Supply Chain Greenhouse Gas Emission Factors for US Industries and Commodities - Appendix 1 - Model Equations

Wesley Ingwersen[*]
Mo Li[†]

July 2, 2020

This document is intended to accompany the full report. Please see the full report for details. USEPA. 2020. Supply Chain Greenhouse Gas Emission Factors for US Industries and Commodities. EPA/600/X-20/001.

We use matrix algebra to represent the steps of creating the EEIO models, using conventions for variable names commonly used in the IO literature and the existing USEEIO model documentation when possible. Capital letters indicate matrices and lower case letters vectors. A "^" over a variable represents the diagonalization of a vector as a matrix. An exponent of "-1" represents an inverse. A "′" represents the transposed (rows and columns switched) form of a matrix or vector.

Model construction

First, the national totals of GHGs by industry must be transformed into a matrix of coefficients in the form of national total per industry, also referred to as the GHG satellite account, represented as B in the following equation, with a subscript of I denoting its industry form, and t denoting the USD year the same year as the IO data.

$$B_{I,t} = E_{I,s}\hat{x}_{s,t}^{-1} \tag{1}$$

In the above equation, E_I is a emission x industry matrix of national totals of each GHG by industy sector in year s, and $x_{s,t}$ is a vector of gross output by industry in year s, given in year t dollars. The industries in the E columns match the industries in x.

For x_s to be in year t USD, x must first be price adjusted in the following equation,

$$x_{i,s} = x_{i,a} * \rho_{i,a->b} \tag{2}$$

where $x_{i,a}$ is the year industry ouput for industry i in the currency year, s, corresponding to the year of the national GHG totals. ρ is a price adjustment factor for industry i from currency year s to t USD.ρ for industry i is prepared using the following equation,

$$\rho_{i,a->b} = \frac{cpi_{i,b}}{cpi_{i,a}} \tag{3}$$

where ρ from currency year a to b is the ratio of the industry chain price index for year b to that of year a. These equations are used for each industry in the given model.

The core of an IO model is a direct requirements matrix,A, representing the dollar inputs from other sectors per dollar output. Two A matrices are derived, one in commodity x commodity form, A_C, and the other in industry by industry form, A_I. Both are created using a combination of normalized forms of the Make, V, and Use, U, tables. The Make table is normalized by the commodity output vector, q, and the Use table

[*]Center for Environmental Solutions and Emergency Response, USEPA Office of Research and Development
[†]GDIT Inc.

by the industry output vector, x. The only difference between A_C and A_I is the order of multiplication, as shown in the following two equations.

$$A_C = U\hat{x}^{-1} * V\hat{q}^{-1} \tag{4}$$

$$A_I = V\hat{q}^{-1} * U\hat{x}^{-1} \tag{5}$$

Another name for $V\hat{q}^{-1}$ is the Market Shares matrix. The GHG satellite table for the commodity models must be transformed by this matrix to be in emission x commodity form, using the following equation.

$$B_C = B_I * V\hat{q}^{-1} \tag{6}$$

The items of the B_c matrix represent direct emission factors for commodities.

The final factors must be calculated by calculating a total requirements matrix. Leontief (Leontief and 1906-, 1953) first showed this could be done by subtracting the A matrix from an identity matrix, and taking its inverse. Supply chain emission factors can be calculated by taking the product of the B matrix of direct emission factors by the total requirements matrix, L, where L is defined as

$$L_C = (I - A_C)^{-1} \tag{7}$$

$$L_I = (I - A_I)^{-1} \tag{8}$$

The following two equations show this calculation for the industry and commodity, where

$$M_I = B_I * L_I \tag{9}$$

$$M_C = B_C * L_C \tag{10}$$

This result in emission x sector matrices, M, of supply chain factors, which include direct and indirect emissions per dollar produced for each sector. The factors in M have to be further transformed into purchaser price using sector-specific marginal cost data. Marginal cost data includes cost for distributions, wholesale and retail of commodity. This can be done by creating a vector or marginal cost ratios, Φ_c, both for commodities and industries for their respective models, as i the following equations.

$$\bar{M_C} = M_C * \hat{\Phi}_c \tag{11}$$

$$\bar{M_I} = M_I * \hat{\Phi_I}' \tag{12}$$

Φ_c is a vector of commodity specific purchaser to producer prices.

$$\Phi_C = \frac{q_{pur,c}}{q_{pro,c}} \tag{13}$$

For use in industry models, commodity margin factors, Φ_c, need to be transformed into industry margin factors. The factors. A commodity x industry "commodity mix" matrix (Miller and Blair, 2009), $V\hat{x}^{-1}$, can be used for this purpose.

$$\Phi_I = \hat{\Phi_C} * V\hat{x}^{-1} \tag{14}$$

where, commodity mix matrix, a commodity x industry matrix representing the mix of commodities being produced by industries.

To identify the root sources of emissions for a given supply chain factor in M, we assume \$1 million USD in demand for the related sector, and calculate a scaling vector, s , as

$$s = (I - A)^{-1}y \tag{15}$$

where A is either A_I or A_C, depending on model type, y a vector of the same length as A with zeroes for all sectors except the sector of interest, where the value is set to \$1 million USD. This is turn used to calculate a matrix of total emissions by root emitting sector, M_d

$$M_d = B\hat{s} \tag{16}$$

We also estimate the potential impacts of the margins activities associated with industries and commodities, using the same margin data along with model estimates for supply chain emissions for each sector comprising the margins. We break Φ down into its margin transportation,t, wholesale,w, and retail,r sectors:

$$\Phi_C = \Phi_C, t + \Phi_C, w + \Phi_C, r \tag{17}$$

For each margin components, we allocate the margin across all margin sectors associated with that type, for example, for retail we need to allocate the retail margins across model retail sectors. For this We develop a margin allocation matrix,MA, with rows for each of the margin components and columns for all model margin sectors. We allocate across sectors of a given margin type based on relative commodity output of the model sectors for that margin type. Note this procedure does not result in commodity or industry-specific margin types. We assemble the sector-specific margin coefficients into a commmodity by coefficient matrix, P.We then multiply the margin coefficients matrix by the margin allocation matrix

$$P_a = P * MA \tag{18}$$

The result P_a represents \$ of margin input per \$ producer price for each model sector by margin sector type. We transform the result, P_a, into the shape of the A matrix, to easily match up with the emissions factors, and name the reshaped form A_p.

We derive emission factors for the margins with a multiplication of the original emission factor matrices, M, by A_p.

$$M_p = M * A_p \tag{19}$$

Finally, M_p is adjusted to be in purchaser prices like what is done with M above.

$$\bar{M_p} = M_p * \hat{\Phi} \tag{20}$$

where $\hat{\Phi}$ is for industries or commodities depending on the model.

Data quality scores for emission factors

Data quality scores of 1-5 are assigned based on the US EPA LCA Data Quality system (USEPA, 2016) to values in B for each of five indicators. These be represented in a matrix of the form of B called $B_{dqi,i}$ created for each indicator i. To scale these scores to put them to create scores for the supply chain emission factors in , this is multiplied by L.

$$M_{dqi,i} = B_{dqi,i} * L \tag{21}$$

where B and L may be either both for industry, B_I and L_I or commodity, B_C or B_I models.

Model validation

The validation of the supply chain factors is performed for each model from which the factors were derived. The validation test checks that the model supply chain GHG factors can be scaled back to total GHG industry emissions.

Mathematically, this validation test is based on the input-output identity that $x = Ly$, where y is the total final demand for commodities and industries. Therefore multiplying B by both sides, to scale output using the direct emission factors, should result in equality.

$$B * x = B * L * y \tag{22}$$

This can be simplied in the form of the following equation.

$$Ei = (M * y)i + E_f i \tag{23}$$

where i is a column vector of 1s that acts as a summation vector. E_f is a matrix of emissions associated with the final user, and final user sectors are not part of M. This can be understood such that the sum of the supply chain GHG factors scaled for total final demand is the same as the sum of national GHG totals for each GHG.

The issue with the model M, derived from the total requirements matrix, is that it includes foreign input requirements, and that final demand also includes imports. To use the model to estimate US-produced GHGs, then M has to be created with domestic requirements only. This requires a recreation of an A matrix with only domestic requriements. To do this, we modify the original Use table, U, to remove foreign requirements using the Import matrix, I_m.

$$U_d = U - I_m \tag{24}$$

and then we subtract the final demand provide in the import matrix, Y' from the final demand in the original Use table, to get domestic final demand, Y_d.

$$Y_d = Y - Y' \tag{25}$$

Then the A matrix with just domestic direct requirements, A_d, can be created using a similar derviation used for A above. Note that normalized Make table transactions are not the same. The follow equation is demonstrated below for the commodity form of the model.

$$A_{d,C} = U_d \hat{x}^{-1} * V \hat{q}^{-1} \tag{26}$$

We can follow from the steps to create M from A and apply them to create M_d using the A_d. This refines the data validation check to compare national GHG totals only to the sum of the supply chain factors based on domestic requirements scaled by the domestic final demand, as in the following equation:

$$Ei = (M_d * y_d)i + E_f i \tag{27}$$

The difference between the left and right terms of the equations is evaluated per GHG and explained as a model error.

References

Leontief, W.W., 1906-, 1953. Studies in the structure of the American economy. http://agris.fao.org/agris-search/search.do?recordID=US201300615417 (accessed 10.16.2019).

Miller, R.E., Blair, P.D., 2009. Input-output analysis: Foundations and extensions. Cambridge University Press.

USEPA, 2016. Guidance on Data Quality Assessment for Life Cycle Inventory Data (Report). US Environmental Protection Agency, National Risk Management Research Laboratory, Life Cycle Assessment Research Center. https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=321834