Modeling regional supply responses using farm-level economic data and a biophysical model:
a case study on Brazilian land-use change

Samuel Ferreira Balieiro

Thünen Report 106

Bereits in dieser Reihe erschie-
nene Bände finden Sie im Inter-
net unter www.thuenen.de

*Volumes already published in
this series are available on the
Internet at www.thuenen.de*

THÜNEN

Modeling regional supply responses using farm-level economic data and a biophysical model: a case study on Brazilian land-use change

Samuel Ferreira Balieiro

Thünen Report 106

Dr. Samuel Ferreira Balieiro
Thünen-Institut für Betriebswirtschaft
Bundesallee 63
38116 Braunschweig

Summary

Estimating farmers' supply responses to changes in framework conditions is important to in-form decision-makers on the expected impacts on production volume as well as the resulting land-use shifts. Existing agricultural supply response models generally require either larger databases with farm-level data for microregional analysis or are implemented with a coarse resolution (e.g., country level) due to the lack of data. While such approaches are suitable for regions with abundancy of data or for global-scale analysis, there is a need for an alternative for micro-level analysis in countries with low data availability. In addition, it is important to include the spatial component in the regional supply response analysis, allowing not only the quantification of the overall change in output but also the likely spatial land-use change.

Against this background, this dissertation aims to answer the research question whether a combination of a biophysical model with farm-level economic data can be used to estimate farm-level profitability of individual crops and respective cropping systems and thereby simu-late farmers' supply responses in countries with limited data availability. To answer this ques-tion, a new modeling approach called **Profitability Assessment Model (PAM)** is developed, tested and validated. This new modeling approach follows the principles of minimum data, focusing on delivering timely and quantitative analyses with satisfactory accuracy to inform decision-makers. That is an important feature since the overall goal of the concept is to limit the data required by the model to a minimum, allowing quick implementation while accepting moderate accuracy.

The PAM is a spatially explicit model with simulation units' size of spatial resolution grid vary-ing between 5 and 30 arcmin (10x10 to 50x50 km in area), following that used by the Global Biosphere Management Model (GLOBIOM). PAM estimates the profitability of each farming alternative at the simulation unit level and allocates the land to maximize farmers' return to land. The model has four main modules developed to account for the different components of the profitability estimation:

- **Plant module:** Using the EPIC-IIASA crop growth model (biophysical), this module es-timates crop yields and fertilizer input use based on a standard production system. The module uses weather and soil data, farm operation schedules and crop growth param-eters to estimate actual daily biomass gain, which is a function of potential daily gains penalized due to stress factors such as lack of nutrients, water, etc.

- **Transport module:** Transport costs are needed primarily to simulate farm-gate prices for both inputs and outputs. The tool endogenously estimates the transport costs be-tween the simulation units and the nearest port. It uses observed data to estimate freight cost per km, which is combined with the estimated travel distances.

- **Economic module:** This is the main innovation of the PAM model. It extrapolates known production cost data per crop from the *agri benchmark* typical farms to all the simulation units, using individual functions for each cost component. Finally, the module combines the information from the previously mentioned modules to calculate the return to land of each farming alternative at the simulation unit level.

- **Land allocation module:** Maximizes the return to land for each simulation unit, allocating land to the most profitable farming alternative. Constraints such current land use and maximum share of each crop are also considered. The output is a land-use map.

The PAM model is developed and calibrated for the Brazilian agricultural sector. Using Brazil as the case study is interesting due to its overall importance in the global production of agricultural commodities as well as the environmental impact of land-use changes. Due to its prominent role as an exporter, changes in international prices and trade agreements are expected to impact Brazilian farmers directly and may cause land-use changes. Besides the empirical relevance of Brazil, the availability of data helps the development of the extrapolation routines and the model validation. Additionally, the predominant cropping systems are relatively simple, supporting the development and implementation of the model.

For this case study, four production system are represented in the PAM model: (a) double cropping of soybeans and maize, (b) soybeans with a cover crop, (c) sugarcane monoculture and (d) beef production. While the profitability of the arable crops is endogenously estimated, beef is considered as an opt-out option, which is modeled based on exogenous return-to-land information. Since soybean, maize and sugarcane production accounts for 84% of the total seeded area in Brazil, the current version of the PAM model represents the most important cropping alternatives to farmers in Brazil, but not all.

An important methodological contribution of the dissertation is the development of routines for the extrapolation of each production cost component from the known typical farms' data to all regions in the country. These routines are based on local expertise as well as existing information on yield levels, prevailing production systems and farming conditions. Each cost component is analyzed individually and, based on theoretical discussions, specific cost functions are proposed following the expected behavior of each cost item – e.g., linear relationship with yields or fixed per ha. That should improve the accuracy of the model in estimating production costs (and finally profitability) while also allowing the model to be adapted to simulate changes in framework conditions that may affect only selected cost items (e.g., a significant increase in fuel prices).

In addition, the PAM model improves on existing models because it accounts for specific cost components such as the transport of sugarcane from farm to mill, which is required due to the perishability of the crop. Besides the important impact of inbound transport cost on the overall profitability of sugarcane production, the endogenous simulation of this cost item allows the model to spatially differentiate among regions depending on the current availability

of mills. That is expected to play an important role regarding farmers' supply responses depending on the time horizon of the analysis since the lack of milling infrastructure increases the inbound transport cost significantly in the short-term (e.g., less than 5 years). In the medium-term (5 to 15 years), however, mills may be built and thereby substantially improve the on-farm competitiveness of sugarcane.

A major constraint for regional profitability analysis is the lack of information regarding farm input and output prices. To overcome this problem, the PAM model provides an interesting alternative by endogenously estimating prices via the transport module. By considering the different transportation costs of each crop and basing the distance estimation on the actual availability of roads, the model allows a straightforward conversion of reference prices to farm-gate prices. For Brazil, reference price is the nearest port because much of its production is exported while the majority of its bulk inputs, such as fertilizer, are imported. For other countries, the reference price point can be adapted (i.e., domestic market) while still following the same approach to calculate the farm-gate price. The ability to endogenously simulate transport cost is a useful feature for the simulation of scenarios based on price shocks.

Apart from the development of the modeling approach, this dissertation focuses on the quantitative model validation as a key step to identify strengths and limitations of the concept. Projected yields are validated against regional statistics and production cost estimates are benchmarked against the two available datasets, with a suitable number of primary typical-farm data. Furthermore, the resulting land-use maps are evaluated against two simplified validation maps representing current land use.

In the business-as-usual scenario, the PAM model estimates a national weighted average of returns to land of 248 USD/ha for double cropping and 188 USD/ha for sugarcane. This relationship, however, is different in the states of Sao Paulo and Minas Gerais, where, on average, sugarcane has a higher return to land than double cropping. Benchmarking PAM's production cost estimates with observed local data shows a satisfactory model accuracy with a relative mean absolute error (rMAE) lower than 14%. The lowest error found in the production cost estimation is in sugarcane (rMAE of 8.7%) and the highest in second-crop maize (rMAE of 14%). Spatially, the model has a better performance in the Center-west region for grains and in the Center-south region for sugarcane. That likely is the result of model parametrization focused on the main production regions and the lack of typical farms in certain regions.

Conversely, yield estimation with the biophysical model is challenging due to the distinct error patterns among the crops in Brazil. Compared with official statistics, the EPIC-IIASA model with standard calibration overestimates soybean yields (rMAE of 52%) while it underestimates sugarcane yields (rMAE of 24%). Since this error pattern artificially changes the relative profitability among the crop alternatives, crop yields are adjusted for this case study using available regional statistics, allowing the development and testing of the remaining modules in the PAM modeling approach.

The validation of the business-as-usual land-use map shows that the PAM model is able to satisfactorily reproduce the current land use in Brazil. The visual and quantitative validation results show a strong correlation between the available land-use maps, with PAM allocating the same crop as observed in 86% of total arable land. For important agricultural states, such as Mato Grosso, the agreement between the PAM's results and observed data is as high as 95%. That is an interesting result showing, on one hand, the importance of carefully calibrating the model with local expertise and, on the other, that the specialization of Brazilian farms with relatively simple crop rotations reduces the complexity of estimating land allocation strategies, thereby increasing the model accuracy.

To test the ability of the PAM model to predict land-use and output changes due to changing framework conditions, a scenario analysis is carried out: What will happen in case yields of key crops change significantly as a consequence of climate change? An extensive literature is available on the consequences of climate change in Brazil, with a strong regional and crop-specific variation regarding expected future yield developments. To simulate the impact of the growers' reaction to yield changes due to climate change, two adaptation scenarios are developed:

- **No double cropping in the tropical region (NoDC)**: Due to the expected yield penalty in short-cycle soybeans grown in double-cropping systems, farmers move out of the double-cropping system to single soybean cultivation in the tropical regions.

- **Double cropping with yield penalty (DCYp)**: Farmers keep producing in the current double-cropping systems while accepting the yield penalty for soybeans and maize.

Due to the strong reduction in the returns to land for grains (i.e., maize and soybeans) in the tropical region – i.e., on average 26% in NoDC and 39% in DCYp - more than 24% of the current arable land is simulated to move from grains to sugarcane production. These results, however, vary significantly in the different regions, where the most affected states are Goiás, Paraná and Mato Grosso, jointly accounting for more than 55% of the total land-use change. The results highlight two important considerations: (a) the reduction in returns to land for grains is usually smaller in the NoDC scenario and (b) sugarcane is expected to benefit if simulated future yield patterns are observed. Both indicate a reduction of grain availability from Brazil with grain producing areas moving to sugarcane. A stronger impact is expected on the production of second-crop maize, with farmers switching from double cropping to single-soybean production systems.

Although the scenario analysis shows an interesting trend, it is important to highlight that such strong yield impacts on maize and soybeans are expected to trigger different dynamics that are not considered. For example, farmers are expected to adjust seeding period, use drought-resistant varieties and irrigation, reducing the magnitude of the change. Moreover, international prices also should react to such a significant change in Brazilian exports and thereby readjust the profitability of the cropping alternatives, which in turn impacts farmers' supply responses.

That shows an important limitation of the PAM approach, which is the lack of an internal demand function. The model is not designed to internally estimate the new market equilibrium, so changes, such as depicted in the scenario analysis, do not affect farm-gate prices and demand. To overcome this problem, the PAM model should be integrated in the more developed market-equilibrium model, benefiting from the complex estimation of macro-economic dynamics (e.g., GLOBIOM) while improving the production cost estimation. Such integration should also improve the simulation of the land use by adding important agricultural sectors (e.g., forestry and livestock), which are important in Brazil.

The PAM model has other limitations and areas for further development that should be carefully considered: (a) Due to the strong dependency on typical farm data, it is important to consider broadening the coverage of typical farms in the entire research region. (b) The tailored construction of the cost function required an in-depth local knowledge of production systems that may not always be readily available. (c) The land allocation module requires further development to account for more complex production systems. (d) The model does not consider conversion costs so it assumes any additional return to land in one alternative relative to another will induce land-use change. (e) It is necessary to further improve the biophysical model or use a locally calibrated alternative.

Further improving the biophysical model is particularly important due to the decisive role that yields play directly in the profitability estimation and, indirectly, as a proxy for the production cost extrapolation. The standard EPIC-IIASA model currently lacks the regional calibration for Brazilian cultivars, where soybeans, for example, are grown in a shorter cycle to allow the double cropping with maize. An ongoing cooperation with local research groups and production system experts (e.g., *agri benchmark* partners) should allow improving the model by using field trials to better calibrate the crop growth parameters and finetune the farm operations' schedule to represent the regional differences in the country.

Conceptually, it also is important to highlight that the PAM model provides a profit-based farmers' supply response analysis for a time horizon between 5 to 20 years. This feature is not given by methodological constraints but mainly by the conceptual difficulties in estimation significantly shorter or longer time horizons. On one hand, short-term price movements are not expected to motivate major changes in production systems (i.e., moving from grains to sugarcane) at regional level since farmers are aware that crop prices fluctuate and changing activities may require knowledge and investments. On the other hand, long-term simulation (more than 20 years) requires a technical adaption of the costing structure, which also should consider whether fundamental changes in the current production system are expected. The suitable time horizon for each analysis should be determined by regional characteristics of crops (perennial vs annual) and the ability to depict the expected changes in production system with the typical farm data.

This dissertation contributes to the overall development of regional farmers' supply response models for countries with limited data availability, showing that it is feasible to combine a biophysical model and farm-level economic data as the basis for the profitability estimation in a high spatial resolution. The ability to estimate individual cost components separately gives the model the required flexibility for the simulation of market- and policy-related questions, providing timely and accurate information for decision-makers. The bottom-up approach based on local expertise is an important strength of the PAM model, avoiding unrealistic parametrization and ensuring that the majority of local features of production systems are included in the estimation. Finally, considering the overall goal of using minimum data, the model accuracy indicates a strong potential of the model to answer research questions, with additional parametrization and integration expected to further improve its performance.

Keywords: Brazil, Land-use change, Biophysical model, Production costs, Supply analysis, Soybeans, Maize, Sugarcane, Climate change, Profitability

Zusammenfassung

Die Schätzung der Angebotsreaktionen von Landwirten auf veränderte Rahmenbedingungen ist wichtig, um Entscheidungsträger über die erwarteten Auswirkungen einer Maßnahme auf das Produktionsvolumen und die daraus resultierenden Landnutzungsänderungen zu informieren. Bestehende Modelle für landwirtschaftliche Angebotsreaktionen erfordern in der Regel entweder größere Datenbanken für mikroregionale Analysen auf Betriebsebene oder werden aufgrund fehlender Daten mit einer groben Auflösung (z. B. auf Länderebene) implementiert. Solche Ansätze eigenen sich für Regionen mit einer hohen Datenverfügbarkeit oder für Analysen auf globaler Ebene. Für die Analyse auf Mikroebene, in Ländern mit geringer Datenverfügbarkeit, bedarf es jedoch eines alternativen Ansatzes. Darüber hinaus ist es wichtig, die räumliche Komponente in die Analyse der regionalen Angebotsreaktion einzubeziehen, um nicht nur die Gesamtveränderung der Produktion sondern auch die anzunehmende räumliche Landnutzungsänderung quantifiziert zu können.

Vor diesem Hintergrund soll in dieser Dissertation die Forschungsfrage beantwortet werden, ob die Kombination eines biophysikalischen Modells mit ökonomischen Daten auf einzelbetrieblicher Ebene dazu verwendet werden kann, um die Rentabilität einzelner Kulturen und entsprechender Anbausysteme auf Betriebsebene zu schätzen. Insbesondere sollen mit diesem Ansatz die Angebotsreaktionen von Landwirten in Ländern mit begrenzter Datenverfügbarkeit simuliert werden. Zur Beantwortung dieser Frage wird ein neuer Modellierungsansatz, das **Profitability Assessment Model (PAM),** entwickelt, getestet und validiert. Dieser neue Modellierungsansatz folgt dem Prinzip des Mindestdatensatzes und konzentriert sich darauf zeitnahe und quantitative Analysen mit zufriedenstellender Genauigkeit zu liefern, um eine gesonderte Informationsbasis für Entscheidungsträger zu schaffen. Dies ist ein wichtiges Merkmal, da das Gesamtziel des Konzepts darin besteht, die für das Modell erforderlichen Daten auf ein Minimum zu beschränken, sodass eine schnelle Umsetzung bei zufriedenstellender Genauigkeit möglich ist.

Das PAM ist ein räumlich explizites Modell, bei dem die Größe der Simulationseinheiten in einem Raster von 5 bis 30 arcmin (10x10 bis 50x50 km Fläche) variiert, in Anlehnung an das *Global Biosphere Management Model (GLOBIOM)*. PAM schätzt die Rentabilität jeder Bewirtschaftungsalternative auf Ebene der Simulationseinheit und weist das Land so zu, dass die Grundrente der Landwirte maximiert wird. Das Modell besteht aus vier Hauptmodulen, die entwickelt wurden, um die verschiedenen Komponenten der Rentabilitätsschätzung zu berücksichtigen:

- *Plant module*: Unter Verwendung des EPIC-IIASA-Pflanzenwachstumsmodells (biophysikalisch) schätzt dieses Modul die Ernteerträge und den Düngemitteleinsatz auf der Grundlage eines Standardproduktionssystems. Das Modul verwendet Wetter- und Bodendaten, Betriebsabläufe und Wachstumsparameter der Pflanzen, um den tatsächlichen täglichen Biomassezuwachs zu schätzen. Dieser ist eine Funktion des potenziellen täglichen Zuwachses, welcher jedoch durch Stressfaktoren wie Nährstoff- und Wassermangel usw. beeinträchtigt wird.

- **_Transport module_:** Die Transportkosten werden in erster Linie benötigt, um die Hoftor-Preise sowohl für die Inputs als auch für die Outputs zu simulieren. Das Tool schätzt endogen die Transportkosten zwischen den Simulationseinheiten und dem nächstgelegenen Hafen. Es verwendet beobachtete Daten zur Schätzung der Frachtkosten pro km, welche mit der geschätzten Entfernung kombiniert werden.

- **_Economic module_:** Dieses Modul ist die wichtigste Neuerung des PAM-Modells. Das Modul extrapoliert für jede Kultur die bekannten Produktionskostendaten der typischen *agri benchmark* Betriebe auf alle Simulationseinheiten. Dabei werden individuelle Funktionen für jede Kostenkomponente verwendet. Schließlich kombiniert das Modul die Informationen aus den zuvor genannten Modulen, um die Grundrente jeder landwirtschaftlichen Alternative auf Ebene der Simulationseinheit zu berechnen.

- **_Land allocation module_:** Dieses Modul maximiert die Grundrente für jede Simulationseinheit und weist das Land der profitabelsten Anbaualternative zu. Einschränkungen, wie die aktuelle Landnutzung und der maximale Anteil der einzelnen Kulturen, werden ebenfalls berücksichtigt. Das Ergebnis stellt eine Landnutzungskarte dar.

Das PAM-Modell wurde für den brasilianischen Agrarsektor entwickelt und kalibriert. Die Verwendung von Brasilien als Fallstudie ist aufgrund der allgemeinen Bedeutung des Landes für die weltweite Produktion von Agrarrohstoffen sowie der Umweltauswirkungen von Landnutzungsänderungen interessant. Aufgrund der überdurchschnittlichen Rolle des Landes als Exporteur ist davon auszugehen, dass sich Veränderungen der Weltmarktpreise und mögliche Handelsabkommen direkt auf die brasilianische Landwirtschaft auswirken und zu Landnutzungsänderungen führen können. Neben der empirischen Relevanz Brasiliens hilft die Verfügbarkeit von Daten bei der Entwicklung der Extrapolationsverfahren und der Modellvalidierung. Außerdem sind die vorherrschenden Anbausysteme relativ einfach strukturiert, was die Entwicklung und Implementierung des Modells erleichtert.

Für diese Fallstudie werden vier Produktionssysteme im PAM-Modell dargestellt: (a) Zweikulturanbausystem von Sojabohnen und Mais, (b) Sojabohnen mit einer Zwischenfrucht, (c) Zuckerrohranbau in Monokultur und (d) Rindfleischproduktion. Während die Rentabilität der Ackerkulturen endogen geschätzt wird, wird Rindfleisch als Ausstiegsmöglichkeit betrachtet, die auf der Grundlage exogener Informationen über die Grundrente modelliert wird. Der Anbau von Sojabohnen, Mais und Zuckerrohr macht 84 % der gesamten Anbaufläche in Brasilien aus, somit stellt die aktuelle Version des PAM-Modells die wichtigsten Anbaualternativen für die Landwirte in Brasilien dar, jedoch nicht alle.

Ein wichtiger methodischer Beitrag der Dissertation ist die Entwicklung von Verfahren für die Extrapolation einzelner Produktionskostenkomponenten der bekannten Daten von typischen Betrieben auf alle Regionen eines Landes. Diese Verfahren beruhen auf lokalem Fachwissen sowie vorhandenen Informationen über Ertragsniveaus, vorherrschende Produktionssysteme und landwirtschaftliche Bedingungen. Jede Kostenkomponente wird einzeln analysiert und wird – auf Grundlage theoretischer Erörterungen – spezifischen Kostenfunktionen zugeordnet, je nachdem wie das erwartete Verhalten jeder Kostenposition ist – z. B. eine lineare

Beziehung zu den Erträgen oder ein Fixwert pro ha. Dies dürfte die Genauigkeit des Modells bei der Schätzung der Produktionskosten (und letztlich der Rentabilität) verbessern und gleichzeitig eine Anpassung des Modells zur Simulation von veränderten Rahmenbedingungen ermöglichen, die sich nur auf ausgewählte Kostenpositionen auswirken können (z. B. ein erheblicher Anstieg der Kraftstoffpreise).

Darüber hinaus verbessert das PAM-Modell die bestehenden Modelle durch die Berücksichtigung spezifischer Kostenpositionen, wie z. B. den Transport des Zuckerrohrs vom landwirtschaftlichen Betrieb zur Mühle, welcher aufgrund der Verderblichkeit des Erntegutes erforderlich ist. Neben dem bedeutenden Einfluss der Transportkosten auf die Gesamtrentabilität der Zuckerrohrproduktion ermöglicht die endogene Simulation dieses Kostenfaktors dem Modell auch, in Abhängigkeit der regionalen Verfügbarkeit von Mühlen, eine räumliche Differenzierung zwischen einzelnen Regionen herzustellen. Es ist zu erwarten, dass dies, je nach Zeithorizont der Analyse, eine wichtige Rolle für die Angebotsreaktionen von Landwirten spielen wird, da eine mangelhafte Mühleninfrastruktur die Kosten für den Transport kurzfristig (z. B. in weniger als 5 Jahren) erheblich erhöht. Mittelfristig (5 bis 15 Jahre) jedoch, können neuen Mühlen gebaut werden und damit die Wettbewerbsfähigkeit von Zuckerrohr auf dem Betrieb erheblich verbessern.

Ein wesentliches Hindernis für die regionale Rentabilitätsanalyse ist der Mangel an Informationen über die Preise der landwirtschaftlichen Inputs und Outputs. Diese Problematik wird durch einen alternativen Ansatz des PAM-Modelles umgegangen, indem es die Preise endogen über das Transportmodul schätzt. Die Berücksichtigung der unterschiedlichen Transportkosten für einzelne Kulturen und die Schätzung der Entfernungen auf Grundlage tatsächlich existierender Straßen ermöglichen dem Modell eine unkomplizierte Umrechnung von Referenzpreisen in Hoftor-Preise. Für Brasilien ist der Referenzpreis der nächstgelegene Hafen, da ein großer Teil der Produktion exportiert wird, während der Großteil der Betriebsmittel wie z. B. Düngemittel importiert werden. Für andere Länder kann der Referenzpreis angepasst werden (d. h. Inlandsmarkt), wobei der Ansatz zur Berechnung des Hoftor-Preises gleichbleibt. Die Möglichkeit, die Transportkosten endogen zu simulieren, ist eine nützliche Funktion für die Simulation von Szenarien, die auf Preisshocks basieren.

Neben der Entwicklung des Modellierungsansatzes liegt der Schwerpunkt dieser Dissertation auf der quantitativen Modellvalidierung, welche ein wichtiges Element für die Ermittlung der Stärken und Limitationen des Konzepts bildet. Die prognostizierten Erträge werden anhand regionaler Statistiken validiert und die Produktionskostenschätzungen mit den beiden verfügbaren Datensätzen verglichen, wobei eine angemessene Anzahl primärer, typischer landwirtschaftlicher Datensätze verwendet wird. Darüber hinaus werden die resultierenden Landnutzungskarten anhand von zwei vereinfachten Referenzkarten validiert und bewertet.

Im Business-as-usual-Szenario schätzt das PAM-Modell den nationalen gewichteten Durchschnitt der Grundrenten auf 248 USD/ha für das Zweikultursystem (Soja und Mais) und 188 USD/ha für Zuckerrohr. Dieses Verhältnis weicht jedoch in den Bundesstaaten Sao Paulo und Minas Gerais ab, da Zuckerrohr hier im Durchschnitt eine höhere Grundrente aufweist als der Zweikulturanbau. Ein Benchmarking der PAM-Produktionskostenschätzungen mit

beobachteten lokalen Daten zeigt eine zufriedenstellende Modellgenauigkeit mit einem relativen mittleren absoluten Fehler (rMAE) von weniger als 14 %. Der geringste Fehler, bei der Schätzung der Produktionskosten, wurde bei Zuckerrohr (rMAE von 8,7 %) und der höchste bei Mais in der zweiten Kultur (rMAE von 14 %) festgestellt. Räumlich gesehen hat das Modell eine bessere Leistung in der zentral-westlichen Region für Getreide und in der zentral-südlichen Region für Zuckerrohr. Dieser Umstand ist wahrscheinlich darauf zurückzuführen, dass die Modellparametrisierung auf die Haupterzeugungsregionen ausgerichtet ist und es in bestimmten Regionen keine typischen Betriebe gibt.

Umgekehrt ist die Ertragsschätzung mit dem biophysikalischen Modell aufgrund der unterschiedlichen Fehlermuster bei den verschiedenen Kulturen in Brasilien eine Herausforderung. Im Vergleich zu offiziellen Statistiken überschätzt das EPIC-IIASA-Modell in der Standardkalibrierung die Sojabohnenerträge (rMAE von 52 %), während es die Zuckerrohrerträge unterschätzt (rMAE von 24 %). Da dieses Fehlermuster die relative Rentabilität zwischen den Anbaualternativen künstlich verändert, werden die Ernteerträge für diese Fallstudie anhand verfügbarer regionaler Statistiken angepasst. Diese Vorgehensweise ermöglicht die Entwicklung und Prüfung der übrigen Module des PAM-Modellierungsansatzes.

Die Validierung der Business-as-usual-Landnutzungskarte zeigt, dass das PAM-Modell in der Lage ist, die aktuelle Landnutzung in Brasilien zufriedenstellend zu reproduzieren. Die visuellen und quantitativen Validierungsergebnisse zeigen eine starke Korrelation zwischen den verfügbaren Landnutzungskarten, wobei PAM für 86 % der gesamten Ackerfläche die tatsächlich zu beobachtenden Kulturen korrekt zuweist. In wichtigen Agrarstaaten wie Mato Grosso beträgt die Übereinstimmung zwischen den PAM-Ergebnissen und den beobachteten Daten sogar 95 %. Dieses interessante Ergebnis zeigt, wie wichtig eine sorgfältige Kalibrierung des Modells mit lokalem Fachwissen ist. Weiterhin verdeutlichts es, dass die verhältnismäßig einfachen Fruchtfolgen der spezialisierten brasilianischen Betriebe die Komplexität der Schätzung der Landnutzungsstrategien verringert und sich damit die Genauigkeit des Modells erhöht.

Um die Fähigkeit des PAM-Modells zur Vorhersage von Landnutzungs- und Ertragsänderungen aufgrund veränderter Rahmenbedingungen zu testen, wird eine Szenarioanalyse durchgeführt: Was passiert, wenn sich die Erträge der wichtigsten Nutzpflanzen infolge des Klimawandels deutlich verändern? Zu den Folgen des Klimawandels in Brasilien gibt es umfangreiche Literatur, wobei die erwarteten zukünftigen Ertragsentwicklungen regional und kulturspezifisch stark variieren. Um die Auswirkungen der Reaktion von Landwirten auf Ertragsänderungen infolge des Klimawandels zu simulieren, wurden zwei Anpassungsszenarien entwickelt:

- **Kein Zweikulturanbau in der tropischen Region (NoDC)**: Aufgrund der zu erwartenden Ertragseinbußen bei kurzzyklischen Sojabohnen, die in Zweikulturanbausystemen eingesetzt werden, gehen die Landwirte in den tropischen Regionen beim Sojabohnenanbau zu einem einfachen Anbausystem über.

- **Zweikulturanbau mit Ertragsverlust (DCYp)**: Die Landwirte produzieren weiterhin in den derzeitigen Zweikulturanbausystemen, nehmen aber einen Ertragsverlust bei Sojabohnen und Mais in Kauf.

Aufgrund des starken Rückgangs der Erträge aus dem Getreideanbau (bei Mais und Sojabohnen) in der tropischen Region – d. h. durchschnittlich 26 % in NoDC und 39 % in DCYp – wird simuliert, dass mehr als 24 % der derzeitigen Anbaufläche von Getreide- auf den Zuckerrohranbau umgestellt wird. Diese Ergebnisse variieren jedoch erheblich in den verschiedenen Regionen, wobei die am stärksten betroffenen Staaten Goiás, Paraná und Mato Grosso sind. Zusammen entfallen auf diese Staaten mehr als 55 % der gesamten Landnutzungsänderung. Die Ergebnisse zeigen zwei wichtige Aspekte auf: (a) Die Abnahme der Erträge aus Getreideanbauflächen ist im NoDC-Szenario in der Regel geringer und (b) Zuckerrohr dürfte vom Klimawandel profitieren, falls die simulierten Ertragseffekten eingehalten werden. Beide Szenarien deuten auf eine verminderte Verfügbarkeit von Getreide aus Brasilien hin, wobei die Getreideanbaugebiete auf Zuckerrohrproduktion umstellen werden. Eine stärkere Auswirkung wird für die Produktion von Mais als Zweitfrucht erwartet, da Landwirte von Zweikulturanbausystemen auf den einfachen Anbau von Sojabohnen umstellen werden.

Obwohl die Szenarioanalyse einen interessanten Trend zeigt ist es wichtig zu betonen, dass solch starke Auswirkungen auf die Erträge von Mais und Sojabohnen weiterführende Dynamiken auslösen dürften, die nicht berücksichtigt wurden. So ist beispielsweise zu erwarten, dass die Landwirte die Aussaatzeit anpassen, dürreresistente Sorten verwenden und Bewässerung einsetzen, wodurch das Ausmaß der Veränderung verringert würde. Darüber hinaus dürften auch die internationalen Preise auf eine solch signifikante Veränderung der brasilianischen Exporte reagieren und wiederum die Rentabilität der Anbaualternativen beeinflussen, was eine Anpassung des Angebots durch die Landwirte nach sich zöge.

Das Fehlen einer internen Nachfragefunktion ist eine wichtige Einschränkung des PAM-Ansatzes. Das Modell ist nicht darauf ausgelegt, ein neues Marktgleichgewicht intern zu schätzen, sodass sich Änderungen, wie sie in der Szenarioanalyse dargestellt werden, nicht auf die Hoftor-Preise und die Nachfrage auswirken. Um dieses Problem zu überwinden, sollte das PAM-Modell in das höher entwickelte Marktgleichgewichtsmodell integriert werden. So könnte es von der komplexen Schätzung der makroökonomischen Dynamik (z. B. GLOBIOM) profitieren sowie gleichzeitig die Schätzung der Produktionskosten verbessern. Eine solche Integration würde auch die Simulation der Landnutzung verbessern, da diese weitere für Brasilien bedeutende landwirtschaftliche Sektoren (z. B. Forstwirtschaft und Viehzucht) einbezöge.

Das PAM-Modell hat weitere Einschränkungen und Bereiche die weiterer Entwicklung bedürfen und sorgfältig geprüft werden sollten: (a) Aufgrund der starken Abhängigkeit von den Daten typischer Betriebe ist es wichtig, eine Ausweitung der Abdeckung durch typische Betriebe in der gesamten Forschungsregion in Betracht zu ziehen. (b) Die angepasste Aufstellung der Kostenfunktion erfordert tiefere Kenntnisse über die lokalen Produktionssysteme, welche nicht immer ohne weiteres verfügbar sind. (c) Das Modul für die Landzuteilung muss weiterentwickelt werden, um komplexere Produktionssysteme einbeziehen zu können. (d) Das Modell berücksichtigt keine Anpassungskosten, sodass jeder Zuwachs der Grundrente in einer Alternative zu Landnutzungsänderung führt. (e) Es ist notwendig das biophysikalische Modell zu verbessern oder regional kalibrierte Alternativen zu verwenden.

Die weitere Verbesserung des biophysikalischen Modells ist besonders wichtig, da die Erträge eine entscheidende Rolle bei der Rentabilitätsschätzung und stellvertretend in der Extrapolation der Produktionskosten spielen. Dem Standardmodell EPIC-IIASA fehlt derzeit die regionale Kalibrierung für brasilianische Kulturpflanzen beispielsweise Sojabohnen, welche in einem kürzeren Zyklus angebaut werden, um den Anbau von Mais als Zweitkultur zu ermöglichen. Eine kontinuierliche Zusammenarbeit mit lokalen Forschungsgruppen und Experten für Produktionssysteme (z. B. *agri benchmark* Partner) sollte es ermöglichen das Modell durch Feldversuche zu verbessern. Die Parameter für das Pflanzenwachstum könnten so besser kalibriert und saisonale Abläufe der landwirtschaftlichen Betriebe möglichst genau abgebildet werden, sodass sich regionale Unterschiede innerhalb des Landes besser wiedergeben lassen.

Aus konzeptioneller Sicht ist es auch wichtig hervorzuheben, dass das PAM-Modell eine gewinnorientierte Analyse der Angebotsreaktion von Landwirten für einen Zeithorizont zwischen 5 und 20 Jahren liefert. Dies ist nicht auf die methodische Notwendigkeit zurückzuführen, sondern vor allem auf konzeptionelle Schwierigkeiten bei der Schätzung deutlich kürzerer oder längerer Zeithorizonte. Es ist nicht zu erwarten, dass kurzfristige Preisschwankungen zu größeren Änderungen der Produktionssysteme (z. B. Umstellung von Getreide auf Zuckerrohr) auf regionaler Ebene führen. Den Landwirten ist generell bewusst, dass die Preise für Kulturpflanzen schwanken und eine Änderung der Aktivitäten weiterführendes Wissen und Investitionen erfordern kann. Eine langfristige Simulation (mehr als 20 Jahre) hingegen erfordert eine technische Anpassung der Kostenstruktur, bei der auch berücksichtigt werden sollte, ob grundlegende Änderungen im derzeitigen Produktionssystem zu erwarten sind. Der geeignete Zeithorizont für jede Analyse sollte durch die regionalen Merkmale der Kulturen (mehrjährig vs. einjährig) und die Möglichkeit bestimmt werden, die erwarteten Veränderungen im Produktionssystem mit den typischen Betriebsdaten abzubilden.

Diese Dissertation leistet einen Beitrag zur allgemeinen Entwicklung regionaler Modelle für die Simulierung von Angebotsreaktion durch Landwirte in Ländern mit begrenzter Datenverfügbarkeit, indem sie veranschaulicht, dass es möglich ist ein biophysikalisches Modell und ökonomische Daten auf Betriebsebene zu kombinieren und als Grundlage für eine Rentabilitätsschätzung mit hoher räumlicher Auflösung zu verwenden. Die Möglichkeit einzelne Kostenkomponenten separat zu schätzen, verleiht dem Modell die erforderliche Flexibilität für die Simulation markt- und politikbezogener Fragestellungen und liefert den Entscheidungsträgern rechtzeitig genaue Informationen. Der Bottom-up-Ansatz, der sich auf lokales Fachwissen stützt, ist eine wichtige Stärke des PAM-Modells, da er eine unrealistische Parametrisierung vermeidet und sicherstellt, dass die meisten lokalen Merkmale der Produktionssysteme in die Schätzung einbezogen werden. In Anbetracht des übergeordneten Ziels möglichst wenige Daten zu verwenden deutet die Modellgenauigkeit auf ein großes Potenzial zur Beantwortung von Forschungsfragen hin, wobei zusätzliche Parametrisierung und Integration die Leistungsfähigkeit weiter verbessern dürften.

Schlüsselwörter: Brasilien, Landnutzungsänderung, Biophysikalisches Modell, Produktionskosten, Angebotsanalyse, Sojabohnen, Mais, Zuckerrohr, Klimawandel, Wirtschaftlichkeit

Table of contents

List of Figures

List of Tables

List of Appendices

List of Acronyms

ABIOVE	Brazilian Association of Vegetable Oil Industries
ACM	AgriCostModel
AgMIP	Agricultural Model Inter-comparison Project
AL	Alagoas State
ANP	National Agency of Oil, Gas, and Biofuels
BA	Bahia State
BAU	Business-as-Usual
CEPEA	*Centro de Estudos Avançados em Economia Aplicada*
CGE	Computable General Equilibrium
CO_2	Carbon Dioxide
CS	Center-South Region
DCYp	Double Cropping with Yield Penalty Scenario
EMP	Econometric Mathematical Programming
EPIC	Environmental Policy Integrated Climate model
EU	European Union
FADN	Farm Accountancy Network
FAO	Food and Agriculture Organization
FOB	Free on Board
GEM	General Equilibrium Model
GHG	Green House Gases
GIS	Geographic Information System
GLOBIOM	Global Biosphere Management Model
GMO	Genetically Modified Organism
GO	Goiás State
GPM	Generalized Proximity Matrix
ha	Hectares
HRU	Homogeneous Response Units
IBGE	Brazilian Institute of Geography and Statistics
IF	Intensification Factor
IFPRI	International Food Policy Research Institute
IIASA	International Institute for Applied Systems Analysis
INPE	National Institute for Space Research
K	Potash
km	Kilometers

KTBL	*Kuratorium für Technik und Bauwesen in der Landwirtschaft*
LUC	Land-use change
MA	Maranhão State
MAE	Mean Absolute Error
MAgPIE	Model of Agricultural Production and its Impact on the Environment
MAPITOBA	Region including MA, PI, TO, BA states
MD	Minimum Data
ME	Mean Error
MG	Minas Gerais State
Mha	Million ha
MS	Mato Grosso do Sul State
MT	Mato Grosso State
N	Nitrogen
NoDC	No Double Cropping in Tropical Zone Scenario
NPLT	National Plan for Logistics and Transportation
OLS	Ordinary Least Squares
P	Phosphorus
PAM	Profitability Assessment Model
PECEGE	*Programa de Educação Continuada em Economia e Gestão de Empresas*
PEM	Partial Equilibrium Model
PI	Piauí State
PMP	Positive Mathematical Programming
PR	Paraná State
RFA	Representative Farm Aggregation model
rMAE	Relative Mean Absolute Error
RMSE	Root Mean Square Error
RO	Rondônia
rRMSE	Relative Root Mean Square Error
RS	Rio Grande do Sul State
SC	Santa Catarina State
SD	Standard Deviation
SIFRECA	Freight Information System (in Portuguese)
Simulation Units	SimU(s)
SOP	Standard Operating Procedure
SP	São Paulo State

SPAM	Spatial Production Allocation Model
t	metric ton
TO	Tocantins States
TRS	Total Recoverable Sugar
UDOP	Brazilian National Bioenergy Association
UNICA	Brazilian Sugarcane Industry Association
USA	United States of America
USD	United States Dollar

1.   Introduction

1.1.   Background

Understanding farmers' responses toward changes in framework conditions related to markets as well as production factors is of key importance for decision-makers. Market changes such as trade liberalization – e.g., European Union (EU) and Mercosur agreement, movements in crude oil prices, as well as production system changes (e.g., long-term reduction in yields), may shift prices and costs ratios between crop alternatives motivating farmers to respond by changing their land allocation to a different cropping alternative. These supply responses when aggregated may also induce more fundamental land-use changes, in the sense of converting idle or grassland into production, for example.

Commonly, supply response models can be classified into two main types: mathematical programming models (process-based) and econometric approaches, with the latter usually subdivided into several classes depending on their estimation strategies (Colman, 1986). The former is usually based on optimization functions depicting resource allocation within the farm and use of aggregation techniques to upscale results to represent sectors. The latter is mainly based on empirical macro-level data measuring the responsiveness of sectors or producers to price changes on acreage (Rao, 1989).

Although macro-level supply analyses with aggregated data are sufficient for several political purposes, for example analyzing total output and food available, farm-level analyses are of importance due to their ability to disaggregate results to regional and enterprise levels allowing a better targeting at specific regions or supply chains. Additionally, micro-level analyses increase the understanding of farmers' behavior and motivation while allocating their resources to produce a specific crop (Rao, 1989).

Process-based supply models for micro-level analysis are usually desirable for modeling regional farmers' behavior but due to the lack of regional statistics, this approach is usually not feasible for countries with weak statistics infrastructure – i.e., most developing and emerging economies. Basic information on production costs, acreage, and output is not always available at regional levels (Chen & Önal, 2012). An alternative for these countries is usually a case study approach in which typical farms are considered to represent the main production system and therefore representing the largest share of the output (Nehring, 2011; Osaki & Batalha, 2014). Although this approach is useful to understand production systems and provide detailed information on production costs, it lacks the regional representativeness that is required to estimate regional supply and land-use changes.

Against this background, there seems to be a need for a concept that allows farm-level supply estimation without requiring detailed farm-level datasets like Farm Accountancy Network (FADN) in Europe but still yields results within an acceptable error margin to inform decision-makers (Antle & Valdivia, 2006).

1.2.   Research scope

This dissertation aims to develop a Minimum Data (MD)[1] modeling approach that estimates the profitability of different crops and land use options at the farm level and mimic farmers' land allocation decisions. Modeling these decisions spatially explicitly allows us to understand land-use changes at regional levels in countries with low data availability. Thereby the following research question will be answered:

Can a combined biophysical and farm-level data model estimate the profitability and land allocation strategies of farmers in countries with low data availability?

I propose a bottom-up modeling approach that integrates crop management and costs data and a calibrated biophysical crop growth model. The basic idea is to use the calibrated biophysical model to generate yields and input use (based on crop management information) estimates, which are then attached to an economic module to estimate the farm-level profitability of each cropping system.

Due to limited time and financial resources, the proposed approach is developed and tested as a case study for Brazil, where data availability is moderate, allowing the validation of the model. Moreover, the Brazilian agriculture sector is of major importance worldwide, being among the major players in soybean and sugar production. In addition, its farmers are consistently faced with the choice between soybeans/maize and sugarcane production and, as given the importance of exports, Brazilian farmers are expected to react quickly to signals from the world market. Finally, land-use changes are of key importance for decision-makers due to the possibility of converting natural vegetation into arable land and so affecting important ecosystems.

To understand the model behavior in predicting supply responses due to changes in framework conditions, I carry out a scenario analysis on the land-use changes caused by the expected impact of climate change on crop yields. This empirical analysis should help the understanding of how climate change may affect the relative profitability of crops and lead to reallocation of agricultural land. The long-term impact of climate change in Brazilian agriculture is a relevant topic due to its importance in supplying world markets with commodities. Additionally, the scenario analysis should improve the understanding of the model's capabilities and limitations regarding expected land-use change caused by changes in framework conditions.

The underlying assumption for the approach is that farmers' cropping decisions are based mainly on the economic returns of each alternative. However, other important factors are expected to affect farmers' decision behavior that may not be fully represented by the

---

[1]   Minimum data approach as discussed by Antle and Valdivia (2006); aiming at timely and quantitative analysis that can provide sufficiently accurate analysis to inform policy decision making.

profitability analysis. For instance, the risk associated with each crop alternative (Liang, Miller, Harri, & Coble, 2011), technical recommendations regarding crop rotations (Arnberg & Hansen, 2012) and the role of perishability/storage of crops (Wright, 2011) are a few examples of additional factors that are expected to influence farmers' cropping decisions. These factors are not directly simulated in the model but are part of the overall discussion on farmers' behavior toward changes in the framework conditions.

1.3.   Dissertation structure

To better understand the different approaches available to estimate farmers' supply responses, the next chapter reviews the current literature on the most common methods used for such estimations. Thereafter, the framework conditions for the Brazil agricultural sector are described, focusing on the most important crops, such as soybeans, maize and sugarcane, as well as the crop rotations frequently employed by farmers.

Keeping the focus on the development of this modeling approach, the methodology chapter describes in detail setting up the different building blocks of the models as well as the procedure to calibrate the model for the case study in Brazil. To make sure the proposed model provides useful and robust results on farming economics and land allocation, the results from the calibration scenario (i.e., Business-as-Usual) is validated against observed data in a structured manner to highlight possible skewness of the model and how its different building blocks – e.g., transportation module or plant module – behave calibrated with information from recent years.

The validation chapter also lays out an interesting procedure to evaluate such models when it comes to assessing the ability of the approach to correctly allocate the available land to certain crops. To achieve a robust validation, I combine qualitative and quantitative methods to avoid, as much as possible, visual bias.

Thereafter, scenarios are proposed based on reviewing the most current literature available regarding the expected effects of climate changes on crop yields in the different climate zones in Brazil. The focus of this chapter is not only empirically understanding the expected impact on farming profitability and land-use changes in Brazil but also highlighting the applicability of the proposed model in answering such key research questions in countries with relatively low data availability.

Finally, the results from the different scenarios as well as the outcomes from the validation process are discussed and compared with the literature, focusing on possible areas for further development of the modeling approach, areas where the model can be useful, and its limitations.

2.   Farmers' supply responses and land use change

This chapter reviews the literature around the idea that farmers' supply responses lead to the reallocation of resources that may cause land-use change (LUC)[2]. Beginning with the estimation of farmers' supply responses, it focuses on the importance of such analysis, existing methods, and the level of analysis – e.g., nationwide vs regional, single vs multi-crop. In addition to the technical review, the following sections aim to understand existing models and how they are currently used.

Furthermore, I look at the modeling of land-use changes with a strong focus on economic-based approaches. A key aspect of this review is the specific estimation of production costs since they are a main driver of farming profitability. Production cost estimating is carried out differently in the models depending on the level of aggregation and techniques used to overcome problems with data availability.

Finally, I review the concept of minimum data and how framework conditions such as data availability, timeliness, and budget affect the complexity of methods when providing scientific information for decision-makers.

An important share of the supporting literature for the development of the proposed modeling approach is presented directly in the technical sections of this dissertation (see chapter 4), with relevant literature also presented in the model validation and scenario analysis chapters (6 and 7 respectively).

2.1.   Importance of farmer's supply responses

Understanding farmers' supply responses to changes in framework conditions such as prices, yields, and inputs should be a key concern of policy-makers. The aggregation of these effects impacts the overall economy, food prices, and availability, as well as the livelihood of those working in the agricultural sector (Rao, 1989). In addition to the overall economic impact (i.e., prices and quantities), a strong focus of the literature is to understand how different drivers such as trade liberalization (McKay, Morrissey, & Vaillant, 1999), governmental programs (Chen & Önal, 2012), and overall investments (e.g., roads and irrigation) affect agricultural supply responses and the stakeholders involved in the sector (Binswanger, 1989). That illustrates how comprehensive and important understanding farmers' reactions to framework conditions is and that policy instruments may impact the underlying conditions, shifting the agricultural supply in various directions. Understanding the magnitude and causes driving such land-use change is key to help decisions-makers make more rational decisions (Noszczyk,

---

[2]   The term land-use change in this work refers to a broader phenomenon including land-cover change – e.g., from grassland to arable land, as well as changes in agricultural practices (i.e., crops grown) within the arable land. The latter is also considered land-use change and, according to Foley et al. (2005), has important implications for the environment and the global food supply.

2019), especially in regions with important natural resources such as tropical forests and a competitive agricultural sector – e.g., Brazil, Indonesia, etc. (Richards, Myers, Swinton, & Walker, 2012).

In summary, estimating farmer's supply responses is key to improve the overall understanding of the linkage between policy and market-driven changes affecting key economic issues such as food availability and agricultural income, as well as broadening the knowledge regarding possible changes in land use and its effect on biodiversity. Due to its key role in economics, several methods are commonly used to estimate agricultural supply responses.

2.2.   Agricultural supply estimation methods

In the following subsection, I review the main characteristics of the several efforts made by researchers to estimate farmers' supply responses according to their parametrization (underlying functions), scope (single vs multiple crops), and spatial coverage (e.g., country, regions, farm). It is important to note that there are innumerable ways of classifying models and techniques and the basic classification and terminology used here comes from the work of Colman (1986) and Rao (1989).

Econometric vs mathematical programming approach

According to Colman (1986), most of the methods used for estimating agricultural output supply can be classified into two categories (a) econometric[3] and (b) mathematical programming. The former includes most of the estimation work, using several functions and parametrization to econometrically estimate output supply response. The resulting elasticities represent the output change caused by movements in output price, for example. This category can be subdivided according to the estimation techniques – see Colman (1986) for a detailed discussion.

The mathematical programming or process-based approach is widely used in different simulation of supply response problems at sector, regional, and farm levels. This is mainly due to its computational efficiency and the bottom-up type of analysis (Chen & Önal, 2012). A strong feature of this approach is the idea of using a representative farm aggregation model (RFA), wherein representative farm characteristics are used to estimate the resource allocation at the farm level, which is later upscaled to a regional or sector level following different strategies (Sharples, 1969). Colman (1986) argues the main attractiveness of this approach comes from the ability to represent complex multi-product relationships as observed in real world farms by starting with a known technology at the farm level and upscaling to represent a sector. The latter is also a major problem with this method since the aggregation is likely to generate biases. There are a diversity of innovative approaches to overcome or minimize the aggregation bias, which can be found in the work of Chen and Önal (2012).

---

[3]   Or "dual systems of supply and input demand equations" according to Heckelei and Wolff (2003).

An interesting enhancement to the mathematical programming models to estimate agricultural supply is the implementation of Positive Mathematical Programming (PMP) promoted by Howitt (1995). It allows the perfect calibration of models at the baseline year based on observed information by adding non-linear terms in the objective function (Heckelei & Britz, 2005). This development has increased the applicability of agricultural supply programming models when tackling complex regional sectors for the purpose of policy analysis. According to Heckelei, Britz, and Zhang (2012), all models using PMP are either developed in Europe or North America, relying strongly on the detailed farm-level information observed in these regions.

Lately, a growing body of literature has focused on combining the econometric and mathematical programming approaches (more specifically PMP) to create Econometric Mathematical Programming (EMP). According to Heckelei and Wolff (2003), this alternative approach estimates simultaneously all the parameters in the model, improving significantly upon the standard PMP. An empirical application of the combined EMP model can be found in the work from Britz and Arata (2019).

Finally, it is important to highlight that the selection of the estimation technique depends strongly on the questions to be answered as well as the availability of information (Colman, 1986). The ability to introduce constraints or shocks, such as new policy or technology, that allow economists to project structural changes makes mathematical programming an interesting tool for policy decision support systems. The introduction of PMP has improved this feature even further, motivating the development of models using PMP (Heckelei & Britz, 2005). This development matches a recent goal of multidisciplinary research connecting, for example, environmental questions with economic supply responses, such as the impact of economic changes on public goods (Heckelei et al., 2012).

Aggregated vs single crop analyses

Another important characteristic of farmers' supply response estimation is the differentiation between single and aggregated crop analyses. Binswanger (1989) argues it is crucial to look at the aggregated crop level because the most important farming inputs (e.g., land, labor and capital) are fixed in the short term. He adds that any growth in a single crop takes up resources from other crops, leading to a much lower supply response at the aggregated level if compared with single crop estimation. Any significant output increase comes only by adding more resources or changes in technology (Binswanger, 1989).

Supporting the idea of including multi-crop analysis is the idea that farmers consider resource allocation based on a rotation instead of single crops, due to technical considerations such as pest and disease pressure as well as to reduce their risk (diversification). Furthermore, considering a mixed-crop alternative minimizes the overspecialization problem commonly found in mathematical programming (Chen & Önal, 2012). Finally, McCarl (1982) argues that if the crop mix approach is employed at the farm-level decision process, upscaling to a sector level should be less problematic.

Estimation level: country, region, and farm

One major feature of agricultural supply response analyses is the scale of the estimation. Technically, both types of supply estimation methods can be used at all levels but, due to the lack of available data at regional and often at sector levels, the mathematical programming model is commonly preferred at the micro-level. The reason is the possibility of using a set of representative farms to calibrate the model and, if enough information is known about the overall population of farms, scale up the results to sector levels (Colman, 1986).

Such a bottom-up type of model (i.e., from farm to sector) is widely used in Europe due to the availability of single farm records from the FADN, the main source for most of the PMP models. It is important to note that in addition to the availability of data, the increased use of PMP models in Europe also is related to the shift in agricultural policy from price support to advanced instruments linked to environmental goods, for example, which are easier to model using mathematical programming approaches (Heckelei et al., 2012; Heckelei & Britz, 2005).

In most policy support systems, it is key to be able to disaggregate the national or macro-level supply response analyses to regional and even micro-regional levels. According to Rao (1989) regional studies allow researchers to (a) break down the overall policy impact to specific regions as well as farm types (e.g., size, income) to better understand possible policy mismatches and allow better fine-tuning, and (b) to identify whether the policy target group is motivated to react based on the selected instrument (e.g., subsidy, payments).

Summary of agri supply estimation methods

Farmers' supply responses models are generally classified into two categories: (a) econometric and (b) mathematical programming based on their estimation techniques. The selection of the estimation method depends primarily on the research questions and the availability of data. PMP models are interesting for complex and multidisciplinary analyses and have been widely used in the European context, in part due to the availability of single-farm records. Supply response analyses should aim at modeling the actual cropping alternatives available to farmers in terms of crop rotations or mix instead of single crop analyses. Finally, it is important to consider the level at which the analysis is carried out. For policy support, it is preferable to disaggregate the result to regions or farm-level to increase the understanding of the impact of the policy instrument on different farm types. That is even more important if the analysis involved the impact on environmental resources, in which the spatial allocation of agricultural production may play a stronger role.

2.3.   Modeling land use

Building on the idea that local decisions affect the allocation of resources, it is interesting to expand the discussion to consider that farmers' supply responses may affect land use. Land-use change (LUC) research is an extremely broad and interdisciplinary field. Noszczyk (2019)

reviews the main methods used to study land-use change, with detailed classification and evaluation of the most common approaches.

Focusing on the economic-based LUC models, the main idea is that local agents (i.e., farmers) aim to allocate their resources (land) to maximize profits (Dang & Kawasaki, 2016). Thereby, market and policy changes affecting farmers' returns may lead to the reallocation of land and, in turn, land-use change. It is important to note that the term "land-use change" is used primarily to indicate a modification in the land cover, such as forest to arable land. However, a broader meaning also should include moving from one cropping system to another (Veldkamp & Lambin, 2001). According to Dang and Kawasaki (2016), most economic-based models lack spatial attributes, focusing mainly on the overall impact and change rate. That represents an important drawback that, however, can be partially compensated for by integrating different models.

Land-use change models commonly address two distinct questions: (a) spatial change or (b) quantity of change. The former is much easier to assess as it is based mainly on linking landscape attributes that may cause land use, such as roads. The latter involves understanding the underlying drivers, which usually are complex systems of macro-economic and policy changes (Veldkamp & Lambin, 2001).

Agro-economic land-use models

Following the recent focus on understanding the drivers and effects of land-use change, new spatially explicit agro-economic models have been developed to consider the economic behavior while understanding the spatial element of land-use changes. Schmitz et al. (2014) review the 10 models within the Agricultural Model Inter-comparison Project (AgMIP), which are commonly partial (PEM) or general equilibrium models (GEM). The authors argue that the spatial dimension is key for modeling economic behavior in land-use change and that lately, more global databases are available to fulfill the data requirements for such a model.

According to Schmitz et al. (2014), the Global Biosphere Management Model (GLOBIOM) and the Model of Agricultural Production and its Impact on the Environment (MAgPIE) are the only agro-economic land-use models constructed with grid-specific optimizations. That allows these models to make use of disaggregated global data. GLOBIOM has a high-resolution resulting in ca. 200,000 simulation units (SimUs) compared with ca. 60,000 from MAgPIE, and only 114 from most of the computable general equilibrium (CGE) models. Schmitz et al. (2014) conclude that the spatial PEM models (i.e., GLOBIOM and MAgPIE) have the strength of fine resolution, accounting for the spatial heterogeneity that is crucial to model land-use changes. However, further research is necessary to better understand important dynamics in supply responses as well as the conversion costs between activities (e.g., pasture to crop land) to reduce the models' dependency on exogenous assumptions.

Summary of modeling land use

Global economic modeling recently has focused strongly on including land-use changes in defined framework due to the relevance of the topic and the need to include agro-economic behavior to improve LUC modeling. Analyses primarily focus on spatial effects of land-use change but quantification is crucial. Finally, the spatial partial equilibrium models such as GLOBIOM and MAgPIE have contributed significantly to the field because their fine spatial resolution accounts for biophysical heterogeneity.

2.4.   Estimating production costs

In the context of farmers' supply responses and economic-based land-use models, an important component is the estimation of production costs (or cost functions) for agricultural activities. Based on the idea that farmers allocate their resources to maximize their profit (returns), it is evident that production costs play a significant role alongside revenues. For econometric agricultural supply response models, the cost function is directly estimated from available data (Chavas & Cox, 1995). Programming models also can be based on econometric estimation if data is available, or on cost information from representative farms (Colman, 1986).

For developed regions such as Europe, large datasets containing single-farm records are available (e.g., FADN) allowing modelers to use them as the basis for the production cost estimations for different agricultural systems (Heckelei & Britz, 2005). Unfortunately, that is not the case for most developing countries, which may have information on representative or average costs but rarely for multiple geographic locations (Antle, Diagana, Stoorvogel, & Valdivia, 2010).

To illustrate the extent of this data availability problem, one of the most spatiality explicit agro-economic models, GLOBIOM, has only recently developed a disaggregated-cost module – AgriCostModel (ACM). It utilizes technical coefficients from a German database (KTBL[4]) as well as statistical data (e.g., FAO) to calculate production costs, which are extrapolated spatially using an "intensification factor". This factor is calculated using the estimated yields and fertilizer application in the different scenarios (Deppermann et al., 2018). Such developments are likely to improve the overall performance of the model but depend on assumptions linking the cost items with the intensification factor[5] for the extrapolation to the spatial units.

---

[4]   www.ktbl.de

[5]   According to Deppermann et al. (2018) nitrogen, phosphorus, and irrigation requirements are explicitly given by the biophysical model (EPIC) at the simulation unit level. All the remaining costs are given at a monetary unit [fix] and extrapolate spatially using the intensification factor - except for labor, fuel, and seed that are fixed for each management option.

Summary of production cost estimation

Production cost datasets with suitable spatial coverage are rare outside of developed countries. That has led to the development of different estimations and/or extrapolation techniques to cope with this data deficit. However, these techniques rely on assumptions of the cost behavior of different components regarding productivity or management intensity, for example. That requires a thorough understanding of production cost behavior to reduce the assumption bias.

2.5.   Minimum data and policy advice

Considering that policy-makers are fundamentally interested in understanding how policy instruments may motivate (or demotivate) farmers to take a specific course of action (i.e., grow a specific crop, expand their land), it is important to consider the effectiveness of each method mentioned in providing policy advice. In essence, all discussed methods are suitable to develop a policy-supporting system to inform policy makers. However, it is crucial to understand that high-resolution economic data (for example) is not widely available, which makes sophisticated modeling challenging.

Minimum Data approach (MD)[6]

Antle and Valdivia (2006) argue that there is a strong demand for timely and quantitative analyses that provide results within an acceptable accuracy to inform policy-makers. They further discuss that this level of accuracy is likely to be lower than expected at scientific publications but should add important quantitative information to rather qualitative analyses carried out otherwise.

The idea is that the underlying requirement for a land-use model is the expected economic returns from each competing alternative at each geographic location. If such information is available, this site-specific data can be used to estimate the necessary equations. However, the authors state that such information is hardly available. They propose a MD approach that allows the use of secondary data (e.g., representative farms) as well as expert judgment, when necessary, to fill the gap. The MD approach, according to the authors, has important characteristics, such as:

- providing timely and policy-relevant analyses for a range of countries with limited data availability,

- comparatively easier to learn and implement in developing countries with relatively low costs,

- availability of the data and training needed to implement in regions such as Latin America and Africa.

---

[6]   The whole subsection is based on the approach developed by Antle and Valdivia (2006).

The authors conclude that economic modeling based on the MD approach can provide meaningful results with sufficient accuracy to help to inform policy-makers (Antle & Valdivia, 2006).

In a recent work, Antle et al. (2010) state that since the introduction of the MD approach, several analyses were made, mainly in developing countries, confirming their hypothesis that such models have a place within agro-economic modeling. The authors further develop the model for subsistence farming and conclude that the MD model also performs for such farming systems.

Summary of the minimum data approach

Against this background, it is important to consider that if the research goal is to advise policy-makers, a trade-off between the level of precision and timeliness is always expected. This effect is likely to be stronger in regions with limited data availability. That should motivate researchers to balance the expected accuracy of results with the level of data available and other framework conditions such as time and budget.

2.6.   Farm-level economic data

Due to the lack of spatially specific statistics on production costs, an alternative solution may be the use of representative farms. The *agri benchmark* network provides typical economic data for several production systems worldwide. The network is non-profit and non-political, coordinated by the Thünen Institute of Farm Economics in Germany.

***agri benchmark* typical farms**

The typical farms in the *agri benchmark* network are collected following an internationally standardized operating procedure (SOP), ensuring that the steps for data collection, validation, and update are applied equally worldwide, safeguarding the quality and comparability of the production system data (Zimmer & Deblitz, 2005). The basic idea is to understand the production system of the most common type of farm in the main production regions in a country, following the argumentation from Elliott (1928) that those are more representative than the average farm for a specific region (i.e., representative farm).

Primary data are collected in focus group discussions with 5 to 10 participants, including farmers, researchers, and regional advisors. The data are not based on an existing farm, but rather collected by reaching a consensus among the participants on the "most common" practice, including fertilizer rates, machinery setup, crop rotations, etc. (Krug, 2013). The construction of a synthetic typical farm instead of selecting an existing farm avoids two problems: (a) the influence of top management on the results, even if the farm is typical in terms of resources (i.e., size, machinery, climate); and (b) the problem in disclosing an individual's sensitive data (Feuz & Skold, 1992).

Finally, the typical farm information is processed by the TYPICROP model, a whole-farm model that computes all major economic indicators, allocating all costs and revenues to each specific crop grown on the farm. A detailed description of the model can be found in the work of Hemme (1999) and Nehring (2011). The model has been adapted for sugarcane analyses, accounting for the differences compared with grains, such as semi-perennial cycle, sugar yields, and varying harvest cycles (Balieiro, Witte, & Weerathaworn, 2016). A detailed explanation of the *agri benchmark* typical farm approach can be found in the work from Chibanda et al. (2020).

Typical farms in supply estimation analyses

Using typical farms to understand farmers' behavior and adaptation to changes in markets and policy is not new and has been discussed by several authors in the agricultural economic field (Carter, 1963; Elliott, 1928; Plaxico & Tweeten, 1963). According to Feuz and Skold (1992), the benefit of using typical farms can be summarized as: (a) lower time and cost required to gather farm data compared with individual data collection; (b) the ability to represent different types of farms within a region compared with averages; (c) possibility to easily model technical and framework changes affecting the production systems. These authors also reinforce the benefits of using typical or model farms instead of average farms for economic analyses as well as the benefit of selecting these farms based on knowledge instead of random selection, which is the main component of the SOP developed within the *agri benchmark* network.

Specifically, analyses of farmers' supply responses using typical or representative farms have been carried out for several decades. The approach gained strong attention in the 1970s due to linear programming and improved computing capacities. This technique, usually referred to as RFA is interesting to reproduce multi-crop analyses and, depending on the farm population, the results can be upscaled to sector levels (Sharples, 1969). This approach has also limitations, which are associated mainly with the "aggregation of results" and "selection of farms" affecting the reliability of the analyses.

Summary of typical farm approach

Using typical farms in agricultural economics is not new. The benefits are the possibility to understand complex systems within the farm (crop rotations) and the ability to use the information to calibrate mathematical programming models. The *agri benchmark* network has a number of typical farms worldwide and the data may be suitable for the estimation of farmers' supply response in countries with a lack of production cost data. To a certain extent, results could be upscaled to sector levels, but the selection of farms is key to reduce aggregation bias.

2.7.   Summary

Considering most developing countries, using a standard econometric agricultural supply response model may not be desirable due to lack of disaggregated data and the accompanying difficulties to simulate framework changes (e.g., taxes, prices, climate change) to inform

decision-makers. Using a mathematical programming model therefore is more suitable but the caveat of data availability prevents the use of more sophisticated models (e.g., PMP) commonly employed in Europe and North America. Including the land-use change component is crucial to understand how farmers' resource allocation behavior may affect the overall landscape. Global agro-economic land-use models have been used extensively for these questions, but the lack of a detailed agricultural production cost estimation is expected to influence the accuracy of the model.

Against this background, using existing information from typical farms to improve the estimation of production costs and profitability while having a biophysical model to enhance the spatial resolution should improve the agricultural supply response estimations for countries with limited data availability. That motivates the development of a new model based on minimum data values to simulate farmers' reactions to changes in framework conditions and resulting land-use changes, providing timely information with sufficient accuracy to decision-makers.

3.   Land use and cropping system in Brazil

Empirically, the proposed supply response model is tested for Brazil, focusing on the main crops produced and possible changes to the current land use. To understand the setting up of the model and simulations, it is important to be aware of recent developments in land use, the spatial distribution of crops, and the production system or options that are available for Brazilian farmers. A deeper understanding of the cropping alternatives is key to ensure that the model is calibrated with real, practical options that farmers face when allocating their land resources.

This chapter briefly explains the current land use in Brazil and describes the importance of the crops that are considered in the case study as well as their spatial distribution. Furthermore, the main drivers of the production expansion are analyzed to identify possible drivers for the future – i.e., yield increase vs area expansion. Finally, I describe the production alternatives that are typical in Brazil, which are the basis for the case study setup.

3.1.   Current land use

Brazil is a country with a total area of approximately 851 million ha. Of this total, Brazil uses approximately 8% for farming (incl. perennial and annual crops) and 13% for managed pasture, while 58% of its territory currently has native vegetation[7]. In absolute terms, ca. 66 million ha are being farmed and can be considered arable land, whereas 113 million ha are classified as managed pasture or grassland (IBGE, 2020). Figure 1 shows the most recent land-use map from the Brazilian Institute of Geography and Statistics (IBGE)[8] in 2016 as well as the administrative regions as defined by IBGE.

---

[7]   It includes native grassland.

[8]   In Portuguese - *Instituto Brasileiro de Geografia e Estatística* (www.ibge.gov.br)

Figure 1:   Land classification and regions in Brazil
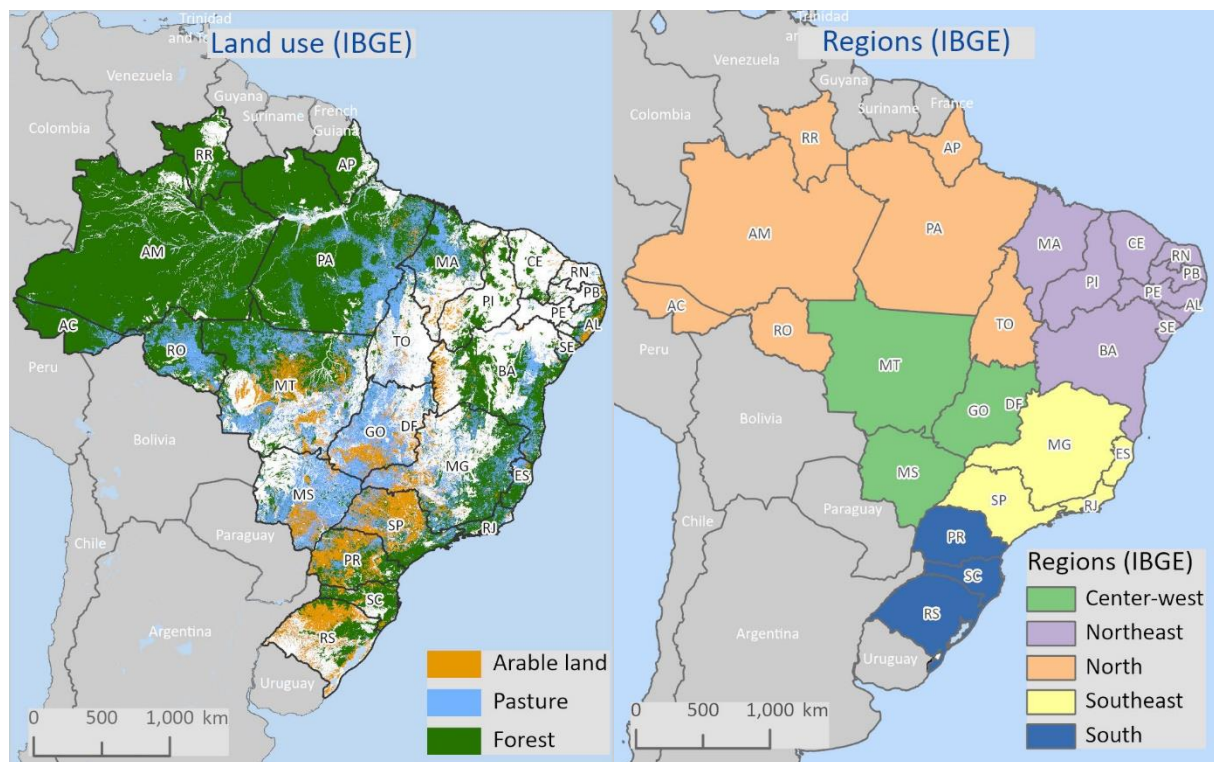
Source: Landcover from IBGE (2018) – created by the author

Currently, most of the arable land is concentrated in the southern regions, including the states of São Paulo (SP), Paraná (PR), and Rio Grande do Sul (RS) as important producers. Another important region is the Center-west[9], including the states of Mato Grosso do Sul (MS), Goiás (GO), and Mato Grosso (MT) . Pasture or grassland is commonly found throughout the central and southern regions, but the highest occurrence is in the Center-west.

Breaking the arable land down to crop levels, from a total of 73.2 million ha of seeded area[10], more than 47% is used for soybean production, 23% for maize and almost 14% for sugarcane (IBGE, 2019). As shown in Figure 2, these three crops play a major role in the Brazilian agricultural sector, currently accounting for roughly 84% of the total seeded area.

---

[9]   Interchangeably called "Cerrado" for this dissertation.

[10]   Including first, second and third season. For example, a hectare with double-cropped soybeans/maize is counted twice. The arable land in total is around 66 million ha.

Figure 2:   **Share of the seeded area as a sum of soybeans, maize, and sugarcane in the total seeded area of each region in Brazil (1989-2019) – in %**
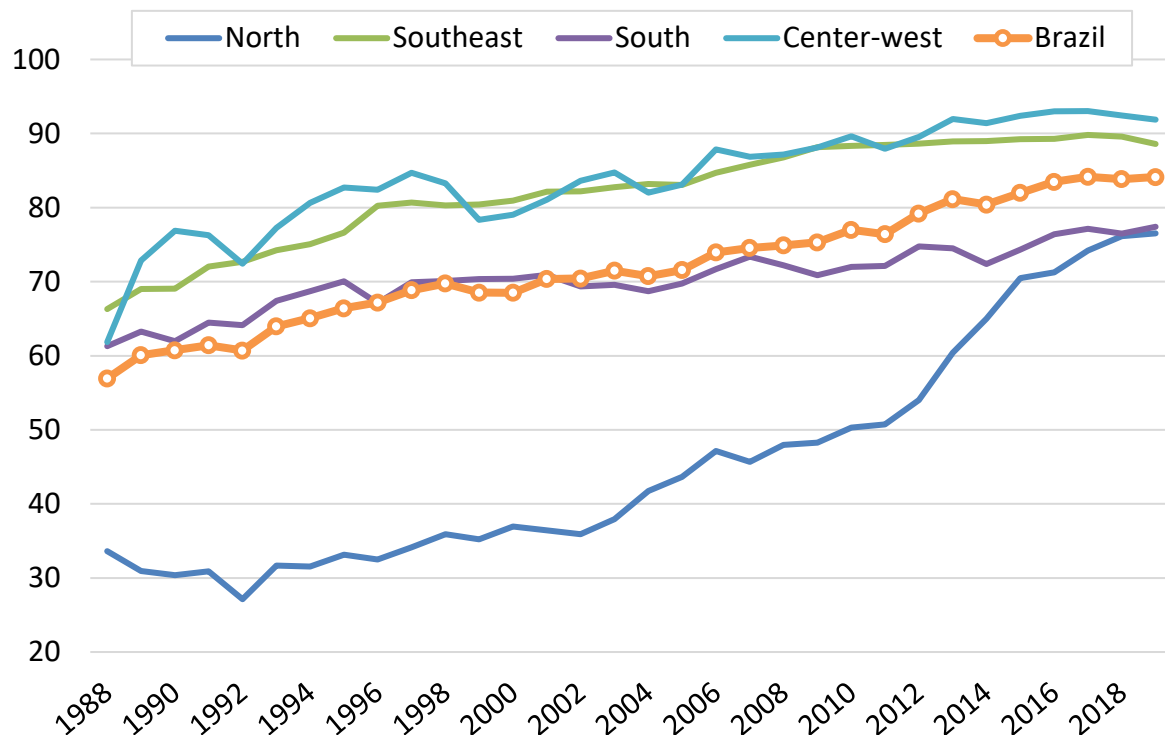
Source: IBGE (2019) – created by the author.

Looking back at the period between 1989 and 2019, besides the overall expansion of arable land, the relative importance of sugarcane, soybeans, and maize have increased in all regions in Brazil. For example, in the Center-west region, the share of total seeded land devoted to these three crops combined went from 62% in 1989 to 92% in 2019. In the same period, total arable land in this region increased by 3.5 fold, from 8 million ha to more than 29 million ha (IBGE, 2019). The combination of these two effects – i.e., overall acreage increase (more area) combined with relative gains in terms of their share in total arable land (displacing other crops) – partially explains the agricultural boom observed in these three crops in Brazil.

Against this background, the focus should be on sugarcane, soybeans, and maize to understand the recent and future land-use changes in Brazil related to crops. Focusing on these three crops has the benefit of using the data currently available from *agri benchmark* as well as avoiding additional complexity to depict crops of minor importance. Moreover, the recent acreage trend indicates that these crops have taken acreage from other crops, highlighting their importance in the agricultural sector in Brazil.

Besides the main crops, the beef sector in Brazil is important, with managed pasture currently occupying 13% of the total area, which represents 113 million ha (IBGE, 2020). Due to the predominant position of the crop alternatives and the beef sector, it is important to understand their regional distribution and how they typically are integrated into production systems.

3.2.   Sugarcane

Brazil is the largest sugarcane producer in the world, responsible for 23% of total global production and more than 50% of sugar exports in recent years (USDA, 2021). Sugarcane production increased from 253 million t in 1989 to 753 million t in 2019. However, the development of sugarcane production in Brazil started in the 1970s, driven mainly by governmental programs to foster ethanol production (Goldemberg, Coelho, & Guardabassi, 2008; Matsuoka, Ferro, & Arruda, 2009).

A second rapid expansion in sugarcane production took place between the years 2000 and 2015 driven mainly by increasing international and domestic demand for ethanol and sugar, higher industrial productivity and the introduction of mechanical harvesting (Cherubin et al., 2021). For this case study, it is important to understand the key geographical and farming features of sugarcane production in Brazil. Figure 3 shows the spatial development of the sugarcane areas, comparing 1989 and 2019.

Figure 3:   Sugarcane seeded area comparing 1989 and 2019 at the municipality level in Brazil (in ha)
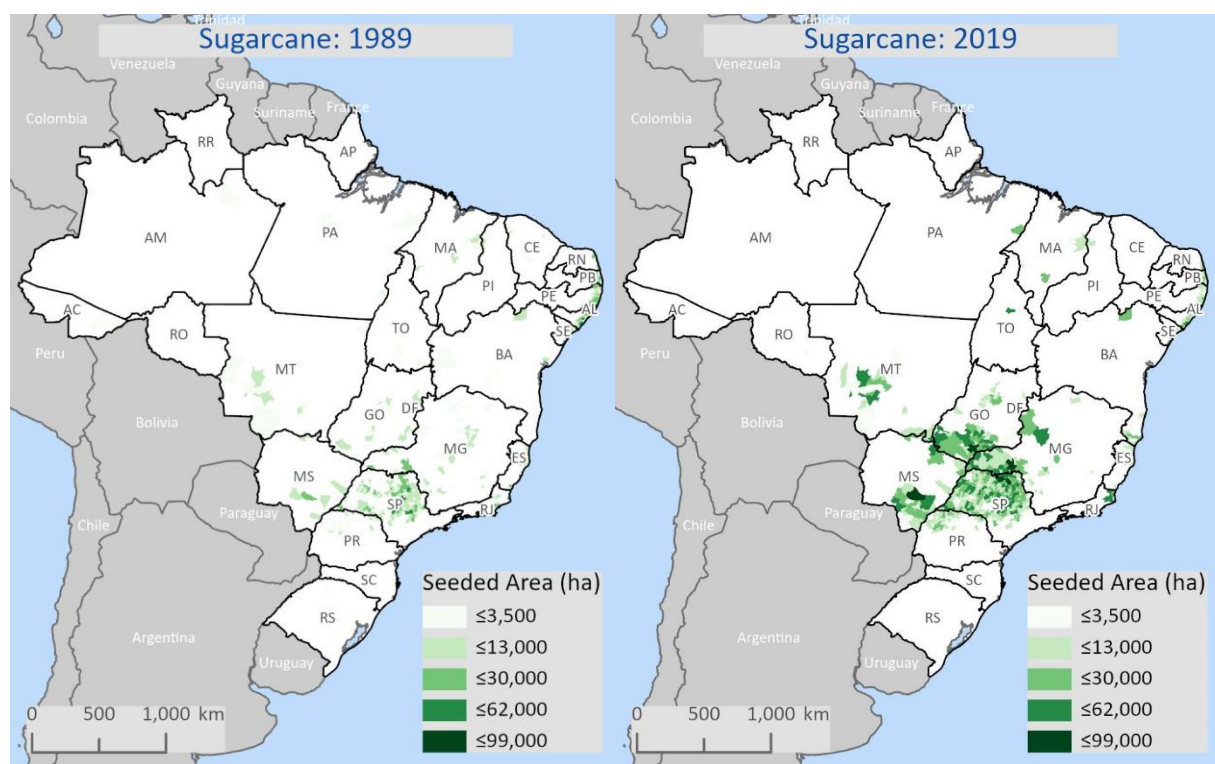
Source: IBGE (2019) – created by the author.

Sugarcane area expanded 2.5 folds in the period between 1989 and 2015, from ca. 4 million ha to more than 10 million ha (IBGE, 2019). According to Cherubin et al. (2021), this expansion took place mainly in pasture areas in the central-south (CS). Since 2015, sugarcane acreage has been stable at around 10 million ha. Spatially, the production stayed concentrated around

the main cluster in the southeast region, with the state of São Paulo (SP) representing ca. 55% of the total area in Brazil (see Figure 3). Nonetheless, it is important to recognize the expansion of sugarcane into the Center-west region mainly into the states of Mato Grosso do Sul (MS) and Goiás (GO) and, to a lesser extent, to Mato Grosso (MT). This is the first indication the land expansion of sugarcane may be putting pressure on other crops competing for land as well as allowing farmers from these new regions to switch to sugarcane, if profitable.

A typical sugarcane production system in the Southeast of Brazil is rainfed (Bordonal et al., 2018), with farmers growing sugarcane, on average, for six years as a semi-perennial crop. Within this cycle, the crop is harvested five times during the crushing season that usually extends from March to December (Balieiro et al., 2016). Due to its multi-annual feature, it is unfeasible to have a proper interannual crop rotation, leading to a monoculture at least for the period between planting and the last harvest. Differently from most other crops, sugarcane is perishable and must be processed locally in sugar mills that produce mainly sugar, ethanol and energy for the domestic and international markets (Cherubin et al., 2021).

3.3.   Soybean and maize

Soybeans

As for sugarcane, Brazil is a major international player in soybean production. In 2019, Brazil was the biggest producer, with an output of 114.2 million t of soybeans, followed by the USA with 96.8 million t. Brazilian soybean production represented ca. 34% of the worldwide production in 2019 with a strong focus on exports (FAOSTAT, 2021). According to the International Trade Center (ITC) in 2020, Brazil exported ca. 83 million t of soybeans to the world, with China buying 73% (i.e., 60.6 million t) of the total exports from Brazil (ITC, 2021).

Until 1960, soybeans were a relatively unimportant crop in Brazil. Due to governmental programs and increasing international prices, soybeans started to expand in the subtropical regions in the south. However, the most notable expansion started in 1980 due to the development of cultivars that were less sensitive to photoperiodic variations, allowing the crop to develop properly in lower latitudes (Cattelan & Dall'Agnol, 2018). This rapid development changed the agricultural sector completely in Brazil and has been the focus of several studies aiming to understand the link between this expansion and land-use changes (deforestation). Macedo et al. (2012) argue that the expansion of soybeans was overwhelmingly into previously cleared pasture areas, which is in line with the finding from Barona, Ramankutty, Hyman, and Coomes (2010). However, both sets of authors suggest that it is possible to have leakage effects of soybeans displacing cattle production into areas close to the Amazon basin, inducing deforestation.

Figure 4 illustrates the expansion of soybeans in Brazil, moving mainly from the southern (subtropical) regions to the tropical Center-west regions centered around the state of Mato Grosso (MT). Furthermore, soybeans have also reached northern states in the region known as

MAPITOBA – i.e., Maranhão (MA), Piauí (PI), Tocantins (TO), and Bahia (BA), which is currently considered the expansion frontier in Brazil (Cherubin et al., 2021).

Figure 4:   Soybeans seeded area comparing 1989 and 2019 at the municipality level in Brazil (in ha)
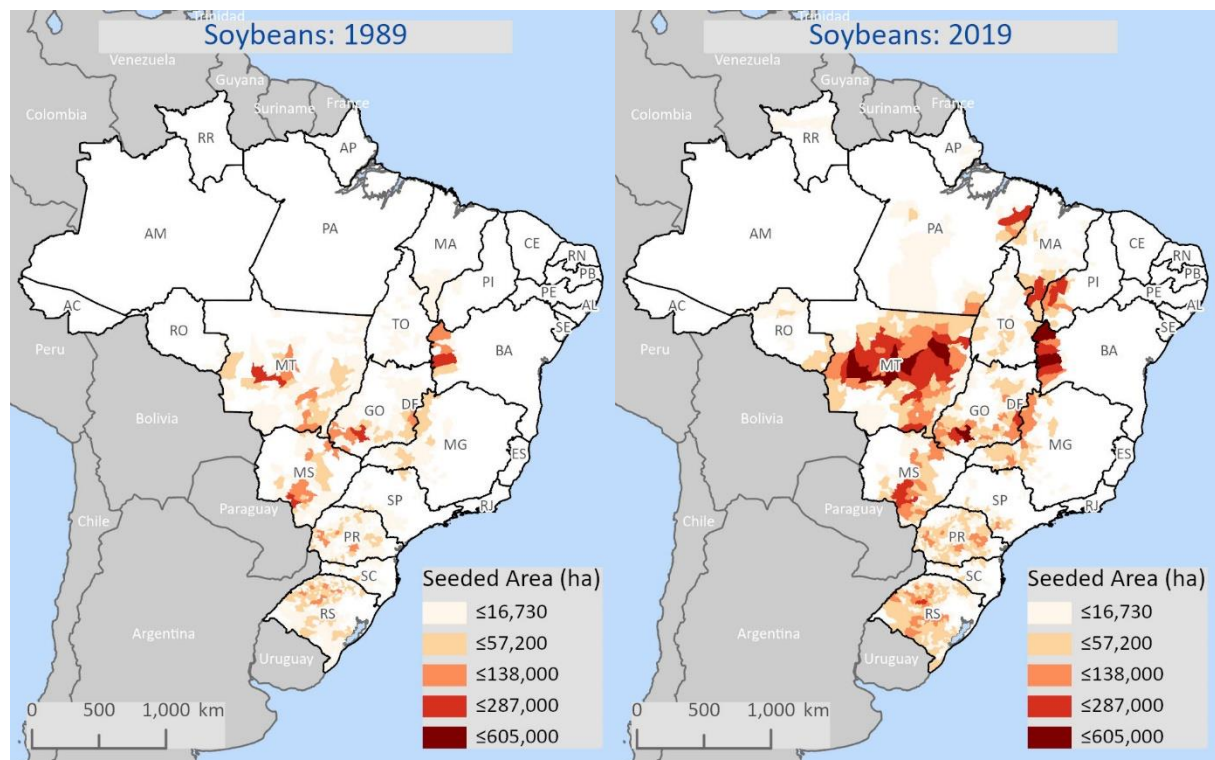
Source: IBGE (2019) – created by the author.

Compared with sugarcane, the spatial expansion of soybeans is much more prominent. Soybeans moved strongly northward as well as increasing in area within the states that were already producing in 1989. It is interesting to note, that the states of Goiás (GO) and Mato Grosso do Sul (MS) experienced a strong increase in area which is in line with the sugarcane expansion, reinforcing the hypothesis that competition for land has intensified. Finally, the maps are based on information at the municipality level (spatial boundaries), which should be not confused with the actual area. That explains why sugarcane and soybeans are strongly represented in the same location.

Double cropping with maize

Double cropping is a very common practice in grain[11] farming in Brazil. It is, to some extent, a result of breeding soybeans for shorter photoperiods, which allows farmers to adapt their planting calendar to better use the wet season (Abrahão & Costa, 2018). A typical practice is

---

[11]   Throughout the dissertation, the term "grain(s)" refers to grain crops in a broader definition, involving cereals and legumes – more precisely, soybeans and maize - van Alfen (2014).

to seed soybeans in summer (October to January/February) followed by maize grown from February to June (Pires et al., 2016). The system can be adapted to other crops such as cotton (Center-west) or wheat in the subtropical region, depending on the growing conditions. The share of the soybean acreage that is seeded with maize in winter depends on several factors, such as precipitation, the cycle length of soybeans, and farmers' machinery setup (Osaki & Batalha, 2014). Figure 5 illustrates a typical double-cropping system with soybeans as the leading crop in summer and maize in winter and the main operations.

Figure 5: Double-cropping system of soybeans and maize and its main operations – based on a typical farm from Mato Grosso (MT)
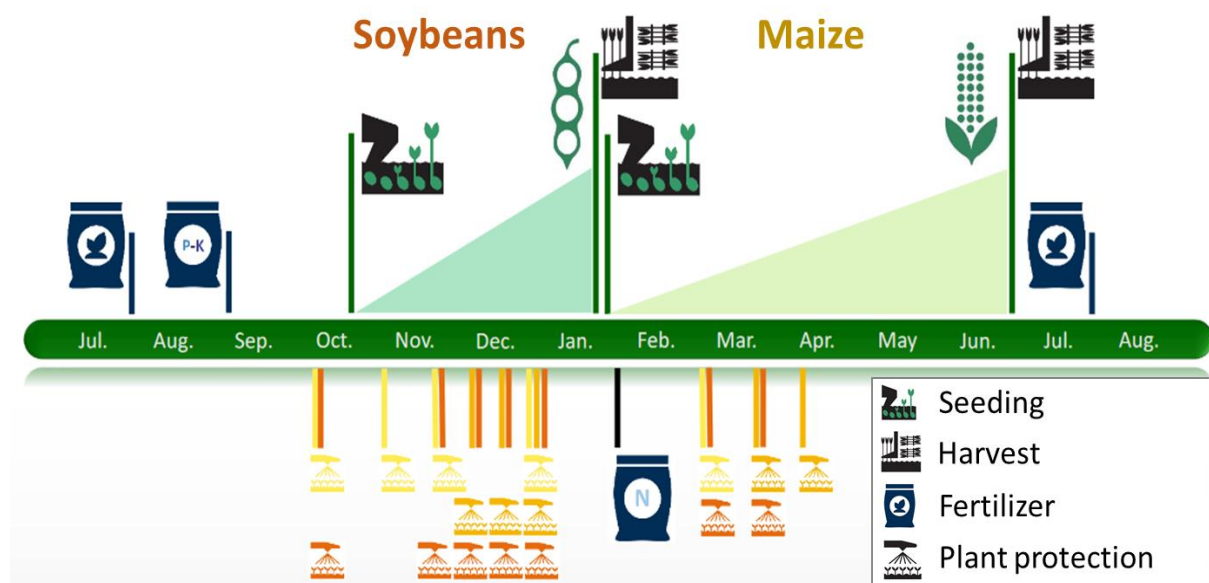
Source: based on the typical farm BR1300MT from agri benchmark (2020) – created by the author.

As Figure 5 highlights the sequence of operations in greater detail, it indicates the intensity of plant protection products during soybeans' short cycle and, more importantly, the peak of operations between the last sprayings in late December, harvest of soybeans, and the simultaneous seeding of maize. The short time window for harvesting soybeans and seeding maize has a strong influence on the share of the soybean area that is sown with maize afterwards. Delaying the seeding of maize due to rainfall or machinery problems increases the risk of weather deficits for the full development of the crop, reducing yields (Pires et al., 2016).

The importance of double cropping soybeans with maize is shown in Figure 6. The information on first-season maize represents the system of growing the crop in the summer season and the second-season stands for maize grown as a following crop – mainly after soybeans.

Figure 6:   Brazilian maize production depending on the growing season (in million t)
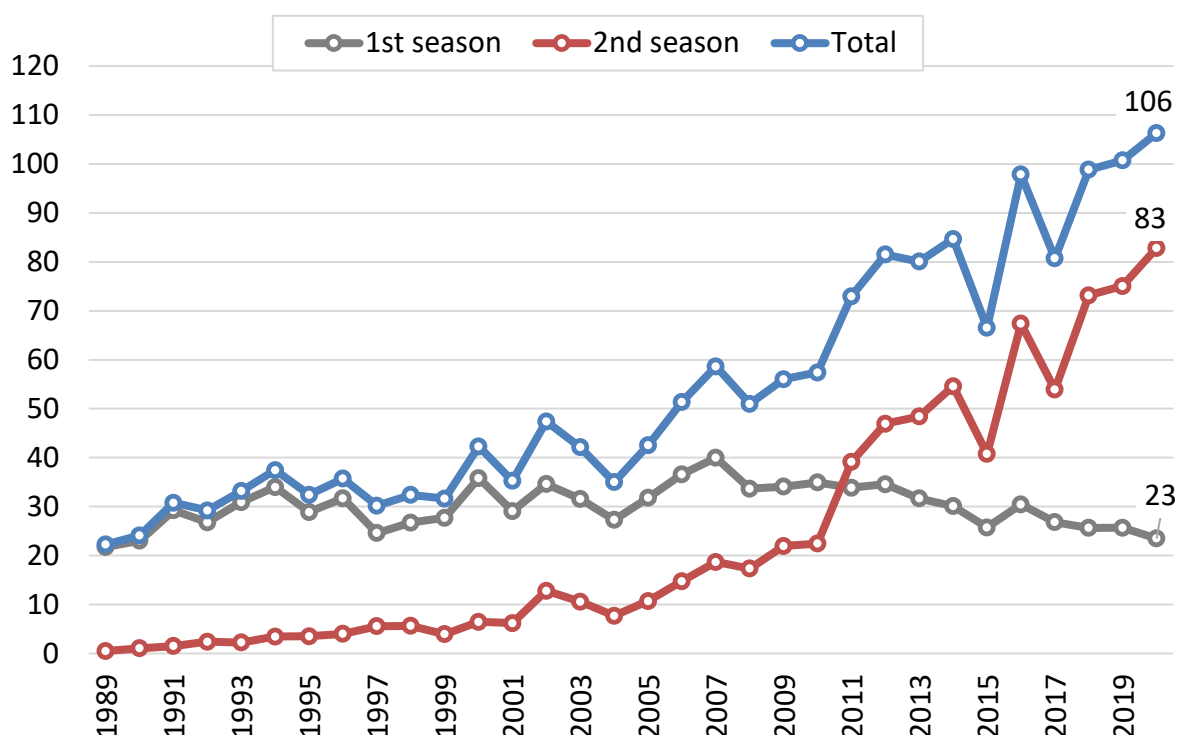
Source: CONAB (2021) – created by the author.

The production of maize as a second crop grew slowly until 2010, reaching approximately 20 million t, after which second-season maize saw a massive expansion, resulting in a more than fourfold increase in one decade, reaching more than 80 million t in 2020 (CONAB, 2021). In 2020, ca. 77% of the total maize production in Brazil was from second-season corn, compared with less than 40% in 2010. It is important to highlight that 66% of the maize grown as a first-season is in the South and Southeast (subtropical) regions, whereas more than 70% of the total maize produced as the second crop is in the Center-west (tropical) region.

In summary, soybean and maize production in Brazil expanded strongly in the past decades, positioning the country as one of the top producers worldwide. Following technical improvements, a highly profitable double-cropping system was implemented, allowing farmers to grow two crops in one season, creating a scenario in which soybean and maize occupies 70% of the total seeded area in Brazil (IBGE, 2019).

3.4.   Beef production

Brazil is a major player in the beef sector, with a cattle herd of around 215 million head in 2019 (IBGE, 2021). While the total volume of beef traded globally increased by 50% between 1990 and 2020, the total export volume from Brazil increased by tenfold, reaching a volume of 2.54 million t. That places Brazil as the largest exporter of beef, accounting for roughly 24% of the total exports. According to the classification of IBGE (2018), most of the 113 million ha

of managed pasture is located close to arable land so understanding the spatial distribution of beef production and its interaction with crop production is important. Figure 7 shows the number of head of cattle comparing the years of 1989 and 2019, at the municipality level.

Figure 7:   Total number of cattle per municipality comparing 1989 and 2019 in Brazil
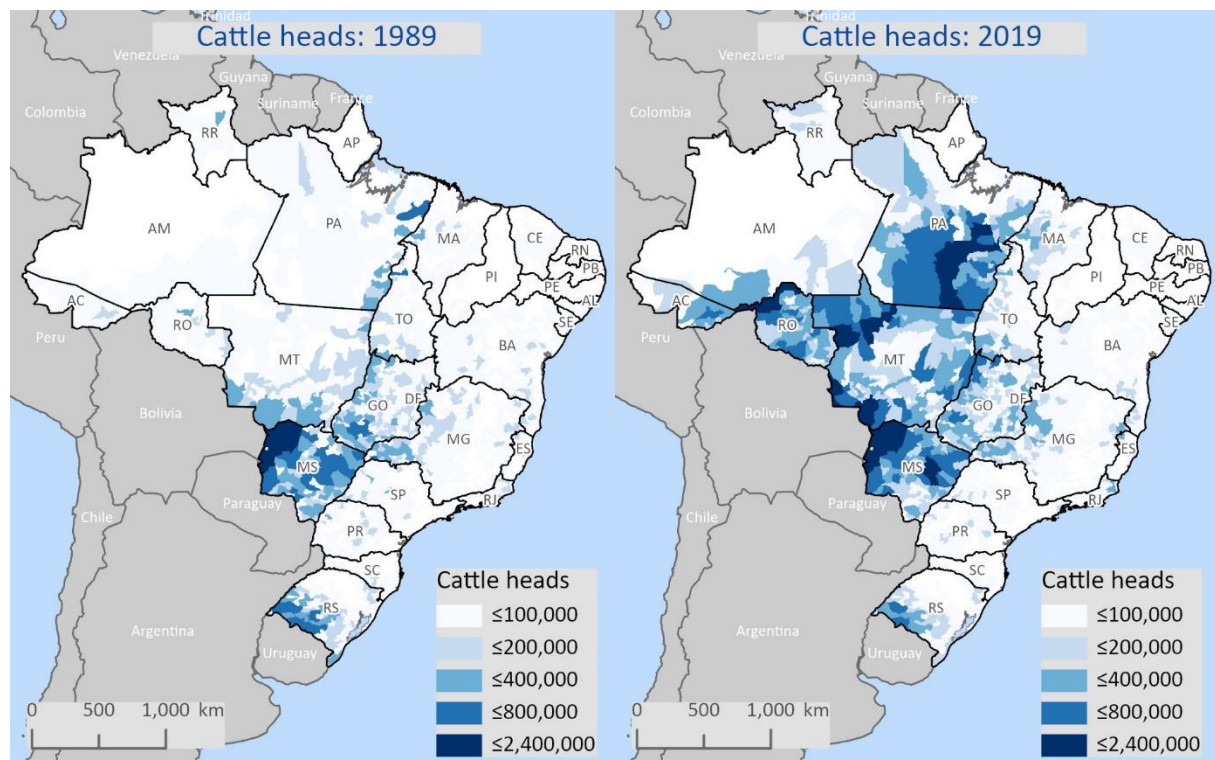
Source: IBGE (2021) – created by the author.

The pattern of expansion of cattle in Brazil is similar to that for soybeans and maize, with a strong increase in production in the Center-west and North regions, mainly in the states of Mato Grosso (MT), Pará (PA), and Rondônia (RO). The significant production in the states Mato Grosso do Sul (MS) and Goiás (GO), which, together with MT, are a main area of grain production, indicates there might be a close interaction between cattle and crop production – i.e., competition for land.

3.5.   Drivers of production expansion: yield gain vs acreage

The rapid production expansion of soybeans, maize, and sugarcane changed considerably the land use in Brazil by converting pasture into arable land as well as displacing other crops in the current farming areas (Barona et al., 2010). Such increases usually are the result of acreage expansion and, to some extent, yield gains. It is important to understand which has been the dominant driving force to align the model to possible future development. For example, if a crop has experienced limited yield gain, additional demand is expected to be meet with increased acreage, leading to land-use change. Observing the past trends may help understand

how the sector has reacted in terms of investment in yields and acreage expansion. Figure 8 shows the development of area and production for soybeans and sugarcane in Brazil from 1990 to 2019, considering 1990 as the reference year.

Figure 8:   Production and area change rate for sugarcane and soybeans from 1990 (baseline) to 2019 in Brazil
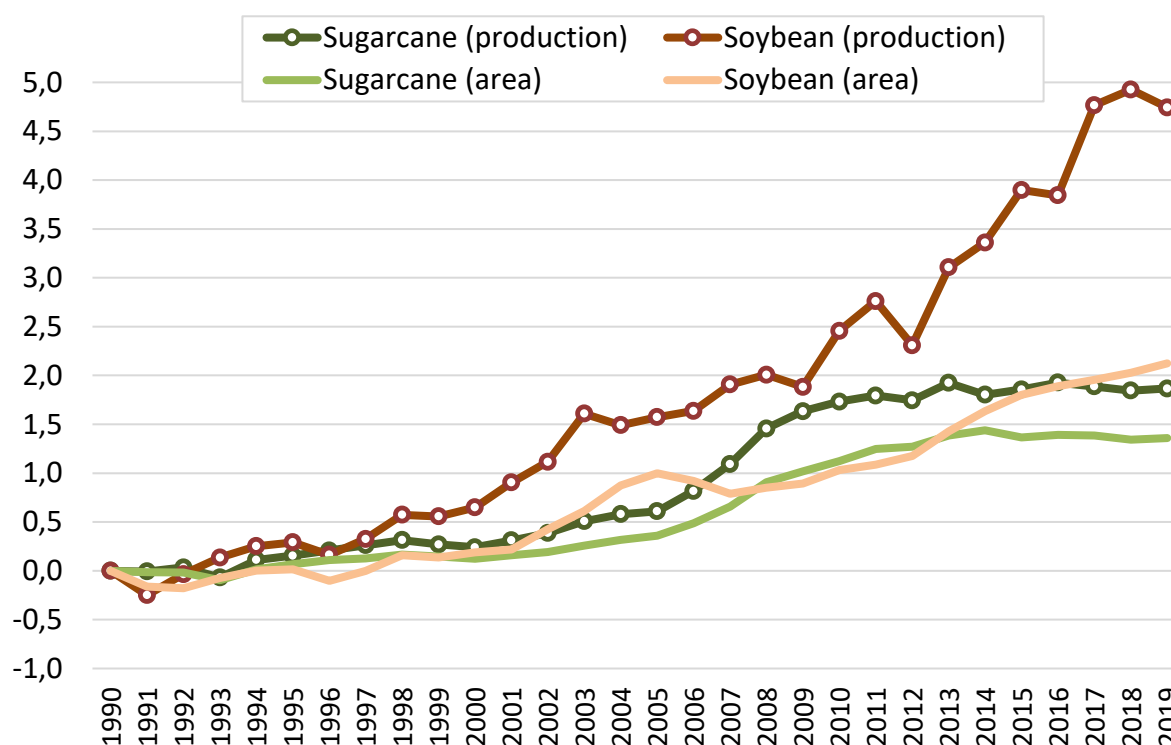
Source: IBGE (2019) – created by the author.

The first interesting observation is that yield gain has played a role for both crops since the overall production increase is stronger than the addition of area. Considering the whole 30-year period, sugarcane yields increased by 21% (0.7% annually) whereas soybean yields increased by 84% (2.1% annually). Focusing on the recent 15 years (2005-2019), the difference between the crops is even more significant. In this period, sugarcane yields improved by 0.2% per year (total 2.5%) while soybeans had impressive gains of 2.4% per year or 43% in the 15 years (IBGE, 2019). That is a first indication that yields have played a more important role in soybean production expansion than in sugarcane and that this difference was more dominant in recent years.

Although yields play an important role in the production expansion of both crops, the main driving force in the past seems to be acreage increase. In the same 30-year period, acreage for sugarcane increased by 136% or 2.9% per year whereas soybean acreage increased by 212%, with an annual rate of 3.9% per year (IBGE, 2019). However, it does not necessarily mean the total area increase is due to adding new arable land. As previously shown, the importance or share of these crops also has increased in recent years.

In summary, the data show that the soybean and sugarcane production increases observed in the past 30 years in Brazil were driven by a combination of strong acreage increases and significant yield gains. Soybeans have had stronger yield gains than sugarcane, which may indicate that, with all other factors constant, this crop has become increasingly more competitive.

Note, however, that even though sugarcane yields at the farm level have stagnated in recent years, industrial efficiency has increased significantly which may counterbalance the farm yield effect and help to improve the overall economic performance of the value chain (Cherubin et al., 2021).

3.6.   Farmers' production system options in Brazil

Considering that soybeans, maize and sugarcane account for 84% of the total seeded area in Brazil (IBGE, 2019), the case study focus on these crops is a logical first attempt to understand land-use dynamics in Brazil. Regionally, other crops such as cotton, beans, and wheat may play an important role but currently they represent only a small share of total seeded area – i.e., 1.6%, 4%, and 2.9% respectively (IBGE, 2019). Since these crops are of minor importance, they are not represented in the *agri benchmark* typical farms network in Brazil, which makes production cost estimation challenging.

There are several production systems involved in growing the major crops depending on climate and regional conditions. The following are the most common systems:

    I.   **Sugarcane monoculture:** unfeasible to have crop rotations since the crop is semi-perennial.

   II.   **Soybean with a cover crop:** soybeans grown in summer and a cover crop in winter to generate biomass for the no-till system.

 III.   **Double cropping soybean-maize:** soybeans grown in summer with maize a second-season crop in winter.

 IV.   **Beef production:** as an alternative for arable farms to diversify or leave the sector.

The focus on no-till systems for the variants with grains is a result of the widespread use of this practice in Brazil. According to FEBRAPDP (2021), considering grain-producing land only, more than 32 million ha are under a no-till system. That is also the system represented by the typical farms from the *agri benchmark* network in Brazil.

The alternative to switch between beef and cropping systems is included in the set of options available to farmers for the case study. As Brazil has currently ca. 113 million ha of pasture located close the main crop production areas, it is important to include cattle as a production system option. Since the focus of this work is on crop production, beef is considered an "opt-out" alternative in case farmers decide to stop growing crops.

Summary

This dissertation focuses on double cropping with soybean as the leading crop in summer with either a cover crop or maize in winter and sugarcane or beef production as the main alternatives to farmers in Brazil. Due to their importance in total agricultural area as well as commodities for international markets, the model should be able to depict the current land use in most regions in Brazil. The model is set with a realistic set of options to farmers, which is key to avoid errors such as allowing the "competition" among crops that are jointly grown in a double-cropping system.

4.   Methodology: developing the modeling approach

To answer the research question, I propose a new modeling approach to test whether the combination of a biophysical model with farm-level economic data provides a reliable estimation strategy on farmers' supply responses. Since the focus of this work is the methodologic development of the modeling concept, this chapter explains in detail the building blocks of the model, using the Brazilian case to illustrate the practical background and applications. The first section focuses on the modeling concept, explaining the information flow between the models. Then, the individual modules are explained in greater detail.

The proposed modeling approach is called **Profitability Assessment Model (PAM)** based on the idea that farmers' supply responses are mainly based on the profitability of the different cropping alternatives in a certain region. The model is written in the program language Python (Python, 2020), using mainly the Pandas library (Reback et al., 2020). Additionally, Geographic Information System (GIS) software (ArcGIS Pro) is used to spatially allocate some of the external information into the model as well as to map and analyze the output.

4.1.   Model structure and information flow

The main idea behind the PAM modeling approach is the combination of a biophysical plant growth model with farm-level economic data and using a minimum data approach for regions with poor data availability. The approach is based on combining different models and layers of information. Figure 9 shows a simplified scheme of the four main modules within the PAM approach – i.e., plant, transport, economic and land allocation.

Figure 9:   Schema of information flow in the different modules of the PAM model
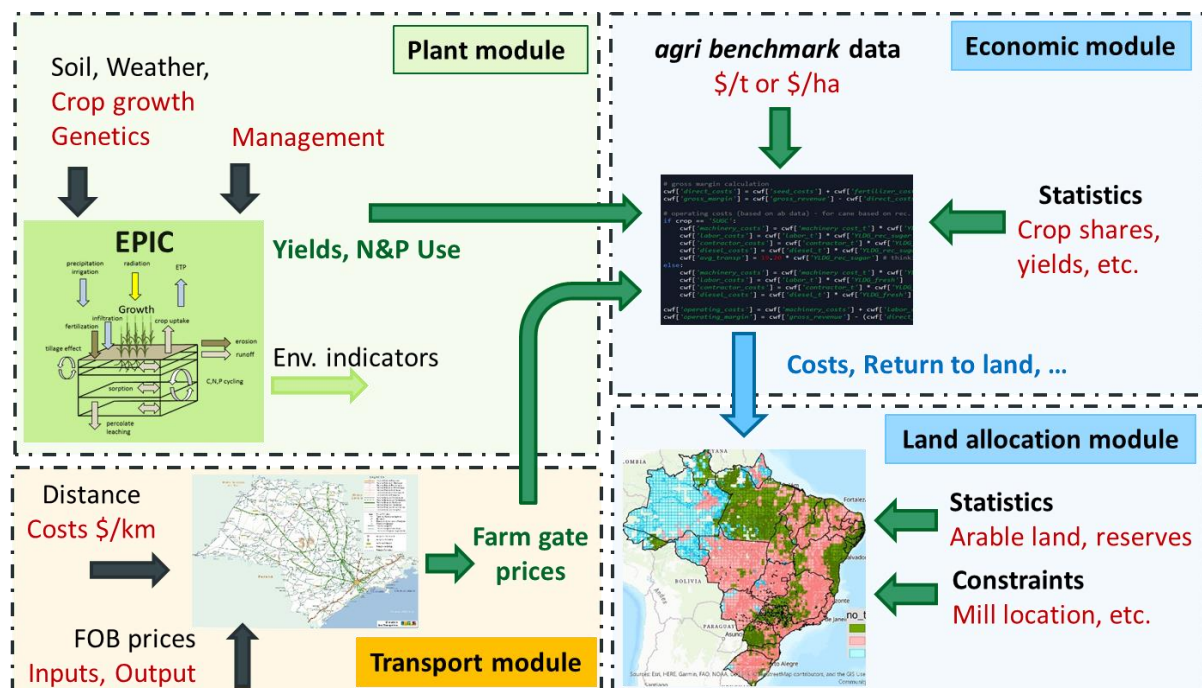
Source: created by the author.

The plant module is a key component, providing mainly crop yields and respective fertilizer input use. That information flows directly into the economic module, which also receives information from the transport module. The latter converts the Free on Board (FOB) prices of input (e.g., fertilizer) and output (e.g., soybeans) to simulation unit levels. The economic module is the main new development within the PAM approach. It combines the information from the two modules with several input layers from statistics and *agri benchmark* data to calculate the profitability of each farming alternative for each simulation unit. Finally, the land allocation module allocates farmland to the different alternatives, maximizing farmers' economic returns. This process, however, is complex, using several layers of input and output so that the following sections focus on further explaining the functionally and setup of modules.

Due to the current importance of each crop (see chapter 3), the current version of the PAM is calibrated for the following alternatives in Brazil:

   − sugarcane monoculture,

   − soybean with a cover crop,

   − double cropping soybean-maize,

   − beef production.

While the first three alternatives are endogenously estimated by the PAM approach, the returns to land from beef production are obtained from the *agri benchmark* Beef network.

4.2.   Model resolution – Simulation Units

Before diving into the details of each module within the proposed approach, it is important to explain the model resolution. The spatial resolution is an important source of uncertainty in large-scale modeling (Mearns, Easterling, Hays, & Marx, 2001). The PAM approach follows the spatial resolution from the GLOBIOM model developed by the International Institute for Applied Systems Analysis (IIASA). Hence, it is possible to consistently use other IIASA models as well as ensure that the outputs of PAM can be implemented easily in other tools. Additionally, the calibration from the biophysical model follows the spatial resolution from IIASA, which is the highest resolution from all components of the PAM model.

The simulation units for the GLOBIOM model are derived as follows:

- The geospatial data required for the biophysical modeling was compiled from several sources by Skalský et al. (2008) and harmonized into Homogeneous Response Units (HRU). These clusters are derived mainly from landscape parameters that do not change over time – i.e., five altitude classes, seven slope classes, and five soil classes.

- In the next step, the HRU are intersected with a spatial grid of 0.5°× 0.5° and country boundaries to draw the resulting SimU that contain information on global climate data, land use, irrigation, etc. (Havlík et al., 2011).

The resulting SimUs are polygons with size that varies between 5 and 30 arcmin (i.e., 10x10 to 50x50 km in area) and they define the spatial resolutions for the PAM modeling approach. The SimU also are defined as the enterprise or farm so the model calculates the profitability of each cropping alternative and the land allocation decisions at this level. Therefore, it is important to understand the scale and characteristics of the SimU.

In Brazil, there are 11,003 SimUs with sizes varying between 7,500 to 300,000 ha. The size of these units depends on the heterogeneity of their conditions as well as administrative boundaries. Those with the largest areas are found in the Amazonia region, where agricultural production currently is negligible and predominant land use is native vegetation. Conversely, the SimUs in other regions are significantly smaller due to the variety of natural conditions. Figure 10 shows a classification map illustrating the variability of SimUs' size across regions (left) as well as the distribution of size (right). The median size of the SimUs is 34,200 ha (18.5 x 18.5 km), as depicted in the histogram, and around 75% of the units are under 100,000 ha (32 x 32 km).

Figure 10:   Map of simulation units' size (left) and frequency distribution of unit size (right) – in ha
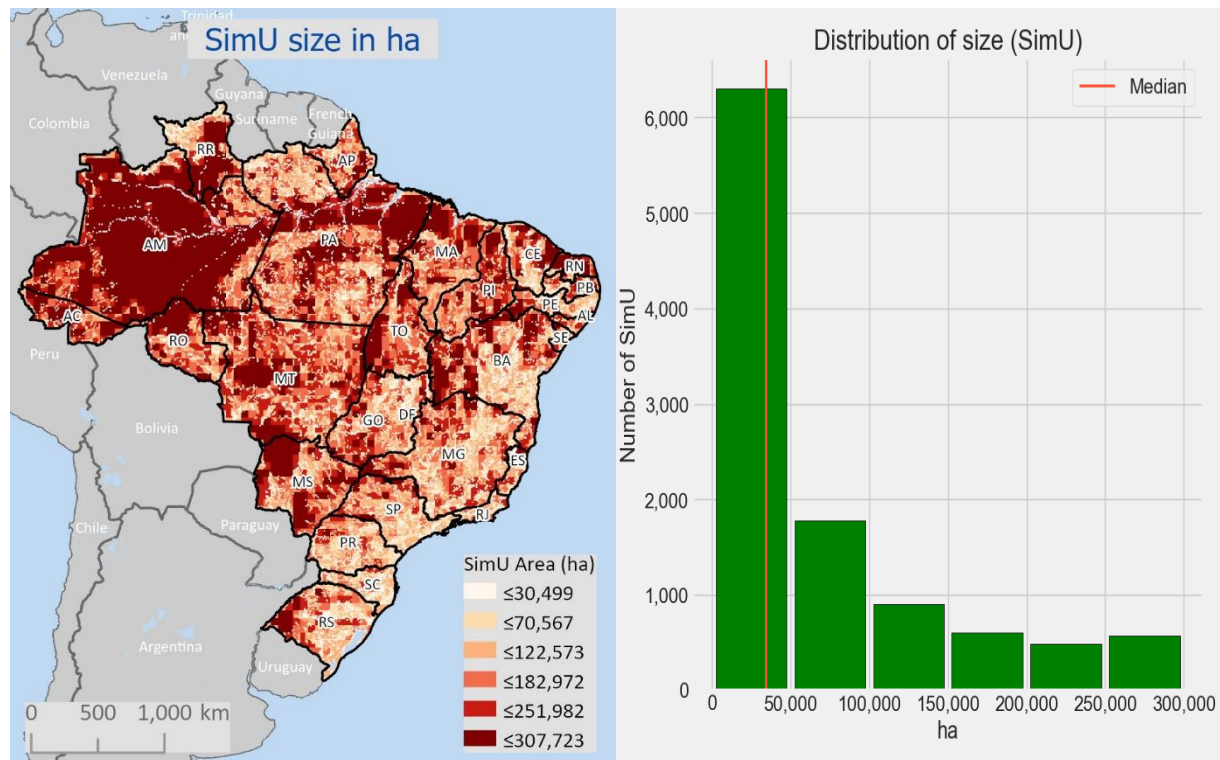
Source: created by the author.

Another important advantage of basing the PAM approach on the SimUs is the ability to integrate official statistics to these units, serving as an input for the model as well as validation of results. This is possible because the National Institute for Space Research (INPE) has developed a more detailed version of the GLOBIOM model for Brazil and, in the process, developed an algorithm that consistently allocated official statistics (IBGE) that are reported at the municipality level (e.g., harvested area), as well as land cover information from PRODES (INPE) to SimU, making this information readily available for the PAM analyses (Câmara et al., 2015). The link between geographic information (SimU) to administrative units (municipalities) is of extreme importance for the calibration and validation process of the model.

In summary, the SimUs define the spatial resolution within the PAM approach with a median area of 34,200 ha (18.5 by 18.5 km). This unit also is considered one enterprise when allocating cropping alternatives. The decision on this specific resolution is mainly due to possible synergies between models as well as the ability to use available structures and information.

4.3.   Plant module: estimating yields and fertilizer use

A main component of the PAM approach is the yield estimation for individual crops because together with costs and farm-gate prices, they determine the competitiveness of each farming

alternative. Yield levels also are used as a proxy for the different types of production systems as well as in the allocation of certain production costs. Therefore, estimating yields using a biophysical model is a major component of the PAM approach.

The yield estimation is necessary due to the lack of statistical information at such high spatial resolution as well as the ability to estimate yields and consequently production costs in areas that currently are not producing these specific (or any) crops. That fits exactly with the goal of the PAM approach in providing robust supply response estimations in countries or regions with limited regional statistics.

4.3.1.   Module description

The Environmental Policy Integrated Climate model (EPIC) is the main component of the plant module (Williams & Singh, 1995). This model is widely used and tested to estimate yields of several plants based on atmospheric and soil interactions affecting biomass production and, consequently, yields. Crop growth is the main process modeled and the basic concept is using a radiation-use efficiency, where a fraction of radiation is intercepted by the plant and converted to biomass (i.e., roots and above-ground). These potential daily gains are penalized due to factors such as carbon dioxide (CO_2) atmospheric concentration, water, and nutrient availability, temperature, etc. reducing the daily biomass gain from the potential to actual (Izaurralde, Williams, McGill, Rosenberg, & Jakas, 2006).

Main inputs into the EPIC model include the crop calibration information based on field trials as well as site-specific variables such as soil, weather, and management information containing all activities such as seeding, fertilizer applications, and harvesting. The first description of the model can be found in the work of Williams, Jones, and Dyke (1984) and the expanded version on Williams, Jones, Kiniry, and Spanel (1989). Figure 11 shows a schematic illustration of the main components and processes modeled by EPIC.

Figure 11:   Main components and flows of the EPIC model[12]

Main components in EPIC:

- crop growth
- weather
- hydrology
- soil temperature
- soil erosion
- C, N, P cycling
- pesticide fate
- tillage
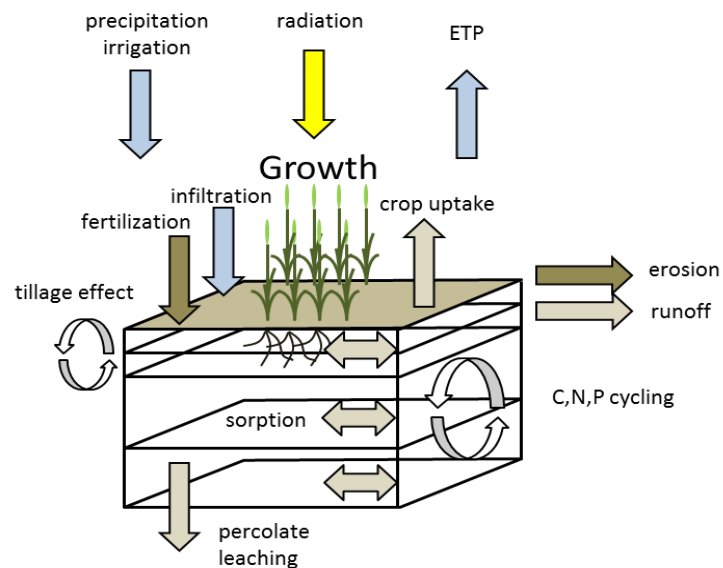
Source: Azevedo (2016, ppt presentation)

EPIC has been improved by IIASA with crop-specific calibration for the main arable crops and regions worldwide. This version has been used and tested in several projects and papers containing detailed calibration parameters for all crops considered in the current version of the PAM approach (Balkovič et al., 2013). An important advantage of IIASA's version of EPIC is the specific crop calibration for the second-season maize grown in Brazil. Since intensive breeding was required to develop varieties that can cope with drier conditions and mature in relatively short cycles, a special calibration of crop parameters within EPIC is required to allow realistic estimation of soybean plus maize production system (Câmara et al., 2015).

The fertilization information of EPIC is of key importance for the PAM approach, providing the fertilizer rates to the crops. EPIC allows different options to define the fertilization scheme, varying mainly between fixed and variable application mode. The *fixed mode* can be divided into two categories: "user schedule" which applies the exact fertilizer mix and rates as defined by the production system, or "automatic" in which the fertilizer mix and rate come from the production system, but the application is triggered if stress levels are reached. The fertilization rate in the *variable application* mode is mainly based on the crop uptake and availability in the soil and the timing is based either on the "user schedule" or is "automatic" when stress levels are reached. The user can also define the maximum amount of fertilizer applied as well as the minimum interval between each application. Thus, the application rates of nitrogen (N) and phosphorus (P) depend on the definitions of production systems and the fertilization mode used. The latter determines how the model reacts to crop needs, soil supply, and the nutrient

---

[12]   C = carbon; N = nitrogen; P = phosphorus; ETP = evapotranspiration.

stress relative to stress factors – water and temperature (Williams, Izaurralde, Williams, & Steglich, 2015).

EPIC has been employed in a range of research, including soil science, climate change and economics. It has been tested extensively against observed data and expanded to model different crops in varying regions worldwide. A detailed description of publications using EPIC and the main findings can be found in Gassman et al. (2005).

4.3.2.   Data input into EPIC

Yield estimation plays a major role in the overall PAM approach so it is important to understand the underlying database and calibration parameters that are used for yield estimation with EPIC. A detailed explanation of the data input and requirements for EPIC can be found in the work of Skalský et al. (2008), from which the short explanations that follow are derived:

- **Weather data:** Global daily weather estimations are available only from 1981 to 2010. Therefore, the EPIC estimations are carried out for these years and the final output is an arithmetic average of the period.

- **Soil and topography:** Detailed information is included on average field size, slope, elevation, and geographic coordinates. For each soil layer, EPIC requires detailed information such as depth, density, sand content, pH, etc.

- **Crop calendar:** Using IIASA's standard operation schedule for the analyzed crops, they include information on the time of operations (e.g., planting, harvesting, applications) as well as the type of fertilizer, amount of irrigation, tillage operation, etc.

- **Scenarios:** Simulations from varying scenarios are available following IIASA's scenarios mainly on maximum nitrogen application and irrigation schemes. For the current version of PAM, these scenarios are not fully used because the fertilization rates follow the automatic used mode, and irrigation economics are not completely integrated into the model.

4.3.3.   Integrating EPIC output into the PAM model

The main contributions of EPIC to the PAM approach are the estimation of yields and application rates of fertilizer. The EPIC output is the starting point for the PAM modeling approach, providing detailed information on biomass production, nutrient balances, and main sources of plant stress.

Calculating commercial yields

The yield output from EPIC is transformed to commercial yields in the PAM by accounting for technical harvest losses as well as including the moisture content typically found in grains and sugarcane. For maize and soybeans, the dry matter of the commercial yield is assumed to be

86% and for sugarcane stalks, 23%. For sugarcane, EPIC estimates the above-ground biomass as forage yield but does not provide information on the actual fresh stalk yield. Therefore, the PAM model assumes that 80% of the above-ground biomass is sugarcane stalks (Marin, Thorburn, Nassif, & Costa, 2015). The data for harvest efficiency are taken from technical reports – maize 8%, soybeans 2% (Zandonadi, Ruffato, & Figueiredo, 2015), and sugarcane 5% (Silva, Corrêa, Cortez, & Furlani, 2008). These values are expected to vary depending on field conditions and machinery so that PAM allows a different assumption (input) if local information is available.

For sugarcane, information on stalk yield is important because several cost such as harvesting and transportation are linked to it. However, the economic productivity is based on a combination of stalk yield and sugar content. Sugar content is the main parameter defined in cane payment system, directly affecting farming economic performance. Unfortunately, EPIC does not report sugar content as do other biophysical models such as the Decision Support System for Agrotechnology Transfer (DSSAT) and the Agricultural Production Systems Simulator (APSIM) that have special modules for sugarcane (Marin et al., 2011). Therefore, to convert fresh stalk yields from EPIC to sugar yield, the plant module uses the Total Recoverable Sugar information (TRS or *ATR* in Portuguese) from CONAB (2019), on state-level averages from seasons 2017/18 and 2018/19.

The following formulas are used to incorporate EPIC yields into the PAM modeling approach:

$$YLD\_sugc_i = \frac{YLDF_i}{DryM_{sugc}} \times ShareStk \times HarvEff_{sugc} \qquad (1)$$

$$YLD\_trs_i = YLD\_sugc_i \times TRS_{uf} \qquad (2)$$

$$YLD\_grain_i = \frac{YLDG_i}{DryM_{grain}} \times HarvEff_{grain} \qquad (3)$$

where YLD is the commercial yield in t/ha; $YLDF$ and $YLDG$ are the EPIC forage and grain yields in t/ha for each SimU i, respectively. The $ShareStk$ is the share of stalks in the forage yield (in %); $HarvEff$ is the harvest efficiency (in %); TRS is total recoverable sugar for each state uf (in %), and $DryM$ is the dry matter in the final product (in %).

Nutrient application rates

Besides yields, EPIC also reports application rates of N and P based on the fertilization scheme, the management information and the actual nutrient use by the plants to produce the estimated yields (see model description). The "automatic" fertilization scheme is used following IIASA's standard estimation, so the fertilization mix is defined by the business-as-usual (BAU) crop management data, but the application is triggered if stress levels are reached. For N, the upper application rate is set (scenarios varying from 0 to 400 kg/ha) and P is applied to relieve plant stress in balance with N and water stresses. Depending on the research scope, it may be desirable to use either the applied fertilizer or the plant uptake; hence PAM allows the user to decide whether the former or the latter should be used for cost estimation.

The current version of the PAM model uses the standard applied quantity N and P for the cost calculation (further explanation under the economic module). Potash (K) usage is not simulated by EPIC. Hence, the model uses results from local field trials on kg uptake K per t of crop produced as the reference to estimate K costs (Mascarenhas, Tanaka, Wutke, Braga, & Miranda, 2004; Oliveira et al., 2010).

Finally, it is important to note that even though the current version of the PAM model uses EPIC as the biophysical model, any other model providing high-resolution information on yields and nutrient use can be employed with a similar approach. This feature is important since, for different regions and crops, other calibrated models may be more suitable.

4.4.   Transport module: generating farm-gate prices

The transport module is an important structure within the PAM approach, allowing the estimation of input and output prices at the farm-gate level. Currently, there are no official statistics on input (e.g., fertilizer) and output (e.g., soybean) prices at the farm level in all states in Brazil. Having an internal module to estimate farm-gate crop and input prices is important specially for large countries such as Brazil, where local prices vary greatly depending on the distance to market.

For Brazil, the transport module is designed to convert FOB crop prices to farm-gate level based on the concept that farmers in exporting countries receive a farm-gate price that is derived from the international prices minus transaction costs to deliver the crop to the port (Freebairn, 1987). The Brazilian agricultural sector has a strong focus on exporting, with 72% of total production of sugar, 64% of soybeans, and 34% of maize delivered to the global markets[13]. Hence the prevailing local price in Brazil is expected to be a function of FOB prices and transport cost of delivering the crop to the port. A similar concept is used to estimate farm-gate fertilizer prices given that more than 80% of the fertilizer used in Brazil is imported

---

[13]   Share of export volume on the total production calculated as an average of 2017-2020 based on the production volume from CONAB (2021) and export volume from ITC (2021).

(ANDA, 2020). Hence, fertilizer prices at farm gate are a function of the international prices plus the transport cost to reach the farm level. Even for domestically produced fertilizer, this pricing approach should apply, considering that price deviations should motivate farmers to switch between origins, creating a balance between domestic and imported inputs.

Technically, the module uses the GLOBIOM-Brazil approach to estimate the travel distances between the SimUs and the nearest port. Coefficients for freight costs are derived from observed information to calculate the transport costs per unit of product (input or output). The combination of these two components provides the final parameter for transport costs. Additionally, this module estimates the inbound transportation costs for sugarcane – i.e., from the fields to mills, which is a major cost component within the sugarcane supply chain. The following subchapters explain in greater detail the two main components of the transport module and the approach developed to estimate inbound transport costs for sugarcane.

4.4.1.   Module description

The current transport module is based on two main components: (a) the GLOBIOM-Brazil version estimating the travel distance in km between the SimU and the nearest port, and (b) regression analysis to estimate the transportation costs (USD/km) based on observed freight costs from research team ESALQ-LOG[14] in Brazil. The combination of these two components results in the estimation of transport costs in USD per t of product to and from the nearest port currently in operation.

Getting the distance between SimU and the ports

The distance between the SimU and the ports is estimated following the Generalized Proximity Matrix (GPM) approach proposed by Aguiar, Câmara, Monteiro, and Souza (2003). The 2012 National Plan for Logistics and Transportation (NPLT) serves as the base for the available roads to connect the starting point (i.e., centroids of the SimU) and the destination (i.e., port or domestic market). Figure 12 shows the road infrastructure in Brazil (left) as well as the nearest port from each SimU (right).

---

[14]   Group of Research and Extension in Agroindustrial Logistics – https://esalqlog.esalq.usp.br/en/

Figure 12:   Road infrastructure in Brazil (left) and the nearest port from each SimU (right)
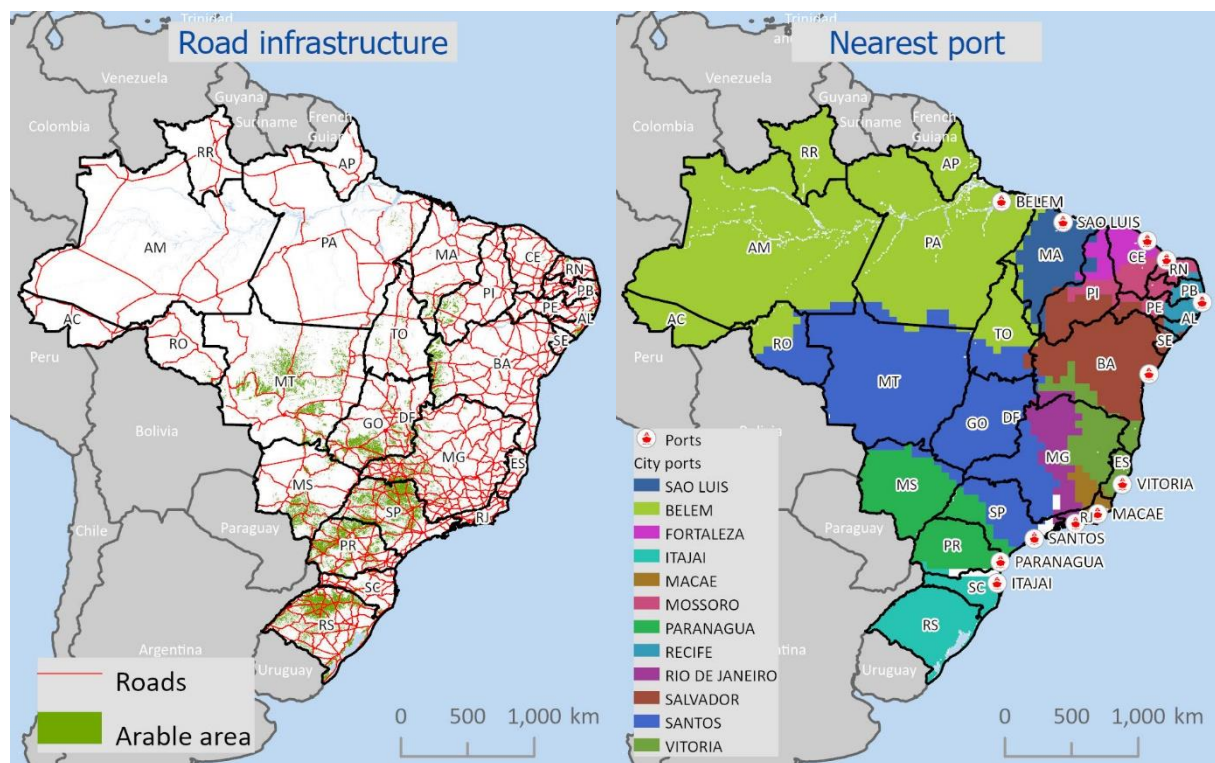
Source: based on NPLT and INPE (2019) – created by the author.

Most SimUs located in the current arable area (see Figure 12) have as their nearest ports Santos and Paranaguá. These ports are the two main outlets for commodities in Brazil, together accounting for 93% of the total volume of sugar and 43% of total volume of soybeans exported in 2020 (ABIOVE, 2021; UNICA, 2021). These ports also handle the majority of the maize exported from Brazil (Souza, Reis, Abraham, & Machado, 2017).

It is interesting to observe the lack of road infrastructure mainly in the northern states in Brazil. That is the result of the very low population density in these regions as well as the largest share of the area being under forest. For commodities, it means that any attempt to grow a crop in these regions will be faced with high (imported) input costs and low output prices due to relatively high transport costs.

Estimating freight costs

Transportation costs under the GLOBIOM-Brazil approach are estimated based on travel time and current conditions of roads resulting in a constant transport cost from each SimU to the destination (Câmara et al., 2015). The differentiation between crops is mainly due to the share of the output sent to the international vs domestic market. This costing approach has important limitations because it does not consider different technical requirements in transporting different commodities – e.g., sugar and soybeans. Sugar requires a different type of trucks and has an associated higher transport costs per km. More importantly, this approach does

not account for the fixed component of the transport costs, since the calculation is based on distance times the average transport costs per km, varying only according to the quality of the roads.

Therefore, although PAM primarily uses the distance estimation approach from GLOBIOM-Brazil, as mentioned, it improves its costing approach by accounting for the crop-specific transportation costs – differentiating between the fixed and variable components, based on observed freight information. Transportation costs are estimated through a regression analysis of observed freight costs (ESALQ-LOG) for each specific product. The following ordinary least squares (OLS) linear regression is used to assess the impact of distance on transport costs:

$$T_i = \beta_0 + \beta_1 D_i + \varepsilon_i \qquad (4)$$

where T_i is the transport cost in USD per t of the product; β_0 is the intercept that can be interpreted as fixed cost in USD/t; β_1 is distance coefficient that represents the variable costs in USD/t per km, and D_i is the observed distance in km from the SimU to the nearest port.

The freight data used in the estimation comes from the Freight Information System (SIFRECA in Portuguese) developed at ESALQ-LOG. Primary transportation costs for each specific product/commodity and route are collected weekly and consolidated to a monthly indicator. The regression analysis uses the monthly indicator for all routes from each specific product to all destinations in Brazil, summing up to 14,626 pieces of freight information for 2017[15]. Table 1 shows the results from the OLS estimation.

---

[15] I have access to the database from ESALQ-LOG for 2017. Based on the main cost component of the road transport (i.e., diesel), the average real diesel prices from 2015-2019 is R$ 3.64/l, whereas the real price in 2017 was R$ 3.67/l which indicates that 2017 can be considered an average year based on diesel costs. A comparison of the real diesel prices can be found in appendix 6.

Table 1:   OLS results: freight costs in Brazil

Variables	Coefficients β	
	Sugar	*Grains**
Intercept (β_0)	12.326	9.997
Distance (β_1)	0.0347	0.0340
R^2	0.84	0.90
N	1,694	12,932

*soybeans and maize

Note: SIFRECA monthly indicator for 2017

Source: own calculations based on ESALQ-LOG (2019)

The intercept values show the importance of analyzing freight costs for different products separately. As expected, sugar freight has a higher fixed cost component, indicating that different equipment or services are required to move this product compared with grains. That is important since SimUs, with a similar distance to ports, will face higher freight costs for transporting sugar than soybeans to the ports. The distance coefficient is similar between the crops due to its main components – diesel, labor, maintenance (Fliehr, Zimmer, & Smith, 2019).

With the freight cost coefficients and the travel distance, the transport costs within the PAM model are derived using the following equation:

$$Transp\_t_{grains_i} = D_i \times \beta_{1_{grains}} + \beta_{0_{grains}} \tag{5}$$

where: D_i is the distance in km to the nearest port from the SimU i and β's are the coefficients for freight costs for grains. A similar equation is used for sugar, changing only the respective coefficients as shown in Table 1.

The structure of the transport module allows an easy update of the freight costs coefficients if changes in inputs such as diesel prices are observed. Moreover, it allows an endogenous estimation of transport costs from all SimUs to ports, meaning that any FOB prices of inputs and outputs can be automatically converted to farm-gate prices, accounting for one of the major framework differences between regions in Brazil.

4.4.2.   Inbound transport costs sugarcane

The transport of sugarcane from farms to the mill for processing is an important cost component of sugarcane production. Differently from soybeans and maize, sugarcane is not commercialized directly from the farm. Hence, transporting cane to a processing facility is a key cost component within the supply chain and needs to be included when modeling regional allocation of sugarcane production.

In the PAM model, sugarcane transport from the farm to the mill is built into the transport module. The model currently allows two alternatives based on costs per t of sugarcane:

(a) Actual distance: using the geodesic distance between the centroid of the SimU and the location of the nearest operating mill in Brazil. This option allows modeling short-term developments so production from any extra hectare must be delivered to an existing mill and therefore incurs relatively high transportation costs. The distance between the SimUs and the current mills is estimated using the GIS software and assuming that each SimU can deliver to its nearest mill. The transport cost per km for this option is based on the work from Françoso, Bigaton, Silva, and Marques (2017) as follows:

$$InbGeo\_t_i = \ GeoD_i \times 0.0582 + 1.02 \qquad\qquad (6)$$

where: GD_i is the geodesic distance between the SimU centroid and an existing mill (2017). The coefficients from Françoso et al. (2017) were converted to USD using the exchange rate of 0.2882 which is the average for 2016. That is important because the regression is based on data from the season 2016/17 in Brazil.

(b) Current average: This option simulates a medium-term perspective in which expansion of sugarcane to new region, where no mills currently are operating, may motivate building new mills and thereby allow the reduction of the transport cost to a level comparable to today. Therefore, the model assumes the average transport distance to be at a maximum of 30 km and the resulting transport cost is calculated using the following equation:

$$InbAv\_t_i = \ \text{if } GeoD_i \geq 30; \text{then } 2.766 \text{ else } InbGeo\_t_i \qquad (7)$$

where: 2.766 is the transport costs at an average of 30 km based on equation 6.

The resulting transport cost per t of sugarcane is converted to cost per ha based on the following equation:

$$InbT\_ha_i = \ YLD\_sugc_i \times InbAv\_t_i \qquad\qquad (8)$$

where YLD_sugc is the fresh stalk yield based on equation 1 and $InbTav\_t$ is the transport cost per t of sugarcane for each SimU i.

Transporting sugarcane is costly (ca. 20 USD/t of recoverable sugar) and hence more than 95% of the production currently takes place within 50 km of a sugar mill. Figure 13 shows the geodesic distance (GIS) from the SimU centroids to the current mills, not considering areas that are currently under forest (left) as well as the share of total production at certain distances to the mill operating in 2017 (right).

Figure 13:   Distance from SimU to the nearest mill operating in 2017 (left) in km and the share of total production at certain distances to an operating mill (right) in %
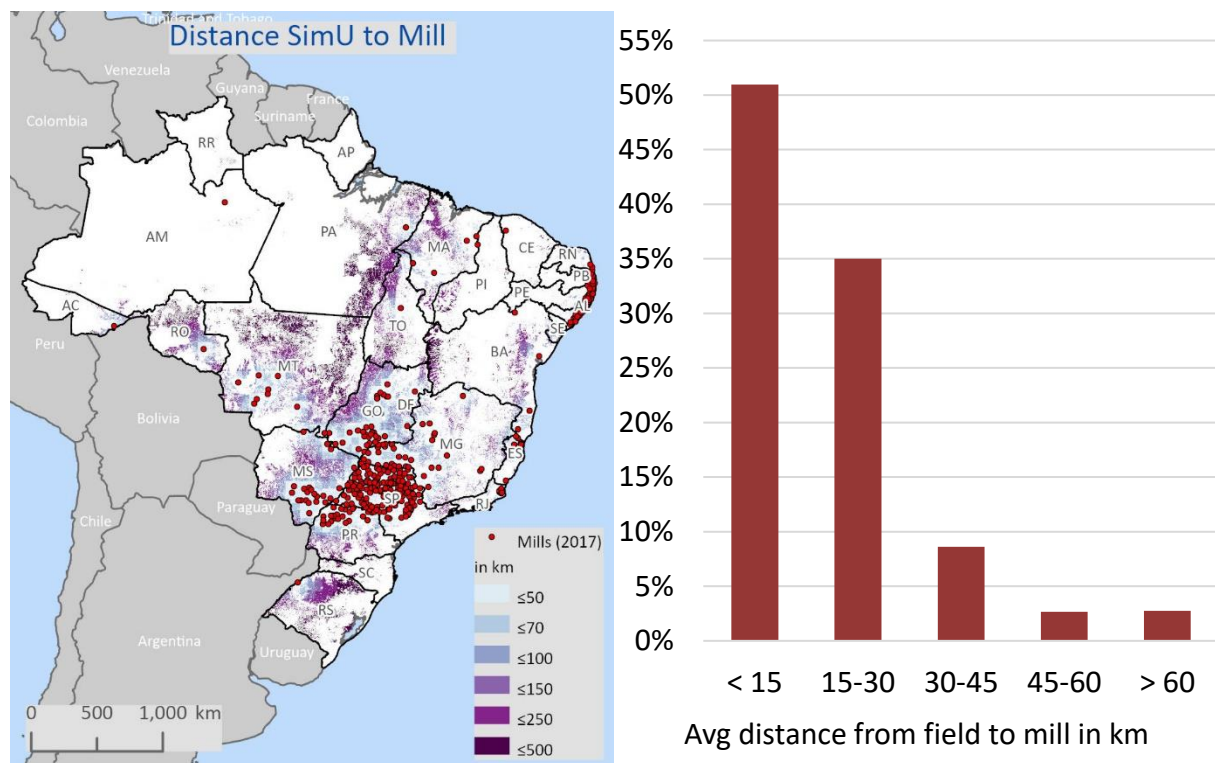
Source: Data for current mills from UDOP (2017) and production from IBGE (2015) – created by the author.

Confirming the expectation, most of the sugarcane production takes place within a radius of 50 km of an operating mill. Areas far away from the main production hub in São Paulo (SP) have distances to an operating mill as high as 500 km, which in turn would mean a transport cost of more than 30 USD/t of fresh cane (equation 7) or 230 USD/t of TRS[16].

In a nutshell, considering the inbound transportation costs for sugarcane is crucial to guarantee a holistic picture when it comes to profitability at the farm level, since the transport cost of sugarcane from farm to mill is commonly paid by the farmer. The choice of different alternatives to include inbound costs into the model allows a variety of analyses considering short- and long-term developments.

4.5.   Economic module: estimating profitability

The economic module is the major innovation within the PAM modeling approach. It has been developed entirely to estimate production costs, revenues, and return to land for each crop alternative at the SimU level. The availability of production cost information is very limited to

---

[16]   Assuming 13% of recoverable sugar (TRS) per t of fresh cane.

the level of detail required to analyze microregional land-use allocation strategies in most developing countries (Rao, 1989). Therefore, the economic module aims to fill this information gap by combining the output of the auxiliary modules (i.e., plant and transport) and other databases, allowing the estimation of the economic performance of the different cropping alternatives at the SimU level.

This module is based on several steps to compute establishment and operating costs as well as revenues, following assumptions regarding crop intensity, yield levels, and observed information from case study farms (i.e., typical farms). The output is the return to land estimations for each cropping alternative for each SimU – under the assumptions and price scenarios used.

The PAM economic module follows the reasoning behind the minimum data approach so the entire development focuses on existing data, mainly from typical farms within the *agri benchmark* network, and national statistics. For Brazil, detailed data of additional typical farms from PECEGE and CEPEA is used for the development of costing extrapolation strategies (production costs). Even though these additional datasets are important for the technical development of the model, they are not used for the simulations. The idea is to assess the model's performance using only the existing data structure of typical farms from *agri benchmark*. The following subsections explain in greater detail the building blocks of the economic module within the PAM approach.

4.5.1.   Module description

The main goal of the economic module is to estimate the returns to land from each cropping alternative at the SimU level. Return to land is the measure for profitability and it is composed of revenues and total costs. The key innovation within the PAM framework is the detailed production cost estimation using a biophysical model and farm-level data. In this module, the aim is to assess the different behaviors of the cost components – e.g., linear to yields or fixed per ha – to realistically estimate production costs based on the current production system in Brazil.

The overriding question is how to estimate the different cost components such as fertilizer, machinery costs, etc., based on few observed production costs data (i.e., typical farms) and using the output from the biophysical model to rationally extrapolate these known costs to regions where such costing information is not available.

In the following subsection, I assess the relationship between the cost items and yield or area, considering important differences between cropping systems regionally as well as the overall productivity of the system based on yields. The observed farm-level production cost comes from the *agri benchmark* network, explained in detail in the next subsection. Finally, this module combines the information from the transportation module to estimate farm-gate prices of the crops produced, allowing the calculation of the main economic indicator – i.e., returns to land for each available farming alternative at the SimU level.

4.5.2.   Production costs data: typical farms

Considering the data availability and the benefits of using representative farms, the PAM model currently uses the existing typical farms from the *agri benchmark* network in Brazil as the basis for most production cost information. The network has three grain-specialized and three sugarcane farms in Brazil. Even though the network has only six farms (five used as input)[17], the local research institutions collect information from many more (validation) typical farms that are not part of the standard database. Figure 14 shows the location for both input and validation farms (see chapter 6) and the areas producing soybeans and sugarcane.

Figure 14:   **Location of typical farms (model input and validation) and production for soybeans (left) and sugarcane (right) in 2016**
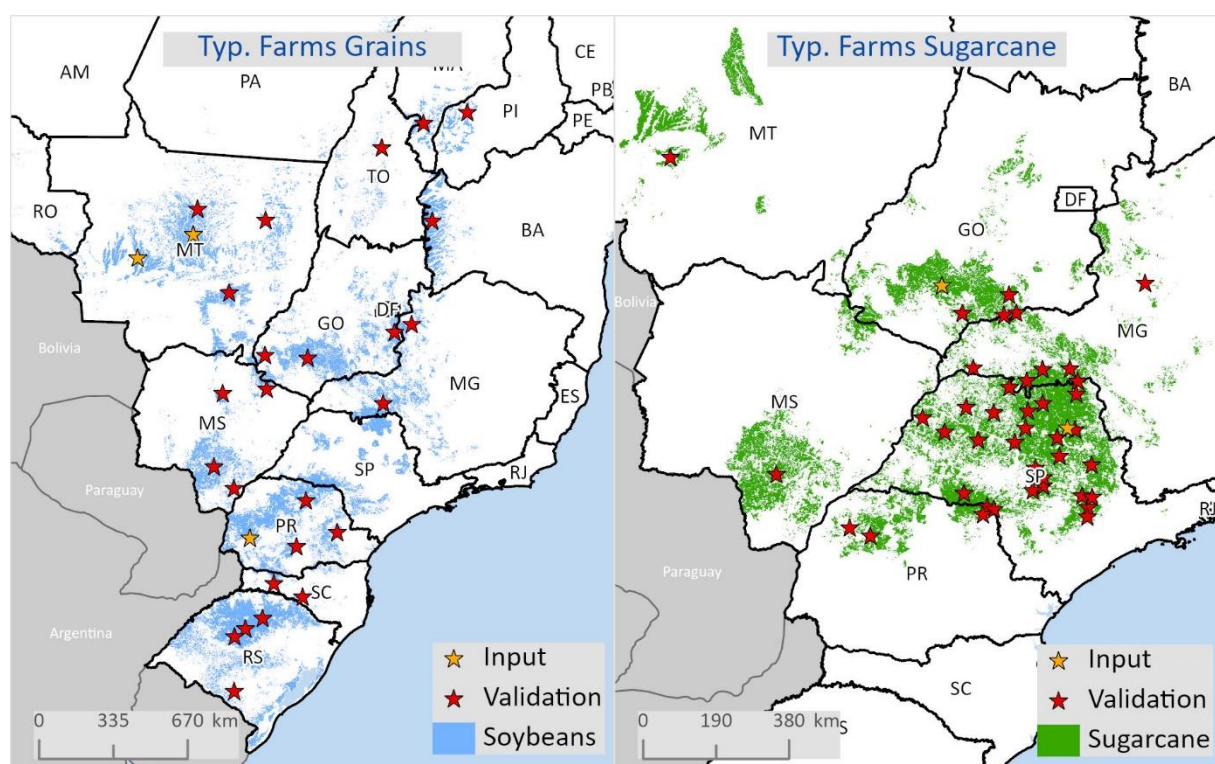
Source: Production from IBGE (2016) – created by the author.

According to Feuz and Skold (1992), it is important to consider the following when selecting the typical farms: (a) the diversity among farms, (b) the level of detail and (c) the criteria regarding how to combine the farms to create a group. The farms selected in Brazil are in the most important production areas, covering the regional differences between production systems as well as representing some of the diversity of farms. For this case study, the level of

---

[17]   For this case study, the sugarcane typical farm BR170RE is not used as input for PAM due to the lack of recent data – data available only until 2016.

detail is regional so that the typical farms should depict the main differences between regions instead of having greater detail within the regions.

To understand the differences among the farms, Table 2 shows the key indicators for all the input farms selected to represent the most important production systems in Brazil.

Table 2:   Key features of the *agri benchmark* typical farms in Brazil

Indicator	Sugarcane			Soybean/Maize		
	BR220ST	BR460RV	BR170RE*	BR1300MT	BR3000MT	BR65PR
Location						
Region	Southeast	Center-west	Northeast	Center-west	Center-west	South
City	SER**	Rio Verde	Recife	Sorriso	CNP**	Cascavel
Climate						
Ann. rainfall (mm)	1,518	1,367	740	1,553	1,776	1,559
Rain distribution	Oct-Apr	Oct-Apr	Apr-Sep	Oct-Apr	Oct-Apr	Oct-Apr
Production						
Farm size (ha)	220	460	170	1,300	3,000	65
Own land (ha)	100	100	100	100	83	100
Full-time labor[1]	2.8	3.0	4.6	5.6	6.3	1.1
Double crop (%)[2]	-	-	-	80	70	70
Cuts per cycle	6	7	7	-	-	-
Mech. Hvst (%)	100	100	0	100	100	100
Data source	PECEGE			CEPEA/CNA		

* Farm data is not used as input for the case study due to the lack of recent data

** SER = *Sertãozinho* and CNP = *Campo Novo dos Parecis*

[1] one full-time labor unit = 2,200 hours per year

[2] share of soybean area with maize as the second crop

Source: own calculations based on agri benchmark (2020) and Climatempo (2020)

It is important to highlight the spatial distribution of the farms, covering the main production areas in Brazil. For sugarcane, the focus is on the Southeast region, which has the highest production within the country, followed by the Center-west and Northeast. Besides the significant difference in the climatic conditions, the share of mechanical harvest indicates the great diversity among the sugarcane areas – the northeast farm having only manual cutting and the remaining farms, 100% mechanical.

For grains, the focus is primarily on the Center-west region (Mato Grosso state) with two farms, followed by a farm in the *traditional* South region. The farm size indicates the regional heterogeneity the farms represent. The farm in Paraná (Cascavel) has only 65 ha compared with Center-west farms size at 1,300 and 3,000 ha, allowing, for example, better understanding of how farm size may impact production costs. The full-time labor endowment is another

indication of varying regional characteristics, extending from 1.1 in Parana to more than 5 in Mato Grosso, which is interesting considering that the latter has 20 times more acreage.

Considering the spatial distribution and the diversity of the key features, the *agri benchmark* typical farms seem to be a suitable basis for representing the diversity of the production systems in Brazil while keeping the sample size feasible to follow the minimum data criteria proposed in this dissertation. The typical farm data is inputted as average for the region where the farms are located as an average from 2016 to 2018 for calibration and the BAU scenario.

Finally, it is important to note that the importance of the typical farm data varies among the different cost components since some key cost items such as fertilizers are derived from the plant module. Throughout the economic module, the user can switch from the *agri benchmark* data as to more detailed costing data if microregional information is available. The importance of typical farm data in each cost component, as well as how these costs are estimated, are the focus of the following sections.

4.5.3.   Crop-establishment costs

The first cost component within the economic module is crop establishment. The definition of crop-establishment costs follows the terminology used within the *agri benchmark* network, where the main components are seeds, fertilizer, and plant protection. The remaining components of direct costs[18] are drying, crop insurance and variable irrigation costs, which are not currently included in the PAM model (Nehring, 2011). According to recent data from agri benchmark (2020), crop-establishment costs account for 59% of the total production costs for soybean, 62% for maize, and 30% for sugarcane on the typical farms in Brazil.

The basic idea is to estimate the physical input use for each crop alternative at SimU level (e.g., kg of N) and the farm-level input prices (e.g., USD/t of fertilizer) to calculate the crop-establishment cost per hectare for each cost component (e.g., fertilizer costs in USD/ha). The physical input use strategy is not valid for crop protection products since only costs per ha are available and it is not meaningful to consider input volume due to the different concentrations and active ingredients. To estimate crop-establishment costs, the PAM model uses the output from the plant model on the input use of fertilizer, combining this information with production costs from the typical farms, following this basic equation:

---

[18]   According to Nehring (2011), direct costs are crop-establishment costs plus drying, insurance and variable irrigation cost.

$$EstCost\_ha_i = SeedCost\_ha_i + PPCost\_ha_i + \\ FertCost\_ha_i \qquad\qquad (9)$$

where $SeedCost\_ha$ are the seed costs per ha; $PPCost\_ha$, the plant protection costs per ha; and $FertCost\_ha$, the fertilizer costs per ha for the SimU i. The following subsections explain in greater detail modeling steps to estimate each subcategory of the crop-establishment costs.

4.5.3.1.   Seed costs

In all the analyzed crops, seed is an important cost item. Of the total crop-establishment cost, they account for 19% in soybeans, 37% in maize and 13% in sugarcane[19], illustrating the importance of correctly estimating this cost component (agri benchmark, 2020). However, it is challenging to link seed costs either to crop output or to consider them as a fixed cost per ha.

The underlying function between seed's physical input in kg/ha and yield is not expected to be linear because the same input use (kg/ha) of two different varieties in different climate conditions may lead to significant differences in yields. Additionally, seed pricing is not directly related to the input volume. A kilogram of genetically modified organism (GMO) soybeans may be valued at twice as much as non-GMO soybeans and they may still produce similar yields, depending on the climate conditions, but leading to quite different cost structure related to crop protection. This costing behavior poses a challenge for modeling seed costs at a regional level.

In the PAM framework, the lack of information on the varieties used by farmers poses another challenge since such data are not publicly available. A farmer is likely to change his/her variety setup every year depending on price expectation and previous experiences – e.g., increase or reduce area with non-GM soybeans. In the process of creating a synthetic typical farm, it is not feasible to agree on a seed portfolio, so that the information collected within the *agri benchmark* network includes only the input use in kg/ha and price. Moreover, the plant module (EPIC) does not provide information on seed usage. It is an input of the model under the crop management definitions affecting biomass only at the start crop cycle.

Against this background, the current version of the PAM model assumes a regional linear seed cost per t of output. The reasoning is that SimUs with much lower yields are likely to have a lower investment in seeds – i.e., the relationship between input costs and output potential, but SimUs with production levels similar to the typical farms are likely to have similar seed costs per ha. The following function is used to estimate seed costs per ha:

---

[19] Based on *agri benchmark* data for 2016 to 2018 in USD/ha for all farms in Brazil. For sugarcane, the whole cycle average is considered so that the high seed cost in the establishment phase is amortized throughout the productive cycle of the crop.

$$SeedCost\_ha_i = SeedCost\_t_{ab_{region}} \times YLD_i \qquad (10)$$

where $SeedCost\_t_{ab_{region}}$ are the regional typical farm average seed costs in USD/t of output and YLD_i is the SimU fresh yields in t/ha, either of grains or sugarcane.

Alternatively, one could use a fixed seed input per ha and spatially correct prices as the ACM model (Deppermann et al., 2018) or keep seed costs as fixed per ha. The caveat regarding keeping seed costs completely fixed is that for areas with significantly lower yields (e.g., -50% compared with typical farms), it is unlikely that farmers would invest the same in seeds per ha as in the relatively high-yielding regions where the typical farms are located. Hence, by adjusting the seed cost with yields, I expect to connect this cost with the expected crop output.

Improving seed cost estimation requires:

- a better understanding of the seed pricing mechanism (trying to answer whether there is a clear connection between seed pricing (cost) and yields).

- additional granular information on the type of seeds that farmers use. That could improve the parametrization of model to account for GMO, non-GMO or farm-saved seeds, for example.

- a better understanding of the relationship between seed input and crop output.

Increasing the complexity of the estimation of seed costs is the fact that farmers may invest in seed in order to optimize other costs such as plant protection. For example, farmers investing in GMO soybeans are willing to spend more for seed while saving on plant protection costs, even if yields are similar to conventional (non-GMO) varieties.

4.5.3.2. Fertilizer costs

Fertilizer costs account for 33% of the total establishment costs in soybeans, 38% in corn, and 52% in sugarcane (agri benchmark, 2020). Following the terminology from *agri benchmark*, the model estimates total fertilizer costs via the following equations:

$$FertCost\_ha_i = Ncost\_ha_i + Pcost\_ha_i + Kcost\_ha_i \qquad (11)$$

where the total fertilizer cost per ha at the SimU i ($FertCost\_ha_i$) is the sum of the nitrogen ($Ncost\_ha_i$), phosphorus ($Pcost\_ha_i$) and potash ($Kcost\_ha_i$) costs per ha. In the following paragraphs, I outline how the individual nutrient costs are estimated.

Nitrogen and Phosphorus input use

For N and P, the PAM model relies on the fertilizer use estimated by EPIC in the plant module. That is a great advantage of EPIC considering actual yield and plant-environment relationships to determine the actual nutrient usage or application. EPIC estimates the total fertilizer applied per ha as well as the total fertilizer used by the crop. PAM allows the user to define whether the actual applied fertilizer amount or the uptake should be considered for the costing calculation.

For this dissertation, the application quantity[20] is used because when choosing between crops farmers are expected to consider the economic returns including input expenses and are unlikely to consider the uptake of natural nutrients present in soils. Due to the lack of long and complex crop rotations, it is plausible to assume that the bulk of nutrition applied is used by the crop in question, differing from regions where fertilization of systems takes place. Finally, the pricing of the nutrients in soils is complex due to the uncertainty regarding their availability (i.e., the share of the total nutrient that the plant can uptake) as well as a natural fixation, for example in soybeans, that can hardly be priced with nitrogen present in the urea.

Potash input use

Potash is the only macronutrient that is not explicitly modeled by EPIC. To estimate K input use within the PAM framework, nutrient uptake by plants is used as a linear function with crop yield – i.e., kg of K per t of output. Uptake per unit of output is kept constant at 20.7 kg/t for soybeans, 6.4 kg/t for maize and 1.5 kg/t for sugarcane, based on the experimental data from Mascarenhas et al. (2004) and Oliveira et al. (2010). According to agri benchmark (2020), typical farms in Brazil apply ca. 90 kg/ha of K for sugarcane and 70 kg/ha for soybeans, indicating the importance of including this fertilizer in production cost estimations.

Fertilizer farm-gate prices

The final step to calculate the fertilizer cost is to estimate farm-gate prices. The standard approach within the PAM model is to use FOB fertilizer prices and convert them to farm level by adding the transportation costs from port to farm. For this case study, I have access to an exclusive database from the research center CEPEA, with observed monthly fertilizer prices for Urea (nitrogen), DAP (phosphorus), and KCL (potash) for 10 major producing states Brazil – averages from 2016 to 2018. This database is the base for the farm-level fertilizer pricing in this version of the PAM model.

The observed fertilizer price information is interpolated to estimate prices for the remaining 16 states that currently do not have a significant production of soybeans and maize. This interpolation uses the price at the nearest state, adding the transport cost to move the fertilizer

---

[20]   The actual amount applied depends on the crop requirements (FTN for N and FTP for P) – not a fixed amount per ha (see section 4.3.3).

to the specific state (see appendix 3). Finally, the fertilizer price in kg of commercialized product is converted to single nutrients on a per kg element basis, following the international standard from *agri benchmark* (Nehring, 2011).

Against this background, the estimation of N and P cost per hectare for each crop follows the equations:

$$Ncost\_ha_i = FTN_i \times Nprice\_t_{uf} \qquad\qquad (12)$$

$$Pcost\_ha_i = FTP_i \times Pprice\_t_{uf} \qquad\qquad (13)$$

where fertilizer costs per ha are a function of the FTP and FTN as the fertilizer applied in kg/ha from EPIC at SimU i, and $Nprice$ and $Pprice$ as fertilizer prices at the state level (uf).

The costs for K are estimated using the following equation:

$$Kcost\_ha_i = Kuptake \times YLD_i \times Kprice\_t_{uf} \qquad\qquad (14)$$

where K cost per ha ($Kcost\_ha$) is a function of the crop yield (YLD), the nutrient update ($Kuptake$) in kg per t of output at the SimU i, and the price ($Kprice\_t$) at the state level (uf).

In a nutshell, fertilizer costs in the PAM model are calculated mainly by combining the input quantity estimated by EPIC and farm-gate prices coming from the transport module. For Brazil, additional fertilizer price information for the main producing states is used to improve the model calibration. Since fertilizer is one of the main cost components within the establishment costs for sugarcane and maize, it is important the estimating is performed regionally, highlighting the importance of the biophysical model.

4.5.3.3. Plant protection costs

Plant protection costs are the final component of crop-establishment costs in the PAM model. Of the establishment costs, plant protection costs account for 48% in soybeans, 24% in maize, and 35% in sugarcane (agri benchmark, 2020). It is, however, the most challenging cost item to estimate regionally.

On one hand, the economic returns for plant protection products are expected to vary strongly regionally and seasonally. For example, fungicide applications are expected to increase yields compared with no application in seasons with higher precipitation and a resulting higher incidence of fungal diseases. However, in drier seasons, farmers may still apply fungicide to safeguard yields despite a limited economic return. On the other hand, a strong diminishing return with increasing application rates is expected. These are strong indications that a linear relationship between plant protection costs and yields is unlikely. Finally, physical input use (e.g.,

kg/ha) is not a suitable parameter due to the large variation between concentration and required grams of active ingredient for different products.

Consequently, the model assumes a constant plant protection expenditure within the region of the *agri benchmark* typical farms due to a relatively loose connection between plant protection expenses and yields. It is important to keep in mind that a regional differentiation is achieved assuming that SimUs within the area of the typical farm receive the plant protection expenses from this farm (or farms) and not the national average.

Additionally, I attempt to improve the accuracy of the plant protection estimation by developing two different plant protection schemes based on the overall productivity (yield) of the SimU. The idea is that farms that have low output potential are likely to reduce their overall expenditures in plant protection to avoid negative margins. The two plant protection schemes are called "intensive" and "extensive" systems and are based primarily on adjustments to the current production systems observed in the typical farms. In the current version, SimUs with yields lower than 50% of the typical farm are classified as extensive and the remaining are considered intensive systems.

It is challenging to develop a plant protection program for the extensive systems because these areas are not under production nowadays. The current production volume in SimU with equal or less than 50% of the average yield of the typical farms represent only 0.1% of the total soybean production, based on IBGE (2015). Under the current economic situation, it is not viable to produce soybeans in such low-yielding areas. Since future scenarios may change this situation, an expert from CEPEA was asked to develop a theoretical plant protection program for such low-yielding areas. According to the expert from CEPEA, the major saving potential is expected to be a reduction of fungicide and insecticide applications. It is not likely that farmers would be able to reduce herbicide costs since desiccation is key for the success of the no-till systems and it is already based on glyphosate application so that cheaper alternatives are not available.

Hence, the two plant protection programs are adaptations of the *agri benchmark* typical farm in Sorriso/MT. In the season 2017/18, GMO soybeans received three fungicide applications and six insecticide applications, which is assumed to be baseline ("intensive system") for the following calculations. The alternative program assumes a lower pressure from insects so that only four insecticide applications are considered, with relatively cheaper products and not using products to control white fly (which are very expensive and not required everywhere). Additionally, only two fungicide applications should enough for such low-yielding regions. This reduction represents a potential plant protection savings of approximately 40% compared with the baseline. Hence, the final equations of estimate plant protection cost in the PAM model are:

$$PPCost\_ha_i = PPCost\_ha_{ab_{region}} \tag{15}$$

If: $YLD_i \leq 0.5\ YLD_{ab_{region}}$ then:

$$PPCost\_ha_i = PPCost\_ha_{ab_{region}} \times 0.60^{[21]} \tag{16}$$

where hectare cost is equal to the plant protection costs from the regional typical farm ($PPCost\_ha_{ab_{region}}$) if yields at the SimU i equal or are higher than 50% of the yield of the typical farms. If this condition is not met, only 60% of the typical farm costs are allocated to the SimU i.

4.5.4.   Operating costs

After the crop-establishment cost, the next important cost item estimated within the economic module is the operating costs. They are an important component of the cost structure of farms in Brazil, especially for sugarcane production, where operating costs account for 56% of total production costs. Conversely, soybean and maize have a relatively higher share of establishment costs, resulting in operating costs of ca. 40% of the total costs (agri benchmark, 2020).

First, it is important to define operating costs. The model follows the terminology employed within the *agri benchmark* network, where the major operating cost components are labor, machinery, diesel and contractors. The following summary illustrates how these cost items are calculated for the typical farms, based on the work from Nehring (2011):

- **Labor costs** account for hired and family labor, considering the former as cash expenditures and the latter as opportunity costs of the family members involved in management and/or production. Costs include actual working hours defined by the typical production system as well as overheads incurred to depict the actual situation on farms.

- **Machinery costs** include depreciation, financing, and repairs. It is noteworthy that the actual lifetime of each machine is used to compute the linear depreciation. If the focus group indicates that certain equipment will be repurchased after its lifetime, the current repurchase price of the new equipment is used to calculate the depreciation (instead of historical prices). Thereby, the model considers that farmers in high-inflation countries need to generate enough returns to repurchase a new machine, including inflation during the depreciation period.

---

[21] For sugarcane, a plant protection cost reduction of 15% is expected for the "extensive" system since the current production system is already based, to a large extent, on biological control, indicating that substantial cost reductions are not likely.

- **Diesel costs** are calculated for each fieldwork operation within the production system. Diesel overheads are then allocated to each crop according to its share of machinery running hours on the total.

- **Contractor costs** are the third-party fee charged for operations. It is inputted per operation including all costs for diesel, labor and machinery as charged by the contractors.

Estimating operating costs is challenging due to the lack of regional-level data as well as information on basic farm machinery setup that could support cost estimation. The alternative approach using typical farm data provides detailed information for the specific regions, but the overarching task is to model operating cost at the unknown locations. The next subsection explores the underlying reasoning behind the estimation steps developed within the PAM framework.

4.5.4.1. Operating costs at sugarcane farms

The first important step is to understand the differences between the farms currently producing grains (soybean/maize) and sugarcane in Brazil. A major difference is the outsourcing of the most important operations to contractors. Currently, around 90% of the machinery costs from the typical sugarcane farms in Brazil[22] are contractor costs. That is the result of outsourcing importing operations such as planting, harvesting, loading, and transportation, mainly to sugar mills that provide these services to growers. Large sugarcane farmers also offer services to smaller growers at fees similar to those charged by mills, indicating that the fee levels are economically set to provide returns on investments, labor, and other cash expenditures. Comparing the harvest cost across all 45 regions from PECEGE, the average cost is USD 9.4/t of fresh cane with a standard deviation of only USD 1.2/t, illustrating the similarity of contractor fees charged in the producing regions in Brazil[22].

Another important feature of sugarcane growing in Brazil is that these operating fees are mainly based per ton of fresh cane, creating a direct link between operating costs and yields. Such costs represent 73% of the total operating costs. The remaining 27% are divided into own machinery (9%), labor (12%) and diesel costs (6%) that are related to ratoon[23] upkeep (spraying and fertilizing), and a few operations during planting and farm management. Therefore, there seems to be a clear relationship between cost structures that are based on yield (largest share) and the remaining based on area.

---

[22] Based on the PECEGE sample of 45 producing regions with typical farms as an average for 2016-2018 seasons.

[23] Ratoon is the common term for cane that emerges from underground buds after the harvest. Commonly the first year is called "plant" and the following five to six harvests "ratoons" - Ramburan, Wettergreen, Berry, and Shongwe (2013)

Accounting for these features in sugarcane production, the PAM model uses a mixed approach combining yield- and area-related costs, using the information from the typical farms as the main parameters:

$$MachCost\_ha_i = MachCost\_ha_{ab_{region}} \tag{17}$$

$$LabCost\_ha_i = LabCost\_ha_{ab_{region}} \tag{18}$$

$$DieCost\_ha_i = DieCost\_ha_{ab_{region}} \tag{19}$$

$$ContCost\_ha_i = ContCost\_t_{ab_{region}} \times YLD_i \tag{20}$$

$$OperCost_{ha_i} = MachCost_{ha_i} + LabCost_{ha_i} \\ + DieCost\_ha_i + ContCost\_ha_i \tag{21}$$

where $MachCost\_ha$ stands for machinery costs per ha, $LabCost\_ha$ for labor costs per ha, $DieCost\_ha$ for diesel costs per ha based on the regional typical farms in USD/ha. $ContCost\_ha$ is a function of contractor cost per t of cane from the regional typical farm ($ContCost\_t_{ab_{region}}$) multiplied by the estimated yield (YLD) at the SimU i.

4.5.4.2.   Operating costs at grain farms

For the grain-producing farms, the situation is more complex because most farms perform field operations themselves, requiring a different approach to correctly estimate their operating costs. To address this issue, I considered different approaches using the complete dataset from CEPEA (all typical farms) to try to correctly parametrize the operating costs estimation. First, a more detailed investigation of the framework conditions in Brazil is required.

Farm size as a driver of operating costs

The first approach considered for the estimation of operating costs for grain-producing farmers is to base the estimation primarily on farm size as the main cost driver. The economic theory offers rich literature on the economics of size as a way of reducing unit cost, by spreading the fixed cost component over a larger number of produced units (Hall & LeVeen, 1978; Raup, 1969).

The marginal effect of the economics of scale is expected to decrease from a certain farm size onwards (Forstner et al., 2018; Hall & LeVeen, 1978). The reasoning is that for the most important fieldwork operations - i.e., seeding and harvesting – operating costs for the first hectares are relatively high since the fixed cost component applies to a small number of produced units, and with an increasing area, a reduction of the unit costs is observed. However, when the full capacity of the first machinery set (i.e., seeder plus tractor and/or combine) is

exhausted, the farmer is motived to expand his/her machinery setup while also increasing acreage accordingly to stay at the most efficient size/machinery combination.

To evaluate the impact of size on the operating cost for grain farms in Brazil, typical farm information from all regions covered by CEPEA is used. The idea is to compare operation costs across different farm sizes to understand whether a clear pattern between farm size and costs can be observed. Figure 15 shows the operating costs including machinery, labor, and depreciation for all 31 typical farms from CEPEA on average from 2016 to 2018 in USD.

Figure 15: **Operating costs USD/ha (left) and USD/t of soybeans (right) for all CEPEA typical farms on average for 2016 to 2018**
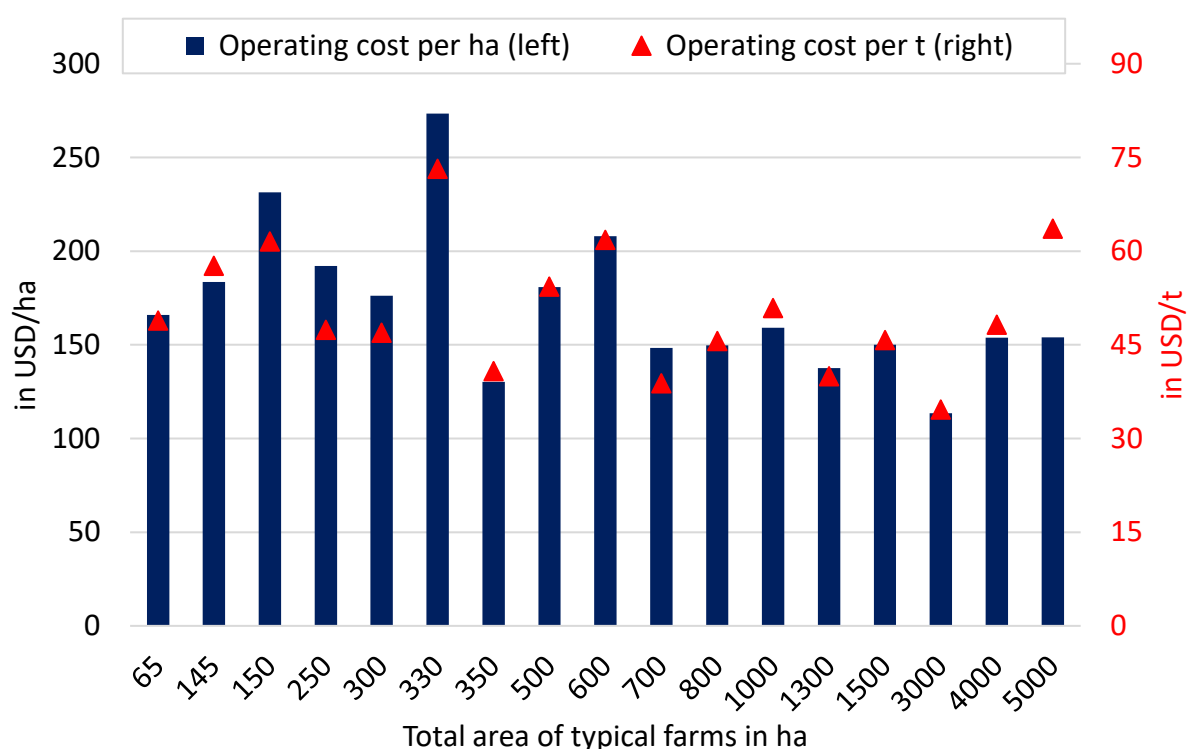
Source: CEPEA (2019) originally in Reais converted to USD using the exchange rate of 1 BRL = 0.3 USD. Created by the author.

Figure 15 indicates that a clear relationship between size and operating costs for soybean producing farms is not found using the only information available on a large scale in Brazil. A possible explanation is that size alone may not explain differences and other factors such as topography, field size, farm structure, and the weather could play a role.

Brazil has well-defined regional segregation of soybean production, dividing the country into two main areas – i.e., *Cerrado* and *South*. That motivates the next attempt to design a tailored modeling approach for the Brazilian grain-producing farmers considering macro-regional differences besides size.

Regional differences and soybean production

Since size alone does not seem to explain differences in operating costs among farms in Brazil, a closer understanding of other factors that are expected to affect machinery performance and farm setup should improve the estimation of operating costs.

Available characteristics that are expected to affect machinery performance and thereby labor efficiency and operating costs mainly include days for operations, topography, location, and shape of fields (Molin, Milan, Nesrallah, Castro, & Gimenez, 2006; Sørensen, 2003). Alone, differences in field shape and size are expected to impact the performance of important operations such as harvesting, via unproductive time lost due to moving machinery between fields and turns.

To cluster areas with similar vegetation, topography, and climate, Brazil has been divided into six **biomes**, indicating that these regions may have different conditions for agriculture. According to the agricultural census of 2017 (IBGE), more than 98% of soybean production is located in two biomes, namely *Cerrado* (56%) and *Mata Atlantica* (42%). Figure 16 shows the biomes in Brazil and soybean production as well as illustrative pictures from the native vegetation of *Cerrado* and *Mata Atlantica*. It is important to highlight that these biomes cover more than 3 million km² with a diversity of vegetation, so the pictures are merely illustrative.

Figure 16:   Brazilian biomes and soybean production
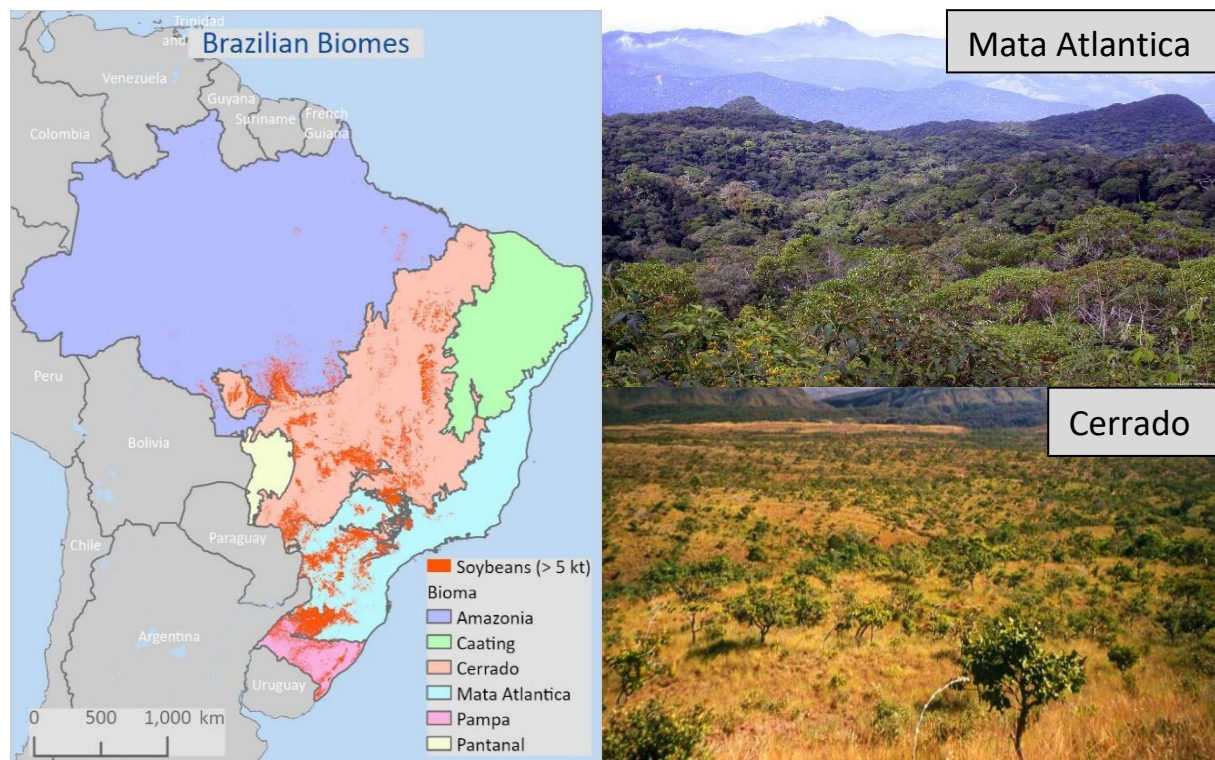
Source: Biomes layer from IBGE (2018) – created by the author; Top picture from www.klimanaturali.org and bottom picture from revistasagarana.com.br (accessed on 31.05.2020)

Differences in topography and climate of the two biomes are significant. Soybean production in the *Cerrado* takes place mainly in areas with a slope lower than 3% (64.3% of the area) and 31.1% with a slope lower than 8%, highlighting the suitability of this region for mechanization (Victoria, Bolfe, & da Silva, 2017). Conversely, *Mata Atlantica* is marked by higher slopes and a greater presence of hills and mountains, increasing the difficulties for mechanical operations.

Rainfall patterns influence the availability of days to carry out field operations so that areas with more frequent precipitation require farmers to invest in more than optimum machinery (as compared with their farm size) to be able to seed and harvest their crops in the available days. Sørensen (2003) argues that available working hours are crucial information in determining machinery size and therefore affect overall operating costs. Furthermore, he concludes that overcapacity is 50% cheaper than under-capacity, illustrating the farmers with unreliable weather conditions are prone to over-mechanize and increase costs compared with areas with greater available working hours. An indication of the availability of days for operations is the number of days with rainfall in the months with a high workload. Soybeans are usually seeded from the middle of September to November and harvested between January and April, followed directly by seeding of maize as the second-season crop. Therefore, the months with higher workloads are expected to be October and February/March as well as July/August, when maize is harvested. Figure 17 shows the historical average share of days with rainfall in a main producing state for each biome.

Figure 17:   Average share of days with rainfall in Paraná (Mata Atlantica) and Mato Grosso (Cerrado) – in %
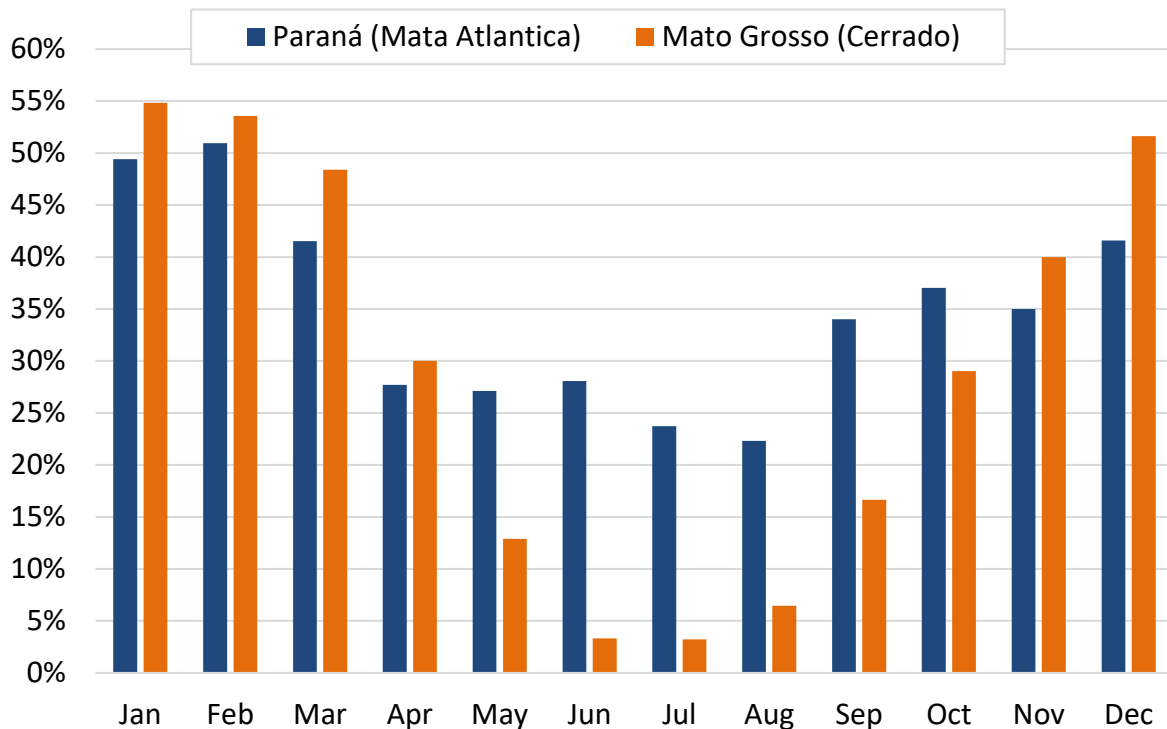
Source: based on Marcuzzo, Oliveira, and Cardoso (2013) for Mato Grosso and on Leite, Adacheski, and Virgens Filho (2011) for Paraná - created by the author.

Figure 17 indicates that the more defined seasonal pattern in Mato Grosso (Cerrado), with a significantly lower number of rainfall days in winter and autumn, indicates that farmers should be able to increase the usage of their machinery for maize harvest and soybean seeding. The latter is crucial to be able to use the rainfall volume in summer. It is important to note that, on average, both regions receive roughly the same 1,760 mm per year but with distinct distribution (Climatempo, 2020).

Finally, the Mata Atlantica biome is characterized as highly populated, with more than 50% of the total population and more than 3,000 cities, indicating that the availability of larger fields is expected to be lower than in the less populated *Cerrado* biome. All these factors suggest that besides size, regional differences may play an important role in machinery setup and efficiency of operations across farmers and therefore impact their operation costs. Hence, the typical farms from CEPEA were divided according to their biome and the results are presented in Figure 18.

Figure 18:   Operating costs USD/ha (left) and USD/t of soybeans (right) for all CEPEA typical farms on average for 2016 to 2018 – classified according to their biome
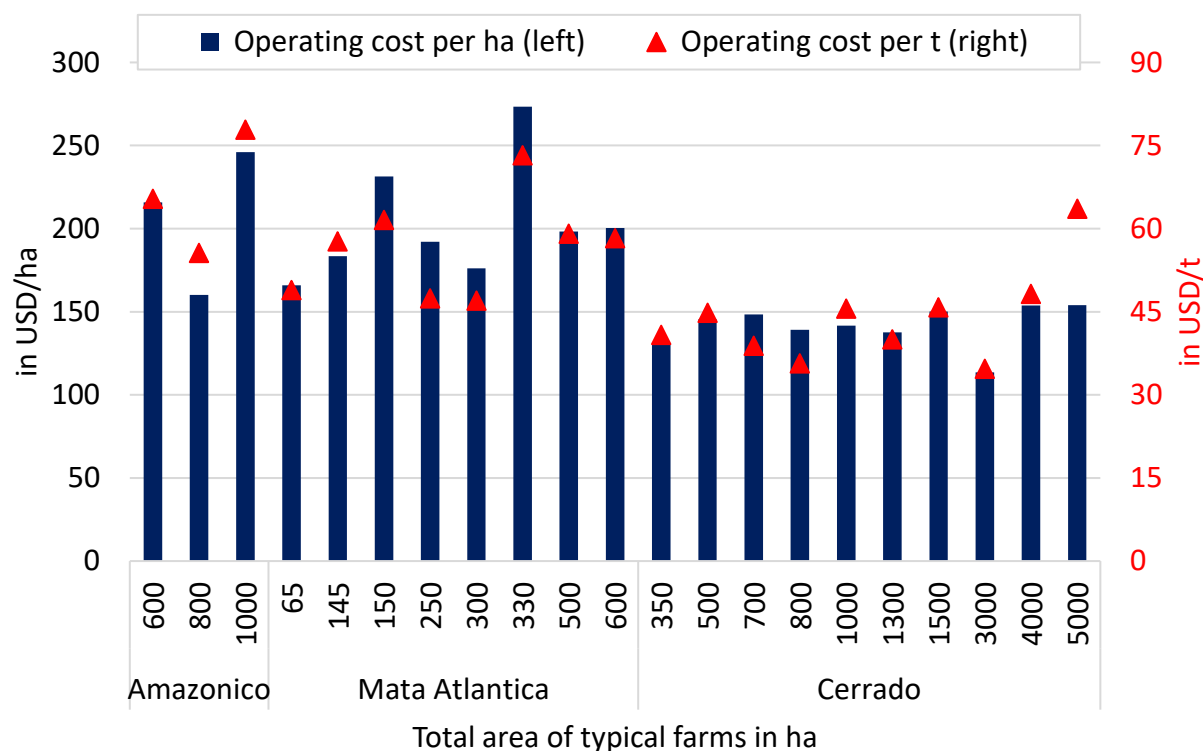
Source: CEPEA (2019) originally in *Reais* converted to USD using the exchange rate of 1 BRL = 0.3 USD. Created by the author.

Figure 18 indicates that, indeed, the combination of the different regional characteristics plays an important role in estimating operating costs for soybean farmers. The farmers in the biome *Mata Atlantica* have higher operating costs even at farm sizes similar to farmers in *Cerrado*. Therefore, to estimate operating costs for grain farms in Brazil, it is important to consider their regional location, suggesting that at least one typical farm in each macro-region should be required to allow robust estimations. There appears to be no close relationship between operating costs and yields (see Figure 18), which is expected since machinery setup and labor are defined for the average level of yield that can be attained in the region, reducing the influence of annual yield changes in the operating costs.

Even though size alone does not seem to fully explain the differences in operation costs, it is noteworthy that farms in these regions have significant differences in size. According to IBGE (2019), more than 83% of the total soybean production in the *Mata Atlantica* biome is produced by farms with less than 200 ha, whereas roughly 60% of the soybean production from *Cerrado* is grown by farms with more than 1,000 ha, indicating farms adopted different optimal operating sizes depending on their growing conditions.

Against this background, the PAM approach divides Brazil into two main regions for the estimation of the operating costs, namely *Traditional* (largely the *Mata Atlantica* biome) and

Expansion (for the *Cerrado* biome). Overlapping the biomes boundaries with the state level to create the regions facilitates the connection between statistical data that is provided at the state level (fertilizer base prices, taxes), which would be difficult to achieve by using the exact biome boundaries. Figure 19 shows the regions for estimation of operating costs for grain farms within the PAM approach as well as the distribution of soybean production according to the size of farms in each region.

Figure 19:   Regions for operating cost estimation in the PAM approach (left) and the share of soybean production per class of size of farms in each region (right)
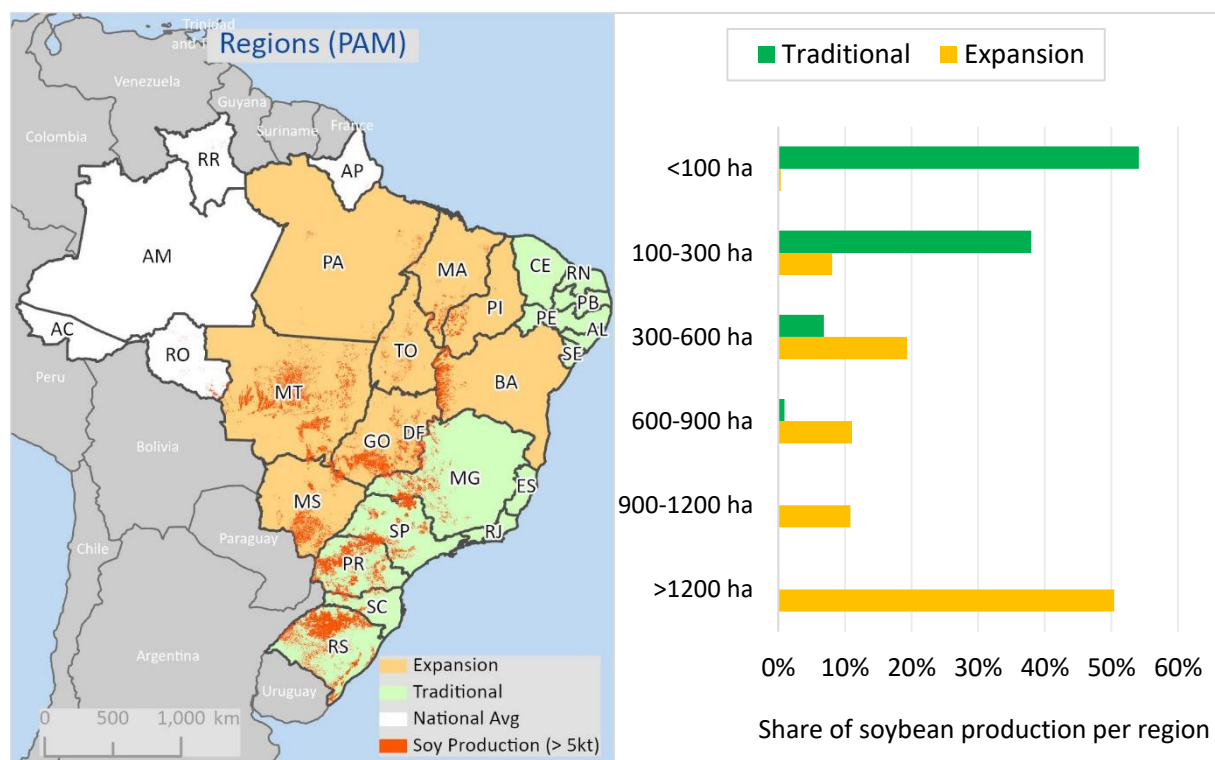
Source: IBGE (2019) – created by the author.

It is important to consider future scenarios including the expansion of planted area to land currently in pasture or native vegetation. Crop land expansion in Brazil has taken place mainly in the MAPITOBA region, with farm structure and growing conditions similar to the main producing areas in *Cerrado* (e.g., Mato Grosso). Therefore, this area is considered part of the *Expansion* region within the PAM approach.

Since yield is not expected to have significant impact on operation costs in the grain farms, the estimation uses per ha cost from the typical farms in the defined regions for the cost estimation. It is important to highlight that the available farms also meet the structural size as

described earlier, with the typical farm from the *Traditional*[24] region having 65 ha and the two typical farms in the *Expansion* Region having 1,300 ha and 3,000 ha. The formula for estimation of operating cost in the grain farms is as follows:

$$OperCost\_ha_i = OperCost\_ha_{ab_{region}} \tag{22}$$

where the $OperCost\_ha$ at the SimU i is equal to the operating cost at the typical farm(s) in the region where the SimU i is located[25].

Summary

Based on the available production costs data, farm size alone does not seem to explain the differences in operating costs in grain-producing farms in Brazil. Considering additional regional characteristics such as topography, climate and field size may improve operating costs estimation. That implies at least one farm for each region is necessary to depict inter-regional differences. Due to the lack of a clear connection between operating costs and yields, the model assumes cost as similar on a per ha basis within the defined regions. Additional information and evaluation of the drivers of operating costs should help improve the estimation even further.

4.5.5.   Crop prices at farm-gate level

A key piece of information required to estimate the economic return at the farm level is the crop price. This information is commonly not available in the high-resolution used by the PAM approach. Therefore, I propose an endogenous estimation of prices based mainly on reference prices and transportation costs. It is important to mention, that by endogenously converting FOB prices to farm-gate prices, the model allows an easy input solution to run simulations with projected or observed prices. Finally, a tailored approach is proposed to estimate farm gate prices for sugarcane since it is not commercialized as fresh cane, resulting in a more complex payment system.

4.5.5.1.   Based on reference price (FOB)

Farm-level prices at the SimU levels for maize and soybeans are estimated by deducting the estimated transport cost (farm to the nearest port) from FOB prices. This assumes that farmers in exporting countries receive international prices minus the costs to reach the market –

---

[24] Within the PAM framework, the regions are renamed as "*Traditional*", mainly representing the Mata Atlantica biome, and "*Expansion*", accounting mainly for the production areas in the Cerrado. This is necessary since the administrative state-level regions are considered instead the biome classification scheme.

[25] The detailed assignment of the typical farms' data to each region for grains and sugarcane production is shown in appendix 11.

i.e., getting the product to the harbor (Freebairn, 1987). Farm-gate prices for grains are esti-mated following the equation:

$$GatePrice\_t_{j_i} = FOBprice\_t_j - Transp\_t_{j_i} \qquad (23)$$

where the $GatePrice\_t$ is the farm-gate price in USD/t, $FOBprice\_t$ is FOB price and $Transp\_t$ is the transport cost, for the crop j in the SimU i. FOB prices come from CEPEA daily in BRL per bag[26]. The information is converted to USD/t and averaged arithmetically.

4.5.5.2.   Farm-gate prices for sugarcane

The transport cost approach to estimate farm-gate prices, as applied for soybeans and maize, is not suitable for sugarcane. The farm output (fresh cane) must be processed and the two main outputs – i.e., sugar and ethanol – are then commercialized. Therefore, an alternative concept is proposed to estimate farm-gate prices for sugarcane production. In Brazil, cane farmers receive shares of the proceeds from the sales of sugar and ethanol at the mill level, according to a predefined formula (CONSECANA, 2006). Therefore, ethanol and sugar prices at the SimU level are necessary to calculate the resulting cane price (TRS price).

Sugar prices at the SimU level are estimated with the same approach as for grains – i.e., FOB minus transport costs. That is realistic since most of the sugar produced in Brazil is exported (UNICA, 2019). Ethanol prices are more challenging since the largest share of the production is consumed domestically. The National Agency of Oil, Gas, and Biofuels (ANP) reports monthly data on ethanol prices at the retail level for roughly 500 municipalities in Brazil (ANP, 2020) – see appendix 4. However, these are gas-station levels while the reference for the TRS formula is at the mill level. Therefore, the model uses CEPEA mill-level ethanol prices for the state of São Paulo, which is converted to the remaining locations by the following steps:

(1)   Removing the sales tax, which differs by state, from the observed retail prices (ANP).

(2)   Calculating the relationship between the average retail prices without tax of each state to São Paulo (ANP).

(3)   Using the resulting relationship (item 2) to convert mill-level prices from São Paulo (CE-PEA) to state-level prices elsewhere.

(4)   The resulting state-level prices are finally converted to the municipality level by the same logic – based on the observed relationship between the state and municipality retail prices.

---

[26] Sugar: reference Santos, Soybeans: reference Paranaguá; maize: reference Campinas (https://ce-pea.esalq.usp.br/en)

(5)   The SimU receives the ethanol price of the nearest municipality, based on the geodesic distance using the GIS software.

The idea is that within the states the major driver of the price difference should be transportation and competition among gas stations[27]. Having the sugar and ethanol prices at the SimU level, the TRS price is estimated as:

$$TRSsugar\_t_i = GatePrice\_t_{sugar_i} \times 0.9528 \times 0.595 \qquad (24)$$

$$TRSeth\_m^3{}_i = GatePrice\_m^3{}_{eth_i} \times 0.5966 \times 0.6210 \qquad (25)$$

$$
\begin{aligned}
TRSfinal\_t_i = \; & (TRS\ sugar\_t_i \times MixSug_{uf}) \\
& + (TRS\ eth\_m^3{}_i \times MixEth_{uf})
\end{aligned}
\qquad (26)
$$

where TRS (e.g., $TRSsugar\_t$) for each product is a function of the price at the SimU i (e.g., $GatePrice\_t_{sugar}$), the chemical conversion between TRS and the product (e.g., 0.9528 for sugar), meaning converting one t of TRS produces 953 kg of sugar. CONSECANA (2006) also defines the share of the final product value that goes to cane growers, which, in the case of sugar, is 59.5%. The final TRS ($TRSfinal\_t$) depends on the production mix, which is share of sugar and ethanol produced at the state level (uf). The production mix data at state level comes from CONAB (2021), as an average for 2016-2018.

Summary

The endogenous estimation of farm-level crop price is a major advantage within the PAM approach, allowing an easy simulation of observed or projected prices. The underlying idea is to use the transportation cost to correct reference prices to farm-gate prices. For sugarcane, the overall payment system is more complex due to the mandatory processing and the different destinations of the final output – i.e., sugar and ethanol. Having a dedicated approach to account for these differences should help improve the model's accuracy significantly.

4.5.6.   Profitability: estimating returns to land

The final step within the economic module is measuring the economic performance of each production alternative at each SimU. Several indicators are suitable to infer the economic performance of farming activities – e.g., gross margin, return to land and profit. The selection of the indicator depends strongly on the research question as well as the temporal scope of the decision process.

---

[27]   The current version of the PAM model for Brazil, allows the user to choose a simplified ethanol price estimation by using observed margins between retail and distribution prices from ANP to convert gas-station prices to the distribution level. Hence, the model uses the distribution prices instead of mill-level prices. That, however, will overestimate prices and may cause biases due to the different levels of tax.

For many farmers, comparing production alternatives at the gross-margin level is sufficient (usually including the variable cost for labor and machinery) because this metric is relatively easy to measure and in the very short-term the impact of fixed costs and the opportunity cost of own-labor and capital may be negligible. However, for more complex and lasting decisions such as to moving to new crops, which may involve investing in new machinery, changing labor endowments or expanding the farming operating, a more detailed indicator such as return to land or profitability is required (Seo & Mendelsohn, 2008).

Within the PAM framework, the goal is to assess the economic performance of different production alternatives within the farm. That motivates the use of return to land instead of profit since the former excludes land costs, which should not affect cropping allocation decisions within the farm. Another important advantage of using return to land instead of profitability is avoiding the complexity of estimating land rent and value. Land values should originate primarily from the expected return to land, but several other factors such as potential real-estate development, taxation and government programs create imbalances between the expected returns and the actual values (Borchers, Ifft, & Kuethe, 2014).

Against this background, the **return to land** is estimated as follows:

$$RetLand\_ha_{j_i} = GrossRev\_ha_{j_i} - TotCost\_ha_{j_i} \qquad (27)$$

where (grains):

$$GrossRev\_ha_{j_i} = YLD_{j_i} \times GatePrice_{j_i} \qquad (28)$$

$$TotCost\_ha_{j_i} = EstCost\_ha_{j_i} + OperCost\_ha_{j_i} + MiscCost\_ha_{j_i} \qquad (29)$$

for sugarcane:

$$GrossRev\_ha_{sug_i} = YLD\_trs_i \times TRSfinal\_t_i \qquad (30)$$

$$
\begin{aligned}
TotCost\_ha_{sug_i} = {}& EstCost\_ha_{sug_i} + OperCost\_ha_{sug_i} + \\
& MiscCost\_ha_{sug_i} + InbT\_ha_{sug_i}
\end{aligned}
\qquad (31)
$$

where the return to land ($RetLand\_ha$) is a function of gross revenue ($GrossRev\_ha$) and total costs without land ($TotCost\_ha$), all per ha for the crop j in the SimU i. For sugarcane, total costs also include transport from farm to mill ($InbT\_ha$).

The model also estimates gross margin to react to different research questions. That may be important for regions where the robust estimation of operating cost is challenging due to the lack of typical farms or unreliability of data.

4.6.   Land allocation module: highest return to land

The land allocation module is the final step of the PAM modeling approach, acting as the decision-maker in selecting the most profitable production alternative for each SimU. This module is programmed separately from the previous modules, combining their output. The basic idea is to identify the production alternative with the highest return to land, accounting for constraints and thereby identifying the "best" option for each SimU.

In the case of Brazil, the land allocation module is relatively straightforward since more complex crop rotations are not typical for the considered crops. Farmers producing soybeans may decide to grow maize as a second-season crop, creating a crop sequence, but it should not be considered crop rotation since, at the same time each year (e.g., summer), the same crop is grown in all areas within the farm. That differs strongly from more complex production systems as observed in Europe, with crop rotation varying from three to many crops increasing the complexity of the decision process. The presence of diverse crop rotations increases the complexity of the estimation because the main assumption of farmers selecting the crop with the highest return to land may not fully depict the reality since several other factors such as technical restrictions (e.g., avoid the same plant protection group) or yield penalty (e.g., wheat after wheat), for example, are considered and affect farmer's decision.

For Brazil, the model allows for full specialization, where farmers may produce only sugarcane, a combination of soybeans and maize (no-till), or beef. These choices are common practice in the main production regions in Brazil. To follow the no-till principles, farmers must grow a second crop after soybeans or maize to provide soil cover and the biomass required to keep the system functioning (Moraes Sá et al., 2015). Therefore, the economic performance of monoculture soybeans needs to include the costs of cultivating a second crop.

The predominant alternative as the second crop is maize due to the possibility of generating at least a positive return to land compared with expenses of growing a cover crop. In the model, the alternative "soybeans" includes either the cost of cultivating a cover crop in winter or maize, depending on the economic performance of the latter. Currently, the model allows a maximum of 80% of maize double cropping because that is the highest observed in the typical farms and higher shares are technically challenging to achieve because of the short window for seeding maize after soybeans.

The main steps from the land allocation module are:

(1)   Combining the results from each crop coming from the previous module into a unique working file.

(2)   Using information from statistics to calculate the current share of soybean acreage grown with maize as the second crop for each SimU ($ShareDC_i$) – see appendix 5.

(3)   Calculating the final return to land for the soybean double-cropping:

If $RetLand\_ha_{maize_i} > 0$ then:

$$RetLand\_ha_{dc_i} = RetLand\_ha_{soy_i} + \qquad\qquad (32)$$
$$(RetLand\_ha_{maize_i} \times ShareDC_i)$$

else:

$$RetLand\_ha_{dc_i} = RetLand\_ha_{soy_i} - CostCover^{28} \qquad (33)$$

(4)   Selecting the production alternative with the highest return to land among all available alternatives for each SimU.

The output from the module is analyzed using the GIS software, enabling the visualities of the regional differences in production alternatives. More importantly, GIS is used to combine the model results with existing information on current land use (e.g., rivers, forest, etc.), allowing the inclusion of geographic restrictions such as not allowing production in the forest, considering only current arable or the combination of arable and pasture.

Finally, it is important to highlight that the PAM model does not have a demand function. The land allocation results should be interpreted carefully since they indicate the alternative with the highest return to land in each SimU using estimated farm-level prices, costs and yields.

---

[28]   The cost of growing a cover crop is estimated at 73 USD/ ha as an average from 2016-2018 (*agri benchmark*).

4.7.   Business-as-usual (BAU)

The focus of this dissertation is the development and evaluation of the proposed PAM modeling approach. The BAU scenario therefore is of major importance because its output is used for the model evaluation. For this scenario, production costs from the typical farms are an average of 2016 to 2018, as is the fertilizer price coming from the CEPEA database as well as the technical parameters at the state level (e.g., the mix of production for sugarcane, TRS, etc.). The main exception is the information used in the regression analysis for freight costs, which are from 2017. However, this should not cause major problems because the relationship between the coefficients is expected to remain stable. In the plant module, the baseline technology for EPIC is used, selecting the production schedule and input use similar to current non-irrigated levels observed in Brazil.

The most important variable for the BAU scenario is crop price. Figure 20 shows the price development for maize, soybeans, sugar, and ethanol in Brazil as well as the price assumption for the BAU scenario in USD/t or m³.

Figure 20:   **Historical and BAU average of FOB prices for soybeans, maize, sugar, and ethanol (in USD)**
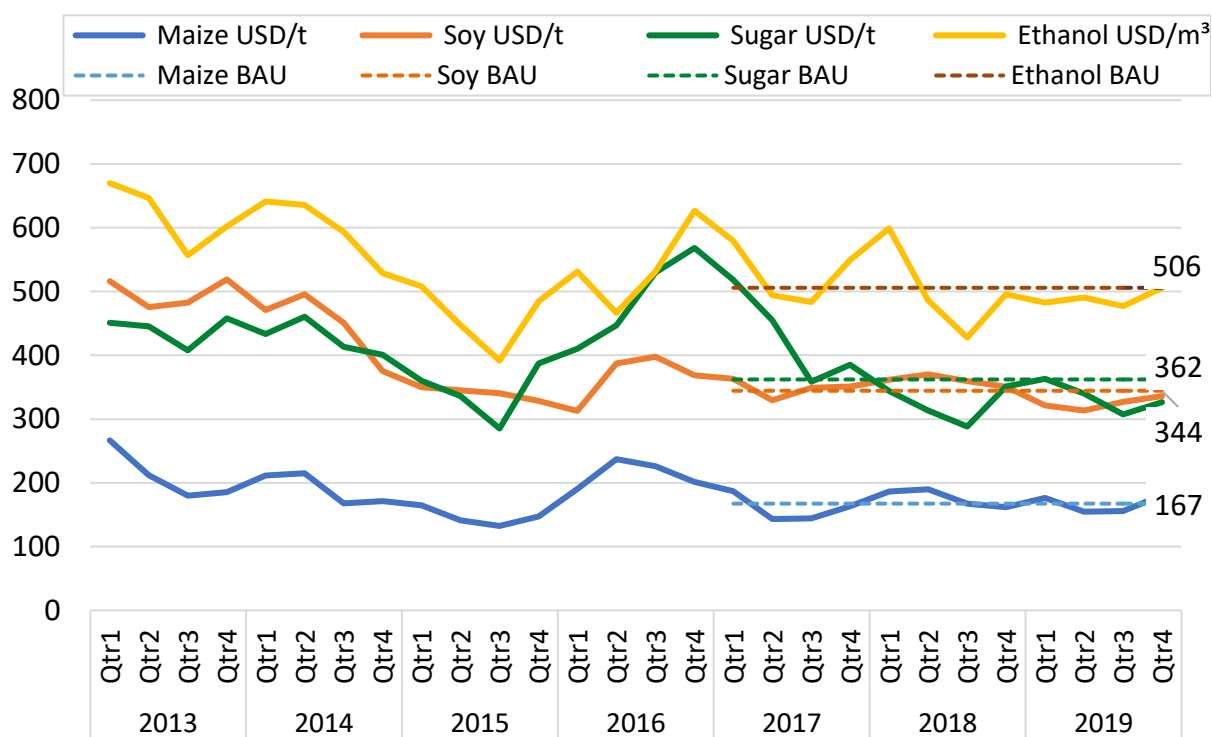
Source: CEPEA (2020) – created by the author.

Prices in 2016 were considerably higher than the long-term average, with strong variations throughout the year. Hence, for the BAU scenario, the average price for 2017 to 2019 is considered as a reference price for the estimation of farm-gate prices. These prices are 344 USD/t for soybeans, 167 USD/t for maize, and 362 USD/t for sugar. The reference ethanol price in São Paulo is 506 USD/m³, which is converted to SimU levels using a different approach (see section 4.5.5.2). Finally, I simulate the BAU scenario with actual inbound transport costs to represent the status quo.

5.   Results of the business-as-usual scenario

This chapter presents the results of the PAM model for the Brazilian case study. The focus is on the results from the BAU scenario while benchmarking results against observed production costs data and available spatial information on the current land allocation of soybeans, maize and sugarcane in Brazil. To explore the variability of the model, the results often include the current grassland, but the main focus is on current arable land. The inclusion of grassland using the same calibration for the BAU provides an interesting discussion on the competitiveness of crops displacing grassland. The BAU results do not include areas currently classified as native vegetation due to the complexity of assuming conversion costs and the lack of typical farms.

Since yield is a major driver of the economic performance of the production alternative, the following section focuses on interpreting the plant module results; production cost results are then presented; and, finally, the output from the land allocation model is shown. It is important to highlight that the production cost results shown in the maps do not indicate that sugarcane or grains currently are grown in these areas. The model estimates production costs for each individual SimU that has arable (and grass) land for the whole country.

5.1.   Yield simulations

Crop yields are the major source of model variation and understanding the plant module's behavior is key to ensuring that the economic module can be interpreted correctly. Yield ratios are an important part of the on-farm competitiveness of crops and therefore achieving a good model fit for yields is ideal.

Soybeans

Figure 21 compares the yields from soybeans estimated by the plant module (EPIC) on a state-level average compared with the official statistics from IBGE (avg. 2013 to 2015)[29] as well as the total harvested area for each state in 2015.

---

[29]   The most recent information available from IBGE that has been allocated to the SimU level.

Figure 21:   **EPIC and IBGE soybean yields on a state-level average (in t/ha) and the harvested area in 2015 from IBGE (in million ha)**
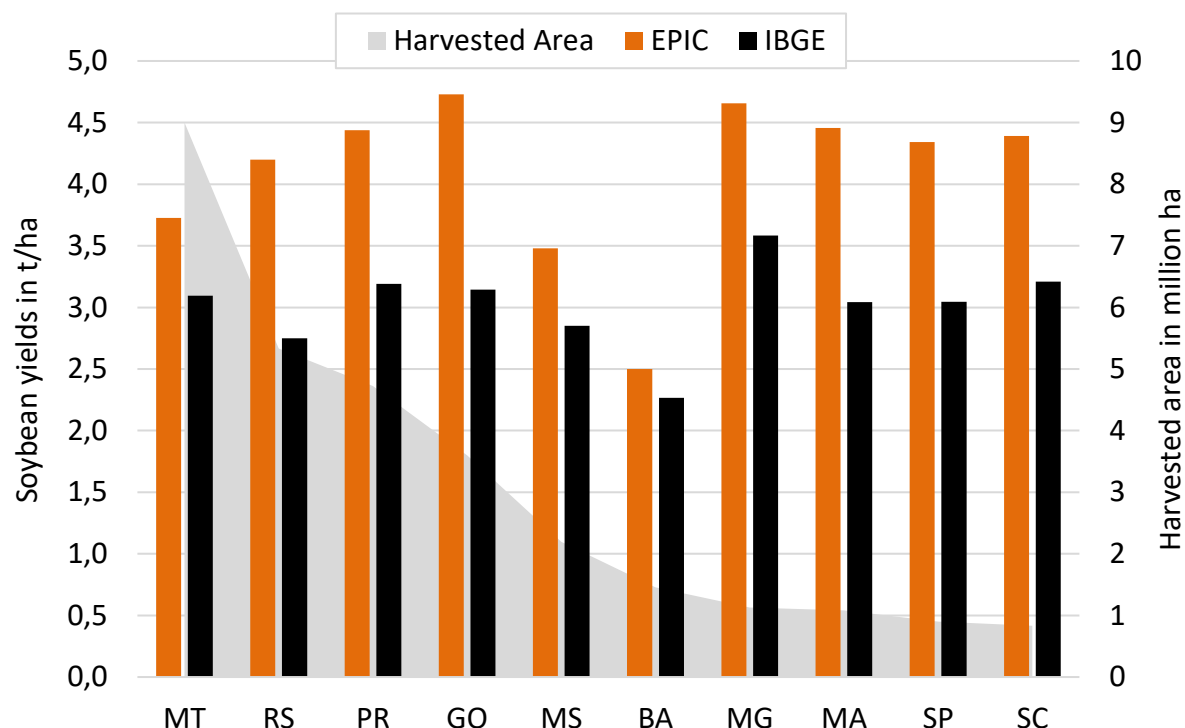
Source: results from EPIC and IBGE official statistics (2015) – created by the author.

The standard EPIC calibration for Brazil tends to overestimate soybean yield when compared with official statistics. That effect is consistent across all the states in Brazil, including the main producing regions such as Mato Grosso (MT) and Paraná (PR). EPIC performs better in the state of Bahia (BA), with yields very close to the observed statistics. The reason behind this trend could be the lack of more regional calibrations in Brazil for the crop growth parameters (field trials) as well as for the different management systems. Also, the absence of yield penalty in EPIC due to pests and diseases may be a strong driver of this estimation behavior. In tropical and in intensive production systems such as those in Brazil, the pressure from pests and diseases is expected to be significant.

Fresh cane yields

Figure 22 shows the comparison between fresh cane yields estimated using the standard IIASA calibration in EPIC and the official statistics allocated to SimU level from IBGE (avg. 2013 to 2015).

Figure 22:   EPIC and IBGE fresh cane yields on a state-level average (in t/ha) and the harvested area in 2015 from IBGE (in million ha)
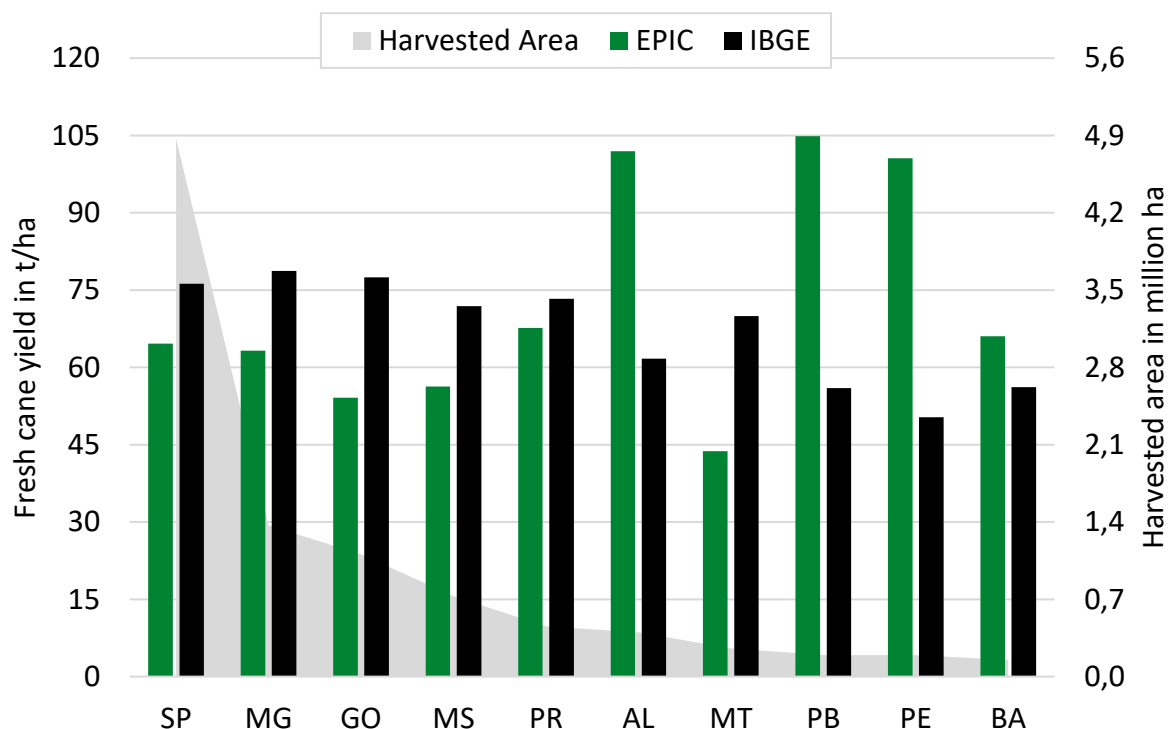
Source: results from EPIC and IBGE official statistics (2015) – created by the author.

For sugarcane, the performance of the yield estimation using the standard calibration in EPIC varies strongly among states. In the four main producing states – namely São Paulo (SP), Minas Gerais (MG), Goiás (Go), and Mato Grosso do Sul (MS), EPIC underestimates sugarcane yields. The biggest difference is in the Goiás states (GO) with EPIC estimating 54 t/ha whereas the official statistic is 77 t/ha. In the Northeast region (AL, PB, PE), EPIC significantly overestimates fresh cane yields. That is an interesting finding given that this region is much drier than the CS region, which can be partly seen by the yield information from statistics. Here it is important to note that no irrigation is allowed in the EPIC simulation used in the BAU scenario.

Second-season maize

Figure 23 shows the comparison between the yields estimated using the standard IIASA calibration for EPIC and the statistics from IBGE (average 2013 to 2015).

Figure 23:   EPIC and IBGE maize second-crop yields on a state-level average (in t/ha) and the harvested area in 2015 from IBGE (in million ha)
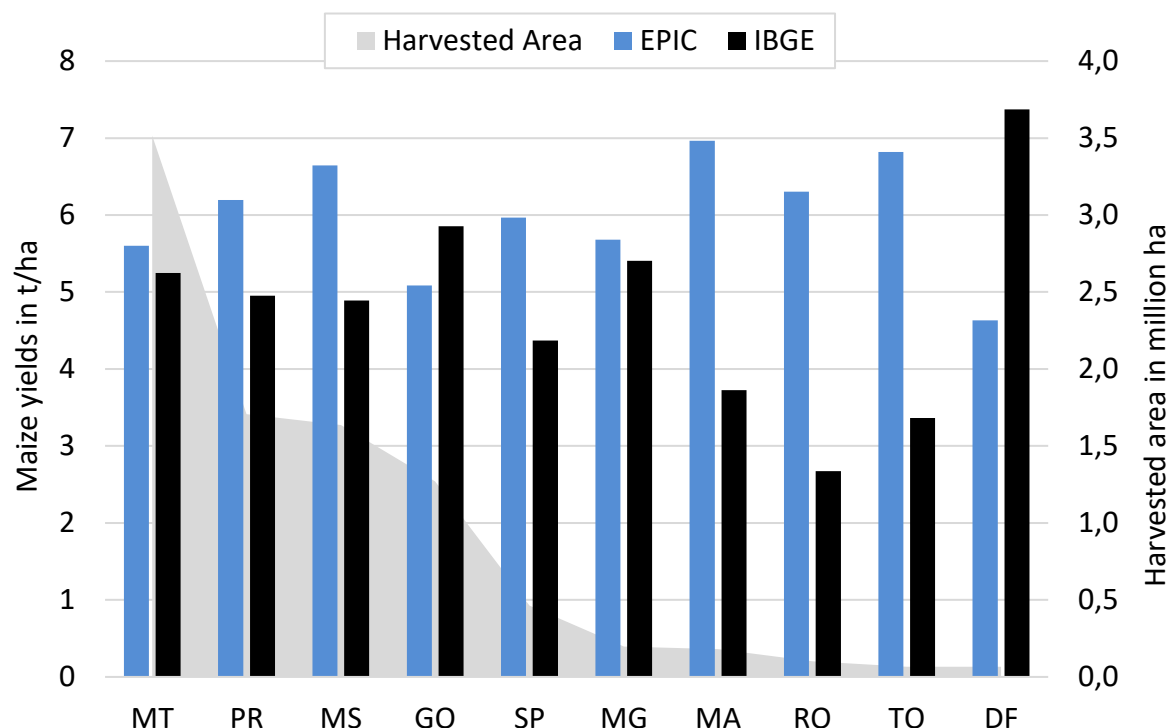
The EPIC estimation for Mato Grosso (MT), the largest producing state is satisfactory, with EPIC estimating 5.6 t/ha and the official statistics at 5.3 t/ha. Since this state alone produces roughly 50% of the total second-season maize output in Brazil, it is possible to argue that EPIC is well suited for the maize yield estimation in the main producing areas. The strong differences found in the states of Maranhão (MA), Rondônia (RO) and Tocantins (TO) should not be overinterpreted. Because these states have a relatively small acreage of maize as a second crop, the statistical data may not fully represent what could be achieved in the state if more areas were in production.

The overall error from EPIC while estimating the yield of maize as a second crop is only 3% if a weighted average (based on harvested area) is used for the comparison, supporting the view that EPIC is well-calibrated for the main producing regions in Brazil. That is likely to be the result of intensive calibration and adjustments within EPIC carried out by IIASA since this cultivar was mainly implemented to depict the double-cropping systems in South America.

Yield adjustments for the Brazilian case study

Considering that standard IIASA calibration of EPIC currently overestimates soybean yields while underestimating sugarcane (for the main producing regions), it is challenging to use the original output from EPIC for the case study. It is expected to create unrealistic yield ratios and therefore affects the competitiveness of the farming alternatives. Improving the EPIC

estimations directly demands time and funding beyond the scope of this project. To enable the simulation for this case study, the PAM model uses the available data from IBGE to adjust the EPIC yields, generating a more realistic picture of the current yield levels in Brazil. If IBGE data are available, PAM uses the yield data from IBGE instead of EPIC for the specific SimU[30]. Such correction is possible only because more than 90% of the SimUs (arable and grassland) in Brazil have crop yield data from IBGE. That is important because for SimUs without IBGE data, the yield information relies on EPIC. Figure 24 shows the yield data availability from IBGE for the area currently classified as arable and grassland.

Figure 24:   Area with available IBGE yield data for pasture and arable land in Brazil
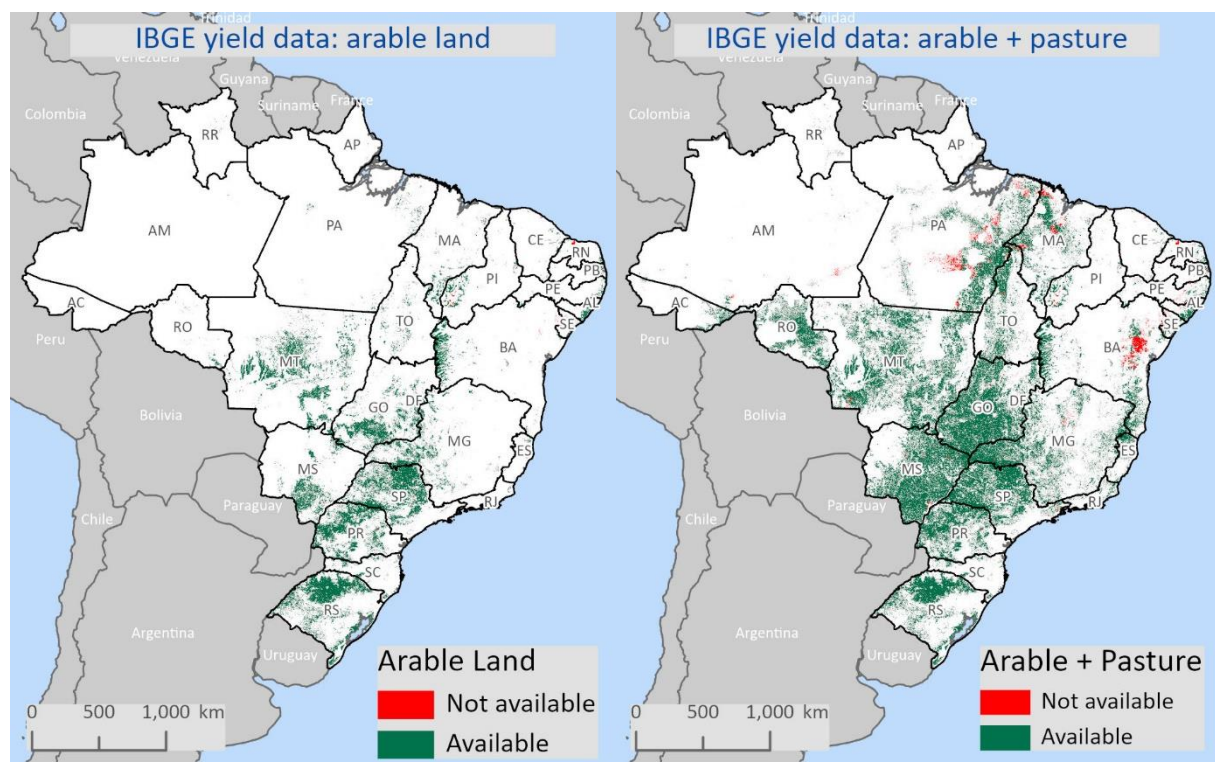
Source: IBGE (2015) – created by the author.

It is important to highlight that this yield correction for Brazil should represent a future version of the PAM model with a more accurate yield estimation either based on EPIC or on any existing crop growth model. The adjusted EPIC yields are used as the base of the BAU scenario, validation and scenario analysis.

---

[30]   For 1,119 SimUs, the sugarcane yield information from IBGE is lower than 30 t/ha, which is extremely low. For these SimUs, the model assumes a yield of 30 t/ha of fresh cane. These SimUs together produced only 0.07% of the total sugarcane output in 2015, which helps explaining why the IBGE data may not be reliable (based on a small area and volume).

BAU yields

Table 3 shows the results from the plant module (adjusted yields) for the main producing states in Brazil sorted by the total planted area in 2019. On average, sugarcane yields around 74 t/ha, whereas soybeans yield around 2.9 t/ha and maize as a second crop, 5.5 t/ha. It is interesting to observe the yield ratio differences among states. Focusing on the largest soybean producing state (MT) and the largest sugarcane producing state (SP), they have similar soybean yields around the national average. However, on average, São Paulo has 7 t/ha higher sugarcane yields. This is a first indication of the natural conditions favoring sugarcane production over soybeans in São Paulo.

Table 3:   Yields from PAM model and planted area per state (2015)

States	Sugarcane	Soybeans	Maize (2nd)	Planted area*
	t/ha			'000 ha
Mato Grosso (MT)	70	3.0	5.7	1,661
Paraná (PR)	75	3.2	5.4	1,057
Rio Grande do Sul (RS)	45	2.8	-	911
São Paulo (SP)	77	2.9	5.0	817
Goiás (GO)	77	2.8	6.2	652
Mato Grosso do Sul (MS)	72	3.0	5.2	579
Minas Gerais (MG)	79	2.7	5.3	560
Bahia (BA)	59	2.6	7.4	407
Maranhão (MA)	69	2.8	4.3	171
Piauí (PI)	53	2.2	4.8	163
Santa Catarina (SC)	50	3.1	4.9	144
Tocantins (TO)	79	2.9	3.3	143
Pará (PA)	78	2.9	3.0	142
Average (weighted)	74	2.9	5.5	

* Includes all arable crops in 2019 (IBGE)

Source: own calculations and IBGE (2015, 2019)

The yields shown in Table 3 are the weighted average of the state but a significant variation within the simulation units is observed. This variation is of key importance for the overall analysis since it represents the information gain of downscaling from administrative (e.g., state or county) level to the local level. Figure 25 shows the yield estimates from PAM at the SimU level, highlighting the yield variation for the most important states producing soybeans (Mato Grosso) or sugarcane (São Paulo).

Figure 25:   PAM yields on sugarcane in São Paulo (left) and on soybeans in Mato Grosso (right) for arable and grassland (in t/ha)
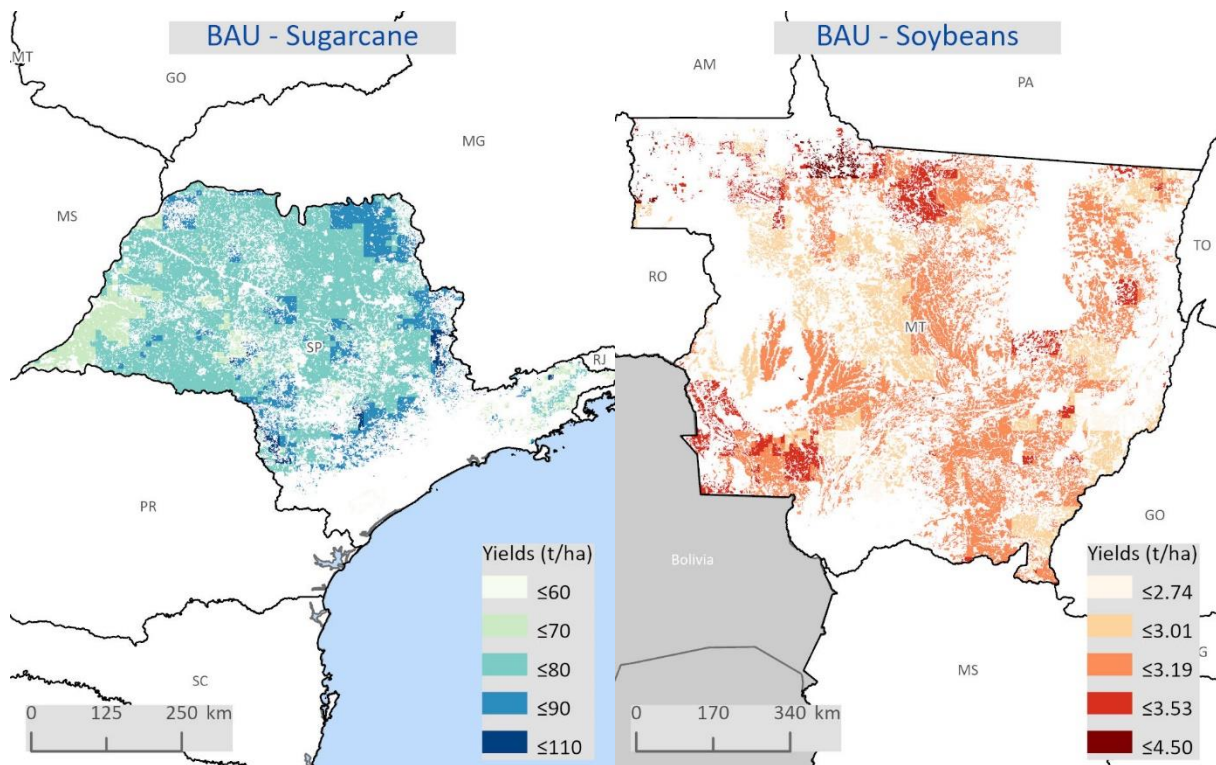
Source: PAM model estimation – created by the author.

It is interesting to observe that for sugarcane, yield in the state of São Paulo varies from less than 55 to more than 100 t/ha, building the weighted average of 77 t/ha. This significant variation illustrates the importance of regional estimation to avoid the oversimplification caused by low-resolution information. A similar picture can be drawn for soybeans in Mato Grosso (MT), with yields varying significantly around the weighted average of 3 t/ha.

Summary

Yield estimation using EPIC alone is currently challenging for the overall modeling approach because the overestimation of soybean yields as well as the error on sugarcane estimation may lead to biased results by artificially changing the yield ratios in some regions. It is important to consider that, if the crop growth model had the same overall error, it may not be a significant issue since yield ratios could be correct. Therefore, for the overall analysis, the EPIC yield estimates are adjusted using the available information from IBGE allocated to SimU.

The weighted average yield for sugarcane is 74 t/ha; for soybeans, 2.9 t/ha; and maize as the second-season crop, at 5.5 t/ha. The disaggregated yield results already show a strong spatial variability within the states and SimUs, highlighting the importance of spatial disaggregation when modeling farmers' supply response.

5.2.   Production costs

The major contribution of the PAM approach is the estimation of production cost at the farm (SimU) level. The results from the BAU scenario, with adjusted yields, are presented in the following subsection. The structure follows the methodology of the production cost components. First, I present the results from crop-establishment costs estimation, focusing mainly on fertilizer and crop protection costs. Next, results from the operating cost estimation are shown, highlighting the diversity of systems and costs observed in the main producing regions. Finally, the results from the sugarcane inbound transport cost, the total costs without land and the most important indicator of the return to land are presented.

The results are obtained for all 11,000 SimU, which makes visualization in graphical (chart) form challenging. Therefore, most of the results are shown in the form of maps, highlighting the variation among the regions as well as a state-level average for the 10 biggest producing states in Brazil.

5.2.1.   Crop-establishment costs

In the PAM economic module, crop-establishment costs are the combination of seed, fertilizer and plant protection. Fertilizer costs play a major role in overall production costs and therefore are the first cost component presented.

Fertilizer costs

Figure 26 shows the fertilizer cost estimated for the BAU scenario for soybeans in Mato Grosso (MT) and sugarcane in São Paulo (SP).

The average fertilizer cost (i.e., N, P, K) is 268 USD/ha[31] for sugarcane in São Paulo, with a variation from ca. 250 to 400 USD/ha depending on the production intensity and yields of the SimU. As expected, the regions in the state with higher yields have also higher fertilizer costs, which can be explained by the following factors:

- The application mode used in EPIC reacts to the need of the plants so that higher yields mean more demand for nutrients.

- Since K application rates are directly linked to output (based on the crop uptake), the higher the yields the higher the costs per ha.

---

[31]   All weighted, based on harvested area for 2015 from IBGE allocated at SimU level by INPE – Câmara et al. (2015).

Figure 26:   Fertilizer costs for sugarcane in São Paulo (left) and soybeans in Mato Grosso (right) for the BAU scenario (in USD/ha)
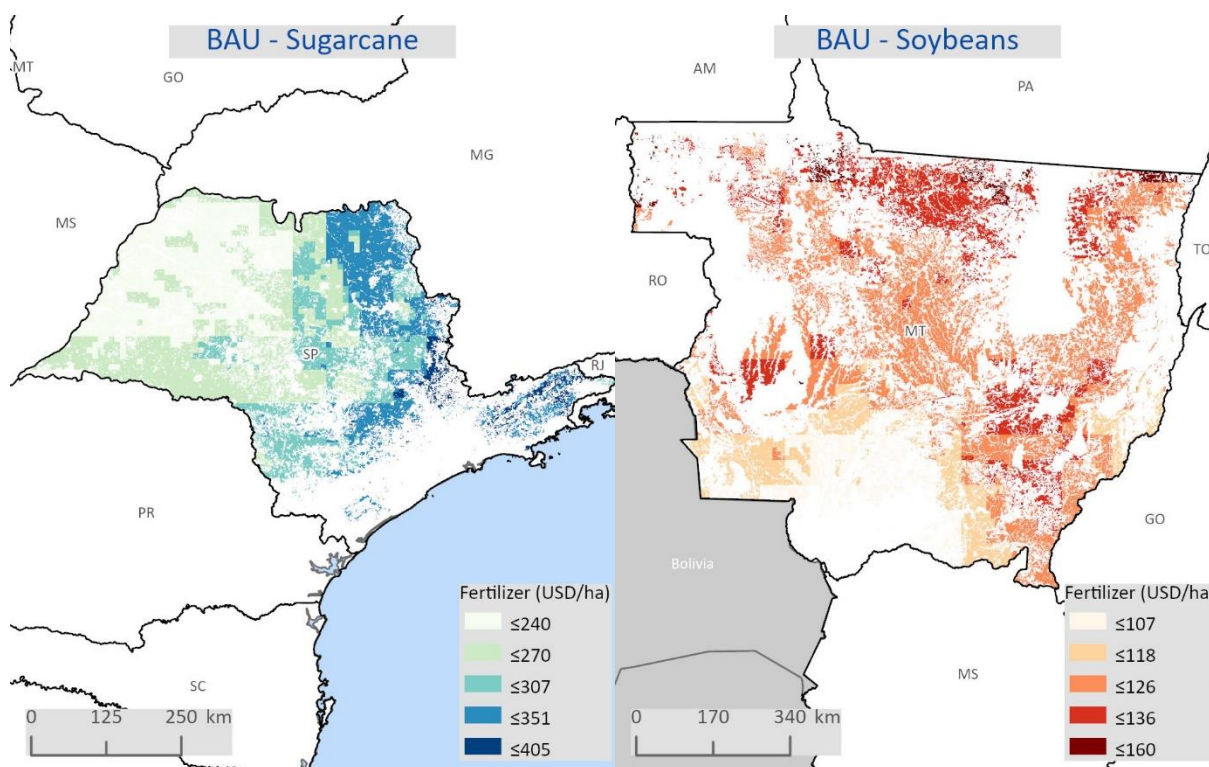
Source: PAM model – created by the author.

For soybeans, fertilizer costs range between 100 and 160 USD/ha in the state of Mato Grosso, with a weighted average of 123 USD/ha. The main cost component is P accounting for 58% of these costs. The importance of P as the main nutrient cost in soybeans also is observed in the typical farm data from *agri benchmark* (65%). That is interesting considering that the information from the typical farms is not used in fertilizer cost estimating within the PAM approach. In comparison to sugarcane, soybeans have significantly lower fertilizer costs, mainly due to the absence of N costs for soybeans due to their biological nitrogen fixation, compared with ca. 100 USD/ha spent for this nutrient in sugarcane production.

Even though the total fertilizer cost is an important indicator of the overall economic performance of a crop, it is important to identify the different drivers of total fertilizer costs – i.e., the price of nutrients vs their input use in kg/ha. Figure 27 shows the estimated pure nutrient input in kg/ha as a weighted average (based on harvested area) for the main soybean and sugarcane producing states in Brazil.

Figure 27: **Estimated pure nutrient input (in kg/ha) for the main soybeans and sugarcane producing states as a weighted average (based on harvested area). The error bars show the standard deviation**
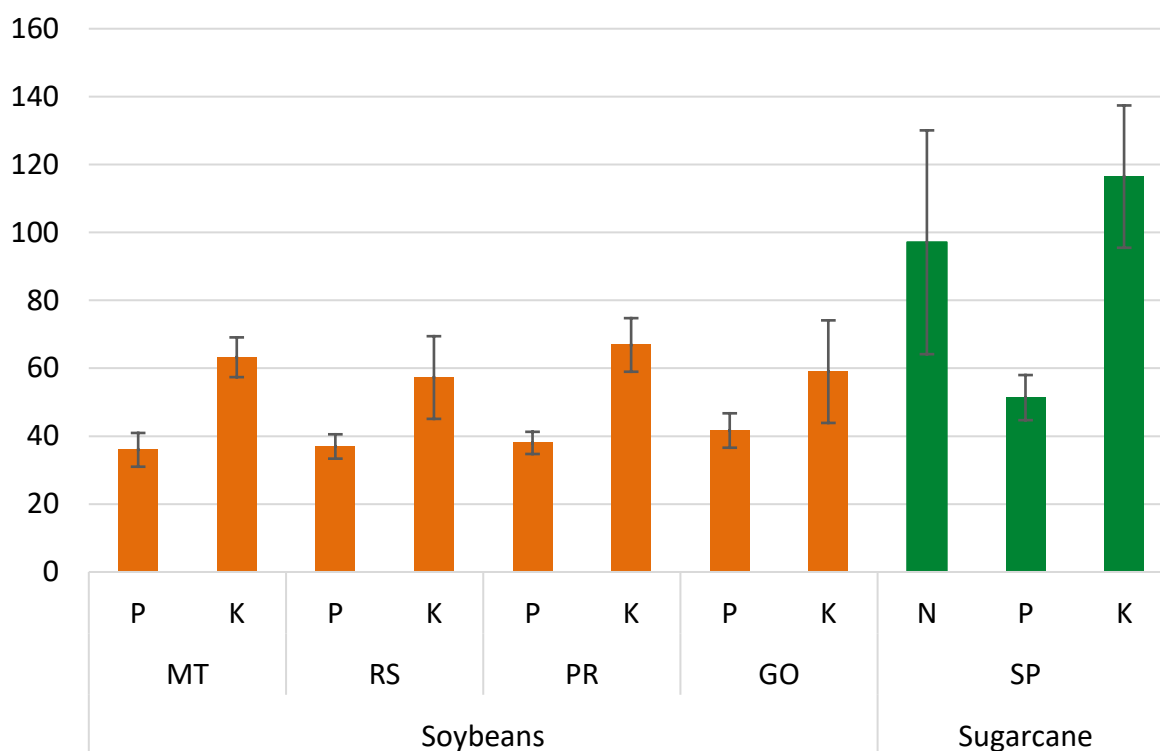
The input use information highlights that sugarcane has a much higher fertilizer use per ha, which is in line with the expectation due to the higher biomass production. Focusing on the average, N input for sugarcane in São Paulo is estimated at 97 kg/ha and K at 116 kg/ha, roughly double the K input required for soybeans. P input in sugarcane is ca. 35% higher than the average for soybeans. The strong standard deviation (SD) in N and P input use for sugarcane is largely due to the high variation in yields since these nutrients are closely linked to the crop output – which increases the SD on a per ha basis.

Even though the yields in the BAU scenario are correct using the IBGE data, the final fertilizer input use per t of output is similar between the observed typical farms' data and the output from the PAM model. For soybeans, the P input from the PAM (BAU) model averages 11.6 kg/t of soybeans; whereas in the typical farms, this value varies between 8.5 and 11.4 kg/t. Nitrogen input in sugarcane averages 1.3 kg/t of fresh cane in the CS region, whereas the typical farm data varies between 1.2 and 1.4 kg/t of sugarcane (BR220ST and BR460RV, respectively). The strongest variation between PAM results and the typical farm data comes from P input in sugarcane, with the former having 0.66 kg/t whereas the latter 0.18 kg/t. That is an unexpected outcome because the yield correction for sugarcane is an increase in yields, which decreases the input per t of output. Hence, EPIC seems to overestimate the application of P. Finally, the yield correction should not influence the K input because it is calculated in the PAM model using the update based on crop yields (not from EPIC).

Since the PAM model explicitly estimates input use as well as farm-gate nutrient prices, it is possible to run scenarios, for example, looking at the impact of fertilizer prices on farming costs and land allocation decisions. According to ANDA (2020), 81% of the total fertilizer used in Brazil in 2020 was imported, indicating the high dependency on the world market. Fluctuation in the currency, for example, is expected to impact fertilizer prices strongly and relatively more in sugarcane production due to the higher input usage per ha. The ability to isolate the impact of nutrient price in the overall fertilizer costs is a major advantage of the PAM economic module.

Plant protection costs

Another important component of the crop-establishment cost is plant protection expenses. Following the methodology proposed, two different plant protection systems are proposed depending on the intensity of production. Most of the areas in São Paulo (SP) and Mato Grosso (MT) have relatively high yields and therefore face similar plant protection costs within the state. For sugarcane, plant protection costs for São Paulo average 190 USD/ha, mainly driven by insecticides and herbicides. The plant protection cost average for Mato Grosso for soybeans is 193 USD/ha, with insecticides representing roughly 50% of the plant protection costs. Comparing the two states, crop protection costs are very similar between the two crops.

Seed costs

The third component of establishment costs is seed expenditures. Due to the methodological limitations on seed input as well as seed price, the results from PAM strongly follow the yield distribution and the difference in the production system of the SimU. For soybeans, the average seed cost for Mato Grosso is 73 USD/ha whereas the average for sugarcane over the whole production cycle is 70 USD/ha. Even though these costs are similar, the perceptions of farmers are very different. For sugarcane, seed costs are a one-time expenditure at the establishment phase. Since most farmers use their cane as seedcane, they commonly underestimate or do not even account for this cost component in their budgeting. The individual decision of farmers further complicates the interpretation of seed costs for sugarcane, since they can change the duration of the sugarcane cycle (i.e., 5 or 6 years) by leaving the crop to ratoon another year. This ability to influence seed costs is more restricted for soybean farmers, whose main option is to use cheaper seeds (e.g., non-GMO), which, in turn, may generate higher plant protection costs or decrease yields, making this option usually uneconomical to farmers.

Total crop-establishment costs

Figure 28 shows the estimation results of the total crop-establishment cost for sugarcane and soybeans, considering all areas that are currently classified as arable and grassland in Brazil, in USD/ha.

Figure 28:   **Crop-establishment costs for sugarcane (left) and soybeans (right) for the BAU
                      scenario (in USD/ha)**
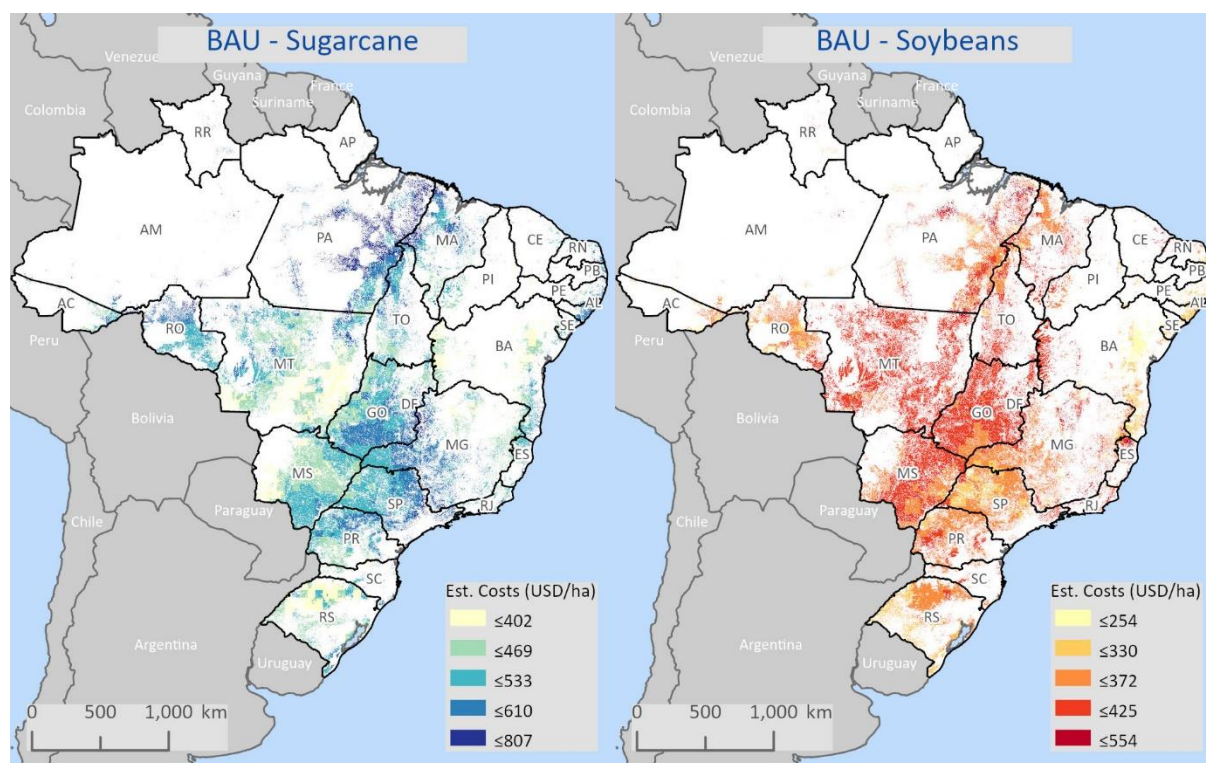
Source: PAM results (2020) – created by the author.

The weighted average of crop-establishment costs for sugarcane in Brazil is 530 USD/ha, with
a strong variation from 223 to 731 USD/ha. In the case of soybeans, the average is 373 USD/ha
with a spread between 220 to 440 USD/ha. This finding indicates that sugarcane is a more
cost-intensive crop at least while looking at crop-establishment costs. It is important to note
that establishment expenditures are expected to play an important role in farmers' risk per-
ception since higher cash requirements also lead to great exposure if the crop fails. Figure 29
shows the disaggregation of crop establishment into its main components as a share of total
crop-establishment costs.

Figure 29:   Composition of estimated crop-establishment costs for soybeans and sugarcane in the five main states producing each crop (in %)
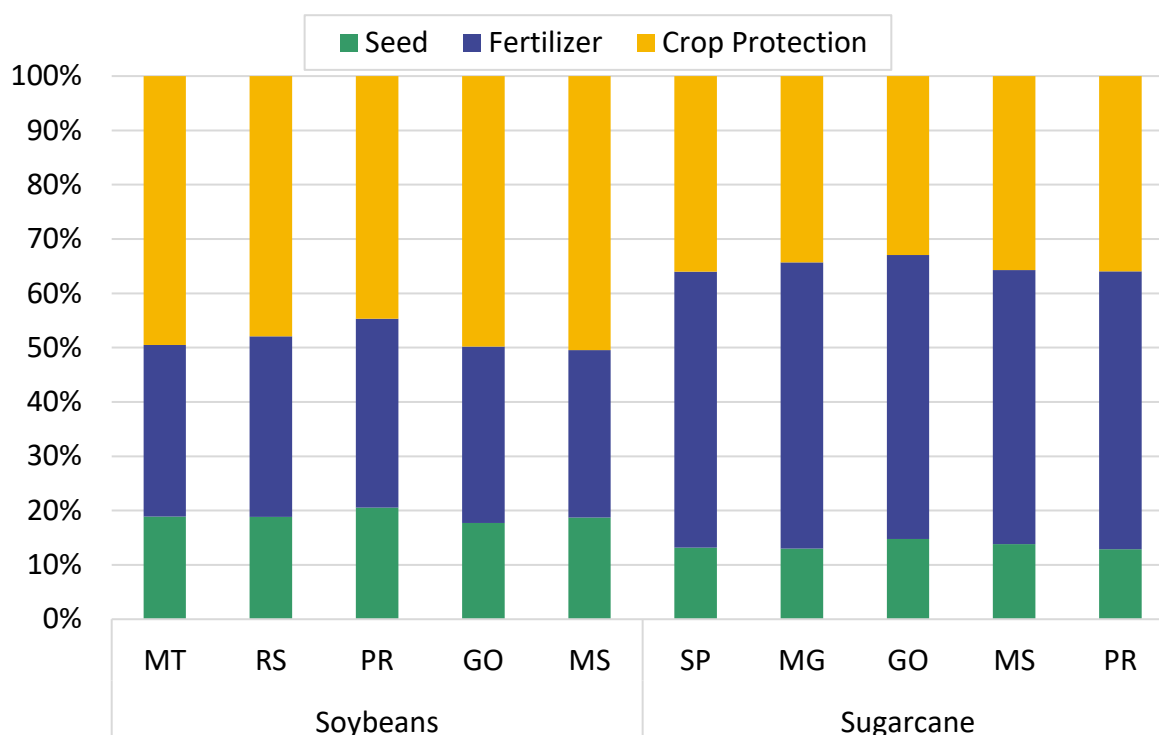
Source: PAM results (2020) – created by the author.

Looking at the main components of the establishment costs, it is interesting that the largest cost component for sugarcane is fertilizer, at 52%, whereas for soybeans it is plant protection, at 49% of the establishment costs. This illustrates the importance of biological N fixation in soybeans, giving this crop a performance edge over sugarcane. The cost component with the lowest share is seed for both crops, representing 13% and 19% for sugarcane and soybeans, respectively.

Even though the focus of the establishment cost section is on the main crops – namely soybeans and sugarcane, it is important to look at some results for maize as a second crop as well because it impacts the overall performance of double cropping. The weighted average for the total crop-establishment cost for maize as the second crop is 342 USD/ha, with a variation between 159 and 563 USD/ha. In contrast to the other crops, seed cost is as important as fertilizer, at 38% of the establishment costs for each, and the remaining 24% as plant protection costs.

To compare the actual cropping alternatives, Figure 30 shows the comparison of estimated crop-establishment costs between sugarcane and a double-cropping system, assuming a 60% share of the soybean area grown with maize as the second crop. The information is on a weighted average (based on harvested area) for the 10 states with the largest share of arable area in Brazil.

Figure 30:   **Crop-establishment costs for sugarcane and double-crop (soybeans + maize) with maize grown on 60% of the soybean area (in USD/ha)**
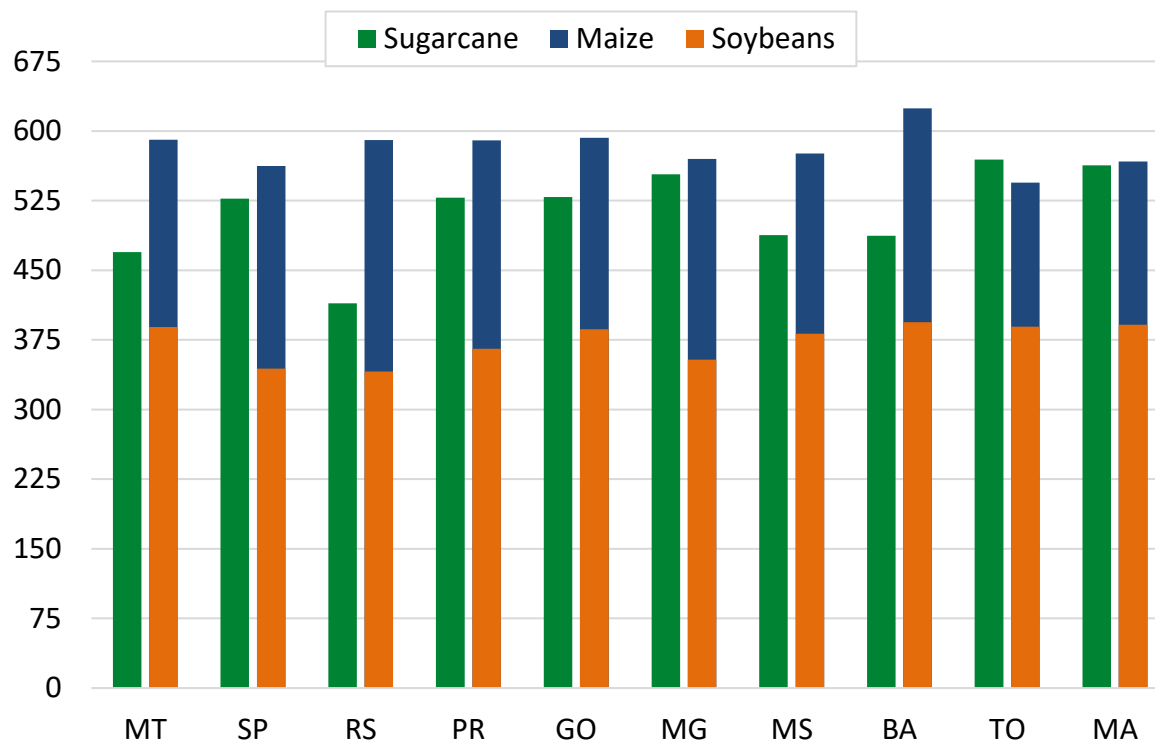
Source: PAM results (2020) – created by the author.

Considering a typical double-cropping system (60%), the total establishment cost as a weighted average for Brazil is 579 USD/ha, with the largest share of cost coming from seeds (45%), fertilizer (35%), and plant protection (20%). As shown in Figure 30, there is a significant variation among the states, with the overwhelming majority having a higher crop-establishment cost for the double-crop compared with sugarcane. The only exception is Tocantins (TO), where double cropping is slightly less expensive than sugarcane.

Summary

In most producing states, double cropping of soybeans and maize has higher crop-establishment costs than monoculture sugarcane: averages of 579 and 530 USD/ha, respectively. Breaking down the cost items, fertilizer is the main cost driver for sugarcane, whereas plant protection and seed costs (maize) are the main drivers for grain production. The ability to explicitly estimate nutrient input as well as fertilizer price at the SimU level is a major advantage of the PAM modeling approach allowing the simulation of a variety of scenarios and how they impact farm-level economics.

5.2.2.   Operating costs

Apart from the establishment costs, the economic module of the PAM model estimates oper-
ating costs at the SimU level. The main components of the operating costs are machinery,
diesel, contractors, and labor. In contrast to crop-establishment costs, the interpretation of
the individual components for operating costs is not straightforward. Some production sys-
tems rely more on labor, others on machinery, so a cross-comparison of machinery costs is
not meaningful. Therefore, the focus in the following is to look at the total operating costs for
sugarcane and soybeans for areas classified as arable and grassland in Brazil. Figure 31 shows
the operating costs estimated by the PAM model for the BAU scenario in USD/ha.

**Figure 31:   Operating costs for sugarcane (left) and soybeans (right) for the BAU scenario
(in USD/ha)**
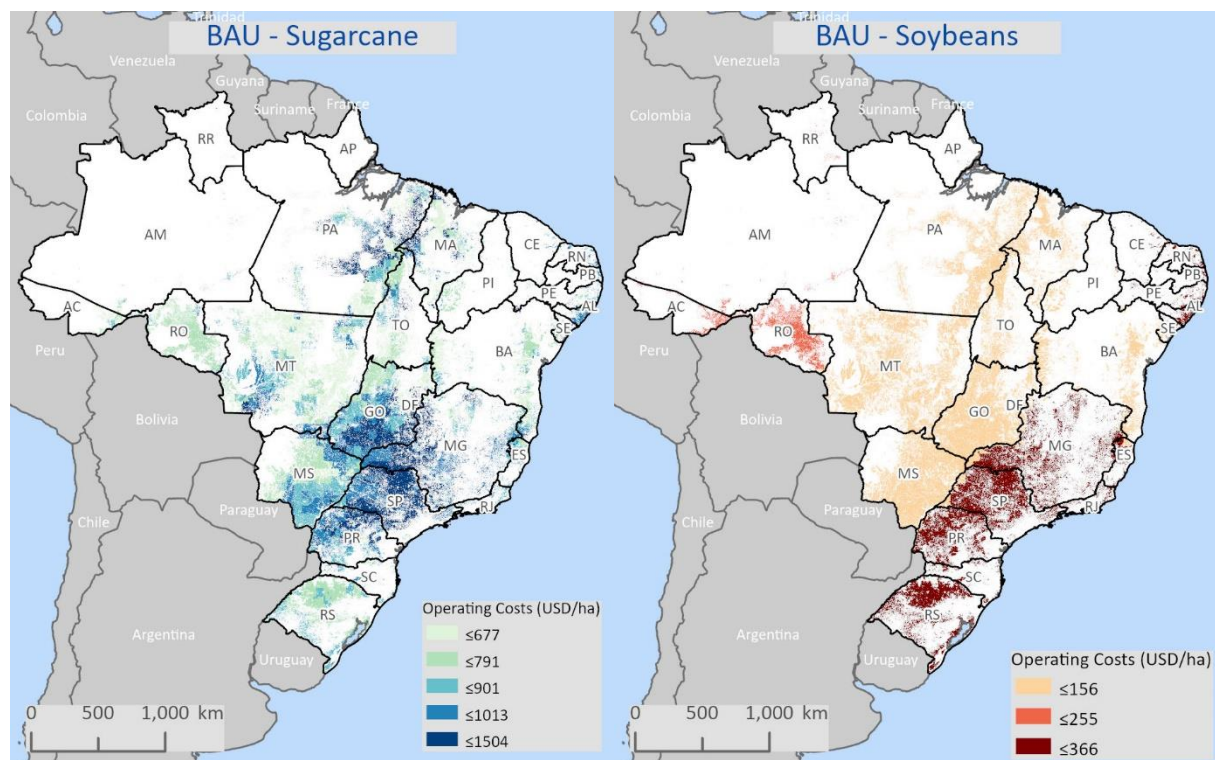
Source: PAM results (2020) – created by the author.

Operating costs for soybeans

In the BAU scenario, considering all regions in the country, the weighted average of operating
costs for soybeans is 241 USD/ha, with a variation from 156 to 366 USD/ha, depending on the
regional conditions – i.e., *expansion* or *traditional*. As clearly observed in Figure 31, the main
driver for differences in operating costs for soybean is the region as defined in the model (see
methodology). The results in this dissertation indicate that the *expansion* region has low op-
erating costs per ha compared with the *traditional* region, which could suggest that the re-
gional characteristics such as suitable field size and topography may be affecting the operating
costs.

Machinery plus diesel costs account for 67% of the total operating costs, with labor being the second most important cost component at 25% of the total operating costs. The remaining 8% are contractor costs. This composition also varies among regions, with farmers in the *traditional* region having proportionally higher machinery and labor costs per ha since the fixed component of these costs is allocated to fewer hectares.

Sugarcane

The operating costs for sugarcane, as shown in Figure 31, are significantly higher than soybeans, with a weighted average of 995 USD/ha. It also is interesting to observe the variation among regions, from 677 to 1,504 USD/ha, which is mainly the result of the correlation between yields and operating costs. Since most of the operations such as harvesting, loading and transportation are contractor fees based on fresh cane yield, high-yielding regions such as SP and PR have higher operating costs per ha. The key components of the operating costs are contractors, accounting for 41%, labor 33% and machinery plus diesel for the remaining 26%.

It is important to highlight that, unlike establishment costs, for which the largest portion of costs are cash expenditures that are highly visible to farmers, operating costs have a significant share of "hidden costs" such as depreciation and opportunity costs. That is more predominant for soybean farmers since the largest cost positions, such as machinery and labor, have a large share of these non-cash costs such as depreciation, opportunity cost of family labor and own capital. This difference may play an important role in farmers' cost perception as these hidden costs frequently are not accounted for during budgeting and the land allocation decision process. Conversely, sugarcane has the largest cost component as "contractors," meaning that operating costs are visible to farmers and most likely affect their resource allocation decisions.

Double Cropping

It is important to also consider the double-cropping alternative while analyzing operating costs. For establishment costs, it is possible to analyze second-season maize and soybeans separately, which is not entirely correct for operating costs. The desire to grow maize after soybeans motivates farmers to invest relatively more into machinery to ensure that soybean harvest followed by maize seeding maize are able to be performed in the necessary short window of time. Therefore, the BAU soybean scenario is likely to slightly overestimate the machinery setup required to only soybeans. This should not render major problems, however, since the double-crop combination is attractive in most regions.

While looking at the combination of summer soybeans and second-season maize on 60% of soybean area, the total operating costs are 363 USD/ha, with maize adding 122 USD/ha to the soybean operating costs. Even in the double-cropping scenario, the operating cost gap between sugarcane and grains is still significant.

Summary

Operating costs for sugarcane, on average, are 995 USD/ha, whereas double cropping (60%) has a total operating cost of 363 USD/ha. Aside from the absolute cost difference, sugarcane has a relatively higher share of cash expenses within operating costs (contractor fees), which is more visible to farmers than hidden costs (e.g., depreciation), which form most of the operating cost for grains. The double effect of higher and more visible costs is expected to play an important role in farmers' resource allocation decisions.

5.2.3.   Inbound transport cost for sugarcane

The last building block of total costs is the inbound transport cost for sugarcane. Different from soybean and maize, sugarcane must be transported from the farm to the factory for processing. Figure 32 shows the transport cost from farm to mill in USD/ha, comparing the "actual distance" from the SimU to an operating mill against the "current average"[32].

Figure 32:   Sugarcane transport from farm to mill (in USD/ha). The actual distance to an operating mill (left) or the average of 30 km of transport distance (right)
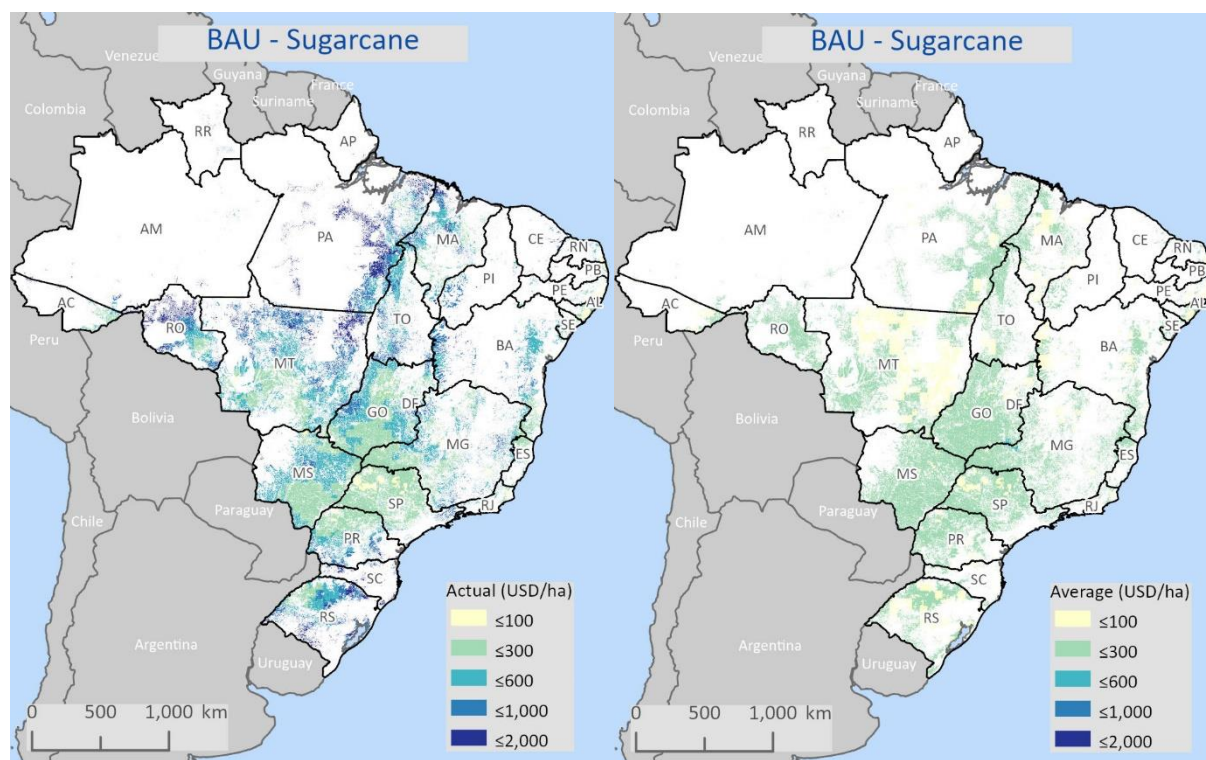
Source: PAM results (2020) – created by the author.

---

[32]   The average is the result of converting the average transport cost to a distance in km, using the cost curve from Françoso et al. (2017). The resulting cost of 2.76 USD/t of fresh cane applies only to SimU with distances greater than 30 km, avoiding artificially increasing the cost for SimU that already are close to an operating mill.

Figure 32 highlights the importance of carefully considering sugarcane transport from farm to mill. If the "actual distance" to an operating mill is considered to estimate the cost of inbound transport, the model results show an arithmetic average cost is ca. 700 USD/ha, highlighting the importance of this cost component in the total cost – i.e., 132% of the total establishment costs. The estimated inbound transport costs vary strongly from under 100 USD/ha for some areas in the state of São Paulo, where most of the mills are located, to regions with a cost close to 2,000 USD/ha, mainly in the north and south of Brazil. Hence, using the actual distance to an operating mill significantly penalizes sugarcane production in the regions far from São Paulo (SP) so the model is expected to allocate most of the remote areas to soybeans and maize. This, indeed, captures the short-term supply capabilities of certain regions because the absence of mills practically nullifies the alternative of growing sugarcane.

When calculating a weighted average based on current production, the inbound transport costs are 162 USD/ha, reflecting the current situation in which more than 95% of the production takes place within a 50-km radius from an existing mill. Even though the actual distance scenario depicts the current and short-term situation, it also is important to consider future development in milling capacity – i.e., construction of new factories. Therefore, the "average scenario" for inbound transport also is simulated.

The idea with the average scenario is to simulate future situations in which the transport cost from farm to the mill is based on the current average observed – i.e., maximum of 2.76 USD/t of fresh cane, which means a travel distance of 30 km from field to mill. The results from the "average scenario" are shown in Figure 32 (right), where the maximum observed cost is around 450 USD/ha, mainly driven by high yields.

Due to the strong impact of the inbound transport cost in the competitiveness of sugarcane, the BAU scenario uses the actual distance to an operating mill to reproduce the current situation. To illustrate the potential of sugarcane production in remote regions, a scenario is included using the current average transport costs while keeping everything else constant. The latter should illustrate the competitiveness of sugarcane if the constraint of having a nearby mill is not considered.

5.2.4.   Total production costs

Total production cost comprises establishment, operating, inbound transport for sugarcane and miscellaneous costs. As previously explained, due to the research question, land cost is not included in the total cost analysis as it should not influence the on-farm competitiveness of cropping alternatives.

Figure 33 shows the total cost for sugarcane and soybeans in USD/ha for the BAU scenario. There is a strong difference in total costs between these crops, with sugarcane being significantly more cost-intensive than soybeans. The extreme values of around 4,000 USD/ha

observed in Figure 33 are the result of the inbound transport cost in the dark blue area, where sugarcane would have to be transported more than 600 km to the nearest mill.

Figure 33:   Total costs for sugarcane (left) and soybeans (right) for the BAU scenario (in USD/ha)
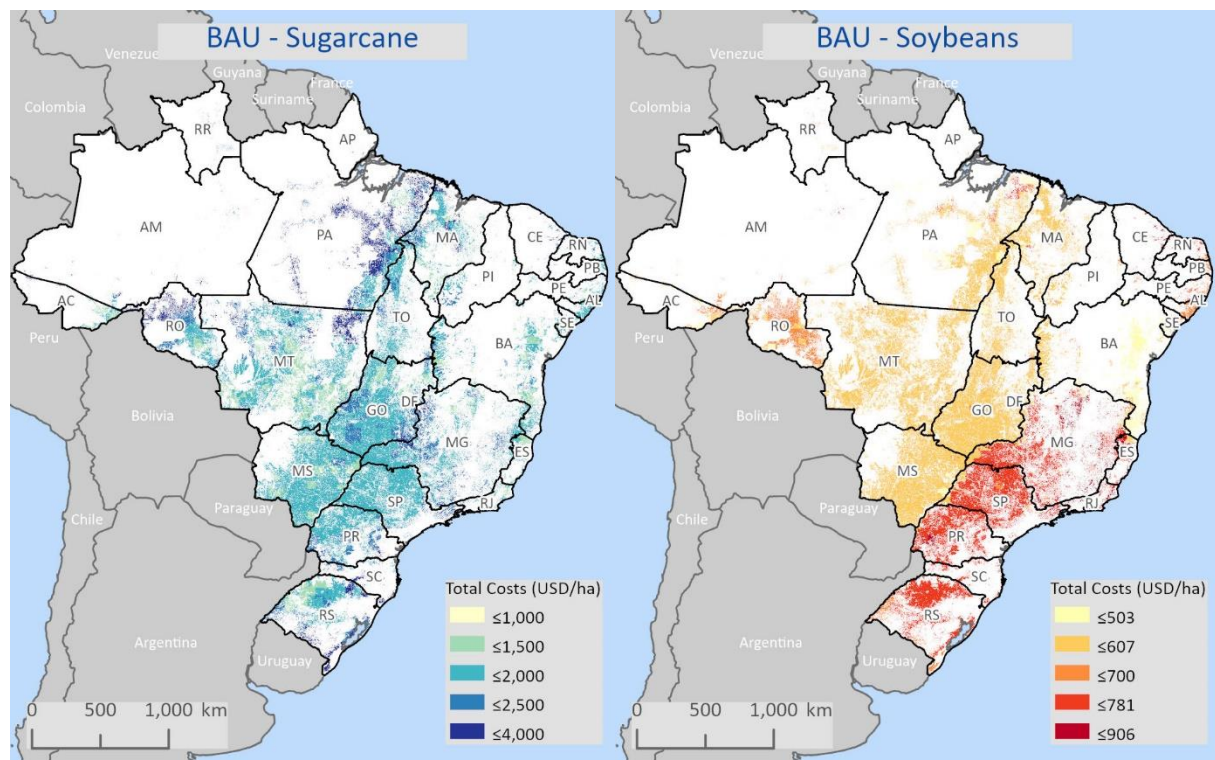
Source: PAM results (2020) – created by the author.

The total costs for soybeans have a strong correlation with the growing areas as defined in the model – *expansion* vs *traditional* – as that factor plays an important role in defining the operating cost component. The variation among the SimUs for soybeans is not as strong as for sugarcane, with the lowest values around 500 USD/ha and the highest 900 USD/ha, since inbound transport cost for soybeans is not required.

The national average total cost for growing sugarcane is 1,765 USD/ha with a variation from 1,480 USD/ha to 2,400 USD/ha. Compared with soybean and maize, sugarcane's total cost is almost three times higher on a per ha basis. However, it is important to keep in mind that in the most important producing regions, double cropping is common and adding the costs for soybean and maize provides a more realistic picture. Figure 34 presents the total production cost for sugarcane and double cropping with 60% of maize after soybeans in USD/ha on a weighted average for the 10 most important producing states.

Figure 34:   Total production costs for sugarcane and double cropping (soybeans + maize) for the BAU scenario (in USD/ha)
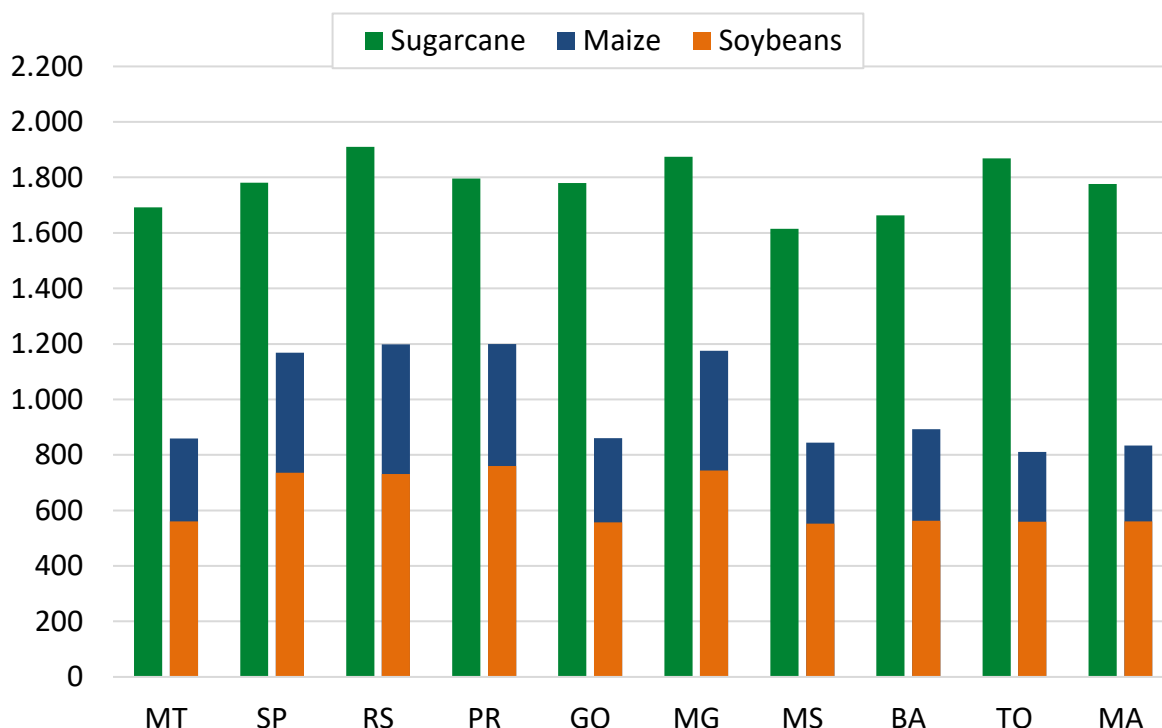
Compared with 60% double cropping (Figure 34), the total cost for sugarcane, on average, is 82% higher per ha. From a pure cost perspective, it is clear that some states (e.g., MT and GO) have a larger gap between the two alternatives than average, indicating cost competitiveness for double cropping over sugarcane.

Comparing the main soybean (MT) and sugarcane (SP) producing states, it is interesting that sugarcane total costs are relatively similar: SP's cost is only ca. 5% higher than MT's, whereas double cropping is 30% more expensive. That indicates cost competitiveness for sugarcane in SP compared with double cropping, with all other factors constant.

Table 4 presents the total production cost results for the BAU scenario as a national weighted average based on the harvested area for all crops and the double-cropping system, breaking the total cost into its main cost components.

Table 4:   BAU results from the PAM model as weighted average on the harvested area

Cost position[1]	Sugarcane		Soybeans		Maize (2nd)		Double Cropping[2]	
	USD/ha	%	USD/ha	%	USD/ha	%	USD/ha	%
Establishment costs	530	30	373	59	342	62	579	60
Operating costs	995	56	241	38	203	36	363	37
Inbound transport costs	162	9	-	-	-	-	-	-
Miscellaneous costs	79	4	19	3	11	2	26	3
Total costs: average	1,765	100	634	100	556	100	968	100

[1] Weighted averages based on the harvested area from IBGE (2019)

[2] 60% maize after soybeans

Source: PAM results (2020) – created by the author.

Besides the total cost analysis, it is important to consider the breakdown of the cost components to better understanding how changes in input items (e.g., nutrient, diesel, etc.) affect the cropping alternatives. While the most important cost component for sugarcane is operating costs, at 56%, the main driver of the total cost for double cropping of soybeans and maize is establishment costs, summing up to ca. 60%. Such analyses help us understand how changes in different cost components affect the on-farm competitiveness of the cropping alternatives. For example, changes in exchange rate directly affecting imported farm inputs such as fertilizer and plant protection products are expected to have a relatively stronger impact on soybeans and maize than on sugarcane since it directly affects their main cost component – i.e., establishment costs. Conversely, changes in capital and labor costs are likely to have a greater impact on the economic performance of sugarcane.

Summary

The results from production cost estimation within the BAU scenario can be summarized as follows:

–   Production costs vary strongly among the observed regions, which highlights the importance of micro-level economic modeling.

–   For sugarcane, the main cost component is operating costs accounting for 56% of the total costs, which are closely correlated to yields.

–   For soybeans, establishment costs are the main cost component and account for 60% of total costs (similar to double cropping).

–   Total costs for sugarcane are ca. 82% higher than a typical double-cropping system (60% of area grown with maize).

–   Accounting for the inbound transport costs for sugarcane from farm to mill is key to realistically represent short- and medium-term land-use changes since the lack of mills practically nullifies sugarcane as an option to farmers in some regions.

–   The ability to identify the driver of each cost item – i.e., price vs quantity - allows a better understanding of specific changes affecting cropping alternatives (e.g., exchange rate changes).

5.3.   Farm-gate prices

In addition to yields and production costs, a major component of crop economic performance is the price at the farm gate. Robust estimation of the price at microregional levels is important for a meaningful supply response tool since factors such as the distance to the port play a major role in the revenue formation at the regional level. In the PAM approach, the transport cost is modeled endogenously, accounting for the difference in the freight of each crop to be transported as well as the availability of roads to reach the ports. This powerful combination allows the estimation of farm-gate prices based on reference FOB prices for the analyzed crops. Figure 35 shows the producer prices for sugar (USD/t) and ethanol (USD/m³) at the SimU level.

Figure 35:   Producer prices for sugar (left) and ethanol (right) – in USD
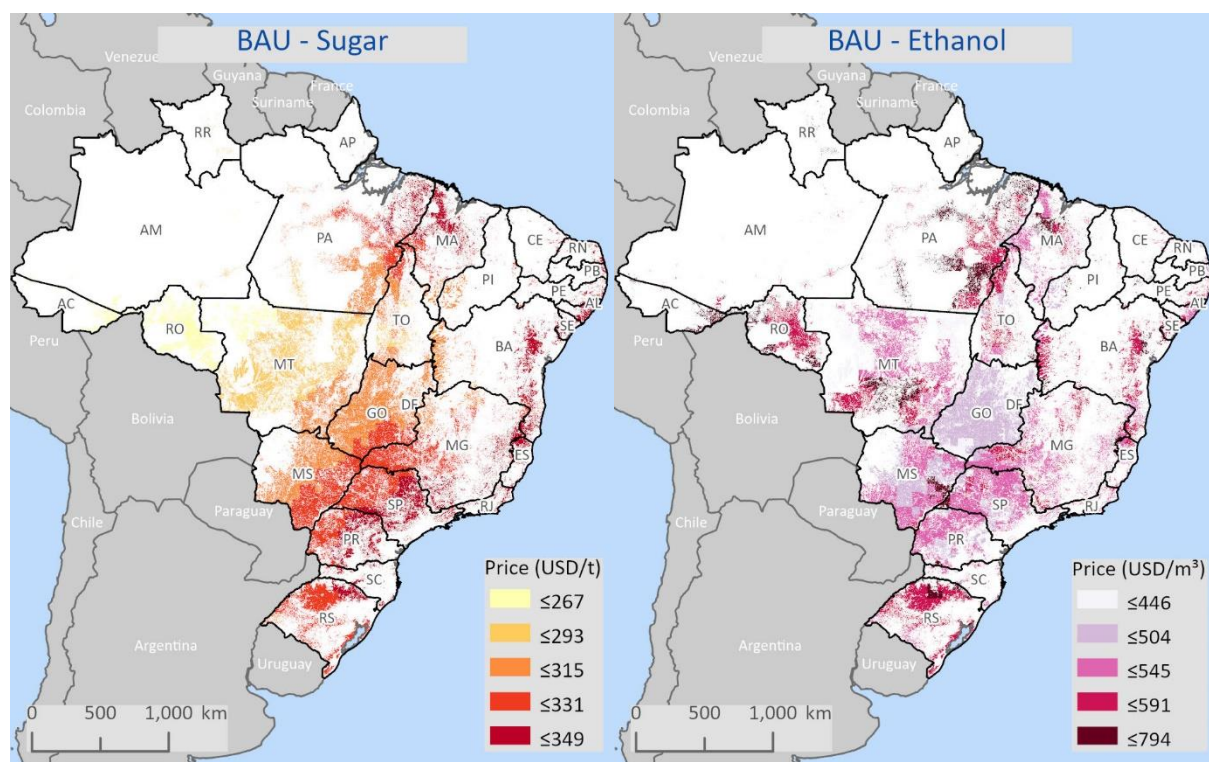
Source: PAM results (2020) and ANP (2020) – created by the author.

Sugar prices, shown in Figure 35 (left), illustrate the direct impact of transport cost on farm-gate prices, where producers in the state of MT face a discount of ca. 77 USD/t or 22% of the FOB price due to the travel distance of roughly 2,000 km to reach the port.

For ethanol, the price formation is different, with most of the production used domestically. Production regions farther from the port therefore can increase their mix of production toward more ethanol because, due to the higher cost of transporting ethanol to these regions, regional prices are attractive, as shown in Figure 35. This effect highlights the complexity of modeling farm-gate prices for sugarcane since the regional prices for sugar and ethanol, as well as the mix of production (share to ethanol and sugar), play significant roles in formulating the TRS price. Changes in framework conditions such as increasing oil prices affect sugarcane prices in different directions by increasing ethanol prices while reducing the price of sugar (i.e., increase in transport costs). This effect can be amplified by millers adjusting their production mix to more ethanol, resulting in greater sugarcane price variation across the country – e.g., price reduction for sugar-producing regions and price increase for remote ethanol-producing regions.

Figure 36 presents the estimated farm-gate prices on a state-level average as well as the reference price (e.g., FOB) for sugarcane (TRS basis), soybeans, and maize.

Figure 36:   Estimated farm gate prices on a state average and reference prices (BAU) for sugarcane (TRS basis), soybeans, and maize (in USD/t)
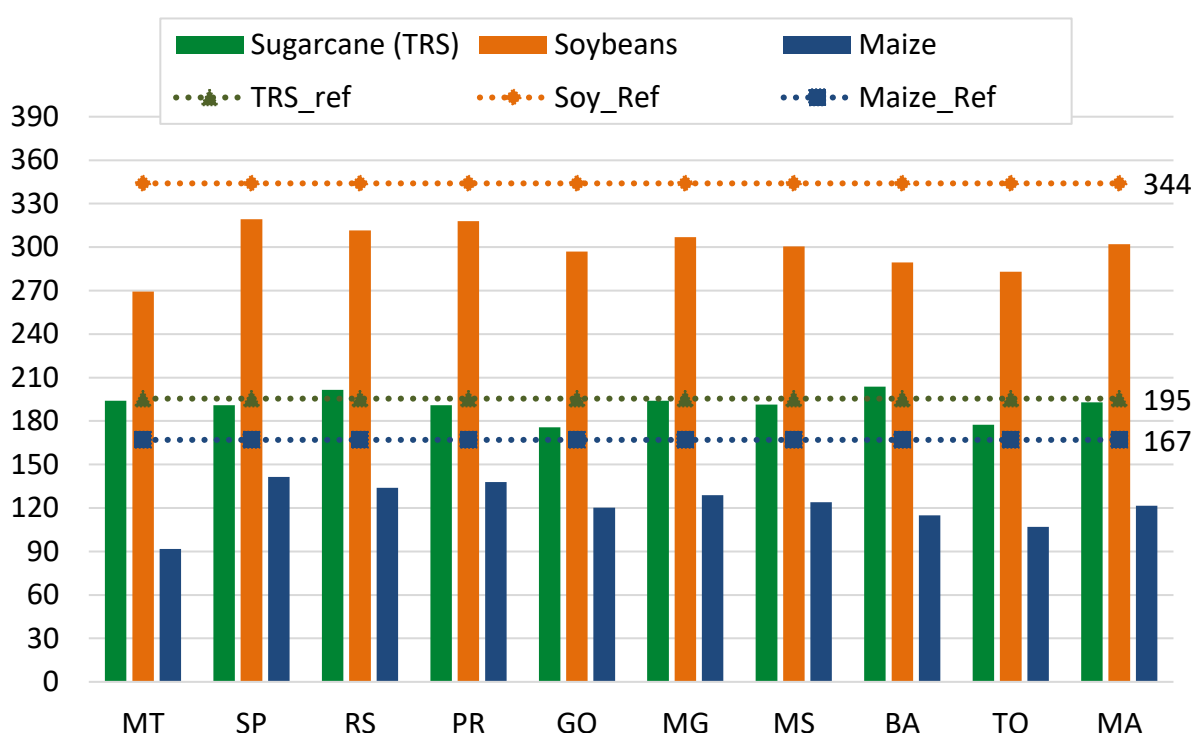
Note: TRS reference price (TRS_ref) is based on a production mix of 45% sugar and 55% ethanol. MT: Mato Grosso; SP: São Paulo; RS: Rio Grande do Sul; PR: Paraná; GO: Goiás; MG: Minas Gerais; MS: Mato Grosso do Sul; BA: Bahia; TO: Tocantins; and MA: Maranhão.

Source: PAM results (2020) – created by the author.

A closer look at the farm-gate prices reveals the importance of the transport cost and how much it affects the profitability of farmers. For soybeans, the national weighted average is 296 USD/t, meaning an average transport cost of 48 USD/t. However, states such as MT have

an average of 269 USD/ha, indicating an extra 30 USD/t discount compared with the national average. If compared with states like SP, the price gap is 50 USD/t, showing the spatial advantage of producers in SP. Conversely, the price for sugarcane based on the TRS value is almost identical between these two states, largely due to the domestic price of ethanol. That shows ethanol's importance in supporting farm-gate prices even in regions far from the ports.

Maize has the highest disadvantage from longer transport distance due to the relatively low product price in relation to the freight costs. For example, on average, producers in the state of MT face a discount of 75 USD/t due to transport costs, representing a discount of 45% from the FOB price for maize compared with only 21% for soybeans. The result is maize farm-gate prices below 100 USD/t.

Summary

It is known that transport costs have an impact on the overall competitiveness of the Brazilian agriculture sector. The findings in this dissertation add a different perspective regarding transport costs also affecting the on-farm competitiveness among crops. Remote regions face much lower farm-gate prices for grains, with discounts of more than 70 USD/t, which impact maize comparatively stronger since the absolute crop price is much lower than that of soybeans, resulting in a discount of 45% of the maize price. Due to the domestic demand for ethanol, on the other hand, remote regions have sugarcane prices similar to less remote areas, at least as long as the regional demand is sufficient to absorb the regional supply. Although sugarcane prices are similar in most states for the BAU scenario, changes in one of the output products alone – e.g., increased oil prices – are expected to affect regions differently since sugarcane millers can adjust their output mix accordingly.

5.4.   Return to land

The results from the yields, production costs, and farm-gate prices estimations form the basis to evaluate the competitiveness of cropping alternatives by generating an estimate for the return to land. This is a single profitability indicator, balancing out all factors such as environmental conditions (i.e., yields), location (i.e., transport costs and prices) and the production costs. Return to land for soybeans and maize grown as the second crop in the BAU scenario are presented in Figure 37, for the combined current arable and grassland.

Figure 37:   Return to land for soybeans (left) and maize grown as a second crop (right) for the BAU scenario (in USD/ha)
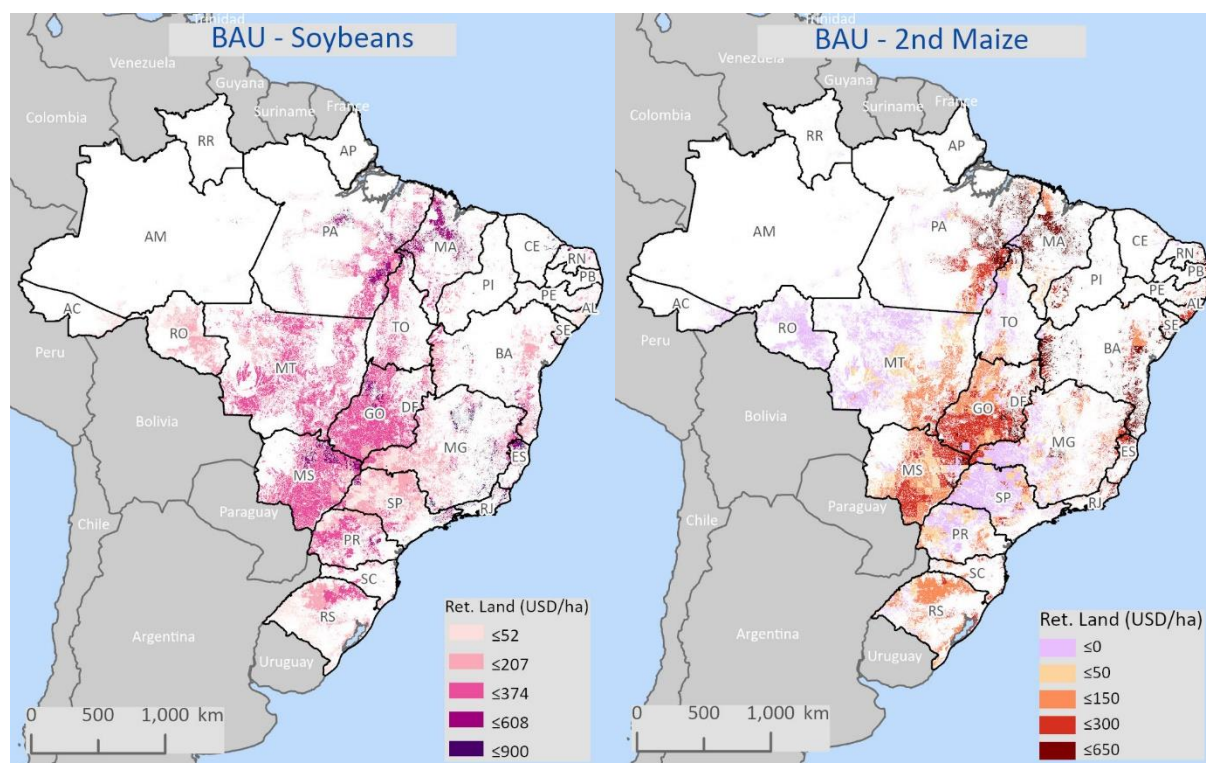
Source: PAM results (2020) – created by the author.

For soybeans, most SimUs have a positive return to land, with values as high as 600 USD/ha. Focusing on the main grain-producing states (e.g., MT), the variability of returns to land for soybeans is much lower, ranging from 117 to 438 USD/ha, with a weighted average of 261 USD/ha. Conversely, the returns to land from maize as a second crop have very different behavior. In addition to high regional variability, a share of SimUs in the important producing states such as MT have negative returns to land. That may be contrary to the expectation since MT alone produces roughly 50% of the total output of maize grown as a second crop (CONAB, 2021).

The decision to grow maize as a second crop is more complex than the simple evaluation of the economic performance. Farmers must produce a second crop to guarantee the efficiency of the no-till system – e.g., soil coverage. Therefore, farmers who cannot or opt not to grow maize as a second crop spend ca. 73 USD/ha[33] growing a cover crop instead. Besides the economic perspective, the option of growing maize as a second crop comes with important advantages such as usage of farm assets and labor as well as improving cash flows. The latter is

---

[33]   Based on the average of all *agri benchmark* typical farms in the state of MT. These are the cost of seeding and supporting operations to grow a cover crop after soybeans.

of key importance for farmers to reduce their use of credit, which is expensive in Brazil, and to finance ongoing expenses such as permanent farm staff.

Finally, it is important to consider that the *agri benchmark* methodology allocates fixed costs proportionally to all crops grown on the farm; hence maize also receives a significant share of family labor, capital costs, etc. However, since the alternative to maize is a cover crop with zero revenue, the fixed cost could be allocated entirely to soybeans. That does not change the aggregated picture seen in Figure 38 but is expected to change the pure return to land analysis for second-crop maize. To account for this complexity and to show the actual crop alternative for grain-producing farmers, returns to land for double cropping (soybeans as the main crop) as well as returns for sugarcane are presented in Figure 38.

Figure 38:   Return to land of sugarcane (left) and for double cropping (right) in the BAU scenario (in USD/ha)
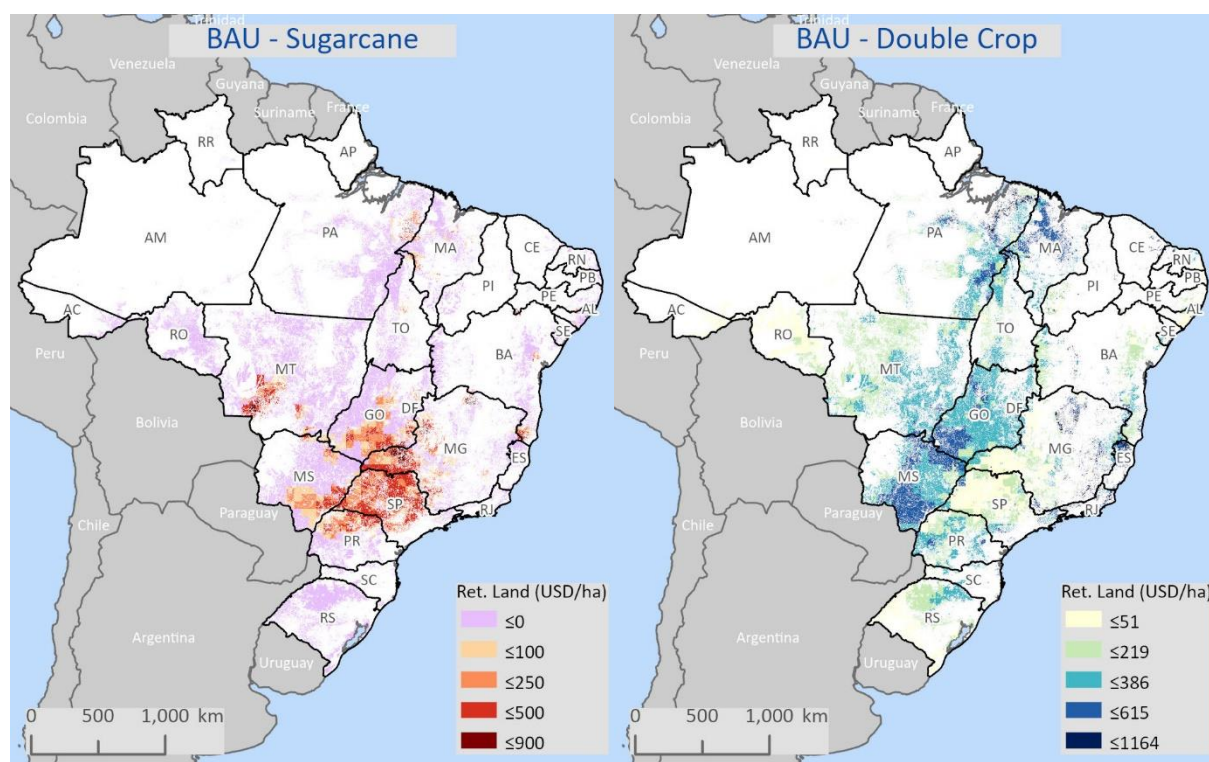
Source: PAM results (2020) – created by the author.

The returns to land the double cropping are mostly positive, with regions having returns higher than 1,000 USD/ha. The national weighted average is 248 USD/ha, indicating double cropping with maize increases the return to land for farmers, on average, by 14 USD/ha. Adding this small gain to the forgone cost of growing a cover crop results in an average net gain of 87 USD/ha, which partly explains why farmers choose to grow maize as a second crop if environmental conditions allow.

Even though the national weighted average of returns to land for sugarcane is 188 USD/ha, most SimU have a negative return to land in the BAU scenario (see Figure 38 – left). An important driver of the negative return is the inbound transport (i.e., farm to mill), which increases production costs of sugarcane far away from current mills. The SimUs with negative returns to land for sugarcane have an average estimated inbound transport cost to the nearest mill of 850 USD/ha, compared with only 250 USD/ha for the SimUs with positive returns to land. That illustrates how strongly the presence of a mill affects the short-term competitiveness of sugarcane and farmers' willingness to switch to this crop. That also reinforces the need to include such constraints in the supply response model to produce meaningful and realistic results.

The focus of this dissertation is the on-farm competitiveness driving supply responses from farmers and Figure 39 distills the average return to land for sugarcane and double cropping estimated in the BAU scenario, in USD/ha.

Figure 39:   Return to land for sugarcane and double cropping (soybeans + maize) as a state-level average for the BAU scenario (in USD/ha)
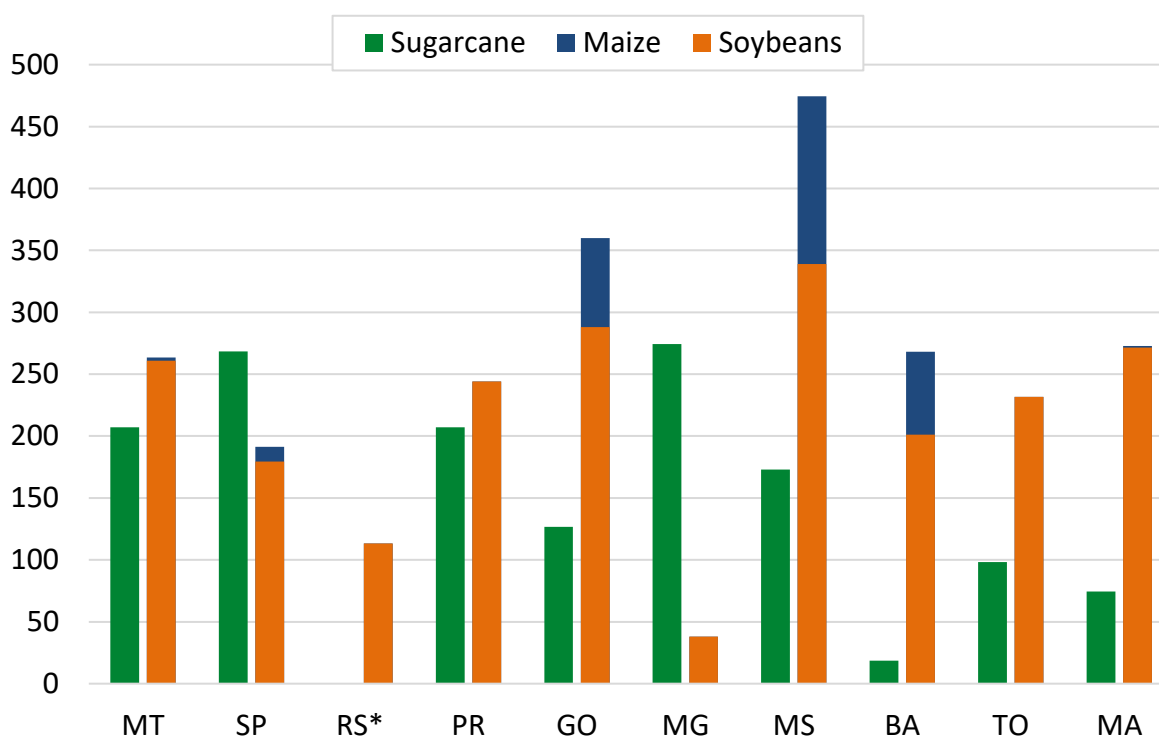
Note: * The return to land for sugarcane is -1,095 USD/ha due to its inbound transport costs. The data are not shown to facilitate the reading of the graphic.

Source: PAM results (2020) – created by the author.

The first interesting finding from the state-level analysis is that 8 of the 10 states with the largest share of acreage, on average, have a higher return to land from double cropping than from sugarcane. The exceptions are SP and MG. That, however, is in line with expectations

given since these two states alone produce more than 70% of the total sugarcane output in Brazil (CONAB, 2021). Secondly, maize contributes to an important share of the return to land for double cropping in the states of GO, MS and BA, while avoiding reductions in returns in the other states if compared with the alternative of a cover crop. Finally, it is interesting to observe that double cropping is the alternative with the highest returns in the states with the greatest expansion in crop production in the recent year (MAPITOBA), which is partly explained by the lack of mills in these regions.

Summary

Considering the national weighted averages, double cropping has 248 USD/ha and sugarcane has 188 USD/ha as return to land. Regionally, this picture is more diverse, with most states still having a higher return to land for double cropping than for sugarcane in the BAU scenario simulations. São Paulo (SP) and Minas Gerais (MG) are the exceptions, with sugarcane leading the economic performance. Maize grown as a second crop contributes an important share of the overall returns while also avoiding additional costs vs the alternative of growing a cover crop. Finally, it is important to highlight that the state-level average should not be overinterpreted because intraregional differences are observed.

5.5.   Land use

The final output from the PAM model is land allocation based on the cropping alternative with the highest return to land while considering regional constraints. Combining the information from all cropping alternatives and the land-use classification by IBGE (2018), the model outputs the best performing alternative at each SimU. It allows the theoretical estimation of land use based on the assumption that producers aim to maximize their return to land. Changes in the framework will impact factors such as yields, prices and costs, leading farmers to change their land allocation to other crops and inducing land-use change. That is the key idea behind regional supply response analysis, focusing not only on the overall crop output change but on possible resource reallocations and resulting land-use change.

As described in the return to land analysis, the transport of sugarcane from farm to mill is a key driver of negative economic results in regions where no processing infrastructure currently exists. To illustrate the importance of this constraint, land-use results are presented first for the BAU scenario and second, keeping all remaining assumptions constant and using the current average inbound transport cost for sugarcane (see methodology), illustrating its impact on the overall land-use results within the model.

5.5.1.   Business as usual (BAU)

The land allocation results from the BAU scenario for current arable land as well as the combination of arable and grassland are presented in Figure 40. It is important to note that besides double cropping (soybean and maize) and sugarcane, beef production is also included as a

possible alternative. In addition, keep in mind that other important supply chains such as oranges, coffee, cotton and rice are not considered in this version of the PAM model.

Figure 40:   **PAM land-use simulations based on the highest return to land for the selected farming alternatives, for arable plus grassland (left) and only arable land (right)**
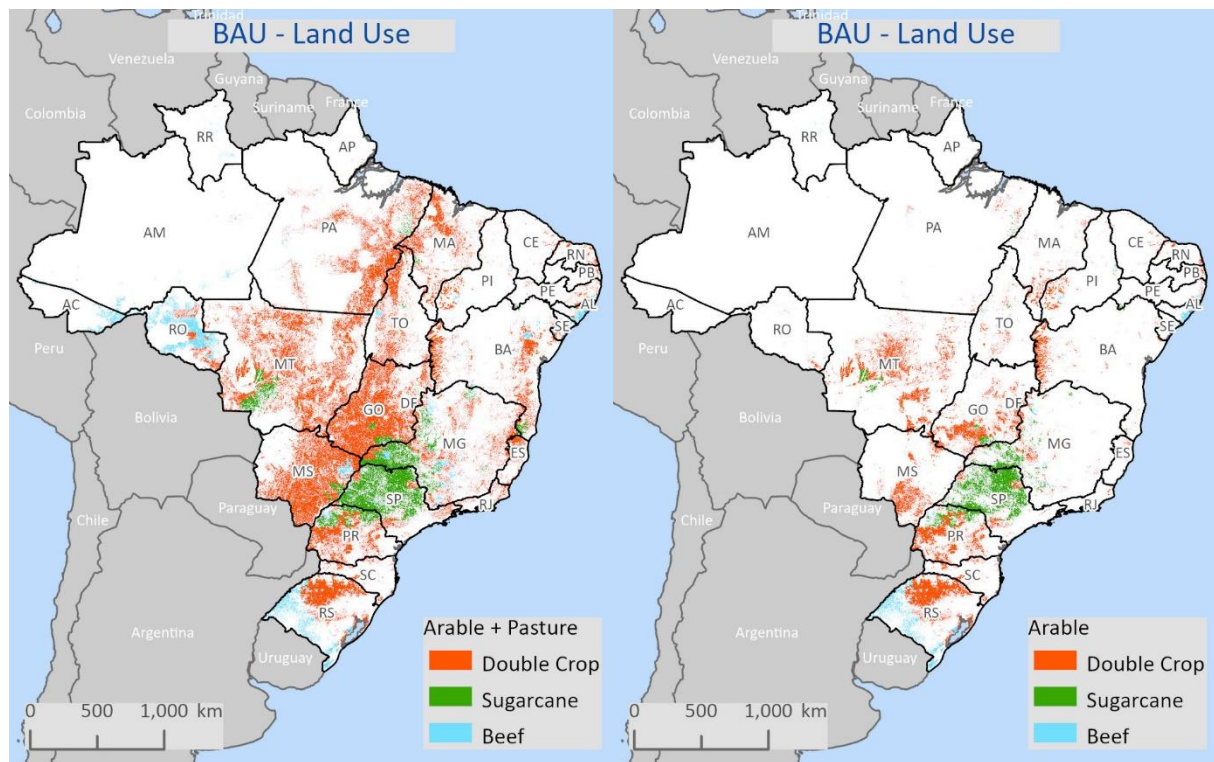
Source: PAM results (2020) and IBGE (2016) – created by the author.

The overall picture is interesting, showing a strong regionalization or clustering of production for the analyzed cropping alternatives. The main cluster producing sugarcane is centered in the state of SP, while the remaining area is predominately allocated to soybean and maize production (i.e., double cropping). This strong regionalization is an indication that a combination of factors such as yields, distance to ports, presence of milling infrastructure, and regional characteristics appear to motivate farmers to specialize in one production system within each region. To a certain extent, the methodology prevents a single SimU from having more than one best-performing crop but the lack of a strong mix of results within a certain region indicates that regional factors are impacting the economic performance of a certain crop and inducing strong regional specialization.

If the model is allowed to expand into the current grassland areas, the regional clustering does not change significantly, with most of the grassland allocated to double cropping in the

Center-west states such as MS, MT, GO and the MAPITO[34] region. Sugarcane would expand mainly to areas in MG and PR states that are neighboring the main cluster in SP.

This result does not necessarily mean that the potential for sugarcane is limited in all current grassland. A possible explanation is the lack of mills currently operating in these regions, significantly increasing the cost to produce sugarcane if it needs to be delivered to a factory in operation. This constrain is likely to be the most restrictive in this case, but it fits the idea of reproducing the current BAU scenario for land use. It is expected that if such an extensive expansion takes place, companies may be willing to invest in new processing facilities, but it is not feasible to assume such developments in the very short term (see section 4.4.2).

The allocation to beef production to only a few areas in the far south (Rio Grande do Sul - RS) and far west (Rondônia - RO) indicates that the observed returns to land for beef are not attractive for the main share of the grassland surrounding the clusters producing grains and sugarcane. That is an interesting finding, indicating that, at least based solely on economic performance, expanding arable production into grassland seems to be profitable. However, several factors may limit or reduce the attractiveness of such an expansion in the medium-term, for example:

- several grasslands are protected by natural parks and the forest code legislation,

- the model does not include the conversion cost and initial investment to convert grassland to arable land,

- there is expected to be a lack of infrastructure such as service providers, traders and roads to allow grain production, which is not included in the model,

- a large-scale expansion needs to be accompanied by a strong increase in international demand; otherwise, the price would sharply decline, and the estimated economic performance could not be achieved.

In the long-term, such limitations may not significantly restrict the expansion into grassland. Hence, strong shocks in demand may induce the expansion of arable land into grassland in the medium- to long-term, which has been observed in recent years (Barona et al., 2010; Macedo et al., 2012).

Without adjusting the yield derived by EPIC, the PAM model allocates most of the arable and grassland in Brazil to double cropping of soybeans due to the EPIC behavior of overestimating soybean yields while underestimating those of sugarcane (see appendix 7).

Finally, beef production performs better than soybean double cropping and sugarcane in the state of RS, even on the current arable land. That indicates that the model may not consider

---

[34] Refers to the states of Maranhão (MA), Piaui (PI), and Tocantins (TO).

regional factors such as local demand for grains based on livestock production, which creates a different pricing structure of the output, or other crops such as wheat and rice that may play an important role in this region. The evaluation of the model's performance in the allocation of land to different cropping alternatives is described in chapter 6.

5.5.2.   BAU with additional cane mills

The impact of the current mill infrastructure on the sugarcane returns to land is expected to be significant. To test this hypothesis, the BAU scenario is calculated considering the current average of inbound transport of sugarcane to mills. Results from the PAM land-use simulations with "additional cane mills" are presented in Figure 41.

Figure 41:   PAM land-use results for the BAU with additional cane mills (left) and difference from BAU for sugarcane (right), current arable and grassland
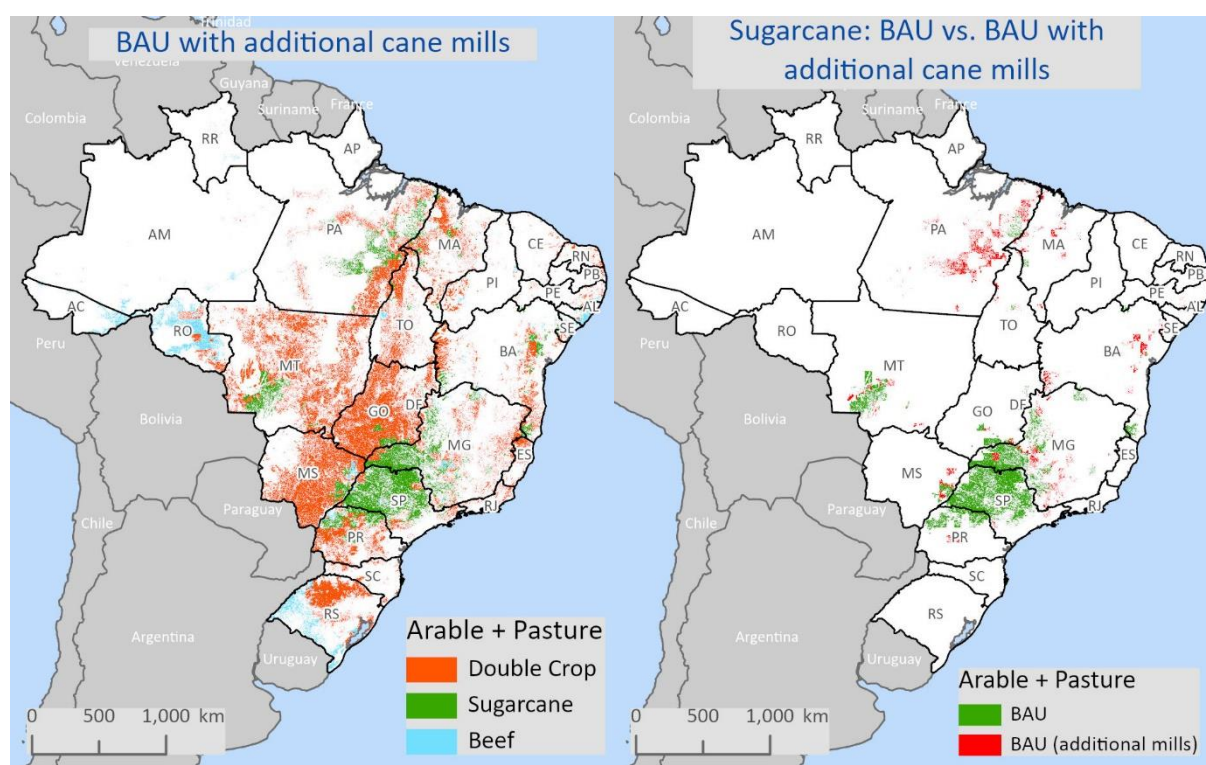
Source: PAM results (2020) – created by the author.

Removing the constraint of having to deliver cane to an existing mill (BAU) increases the return to land for sugarcane in 90% of the SimUs. That leads to an overall change on 7% of total acreage including arable and pasture – i.e., ca. 11.5 out of 175.5 million ha. The major grain-producing states such as MT, MS and Santa Catarina (SC) still have a higher return to land for double cropping, resulting in a similar outcome between the two scenarios for the CS region. This finding supports the idea that this region has a competitive framework for grain-producing due to a combination of yield ratios, farming cost structure and crop output prices.

Surprising is that 65% of the land allocation change to sugarcane is found in the states of Pará (PA), Maranhão (MA), and Bahia (BA), which jointly account only 2% of the national production (CONAB, 2021), which in turn indicates that, at least on the farm level, sugarcane appears to be competitive against beef and double cropping of soybeans and maize. The lack of sugarcane production in these areas may be driven by:

- lack of investments: there is minimum acreage required to justify setting up new mills,

- risk perception or other constraints to enter the market, such as lack of expertise, or

- the presence of more profitable agricultural alternatives not considered in the PAM model.

Summary

The land allocation results from the PAM model for the BAU scenario show a strong regionalization of production, with sugarcane having the highest returns around the state of SP, while double cropping is the dominant alternative elsewhere. Allowing the expansion into the current grassland does not change the picture dramatically, highlighting the hypothesis of regional characteristics affecting the profitability of each cropping alternative. Removing the constraint of having to deliver to an existing mill leads to an overall change in 7% of the total arable and grassland with 65% of the total changes occurring in the Northern region. That is an interesting finding that indicates a competitive edge for sugarcane in these areas.

6.   Model evaluation

The main goal of this dissertation is to develop a modeling concept to estimate farmers' supply responses and land allocation for countries with low data availability. A key step within this development is the evaluation of the model results to better understand its accuracy, possible skewness and the main areas for further development. In this chapter, I focus on the evaluation of the PAM modeling approach to answer questions on farmers' supply responses and the consequent land-use change, thereby informing decision-makers. According to Gardner and Urban (2003), the substantial increase in efforts to develop new models has not been met by the same investment in defining methods for model evaluation, even though it is as important as model development itself.

In the following subsections, the focus is on model evaluation as defined by Richter, Atzberger, Hank, and Mauser (2012), in which model performance is benchmarked against realistic and independent data. After the evaluation design is explained, the results from the evaluation of the yield and production costs estimations are presented. The final step within the evaluation is to compare the land allocation to individual cropping alternatives (i.e., land-use maps) to observed data from an independent source, illustrating the performance of the PAM model in depicting current land use.

It is important to keep in mind that the entire approach is based on a minimum data (MD) requirement and that an error margin is expected due to the complexity involved in estimating farm-level production costs and economic behavior. Any model is a simplification of reality and several factors affecting farmers' behavior, such as risk perception, management capability and know-how, are not included in the model (Gardner & Urban, 2003).

6.1.   Evaluation design

The evaluation process aims to determine the robustness of the concept and modeling approach, focusing mainly on validating the main innovation within the PAM approach – the estimation of production costs and returns to land. Nonetheless, the evaluation process considers the results of the final land allocation module as well as the performance of EPIC in providing reliable yield estimations.

The selection of the statistical measure used in the model evaluation is of key importance as discussed by Pontius and Schneider (2001). The idea of visually comparing the results from the land allocation module to observed land-use seems a straightforward approach for evaluation. However, basing the evaluation entirely on the graphical comparison may be misleading since results that "look good" may not depict the dynamics that the model intends to represent (Gardner & Urban, 2003). Therefore, a combination of statistical measures (quantitative) and graphical comparison (qualitative) is used for the model evaluation process.

6.1.1.   Statistical measures

The following three main statistical measures are used to evaluate the results of the several building blocks of the PAM model, following an approach similar to that proposed by Balkovič et al. (2013) and Nendel et al. (2011):

Mean Error (ME) =
$$\frac{1}{n} \sum_{i=1}^{n} y_i - x_i \qquad (34)$$

Mean Absolute Error (MAE) =
$$\frac{1}{n} \sum_{i=1}^{n} |y_i - x_i| \qquad (35)$$

Root Mean Square Error (RMSE) =
$$\sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2} \qquad (36)$$

where y_i is the predicted value by the model and x_i is the observed value for i = 1, 2, …, n SimUs or regions depending on the dataset[35]. The statistic measures also are presented in a normalized manner by dividing the measure by the mean of the observations. These relative measures are denoted by r as in rMAE. Even though ME is not a robust error measurement since errors with different signs cancel each other, it may be useful to identify systematic error' patterns of the estimation; for example, if the model systematically over- or underestimates yields of a certain crop.

When it comes to assessing the average model performance, researchers have argued about the benefits and drawbacks of using the two most common indicators – namely RMSE and MAE (Chai & Draxler, 2014). While the MAE gives each error the same weight, RMSE penalizes strongly larger absolute errors. Another important characteristic of RMSE is the difficulty to compare values coming from different sample sizes since it affects the magnitude of the RMSE (Willmott & Matsuura, 2005). Given the calibration of the main PAM modules – i.e., plant and economic, is based mainly on the most important production regions in the country, it is expected that marginal areas will have a higher estimation error. Due to the RMSE characteristic to overpenalize these "outliers" as well as due to the different sample sizes for each crop for the validation of production costs, the main statistical measure used in the evaluation of the PAM results is the MAE. Nonetheless, the RMSE and standard deviation (SD) also are presented.

---

[35] n is SimUs for the yield validation; for production cost it is the regions from CEPEA/PECEGE; and for land allocation it is the pixels in the land use maps.

6.1.2.   Yield and production cost evaluation design

Yield validation design

The yield estimation for sugarcane, soybeans and maize as the second crop is evaluated by comparing the EPIC results to official statistical data coming from IBGE. The official harvested area and production volume are allocated to each SimU following the work from Câmara et al. (2015), from which the crop yield for each SimU is calculated. The average yield data contains the three most recent years available – 2013 to 2015.

Production cost validation

A similar approach using regional data for the yield validation would be ideal to evaluate the production cost estimations within the PAM model. Unfortunately, such detailed and spatially explicit information currently is not available. An alternative may be to collect primary production cost data in several regions to benchmark the observed data to the model results. However, such an approach is time-consuming and cost-intensive, making it beyond the scope of this dissertation. Against this background, I propose the alternative of using data from two research institutions in Brazil – namely CEPEA[36] and PECEGE[37]. They collect a large sample of typical farms' data in the main producing regions in Brazil, following a SOP similar to that of *agri benchmark* (see section 2.6)*. This information is not publicly available, but it was obtained through a cooperation agreement for the evaluation process.

The production cost data for each farm is used as an average from the seasons 2016 to 2018, originally collected in Brazilian Reais and converted to USD with the same exchange rates used in the PAM model. Figure 42 shows the typical farms from PECEGE in the main sugarcane producing regions in Brazil as well as the methodological approach in the GIS software to allocate municipality level information (PECEGE) to SimU levels (PAM results). In a nutshell, the two information layers are spatially joined and a simple average of all SimUs within the boundaries of each municipality is calculated and then compared with the observed data.

---

[36] Centro de Estudos Avançados em Economia Aplicada (CEPEA) - www.cepea.esalq.usp.br

[37] Programa de Educação Continuada em Economia e Gestão de Empresas (PECEGE) - https://pecege.com/

Figure 42:   Typical sugarcane farms for validation (left) and schema of allocation of Si-mUs to municipality level (right)
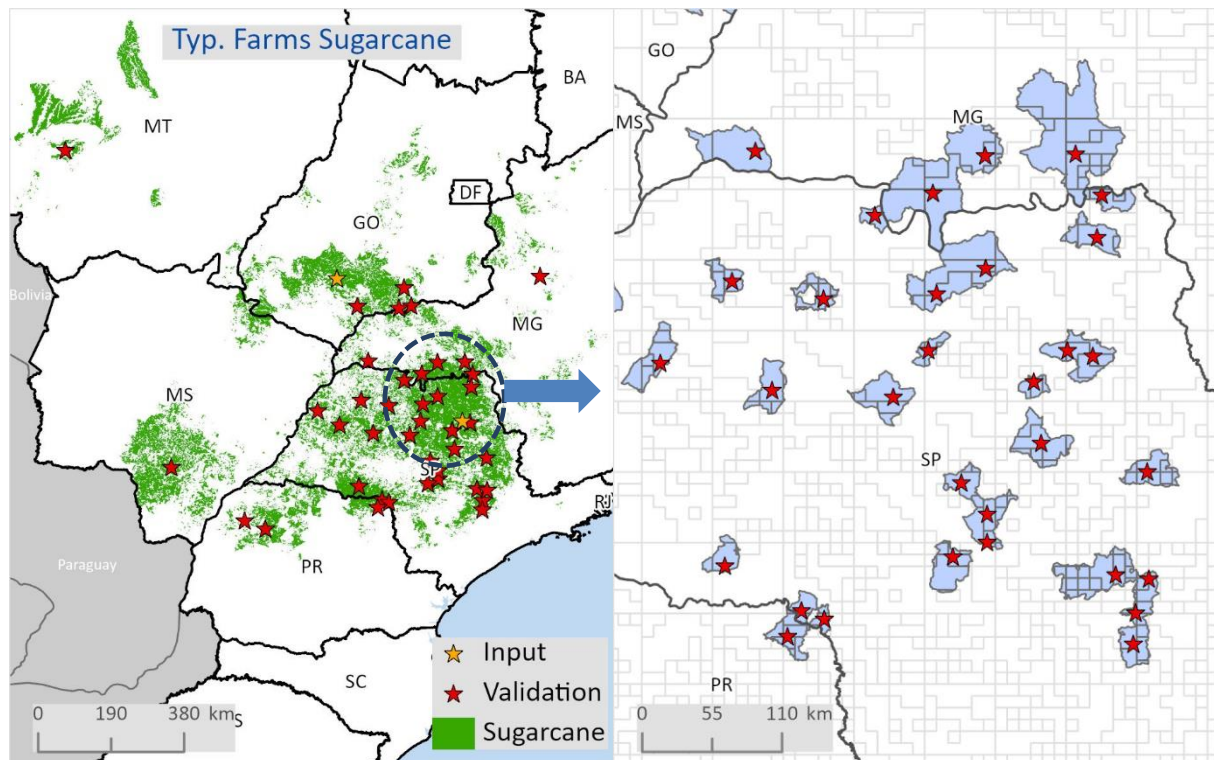
Source: IBGE (2016) – created by the author.

It is important to note that the sample size and distribution of typical farms from PECEGE (Figure 42) aim to represent the most important production regions in Brazil so that the majority of 45 farms are found in the main producing states of São Paulo (SP) and Minas Gerais (MG). The relatively small number of typical farms in other states is expected to decrease the validation robustness, because a single particularity coming from a focus group may cause strong regional deviation (e.g., duration of sugarcane ratoon, seedcane input, etc.). A similar distribution of farms is observed in the data coming from CEPEA for grains.

Although the data collection method of PECEGE and CEPEA is similar to that of *agri benchmark* used in the PAM model, differences are expected, such as the aggregation methodology from individual cost items, allocation of overhead, computing depreciation, etc. To minimize such distortions, the total costs without land is the indicator used in the evaluation of the PAM performance in estimating production costs. In this way, in all three methodologies, costs that may be allocated to different subgroups (e.g., operating vs overhead costs) are certainly included in the total costs.

One important uncertainty that is not ruled out by using total costs as the main indicator for the evaluation is the allocation of non-crop-specific costs such as overhead and family labor (i.e., management) to the individual crops within the farm since the allocation procedure may differ among the research groups. Even though this may cause divergencies in the single-crop

comparison, it should not have a strong impact when comparing the actual crop alternative (i.e., double cropping) against sugarcane since costs from soybeans and maize are combined. For sugarcane, the different allocation methodologies for overhead are not expected to impact the evaluation significantly, because the typical sugarcane farm does not grow any other commercial crop.

For the production cost evaluation, it is important to consider at which unit the costs should be benchmarked – per t of output or per ha. For grain farmers, the logical approach seems to be using cost per ha since the total costs are not closely related to crop output (see 4.5.4.2). However, for sugarcane, important cost items such as harvesting, loading and transportation are linked to yields (i.e., contractors' fees in USD/t). That is important because the yield data from PECEGE come from the focus group, which is likely to represent a combination of experiences from the participants of the panel, agreeing on the most typical yield observed in the region for the season. That is expected to be influenced by the exact location of farmers forming the panel and is not necessarily representative of the environmental conditions of the region. Hence, the validation for sugarcane uses both units to better understand whether the differences come from yield or production cost information.

6.1.3.   Land allocation evaluation design

The final output from the PAM model is the allocation of the cropping alternative with the highest return to land to each SimU. This output is a map based on the geographical location of each SimU and the spatial constraints (i.e., current land use). For the validation, it is desirable to compare the output from the PAM model to the recent real land-cover map, which has not been used for the calibration (Pontius & Schneider, 2001).

Land-cover maps classifying arable and grassland are readily available (IBGE, 2018), but they usually lack information on which crop is grown in each pixel. Currently, such detailed land use maps of observed data, including the crop information per pixel, are not publicly available for the whole country. I propose two alternatives to create validation maps:

- using statistical data at the municipality level from IBGE, and

- results from the Spatial Production Allocation Model (SPAM) from IFPRI.

The following subsections highlight how these two simplified land-use maps were developed for the validation of the PAM model.

Official statistics data at municipality level (IBGE)

Since 1974, the Brazilian Institute of Geography and Statistics (IBGE) annually collects detailed information for the total and harvested area, production volume, yields and production value of all crops grown in Brazil. This information is at the municipality level and is, to the best of my knowledge, the statistical data with the highest spatial resolution publicly available in

Brazil (IBGE, 2019). The information for more than 5,500 municipalities is used to create the validation map based on the IBGE data.

Using GIS software, the information from IBGE's survey is allocated to the total geographic area of each municipality. The resulting layer is spatially combined to the high resolution (pixel of ca. 100 ha each) land classification map from IBGE (2018) as shown in Figure 43 (right), which provides the information on the actual grassland and arable land within the municipality boundaries. The allocation of the results from the PAM model's BAU scenario to the grass and arable land is similar. Hence, I make sure that all information used throughout the land allocation validation is at the same spatial resolution (ca. 100 ha pixels).

Figure 43: **Spatial resolution of different data sources: output from PAM (left), municipality level data from IBGE (middle) and the land-use map from IBGE (right)**
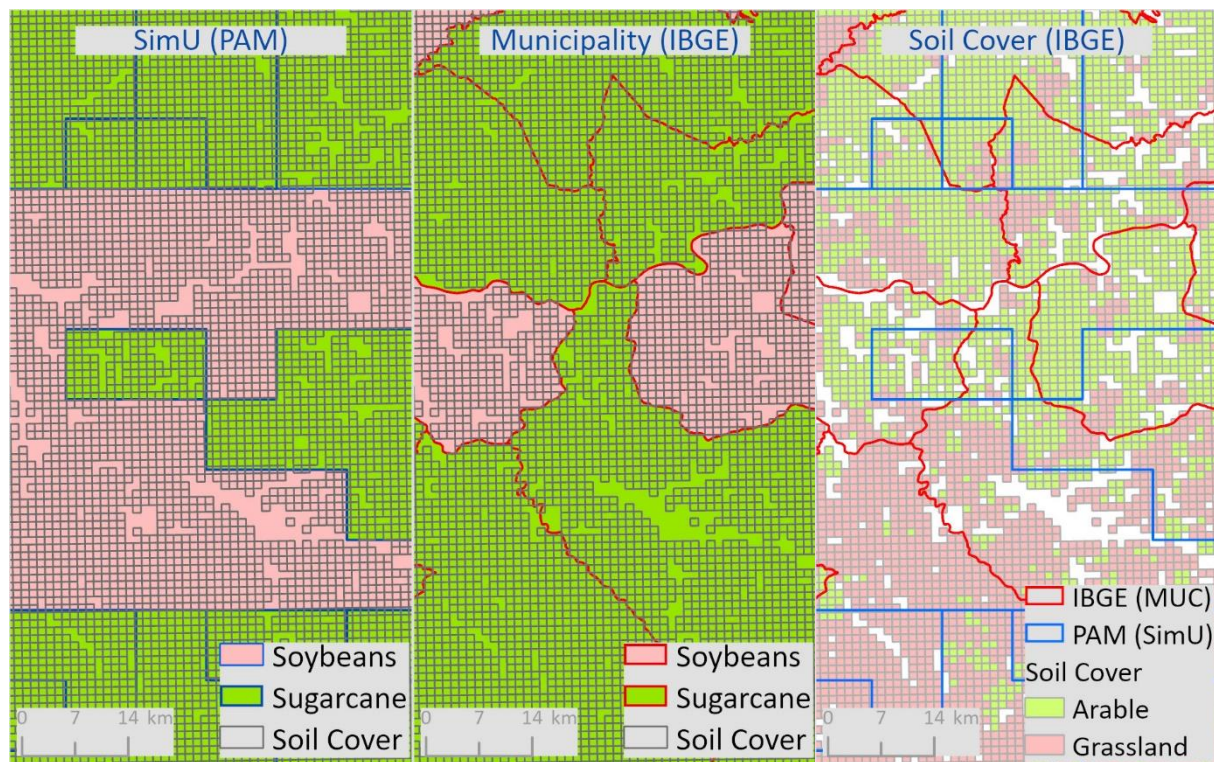
Source: IBGE (2018 and 2019) – created by the author.

The final validation map with IBGE information has more than one crop grown per pixel since several municipalities have multiple crops grown on their arable land – not necessarily by the same farms. Since the optimization module within the PAM model allocates the whole SimU to a single crop alternative – e.g., either sugarcane or double cropping – a similar adjustment is needed for the IBGE data to allow the validation to be performed. Therefore, the crop with the highest acreage (share) in each municipality receives the entire arable land in it.

One may argue that it would be feasible to allocate the arable land within the municipality proportionally to the share of area for each crop from the IBGE data. That is technically

possible but without further information on where to allocate this crop (e.g., satellite imagery), it only can be allocated randomly. That creates biased precision and could be manipulated to fit the PAM outcome, for example. Therefore, the rule of allocating the whole acreage for the most important crop is expected to create an oversimplification of production systems for some regions but it is likely to produce more robust results compared with a random allocation.

This limitation is expected to impact mainly the states of Goiás (GO), Mato Grosso do Sul (MS), and Minas Gerais (MG), which have large municipalities producing both sugarcane and grains. Nonetheless, the IBGE database is highly detailed and contains real observed information, which increases its value significantly for the model evaluation process.

Spatial Production and Allocation Model (SPAM)

To broaden the scope of the land allocation validation, data from SPAM developed by the International Food Policy Research Institute (IFPRI) is used to create a simplified land-use map. The SPAM model uses a variety of inputs ranging from national statistics to biophysical crop suitability assessments, allocating all the information to a detailed spatial grid of 10 by 10 km (IFPRI, 2019).

The simplified land-use map for Brazil is based on the physical area of soybeans and sugarcane from the SPAM 2010 dataset (most recent available). To make the resulting map comparable to the output from the PAM model, the pixel information from the SPAM model in the simplified map is allocated to either sugarcane or soybeans depending on the crop with the highest share of acreage within the pixel (like the IBGE data). This simplification is necessary since PAM allows only one cropping alternative per pixel. The resulting map is spatially combined with the land classification map from IBGE (2018) for arable land, using the same method as for the information from IBGE. The resulting validation map based on the SPAM data for 2010 is presented in Figure 44.

Figure 44:   Simplified land-use map based on SPAM results for land allocation validation – the whole country (left) and main production areas (right)
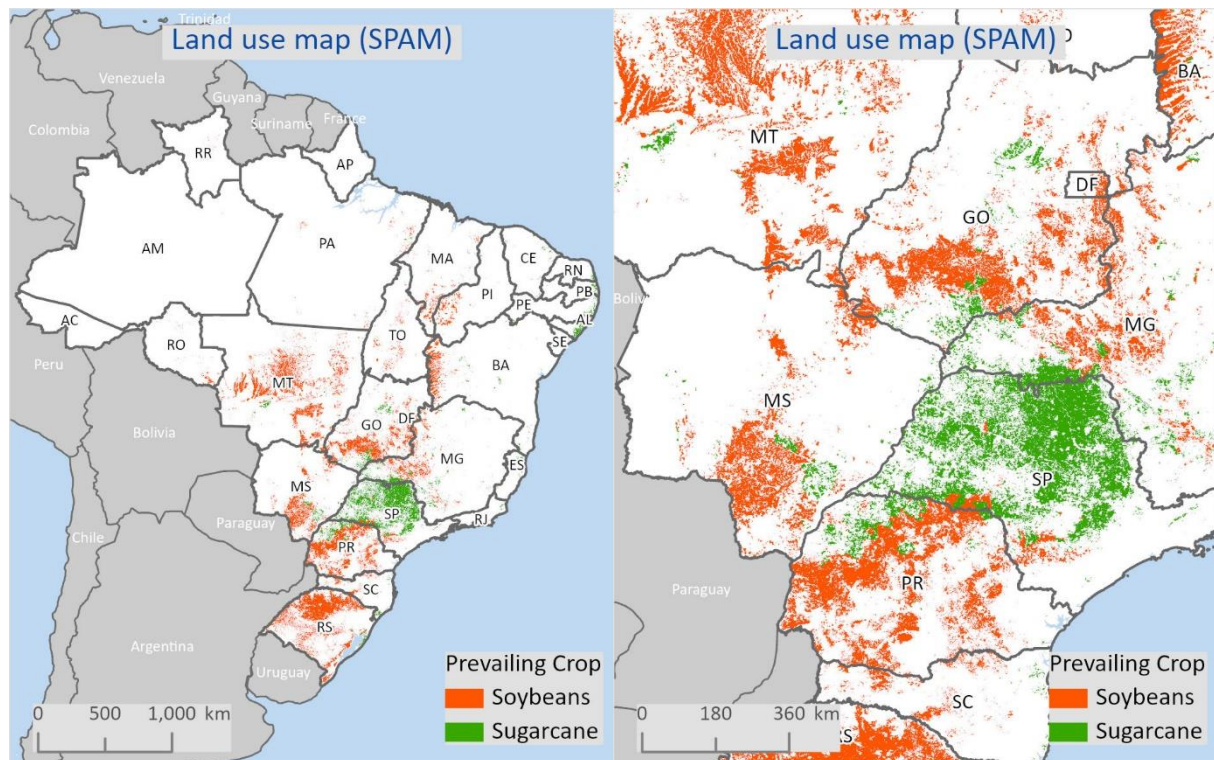
Source: SPAM (2010) – created by the author.

A map similar to shown in Figure 44 is obtained with the information from IBGE (municipality level). However, due to the spatial characteristics of the SPAM model, the level of detail in the resulting validation map is considerably higher since the drawback of using administrative-level information can be avoided. The detailed map (Figure 44 – right) already shows the regions where differences in spatial resolution may play a bigger role in designing and evaluating land-use models. The states of MS, MG, PR and GO are more diversified, with a regional mix of soybeans and sugarcane production, compared with the remaining states, so the relatively coarse spatial resolutions may not be able to accurately portray the variations observed in the land use in these regions.

The performance indicators for the land use allocation

Having the three maps at the same spatial resolution – i.e., PAM (BAU), IBGE and SPAM for the current arable land in Brazil – I propose a three-step method to assess the performance of the PAM model in representing the observed land use for sugarcane and the double-crop of soybeans and maize. This method should allow a qualitative and quantitative assessment of the model's performance.

The procedural steps can be described as follows:

(1) **Graphical comparison (qualitative):** Visually comparing the arable land that the PAM model allocates to each crop with the simplified land-use validation maps obtained from IBGE and SPAM.

(2) **Matching score (quantitative):** Comparing the results from the PAM model to the observed (IBGE) and estimated (SPAM) information for 640,000 pixels in the final validation map. The performance indicator is a percentage value based on the matching pixels between the estimation and observed – e.g., PAM allocates sugarcane to a pixel where the prevailing crop in the validation maps also is sugarcane.

(3) **Performance visualization (qualitative):** Creating a matching map with the current arable land highlighting the pixels where the PAM model does or does not provide results that match the validation maps.

The combination of qualitative and quantitative components is crucial to reduce biased results based solely on visual comparisons (output maps) while allowing the identification of possible regions where the model underperforms (matching maps). It is important to check, for example, whether most of the non-matching pixels are in the same region due to the need to include more crops or regional-specific characteristics or are evenly distributed across the country. Such systemic error may affect substantially the interpretation of the validation results and therefore needs to be identified correctly.

6.2.   Evaluation results

The PAM modeling approach should help us understand farmers' supply responses and their consequence on land-use changes in regions where data availability is limited. Therefore, a thorough modeling evaluation scheme is required to identify possible limitations of the model for policy advisory. The results from the validation process are presented in this subchapter.

To better understand the performance of the main building blocks of the PAM model, the following sections present the validation results for the yield estimation, production costs and land allocation. Hence, I am able to identify where the main sources of uncertainty originate and how to address them properly while interpreting the results. Finally, the validation process helps identifying steps for the further development of the modeling approach.

6.2.1.   Yields

The idea of combining a biophysical model and a production cost data is a main part of the entire PAM modeling approach since yields play a crucial role in the competitiveness of each cropping alternative at the farm level. Hence, the performance of the biophysical model (i.e., EPIC) in estimating yields has a major impact on the model output.

Fortunately, for Brazil, the availability of yield statistics allocated to the lowest geographic level (i.e., SimU) is ideal for the evaluation process, reducing the methodological problems of comparing spatial information to statistical information at an aggregated level such as states, etc. Table 5 presents the results from the yield validation for sugarcane, soybeans, and maize.

Table 5:   Evaluation parameters for EPIC yields compared with national statistics

Indicator	Sugarcane		Soybeans		2nd Maize	
	t/ha	%[1]	t/ha	%[1]	t/ha	%[1]
Yield Mean (IBGE)[2]	74		2.9		5.5	
Mean Error (ME)[2]	-8	-11	1.5	51	0.0	0
Mean Absolute Error (MAE)[2]	18	24	1.5	52	1.1	20
Standard Deviation (SD)	26	35	1.1	38	2.0	36
Root Mean Square Error (RMSE)	29	39	1.7	60	2.3	41
Number of SimU (n)	6,491		4,811		3,095	

[1] normalized based on the mean of the observations

[2] weighted on the harvested area (2015)

Source: own calculations based on IBGE (2019)

The first interesting indicator is the ME showing different signs when comparing sugarcane and soybeans. That indicates that EPIC tends to underestimate sugarcane yields while overestimating soybeans and maize. Although this information may be useful to identify the first tendencies of the model, the more reliable error indicator is the MAE (see 6.1.1). The MAE is presented as the average for the whole country weighted on the harvested area of each SimU in 2015. The weighting reduces the importance of larger errors in areas with current low production.

For sugarcane, the MAE is 18 t/ha, resulting in a rMAE of 24%, while for soybeans, the MAE is 1.5 t/ha or 52% in relative terms. The lowest rMAE observed is on second-season maize with 20%. To further understand the performance of EPIC, Table 5 also presents the results from the RMSE for all crops. As previously discussed, this indicator strongly penalizes larger errors. The RMSE for sugarcane is 29 t/ha and rRMSE 39% of the average yield; soybeans 1.7 t/ha (60%). These results are in line with expectations since the model is expected to have a lower performance in marginal areas, where the current calibration may not apply.

A closer look at the SimU shows that EPIC tends to underestimate sugarcane yields in high-producing regions and overestimate in low-yielding areas. That can be partly observed by comparing the arithmetic mean error against the weighted. The former is 13 t/ha and the latter -8 t/ha. Balkovič et al. (2013) found similar EPIC behavior of underestimation in high-yielding areas and overestimation in low-yield regions.

Slightly different behavior is observed for soybeans, with EPIC mostly overestimating yields across all areas. It indicates that constraints such as the lack of yield penalty due to pests and diseases may be affecting its overall performance (see 4.3.1 for more details).

The best results are obtained with maize as the second crop. The MAE is the lowest in the comparison and even though the RMSE is slightly higher than for sugarcane, which is probably driven by lower sample size and the fact that second-season maize is mainly grown in few regions in Brazil so that marginal areas are likely to produce strong outliers impacting the RMSE. This relatively good performance may be partly explained by the intensive work from IIASA to adapt the crop calibration in EPIC to mimic specifically the second-season maize production in South America.

Considering the relatively simple calibration and the lack of regional production system information inputted into EPIC, it is feasible to argue that the results from sugarcane and maize estimations are reasonable but less satisfactory for soybeans. The error behavior being different between sugarcane and soybeans, when it comes to over- or underestimation, further complicates the usability of EPIC results as input for the PAM approach. If error patterns were in the same direction – e.g., always overestimating – it should be less problematic because the relationship between the crops would still fit. However, with different patterns, it becomes difficult to use EPIC without further calibration for the regional production systems in Brazil. Finally, it is important to highlight that for this case study, the results from EPIC were adjusted using the IBGE data (BAU scenario) to represent a situation in which a better calibration is achieved (see 5.1).

6.2.2.   Production cost evaluation

The production cost estimation approach within PAM is the most important innovation. Therefore, a detailed evaluation process is key to highlight the benefits and drawbacks of using such a modeling approach. Unfortunately, the benchmark information to allow a validation process similar to that for yields is not currently available, so a different strategy is proposed using different data sources to allow the validation. The results from the production cost evaluation are presented first for sugarcane and grains – i.e., soybeans and maize as the second crop – on state-level averages, and later with a summary table comparing the statistical measures proposed in the methodology.

6.2.2.1.   Sugarcane production cost evaluation

Following the proposed methodology, the data from the 45 typical farms for the year 2016-2018 from PECEGE (n=125) is compared with the results from the BAU scenario from PAM. Figure 45 summarizes the results from this comparison in USD per t of sugarcane, using a simple average to compose the state average. It is important to keep in mind that the sample from PECEGE follows the current importance of each state for the national production (Figure

45 in brackets). Therefore, the average for the state of São Paulo (SP) comes from 28 regions, whereas the average for Mato Grosso do Sul (MS) represents one farm. Hence, the average presented should not be interpreted as representative of all areas in each state but as an average of the main production regions within the state.

Figure 45:   Validation sugarcane cost estimates PAM (BAU) vs PECEGE in USD/t[38]
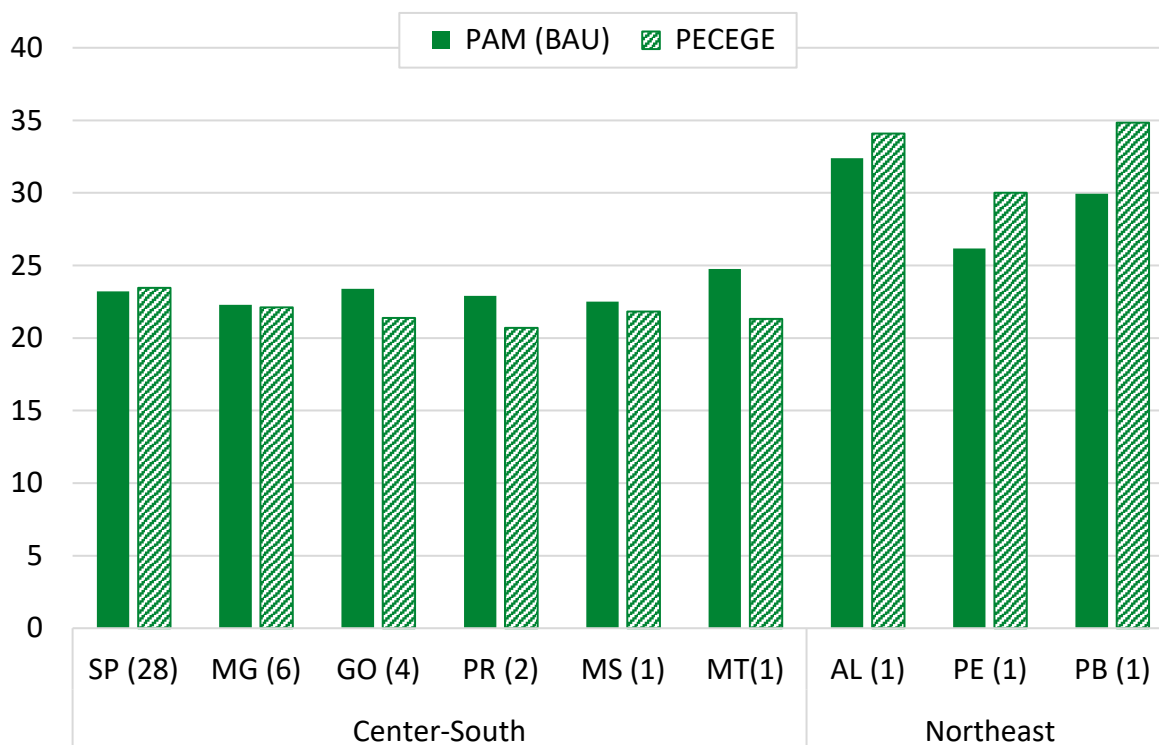
Source: PECEGE (2019) – created by the author.

Considering the most important production region for sugarcane – i.e., CS – PAM estimates a production costs' mean of 23 USD/t compared with 22.5 USD/t from the observed data from PECEGE. The result is a ME of 0.49 USD/t and a rMAE of 8.3%. It is interesting to observe that for the main production states of SP, MG and MS, PAM total cost estimation is almost identical to observed data from PECEGE, whereas, in the other states in the CS region, a relatively larger deviation between observed and estimated can be found, with PAM consistently overestimating production costs.

Focusing on the relatively minor producing states in the Northeast region, the PAM model tends to underestimate production cost per t with a mean of 29.5 USD/t, whereas the three data points from PECEGE have a mean of 33 USD/t, representing a ME of -3.48 USD/t and a rMAE of 10.5%. Such results are expected due to the relatively complex production systems in the Northeast states and the lack of *agri benchmark* typical farms to represent this region
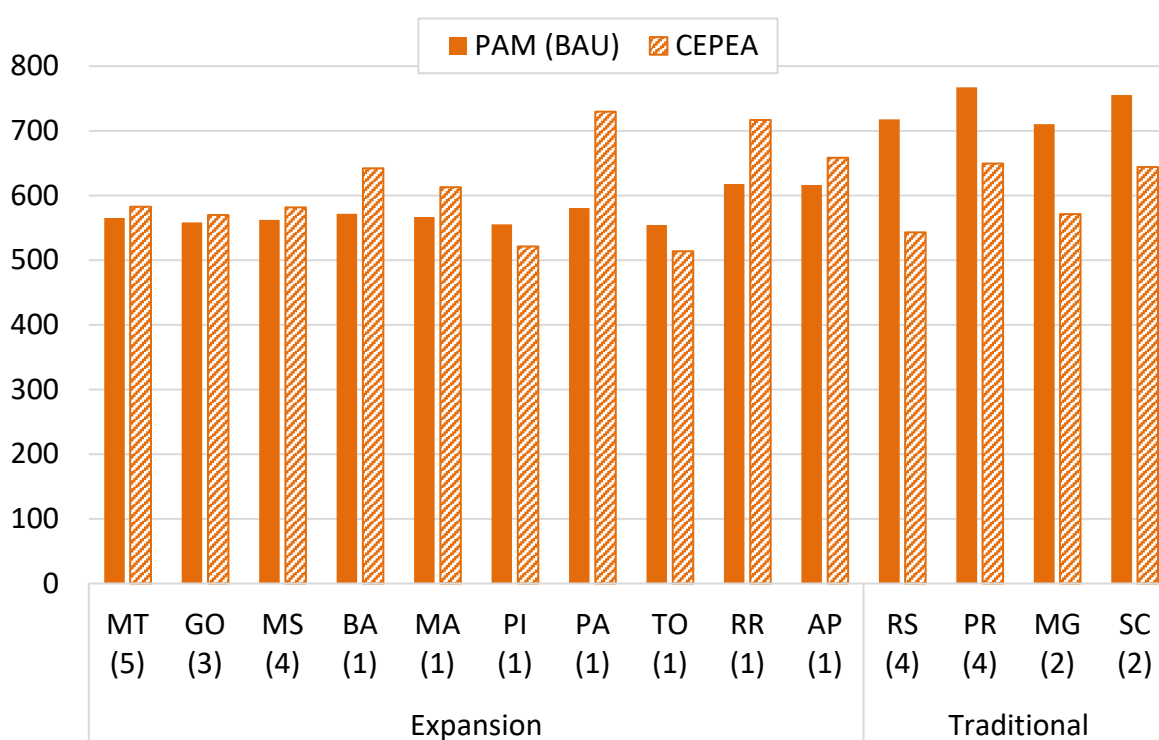
---

[38]   Brackets show the number of municipalities collected by PECEGE used in the average for each state.

in the model. Nonetheless, it is important to note that this region currently accounts for less than 8% of the national output for sugarcane, reducing the impact of this relatively poor performance on the overall output (CONAB, 2021).

6.2.2.2.   Soybeans and maize production cost evaluation

The results from the production cost validation for soybeans are presented in Figure 46, dividing the states according to the PAM classification into *Expansion* and *Traditional* regions. Similar to the data from PECEGE, CEPEA typical farms' location follows the current production patterns so that relatively newer production regions in the *expansion* region have a lower coverage, resulting in state averages coming from only one typical farm.

Figure 46:   Validation soybeans cost estimates PAM (BAU) vs CEPEA in USD/ha[39]

Source: CEPEA/CNA (2020) – created by the author.

The total production costs average estimated by the PAM model for soybeans is 636 USD/ha, compared with 605 USD/ha from the CEPEA data, considering all regions. The model simulations have less than 3% deviation in the three most important producing states (i.e., MT, GO and MS), which together account for 58% of the current soybean acreage in Brazil (CONAB, 2021). The performance for the remaining states in the *Expansion* regions is less satisfactory

---

[39]   Brackets show the number of municipalities collected by PECEGE used in the average for each state.

(standard deviation of 10%), with PAM mostly underestimating production costs. It is interesting to observe that due to the low number of typical farms in these regions, the state averages have a much stronger variation than in the more consolidated states.

The strongest deviation between estimated and observed production cost for soybeans comes from the *Traditional* region, where PAM consistently overestimates production costs in all states. An important characteristic that may partly explain the relatively poor performance of the PAM model in the region is that only one typical farm was used as input for the whole region. Therefore, features of this single typical farm – e.g., small acreage, excessive machinery or assumptions on family labor (wages) are expected to significantly impact the overall performance of the PAM model in this region.

Maize as second-season crop

Figure 47 presents the comparison between the estimated total production costs from the PAM model to the observed data from CEPEA for maize as a second-season crop. PAM estimates a mean of total costs of 533 USD/ha compared with 487 USD/ha from the CEPEA typical farms, including all regions. As for soybeans, PAM performs relatively well in important states such as GO, MS and PA while overestimating production costs in most of the remaining states. The strongest deviation is again found in the *traditional* region – namely in the state of PR – an error pattern similar to that observed for soybeans.

Figure 47:   Validation maize 2nd season cost estimates PAM (BAU) vs CEPEA in USD/ha[40]
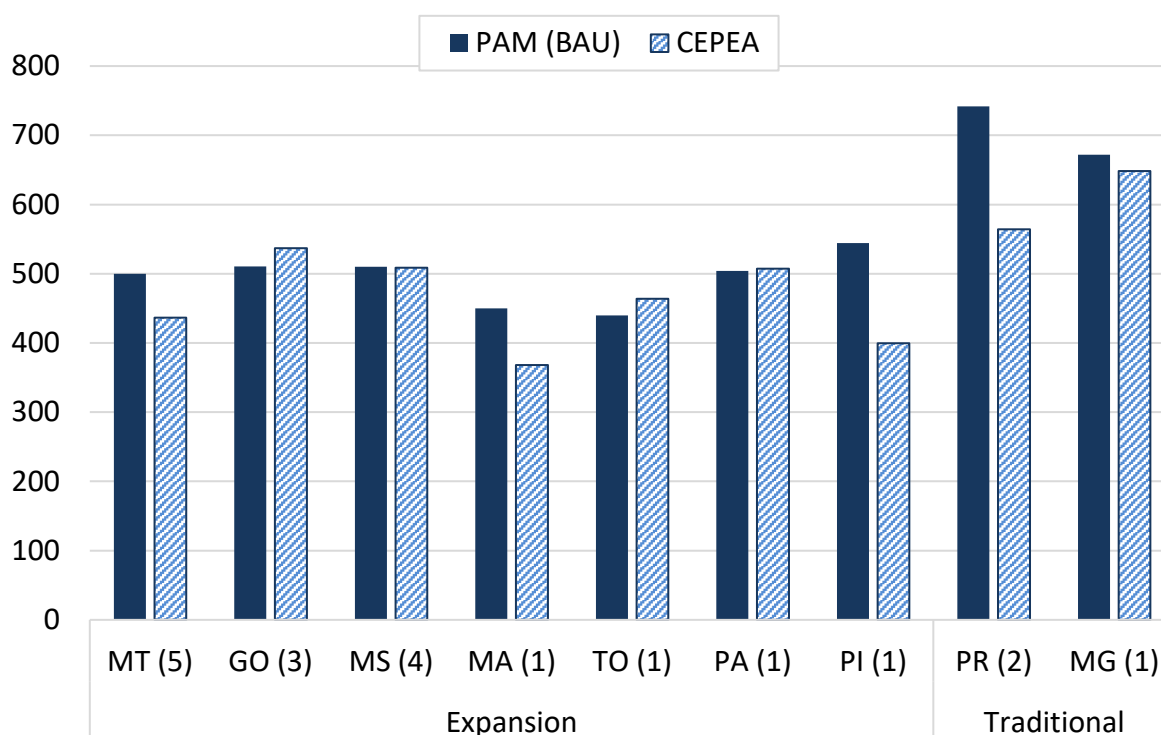
Source: CEPEA/CNA (2020) – created by the author.

Although this visual comparison of the absolute production costs estimates vs observed costs is helpful to identify possible model skewness as well as patterns, it is important to better quantify the PAM performance based on the statistical measures proposed in the methodology.

6.2.3.   Summary of the production cost evaluation

To better quantify and understand the behavior of the PAM economic module in estimating total production costs, Table 6 summarizes the main indicators used to evaluate the model's performance for all crops analyzed, considering all SimU in Brazil.

[40]   Brackets show the number of municipalities collected by PECEGE used in the average for each state.

Table 6:   Model evaluation – benchmark indicators comparing PAM (BAU) total costs in USD/ha results to observed data

Indicator	Sugarcane		Soybeans		2nd Maize	
	USD/ha	%[1]	USD/ha	%[1]	USD/ha	%[1]
Total Costs Mean[2]	1,766		605		487	
Mean Error (ME)	25	1.4	31	5.1	47	9.6
Mean Absolute Error (MAE)	153	8.7	78	12.9	68	14.0
Standard Deviation (SD)	183	10.4	97	16.0	74	15.3
Root Mean Square Error (RMSE)	183	10.4	100	16.5	86	17.7
Number of municipalities (n)	45		31		19	
Data source	PECEGE		CEPEA/CNA		CEPEA/CNA	

[1] normalized based on the mean of observations

[2] mean of the validation typical farm data (observed)

Source: own calculations based on PECEGE (2020) and CEPEA/CNA (2020)

Starting with sugarcane, the average total production cost per ha estimated by the PAM model is 1,790 USD while the observed data from PECEGE average 1,766 USD/ha. The resulting MAE for sugarcane is 153 USD/ha, representing a rMAE of 8.7%. If compared with the 8.5% rMAE obtained with the validation using USD/t as an indicator, it is safe to assume that yields observed from PECEGE and those used in the PAM estimations are similar and not causing major biases. This is an important finding showing that the magnitude and direction of error using the different indicators (per t or ha) are similar, with the PAM model slightly overestimating total production costs on average.

For soybeans, the average total production cost (Table 6) estimated in the BAU scenario is 636 USD/ha compared with observed at 605 USD/ha, resulting in an MAE of 78 USD/ha and rMAE 12.9% of total costs. The rRMSE of 16.5% further highlights the relative robustness of the production costs estimate for soybeans within the PAM modeling approach.

The MAE for maize as the second crop is 68 USD/ha and rMAE 14% from a total average cost of 487 USD/ha – results comparable to the performance of the model for soybeans. It is important to note that, in all crops, PAM on average overestimates production costs at a comparable scale, indicating that the relationship between the cropping alternatives should be close to the reality, even if the model estimates higher absolute values.

Considering (a) the focus on accurate production costs estimation while relying mainly on currently available data, (b) the limited number of typical farms, (c) the possible divergences in costing methodology (e.g., cost allocation, depreciation, etc.) and (d) the overall challenge of estimating production costs, the rMAE of 8.7% for sugarcane, 12.9% for soybeans, and 14% for maize can be considered satisfactory for micro-regional economic analyses. Balkovič et al.

(2013) argue that, for studies of such scale, relative errors of less than 30% can be considered reliable.

Finally, it is interesting that while the overall performance is satisfactory, the *traditional* region for grains as well as the Northeast region for sugarcane may be considered for further development, perhaps based on establishing new farms or refining the current *agri benchmark* typical farms in these regions.

6.2.4.   Land allocation and land-use evaluation

The final output of the PAM modeling approach is the spatial allocation of the crops with the highest return to land to each SimU, creating a simplified land-use map. This subchapter focuses on the validation process of this output following the proposed methodology.

Graphical comparison

The first step of the three proposed for the evaluation of the PAM land allocation module is a graphical comparison between estimated and observed land-use maps. Figure 48 shows the final allocation map from the PAM model for the BAU scenario (left) compared with the simplified land-use map based on IBGE information at the municipality level (right).

Figure 48:   **Graphical evaluation of the land use estimation: PAM land-use map (left) compared with official statistics from IBGE (right)**
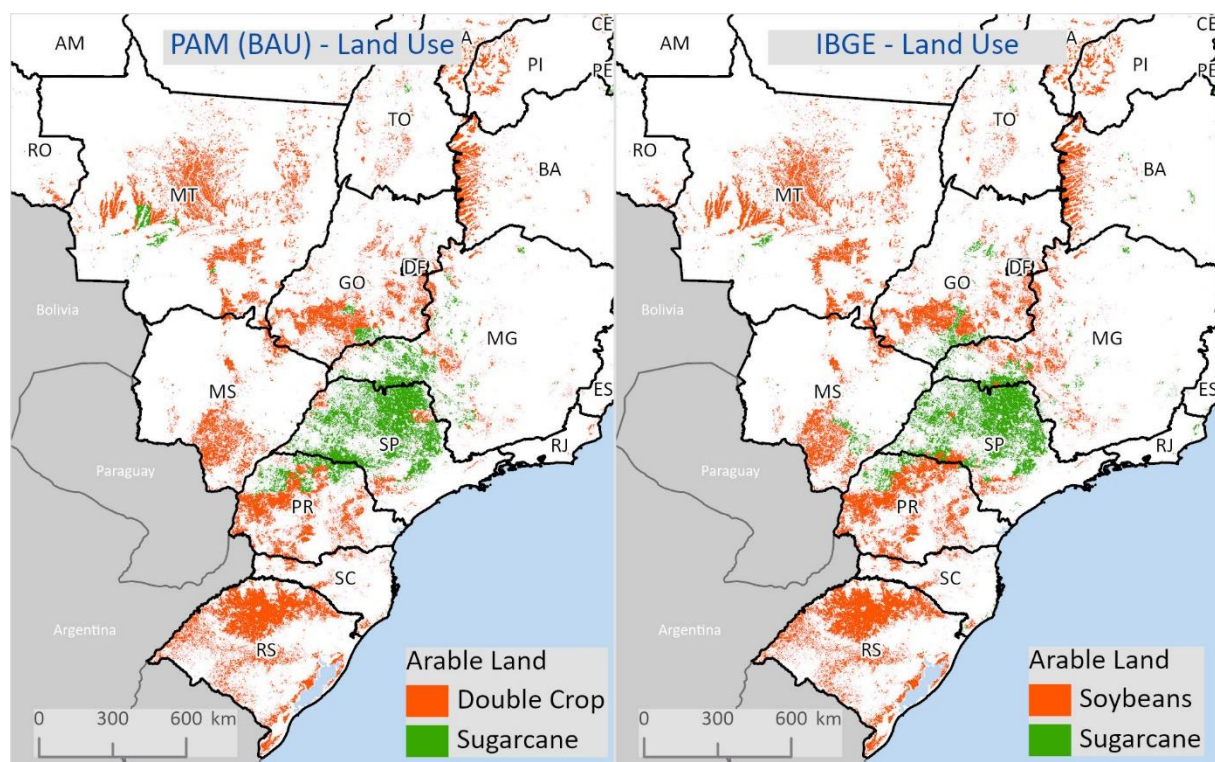
Source: PAM results (2020) and IBGE (2019) – created by the author.

The first important finding is that the overall allocation pattern from PAM is similar to the observed data from IBGE, with the main cluster of sugarcane production centered in the state of SP, expanding into parts of MG, GO and MS. Most of the remaining areas are allocated to double cropping of soybeans and maize. Visually, the main divergences between the model output and the observed data are found in MS and MG. PAM allocates the entire arable land in MS to double cropping whereas the northeastern area of the state has sugarcane as its main crop based on the map from IBGE. For the MG, the situation is the opposite, with PAM allocating most of the western part of the state to sugarcane, whereas IBGE shows soybeans dominating this area.

The overall land allocation pattern from the PAM model can be considered satisfactory since there is a strong visual overlap between the land-use estimation in the BAU scenario and the simplified land-use map from IBGE (and SPAM). Nonetheless, it is important to quantify this performance with statistical measures since it is challenging to evaluate the model based on visual evidence alone.

Matching score

To better quantify the overall performance of the PAM model in reproducing the simplified land-use maps from IBGE and SPAM 2010, the matching score illustrates the number of pixels in these maps with matching results – i.e., same crop alternative. Table 7 shows the results from the comparison as a percentage of matching pixels in each state in Brazil. The table also shows the importance of each state in the national production of sugarcane and soybeans as well as their share of total national acreage (i.e., including all crops).

Table 7:   **Matching score results measuring PAM performance compared with IBGE (2019) and SPAM/IFPRI (2019)**

State	Share of the crop in state arable land (%) [1]		National acreage	PAM performance	
				IBGE	SPAM
	Sugarcane	*Soybeans*	*%*	*%*	
Mato Grosso (MT)	2	60	18.0	95	94
São Paulo (SP)	70	13	15.9	88	91
Rio Grande do Sul (RS)	0	64	14.8	100	98
Paraná (PR)	6	51	11.3	83	83
Goiás (GO)	15	52	10.2	81	79
Minas Gerais (MG)	21	36	7.8	63	59
Mato Grosso do Sul (MS)	13	50	6.9	89	90
Bahia (BA)	2	50	4.6	94	91
Tocantins (TO)	3	66	1.9	98	92
Maranhão (MA)	3	52	1.7	98	93
Piauí (PI)	1	46	1.7	96	94
Santa Catarina (SC)	0	46	1.2	95	81
Pará (PA)	1	46	0.9	77	72
Alagoas (AL)	69	0	0.7	58	58
Rondônia (RO)	0	51	0.5	99	87
Others	-	-	1.8	23	21
Total (Brazil)	14	47	100	87*	86*

[1] average from 2017-2019 (IBGE)

[2] CE, PE, RN, ES, SE, PB, DF, RJ, RR, AP, AM, and AC

* based on the total matching pixels regardless of the state (same results if using weighted average)

Source: own calculations based on IBGE (2019) and IPFRI (2019)

The overall performance of the PAM allocation module varies significantly among the states. On one hand, states such as MG and GO have the lowest matching score – between 60% and 80% – considering the main producing states in Brazil. To a certain extent, that is expected since these states have an important share of soybeans (and maize) as well as sugarcane. Therefore, limitations such as the ability to allocate only one crop to each SimU oversimplify the actual structure of current land use. Other uncertainties come from the validation method to create the benchmark maps by allocating the whole area of a pixel to the crop with the largest share of acreage. That is expected to create larger biases in states with more mixed combinations of crops. Relatively poor performance also observed in the less important states such as AL and "others." That is likely to be a combination of different production systems, the importance of other crops and the lack of typical farms to better calibrate the PAM model.

On the other hand, the PAM land allocation module performs well in the most important pro-ducing states such as MT, SP, RS and PR, which together account for more than 60% of the total acreage in Brazil. The overall performance of PAM in allocating crops that match with the observed data is 86% or 87% depending on the source used as the benchmark. It also is important to highlight that no major differences are observed between the validation using

the IBGE data and the SPAM 2010 data. That is an additional indication of the robustness of the PAM model in reproducing the current land use for these crop alternatives.

This finding supports the observations from the graphical comparison results, showing that the land allocation from the PAM model has an overall good performance, matching more than 86% of the pixels when compared with the benchmark land-use maps. Nonetheless, it is important to understand the model behavior in areas with more complex and mixed production systems, in which the model performs less satisfactorily (mainly MG). For these states, higher resolution modeling should improve the land-use estimation performance.

Performance visualization

The last step in the land allocation validation process is to combine the matching information from the previous section with the graphical visualization. Figure 49 shows the performance visualization maps comparing the results from the PAM (BAU scenario) against the simplified land-use map from IBGE.

Figure 49:   Validation BAU (PAM) against IBGE simplified land use data
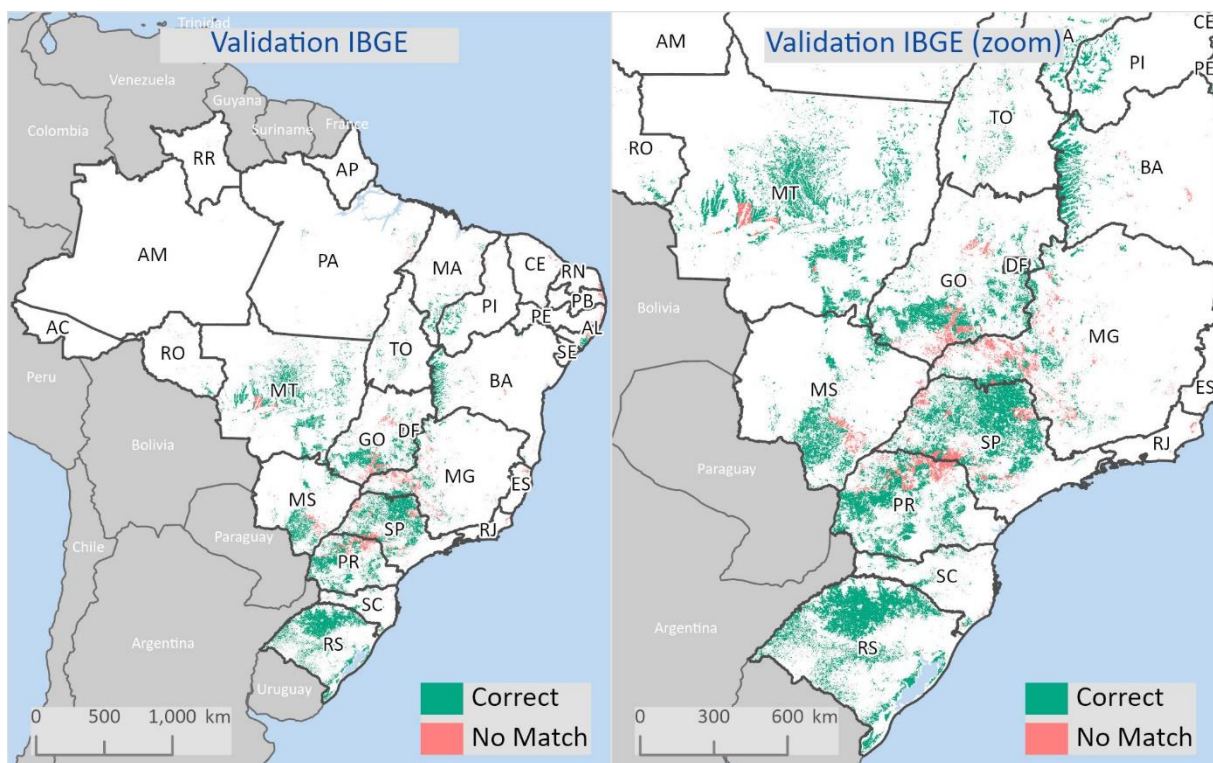
Source: IBGE (2019) – created by the author.

Supporting the previous results, Figure 49 highlights the overall good match between the results from the PAM model and IBGE data. The results using the SPAM as the benchmark map are very similar (see appendix 8). The pixels with non-matching results are mainly found in the MG, GO, and PR. This visualization helps us understand the limitations of the land allocation module within the PAM model.

Summary

The three-step evaluation process for the land allocation module provides interesting insights into how the model behaves. The overall land allocation performance of 86% is satisfactory considering the scope and complexity of estimating land-use based solely on economic profitability. The evaluation process also indicates where the model has its strongest problems (e.g., MG state) due to the diversity of the region and the limitations on spatial resolutions. A thorough evaluation process is imperative to inform users of the limitations and strengths of the model and to highlight areas for further development.

7.   Climate change and the competitiveness of Brazilian crops

Climate change is expected to have the greatest impact on the agricultural sector and especially so in developing countries due to relatively stronger expected atmospheric variation as well as relatively higher dependence of the economy of these regions on the agricultural sector (Mendelsohn, 2008). A common approach to model the impact of climate change on the agricultural sector is to simulate the change in crop yield. That should help forecast future land-use changes in important producing countries due to the reallocation of land to different crops. Several approaches have been used to quantify the impacts of changes in the atmosphere on crop cycles, water availability and, finally, yields (Knox, Hess, Daccache, & Wheeler, 2012). An interesting outcome of such models is the expected changes in yields, depending on the emission patterns in the future. It seems, however, that these atmospheric changes may not affect all crops and regions equally. That brings an important question regarding the possible cropping choices for farmers in the future, depending on how climate change affects the economic performance of each crop at the farm level in different regions.

From this perspective, the PAM modeling approach seems to be an interesting tool to analyze the impact of yield change due to climate change on the economic performance of the major cropping systems in Brazil – namely double cropping (soybean-maize) or sugarcane – and, ultimately, on the land-use structure. Such scenario analyses are important to observe the PAM model behavior with changes in inputs (i.e., crop yields), to identify possible limitations and, more importantly, to illustrate how this new modeling approach can be used empirically. The following subchapters review the literature on the expected impact of climate change in different crops and regions in Brazil, followed by the scenarios implementation strategies, and finally present the results of the scenario simulations.

7.1.   Background: climate change impact on Brazilian cropping pattern

The general understanding is that climate change is expected to impact tropical and subtropical regions more strongly than temperate zones. Narrowing down to Brazil, the prevailing expectation is that crops in the southeastern regions will benefit whereas northeastern and the Amazon regions will most likely suffer from climate change (Mendelsohn, 2008). The main changes are expected to be an increase in average temperature (stronger in the central regions), changes in total and distribution of precipitation and the number of days with temperature above 34°C, among others (Zilli et al., 2020).

Arvor, Dubreuil, Ronchail, Simões, and Funatsu (2014) argue that for the main producing regions in the Cerrado (e.g., MT), a lengthening of the dry period, as well as changes in spatial and temporal rainfall distribution, is expected. Besides global climate changes, deforestation on the Amazon basin may cause further local climate changes in neighboring regions, potentially reducing rainfall in extreme cases by up to 25% (Nobre, Sellers, & Shukla, 1991). Conversely, in the South of Brazil, the level of deforestation is minimal which combined with a relatively small change in precipitation, indicates a positive development of soybean yields

mainly due to higher levels of CO_2 (Pires et al., 2016). Against this regional background, it is important to understand how the different crops may react to such climatic changes.

Sugarcane

First, it is important to highlight that only a few studies have been carried out to quantify the impact of climate changes on sugarcane yields, even though it is a major crop in tropical regions (Knox et al., 2012). Fortunately, most of the studies are in Brazil and point in the same direction (Linnenluecke, Nucifora, & Thompson, 2018). For the main producing region (e.g., SP), most results indicate an increase in yields attributed mainly to increasing CO_2 concentration and temperature being able to compensate for possible losses due to a reduction in precipitation (Marin et al., 2013; Singels, Jones, Marin, Ruane, & Thorburn, 2014). The magnitude of the climate change impact on sugarcane yields in the southern region in Brazil varies strongly among the authors, with values as high as 20% to 58% by 2030 and 59 to 82% by 2090 in the state of SP (dos Santos & Sentelhas, 2014).

When looking at different climate zones in Brazil, the availability of results is very limited. One paper from Carvalho et al. (2015) suggests that sugarcane yields are expected to decrease in the Northeast region (Pernambuco) due to higher temperatures and lower water availability. This effect of yield losses as a result of the strong increase in temperature that cannot be compensated by higher CO_2 availability also is described by Berg, Noblet-Ducoudré, Sultan, Lengaigne, and Guimberteau (2013) for arid regions.

For the tropical region in Brazil, few studies focus on the impact of climate change in sugarcane (not the main producing region). The exception is the work from Marin et al. (2013), in which they compare two regions in the state of SP, including areas with climate regions classified as Aw and As (tropical) based on the Köppen classification (Alvares, Stape, Sentelhas, Gonçalves, & Sparovek, 2013). Most regions in the Cerrado biome belong to the same climate category. In their study, all areas (tropical and subtropical) had a positive overall response to climate change, with sugarcane being able to profit from higher water use efficiency and higher CO_2 availability.

Recently, Flack-Prain et al. (2021) contested the majority of sugarcane yield projections under climate change scenarios, arguing the mechanism of the main models used (i.e., DSSAT and APSIM) may be misrepresenting the complex interactions between sugarcane ecophysiology and other climate indicators. That highlights the fact that any projections are dependent on model specifications, climate scenarios and calibration.

Soybeans and Maize

For soybeans and maize, it is crucial to focus on the "two" main climate zones in Brazil when describing possible changes in yield due to climate change. Moreover, it also is important to consider the combination of crops (i.e., double-crop) instead of a single crop since it is the prevailing system in Brazil. Double-cropping is responsible for the expansion of Brazilian maize and soybeans, allowing farmers to grow two crops in one season. That, however, strongly

depends on the climatic conditions, with the most important indicator of the presence of double cropping being the duration of the wet season (Arvor et al., 2014).

Pires et al. (2016) argue that in most double-cropping regions, the wet season lasts between six and seven months, allowing growers to seed early soybeans right after the ending of the sanitary break[41] and having enough time to harvest soybeans and cultivate second-season maize afterward. Arvor et al. (2014) show a close link between the early onset of the wet season and total precipitation with the total double-cropping area in MT. Therefore, while focusing on the double-cropping system in the Cerrado region (tropical zones), the expected effect of climate change is delayed onset of the wet season, which is expected to delay soybean seeding and thereby reduce farmers' ability to grow second-season maize (Pires et al., 2016).

The system is further complicated by the expected reactions of maize to climate change in tropical regions. Models using the thermal time to compute cycle length indicate that increases in temperature shorten the crop cycle by speeding up phenological phases, reducing biomass growth and yields (Berg et al., 2013). Moreover, the increasing temperature is expected to increase cereal (e.g., maize, wheat, etc.) respiration rates, thereby decreasing yields (Wang, Vanga, Saxena, Orsat, & Raghavan, 2018). Therefore, maize grown after soybeans is expected to have a yield decrease up to 28%, if no technical changes are implemented (Hampf et al., 2020).

For the Southern (sub-tropical) region, an increase in soybean yields is expected, even for short-cycle varieties usually used in the double-cropping systems. The reasons are an expected higher CO_2 concentration and relatively modest changes in precipitation (Pires et al., 2016). For maize, a modest yield reduction is expected in the near future with relatively high yield penalties projected between 2071 and 2100 (Camilo et al., 2018).

7.2.   Scenario implementation in the PAM model

### 7.2.1.   Description of the scenarios

Climatic zones

The PAM approach is based on regional information and, therefore, it is important for a robust implementation of the scenarios to divide the production regions according to their climate zone classification. This dissertation uses the climate classification from Alvares et al. (2013), in which Brazil is classified according to the Köppen zones, based mainly on the total and distribution of rainfall and temperature. The detailed results from this classification in GIS format

---

[41]   The sanitary break is adopted by Paraguay and Brazil between June and September when soybean production is not allowed, to reduce the pressure from Asian soybean rust - Pires et al. (2016).

is used to divide the country into three major climate zones (i.e., simple classification), assigning each SimU to the prevailing climate type based on its geographic location. The resulting classification, as well as the original map, are shown in Figure 50.

Figure 50:   Köppen climate classification for Brazil (left) and simplified version (right)
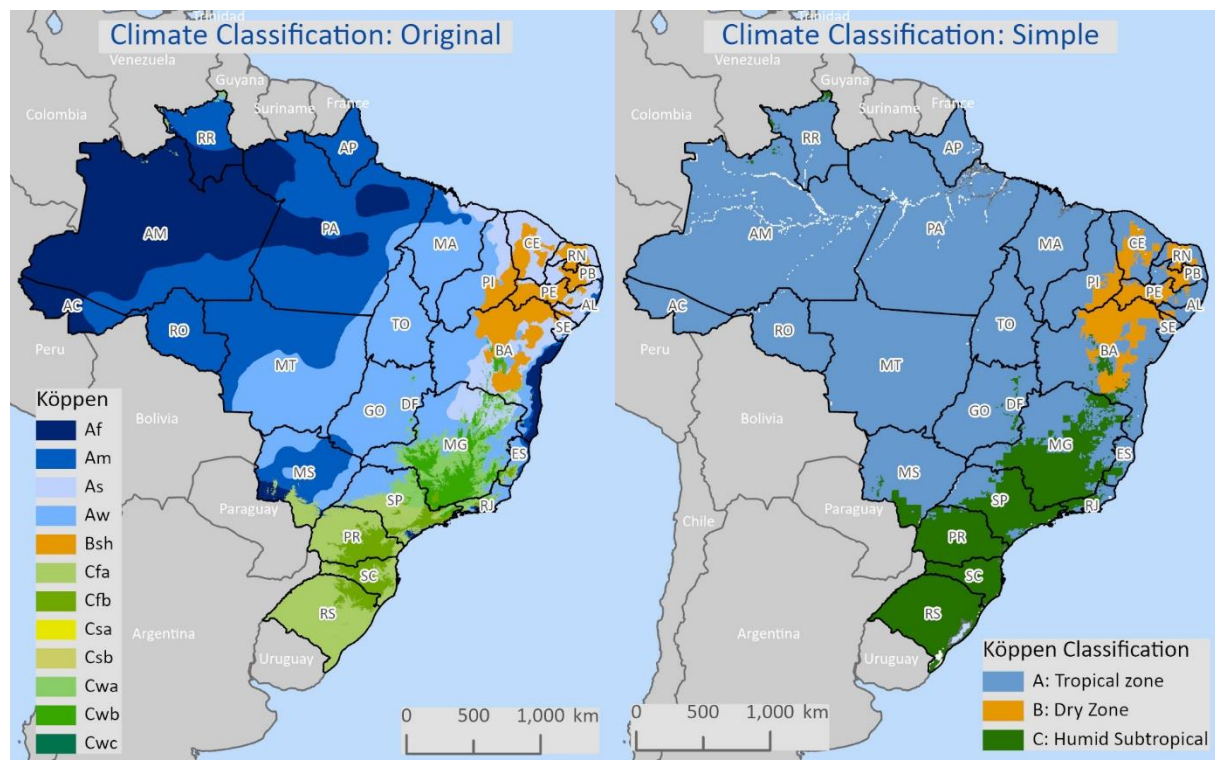
Source: Alvares et al. (2013) – created by the author.

Since no significant crop production takes place in the dry zone (right, B), the most important regions for the scenario implementation are the tropical (A) and the humid subtropical (C) zones. The tropical zone (A) covers most (81.4%) of the Brazilian territory and has a total rainfall variation between 1,000 mm (As, left map) to ca. 3,000 mm (Af). Most of the crop production in the tropical zones takes place in the climatic zone Aw, which receives more than 1,300 mm and has a distinct seasonality – wet summer and dry winter – and mean temperature around 24 °C. The humid subtropical zone (C) covers 13.7% of the total territory, with the Cfa classification covering most of the crop production within this zone. Its total rainfall is similar to the Aw region but with more rainfall in winter and considerably lower temperatures (Alvares et al., 2013).

Yield impact due to climate change

The most important impact of climate change considered in this work is on yields. Therefore, most of the adaptation of the PAM model is to implement yield changes based on available projections with a time horizon until 2050. For sugarcane, this dissertation follows the work from Marin et al. (2013), assuming for SP an average 22% yield increase for climate region C

(humid tropical) and 15% for climate region A (subtropical) since this region has higher mean temperatures. The latter is an adaptation based on the uncertainty regarding the role of increasing temperature and its effect on yields for this region, considering the strong negative impact found in the work from Carvalho (2015). These regional yield increases are kept constant for all soybean-maize scenarios.

For the double-cropping system, I propose two scenarios based mainly on the work from Pires et al. (2016) and Hampf et al. (2020).

No Double Cropping in Tropical Zone (NoDC):

This scenario is mainly based on the work from Pires et al. (2016), which uses a realistic climate change scenario including atmospheric change due to global greenhouse gas (GHG) emissions combined with regional climatic changes due to deforestation in the Cerrado and Amazon (i.e., LUCID+PC13). They estimate a negative yield impact on short-cycle (i.e., 100 days) soybeans, which primarily are used in the combination of second-season maize. Following the argument that short-cycle soybeans in the central-northern (tropical) region of Brazil will suffer significantly from climate change (yield losses from 12% to 50%), it is assumed that farmers in this region most will likely forgo the returns from maize as the second crop and instead grow longer-cycle (i.e., 130 days) soybeans. According to Pires et al. (2016), the expected yield increase for longer-cycle soybeans varies between 2.2% in MT and 3.9% in MATOPIBA. Due to the relative importance of MT, the soybean yield increase assumed in this scenario is 2.2% for the SimUs in the tropical zones.

Conversely, farmers in the humid-subtropical zones are assumed to continue to pursue double cropping with a yield increase of optimum cultivar of soybeans of 15.7% based on Pires et al. (2016). Maize yields in this subtropical zone are assumed to decline by 8.7% following the work from Camilo et al. (2018). Hence, it is feasible to assume that farmers in the humid-subtropical region may continue to practice double-cropping.

Double Cropping with Yield Penalty (DCYp):

This scenario includes a possible adaptation by farmers to changes in the onset of the wet season in the tropical zones caused by climate change. Following the work of Hampf et al. (2020), farmers may react by postponing soybean seeding and therefore delaying the sowing of second-season maize. Even though delaying soybean sowing causes minor changes in yields, the impact on second-season maize yields is notable. Due to the expected changes in rainfall and temperature, the average maize yield penalty without technical adaptation is 28%. Besides the yield penalty, it is important to consider that delaying soybean sowing increases the risk of not being able to grow a second season crop due to the time required for soybeans to mature and the short time window for harvesting soybeans and sowing maize.

Since this effect is projected only for the tropical zones (A) in Brazil, the SimUs in the humid-subtropical zones (C) area are assumed to keep double cropping with a soybean yield increase of 15.6% based on Pires et al. (2016) and a yield penalty for maize similar to the NoDC scenario.

Table 8 summarizes the relative yield changes compared with the BAU scenario for all climate change adaptation scenarios, for the two main climate zones in Brazil. It is important to note that for sugarcane, the yield variation is based on the fresh stalk yield. The sugar content (TRS) per t of fresh cane is the same as for the BAU scenario.

Table 8:   **Assumed yield changes in the climate change scenarios vs the BAU (in %)**

Climate Zones	NoDC		DCYp		All Scenarios
	Soybeans[1]	*Maize*[2]	*Soybeans*[1]	*Maize*[2]	*Sugarcane*[4]
Tropical Zone (A)	2.2	-	-11.8	-28.0[3]	15.0
Humid Subtropical (B)	15.7	-8.7	15.6	-8.7	22.0

[1] Based on Pires et al. (2016)
[2] Based on Camilo et al. (2018)
[3] Based on Hampf et al. (2020)
[4] Based on Marin et al. (2013)

7.2.2.   Changes in the PAM modules

To run the climate change adaptation scenarios, the following changes are required to ensure consistent results that can be benchmarked against the BAU scenario.

Yield changes

The proposed climate change scenarios directly impact the plant module due to the changes in crop yield. Therefore, the yields of the BAU scenario are modified using the proposed changes presented in Table 8.

Transportation module

The basic transportation module is kept unchanged from that in the BAU scenario because external changes in infrastructure or transportation costs (i.e., diesel) are expected to affect all cropping alternatives similarly. The only important difference between the BAU and the scenarios is the assumption on the inbound transport costs for sugarcane from field to mill. For the climate change scenarios, I assume the current average inbound transport costs (see section 4.4.2 for more details) from farm to mill based on the idea that for future scenarios, the constraint of having to deliver cane to a current mill is not realistic so the model assumes that farmers need to cover only the current average transport cost (based on a 30-km distance).

Changes in production costs – economic module

An important topic is the related changes in production costs triggered by the yield changes. For sugarcane, the model is already reacting to changes in yields since there is a direct relationship between yield and the main costs items. Therefore, the expected yield increase directly translates into higher operating costs, the main cost component. This feature holds true

for K costs as well, which are already modeled considering nutrient uptake. The main uncertainty is N since the current model setup uses the output from EPIC on nutrient requirements to calculate costs. Hence, a linear adjustment is implemented using increasing nutrient use based on expected yield increase.

For grains, the production costs are kept the same as for the BAU. Due to the limited relationship between operating costs and yields, it is unfeasible to assume that yield increases will automatically translate into higher operating costs. For establishment costs, the required adjustment in K is already implemented in the model – i.e., based on crop uptake. Since soybeans do not currently receive any N fertilizer, this cost component should not change the outcome of the scenarios. The reasoning behind keeping the production costs similar for second-crop maize is based on the already minimal applications of N and similar constraints in relating its operating costs to yields. The remaining uncertainty is the change in P application that may be triggered by the expected yield decrease.

Crop prices at FOB and farm-gate levels are kept unchanged from the BAU scenario. The idea is to isolate the effect of expected changes in yield (ratios) caused by climate change from possible movements in commodity prices. Nonetheless, it is expected that such strong yield developments in a major producer of all three crops are likely to affect international prices. However, since PAM currently is only a regional supply model, it is out of the scope of this case study to try to estimate such movements. A model consortium with a global equilibrium model could be a way forward to tackle such developments more holistically.

Finally, the scenarios are calculated considering the current arable and grassland in Brazil (IBGE, 2018). The idea is to allow for possible land conversion from grassland to cropland as well as understanding which crops may prevail regionally following such yield change.

7.3.   Results from the climate change adaptation scenarios

7.3.1.   Changes in return to land for grains

Figure 51 presents the estimated returns to land for grains in the different climate zones and states, comparing the climate change adaptation scenarios to the BAU, in USD/ha.

Figure 51:   Returns to land for grains for the climate change adaptation scenarios and the BAU scenario (in USD/ha)
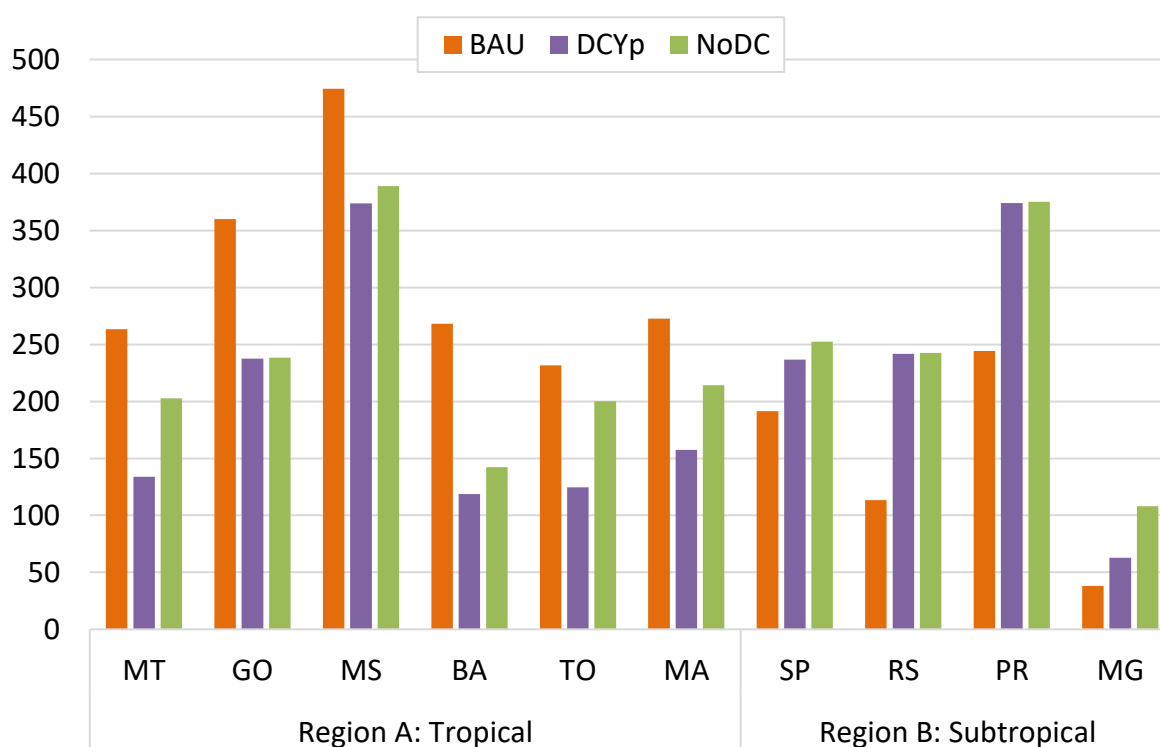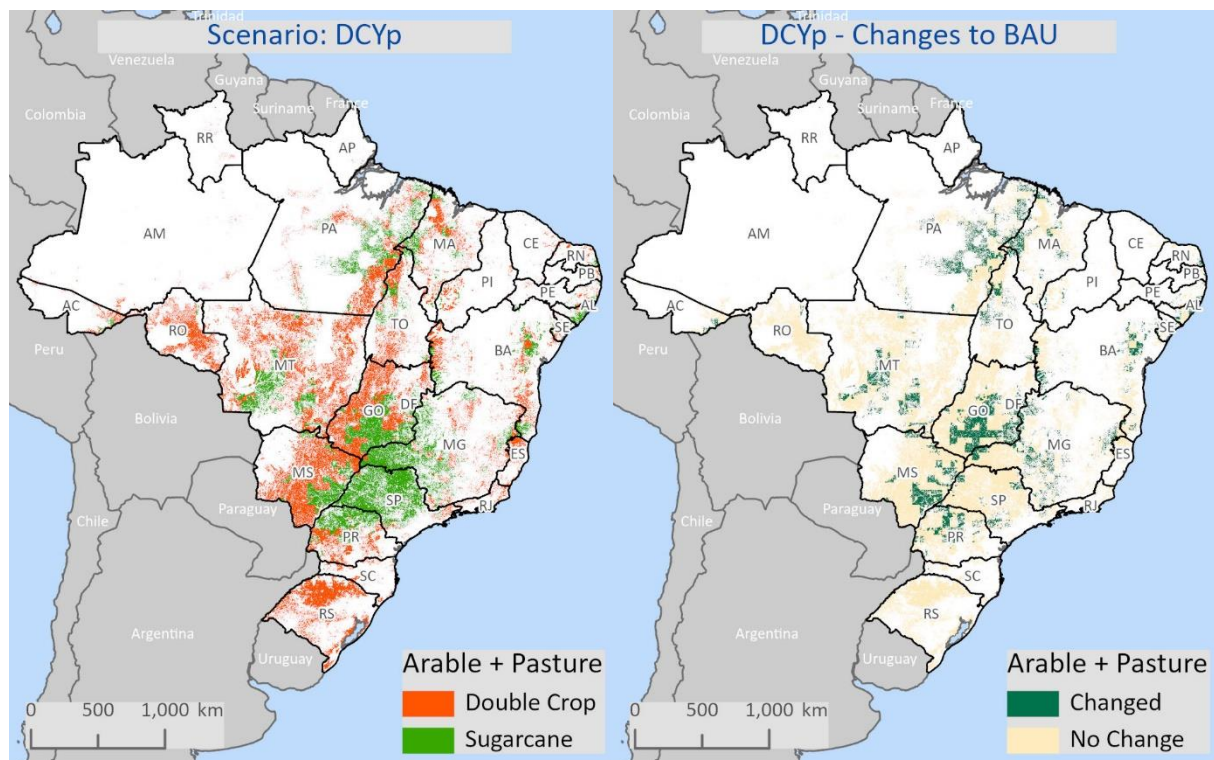

Source: PAM results (2020) – created by the author.

The overall picture is a strong decline in the returns to land in the tropical region (A), from a six-state simple average of 312 USD/ha in the BAU to 191 USD/ha (-39%) in the DCYp scenario and 231 USD/ha (-26%) in the NoDC scenario. The state with the highest reduction on return to the land is BA in both scenarios. Mato Grosso (MT) faces a decrease of 49% in the DCYp scenario but only 23% in the NoDC scenario, already indicating that if farmers face such climate change, they are likely to skip maize as a second crop and move toward longer-cycle soybeans. Conversely, the states in the subtropical region (B) face a significant increase in return to land for grains: on average, 56% for the DCYp scenario and 67% for the NoDC scenario.

7.3.2.   Changes in land use

The more interesting analysis is on the changes in relative profitability at the farm-level, which motivate farmers to reallocate their land, affecting the current land use. Figure 52 shows the

resulting land allocation based on the return to land from each cropping alternative for the DCYp scenario (left) and the observed changes compared with BAU (right).

Figure 52:   Land allocation results from the scenario double-cropping with yield penalty (left) and the changes compared with BAU (right)
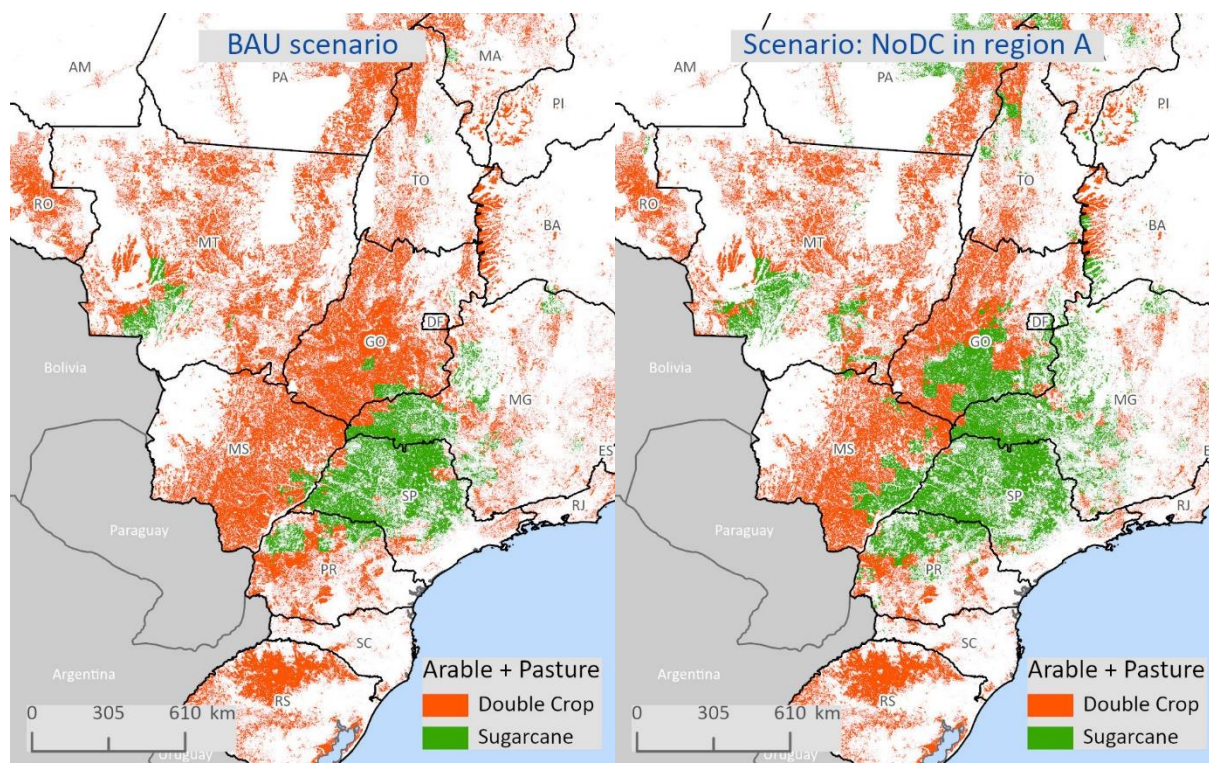

Source: PAM results (2020) – created by the author.

On a national level, most changes occurred in two main regions in Brazil – i.e., North (MA, PA, TO) and Center-west (MT, MS, GO). Both regions experience a strong change from the current double-cropping system of soybean and maize to sugarcane. The important difference be-tween these two regions is that the Center-west region is key for the current grain production in Brazil whereas the northern regions are still largely dominated by grasslands. That indicates that if future crop yields are affected in line with the current expectation, sugarcane should move into current grain-producing areas which, in turn, may result in further expansion into grassland if demand stays relatively stable. It is interesting to further understand whether the changes are only from grains to sugarcane or also the reverse, considering the strong yield increase projected for grains in the humid subtropical regions.

Figure 53 shows the simulated land-use maps for the BAU scenario as the benchmark and the results from the NoDC scenario in region A scenario, focusing on the main producing states.

Figure 53:   **Simulated land-use maps for the BAU scenario (left) and no double cropping in region A scenario (right) – zoom in the main producing states**

Source: PAM results (2020) – created by the author.

For NoDC in region A scenario, the major land-use change in the tropical region is indeed the expansion of sugarcane in the neighboring areas of the state of SP, while no major changes are seen visually in the state of MT. That is an interesting finding showing that, based on profitability, soybeans still are more profitable than sugarcane even without the contribution from maize. In the subtropical region (C), even the higher increases in returns to land for grains (67%) are not able to compete with the increased returns to land from sugarcane, resulting in the expansion of the latter into parts of PR. The state of SP faces a 61 USD/ha (+32%) increase in return to land for grains in the NoDC scenario while the returns from sugarcane increase by 231 USD/ha (+86%), moving all the arable land to sugarcane. Similar results at the local level are found for the DCYp scenario, with sugarcane having higher returns in SP and the neighboring regions compared with grain production (see appendix 9).

Measuring the overall land-use change

Even though the visual interpretation of results is important to understand spatial changes in the landscape, it is also crucial to try to quantify such developments. Table 9 presents the relative change from the scenarios compared with the BAU at the national level as well as the individual contributions of each state to the overall change. It is important to note that, since the current setup of the model represents only these three main crops, all available land (i.e., arable and grassland) is allocated to one of the cropping options. Therefore, the percent

changes may not reflect the actual distribution of these crops for each region but show the crop with the highest returns in each region.

Table 9:   **Total simulated land-use change in Brazil for the climate change adaptation scenarios compared with BAU scenario (in %)**

| State (UF)[1] | Total change and the contributions from each state (%) | | | |
| | Arable Land | | Arable + Grassland | |
	NoDC	*DCYp*	*NoDC*	*DCYp*
Goiás (GO)	32.8	29.1	22.6	21.5
Paraná (PR)	14.7	13.2	6.5	5.7
Mato Grosso (MT)	10.3	17.3	8.1	12.0
Minas Gerais (MG)	9.8	8.8	10.0	9.1
Bahia (BA)	7.2	6.3	8.2	6.9
Mato Grosso do Sul (MS)	7.2	7.5	8.5	10.1
São Paulo (SP)	5.7	5.1	3.0	2.6
Distrito Federal (DF)	2.7	2.5	1.2	1.0
Maranhão (MA)	2.7	3.3	7.9	7.3
Alagoas (AL)	1.7	1.8	1.1	1.0
Tocantins (TO)	1.4	1.6	3.0	3.8
Pará (PA)	1.4	1.3	18.3	16.6
Pernambuco (PB)	0.9	0.9	0.5	0.5
Piauí (PI)	0.5	0.5	0.2	0.2
Total (Brazil)[2]	**24.0**	**26.8**	**22.5**	**25.7**

[1] Sorted based on the change in arable land

[2] Total changes compared with business as usual (BAU)

Source: own calculations (2021)

Focusing firstly on current arable land, the overall simulated change in Brazil is estimated at 24% for the no double crop (NoDC) scenario and 26.8% for the double cropping with yield penalty (DCYp) scenario. Confirming the expectations, the yield reduction in the short-cycle soybeans is stronger resulting in worse economic performance than moving to a single crop system, growing only long-cycle soybeans. That partly explains why the NoDC scenario has lower land-use changes than the DCYp scenario.

Breaking down the national results to individual states, the main changes in land use are observed in the states in the tropical zone – i.e., GO, MT and MG, as well as PR in the subtropical zone. The state of GO alone contributes more than 30% to the overall change in arable land, indicating this state has the highest responsiveness to such yield development. That indicates a relatively close economic performance between the two cropping alternatives in the current scenario so that changes in framework conditions would lead to a strong supply reaction.

The inclusion of the current grassland does not significantly change the overall simulated land-use change, with levels of 22.5% for the NoDC scenario and 25.7% for the DCYp scenario. An important difference is the relative contribution of each state. States from the northern region (tropical) – mainly PA and MA – contribute substantially more when grassland is included than in the arable land-only comparison. That can be explained by the relatively low share of arable land in these states currently. However, the findings indicate that these regions under climate change may have a large share of area, where sugarcane has an economic advantage over double cropping of soybeans and maize.

7.3.3.   Summary and conclusion from the scenario analysis

The PAM model is adjusted to simulate the impact of climate change in the returns to land for the individual cropping alternatives, for specific climatic regions in Brazil. Based on that, the resulting land-use change for two scenarios is calculated: DCYp and NoDC (in region A). The results can be summarized as:

- The return to land for grains decreases by 39% (DCYp) and 26% (NoDC) in the tropical region (A). Conversely, returns to land increase significantly – by more than 50% – for the subtropical region.

- On the resulting land use, both scenarios have relatively similar results showing an overall increase in sugarcane mainly in the tropical region due to better economic performance.

- The state of GO has the strongest simulated land-use changes in both scenarios. Approximately 30% of the total change in arable land and more than 20% of the change in combined arable and grassland happens in this state. That would be a significant move from grains to sugarcane production.

- In the humid subtropical-climate zones (i.e., CS region) the return to land from sugarcane increases more strongly than grains, which is mainly visible in PR. That is interesting since the expected climate change portends yield increases for both cropping alternatives in this zone.

- Under the climate change adaptation scenarios, the northern region is expected to have a large share of the area with sugarcane economically outperforming double-cropping.

- Climate change is expected to decrease the availability of maize as a second crop since most farmers are likely to switch to a single crop system with longer-cycle soybeans in the Center-west (*Cerrado*) region.

Technically, the scenario analysis illustrates the flexibility of the PAM model in estimating changes in overall framework conditions that are not simple crop price movements. The endogenous estimation of production costs allows the introduction of changes such as the maximum amount of fertilizer allowed, yield penalties and changes in transport costs, with the model estimating the overall impact on farming profitability and the resulting allocation of land.

8.   Discussion and conclusions

This dissertation aims to answer the question whether a combined biophysical and farm-level economic data model can satisfactorily estimate the profitability and land allocation strategies of farmers in countries with low data availability. This chapter discusses the development of the PAM model, the main finding for the case study on Brazilian land-use change and the results from the climate change scenarios. The idea is to compare the PAM modeling approach to existing models, mainly focusing on production cost estimation, as well as outlining the model's limitations. That should help the interpretation of the results and identify areas where further research is required to improve model performance.

The discussion chapter is structured as follows: The first subchapters focus on the main building blocks of the PAM approach, namely the biophysical and economic modules. Thereafter, the main results and the model evaluation are discussed, comparing the PAM results to the literature. Finally, I focus on the simulations of the impact of climate change on the Brazilian supply of grains and sugar as well as the possible areas where the PAM model should help inform decision-makers.

8.1.   The biophysical model

The basis of the PAM modeling approach is the combination of a biophysical model with economic farm-level data. The former estimates primarily yield and input use while the latter is the basis for the production costs and profitability estimations. Besides the evident function of the biophysical model to simulate yields for each cropping alternative, this model also is responsible for the spatial variability component. The PAM spatial resolution is primarily determined by the biophysical model since the other components, such as the farm-level data, comes with a much lower spatial resolution. The current version of the PAM model uses the EPIC model calibrated by the IIASA institute (Balkovič et al., 2014; Skalský et al., 2008).

8.1.1.   Accuracy of the yield estimation

The performance of EPIC in estimating yields at the SimU level is a key determinant of overall PAM simulation accuracy. Yields not only are important for the profitability calculation (revenues) but also are used for the extrapolation of the known farm-level economic data to regions without typical farms. For the case study, I use the standard model calibration from EPIC-IIASA for sugarcane, soybeans and second-crop maize.

Grains

The yield estimation results indicate that EPIC overestimates soybean yields in all regions in Brazil compared with national statistics, allocated to the SimU level (Câmara et al., 2015). At the national level, the mean absolute error (MAE) is 1.5 t/ha while the root mean square error (RMSE) is 1.7 t/ha. That represents a relative mean absolute error (rMAE) 52% when compared with the weighted average yield of 2.9 t/ha (IBGE, 2015). This error is significantly higher

than other crop-growth models calibrated for the specific conditions of soybean production in Brazil. Battisti, Sentelhas, and Boote (2017) compared five different crop growth models for soybeans in Brazil, showing a significant increase in model accuracy when calibrated with local parameters. These locally calibrated models had an RMSE lower than 650 kg/ha, which is less than half of the error using the current version of the EPIC-IIASA model.

Conversely, the EPIC-IIASA model performs significantly better in the estimation of yields for second-season maize, with an MAE of 1.1 t/ha, which represents a rMAE of 20% if compared with the national weighted average of 5.5 t/ha. Duarte and Sentelhas (2020) experimented with three crop growth models in Brazil, namely AEZ-FAO, DSSAT-CERES-Maize, and APSIM-Maize, starting with default values and gradually increasing the locally calibrated parameters based on 79 experimental sites in Brazil. They found that the default calibration produced MAE as high as 6.5 t/ha but with extensive calibration, the MAE was reduced, reaching an average between 727 kg/ha and 1.37 t/ha. That highlights that EPIC-IIASA has a well-cali-brated cultivar parameters for second-season maize, with an MAE in the range of values found in locally calibrated models.

Sugarcane

In Brazil, the EPIC-IIASA model underestimates sugarcane yields in highly productive regions and overestimates it in low-yielding areas. Similar behavior is found by Marin et al. (2011) using the DSSAT/CANEGRO calibrated for Brazil. The current EPIC-IIASA has an MAE of 18 t/ha for fresh stalks, which is the same as found by Marin et al. (2015), as an average of the locally calibrated DSSAT/CANEAGRO and APSIM-Sugar results. The sugarcane results from EPIC also are satisfactory if compared with MAE results of ca. 15 t/ha from dos Santos Vianna and Sentelhas (2016) and > 29 t/ha from Dias and Sentelhas (2017).

An important shortcoming of the EPIC-IIASA is the lack of specific functions to estimate sugar content, which is crucial for the evaluation of the economic performance of sugarcane be-cause the payment systems are based mainly on sugar yield instead of fresh-stalk yield. Other models such as DSSAT/CANEAGRO and APSIM-Sugar already have such features (Marin et al., 2015).

Overall performance

Although the EPIC-IIASA model performs well for the yield estimation of second-season maize and sugarcane, the main concern is that the estimation error pattern differs strongly between the crops. If the crop growth model would overestimate (or underestimate) all crops in the study, the yield relationship between them would still be realistic. However, since EPIC-IIASA overestimates soybean yields while underestimating sugarcane's, there is expected to be a shift in the relative profitability of the cropping alternatives, artificially introduced by the crop growth model estimation. That needs to be addressed carefully when using the EPIC-IIASA as the biophysical model for studies of on-farm competitiveness with the PAM model.

Due to this error behavior, yields were corrected for this dissertation using available statistics from IBGE in Brazil (see 5.1 for more details). The yield correction should help illustrate the potential of the PAM economic module if a well-calibrated crop growth model is available. A key step for the further development of the PAM modelling approach is improving the bio-physical component, at least for Brazil.

8.1.2.   Improving the biophysical model estimation

The main challenge of calibrating a biophysical model (e.g., EPIC-IIASA) is that it requires a large number of parameters such as crop management, operating calendar, etc., that ideally should come from local trials with independent data. Such databases, however, are not avail-able on a global scale in a suitable resolution to allow the field-scale calibration of EPIC (Flach et al., 2020). Hence, IIASA has focused on the methodology developed by Balkovič et al. (2013), which allows a meaningful calibration based on the default values from EPIC adjusted to the studied cultivars and sensitivity analysis.

Conversely, local research groups can fine-tune the biophysical model including specific pa-rameters of local cultivars as well as typical farming operational schemes. For soybeans in Brazil, adapting the crop growth model for different seeding times as well as specific cultivar phenological characteristics improves the yield estimation considerably (Battisti et al., 2017). Battisti, Bender, and Sentelhas (2019) discuss the importance of representing the three main maturity groups for soybeans in Brazil, resulting in a crop cycle that varies between 110 and 130 days. The lack of such detailed local calibration may explain the poor performance of the EPIC-IIASA model for the estimation of soybean yields in Brazil. That is important to consider because the EPIC-IIASA may perform considerably better for soybeans in other countries, where the standard calibration fits the cultivar characteristics and practices used.

There is expected to be a trade-off between global scale versus local adaptability of the crop-growth models. The idea is that most models should perform reasonably if well calibrated to local conditions and practices. However, if the overall goal is to simulate yields at the SimU level covering the entire world, it is not feasible to calibrate the model to all characteristics of every single region and cultivar. Hence, depending on the scale of the research project and the question to be answered, different calibration strategies are required.

In the PAM modeling approach, the selection of the biophysical model depends on the scope of the project and the overall research question. The ability to use the SimU's structure from IIASA (GLOBIOM) is a major benefit of using the EPIC-IIASA model, ensuring that spatial reso-lution between the existing IIASA models and PAM is consistent. However, depending on the crop in question and the spatial coverage, it is possible to switch from the EPIC-IIASA model to more locally calibrated models. For this case study, using the EPIC-IIASA model was im-portant due to:

- the high spatial resolution and the coverage of all areas in Brazil,

- the ability to use the data from GLOBIOM-Brazil regarding transport costs,

- the ability to allocate national statistics to the SimU level,

- using EPIC's nutrient estimations for the cost calculations,

- the overall goal of using the PAM model in other countries.

The PAM modeling approach was developed to be easily adapted to carry out analysis on profitability in different countries and that is likely the main reason why the EPIC-IIASA model is used – global coverage. An important future development to improve the overall PAM modeling approach is to use the existing *agri benchmark* typical farms as the basis for the calibration of crop calendar and operations schemes in EPIC, working together with local partners to improve the calibration of phenological parameters as well. Such development, however, is time-consuming and cost-intensive which may motivate the use of alternative models for specific research questions as long as such joint development is under construction. The goal of this dissertation was primarily the development of the economic, transport and land-allocation modules based on the existing simulation from the EPIC-IIASA model. Further development of the PAM model should include improving the interaction between these modules and the plant module as well as fine-tuning the cultivar calibration of the crop growth model.

Finally, for sugarcane, the EPIC-IIASA model needs further development to move from a pure estimation of forage yield into the specific partition of above-ground biomass into stalk and leaves as well as the direct estimation of sugar content. Since cultivars differ in their ripening process and sugar concentration due to weather conditions (Cardozo & Sentelhas, 2013), it is important to refine the crop growth model to simulate the sugar yield, which is the main driver of revenues in sugarcane production.

8.2.   Farm-level profitability analysis

The main innovation within the PAM modeling approach is the estimation of production cost and farm profitability at a microregional level. Using existing typical farm information from the *agri benchmark* network as well as regionally available statistics, the economic module calculates the profitability of each cropping alternative at the SimU level. In this case study, profitability is measured using the return to land. In addition to production costs, the PAM approach has an endogenous mechanism to estimate farm-level input and output prices, which is a main advantage for countries with limited data availability. Since the economic module relies on important assumptions regarding production systems and regional characteristics, it is useful to focus on the impact of these assumptions in the model's performance as well as to discuss alternatives.

8.2.1. Estimation of production costs

Detailed production cost estimation often is not the focus of the main large-scale agro-economic models largely due to the complexity of estimating operating costs (labor, machinery, etc.) as well as due to the lack of regional data on a global scale. Besides European and North American models, which profit from large datasets on production costs (e.g., FADN), most of the models used in countries with poor data availability have to rely on assumptions, fixed cost structures, or some form of cost approximation (Heckelei & Britz, 2005). The exception appears to be the IIASA-ACM model developed to estimate production costs in a spatially explicit manner, in regions with low data availability (Deppermann et al., 2018). Hence, it is important to compare the assumptions and discuss the different techniques used by the IIASA-ACM and PAM models.

Extrapolation mechanisms

The basis for explicit modeling of production costs in high-resolution spatial models is likely to involve an extrapolation technique, starting from known costing information and estimating the main cost variables to each SimU. The IIASA-ACM and the PAM models use the same model structure (IIASA-EPIC) for yields and nutrient (N and P) use estimations, so these costs are already estimated in the desired spatial resolution (SimU). The main difference comes from the overall extrapolation mechanisms for the costs that are not based on EPIC's output.

The IIASA-ACM model uses an intensification factor (IF) that is calculated according to the difference between the current simulated yield and fertilization levels and the optimum level simulated by EPIC. The IF varies between 0 and 1 and is used to extrapolate plant protection, financing, machinery and building costs from the KBTL information to the SimUs. The remaining costs are kept constant for each management option (Deppermann et al., 2018). The PAM model uses a different strategy, mainly focusing on the regional characteristics of the farms and the crops grown.

Using operating costs as an example, for sugarcane most operations are carried out by contractors based on fresh-cane yields while grain farmers carry out most operations themselves, with their own machinery. That feature, combined with the structural differences – e.g., topography and field size – is expected to strongly impact the operating cost components in the different regions in Brazil. That highlights the importance of developing mechanisms that allow a tailored extrapolation of costs considering regional differences in production systems.

While the ACM's IF has its benefits, such as being able to rely on the KTBL information and extrapolate this information to all regions, it is challenging when considering costs such as plant protection. The IF for an extensive system (low input and yield ratios) would lead to lower spraying costs for the crop. One may argue that the technical application costs should be similar regardless of farming intensity, with the main variations in plant protection costs coming from the number of applications and product used. However, it is likely that extensive systems in regions with challenging environmental conditions still face high pressure of pest

and diseases, resulting in a high frequency of applications to be able to yield any output. Such structural differences become more evident if the model extrapolates European costs (KBTL) to tropical or subtropical regions, where the pressure from pests and diseases is structurally different from the basis data. In such cases, the use of the IF as the extrapolation mechanism may be more questionable.

The PAM model uses local information from the *agri benchmark* typical farms, considering characteristics of the most common production systems. That should reduce the constraints of basing the simulations on one specific (European) data source. However, it is still challenging to extrapolate the data to all SimU within a region. This local knowledge helps to cluster SimUs with similar farming conditions and to create a theoretical scheme in which costs are related to the known parameters such as yields and farm size or are considered fixed within the region. While this approach is expected to create a tailored extrapolation mechanism, it requires in-depth understanding of the regional characteristics and how they affect production costs. The main limitation of the PAM extrapolation approach is that creating extrapolation schemes for each region is time-consuming, which may limit its use for global-scale production-cost modeling.

Regardless of the model, a better understand of the drivers of the production costs – i.e., the connection to known parameters – is extremely important to improve the extrapolation mechanism and estimation performance. That is especially the case for costs such as plant protection, seeds, operating and management, which are not expected to have a linear relationship to yields or farm size. Further development in this area will certainly improve the performance of the PAM model in estimating production costs in high spatial resolution.

Using typical farms

Since the PAM model relies strongly on the production cost information from regional typical farms, it is important to consider the benefits and limitations of basing the cost estimation on them. One of the important advantages of the typical farms from the *agri benchmark* network is the ability to represent actual production systems in the main production regions in the country, covering the most important crops grown in each region. This local knowledge is crucial to allow the introduction of complex production systems (e.g., double cropping) in the model, avoiding farmers' options unrealistically depicted in the model – i.e., choosing between soybeans or maize instead of double-cropping.

Furthermore, the detailed data, including physical (e.g., N input) and monetary (e.g., N price) allows the disentangling of the usually available information in the format of crop budgets (e.g., $/ha) to each specific cost component – input quantity and price. That is especially important in the PAM approach since the nutrient input, for example, is separately simulated by the biophysical model, allowing PAM to use only the nutrient price from the typical farm.

Another important advantage of using typical farms from the *agri benchmark* network is the ability to exchange information with local experts, helping adapt the model to realistic

practices at the farm level as well as develop mechanisms for the cost extrapolation. In the case study, for example, the local expert in Brazil helped create the two plant protection schemes to depict extensive and intensive systems in Brazil. Moreover, the cost data is up-dated annually, making it possible to update the PAM simulations on an ongoing basis. Finally, since all typical farms are technically processed in the same way (i.e., questionnaire, calcula-tions), it is easy to adapt the PAM economic module to different countries.

One important consideration when using typical farm data for the PAM model is the ability to represent the main producing regions in the country. That requires a minimum number of typical farms to be able to create the costing spatial variability, ensuring that at least one typ-ical farm is available for each region with specific characteristics. The number of typical farms can be a problem depending on the country and the complexity of production systems. For Brazil, the current coverage of the *agri benchmark* network allows the representation of two main producing regions for grains (South and Center-west) as well as the sugarcane regions (*Traditional*, *Expansion* and Northeast). That can partly explain the relatively good perfor-mance of the model in depicting current production costs.

Conversely, for other countries, the availability of typical farms may be a constraint. Establish-ing typical farms is a time-consuming and cost-intensive activity, so each additional data point represents a high workload for national partners and therefore is limited to a reasonable amount. A limited number of typical farms also may be a problem since any strong variation in one typical farm may influence the whole cost estimation for that specific region, especially when only one typical farm is available. That may partly explain the relatively poor perfor-mance of the PAM model in estimating production costs in the South region in Brazil, where only one typical farm exists. It also is important to mention that having only one typical farm in a certain region limits the ability of the model to consider variations within the region for cost positions that are based strictly on the typical farm data (e.g., plant protection, seeds). The larger and more diverse the region, the larger the expected estimation error should be.

Finally, one important limitation of the typical farms approach is the ability to simulate pro-duction cost for crops that are not currently grown in the region. That is expected to be chal-lenging for studies looking into possible changes in crop rotations due to the inclusion of new crops that may become more profitable because of changes in the framework conditions – e.g., governmental programs. One short-term solution is to adapt the production costs for this new crop to the study region using data coming from other regions (and even countries), based on local expertise and on the input use from the crop growth model. For more long-term research projects, it is feasible to use focus groups to create a hypothetical costing struc-ture for the "new" crop or practice in question.

In a nutshell, the possibility of using typical farm production cost data is an important ad-vantage of the PAM model, allowing the inclusion of detailed regional characteristics into the model and the ability to use the local expertise of the national partners to adapt the model to different countries. A minimum number of typical farms, however, is required to depict the

diversity of the regions and to minimize errors coming from specific features of a single typical farm. The expansion of the network to more countries, regions, and crops will certainly increase the estimation performance and applicability of the PAM model.

Crop-specific cost elements

Another important feature of the PAM model is the built-in function to estimate inbound transport costs – i.e., from farm to mill – for sugarcane. That is a good example of required adaptations in the production cost estimation to correctly include all cost items that influence farmers' profitability and supply decisions. Depending on the research area and the prevailing cropping systems, it is important to adapt the economic module of the PAM model to include additional costs that have a different structure and cannot be directly derived from the typical-farm data. Following the example of inbound transport cost for sugarcane, the information from the typical farm is only a static monetary expense but, in reality, each SimU has a specific distance to an operating mill. The ability to use GIS software to measure the travel distance as well as having local information on transport costs, allows the PAM model to depict this cost component more realistically.

When such additional costs are dependent on infrastructure (e.g., presence of a mill), it is desirable to adapt the PAM model for different scenarios depending on the time horizon of the analysis. For short-term supply analysis of sugarcane in Brazil, the presence of a nearby mill is decisive for the profitability at the farm level. However, in the long run, the establishment of new mills should alleviate this constraint so the ability to switch on and off such cost components is key to simulate changes in different time horizons.

Production cost changes over time

An important consideration when analyzing long-term development in farm profitability and supply responses is expected changes in production costs. The PAM model has a static production cost estimation based on the current information of the typical farms as well as local statistics. For future scenarios – for example, on changes in the crop mix (e.g., single soybeans production instead of double-cropping) – the model is not able to account for the operational changes at the farm setup. That is mainly due to the difficulties in predicting such changes for future scenarios without relying on strong assumptions based on expert knowledge. Such information also is not available at a high spatial resolution without extensive investments in questionnaires and/or focus groups. These investments are not compatible with the idea of the minimum data approach, which is the basis for the development of the PAM model.

One also may argue that in the long-term (+20 years), input costs and farm endowments may change, affecting the results of basing the analysis on the current production cost information. In the PAM model, it is feasible to adjust such developments if they are known – for example adding a linear cost increase of diesel – but such adjustments are challenging because machinery also may develop in the period (e.g., improving the fuel use) which then would also need to be accounted for. Since the main idea is to understand how production costs and farm profitability drive the land allocation decision within the farm, one may argue that as long as

the cost changes are symmetrical among the crops, the overall competitiveness of the cropping alternatives may still be valid. It is clear, however, that if any cost item changes dramatically, affecting one crop more than another, the PAM model is not currently able to capture such distortions without further parametrization.

Finally, it is important to consider that if the research question is to understand how specific input use or price affects the on-farm competitiveness and resulting land use, it is feasible to combine focus group discussions within the *agri benchmark* network to gather the expected farm-level adjustments and introduce such changes in the PAM economic module. That is a clear benefit of estimating all cost components separately, allowing such adaptation for scenario analysis.

8.2.2.   Profitability analysis

The ultimate goal of the economic module is to estimate crop level profitability at SimU. Building on the production cost calculations, the PAM model also estimates the farm-gate prices of each commodity based on the transport module (see 4.5.5). Farm-gate prices are crucial for the revenue calculation so getting the correct parameterization of the model is essential. Revenues and production costs are finally used to calculate returns to land, which is the main profitability indicator in the PAM modeling approach. Since farm-gate prices are decisive for crop profitability, it is key to try to estimate them as realistically as possible.

Relevance of transport cost estimations

The PAM model attempts to improve the estimation of transport costs when compared with the GLOBIOM-Brazil model (Câmara et al., 2015) by: (a) empirically estimating the parameters of the transport cost functions based on local data; and (b) splitting the freight cost into its fixed and variable components. The former helps capture technical requirements of each crop, which may create transport cost differences that are not based on the travel distance. The latter should represent the reality of freight economics, in which SimUs with relatively short travel distances face a higher transportation cost per km due to the proportional importance of the fixed-cost components.

This improvement is possible only due to the access of exclusive data from crop-specific freight coming from more than 14,000 observations from ESALQ-LOG. In other countries, such detailed databases likely are not available so a more deterministic approach based on technical parameters may be an alternative. A possible path to further improve the transport cost estimation is to break down the freight components into individual costs (fuel, labor and equipment). In this way, it should be possible to simulate scenarios with increasing oil prices, for example, positively affecting sugarcane prices while increasing transportation costs (i.e., fuel). Such a level of detail, however, requires a thorough understanding of the local transport industry conditions and access to large local databases, which may not be feasible for most regions.

Additionally, it should be possible to include the GLOBIOM-Brazil information on the quality of roads to further improve the transport cost estimation. That is expected to be important for remote regions where road infrastructure is poor (e.g., Amazon basin), changing the overall transport cost function. Such a combined approach needs careful implementation, which could be achieved by estimating several cost parameters (see section 4.4.1) according to the specific route (origin and destination) and differentiating the cost parameters depending on the geographical location of the SimU.

Besides the freight cost estimations, the model also needs to estimate the travel distance between the SimU and the destination (e.g., port or domestic market). It is a major benefit of the PAM model for Brazil to rely on the estimation approach from Aguiar et al. (2003), which uses the actual road network to simulate the travel distances for each SimU. For countries such as Brazil, the availability of roads differs strongly among regions so a pure straight-line (Euclidean) distance estimate may lead to underestimation of actual travel distance in more remote regions.

Most geographic studies use the straight-line distance due to the relatively easy calculation. If the PAM model should be implemented in a country with lower availability of the road network, one could follow the work from Boscoe, Henry, and Zdeb (2012), which uses a detour index of 1.417 to convert the straight-line distance estimation (GIS software) to an approximated travel distance. The authors were able to find a R² of 0.94 when comparing corrected straight-line to actual travel distances[42]. One important caveat of this approach is that outliers (e.g., Alaska, Amazon Forest) increase the error of such estimation techniques considerably.

In summary, improving transport cost estimation is essential for a robust estimation of farm profitability (i.e., revenues) for export-driven countries. Including crop-specific freight-cost parameters and splitting the cost component into variable and fixed costs should increase the model's performance. When it comes to estimating the travel distance, if actual travel distances can be directly estimated, it is certainly preferable. In other cases, using the easy-to-estimate straight-line (Euclidean) distance with a detour index may generate satisfactory results, at least for areas with relatively normal road infrastructure.

The complexity of sugarcane and other processed crops

Using the transportation cost module to convert crop reference price to farm gate is a useful and flexible approach that is expected to be suitable for most of the agri-commodities. However, for some specific crops such as sugarcane, a more complex farm-gate price estimation process is needed since the farm output (i.e., fresh sugarcane) must be processed before commercialization. That means that the farm-gate price of the final products (i.e., ethanol and sugar) are the determinants of the sugarcane revenue. Increasing the complexity is the fact

---

[42] Data from 66,011 routes for the fifty states of the United States, District of Columbia, and Puerto Rico - Boscoe et al. (2012).

that, at least in Brazil, almost the entire ethanol production is consumed domestically whereas, for sugar, the majority of the output is exported.

The PAM model for Brazil has a built-in structure to estimate mill-level sugar and ethanol prices as well as using regional statistics to infer the mix of production for each region (see 4.5.5.2). This feature is important since the competitiveness of ethanol and sugar varies regionally due to logistics costs and tax incentives, allowing the model to assess the realistic resulting sugarcane price for each region. This detailed breakdown of the components of sugarcane prices makes the model suitable to analyze scenarios in which only one of the outputs changes – e.g., a governmental program to foster biofuels, that is expected to affect regions differently depending on the competitiveness between sugar and ethanol.

The example of sugarcane is only one among the crops that may need more complex farm-gate price estimation systems, depending on the characteristics of the value chain in which they are produced. Adapting the model to such specific features is time-consuming and requires in-depth local expertise, which may not be possible for global-scale models. Against this background, the ability to use the *agri benchmark* network of local experts is a key advantage of the PAM model, allowing a local adaptation of the model to the specific conditions of these special crops worldwide.

8.2.3.   Land allocation strategy

Based on the profitability of each farming alternative at the SimU level, the PAM land allocation module selects the alternative with the highest return to land. This module accounts for the most common production systems or combinations of crops to avoid results that contradict the reality of regional farms. For example, in Brazil, soybeans and maize do not compete for land in most regions since farmers can produce both in a double-cropping system (Pires et al., 2016).

One important limitation of the current version of the PAM model is the allocation of the entire SimU to only one cropping alternative. That is not a major constraint in Brazil because farms are specialized, but for more diverse production systems (i.e., complex crop rotations) an adaptation of the land allocation module is necessary. Operational constraints such as peak working hours, preceding crop value and the maximum share of a single crop in the rotation are a few examples of technical restrictions that need to be implemented in the PAM model for regions with more complex cropping systems. Another development could be considering multi-year rotations (e.g., wheat-sugar beets-wheat-rapeseed) as cropping alternatives instead of single crops. That is expected to be more in line with producing patterns from farmers elsewhere in the world since the allocation of the area to a single crop is usually restricted by rotational constraints as well as avoiding overspecialization in supply responses (Chen & Önal, 2012).

Another limitation of the land allocation module is the lack of explicit modeling of the conversion costs incurred by switching between alternatives (Zhao, Calvin, & Wise, 2020). In the current version of the PAM, any economic advantage leads to the allocation of the SimU to the specific cropping alternative. However, the reality is that farmers are likely to consider the costs involved in changing between alternatives such as the investment in new equipment, the depreciation of idle machinery, expenses into learning the "new" crop, etc. Hence, the land allocation module should evolve to account for current land use and use a technical parameterization to try to mimic the expenses incurred to farmers if they wish to switch between crops. Such conversion costs are even more important if the model is used to analyze the potential conversion of non-agricultural land to cropland. Costs of converting pasture to arable land are expected to be more significant than the conversion between sugarcane and grains, for example. In Brazil, other important constraints that may be included in the land allocation are the forest code and soy moratorium, which limit the extent of conversion of forest to arable land (Soterroni et al., 2018; Soterroni et al., 2019).

Besides the technical conversion costs, it is important to highlight that short-term price movements (1 to 2 years), for example, may not cause the land-use changes that the PAM model would estimate. The model does not account for other factors such as the farmers' risk perception and their knowledge that commodity prices usually fluctuate. For a significant regional land-use change, it is expected that structural changes in framework conditions are required, which are reasonable to estimate using PAM for a time horizon between 5 to 20 years. That time period may change depending on the characteristics of the crops (annual vs perennial) and the conceptual analysis on how much the current production system (typical farm data) could change in the future.

The current version of the PAM model has the advantage of using regional statistics to determine the share of second-season maize, which is an interesting example of how to include additional country or region-specific information to improve the performance of the model in representing current land use. Moreover, by using GIS software, several constraints are included such as limiting results to current arable land only or grassland, depending on the research question to be addressed.

A future development of the PAM model may be connecting the results from the economic module (i.e., returns to land for each cropping alternative) to a more developed land allocation and optimization model such as the GLOBIOM. The integration should be feasible since both models use the same spatial resolution (SimU). An additional advantage would be profiting from the extensive parameterization work carried out by IIASA to adapt the GLOBIOM for the Brazilian agricultural sector, by including other activities such as forestry and animal husbandry, for example (Zilli et al., 2020). Connecting the two models also should help improve the overall performance of the land-use simulation for future scenarios since the PAM model does not have an internal demand function to balance out how land expansion caused by improved profitability may trigger changes in price ratios, which needs to be accounted for. If, however, the goal of the research is to simulate current land use based on the assumption

that farmers optimize their returns to land, the current version of the PAM land allocation module is expected to provide meaningful results.

8.3.   Case study Brazil: Land-use change

Empirically, the newly developed PAM model is tested for Brazil, aiming to understand current and future land uses. The focus of this work is on the main crops – i.e., sugarcane, soybeans and maize, which together account for ca. 84% of the total seeded area in Brazil (IBGE, 2019). The production cost results from the PAM model are compared with an exclusive database of typical farms coming from CEPEA and PECEGE in Brazil while the land allocation (land-use) results are validated against municipality level data from IBGE (2019) and the results from the SPAM model (IFPRI, 2019). In the next subsections, I discuss the results from the case study in Brazil, focusing not only on the current profitability and land use but also on expected changes coming from the impact of climate change on yields in the different climatic regions.

8.3.1.   Profitability of crops

The basis for the land-use analysis of the case study in Brazil is the idea that farmers allocate their resources (i.e., land) to optimize their economic returns. Within the PAM model, profitability is measured based on the return to land, accounting for all production costs except land.

Yields of grain and sugarcane in Brazil

A key driver of a crop's economic performance is yield. The results show a weighted average[43] of sugarcane yields in Brazil of 74 t/ha of fresh cane, while soybeans and second-season maize have an average of 2.9 t/ha and 5.5 t/ha, respectively. Regionally, the yield variation is significant, with state-level averages varying from 45 t/ha up to 79 t/ha for sugarcane, for example. One interesting proxy for the competitiveness of crops within a region is the yield ratios. It indicates the relative advantage of the crop over the competing alternatives and, for an on-farm analysis, it is more important than the absolute difference among the regions (Egli, 2018).

The largest sugarcane-producing state (São Paulo) and the largest soybean-producing state (Mato Grosso) have yield ratios from sugarcane/soybeans of 27 t/ha and 23 t/ha, respectively. The higher the ratio, the higher is the competitiveness of sugarcane over soybeans since by producing one ton of soybeans in São Paulo, growers forgo the return of producing 27 t/ha sugarcane instead. The same ratio is 14% lower in Mato Grosso, showing that based solely on yields, São Paulo has a higher competitive advantage in producing sugarcane instead of

---

[43] All averages discussed in this section are weighted average based on the harvested area of the crop in the SimU in relation to the total harvested area in 2015. Harvested area data are from IBGE (2015) and are allocated to the SimU based on the methodology from Câmara et al. (2015).

soybeans. A closer look at the state of Rio Grande do Sul (RS) shows a sugarcane-to-soybeans ratio of 16 t/ha, which partly explains why this state has negligible sugarcane production[44] while it is the second-largest producing state of soybeans in Brazil (CONAB, 2021). It is important to highlight that the yield ratios differ much more when looking at the SimU level, which is an indication of the necessity of running on-farm competitiveness analysis at micro-regional levels.

Against this background, it is expected that changes in the yield ratios caused by technical progress (i.e., breeding), new pests and disease, or long-term changes in climatic conditions (e.g., climate change) will affect the intraregional competitiveness of sugarcane and grains and may lead to changes in land use.

Production costs and their impact on farm-level competitiveness

Even though yield ratios are a good indication of the on-farm competitiveness of crops, understanding differences in production costs is crucial due to their significant impact on the economic performance of the cropping alternatives. The results show that sugarcane has 45% higher production costs than the double-cropping of soybeans and maize. More interestingly is that, for sugarcane, operating costs represent 56% of the average total costs, while for grains, the main component is establishment costs, accounting for ca. 60% of the total costs. This result highlights the importance of understanding the building blocks of the total costs because changes in framework conditions such as exchange rate and ban of plant protection products are expected to affect the production cost of each crop differently.

Another important consideration is how certain cost components may affect farmers' perception of costs. Operating costs are a good example when comparing sugarcane to double-cropping in Brazil. While more than 40% of the operating cost for sugarcane is represented by contractors, in which the fees are seen as direct cash expenses, grains have roughly 70% of operating costs as machinery and diesel, which have a much higher share of "hidden" cost such as depreciation, capital, etc. This difference in cost perception is expected to influence farmers' supply responses since they may not consider depreciation and their own capital costs during acreage planning. Such cost perceptions are not included in the PAM model, but they are important to understand why farmers may behave differently than the model estimations suggest.

One key determinant of the cost competitiveness of sugarcane is the inbound transport costs. The PAM model estimates a weighted average of 162 USD/ha, which represents ca. 9% of the total costs. Considering average yields, these costs can be converted to an average travel distance to the mill of 20 km. This distance is higher than observed in the literature of 25 km (agri benchmark, 2020; Françoso et al., 2017). This difference is expected to be the result of:

---

[44] According to CONAB (2021), Rio Grande do Sul produced 31,000 t of sugarcane in 2019/20, which represents 0.005% of the national output.

- Difficulty of estimating actual travel distance from straight-line (Euclidean) distance as already discussed, which should be even more challenging for off-road transportation.

- The assumption that the SimU delivers sugarcane to the nearest mill, which may not be the case for all regions. Farmers in reality, may deliver cane to different mills depending on their supply contracts as well as the total processing capacity of the mills.

In a nutshell, I argue that it is important to understand the building blocks of the production costs since changes in the framework conditions such as exchange rate may affect cost components of the cropping alternatives differently. Another important consideration is the effect of cost perceptions affecting supply decisions depending on how farmers perceive costs as direct cash expenses or hidden costs.

Model performance in estimating production costs

The model validation shows that the PAM model has satisfactory accuracy to estimate production costs, with rMAE lower than 14% for all crops (see section 6.2.3). Regionally, the model performs better in the *Center-south* region for sugarcane and in the *Expansion* region for grains. One likely explanation for this error pattern is the lack of typical farms in the *Northeast* region for sugarcane and in the *Traditional* region for grains. Due to the methodology, if only one typical farm is available in a region, any particularities of this farm affect the production cost estimation significantly. Hence, efforts to improve the local coverage (e.g., including more typical farms) in these regions should improve the model's performance.

Validating the production cost estimation is a key step in improving model performance and having the database from CEPEA and PECEGE for Brazil was crucial for the PAM validation. However, the regional distribution of the validation typical farms raises the following concerns for the interpretation of the results:

- Some regions and states have only one typical farm, which may limit the robustness of the validation results.

- It is not possible to assess the accuracy of the model in areas where there are no typical farms to benchmark results.

Considering the idea of minimum data, the scarcity of data and the challenges of estimating production costs (Antle & Valdivia, 2006), the validation results suggest that the PAM model is accurate for the estimation of production costs in the main production regions in Brazil. Improvements in regional adaptations of the model through the establishment of new typical farms are expected to further improve the model. Such adaptation requires ongoing resources for the establishment and updating of typical farms and therefore it is important to consider whether the additional model accuracy justifies such investments.

8.3.2.   Estimating Brazilian land use

By spatially modeling the land allocation of these crops at the microregional level, the PAM model attempts to partially explain the current land use in Brazil. The overall performance of the land allocation module (based on the matching score approach) is ca. 86%. That means that the PAM output matches (i.e., same cropping alternative) in 86% of total arable land (640,000 pixels), when compared with simplified land-use maps from IBGE and the SPAM model. For important production states, the accuracy is above 88%, reaching as high as 95% in Mato Grosso.

Such a high level of agreement between current land use and the output from the PAM model, in part, is unexpected considering that the model bases the land allocation decision entirely on the crops' profitability. The results suggest that Brazilian farms are indeed optimizing their economic returns and that underlying features of a region such as yields, costs and price ratios appear to favor a cropping alternative over others. The land-use results reinforce the leading assumption of the economic-based LUC models that the local agent's decisions are based on profit maximation (Dang & Kawasaki, 2016).

An interesting pattern of land use in Brazil is the concentration of sugarcane around the main producing state (São Paulo), creating a production cluster. One may argue that the concentration of sugarcane around São Paulo in the PAM results is due to the inbound transport costs (from farm to mill), which increases the returns to sugarcane with high mill infrastructure. This argument is valid to a certain extent but even when this constraint is removed (i.e., considering average inbound transport costs), the clustering effect around São Paulo still prevails (see section 5.5.2). That is an important indication that other features such as output prices and yield ratios tend to favor sugarcane in this region. One surprising finding from removing the constraints of having a mill nearby is the allocation of sugarcane to several areas in the North of Brazil (MAPITOBA region), indicating that at least from the economic returns' perspective, sugarcane is competitive, if milling infrastructure is available.

The land allocation module performs relatively poorly in states with a more diverse production patterns, located on the boundaries of the sugarcane clusters. An example is the state of Minas Gerais, which has 21% of the state in sugarcane and 36% in soybeans, whereas PAM matched only ca. 60% of the pixels. PAM allocates most of the cropping area of the state to sugarcane while the simplified land-use maps show a balance between sugarcane and soybeans. In these regions, with high competition between sugarcane and soybeans, it is challenging to precisely estimate land allocation strategies, because farmers that have been producing grains, for example, could be reluctant to switch to sugarcane due to the lack of experience, perception of risk and the current farm setup (Chen & Önal, 2012).

One technical challenge of the land-use validation process for more diverse regions is the spatial resolution differences between the output from PAM and the data used to create the validation maps – i.e., IBGE and SPAM. The IBGE data, for example, is on a municipality level, and

in these diverse regions, several municipalities grow sugarcane and soybeans. In creating the simplified land-use maps, I allocated all the municipality area to the crop with the highest share of acreage to avoid biased allocation (see section 6.1.3). Therefore, it is likely that knowing more precisely the actual production location (spatially) instead of relying on the administrative boundaries would allow a more detailed validation of the PAM results. Such information is not available, however, with most models (e.g., SPAM) allocating the administrative information to match pre-defined constraints when satellite imagery, for example, is not available (IFPRI, 2019).

Besides technically improving the land allocation module as discussed in section 8.2.3, adding crops that have a regional importance such as rice, beans and cotton is expected to increase the performance of the model in reproducing regions with more diverse land use. The common bottleneck for the introduction of new crops is the ability to estimate production costs. EPIC already is calibrated for all these crops but, due to their limited importance, they currently are not covered within *agri benchmark* typical farms. Besides increasing the cropping portfolio, it is desirable to include other sectors, such as forestry and livestock, to be able to represent the agricultural sector more holistically. Such development fits with the suggested approach of combining the PAM model with GLOBIOM, for example, which already covers 18 crops, 5 forestry and 6 livestock products in Brazil (Câmara et al., 2015).

In summary, the findings show an 86% match between the land-use map from PAM and independent land-use validation maps. Sugarcane is allocated mainly to a cluster around the São Paulo state, indicating the relative economic competitiveness of this crop in the region. States with more diverse land use show a lower agreement between PAM model results and validation maps, which may be explained by the close competitiveness of both cropping alternatives and by technical challenges arising from the validation methodology.

8.3.3.   Climate change impact on land use in Brazil

The climate change adaptation scenarios proposed in this dissertation are based on the expected yield changes caused by modifications in the rainfall distribution, temperature and CO_2 concentration, affecting the competitiveness of the cropping alternatives. The two proposed scenarios cover changes in the current production systems for grains: (a) moving to single soybean cultivation (i.e., NoDC), or (b) continuing with the double-cropping system but accepting a yield penalty for second-season maize (DCYp). Conversely, sugarcane yields are assumed to increase due to higher CO_2 concentration and temperature, compensating for the reduction in precipitation. The expected yield changes are derived from the literature based on models calibrated for Brazil (Camilo et al., 2018; Hampf et al., 2020; Marin et al., 2013; Pires et al., 2016).

Empirical discussion of the climate change scenarios

The scenario results show a 39% reduction of returns to land for grain production in the DCYp scenario and a 26% reduction in the NoDC scenario in the SimUs located in the tropical region. That suggests a strong impact on the current double-cropping systems in the tropical region in Brazil, where farmers are expected to forgo the maize economic returns to focus on long-cycle soybeans. This effect should lead to a significant reduction in maize supply from Brazil due to climate change, considering that the majority of current production is under a double-cropping system.

The resulting land-use change is significant, with more than 24% of current arable land moving from grain to sugarcane production in both adaptation scenarios. These results are in line with the work from Zilli et al. (2020), showing a reduction in soybean production between 6.3% and 36.5% and maize, 12.9% and 29.4%, depending on the climate change scenario. The authors estimate that in the eastern Cerrado and MAPITOBA regions, livestock is replacing grains and in the Southern regions, sugarcane. That is a similar spatial land-use change observed in my results, with the most significant difference being that the PAM model sees an expansion of sugarcane instead of livestock in the MAPITOBA region. This effect may be explained by the assumption that milling infrastructure is not a constraint in the climate change adaptation scenarios, which may not be the case in the work of Zilli et al. (2020).

Technically, however, it is possible that farmers may react to such drastic yield reductions by changing sowing dates and using maize cultivars that are more adapted to a drier climate. The latter is likely to be fostered by breeding companies reacting to changes in climatic conditions. To a certain extent, such technical changes should alleviate the impact of climate change on maize yields, as discussed by Hampf et al. (2020), but it is uncertain whether such a yield gap can be closed entirely by technical innovations.

Considering that the most important constraint to the second-season maize yield in the climate change scenario is the availability of water, one alternative may be investing in more irrigation systems that could alleviate this effect. Although ANA (2020) estimates that Brazil could reach 12.4 million ha of irrigation area in 2040, it is uncertain whether it is economically viable to irrigate maize in the Brazilian tropical region, considering that other high-value crops may be more interesting[45]. Future scenario analysis may include the yield changes considering the implementation of irrigation systems. For that, however, the PAM economic module needs to be adapted to simulate irrigation costs, which currently are not a common practice depicted by typical farms.

---

[45]   Currently, 44.5% of the irrigated area is cultivated with sugarcane (3.65 Mha), 15.9% with rice (1.3 Mha) and 5.5% with coffee (0.45 Mha), showing that 66% of the current irrigated area is used for high-value crops – ANA (2020).

A conceptual discussion of the scenario analysis

The climate change scenario shows the flexibility of the PAM model, allowing simulation of farmers' supply responses incurred due to technical changes such as yield development and the more common shocks due to price movements. In addition, the model allows the adaption of the costing structures to capture possible farming practice changes such as a ban on plant protection products. However, it is important to consider the following:

(1)   The lack of demand function limits the ability of the PAM model to capture the new market equilibrium. Such drastic yield changes, for example, are likely to impact overall commodity prices, changing the baseline equilibrium (prices). A possible way forward may be integrating the PAM model with more complex general or partial equilibrium models to be able to capture market changes endogenously.

(2)   Implementing the expected climate changes directly into the biophysical model should generate more reliable results than changing yields ex-post, based on assumptions. Due to the limited accuracy of the current version of IIASA-EPIC in Brazil, intense calibration work and adaptations are necessary for such implementation, which is out of the scope of this dissertation.

(3)   For land-use analysis, it is desirable to include more sectors and crops in the model to be able to mimic complex framework changes. The significant change in yields may lead to farmers moving to different crops that currently are not included in the model. The lack of crops and sectors reduces the ability of the model to capture all possible dynamics.

(4)   An interesting approach to model farmers' reaction to scenario analysis is to run focus group discussions with regional farmers and advisors. By confronting the participants with the theoretical outcomes of the model, it should be possible to obtain a more robust set of adaptation options that then could be implemented in the model. Such multi-stage analysis requires on-the-ground research, which was not possible during the work on this dissertation[46].

Conclusions

Climate change is expected to affect the current cropping systems in the Brazilian tropical region mainly by shortening or delaying the onset of the wet season. That reduces the ability of farmers to use the common double-cropping system with soybeans and maize. The results indicate that farmers are likely to forgo maize returns and move to a single soybean cultivation (with cover crop). In addition, the model estimates that 24% of the current grain area would move to sugarcane, based solely on the returns to land. Internationally, such developments are important since ca. 77% of the total maize output from Brazil comes from the second-

---

[46] Focus group discussions to validate the scenario outputs and possible adaptations of farmers were planned for 2020 and 2021 in Brazil. However due to the COVID-19 pandemic, on-the-ground research was not possible within this timeframe. See WHO (2021) for more information.

season crop (CONAB, 2021), which should lead to an important supply shock globally. Technical adaptation is likely to occur, with farmers changing sowing dates, using more suitable cultivars and investment in irrigation, if economically viable. These adaptations are expected to alleviate the impact of climate change and reduce the magnitude of estimated land-use changes.

The PAM model is a flexible tool allowing the supply response simulations incurred due to technical changes (e.g., yields, inputs) as well as external price movements due to its internal estimation of farm-gate prices. However, the model has important limitations when it comes to balancing the supply changes with demand reaction as well as missing some important agri-sectors such as livestock and forestry. Further model parametrization, as well as its combination with existing macro-level models should improve the model's performance and reduce the uncertainty of results.

8.4.   The PAM model applications

The main idea behind the development of the PAM model is using a combination of a biophysical model and farm-level economic data to estimate profitability and to simulate land allocation strategies of farmers. An important goal of the model is also to be able to rely on existing data, following the idea of a minimum data approach, which should provide timely and relevant information to decision-makers with satisfactory accuracy (Antle & Valdivia, 2006).

Based on the results from the case study in Brazil, I argue that the PAM modeling approach can provide reliable information on farming profitability and land allocation strategies in countries with limited data availability, thereby answering the research question. The performance and applicability of the PAM model, however, depend on the following requirements (not a comprehensive list):

- Availability of a well-calibrated biophysical model for the yield and nutrient use estimations (e.g., EPIC, DSSAT, APSIM).

- The presence of at least one typical farm with production cost information on the main crops in the region (e.g., *agri benchmark* database). Having more typical farms is expected to improve model accuracy and reduce uncertainties.

- Information on transport costs as well as travel distance between the SimU and the crop destinations. The former is more challenging to obtain while the latter could be approximated by using straight-line distances and a detour factor (Boscoe et al., 2012).

- Local expertise on the actual cropping alternatives and systems of farmers in the different regions (e.g., *agri benchmark* national partners). That is essential to avoid unrealistic parametrization of the model and to ensure that all costs are considered in the profitability analysis.

Considering the global coverage of the *agri benchmark* network, with information not only on crop farmers but also on livestock (Chibanda et al., 2020), as well as the broad crop coverage of the IIASA-EPIC (Balkovič et al., 2013), the PAM model has a strong potential to simulate micro-level profitability and land allocation strategies based on on-farm competitiveness. It should be an interesting addition to the already existing macro-economic model groups such as the Thünen Modelling Network[47], benefiting from the complex interactions at the macro-level already simulated by the existing models while adding the micro-level profitability component for regions where data availability is limited (i.e., outside of Europe).

Due to its bottom-up approach, the PAM model can capture a variety of technical, policy and market changes such as the ban of agricultural inputs, increase in global commodity prices the introduction of biofuel policies. These changes affect one or all components (i.e., yields, costs, prices) of farm profitability, which can be simulated by the PAM model at high spatial resolution. The model flexibility is an important feature that should help adapt the model to answer a variety of research questions in a timely manner and provide reliable information to decision-makers. Nonetheless, the model has important limitations such as the lack of a demand function and an internal mechanism to cope with complex crop rotations. These issues need to be carefully considered for utilization of the model as well as for further development of the modeling approach.

---

[47]   https://www.thuenen.de/en/infrastructure/the-thuenen-modelling-network/

References

ABIOVE (2021). *Brazilian Association of Vegetable Oil Industries: Monthly export statistics*. Sao Paulo. Retrieved from https://abiove.org.br/en/statistics/

Abrahão, G. M., & Costa, M. H. (2018). Evolution of rain and photoperiod limitations on the soybean growing season in Brazil: The rise (and possible fall) of double-cropping systems. *Agricultural and Forest Meteorology*, *256-257*, 32–45. https://doi.org/10.1016/j.agrformet.2018.02.031

Agri benchmark (2020). *Result Database*. Braunschweig.

Aguiar, A. P., Câmara, G., Monteiro, A., & Souza, R. C. (Eds.) (2003). *Modeling Spatial Relations by Generalized Proximity Matrices*. Retrieved from http://www.geoinfo.info/portuguese/geoinfo2003/papers/geoinfo2003-11.pdf

Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, *22*(6), 711–728.

ANA (2020). *Irrigation Atlas: Atlas Irrigação*. Brasilia. Retrieved from http://atlasirrigacao.ana.gov.br

ANDA (2020). *Main indicators of the fertilizer sector*. Retrieved from Associação Nacional para Difusão de Adubos website: https://anda.org.br/

ANP (2020). *Time series of fuel prices: Série histórica do levantamento de preços*. Brasilia. Retrieved from http://www.anp.gov.br/precos-e-defesa-da-concorrencia/precos/levantamento-de-precos/serie-historica-levantamento-precos

Antle, J. M., Diagana, B., Stoorvogel, J. J., & Valdivia, R. O. (2010). Minimum-data analysis of ecosystem service supply in semi-subsistence agricultural systems. *Australian Journal of Agricultural and Resource Economics*, *54*(4), 601–617.

Antle, J. M., & Valdivia, R. O. (2006). Modelling the supply of ecosystem services from agriculture: A minimum-data approach *Australian Journal of Agricultural and Resource Economics*, *50*(1), 1–15.

Arnberg, S., & Hansen, L. G. (2012). Short-run and long-run dynamics of farm land allocation: panel data evidence from Denmark. *Agricultural Economics*, *43*(2), 179–190. https://doi.org/10.1111/j.1574-0862.2011.00575.x

Arvor, D., Dubreuil, V., Ronchail, J., Simões, M., & Funatsu, B. M. (2014). Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in Mato Grosso (Brazil). *International Journal of Climatology*, *34*(8), 2622–2633. https://doi.org/10.1002/joc.3863

Balieiro, S., Witte, T. de, & Weerathaworn, P. (Eds.) (2016). Using production-cost analysis to understand the competitiveness of sugarcane production: a comparison among Thailand, Vietnam, South Africa and Brazil. Thailand: ISSCT - International Society of Sugar Cane Technologists.

Balkovič, J., van der Velde, M., Schmid, E., Skalský, R., Khabarov, N., Obersteiner, M., . . . Xiong, W. (2013). Pan-European crop modelling with EPIC: Implementation, up-scaling and regional crop yield validation. *Agricultural Systems*, *120*, 61–75. https://doi.org/10.1016/j.agsy.2013.05.008

Balkovič, J., van der Velde, M., Skalský, R., Xiong, W., Folberth, C., Khabarov, N., . . . Obersteiner, M. (2014). Global wheat production potentials and management flexibility under the representative concentration pathways. *Global and Planetary Change*, *122*, 107–121. https://doi.org/10.1016/j.gloplacha.2014.08.010

Barona, E., Ramankutty, N., Hyman, G., & Coomes, O. T. (2010). The role of pasture and soybean in deforestation of the Brazilian Amazon. *Environmental Research Letters*, *5*(2), 24002. https://doi.org/10.1088/1748-9326/5/2/024002

Battisti, R., Bender, F. D., & Sentelhas, P. C. (2019). Assessment of different gridded weather data for soybean yield simulations in Brazil. *Theoretical and Applied Climatology*, *135*(1-2), 237–247. https://doi.org/10.1007/s00704-018-2383-y

Battisti, R., Sentelhas, P. C., & Boote, K. J. (2017). Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil. *Field Crops Research*, *200*, 28–37. https://doi.org/10.1016/j.fcr.2016.10.004

Berg, A., Noblet-Ducoudré, N. de, Sultan, B., Lengaigne, M., & Guimberteau, M. (2013). Projections of climate change impacts on potential C4 crop productivity over tropical regions. *Agricultural and Forest Meteorology*, *170*, 89–102. https://doi.org/10.1016/j.agrformet.2011.12.003

Binswanger, H. (1989). The Policy Response of Agriculture. *The World Bank Economic Review*, *3*(suppl_1), 231–258. https://doi.org/10.1093/wber/3.suppl_1.231

Borchers, A., Ifft, J., & Kuethe, T. (2014). Linking the Price of Agricultural Land to Use Values and Amenities. *American Journal of Agricultural Economics*, *96*(5), 1307–1320. https://doi.org/10.1093/ajae/aau041

Bordonal, R. d. O., Carvalho, J. L. N., Lal, R., Figueiredo, E. B. de, Oliveira, B. G. de, & La Scala, N. (2018). Sustainability of sugarcane production in Brazil. A review. *Agronomy for Sustainable Development*, *38*(2). https://doi.org/10.1007/s13593-018-0490-x

Boscoe, F. P., Henry, K. A., & Zdeb, M. S. (2012). A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals. *The Professional Geographer : The Journal of the Association of American Geographers*, *64*(2), 10.1080/00330124.2011.583586. https://doi.org/10.1080/00330124.2011.583586

Britz, W., & Arata, L. (2019). Econometric mathematical programming: an application to the estimation of costs and risk preferences at farm level. *Agricultural Economics*, *50*(2), 191–206. https://doi.org/10.1111/agec.12476

Câmara, G., Soterroni, A., Ramos, F., Carvalho, A., Andrade, P., Souza, R. S., . . . Pena, M. (2015). *Modelling land use change in Brazil: 2000–2050*. São José dos Campos, Brasília, Laxenburg, Cambridge. https://doi.org/10.22022/REDD/08-2016.12115.

Camilo, J. A., Andrade, C. L. T., Amaral, T. A., Tigges, C. H. P., Melo, M. L. A., Chan, C. S., & Garcia, A. G. (Eds.) (2018). *Impact of Climate Change on Maize Grown in the Brazilian Cerrado*. USA: ASABE Annual International Meeting. Retrieved from https://elibrary.asabe.org/abstract.asp?aid=49441

Cardozo, N. P., & Sentelhas, P. C. (2013). Climatic effects on sugarcane ripening under the influence of cultivars and crop age. *Scientia Agricola*, *70*(6), 449–456. https://doi.org/10.1590/S0103-90162013000600011

Carter, H. O. (1963). Representative Farms: Guides for Decision Making? *Journal of Farm Economics*, *45*(5), 1448. https://doi.org/10.2307/1236842

Carvalho, A. L. de, Menezes, R. S. C., Nóbrega, R. S., Pinto, A. d. S., Ometto, J. P. H. B., Randow, C. von, & Giarolla, A. (2015). Impact of climate changes on potential sugarcane yield in Pernambuco, northeastern region of Brazil. *Renewable Energy*, *78*, 26–34. https://doi.org/10.1016/j.renene.2014.12.023

Cattelan, A. J., & Dall'Agnol, A. (2018). The rapid soybean growth in Brazil. *Embrapa Soja-Artigo Em Periódico Indexado (ALICE)*.

CEPEA (2019). Production cost of typical farms in Brazil - CNA. Piracicaba.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250.

Chavas, J.-P., & Cox, T. L. (1995). On nonparametric supply response analysis. *American Journal of Agricultural Economics*, *77*(1), 80–92.

Chen, X., & Önal, H. (2012). Modeling Agricultural Supply Response Using Mathematical Programming and Crop Mixes. *American Journal of Agricultural Economics*, *94*(3), 674–686. https://doi.org/10.1093/ajae/aar143

Cherubin, M. R., Carvalho, J. L. N., Cerri, C. E. P., Nogueira, L. A. H., Souza, G. M., & Cantarella, H. (2021). Land Use and Management Effects on Sustainable Sugarcane-Derived Bioenergy. *Land*, *10*(1), 72.

Chibanda, C., Agethen, K., Deblitz, C., Zimmer, Y., Almadani, M. I., Garming, H., . . . Lasner, T. (2020). The Typical Farm Approach and Its Application by the Agri Benchmark Network. *Agriculture*, *10*(12), 646. https://doi.org/10.3390/agriculture10120646

Climatempo (2020). *CLIMATOLOGIA*. Retrieved from Climatempo website: https://www.climatempo.com.br/

Colman, D. (1986). A Review of the Arts of Supply Response Analysis. *Review of Marketing and Agricultural Economics*, *51*(3), 201–230.

CONAB (2019). *Sugarcane Crop Bulletin*. Brasilia. Retrieved from CONAB website: https://www.conab.gov.br/info-agro/safras/cana/boletim-da-safra-de-cana-de-acucar

CONAB (2021). *Serie histórica das Safras*. Brasilia. Retrieved from CONAB website: https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras?start=20

CONSECANA (2006). *Manual of instructions*. Piracicaba. Retrieved from CONSECANA website: http://www.oricana.com.br/novosite/manual_consecana.pdf

Dang, A. N., & Kawasaki, A. (2016). A review of methodological integration in land-use change models. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, *7*(2), 1–25.

Deppermann, A., Balkovič, J., Bundle, S.-C., Di Fulvio, F., Havlik, P., Leclère, D., . . . Schepaschenko, D. (2018). Increasing crop production in Russia and Ukraine—regional and global impacts from intensification and recultivation. *Environmental Research Letters*, *13*(2), 25008. https://doi.org/10.1088/1748-9326/aaa4a4

Dias, H. B., & Sentelhas, P. C. (2017). Evaluation of three sugarcane simulation models and their ensemble for yield estimation in commercially managed fields. *Field Crops Research*, *213*, 174–185. https://doi.org/10.1016/j.fcr.2017.07.022

Dos Santos, D. L., & Sentelhas, P. C. (2014). Climate Change Scenarios and Their Impact on Water Balance and Sugarcane Yield in Southern Brazil. *Sugar Tech*, *16*(4), 356–365. https://doi.org/10.1007/s12355-013-0293-y

Dos Santos Vianna, M., & Sentelhas, P. C. (2016). Performance of DSSAT CSM-CANEGRO under operational conditions and its use in determining the 'Saving irrigation' impact on sugarcane crop. *Sugar Tech*, *18*(1), 75–86.

Duarte, Y. C. N., & Sentelhas, P. C. (2020). Intercomparison and Performance of Maize Crop Models and Their Ensemble for Yield Simulations in Brazil. *International Journal of Plant Production*, *14*(1), 127–139. https://doi.org/10.1007/s42106-019-00073-5

Egli, D. B. (2018). Stability of Corn and Soybean Yield Ratios in Three Midwestern Environments. *Agronomy Journal*, *110*(1), 311–318. https://doi.org/10.2134/agronj2017.07.0403

Elliott, F. F. (1928). The "Representative Firm" Idea Applied to Research and Extension in Agricultural Economics. *Journal of Farm Economics*, *10*(4), 483. https://doi.org/10.2307/1229927

FAOSTAT (2021). *Crop dataset*. Rom. Retrieved from FAO website: http://www.fao.org/faostat/en/#data/QC

FEBRAPDP (2021). *Área sob Plantio Direto: Area under no-till system*. Retrieved from https://febrapdp.org.br/area-de-pd

Feuz, D. M., & Skold, M. D. (1992). Typical farm theory in agricultural research. *Journal of Sustainable Agriculture*, *2*(2), 43–58.

Flach, R., Skalský, R., Folberth, C., Balkovič, J., Jantke, K., & Schneider, U. A. (2020). Water productivity and footprint of major Brazilian rainfed crops – A spatially explicit analysis of crop management scenarios. *Agricultural Water Management*, *233*, 105996. https://doi.org/10.1016/j.agwat.2019.105996

Flack-Prain, S., Shi, L., Zhu, P., Rocha, H. R., Cabral, O., Hu, S., & Williams, M. (2021). The impact of climate change and climate extremes on sugarcane production. *GCB Bioenergy*, *13*(3), 408–424. https://doi.org/10.1111/gcbb.12797

Fliehr, Zimmer, & Smith (2019). Impacts of Transportation and Logistics on Brazilian Soybean Prices and Exports. *Transportation Journal*, *58*(1), 65. https://doi.org/10.5325/transportationj.58.1.0065

Foley, J. A., Defries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., . . . Snyder, P. K. (2005). Global consequences of land use. *Science (New York, N.Y.)*, *309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Forstner, B., Duden, C., Ellßel, R., Gocht, A., Hansen, H., Neuenfeldt, S., . . . Witte, T. de (2018). *Wirkungen von Direktzahlungen in der Landwirtschaft-ausgewählte Aspekte mit Bezug zum Strukturwandel: (Thünen Working Paper)*. Braunschweig. Retrieved from Thünen Institute website: https://literatur.thuenen.de/digbib_extern/dn059808.pdf

Françoso, R. F., Bigaton, A., Silva, H. J. T. d., & Marques, P. V. (2017). Ratio of the cost of transportation of sugarcane due to the distance. *Revista IPecege*, *3*(1), 100. https://doi.org/10.22167/r.ipecege.2017.1.100

Freebairn, J. W. (1987). Implications of wages and industrial policies on the competitiveness of agricultural export industries. *Review of Marketing and Agricultural Economics*, *55*(430-2016-31512), 79–87.

Gardner, R. H., & Urban, D. L. (2003). Model validation and testing: past lessons, present concerns, future prospects. *Models in Ecosystem Science*, 184–203.

Gassman, P. W., Williams, J. R., Benson, V. W., Izaurralde, R. C., Hauck, L. M., Jones, C. A., . . . Flowers, J. D. (2005). *Historical development and applications of the EPIC and APEX models* (CARD Working Papers No. 442). Retrieved from Iowa State University website: http://lib.dr.iastate.edu/cardworkingpapers/422

Goldemberg, J., Coelho, S. T., & Guardabassi, P. (2008). The sustainability of ethanol production from sugarcane. *Energy Policy*, *36*(6), 2086–2097. https://doi.org/10.1016/j.enpol.2008.02.028

Hall, B. F., & LeVeen, E. P. (1978). Farm size and economic efficiency: The case of California. *American Journal of Agricultural Economics*, *60*(4), 589–600.

Hampf, A. C., Stella, T., Berg-Mohnicke, M., Kawohl, T., Kilian, M., & Nendel, C. (2020). Future yields of double-cropping systems in the Southern Amazon, Brazil, under climate change and technological development. *Agricultural Systems*, *177*, 102707. https://doi.org/10.1016/j.agsy.2019.102707

Havlík, P., Schneider, U. A., Schmid, E., Böttcher, H., Fritz, S., Skalský, R., . . . Obersteiner, M. (2011). Global land-use implications of first and second generation biofuel targets. *Energy Policy*, *39*(10), 5690–5702. https://doi.org/10.1016/j.enpol.2010.03.030

Heckelei, T., & Britz, W. (Eds.) (2005). *Models based on positive mathematical programming: state of the art and further extensions*. Italy: 89th Seminar - European Association of Agricultural Economists. Retrieved from https://ideas.repec.org/p/ags/eaae89/234607.html

Heckelei, T., Britz, W., & Zhang, Y. (2012). Positive mathematical programming approaches–recent developments in literature and applied modelling. *Bio-Based and Applied Economics Journal*, *1*(1050-2016-85729), 109–124.

Heckelei, T., & Wolff, H. (2003). Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy. *European Review of Agricultural Economics*, *30*(1), 27–50. https://doi.org/10.1093/erae/30.1.27

Hemme, T. (1999). Ein Konzept zur international vergleichenden Analyse von Politik-und Technikfolgen in der Landwirtschaft - A concept for international analysis of the policy and technology impacts in agriculture (Ph.D.). University of Göttingen, Göttingen. https://doi.org/10.13140/2.1.4534.0803

Howitt, R. E. (1995). Positive mathematical programming. *American Journal of Agricultural Economics*, *77*(2), 329–342.

IBGE (2015). *Municipality Crop Production Survey (PAM)*. Brasilia. Retrieved from IBGE website: https://sidra.ibge.gov.br/pesquisa/pam/tabelas

IBGE (2018). *Land cover and land use in Brazil in 2016*. Brasilia. Retrieved from IBGE website: https://www.ibge.gov.br/apps/monitoramento_cobertura_uso_terra/v1/

IBGE (2019). *Systematic Survey of Agricultural Production*. Brasilia. Retrieved from https://sidra.ibge.gov.br/home/lspa

IBGE (2020). *Land Use and Cover in Brazil 2018: Monitoramento da Cobertura e Uso da Terra 2018*. Brasilia. Retrieved from IBGE website: https://www.ibge.gov.br/geociencias/informacoes-ambientais/cobertura-e-uso-da-terra/15831-cobertura-e-uso-da-terra-do-brasil.html?=&t=o-que-e

IBGE (2021). *Municipal Livestock Survey (PPM)*. Brasilia. Retrieved from IBGE website: https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9107-producao-da-pecuaria-municipal.html?=&t=resultados

IFPRI (2019). Global Spatially-Disaggregated Crop Production Statistics Data for 2010 Version 2.0 (DRAFT VERSION). MapSPAM: Harvard Dataverse. https://doi.org/10.7910/DVN/PRFF8V

ITC (2021). *Trade statistics for international business development*. Retrieved from Intracen website: https://www.trademap.org

Izaurralde, R. C., Williams, J. R., McGill, W. B., Rosenberg, N. J., & Jakas, M. Q. (2006). Simulating soil C dynamics with EPIC: Model description and testing against long-term data. *Ecological Modelling*, *192*(3-4), 362–384. https://doi.org/10.1016/j.ecolmodel.2005.07.010

Knox, J., Hess, T., Daccache, A., & Wheeler, T. (2012). Climate change impacts on crop productivity in Africa and South Asia. *Environmental Research Letters*, *7*(3), 34032.

Krug, J. (2013). Perspektiven ackerbaulicher Grenzstandorte in Nordostdeutschland: Übertragbarkeit extensiver Produktionssysteme überseeischer Trockenstandorte (Ph.D. Thesis). Johann Heinrich von Thünen-Institut, Braunschweig. Retrieved from http://nbn-resolving.de/urn:nbn:de:gbv:253-201310-dn052166-8 https://doi.org/10.3220/REP_6_2013

Leite, M. D. L., Adacheski, P. A., & Virgens Filho, J. S. d. (2011). Análise da frequência e da intensidade das chuvas em Ponta Grossa, Estado do Paraná, no período entre 1954 E 2001. *Acta Scientiarum. Technology*, *33*(1). https://doi.org/10.4025/actascitechnol.v33i1.6957

Liang, Y., Miller, J. C., Harri, A., & Coble, K. H. (2011). Crop Supply Response under Risk: Impacts of Emerging Issues on Southeastern U.S. Agriculture. *Journal of Agricultural and Applied Economics*, *43*(2), 181–194. https://doi.org/10.1017/S1074070800004156

Linnenluecke, M. K., Nucifora, N., & Thompson, N. (2018). Implications of climate change for the sugarcane industry. *Wiley Interdisciplinary Reviews: Climate Change*, *9*(1), e498. https://doi.org/10.1002/wcc.498

Macedo, M. N., DeFries, R. S., Morton, D. C., Stickler, C. M., Galford, G. L., & Shimabukuro, Y. E. (2012). Decoupling of deforestation and soy production in the southern Amazon during the late 2000s. *Proceedings of the National Academy of Sciences*, *109*(4), 1341–1346.

Marcuzzo, F., Oliveira, N. D. L., & Cardoso, M. R. D. (2013). Tendência do Número de Dias de Chuva no Estado do Mato Grosso do Sul e Sua Relação com o Fenômeno ENOS (Trend in Number of Days of Precipitation in the State of Mato Grosso do Sul and Its Relationship with ENSO Phenomenon). *Revista Brasileira De Geografia Física*, *5*(5), 1133. https://doi.org/10.26848/rbgf.v5i5.232824

Marin, F. R., Jones, J. W., Royce, F., Suguitani, C., Donzeli, J. L., Filho, W. J. P., & Nassif, D. S. (2011). Parameterization and Evaluation of Predictions of DSSAT/CANEGRO for Brazilian Sugarcane. *Agronomy Journal*, *103*(2), 304. https://doi.org/10.2134/agronj2010.0302

Marin, F. R., Jones, J. W., Singels, A., Royce, F., Assad, E. D., Pellegrino, G. Q., & Justino, F. (2013). Climate change impacts on sugarcane attainable yield in southern Brazil. *Climatic Change*, *117*(1-2), 227–239. https://doi.org/10.1007/s10584-012-0561-y

Marin, F. R., Thorburn, P. J., Nassif, D. S., & Costa, L. G. (2015). Sugarcane model intercomparison: Structural differences and uncertainties under current and potential future climates. *Environmental Modelling & Software*, *72*, 372–386. https://doi.org/10.1016/j.envsoft.2015.02.019

Mascarenhas, H., Tanaka, R., Wutke, E., Braga, N., & Miranda, M. (2004). *Potássio para a soja (Potash for soybeans)*. Piracicaba. Retrieved from KP Potafos website: http://www.ipni.net/publication/ia-brasil.nsf/0/ED5BEE124EE957B483257AA2005AC216/$FILE/Jornal105.pdf

Matsuoka, S., Ferro, J., & Arruda, P. (2009). The Brazilian experience of sugarcane ethanol industry. *In Vitro Cellular & Developmental Biology - Plant*, *45*(3), 372–381. https://doi.org/10.1007/s11627-009-9220-z

McCarl, B. A. (1982). Cropping Activities in Agricultural Sector Models: A Methodological Proposal. *American Journal of Agricultural Economics*, *64*(4), 768–772. https://doi.org/10.2307/1240588

McKay, A., Morrissey, O., & Vaillant, C. (1999). Aggregate supply response in Tanzanian agriculture. *The Journal of International Trade & Economic Development*, *8*(1), 107–123. https://doi.org/10.1080/09638199900000008

Mearns, L. O., Easterling, W., Hays, C., & Marx, D. (2001). Comparison of Agricultural Impacts of Climate Change Calculated from High and Low Resolution Climate Change Scenarios: Part I. The Uncertainty Due to Spatial Scale. *Climatic Change*, *51*(2), 131–172. https://doi.org/10.1023/A:1012297314857

Mendelsohn, R. (2008). The Impact of Climate Change on Agriculture in Developing Countries. *Journal of Natural Resources Policy Research*, *1*(1), 5–19. https://doi.org/10.1080/19390450802495882

Molin, J. P., Milan, M., Nesrallah, M. G. T., Castro, C. N. de, & Gimenez, L. M. (2006). Utilização de dados georreferenciados na determinação de parâmetros de desempenho em colheita mecanizada: Using georreferenced data for determination of harvesting parameters. *Engenharia Agrícola*, *26*(3), 759–767.

Moraes Sá, J. C. de, Séguy, L., Tivet, F., Lal, R., Bouzinac, S., Borszowskei, P. R., . . . Bertoloni, C. G. (2015). Carbon depletion by plowing and its restoration by no-till cropping systems in oxisols of subtropical and tropical agro-ecoregions in Brazil. *Land Degradation & Development*, *26*(6), 531–543.

Nehring, K. (2011). Farm level implications of high commodity prices: An assessment of adaption strategies and potentials in selected regions in Australia and Germany (Ph.D.). University of Göttingen, Göttingen.

Nendel, C., Berg, M., Kersebaum, K. C., Mirschel, W., Specka, X., Wegehenkel, M., . . . Wieland, R. (2011). The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecological Modelling*, *222*(9), 1614–1625.

Nobre, C. A., Sellers, P. J., & Shukla, J. (1991). Amazonian deforestation and regional climate change. *Journal of Climate*, *4*(10), 957–988.

Noszczyk, T. (2019). A review of approaches to land use changes modeling. *Human and Ecological Risk Assessment: An International Journal*, *25*(6), 1377–1405. https://doi.org/10.1080/10807039.2018.1468994

Oliveira, E. C. A. de, Freire, F. J., Oliveira, R. I. de, Freire, M. B. G. d. S., Simões Neto, D. E., & Silva, S. A. M. d. (2010). Extração e exportação de nutrientes por variedades de cana-de-açúcar cultivadas sob irrigação plena. *Revista Brasileira De Ciência Do Solo*, *34*(4), 1343–1352. https://doi.org/10.1590/S0100-06832010000400031

Osaki, M., & Batalha, M. O. (2014). Optimization model of agricultural production system in grain farms under risk, in Sorriso, Brazil. *Agricultural Systems*, *127*, 178–188. Retrieved from http://www.sciencedirect.com/science/article/pii/S0308521X14000274

Pires, G. F., Abrahão, G. M., Brumatti, L. M., Oliveira, L. J., Costa, M. H., Liddicoat, S., . . . Ladle, R. J. (2016). Increased climate risk in Brazilian double cropping agriculture systems: Implications for land use in Northern Brazil. *Agricultural and Forest Meteorology*, *228-229*, 286–298. https://doi.org/10.1016/j.agrformet.2016.07.005

Plaxico, J. S., & Tweeten, L. G. (1963). Representative Farms for Policy and Projection Research. *Journal of Farm Economics*, *45*(5), 1458. https://doi.org/10.2307/1236844

Pontius, R., & Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment*, *85*(1-3), 239–248. https://doi.org/10.1016/S0167-8809(01)00187-6

Python (2020). Python (Version 3.7.7) [Computer software]: https://www.python.org/. Retrieved from https://www.python.org/

Ramburan, S., Wettergreen, T., Berry, S. D., & Shongwe, B. (2013). Genetic, environmental and management contributions to ratoon decline in sugarcane. *Field Crops Research*, *146*, 105–112. https://doi.org/10.1016/j.fcr.2013.03.011

Rao, J. (1989). Agricultural supply response: A survey. *Agricultural Economics*, *3*(1), 1–22. https://doi.org/10.1016/0169-5150(89)90036-4

Raup, P. M. (1969). Economies and diseconomies of large-scale agriculture. *American Journal of Agricultural Economics*, *51*(5), 1274–1283.

Reback, J., McKinney, W., Jbrockmendel, van Bossche, J. den, Augspurger, T., Cloud, P., . . . H-Vetinari (2020). pandas-dev/pandas: Pandas 1.2.3 [Computer software]: Zenodo.

Richards, P. D., Myers, R. J., Swinton, S. M., & Walker, R. T. (2012). Exchange rates, soybean supply response, and deforestation in South America. *Global Environmental Change*, *22*(2), 454–462. https://doi.org/10.1016/j.gloenvcha.2012.01.004

Richter, K., Atzberger, C., Hank, T. B., & Mauser, W. (2012). Derivation of biophysical variables from Earth observation data: validation and statistical measures. *Journal of Applied Remote Sensing*, *6*(1), 063557-1. https://doi.org/10.1117/1.JRS.6.063557

Schmitz, C., van Meijl, H., Kyle, P., Nelson, G. C., Fujimori, S., Gurgel, A., . . . Valin, H. (2014). Land-use change trajectories up to 2050: insights from a global agro-economic model comparison. *Agricultural Economics*, *45*(1), 69–84. https://doi.org/10.1111/agec.12090

Seo, S. N., & Mendelsohn, R. (2008). An analysis of crop choice: Adapting to climate change in South American farms. *Ecological Economics*, *67*(1), 109–116. https://doi.org/10.1016/j.ecolecon.2007.12.007

Sharples, J. A. (1969). The Representative Farm Approach to Estimation of Supply Response. *American Journal of Agricultural Economics*, *51*(2), 353–361. https://doi.org/10.2307/1237585

Silva, R. P. d., Corrêa, C. F., Cortez, J. W., & Furlani, C. E. A. (2008). Controle estatístico aplicado ao processo de colheita mecanizada de cana-de-açúcar. *Engenharia Agrícola*, *28*(2), 292–304.

Singels, A., Jones, M., Marin, F., Ruane, A., & Thorburn, P. (2014). Predicting Climate Change Impacts on Sugarcane Production at Sites in Australia, Brazil and South Africa Using the Canegro Model. *Sugar Tech*, *16*(4), 347–355. https://doi.org/10.1007/s12355-013-0274-1

Skalský, R., Tarasovicová, Z., Balkovic, J., Schmid, E., Fuchs, M., Moltchanova, E., . . . Scholtz, P. (2008). *Global database for bio-physiscal modeling v. 1.0: (Concept. methodologies and data)*. Laxenburg.

Sørensen, C. G. (2003). Workability and machinery sizing for combine harvesting. *Agricultural Engineering International: CIGR Journal*.

Soterroni, A. C., Mosnier, A., Carvalho, A. X. Y., Câmara, G., Obersteiner, M., Andrade, P. R., . . . Ramos, F. M. (2018). Future environmental and agricultural impacts of Brazil's Forest Code. *Environmental Research Letters*, *13*(7), 74021. https://doi.org/10.1088/1748-9326/aac-cbb

Soterroni, A. C., Ramos, F. M., Mosnier, A., Fargione, J., Andrade, P. R., Baumgarten, L., . . . Polasky, S. (2019). Expanding the Soy Moratorium to Brazil's Cerrado. *Science Advances*, *5*(7), eaav7336. https://doi.org/10.1126/sciadv.aav7336

Souza, A. E., Reis, J. M., Abraham, E. R., & Machado, S. T. (2017). Brazilian Corn Exports: An Analysis of Cargo Flow in Santos and Paranagua Port. In H. Lödding, R. Riedel, K.-D. Thoben, G. von Cieminski, & D. Kiritsis (Eds.), *Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing* (pp. 105–112). Cham: Springer International Publishing.

UDOP (2017). Usinas de açúcar e álcool no Brasil - SHP: GISMAPS - www.gismaps.com.br. Piracicaba.

UNICA (2019). *Brazilian Sugar Exports by State*. Sao Paulo. Retrieved from UNICA website: http://www.unicadata.com.br/listagem.php?idMn=43

UNICA (2021). *Sugar and Ethanol exports: Volume per port in 2020*. Sao Paulo. Retrieved from https://observatoriodacana.com.br/listagem.php?idMn=143&idioma=1

USDA (2021). *Production, Supply and Distribution (PSD)*. United States. Retrieved from USDA website: https://apps.fas.usda.gov/psdonline/app/index.html#/app/downloads

Van Alfen, N. K. (2014). *Encyclopedia of agriculture and food systems: 5-volume set* (2nd ed.). Amsterdam: Elsevier/Academic Press. Retrieved from http://www.sciencedirect.com/science/referenceworks/9780080931395

Veldkamp, A., & Lambin, E. (2001). Predicting land-use change. *Agriculture, Ecosystems & Environment*, *85*(1), 1–6. https://doi.org/10.1016/S0167-8809(01)00199-2

Victoria, C., Bolfe, L., & da Silva, S. (Eds.) (2017). *Pastagens plantadas com potencial para expansão da agricultura anual no Cerrado*. Santos: XVIII Simpósio Brasileiro de Sensoriamento Remoto. Retrieved from https://proceedings.science/sbsr/papers/pastagens-plantadas-com-potencial-para-expansao-da-agricultura-anual-no-cerrado

Wang, J., Vanga, S., Saxena, R., Orsat, V., & Raghavan, V. (2018). Effect of Climate Change on the Yield of Cereal Crops: A Review. *Climate*, *6*(2), 41. https://doi.org/10.3390/cli6020041

WHO (2021). *Coronavirus disease (COVID-19) pandemic*. Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Williams, J. R., Izaurralde, R. C., Williams, C., & Steglich, E. M. (2015). *Agricultural Policy/Environmental EXtender Model: Theoretical Documentation - Version 0806*. Texas. Retrieved from http://agrilife.org/epicapex/files/2017/03/THE-APEX0806-theoretical-documentation-Oct-2015.pdf

Williams, J. R., Jones, C. A., & Dyke, P. T. (1984). A Modeling Approach to Determining the Relationship Between Erosion and Soil Productivity. *Transactions of the ASAE*, *27*(1), 129–144. https://doi.org/10.13031/2013.32748

Williams, J. R., Jones, C. A., Kiniry, J. R. [J. R.], & Spanel, D. A. (1989). The EPIC Crop Growth Model. *Transactions of the ASAE*, *32*(2), 497–511.

Williams, J. R., & Singh, V. P. (1995). Computer models of watershed hydrology. *Chap. The EPIC Model, Water Resources Publications, Highlands Ranch, CO*, 909–1000.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82.

Wright, B. D. (2011). The Economics of Grain Price Volatility. *Applied Economic Perspectives and Policy*, *33*(1), 32–58. https://doi.org/10.1093/aepp/ppq033

Zandonadi, R. S., Ruffato, S., & Figueiredo, Z. N. (2015). Perdas na colheita mecanizada de soja na região Médio-Norte de Mato Grosso: Safra 2012/2013. *Nativa*, *3*(1), 64–66.

Zhao, X., Calvin, K. V., & Wise, M. A. (2020). The critical role of conversion cost and comparative advantage in modeling agricultural land use change. *Climate Change Economics*, *11*(PNNL-SA-142333).

Zilli, M., Scarabello, M., Soterroni, A. C., Valin, H., Mosnier, A., Leclère, D., . . . Ramos, F. M. (2020). The impact of climate change on Brazil's agriculture. *The Science of the Total Environment*, *740*, 139384. https://doi.org/10.1016/j.scitotenv.2020.139384

Zimmer, Y., & Deblitz, C. (2005). *agri benchmark cash crop: a standard operating procedure to define typical farms*. Braunschweig. Retrieved from Thünen Institute of Farm Economics website: https://literatur.thuenen.de/digbib_extern/dk038513.pdf

Appendices

Appendices

Appendix 1: Historical sugarcane production in the main regions in Brazil (in million t)
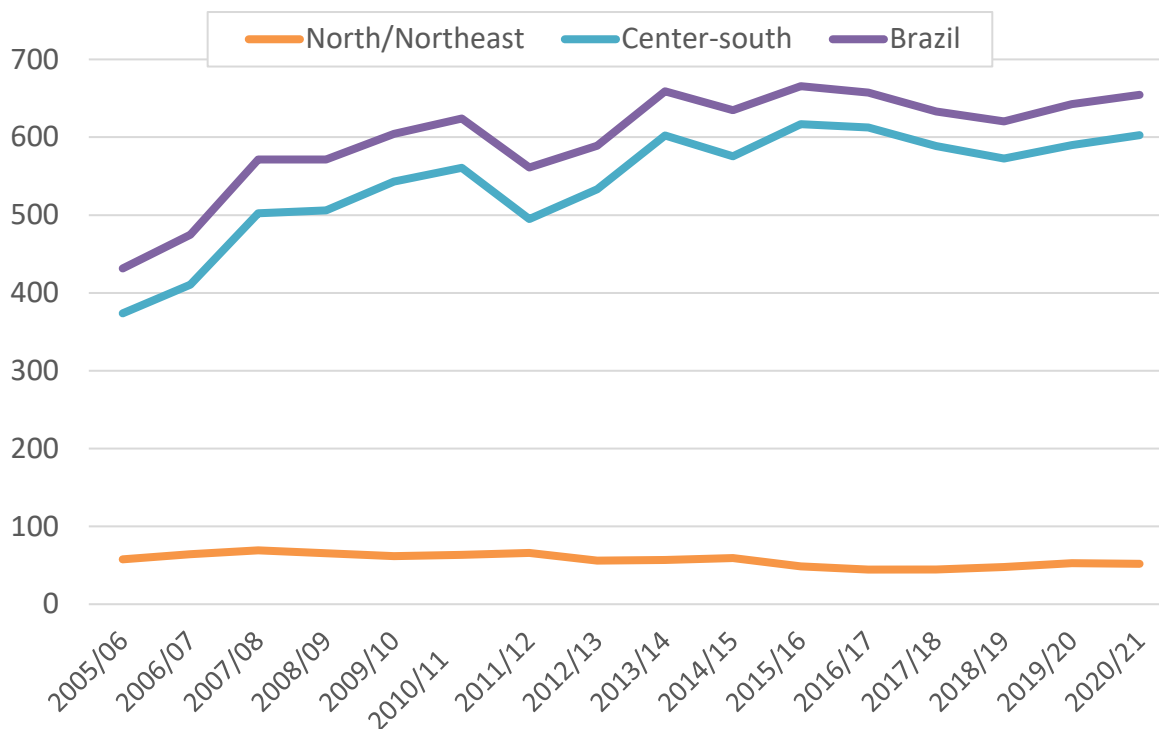
Source: CONAB (2021) – created by the author.

Appendix 2:   Historical soybean production in the main regions in Brazil (in million t)
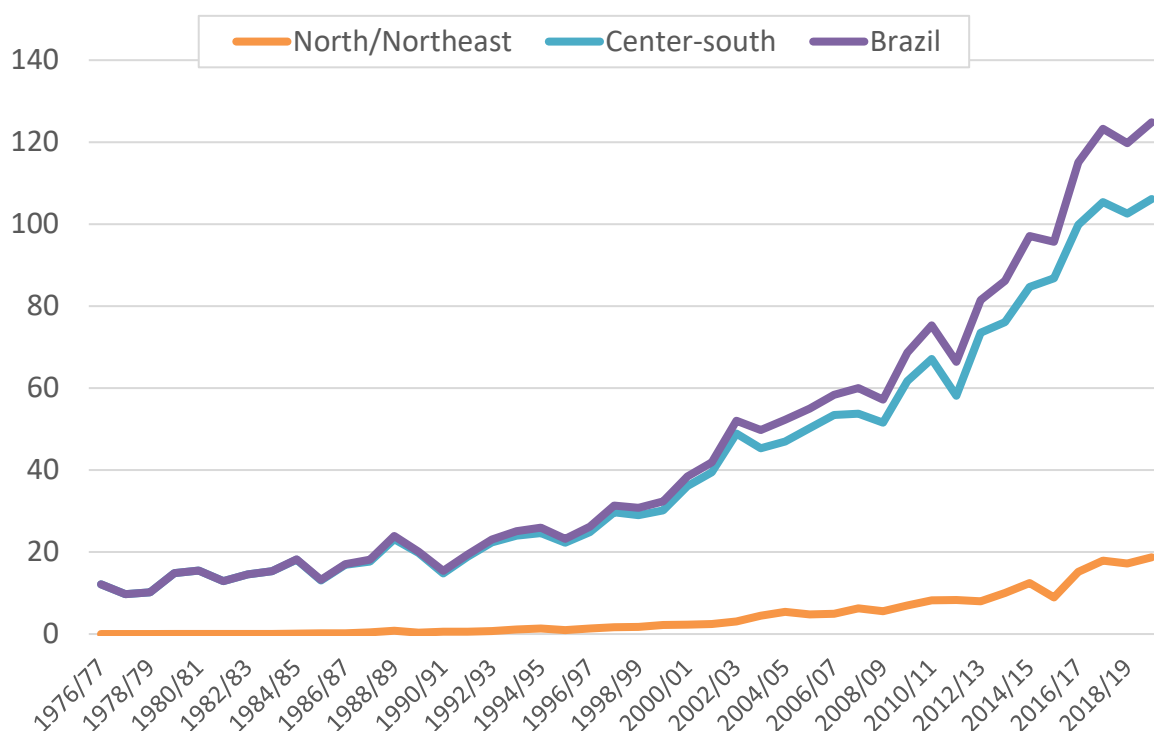
Source: CONAB (2021) – created by the author.

Appendix 3:   Nitrogen and Potash prices (in R$/t) of pure nutrient on average 2016-2018
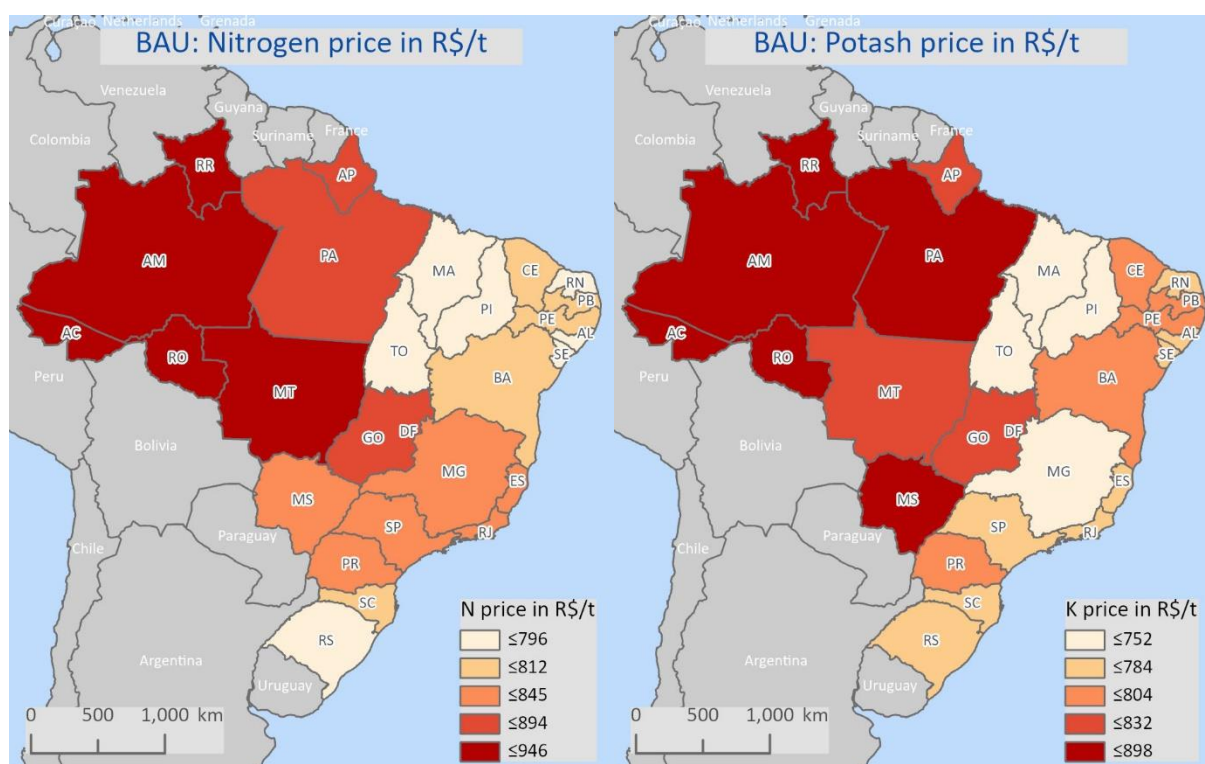
Source: CEPEA (2019) – created by the author.

Appendix 4:   Municipalities with ANP data (left) and allocation of available data to other municipalities (right)
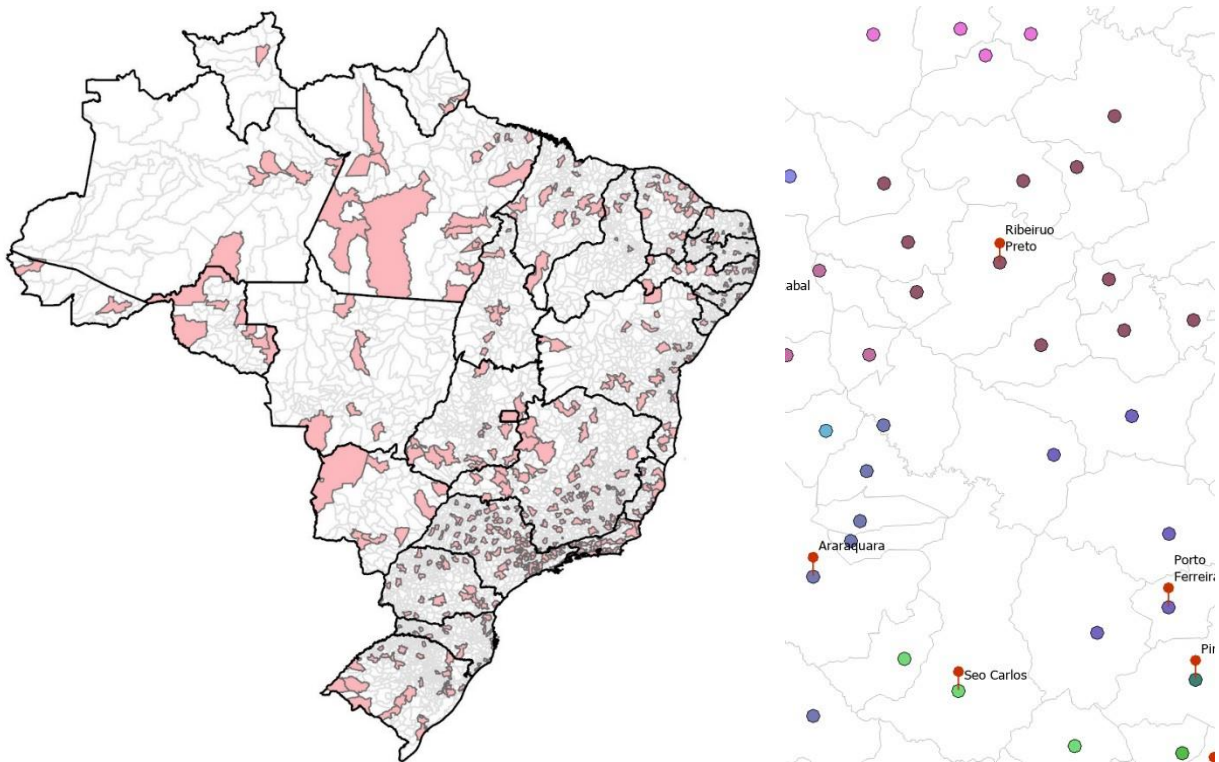
Source: ANP (2020) – created by the author.

Appendix 5:   Share of second-crop maize in the double-cropping system
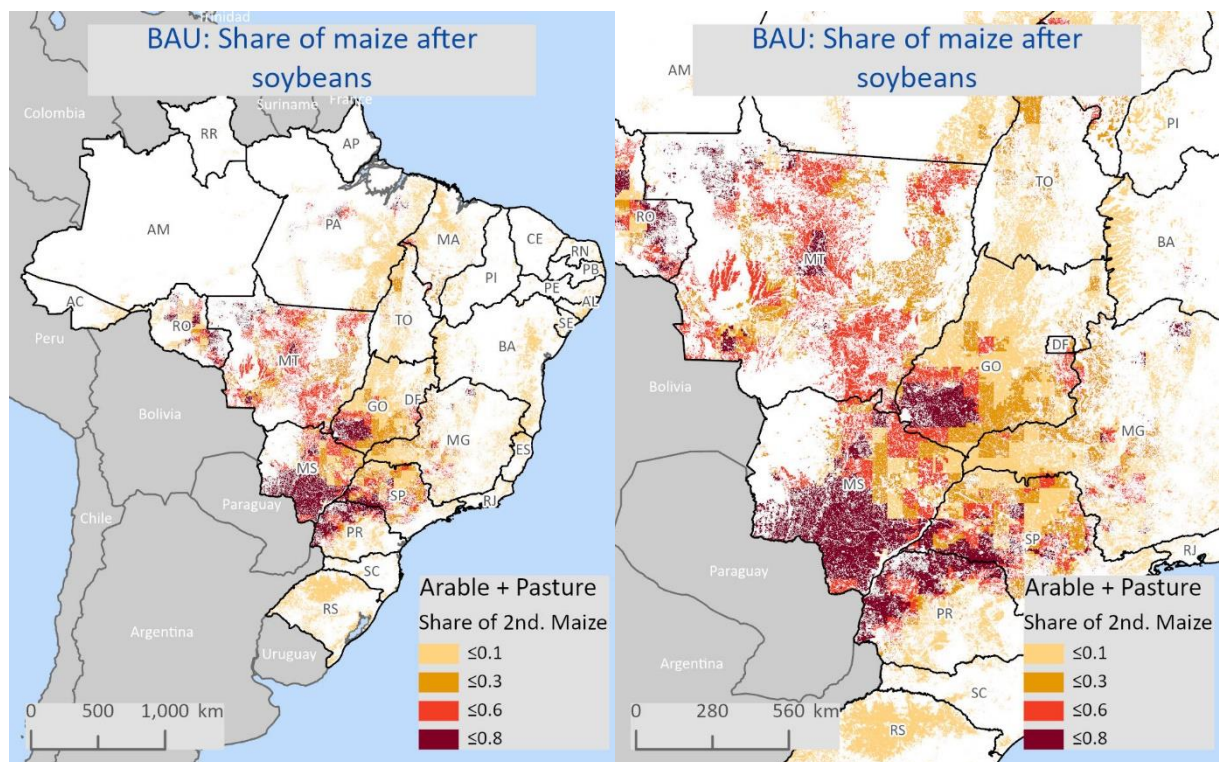
Source: IBGE (2015) – created by the author.

Appendix 6:   Diesel prices in Brazil (in R$/l) deflated using the Extended National Consumer Price Index (IPCA)
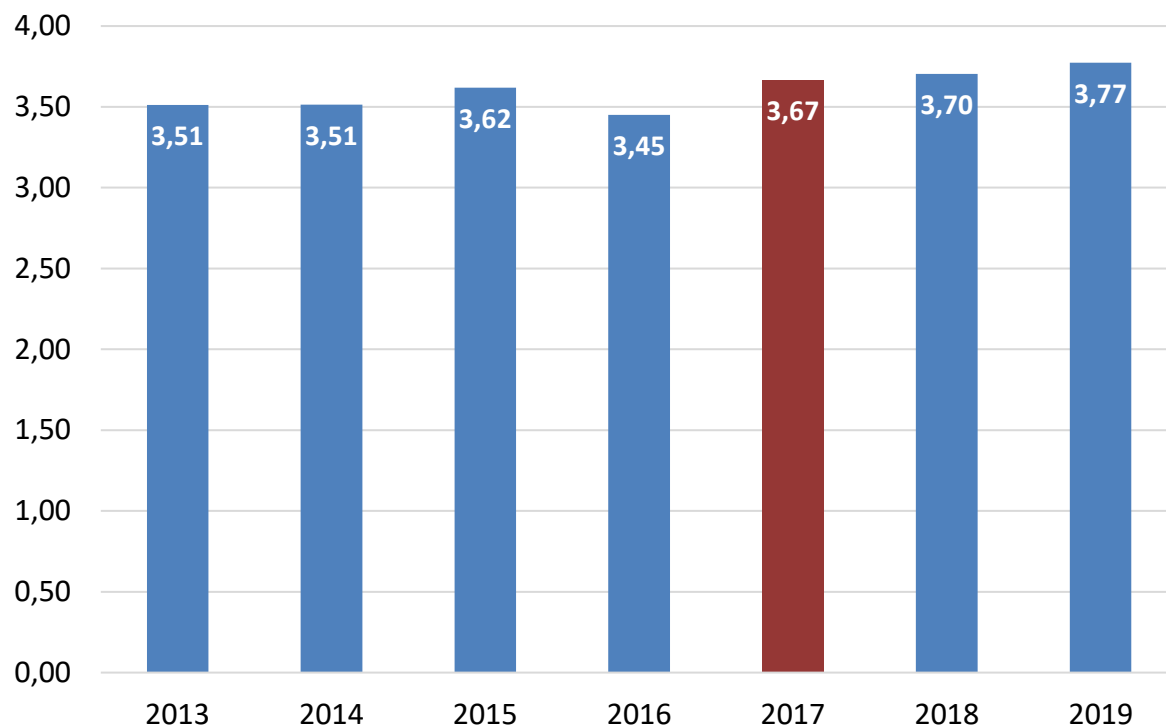
Source: own calculations based on ANP (2020).

Appendix 7:   BAU land use with (left) and without (right) yield correction
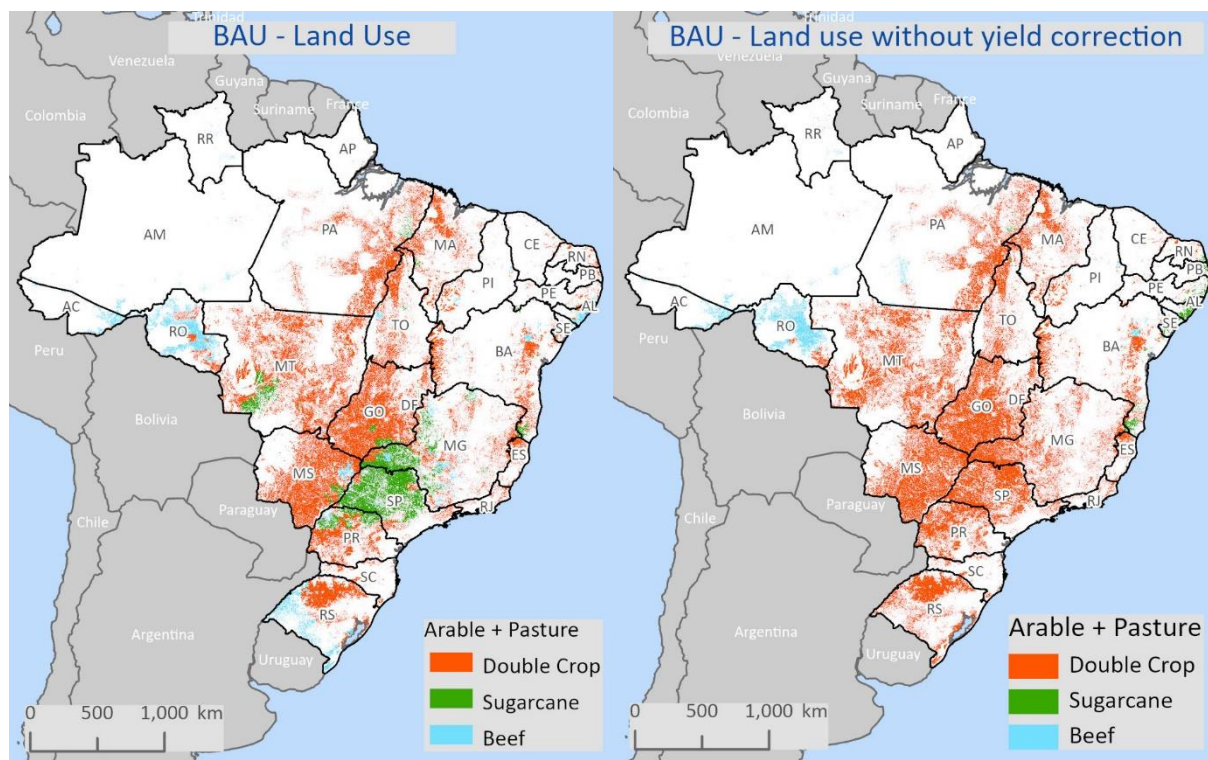
Source: PAM results (2020) – created by the author.

Appendix 8:   Comparison validation match PAM (BAU) vs IBGE (left) and SPAM (right)
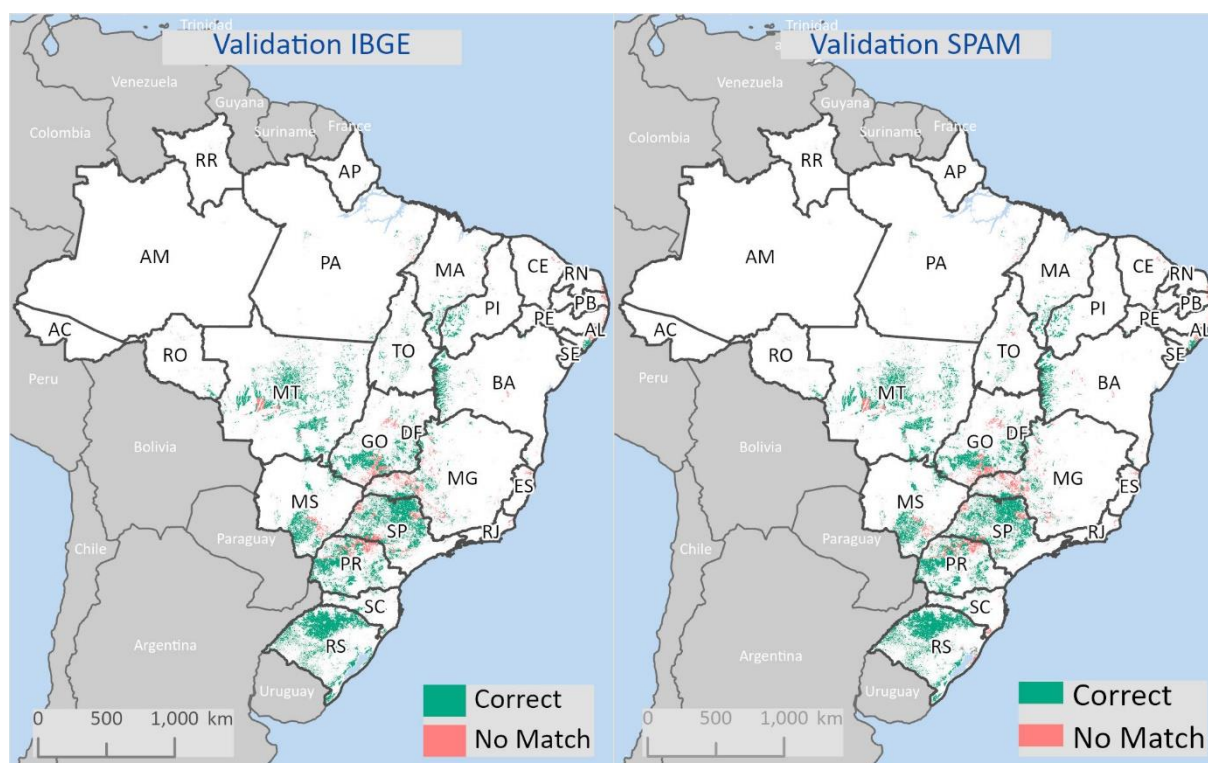
Source: PAM results (2020), IBGE (2019) and IFPRI (2019) – created by the author.

Appendix 9:   Simulated land-use maps for the BAU scenario (left) and double cropping with yield penalty (right) – zoom in the main producing states
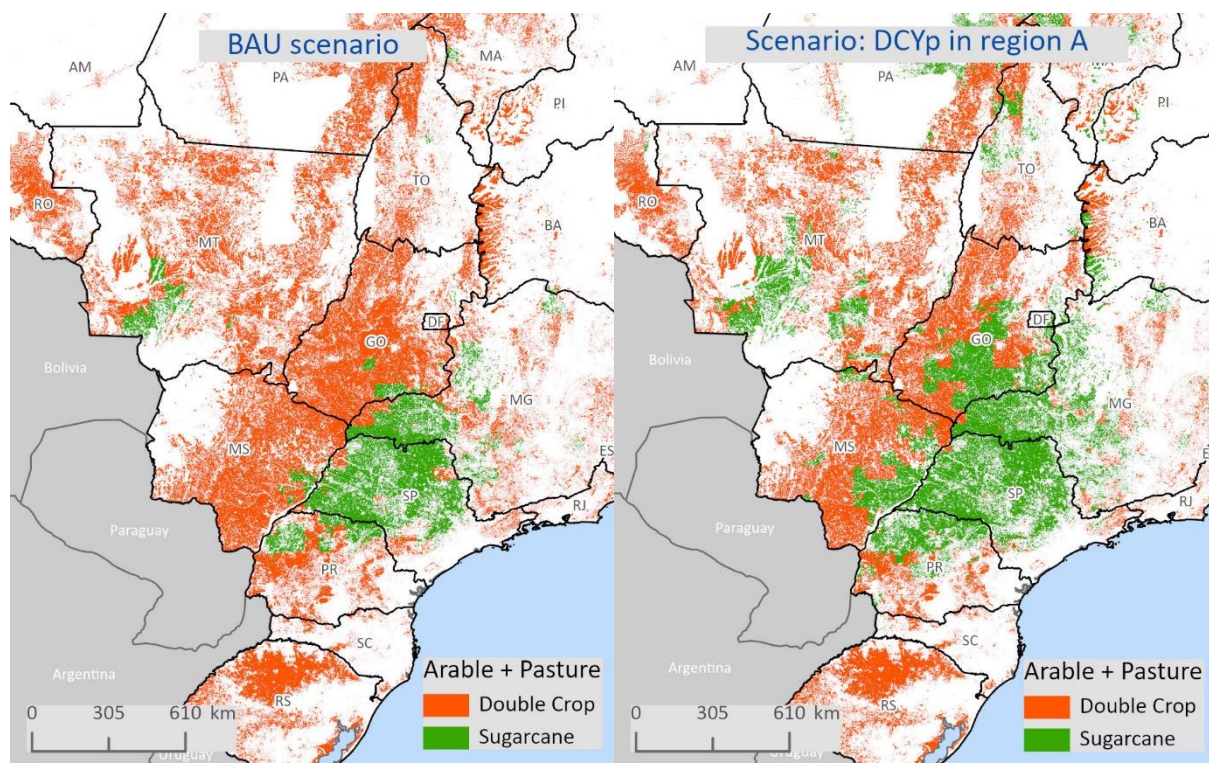
Source: PAM results (2020) – created by the author.

Appendix 10: Schematic representation of the *agri benchmark* costing model
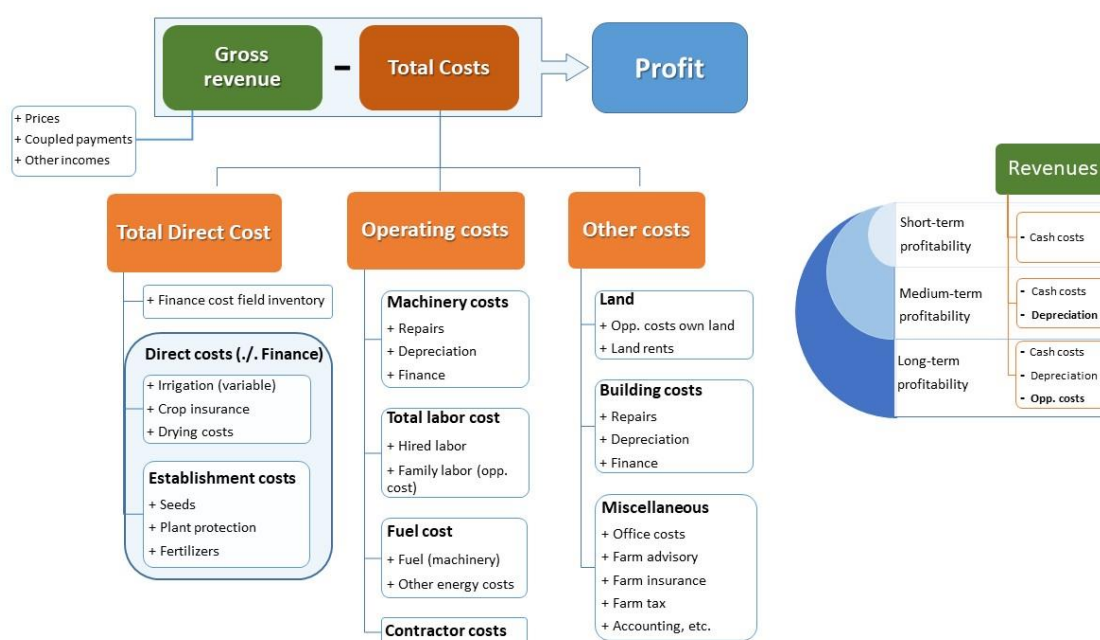
Source: created by the author (2014).

Appendix 11: Assignment of the *agri benchmark* typical farms' data to each state in Brazil for grains (left) and sugarcane (right)
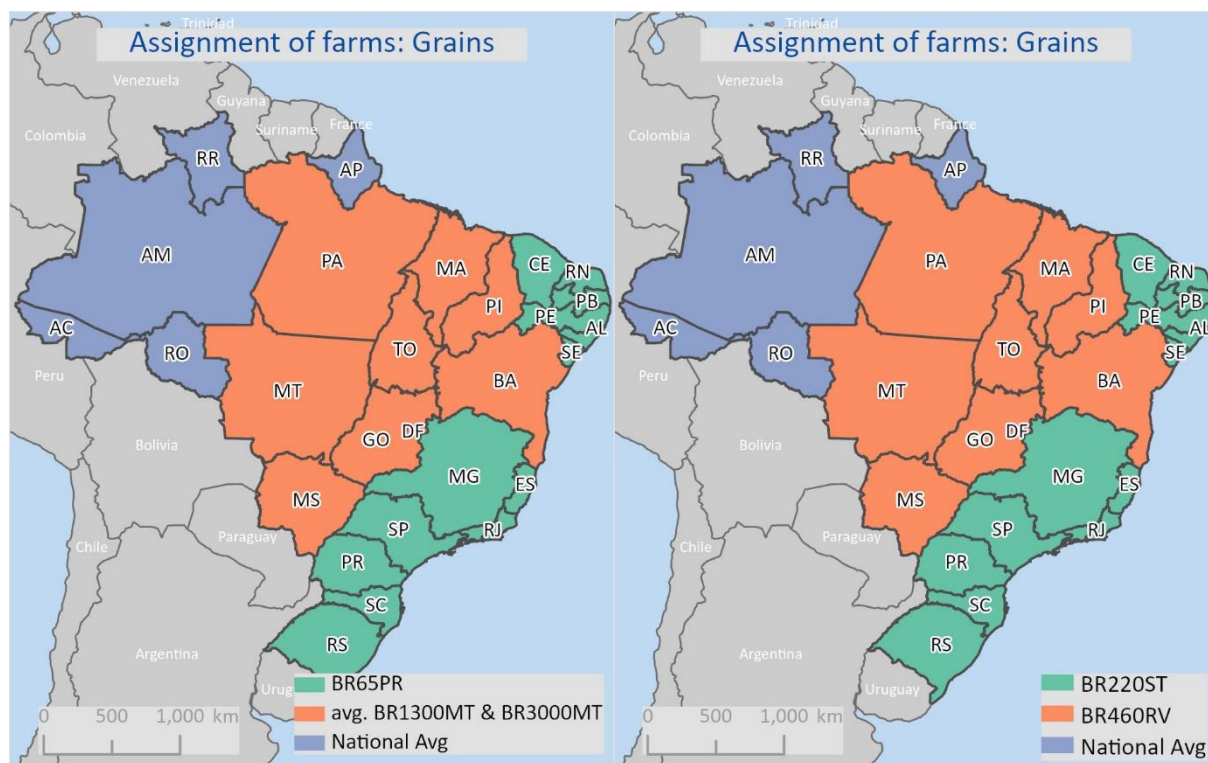
Source: created by the author.

Thünen Report

Bereits in dieser Reihe erschienene Hefte – *Volumes already published in this series*

THÜNEN