive computation times. And, modelling Goiás only is not sufficient, because the reallocation can take place anywhere in the world. Yet, it is obvious that if iLUC could be included in the optimization of the BAU scenario, 1) it is very likely that higher GHG emissions would be obtained and 2) placing sugar cane on croplands would become less favourable, because the crop production has to be moved to elsewhere and eventually natural vegetation will probably be converted.[6] This means that the position and shape of the Pareto frontier would change, as well as the position of the scenario projection relative to the frontier. This might change the improvement potentials and the robust locations identified in our case study. The omission of iLUC in the calculation is of less relevance when working with a scenario in which iLUC prevention is ensured through policies (e.g. Gerssen-Gondelach et al., 2017). Another option to avert (but not prevent) iLUC is to include local food security as an additional impact indicator.

Another potential process representation limitation to this study is that our BAU scenario projection assumes that current trends in the land use system will continue into 2030, while in reality a system can change quite abruptly (Verstegen et al., 2016a). Therefore, the low GHG emissions of the scenario projection are uncertain and by no means undermine the potential benefits of a carbon pricing system in Brazil. Another uncertainty is in the ability of the GA to find the optima. A GA is a metaheuristic, meaning that

[6] The area of natural vegetation eventually converted is often not equal to the area of sugar cane allocated, for example because the location where the agricultural land has been moved, has a higher or lower productivity or different management practices. Or because of price effects, which can only be determined if an economic equilibrium model is used. These things are additional complications when trying to include iLUC in the optimization.

finding the optimum is not guaranteed. Yet, we are confident that we have found near-optima, because 1) the spatial configurations of the individuals in the final population were similar to each other yet not the same (Fig. 6), indicating proper convergence, and 2) running the GA multiple times with the same settings always resulted in almost the same value of the objective function (Appendix C).

Finally, the optimization technique we used is relatively simple. We combined the two impact indicators into a single objective using a carbon price. It would be interesting to assess more impact indicators, such as mean species abundance for ecologic impact, or employment for socio-economic impacts, or food security for local market impact. With more objectives, the conversion of all objectives to a common unit becomes harder. To solve optimization problems of this nature, more complex, multi-objective optimization methods exist (see e.g. Lautenbach et al., 2013), and these methods result in a Pareto frontier of two or more dimensions. So, despite the limitations and uncertainties of our case study, we believe to have shown how the addition of a Pareto frontier to a land use change scenario projection allows one to assess the potential for improvement in different impact indicators.

4. Conclusion

In the future, commodity production chains will ideally combine economic benefits with positive environmental impacts. This study aimed to show how the value of an impact assessment performed via scenario projection can be increased by adding a Pareto frontier, because together they provide information on the attainability of that ideal. To illustrate this, production costs and greenhouse gas (GHG) emissions were assessed given a projected increase in the ethanol supply of 10.2 million m^3 ethanol in the state Goiás, Brazil, in 2030. This was done by complementing the results of a Business-as-Usual (BAU) scenario projection (Verstegen et al., 2016b) with the results of a spatial optimization performed using a genetic algorithm (GA). The spatial optimization was carried out for different carbon prices (\$/tonne CO_2-eq) to establish different weights for the production costs and GHG emission objectives, together forming the Pareto frontier.

The BAU scenario projection of land use change in Goiás towards 2030 resulted in production costs and GHG emissions of 729 \$/m^3 ethanol and 40 kg CO_2-eq/m^3 ethanol, respectively. The Pareto frontier addition allowed for the analysis of two pieces of new information. First, we were able to analyse to what degree the BAU scenario projection could be improved in terms of production costs and GHG emissions of ethanol production. The constrained improvement potential (keeping the other impact indicator(s) at the same value as in the BAU scenario projection) for ethanol production costs was found to be a reduction of about 50 \$/m^3 ethanol, and for GHG emissions a reduction of about 250 kg CO_2-eq/m^3 ethanol. Maximum improvements, i.e. without trying to keep the other indicator constant, were found to be a reduction of 73 \$/m^3 ethanol for production costs (46% additional reduction) and a reduction of 439 kg CO_2-eq/m^3 ethanol for GHG emissions (76% additional reduction).

Second, we assessed the spatial differences between the optimized and projected land use configurations in order to design and suggest policies to reach these improvements. The optimization results are able to identify robust ethanol production locations that perform well for all impact indicators, in our case production costs and GHG emissions. If no ethanol production is projected by the scenario at these locations (as was the case for many robust locations in our case study), policies need to be redesigned to stimulate production there, resulting in reduced impacts. Our paper outlined two different instruments for instituting such policies, namely location-specific (targeting the robust locations themselves) and characteristic-based (targeting the general characteristic of robust locations) instruments. In our case study, the characteristics of robust locations were high yields and current land use as cropland or grass and shrub land; many of the locations selected for ethanol production via the scenario projection already had the latter characteristic, but not the high yield.

Although our case study has some drawbacks, in particular the inability to take into account the costs and GHG emissions of indirect land use change, this study shows that complementing land use change scenario projections with a Pareto frontier offers significant added value for policy making. The developed methodology is general and can be applied to other regions, scales, objectives and commodities.

Acknowledgements

This work was partly carried out within the BE-Basic R&D Program, which was granted an FES subsidy from the Dutch Ministry of Economic affairs, agriculture and innovation (EL&I). We thank the MSc students Luis Manuel de Vries and Sanne Hettinga for the Python implementation of the genetic algorithm and the explorative study on optimizing the locations of sugar cane mills, respectively. Celeste Brennecka is thanked for improving the English in this article.

Software/data availability

The Automatic Model Optimization Reference Implementation, AMORI, optimization software is freely available online: https://sourceforge.net/projects/amori/. Developer: L.M. de Vries, email lmdevries@barcelonascience.com.

The land use model can be downloaded with input data from an earlier case study: https://github.com/JudithVerstegen/PLUC_Mozambique. Developer: J.A. Verstegen, first author, see contact details on first page. Running the model requires Python version 2.7 with NumPy version 1.8, and PCRaster version 4.1.0 (current release), downloadable for free at: http://pcraster.geo.uu.nl/downloads/.

The sources of all datasets and parameter values are provided in Appendix B.

Appendix A. Equations of the cost and emission components

A.1. General

This section provides all equations used for the cost and emission components of the allocation model. In the equations the following notations are used throughout. Italic variables are non-spatial and bold variables are spatial, i.e. a map of values at raster cells instead of a single value. The subscript i for a spatial variable indicates the selection of a cell in the map on which sugar cane is allocated, for $i = 1, 2, \ldots I$. The value for I, the total number of cells with sugar cane, varies per model run (per individual in the genetic algorithm (GA) population), depending on whether sugar cane is allocated on high yielding (low I) or low yielding (high I) cells.

Many of the equations contain the yield of sugar cane. Sugar cane is a semi-perennial crop. This means that after planting, it can be harvested for some consecutive years. In Brazil, a 6-year cycle is most common: the first harvest takes place 12 or 18 months after planting and the subsequent harvests once every year for four years (Macedo et al., 2008). During these four years, the yield gradually decreases. In the next year, the field is renewed. The yield we use, is the average yield over this 6-year cycle. The exception are the cultivation costs, where we do specifically account for the changing yield over the cycle, because here it influences the total costs. The yield of sugar cane in a cell, $\mathbf{y}_i$ (tonne cane), is constructed from a map showing the relative yield distribution over space and an average (over the cycle) maximum attainable yield:

$$\mathbf{y}_i = \mathbf{f}_i \cdot m \cdot a \quad \text{for each } i \tag{A.1}$$

In equation (A.1), m (tonne cane/ha) is the maximum attainable yield, in our case for 2030. Furthermore, $\mathbf{f}_i \in [0,\ 1]$ $(-)$ is the fraction of the maximum yield that can be obtained. And a (ha) is the cell area, which is constant over space since we use an Albers Equal Area map projection.

Many of the cost and emission components are expressed per tonne cane. Because we are interested in the costs of the end product, ethanol, all components are converted to costs and emissions per m^3 ethanol, using the conversion efficiency η (m^3 ethanol/tonne cane). We assume a single conversion efficiency for all mills allocated.

A.2. Costs

All cost equations are derived from the equations by Jonker et al. (2015). For more detailed information we refer to that study. The land costs, c_l (US\$$_{2014}$/$m^3$ ethanol), consist of two parts: the costs to buy the land and the land conversion costs to make the land cultivatable for sugar cane. Both parts contain area dependent costs and yield dependent costs:

$$c_l = \frac{1}{I \cdot \eta} \cdot \sum_{i=1}^{I} \frac{\alpha \cdot a \cdot (a_l \cdot (\mathbf{b}_i + \mathbf{l}_i) + \ b_l \cdot \mathbf{y}_i \cdot (\mathbf{b}_i + \mathbf{l}_i))}{\mathbf{y}_i} \tag{A.2}$$

In equation (A.2), $\mathbf{b}_i$ (US\$$_{2014}$/ha) are the costs to buy the land and $\mathbf{l}_i$ (US\$$_{2014}$/ha) are the land conversion costs. Both vary over space based on region and/or current land use type. The factors a_l $(-)$ and b_l (ha/tonne cane) distinguish between the area dependent costs and yield dependent parts. The factor α $(-)$ is the annuity factor that transforms the total costs to yearly costs: $\alpha = r/(1-(1+r)^{-L})$. Herein, r $(-)$ is the discount rate, i.e. the time value of money according to the theory of time preference, and L (years) is the lifetime or amortization period (Blok, 2006, p. 195, equation 11.2b). Again, a is the cell area.

The cultivation costs, c_c (US\$$_{2014}$/$m^3$ ethanol), contain many factors, like fertilizer application and machinery use. A simplified equation, summing all cost components in area-dependent and yield dependent cultivation costs, is derived from the equation given by Jonker et al. (2015). The initial investment costs are annualized, but this cannot simply be done using α as in equation (A.2), because the yearly costs vary over the 6-year sugar cane cultivation cycle:

$$c_c = \frac{1}{I \cdot \eta} \cdot \sum_{i=1}^{I} \left(\frac{\sum_{t=1}^{6} (a_{c,t} \cdot a) \Big/ \sum_{t=1}^{6} (1+r^t)}{\sum_{t=1}^{6} (\mathbf{y}_i \cdot y_t) \Big/ \sum_{t=1}^{6} (1+r^t)} + \frac{\sum_{t=1}^{6} (b_{c,t}) \Big/ \sum_{t=1}^{6} (1+r^t)}{\sum_{t=1}^{6} y_t \Big/ \sum_{t=1}^{6} (1+r^t)} \right) \tag{A.3}$$

In equation (A.3), y_t $(-)$ is the yield factor for year $t = 1, 2, 3, 4, 5, 6$, decreasing over the 6-year sugar cane cultivation cycle, $a_{c,t}$ (US\$$_{2014}$/ha) are the area-dependent cultivation costs at year t, for example the costs of machinery, $b_{c,t}$ (US\$$_{2014}$/tonne cane) are the yield dependent cultivation costs at year t, for example for fertilizers. Again, r $(-)$ is the discount rate.

The costs of transporting the sugar cane to the mill, c_t (US\$$_{2014}$/$m^3$ ethanol), are calculated over the road network. It is assumed that the truck will take the fastest route. The fastest route is determined by applying a least cost path algorithm on the road map with speed differing per road type and off-road, where the 'costs' per kilometre are one divided by the speed (higher speed are lower 'costs' because it is faster). The 'costs' themselves are not used, only the route, to compute the average speed and diesel use along the routes and total distance of the routes per sugar cane cell:

$$c_t = \frac{1}{I \cdot \eta} \cdot \sum_{i=1}^{I} (a_t / \mathbf{v}_i + \mathbf{b}_i) \cdot \mathbf{d}_i + o_t \tag{A.4}$$

In equation (A.4), $\mathbf{v}_i$ (km/hour) is the average truck speed, and $\mathbf{d}_i$ (km) is the total distance of the fastest route to the nearest (time-wise) mill. Furthermore in equation (A.4), a_t (US\$$_{2014}$/tonne cane-hour) are the annual costs of the truck, $\mathbf{b}_i$ (US\$$_{2014}$/tonne cane-km) are the diesel costs per tonne cane that differ per field depending on the road types in the route to the mill, and o_t (US\$$_{2014}$/tonne cane) are the costs of loading and unloading the truck. When, during the allocation process of the fields, one of the mills has reached the maximum capacity, the spatial variables are updated for all cells that had this 'full' mill as their closest mill, according to the fastest route to the next nearest mill.

The costs of processing the sugar cane (converting it to ethanol), c_p (US\$$_{2014}$/$m^3$ ethanol), include capital depreciation, operational costs and revenues from electricity generation. In contrast to the other cost components, the processing costs are calculated per mill instead of per field, because the costs depend on the scale of the mill. Therefore, the total processing costs are obtained by summing over all *active* mills, $j = 1, 2, \ldots J$. The notion 'active' indicates all mills to which fields are assigned. Mills to which no fields are assigned, do have a location in theory (they have a x and y coordinate in the GA), but do not contribute to the total costs of the individual and are thus excluded from the analysis:

$$c_p = \frac{1}{J \cdot \eta} \cdot \sum_{j=1}^{J} \frac{(\alpha \cdot (a_p \cdot \mathbf{s}_j) + b_p)}{\mathbf{s}_j \cdot q} + c_o - g_e \cdot r_e \tag{A.5}$$

In equation (A.5), $\mathbf{s}_j$ (tonne cane/hour) is the scale of the mill, indicating the sugar cane processing capacity, and $\mathbf{s}_j \cdot q \cdot \eta$ (m^3 ethanol) is the annual output of the mill, in which q (hours) is the number of hours per year the mill is in running. In this study, q is assumed to be the same for all mills. Furthermore, a_p (US\$$_{2014}$-h/tonne cane) is a cost factor that decreases with the scale of the mill, representing the advantages of economies of scale, while b_p (US\$$_{2014}$) is a fixed cost factor. Moreover, c_o (US\$$_{2014}$/$m^3$ ethanol) are the fixed operation costs, and g_e (kWh/m^3 ethanol) is the electricity surplus. This electricity is generated from bagasse, a fibrous product left over after the sugary juice is extracted from the sugar cane. The electricity surplus, the part of the generated electricity the mill does not need for the sugar cane processing, can be sold to the grid; r_e (US\$$_{2014}$/kWh) are the revenues obtained for this surplus. Again, as with land costs, α $(-)$ is the annuity factor that transforms the total costs to yearly costs.

A.3. Emissions

The land emissions, e_l (tonne CO_2-eq/m^3 ethanol), from carbon stock changes are calculated using the IPCC approach (IPCC, 2006). This approach involves five carbon pools: above ground biomass, below ground biomass, dead wood, litter, and soil organic carbon (SOC). In line with the Tier 1 approach of the IPCC, an equilibrium is assumed in the dead wood and litter stocks, i.e. they are considered not to change:

$$e_l = \frac{44}{12} \cdot \frac{1}{I \cdot \eta} \cdot \sum_{i=1}^{I} \frac{(\mathbf{s}_{i,2006} - \mathbf{s}_{i,c} + \mathbf{b}_{i,2006} - \mathbf{b}_{i,c}) \cdot a}{h \cdot \mathbf{y}_i} \tag{A.6}$$

In equation (A.6), $\mathbf{s}_{i,2006}$ (tonne C/ha) is the mineral soil organic carbon for 2006, the reference year. The mineral SOC is calculated given the soil type, climate, land use and management (IPCC, 2006). Next, $\mathbf{s}_{i,c}$ (tonne C/ha) is the mineral soil organic carbon when all cells $i = 1, 2, \ldots I$ are converted to sugar cane. This means that land use and management are changed, while soil type and climate remain the same. Organic SOC is not considered in equation (A.6) because our study area does not contain organic soils. Furthermore, $\mathbf{b}_{i,2006}$ (tonne C/ha) is the above and below ground biomass for 2006, based on land use and productivity. In the same fashion as with $\mathbf{s}_{i,c}$, $\mathbf{b}_{i,c}$ (tonne C/ha) is the above and below ground biomass in all cells $i = 1, 2, \ldots I$ where sugar cane is cultivated. The total carbon stock changes are divided over a time horizon h (years). The factor 44/12 is to convert from C to CO_2-eq.

The cultivation emissions, e_c (tonne CO_2-eq/m^3 ethanol), originate from the use of diesel (for machinery), fertilizers, agrochemicals and other chemicals:

$$e_c = \frac{1}{I \cdot \eta} \cdot \sum_{i=1}^{I} \frac{l_c \cdot a + \mathbf{y}_i \cdot k_c}{\mathbf{y}_i} \tag{A.7}$$

In equation (A.7), k_c (tonne CO_2-eq/ha) are the diesel emissions from the machinery, and l_c (tonne CO_2-eq/tonne cane) are the yield-dependent emissions, including primarily fertilizer emissions. These are mainly N_2O emissions, converted to CO_2-eq.

The transport emissions, e_t (tonne CO_2-eq/m^3 ethanol), are diesel emissions from the trucks transporting the sugar cane to the mill:

$$e_t = \frac{1}{I \cdot \eta} \cdot \sum_{i=1}^{I} \frac{\mathbf{k}_i \cdot \mathbf{d}_i}{l_t} \tag{A.8}$$

In equation (A.8), $\boldsymbol{k}_i$ (tonne CO_2-eq/km) are the diesel emissions per tonne cane that differ per field depending on the road types in the route to the mill, including a factor to correct for the empty return of the truck, and l_t (tonne cane) is the load of a full sugar cane truck.

The processing emissions, e_p (tonne CO_2-eq/m^3 ethanol), are assumed not to differ per mill, in contrast to the processing costs; it is a fixed emission per tonne cane and thus per m^3 ethanol:

$$e_p = \frac{1}{\eta} \cdot k_p - l_p \tag{A.9}$$

In equation (A.9), k_p (tonne CO_2-eq/tonne cane) includes all processing emissions and l_p (tonne CO_2-eq/tonne cane) are the emissions avoided by electricity production.

Appendix B. Input data

This section describes the data used for the Goiás case study within the equations given in the previous section. The values of non-spatial variables are given in Table B.1 for costs and in Table B.2 for emissions. The data sources for all maps are given in Table B.3.

One of the most important variables in the cost calculations is the maximum attainable yield, m. The value of m is determined for 2012 by finding the m for which $\sum_{i=1}^{I} \mathbf{y}_i = q$, where $i = 1, 2, \ldots I$ are in this case all cells that are projected to be sugar cane for 2012 by a combination of data from the Canasat project (Rudorff et al., 2010) and a model projection from the PCRaster Land Use Change model PLUC (van der Hilst et al., in prep.). Furthermore, q is the total sugar cane production reported by the Brazilian Sugarcane Industry Association UNICA (UNICA, 2015). The value of m for 2030 (Table B.1) is found by applying a yield trend over time from Jonker et al. (2015) to the 2012 value.

Regarding land emissions we assume, in line with the IPCC method (IPCC, 2006), that the above and below ground biomass of cropland is zero, because the crops are fully harvested each year. For planted pasture, it is assumed that all above ground biomass is eaten by the livestock each year, so the biomass stock of pasture is only its below ground biomass. Rangelands have a stocking rate of about 70% lower than pastures (Aguiar and d'Athayde, 2014), so we assume that only 30% of the above ground biomass is eaten each year, 70% remains in stock. Along similar lines, we assume that the above ground sugar cane is harvested each year and that the roots remain intact. The biomass stock of planted forest is also its below ground biomass only, as all carbon in the above ground stock is eventually harvested.

Table B.1
Non-spatial data for the sugar cane production costs. All values are for the year 2030 and expressed in US$$_{2014}$; if values in the source were in another monetary unit, they are converted using the IGP-DI index (Banco Central do Brasil, 2015).

Cost component	Variable	Unit	Symbol	Value		Source
general	maximum yield	tonne cane/ha	m	212		Rudorff et al., 2010, see explanation in main text, van der Hilst et al., in prep., UNICA, 2015
	cell area	ha	a	2500		–
	conversion efficiency	m^3 ethanol/tonne cane	η	0.09		Jonker et al., 2015
land	annuity factor	–	α	0.13 *		Jonker et al., 2015
	correction factor	–	a_l	0.33		FNP Informa economics, 2012
	correction factor	ha/tonne cane	b_l	$6.67 \cdot 10^{-3}$		FNP Informa economics, 2012
cultivation	yield factor in year t	–	y_t	t	y_t	Macedo et al., 2004
				1	0	
				2	1.29	
				3	1.09	
				4	0.95	
				5	0.87	
				6	0.81	
	area-dependent cultivation costs in year t	US$$_{2014}$/ha	$a_{c,t}$	t	y_t	Jonker et al., 2015
				1	2577	
				2	1124	
				3	1124	
				4	1124	
				5	1124	
				6	1050	
		US$$_{2014}$/tonne cane	$b_{c,t}$	t	y_t	Jonker et al., 2015

Table B.1 (*continued*)

Cost component	Variable	Unit	Symbol	Value	Source
	yield dependent cultivation costs in year t			1 8.3 2 15.0 3 15.9 4 16.3 5 16.9 6 17.2	
	discount rate	–	r	0.12	Jonker et al., 2015
transport	capital depreciation of the truck	$US\$_{2014}$/tonne cane-hour	a_t	2.68	Jonker et al., 2015
	truck loading and unloading	$US\$_{2014}$/tonne cane	o_t	2.00	Jonker et al., 2015
processing	annuity factor	–	α	0.13 *	Jonker et al., 2015
	period per year the mills runs	hours	q	$170 \cdot 24$	Dias et al., 2011
	cost factor decreasing with scale	$US\$_{2014}$-h/tonne cane	a_p	75.57 ** 59.09 ***	Jonker et al., 2015
	fixed cost facor	$US\$_{2014}$	b_p	$40 \cdot 10^6$ ** $100 \cdot 10^6$ ***	Jonker et al., 2015
	operation costs of the mill	$US\$_{2014}/m^3$ ethanol	c_o	98.67	Jonker et al., 2015
	electricity surplus	kWh/m^3 ethanol	g_e	906.67	Jonker et al., 2015
	revenues from electricity	$US\$_{2014}$/kWh	r_e	0.07	Jonker et al., 2015

* Calculated for an amortization period L of 20 years with a 12% interest rate r (Blok, 2006, p. 195, equation 11.2b).
** For mills with a scale smaller than 1000 tonne cane/hour.
*** For mills with a scale equal to or larger than 1000 tonne cane/hour.

Table B.2
Non-spatial data for the sugar cane production emissions. All values are for the year 2030.

Emission component	Variable	Unit	Symbol	Value	Source
land	time horizon	years	h	20	IPCC, 2006, European Parliament and Council of the European Union, 2009
cultivation	yield-dependent emissions	tonne CO_2-eq/tonne cane	(290) k_c	$15.22 \cdot 10^{-3}$	emission per component: Macedo et al., 2008, quantities: Jonker et al., 2015
	area-dependent emissions	tonne CO_2-eq/ha	l_c	$365.91 \cdot 10^{-3}$	emission per component: Macedo et al., 2008, quantities: Jonker et al., 2015, Seabra et al., 2011
transport	truck load	tonne cane	l_t	30	Jonker et al., 2015, CTBE, 2012
processing	processing emissions	tonne CO_2-eq/tonne cane	k_p	$4.45 \cdot 10^{-3}$	Jonker et al., 2015, Seabra et al., 2010
	emissions avoided by electricity production	tonne CO_2-eq/tonne cane	l_p	$10.44 \cdot 10^{-3}$	surplus quantity: Jonker et al., 2015, energy mix of Brazil (excluding bagasse) and related emissions: IEA, 2013

Table B.3
Sources of all spatial data for sugar cane production costs and emissions.

Component	Variable	Unit	Symbol	Source
general	yield fraction	tonne/ha	$\mathbf{f}_i$	Tóth et al., 2012
land costs	land tenure costs	$US\$_{2014}$/ha	$\mathbf{b}_i$	FNP Informa economics, 2012*
	land conversion costs	$US\$_{2014}$/ha	$\mathbf{l}_i$	FNP Informa economics, 2012**
transport costs	speed	km/hour	$\mathbf{v}_i$	speeds on different road types adapted from (to correct for trucks going slower): de Souza Soler and Verburg, 2010
	diesel costs	$US\$_{2014}$/tonne-km	$\mathbf{b}_i$	adapted from for different speeds: Jonker et al., 2015
	distance	km	$\mathbf{d}_i$	calculated over road network from: UFG, 2015
conversion costs	scale	tonne cane/hour	$\mathbf{s}_j$	determined by the model, max scale for 2030 set at 1348 tonne cane/hour (5.5 Mtonne cane/year) (MME, 2013)
land emissions	mineral soil organic carbon in 2006	tonne C/ha	$\mathbf{s}_{i,2006}$	values for carbon dependent on land use type, soil, climate and management level: IPCC, 2006, 2006 land use map: Verstegen et al., 2015, soil map: Batjes, 2010, climate map: Hijmans et al., 2005, Bernoux et al., 2006
	mineral soil organic carbon when sugar cane is cultivated	tonne C/ha	$\mathbf{s}_{i,c}$	values for carbon dependent on land use type, soil, climate and management level: IPCC, 2006, soil map: Batjes, 2010, climate map: Hijmans et al., 2005
	total (above + below ground) biomass stock for 2006	tonne C/ha	$\mathbf{b}_{i,2006}$	ratio below to above ground biomass: IPCC, 2006, above ground biomass: maximum yield assumptions by the authors together with yield fraction map by Tóth et al. (2012), Jangpromma et al., 2012, de Miranda et al., 2014, Epron et al., 2013, initial land use map: Verstegen et al., 2015
	total biomass stock when sugar cane is cultivated	tonne C/ha	$\mathbf{b}_{i,c}$	ratio below to above ground biomass: IPCC, 2006, above ground biomass: maximum yield assumptions by the authors together with yield fraction map by Tóth et al. (2012), Jangpromma et al., 2012, initial land use map: Verstegen et al., 2015
transport emissions	diesel emissions	tonne CO_2-eq/km	$\boldsymbol{k}_i$	Macedo et al., 2008, quantities: Jonker et al., 2015, Hamelinck et al., 2005b

* The FNP (FNP Informa economics, 2012) specifies land value per micro region in Goiás per land use type (distinction between natural vegetation, pasture and cropland). A land value map was made using the map of micro regions in Brazil and the land use map of 2006 (Verstegen et al., 2015). In addition the FNP indicates a higher land value around the cities of Rio Verde and Santa Helena de Goiás. This higher land value was assigned to all grid cells within a buffer of 50 km around these cities.
** The FNP (FNP Informa economics, 2012) specifies land conversion costs separately for nature and agriculture. A conversion cost map was made by linking these values to the land use map of 2006 (Verstegen et al., 2015).

Appendix C. Genetic algorithm

A GA searches the solution space by mimicking evolutionary processes. It starts with a population of *N* candidate solutions, also called individuals (Fig. C1). Each individual has a genotype, consisting of a bit-string of genes representing the control variables of the problem, and a phenotype, the 'appearance' resulting from the genotype (Bennett et al., 1998). The fitness of each individual in the population is calculated by evaluating this phenotype against the objective(s). The best-performing individuals (a predetermined fraction of the population) are selected to 'reproduce'. This is done by crossover, also called recombination, and mutation (Blum and Roli, 2003). Crossover is the process of taking genes from two parents and combining them into a new genotype. Mutation alters a bit in one or more randomly selected genes. The new generation, the parents and the children together, generally has a higher fitness than the previous generation. The GA is configured to terminate when the optimum has been found.

too low, the objective value stabilizes too early, before the optimum is reached, because too little variation remains in the population. If the fraction is too high, individuals with a relatively low fitness reproduce, thereby not improving the fitness of the next generation. The mutation rate is the fraction of the total population that will be mutated. If the mutation rate is too low, the GA can become stuck at a local optimum, while if it is too high, the genotypes of the individuals with a high fitness change too much and there is no convergence towards the optimum objective value (e.g. Bennett et al., 1998). For our optimization problem the fastest conversion towards the lowest minimum reached was with a population fraction to reproduce of 0.1 and a mutation rate of 0.3. During cross-over, individuals are split at two locations in the bit-string. The maximum number of bits to mutate in a single individual is two.

Next, we increased the population size and number of generations with these settings until no improvement in the objective function was reached anymore. This was at a population of 1000

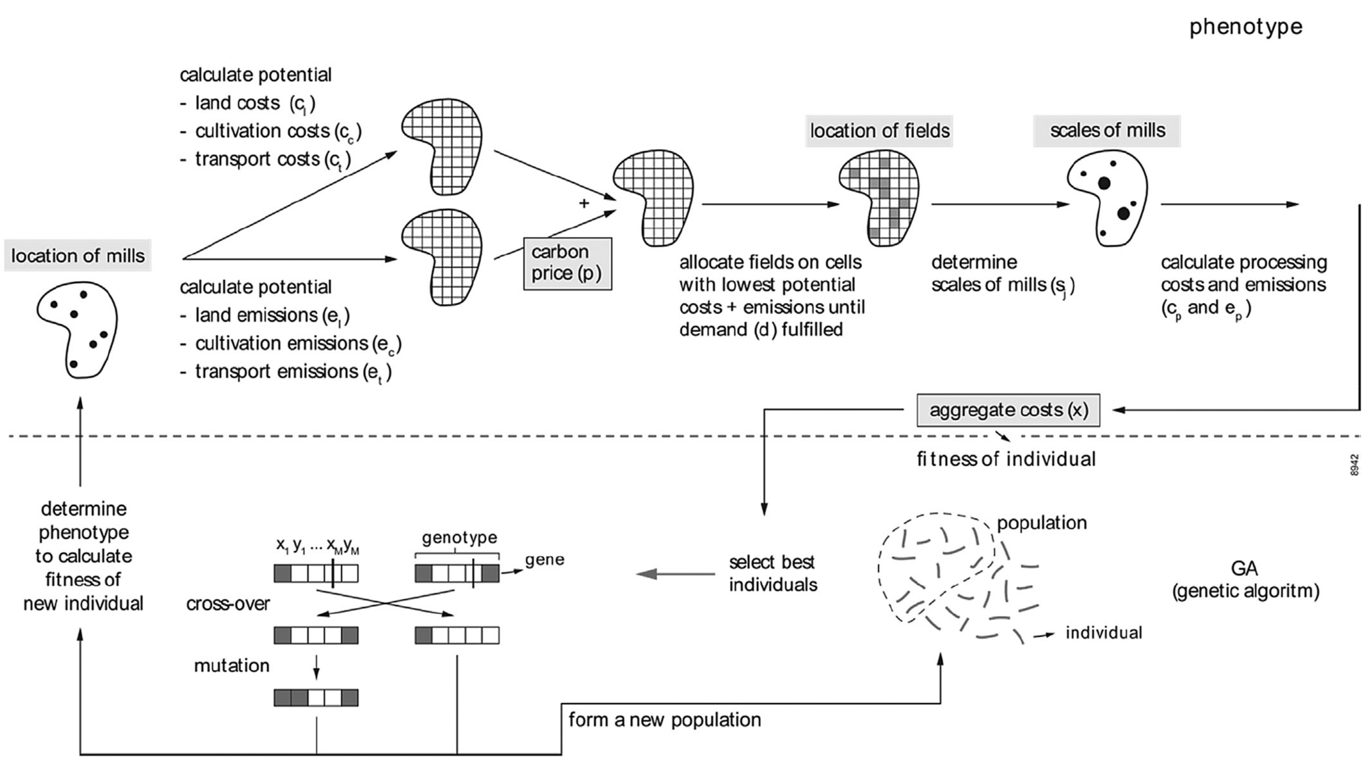

Fig. C1. Conceptual model of the optimization with the genetic algorithm (below the dashed line) and the control variables and calculation of the objective values (above the dashed line).

The settings of the parameters of our GA are tuned by systematic variation and monitoring the effect on the variance in the population and on the objective value of the best individual of the final population. First the fraction of the population to reproduce and the mutation rate are optimized for a population of 100 individuals. A fraction of the population to reproduce of 0.2 means that the best 20% of the population is progressed to the next generation and this 20% creates the new 80% of the population by cross-over. When the fraction of the population to reproduce is

(no improvement anymore for 10 000) for 24 generations (no improvement anymore for 25). One run takes 28 h on a Linux server with 16 GB RAM and 24 cores with 2 GHz Intel Xeon processors. Running the GA five times with these parameter settings for a single carbon price of 100 $US\$_{2014}$/tonne CO_2-eq proved that at generation 24 the objective value of the best final individual is always stable for some consecutive generations, and has a maximum variation of 2.8 $US\$_{2014}$/tonne CO_2-eq between the different runs (Fig. C2).

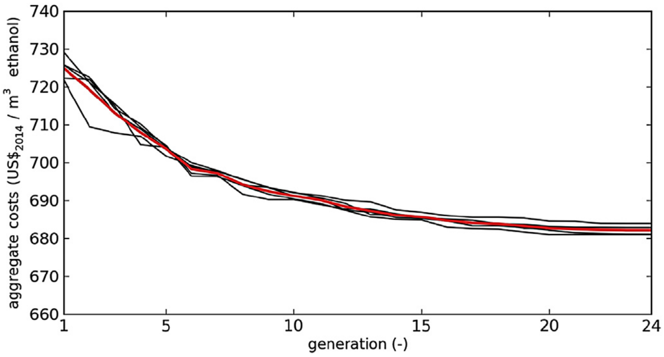

Fig. C2. Development of the objective value *x* over the generations of the GA for five different runs (black lines) at a carbon price *p* of 100 $US\$_{2014}$ / tonne CO_2-eq. The red line indicates the mean over the five runs.

References

Adami, M., Rudorff, B.F.T., Freitas, R., Aguiar, D.A., Sugawara, L.M., Mello, M.P., 2012. Remote sensing time series to evaluate direct land use change of recent expanded sugarcane crop in Brazil. Sustainability 4, 574–585.
Aguiar, G.A.M., d'Athayde, H.P., 2014. Information about Cattle Production and Agriculture in Brazil.
Aguiar, D.A., Rudorff, B.F.T., Silva, W.F., Adami, M., Mello, M.P., 2011. Remote sensing images in support of environmental protocol: monitoring the sugarcane harvest in São Paulo State, Brazil. Remote Sens. 3, 2682–2703.
Akgul, O., Shah, N., Papageorgiou, L.G., 2012. An optimisation framework for a hybrid first/second generation ethanol supply chain. Comput. Chem. Eng. 42, 101–114.
Almeida, J., De Meyer, A., Cattrysse, D., Van Orshoven, J., Achten, W.M.J., Muys, B., 2016. Spatial optimization of Jatropha based electricity value chains including the effect of emissions from land use change. Biomass Bioenergy 90, 218–229.
AMORI, 2009. Automatic model optimization reference implementation. Available online at: http://sourceforge.net/projects/amori.
Arancibia, A.L., Marques, G.F., Mendes, C.A.B., 2016. Systems capacity expansion planning: novel approach for environmental and energy policy change analysis. Environ. Model. Softw. 85, 70–79.
Banco Central do Brasil, 2015. Price Indices in Brazil; with Information up to March 2015.
Batjes, N.H., 2010. IPCC Default Soil Classes Derived from the Harmonized World Soil Data Base (Ver. 1.1). Report 2009/02b.
Bennett, D.A., Armstrong, M.P., Wade, G.A., 1998. Exploring the solution space of semi-structured geographical problems using genetic algorithms. Trans. GIS 3, 51–71.
Bernoux, M., Cerri, C.C., Cerri, C.E.P., Siqueira Neto, M., Metay, A., Perrin, A., Scopel, E., Razafimbelo, T., Blavet, D., Piccolo, M.D.C., Pavei, M., Milne, E., 2006. Cropping systems, carbon sequestration and erosion in Brazil, a review. Agron. Sustain. Dev. 26, 1–8.
Blok, K., 2006. Introduction to Energy Analysis. Techne Press, Amsterdam, The Netherlands.
Blum, C., Roli, A., 2003. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. ACM Comput. Surv. 35, 268–308.
Chen, X., Khanna, M., Yeh, S., 2012. Stimulating learning-by-doing in advanced biofuels: effectiveness of alternative policies. Environ. Res. Lett. 7.
Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20 (1), 37–46.
Cotter, M., Berkhoff, K., Gibreel, T., Ghorbani, A., Golbon, R., Nuppenau, E.-A., Sauerborn, J., 2014. Designing a sustainable land use scenario based on a combination of ecological assessments and economic optimization. Ecol. Indic. 36, 779–787.
CTBE, 2012. CanaSoft (Version 0.15). Technological Assessment Program. Brazilian Bioethanol Science and Technology Laboratory (CTBE).
Dahan, L., Rittenhouse, K., Francis, D., Sopher, P., Schwartz, J., De Clara, S., 2015. The World's Carbon Markets: a Case Study Guide to Emissions Trading - Brazil: an Emissions Trading Case Study.
de Meyer, A., Cattrysse, D., Rasinmäki, J., Van Orshoven, J., 2014. Methods to optimise the design and management of biomass-for-bioenergy supply chains: a review. Renew. Sustain. Energy Rev. 31, 657–670.
de Miranda, S.D.C., Bustamante, M., Palace, M., Hagen, S., Keller, M., Ferreira, L.G., 2014. Regional variations in biomass distribution in Brazilian savanna Woodland. Biotropica 46, 125–138.
de Souza Soler, L., Verburg, P.H., 2010. Combining remote sensing and household level data for regional scale analysis of land cover change in the Brazilian Amazon. Reg. Environ. Change 10, 371–386.
Dias, M.O.S., Cunha, M.P., Jesus, C.D.F., Rocha, G.J.M., Pradella, J.G.C., Rossell, C.E.V., Maciel Filho, R., Bonomi, A., 2011. Second generation ethanol in Brazil: can it compete with electricity production? Bioresour. Technol. 102, 8964–8971.
Epron, D., Nouvellon, Y., Mareschal, L., Moreira, R.M.E., Koutika, L.-., Geneste, B., Delgado-Rojas, J.S., Laclau, J.-., Sola, G., Gonçalves, J.L.D.M., Bouillet, J., 2013. Partitioning of net primary production in Eucalyptus and Acacia stands and in mixed-species plantations: two case-studies in contrasting tropical environments. For. Ecol. Manag. 301, 102–111.
European Parliament and Council of the European Union, 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the Promotion of the Use of Energy from Renewable Sources and Amending and Subsequently Repealing Directives 2001/77/EC and 2003/30/EC.
Fargione, J.E., Plevin, R.J., Hill, J.D., 2010. The ecological impact of biofuels. Annu. Rev. Ecol. Evol. Syst. 41, 351–377.
FNP Informa economics, 2012. Agrianual 2012, 2012 ed (Sao Paulo).
Gaddis, E.J.B., Voinov, A., Seppelt, R., 2014. Spatial optimization of best management practices to attain water quality targets. Water Resour. Manag. 28, 1485–1499.
Gerssen-Gondelach, S.J., Wicke, B., Faaij, A.P.C., 2017. GHG emissions and other environmental impacts of indirect land use change mitigation. GCB Bioenergy 9 (4), 725–742.
Gibbs, H.K., Rausch, L., Munger, J., Schelly, I., Morton, D.C., Noojipady, P., Soares-Filho, B., Barreto, P., Micol, L., Walker, N.F., 2015. Brazil's Soy Moratorium - supply-chain governance is needed to avoid deforestation. Science 347 (6220), 377–378.
Hamelinck, C.N., Suurs, R.A.A., Faaij, A.P.C., 2005a. International bioenergy transport costs and energy balance. Biomass Bioenergy 29, 114–134.
Hamelinck, C.N., Van Hooijdonk, G., Faaij, A.P.C., 2005b. Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. Biomass Bioenergy 28, 384–410.
Haupt, R.L., Haupt, S.E., 2004. The binary genetic algorithm. In: Anonymous Practical Genetic Algorithms, 2 ed. Wiley, Hoboken, New Jersey, pp. 27–66.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.
IBGE, 2013. Produção Agrícola Municipal 2006–2012.
IEA, 2013. World Energy Outlook 2013.
IPCC, 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories.
Jangpromma, N., Thammasirirak, S., Jailsil, P., Songsri, P., 2012. Effects of drought and recovery from drought stress on above ground and root growth, and water use efficiency in sugarcane (Saccharum officinarum L.). Aust. J. Crop Sci. 6, 1298–1304.
Jonker, J.G.G., van der Hilst, F., Junginger, H.M., Cavalett, O., Chagas, M.F., Faaij, A.P.C., 2015. Outlook for ethanol production costs in Brazil up to 2030, for different biomass crops and industrial technologies. Appl. Energy 147, 593–610.
Jonker, J.G.G., van der Hilst, F., Junginger, H.M., Verstegen, J.A., Lin, T., Rodríguez, L.F., Faaij, A.P.C., 2016. Supply chain optimization of sugarcane first generation and eucalyptus second generation ethanol production in Brazil. Appl. Energy 173, 494–510.
Karssenberg, D., Schmitz, O., Salamon, P., de Jong, K., Bierkens, M.F.P., 2010. A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. Environ. Model. Softw. 25, 489–502.
Lambin, E.F., Meyfroidt, P., 2011. Global land use change, economic globalization, and the looming land scarcity. Proc. Natl. Acad. Sci. U. S. A. 108, 3465–3472.
Lapola, D.M., Schaldach, R., Alcamo, J., Bondeau, A., Koch, J., Koelking, C., Priess, J.A., 2010. Indirect land-use changes can overcome carbon savings from biofuels in Brazil. Proc. Natl. Acad. Sci. U. S. A. 107, 3388–3393.
Lautenbach, S., Volk, M., Strauch, M., Whittaker, G., Seppelt, R., 2013. Optimization-based trade-off analysis of biodiesel crop production for managing an agricultural catchment. Environ. Model. Softw. 48, 98–112.
Law, E.A., Bryan, B.A., Meijaard, E., Mallawaarachchi, T., Struebig, M.J., Watts, M.E., Wilson, K.A., 2017. Mixed policies give more options in multifunctional tropical forest landscapes. J. Appl. Ecol. 54 (1), 51–60.
Li, X., Yeh, A.G., 2005. Integration of genetic algorithms and GIS for optimal location search. Int. J. Geogr. Inf. Sci. 19, 581–601.
Macedo, I.C., Leal, M.R.L.V., Silva, J.E.A.R., 2004. Assessment of Greenhouse Gas Emissions in the Production and Use of Fuel Ethanol in Brazil.
Macedo, I.C., Seabra, J.E.A., 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 Bioenergy 32, 582–595.
MME, 2013. Plano decenal de expansão de energia 2021.
Newton, P., Gomez, A.E.A., Jung, S., Kelly, T., de Araújo Mendes, T., Rasmussen, L.V., dos Reis, J.C., de Aragão Ribeiro Rodrigues, R., Tipper, R., van der Horst, D., Watkins, C., 2016. Overcoming barriers to low carbon agriculture and forest restoration in Brazil: the Rural Sustentável project. World Dev. Perspect. 4, 5–7.
Python software foundation, 2014. Python language Reference available online at:, version 2.7. http://www.python.org.
Rudorff, B.F.T., Aguiar, D.A., Silva, W.F., Sugawara, L.M., Adami, M., Moreira, M.A., 2010. Studies on the rapid expansion of sugarcane for ethanol production in São Paulo state (Brazil) using landsat data. Remote Sens. 2, 1057–1076.
Seabra, J.E.A., Tao, L., Chum, H.L., Macedo, I.C., 2010. A techno-economic evaluation of the effects of centralized cellulosic ethanol and co-products refinery options with sugarcane mill clustering. Biomass Bioenergy 34, 1065–1078.
Seabra, J.E.A., Macedo, I.C., Chum, H.L., Faroni, C.E., Sarto, C.A., 2011. Life cycle assessment of Brazilian sugarcane products: GHG emissions and energy use. Biofuels, Bioprod. Biorefining 5, 519–532.
Seppelt, R., Lautenbach, S., Volk, M., 2013. Identifying trade-offs between ecosystem

services, land use, and biodiversity: a plea for combining scenario analysis and optimization on different spatial scales. Curr. Opin. Environ. Sustain. 5, 458–463.

Seppelt, R., Manceur, A.M., Liu, J., Fenichel, E.P., Klotz, S., 2014. Synchronized peak-rate years of global resources use. Ecol. Soc. 19 (4) art50.

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., Rice, C., Scholes, B., Sirotenko, O., Howden, M., McAllister, T., Pan, G., Romanenkov, V., Schneider, U., Towprayoon, S., Wattenbach, M., Smith, J., 2008. Greenhouse gas mitigation in agriculture. Philosophical Trans. R. Soc. B Biol. Sci. 363, 789–813.

Stewart, T.J., Janssen, R., van Herwijnen, M., 2004. A genetic algorithm approach to multiobjective land use planning. Comput. Operations Res. 31, 2293–2313.

Tempels, Van den Belt, 2016. Once the rockets are up, who should care where they come down? The problem of responsibility ascription for the negative consequences of biofuel innovations. SpringerPlus 5 (1).

The Nature Conservancy, 2012. Soy - Good Agricultural Practices and Socio-environmental Certification in Brazil: towards Sustainability, ISBN 978-85-60797-13-4.

Tóth, G., Kozlowski, B., Prieler, S., Wiberg, D., 2012. Global Agro-ecological Zones (GAEZ v3.0).

UFG, 2015. Laboratório de Processamento de Imagens e Geoprocessamento (LAPIG) do Instituto de Estudos Sócio-Ambientais (IESA) da Universidade Federal de Goiás (UFG).

UNICA, 2015. UNICADATA, Cane Harvest Reports, Production History, Prices & Quotes, Exports & Imports, Fuel Consumption, Auto Sales & Fleet Size.

van der Hilst, F., Verstegen, J.A., Woltjer, G., Smeets, E.M.W., Faaij, A.P.C., in prep. Mapping direct and indirect land use changes resulting from biofuel production and the effect of ILUC mitigation measures.

Verstegen, J.A., van der Hilst, F., Woltjer, G., Karssenberg, D., de Jong, S.M., Faaij, A.P.C., 2015. What Can and Can't We Say about Indirect Land-use Change in Brazil Using an Integrated Economic - Land-use Change Model? GCB Bioenergy Early View.

Verstegen, J.A., Karssenberg, D., van der Hilst, F., Faaij, A.P.C., 2016a. Detecting systemic change in a land use system by Bayesian data assimilation. Environ. Model. Softw. 75, 424–438.

Verstegen, J.A., van der Hilst, F., Woltjer, G., Karssenberg, D., de Jong, S.M., Faaij, A.P.C., 2016b. 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, 561–578.

Walter, A., Galdos, M.V., Scarpare, F.V., Leal, M.R.L.V., Seabra, J.E.A., da Cunha, M.P., Picoli, M.C.A., de Oliveira, C.O.F., 2014. Brazilian sugarcane ethanol: developments so far and challenges for the future. Wiley Interdiscip. Rev. Energy Environ. 3, 70–92.

Wang, L., Quiceno, R., Price, C., Malpas, R., Woods, J., 2014. Economic and GHG emissions analyses for sugar cane ethanol in Brazil: looking forward. Renew. Sustain. Energy Rev. 40, 571–582.

Wicke, B., Verweij, P., Van Meijl, H., Van Vuuren, D.P., Faaij, A.P.C., 2012. Indirect land use change: review of existing models and strategies for mitigation. Biofuels 3, 87–100.