trase | TRANSPARENCY FOR SUSTAINABLE ECONOMIES

JUNE 2020

Trase 'SEI-PCS Brazil soy v2.5.0' supply chain map: Data sources and methods

Trase maps supply chains for agricultural commodities, making it possible to link products and supply chain actors with specific areas of production, and associated sustainability risks and opportunities. It uses an approach called Spatially Explicit Information on Production to Consumption Systems (SEI-PCS) as the basis for this work (see this webpage or our manual for more detail). This document describes the data and methods that Trase used to map the subnational supply chain for Brazilian soy exports from 2003-2018, using a model called **'SEI-PCS Brazil soy v2.5.0'**.

For all soy exports, this model determined the likely municipality in which the soy was produced. It used trade and production data, as well as information on tax registrations, company asset ownership, and logistics. The approach links exports back to municipal locations of taxation that can correspond to farms, silos, crushing facilities or wholesale retailing. To link these logistic hubs with the municipalities of production where the soy was most likely produced, we used a linear programming approach. This optimised the allocation of soy from production municipalities to logistic hubs based on company asset ownerships, economic activities in each municipality, and distance. Table 1 provides an overview of key statistics.

Table 1. Summary statistics

	2005	2010	2015	2016	2017	2018
Soy exports (million tonnes*)	39	45	71	68	84	100
Domestic market (% share of production)	23	35	27	30	27	16
Number of exporter groups	337	287	345	290	311	297
Number of importing countries	77	71	81	78	72	94
Exports with unknown source of origin (%)	21.3	7.5	8.1	9.1	11	19

* = metric tons

Data and sources

Trade data

The model used per-shipment data (such as customs data, bills of lading or cargo manifests) for 2003-2018, covering all exports of soy classified under the 'HS' customs codes in Table 2. These products include soybeans, soy cake and soy oil. We used weight-based commodity-equivalence factor (see Table 2) to convert different products to a standard commodity equivalent, in this case soybeans. We confirmed the quality of the data by comparing it with other data sources and with data in different aggregated forms (including COMTRADE, USDA, and reports from the Brazilian private soy sector).

Table 2. Soy products, their HS customs codes, commodity-equivalence factors and the share of each product in total exports in 2018

Product	HS code	Commodity equivalence factor
Soybean cake	120810, 230400	1.031
Soybean oil	150710, 150790	1.031
Soybeans	120100, 120110,120190	1

Domestic demand

We mapped the location and the processing capacity of soy processing centres, including all crushing facilities producing livestock feed and oil refineries for biofuel and other industrial applications. By subtracting municipal demand for exports from municipal production, we obtained the total domestic demand as well as its approximate location, given that soy is domestically used almost exclusively for feed (crushing facilities) and for biofuels (oil refineries).

Production data

We obtained production data for 2003-2018 from the Brazilian Institute of Geography and Statistics (IBGE). These data provide statistics on the volume of soy produced in each municipality.

Supply chain data

Asset data

Exporter assets

The National Registry of Legal Entities (CNPJ) provides the data needed to identify all the assets — such as farms, silos, warehouses, etc. —owned by a given exporter, and all the assets associated with a given activity, such as soy farming or soy crushing.

Soy crushing and processing facilities

ABIOVE (the Brazilian Association of Vegetable Oil Industries) provides information on the location and ownership of soy crushing and processing facilities. Their capacities were estimated from a variety of sources. The National Registry of Storage Facilities (SICARM) of the National Supply Company (CONAB) provides information on soy silos. The Ministry of Agriculture keeps a list of Brazilian company assets that are allowed to export to specific countries, based on sanitary permissions granted by importing countries.

Transportation data

Road network

We used the transportation network data from the Ministry of Transport.

Company data

We used the National Registry of Legal Entities (CNPJ) for company data. This enabled us to link customs records with taxation records, and to identify the subsidiaries of the main traders (e.g. Bunge or Cargill), which we then grouped according to their parent company.

Boundaries

We used the municipal boundaries from the Brazilian Institute of Geography and Statistics (IBGE) 2017 data.

SEI-PCS implementation

We used a logic-based decision tree to link exports back to soy silos and crushing facilities, and farms in municipalities of production. The decision tree allocated exports to logistics hubs based on a series of conditional rules.

It triangulated information in the per-shipment export data against asset-level tax registration numbers. This made it possible to link exports to assets in production municipalities where assets are linked to production activities (e.g. soy farms) or to logistics hubs such as soy storage and crushing facilities. We labelled exports as having an 'unknown source of origin' if we could not link them to a soy storage or crushing facility with a sufficient degree of certainty.

We used an optimised allocation using linear programming techniques to link exports allocated to soy silos and crushing facilities (logistics hub municipalities) to the municipality in which the soy was grown. The linear programming maximised the allocation of soy from municipalities where the exporter company has taxation records associated to soy related activities, while minimising the overall distance between the soy production municipalities and the logistic hubs for exports and domestic demand. Supply nodes were constrained by production and demand nodes by annual processing capacity.

Subnational and company sustainability indicators

Once we had allocated soy from municipalities of production to logistic hubs and therefore though the trade data to ports, companies and countries of destination, we applied the ratios of this volume allocation to municipal sustainability indicators associated with soy production, such as deforestation and carbon emissions. In other words, we assigned risk in each municipality proportionally to each buyer. For example, if a trader exported 20% of the soy produced in a given municipality, 20% of the soy deforestation of that municipality would be allocated to that trader and those specific exports.

The Trase indicator manual for Brazil describes the data sources and methods to calculate these municipal level sustainability indicators including soy deforestation risk. The indicators cover agriculture, environment, territorial governance, actor commitments, and socio-economic and contextual indicators (e.g. biomes).

Changes from previous version(s)

Version	Release date	Changes from previous version
2.5.0	June 2020	• As v2.4 but includes 2018 trade and an improvement to the estimate of deforestation for soy using annual soy maps as well as deforestation data now available in all biomes, except Caatinga and Pampa. A new definition and calculation has also been applied for soy deforestation and soy deforestation risk.
2.4	December 2019	• As v2.3 but with improvements to the decision tree.
2.3	December 2018	• As v.2.2 but with significant improvements in the accuracy with which the sourcing regions of individual export shipments are mapped, due in particular to mapping of domestic demand centres, and improved data coverage on companies and assets.
2.2	September 2017	• As v 2.1 but links per-shipment information to localities of production and logistics facilities based on common asset-level tax registration numbers (eliminating potential uncertainties in determining the plausibility of reported logistics hubs per trade record). Hard-to obtain self-declarations become obsolete and are removed.
2.1	March 2017	• As v2.0 but with added linear program (minimum distance) to link logistics hubs to production municipalities.
2.0	November 2016	• First release.

How to cite this document

Trase. 2020. Trase 'SEI-PCS Brazil soy v2.5.0' supply chain map: Data sources and methods. Available at www.trase.earth.

Annex: Glossary

Term	Definition	Example
Asset	In the context of Trase, a physical or material resource owned by a business or an economic entity that relates to the production, storage or processing of a commodity.	Soy silo, slaughterhouse, refinery, mill, farm.
Commodity equivalent	Measure used to relate the trade flows of different products to a commodity equivalent. This is obtained by using the commodity equivalence factor.	Soy oil and cake products are converted into soybean equivalents.
Commodity-equivalence factor	Factor used to convert the amount of a product into a commodity equivalent.	1 kg of soy meal and oil are equivalent to 1.031 kg of soybeans (3 g are waste).
Decision tree	Outlines the conditional filtering of trade data in order to link commodity exports to a logistic hub.	Each supply chain map manual contains a figure of their respective decision trees.
Distance matrix	The distances between different demand and supply nodes. This is used in the linear programming step to solve the problem of minimising the total distance incurred in meeting all of the demand.	Supply nodes are jurisdictions of production. Demand nodes include exports from ports and domestic demand nodes such as chicken farms for Brazil soy. Distances are based on the available road networks.
HS code	Unique code from the Harmonized System (HS) which describes the nature of the products being traded internationally.	1201: Soya beans, whether or not broken 120110: Soya beans, seed; whether or not broken
Jurisdiction	The territorial administrative units into which a country is divided.	Municipality in Brazil, kabupaten (district) in Indonesia, department in Argentina, department in Paraguay (lower resolution, with departments comprised of districts).
Linear programming	Linear programming (LP, also called linear optimisation) is a method to achieve the best outcome (such as maximum profit or lowest cost) in a mathematical model whose	Use linear program to minimise the distance between logistic hubs and production municipalities.

	requirements are represented by linear relationships.	
Logistics	Activities related to the production, storage, processing, transport, trade, etc., of commodities in supply chains.	Chicken rearing, cattle slaughtering, soybean crushing, palm oil bulking, shipping.
Logistics hub	Jurisdiction containing one or more assets that are nodes in the commodity supply chain.	Municipality, department of silo location, slaughterhouse, palm oil mills.
Node	Jurisdiction, asset, trader or country representing a point of aggregation or transfer of a commodity through its supply chain.	
Supply chain	Sequence of nodes linking a jurisdiction of production to a country of import.	