MASSACHUSETTS INSTITUTE OF TECHNOLOGY

15.093: Optimization Methods

United Nations
Annexation of Food Distribution

Proposal and Analysis

April 10, 2023

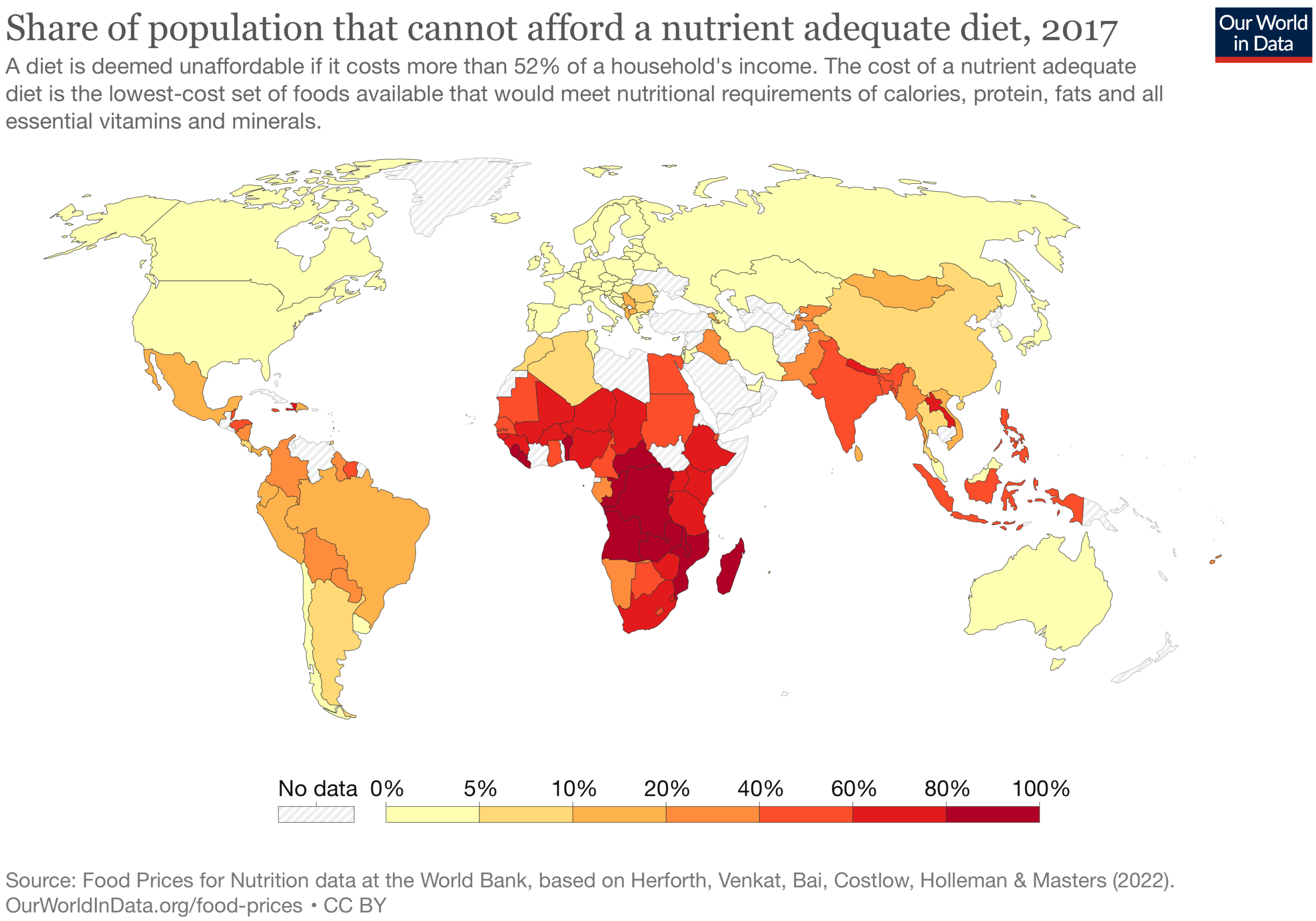

Figure 1: Undernourishment by Country

1 Motivation

From 2019 to present day: The number of those facing food insecurity has risen from 135 million to 345 million. Food insecurity is not normally distributed across the world. Areas such as Sub Saharan Africa and South Asia have large population sizes, with food production constraints. In addition to this inequality of resources, we are now approaching a possible economic recession with still-tightened financial resources due to the COVID 19 pandemic.

1.1 Problem Statement

Our goal is to minimize experienced hunger worldwide in the most cost-efficient manner possible. We take the stance that a global body has annexed the distribution process of all produced goods world-wide.

2 Data

The straightforward approach to this problem would be to consider total food production and demand from every country, and aim to transport food to best meet demand worldwide while minimizing cost of transportation. While this general concept is valid, in reality food production and demand require more nuance than simply observing weight of food produced. Firstly, undernourishment is not just based on total calories consumed, but on the nutrients of foods consumed; a diet of only fats or only carbohydrates would not adequately address nutritional needs. Secondly, each country has a unique distribution of food production, and this distribution must be considered. In the following section, we detail the large variety of data sources we utilize to create an optimization problem rooted in the reality of human nutritional needs.

2.1 Data Sources

Data was obtained from five unique sources:

- Food and Agriculture Organization: Production of 74 food items for 168 countries for the year 2013. Also contains geographic coordinates for every country.
- USDA Food Data Central: Nutritional content (fats, carbohydrates, proteins, fruits and vegetables) for every food item.
- US Department of Health: Dietary guidelines for average American adult (fats, carbohydrates, proteins, fruits and vegetables).
- US Census Bureau: Population for every country in 2013.
- Calorie requirements: Minimum calories required per person in each country; used to scale the dietary guidelines for each country.

In combination, these datasets allow us to create a singular data frame of size 168x74x4 with the total nutritional production for every country, and the total nutritional requirements for every country.

3 Model Formulation

Our initial model (noted as simple model) only focused on direct shipments, i.e. shipments of food from one country to another, with no consideration for food being shipped through countries as "hubs", and no consideration for robustness. We developed further models considering both a network distribution and robustness, as noted in the following sections. For the sake of clarity and brevity, this report contains the formulation for our full model which considers both a network distribution and robustness.

3.1 Network Distribution

3.1.1 Motivation

Direct shipments of food products from every source to every destination country is expensive, unrealistic, and environmentally costly. To address this, we consider distribution hubs. To motivate the use of distribution hubs, we consider that, for large volume shipments along the same route there are often significant discounts in per unit cost due to the certainty and security of those continued shipments. This is a reasonable assumption for a real-world application.
By considering hub shipments, we can further reduce the average trip distance, and (by accounting for discounted bulk shipments) we can further reduce the total cost of transportation.

3.1.2 Implementation

We implement this by adding three new decision variables: Y, Z, and gamma. Y is the quantity shipped as a source-hub shipment, and Z is the quantity shipped between two countries as a hub-destination shipment. Gamma is a binary indicator for whether a path contains a high enough volume to be considered a hub shipment (past the set volume threshold, there is a 20% discount in transportation cost).

3.2 Robustness in Supply & Demand

3.2.1 Motivation

Food supply is an unpredictable thing. Production can be affected by a vast number of stochastic factors from environmental disaster, to bad luck, to war and man-made challenges. For this reason, it is clear that it is unrealistic to expect food production to hold steady at the levels from the time of the collection of this data set.

Food consumption, while not initially a value which should be robust, can exhibit unpredictable fluctuations as well. Influxes of populations, food disasters, and domesticated animal consumption can all be causes of uncertain consumption.

3.2.2 Implementation

Following implementations in lecture, we instantiate 50 different matrices of the same dimensions of supply filled with random values normally distributed in the range [0.9, 1]. These matrices are multiplied element-wise by the supply matrix to produce scenarios of uncertain production.

This method gives negatively skewed results to present a more conservative scenario (that is, approximately 90% average production of the real-world values measured in 2013[1]).

For demand, a nearly congruent approach is taken, except perturbations in demand exist in [0.9, 1.1]. This is a moderately conservative approach - a more conservative approach would limit food allocation and reduce the likelihood of obtaining interpretable results from our model.

3.2.3 Sets

- N = [168]: Countries considered
- K = [74]: Food production classes
- Z = [50]: Uncertainty set size (number of scenarios considered)
- T = [4]: Types of nutrition (carbohydrates, protein, fat, fruit & vegetables)

3.2.4 Decision Variables

- $X_{ijk} \in \mathbb{R}^{n,n,K}$: decision variables that denote the amount of food type k to transport from source country i to destination country j, in units 1000 tonnes
- $Y_{ilk} \in \mathbb{R}^{n,n,K}$: decision variables that denote the amount of food type k to transport from source country i to transportation hub l, in units 1000 tonnes
- $Z_{ljk} \in \mathbb{R}^{n,n,K}$: decision variables that denote the amount of food type k to transport from transportation hub l to destination country j, in units 1000 tonnes
- γ_{ij}: a binary variable that indicates whether country j is a hub for country i

[1]This is the most recent data collected and available from the United Nations.

3.2.5 Parameters

- s^z_{ik}: The amount of supply for each food type k in source country i in scenario z
- d^z_{it}: The amount of demand for each nutrition type t in source country i in scenario z
- c_{ij}: The arc length of the circle which intersects both countries i and j
- W_{kt}: the amount of nutrition t in food type k
- λ: scaling factor; the penalty of leaving 1 person hungry[2]
- y: a 4 by 1 matrix that denotes the inverse of the amount of nutrients needed per person per day, $\begin{bmatrix}\frac{1}{k} & \frac{1}{p} & \frac{1}{f} & \frac{1}{\nu}\end{bmatrix}^T$
- P_i: the population size for country i
- G: the constant which converts 1000 metric tonnes to 100 grams, 1e+7
- T: The threshold for considering a country to be a hub relative to another country
- α: This proportion weights the two objectives inversely so that we can solve the model depending on our set priorities
- H: This is the unit hourly cost of operating a Boeing 747 per km per tonne
- F: auxiliary variable that denotes the maximum total number of people left hungry

3.3 Formulation

$$\min_{X,Y,Z,\gamma,F} \quad \alpha H\left[\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij} * \sum_{k=1}^{K} X_{ijk}\beta_D + Y_{ilk}\beta_H\gamma_{il} + Z_{ljk}c_{lj}\beta_D\gamma_{lj}\right] + (1-\alpha)\lambda F$$

$$\text{s. t.} \quad \frac{1}{P}\left[d^z_{jt} - G\left(\sum_{i=1}^{N} X_{ijk} + \sum_{l=1}^{N} Y_{ljk}\right) W_{kt}\right]\frac{1}{y} \leqslant F \qquad \forall j \in [N], t \in [4] \tag{1}$$

$$\frac{1}{T}\left[\sum_{k=1}^{K} X_{ijk} + Y_{ijk} + Z_{ijk}\right] \geqslant \gamma_{ij} \qquad \forall i \in [N], j \in [N] \tag{2}$$

$$\sum_{i}^{n} X_{ijk} + \sum_{i}^{n} Y_{ilk} \leqslant s^z_{ik} \qquad \forall j \in [n], l \in [n], k \in [K], z \in [Z] \tag{3}$$

$$\sum_{i}^{n} Y_{ijk} = \sum_{l}^{n} Z_{jlk} \qquad \forall j \in [n], l \in [n], k \in [K] \tag{4}$$

$$M * \gamma_{ij} \geqslant \sum_{k} Y_{ljk} \qquad \forall j \in [n], l \in [n] \tag{5}$$

$$\sum_{j}^{N} \gamma_{ij} \leqslant 5 \qquad \forall i \in [N] \tag{6}$$

$$\gamma_{ij} \in \{0,1\} \qquad \forall i \in [n], j \in [n] \tag{7}$$

$$F \in \mathbb{N}^+, \qquad X \in \mathbb{R}^+ \tag{8}$$

Interpretations for the formulation are as follows:

[2]Optimistically, this can be considered the value added of feeding each person

- **Objective function**: minimize the sum of total transportation costs and the penalty of leaving people hungry
- **Constraint (1)**: Auxiliary constraint; this constraint computes the maximum number of people that can be left hungry per day by this food distribution plan. This considers demand under uncertainty.
- **Constraint (2)**: A source country should only select hubs of distribution if those hubs exceed a certain threshold of mass shipped along that particular path.
- **Constraint (3)**: Supply constraint; the quantity of food utilized originating from a certain country should not exceed the supply capabilities of that country under robustness.
- **Constraint (4)**: Hub constraint; for simplicity, we only allow a country to select up to five other hubs for network distribution.
- **Constraint (5) & (6)**: Bounds and dimensional constraints. Quantity shipped is positive, counts of people are not continuous, and we must have γ as a binary indicator variable.

4 Results

While transportation costs are considered in our objective function, we do not want cost to be the limiting factor in addressing undernourishment; our problem assumes that addressing this issue involves taking on cost to begin with. Thus, we report the number undernourished as our primary outcome.

Our simple model that considers all 168 countries indicated that, contrary to our initial intuition[3], there is enough food production to feed all persons when robustness is not considered. Today, according to a UN report, there are 828 million people suffering from nutritionally inadequate diets. Our first model was able to reduce the number of people suffering from nutritionally inadequate diets to 0.

From here, we advance to the formulation previously shown which accounts for network distribution and robustness through uncertainty in both supply & demand. The complexity of this formulation is high and the computational demands are intense; due to practicality of the computing demand, we implemented our robust-network model on the first 40 countries in our data set, and compared the results of this implementation to the simple model for the same 40 countries, as well as a network model for the same 40 countries. Our results from this advancement are as follows[4]:

Formulation	Cost	Tonnes sent via Direct Shipments	Tonnes sent via Hub Shipments	Number of People Left hungry
Simple	\$6,088,625	11,024,159	————————	630,624
Network	\$5,768,030	127,337	10,896,822	630,624
Robust-Network	\$6,003,435	169,940	10,179,521	695,482

[5] This table shows the relative costs of each practice.

Firstly, these results are only for a subset of 40 countries - thus, the rationale for the presence of undernourished citizens in these models (as compared to no undernourished citizens in the larger simple model with all 168 countries) is that the 40 countries did not contain certain countries with high production that are necessary to nourish all individuals.

[3]and in line with popular knowledge apparently

[4]These results are computed using hyper-parameters discussed in 5.1

[5]This calculation is based on the same parameters chosen for our model, and only accounts for the cost incurred by hunger; this neglects the global transportation cost of food as this information is not reliably available.

For reference, in transportation, the United States alone spends $4 billion USD per year on international food assistance[6]. If we evaluate our costs (for these models considering only forty countries) with no context, we arrive at the conclusion that this robust network flow model can provide global transportation routes at a minor price while reducing hunger by a large factor[7] **with only current production levels**, all with significantly higher levels of equity in the distribution of food.[8]

However, it must be considered that due to the arbitrary nature of our hyper parameters, our costs cannot be evaluated in comparison to a real-life setting. Indeed, our costs can only be utilized to compare the efficacy of each model in comparison to one another. However, the presence of the hyper parameters in our model allows for efficient and simple parameter tuning to fit the model to a real life application, such as a non-profit organization aiming to most effectively distribute food while considering changing budget constraints.

The amount of food sent via direct shipments, when compared with those via hub shipments in our models, indicates that network flow adds value to our model; many tons are sent via hub shipments, which allow for more efficient distribution.

We can observe in Figures 3 and 4 (which represent our robust-network model) that certain countries are large food suppliers, while others are "sinks" - for example, China (while having high food production) requires a large influx of nutrition due to the massive size of its populations. Countries such as Brazil have a surplus of goods and are apply to positively impact other countries through their exports. A global approach would require investment into purchasing goods from areas such as Brazil.

We conclude this is indeed a tractable problem for optimization, and that, in the context of the implementation space, it would make sense to leverage this approach. Additionally, we conclude that it would make humanitarian sense for a globally governing body to annex the task of transportation of surplus food production from each country to feed the disadvantaged. Furthermore, global humanitarian organizations can leverage our approach to more effectively address unique nutritional deficiencies in individual countries, rather than taking a generalized approach.

4.1 Extensions

If given more time for this project, we agree that an interesting extension would be to attempt to implement Bender's decomposition in this scenario. Additionally, this implementation only accounts for single hub transportation, but global networks could benefit from multi-hub transportation. The main limiting factor in this approach is time and computational resources.

Furthermore, factoring in global food production costs and better informed costs for other considerations would allow for an incredibly high performing model capable of being implemented today to make tangible changes.

[6]in both emergency food assistance to avert humanitarian crises and development assistance to support agriculture and related sectors

[7]828 Million people go hungry every day as per 2022 UN World Hunger report

[8]Equity was not a factor we controlled for, nor one that we had time to implement as an objective or constraint; however, this is an unintentional effect of our model that we see in the results.

5 Appendix

5.1 Hyper-Parameters of results

The hyper-parameters chosen in our results are as follows:

Hyper-Parameter	Value	Units
λ	250 * 365.25	$USD * Days
T	100	Metric Tonnes
α	0.95	[None]
β_D	1	[None]
β_H	0.6	[None]

The reasoning for choosing these values as such is as follows:

- λ: The value of 250 as a factor in this constant is chosen as the daily cost of not feeding an individual. There is no theory, nor agreed upon value, for such an idea. This value was chosen in concert with other values because our empirical evidence shows that it works to create a solution in which undernourishment is addressed.
- T: This is the capacity of 15 Boeing 747-400 cargo planes.
- α: We want to minimize transportation cost but not at the expense leaving individuals hungry.
- β_D: The discount in price is relative, so we do not discount direct shipments.
- β_H: We discount the cost of hub shipments to promote the use of hub shipments indirectly.

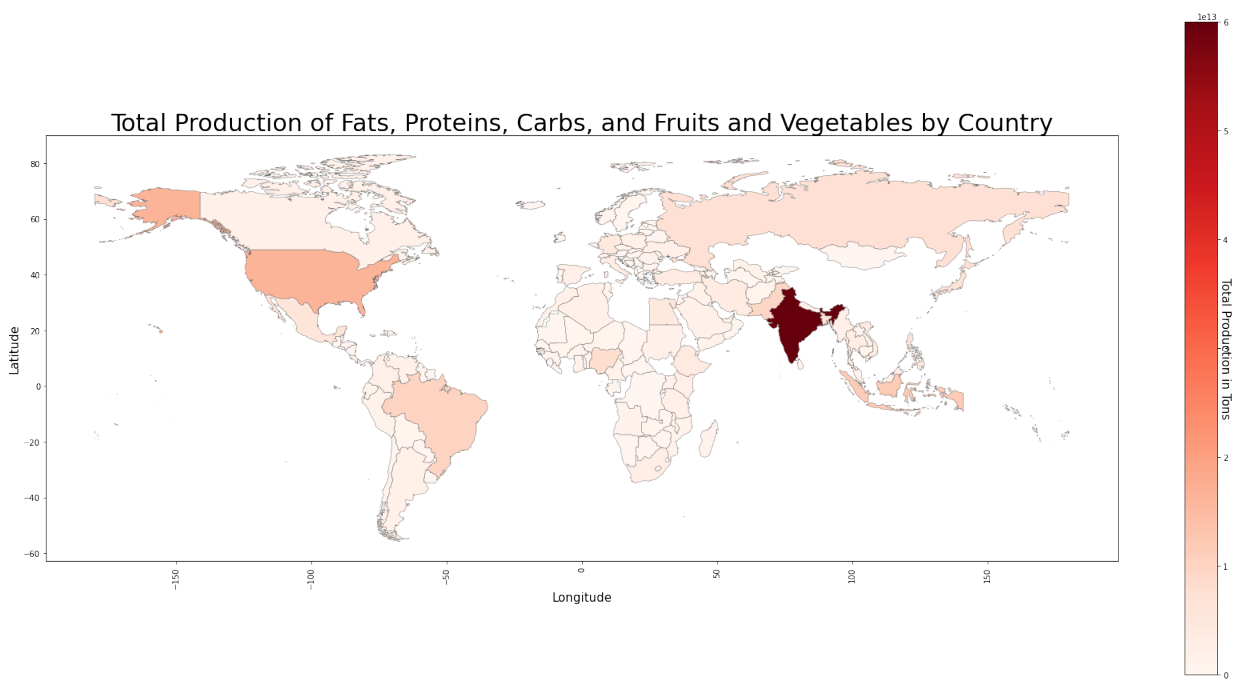

Figure 2: Relative Food Production by Country

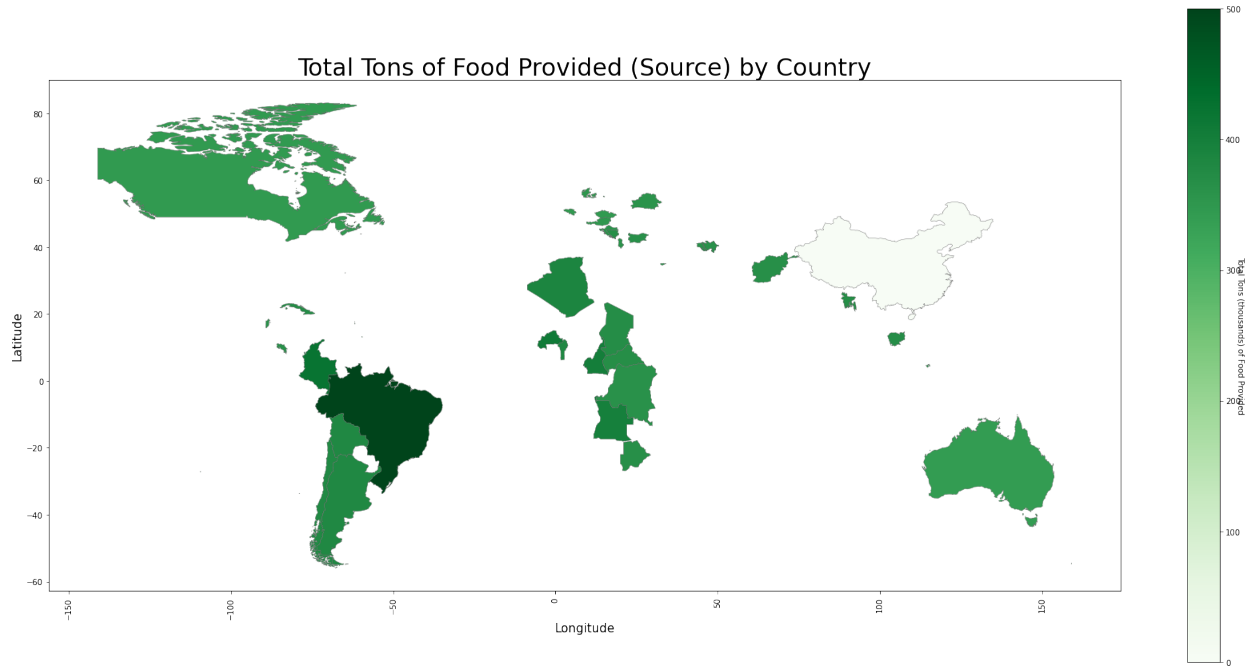

Figure 3: Source by Country in Subset Model

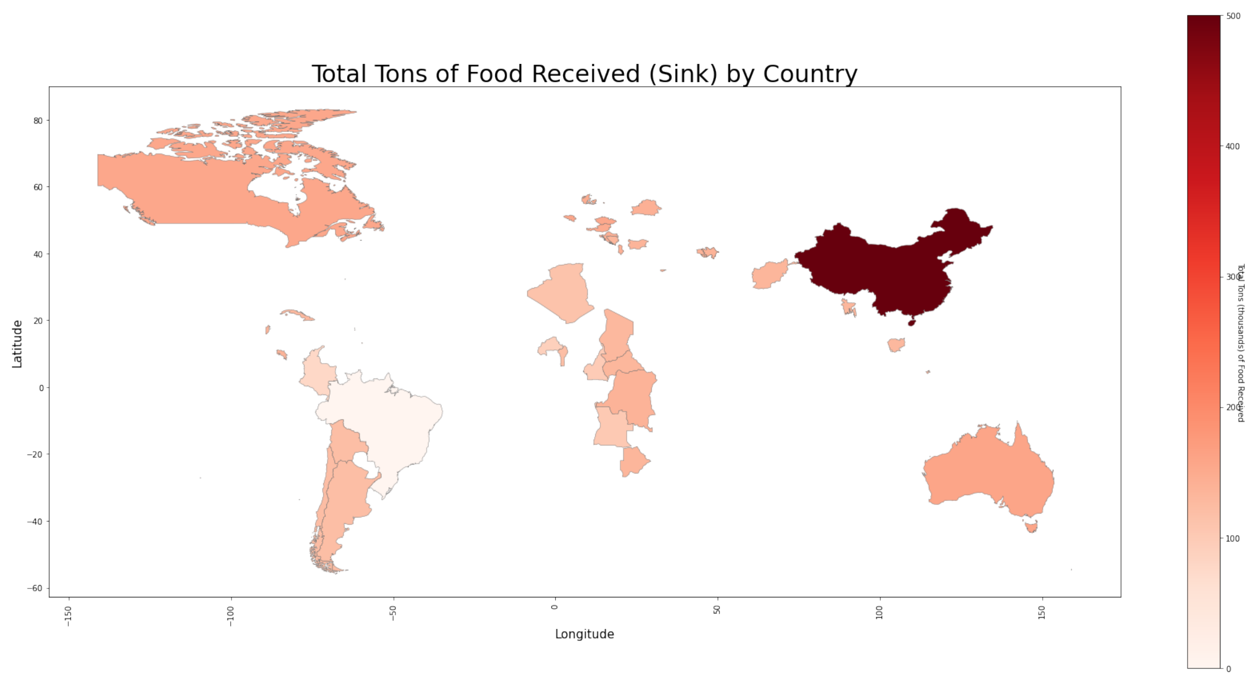

Figure 4: Sink by Country in Subset Model