INFERRING GEOSPATIAL PROPERTIES OF THE GLOBAL FOOD SYSTEM NETWORK

MAXWELL KAYE, School of Earth and Planetary Sciences

McGill University, Montreal

AUGUST, 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Master of Earth and Planetary Science

Abstract

Food enters the earth system through production, moves through geospatial trade fluxes, and then is inevitably consumed by humans, used for feed or biofuels, or lost. The diversity of food types, and complexity of the potential pathways that food can take from production to human consumption, make it challenging to gain a global understanding of geospatial flows, inhibiting progress toward sustainably meeting global food consumption demand. This thesis presents a holistic, globally gridded, energetic perspective on the global food system. We take a global account of all aggregated metabolic energy available in our food supply, and put it in one-to-one correspondence with crop and livestock production. We utilize energy rates of both production and consumption at one degree resolution to infer food trade quantities between grid cells. We construct four idealized models to explore the topology of global food trade, provide new metrics for measuring regional participation in the globalized food system, and establish preliminary upper and lower bound estimates on topological features that may constrain possible food system reform.

Abrégé

Les aliments entrent dans le système terrestre par le biais de la production, se déplacent à travers les flux commerciaux géospatiaux, puis sont inévitablement consommés par les humains, utilisés pour l'alimentation animale ou les biocarburants, ou perdus. La diversité des types d'aliments et la complexité des voies potentielles que les aliments peuvent emprunter de la production à la consommation humaine rendent difficile l'acquisition d'une compréhension globale des flux géospatiaux, ce qui entrave les progrès visant à répondre durablement à la demande mondiale en matière de consommation alimentaire. Cette thèse présente une perspective énergétique holistique et quadrillée à l'échelle mondiale du système alimentaire mondial. Nous prenons en compte l'ensemble de l'énergie métabolique agrégée disponible dans notre approvisionnement alimentaire et nous la mettons en correspondance avec la production végétale et animale. Nous utilisons les taux d'énergie de la production et de la consommation à une résolution d'un degré pour déduire les quantités d'aliments échangés entre les cellules de la grille. Nous construisons quatre modèles idéalisés pour explorer la topologie du commerce alimentaire mondial, fournir de nouveaux paramètres pour mesurer la participation régionale au système alimentaire mondialisé et établir des estimations préliminaires des limites supérieures et inférieures des caractéristiques topologiques qui peuvent limiter une éventuelle réforme du système alimentaire.

Acknowledgements

I am deeply grateful to each and every individual who has helped me along this research journey. First and foremost, thank you to Eric Galbraith for not only providing guidance as I stumble through the maze of academia, but also for being a constant source of wisdom, curiosity, and energy which has challenged me to become a better student and person. For not only tolerating my incessant curiosity and skepticism, but encouraging these traits, you have my deepest gratitude.

Thank you to the the McGill EPS Department for their support and generous funding which made this work possible. Thank you to Graham McDonald for providing insightful feedback and ideas on the advisory committee. The IESD team has been instrumental in getting this research off the ground. Particularly, thank you to my colleagues Abdullah-Al- Faisal, William Fajzel, and Tanya Matitia. Finally, thank you to my friends and family, whom I love dearly, and to all of those who have supported me during the last two years.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

As climate change and international instability threaten the lives and livelihoods of billions around the world, the need for a more sustainable and resilient food system becomes increasingly apparent [Li et al., 2023, Poore and Nemecek, 2018, Ritchie et al., 2022]. The problem is an urgent one, with some research projecting that, without innovation or reform, the environmental impacts of the food system will increase between 50 and 90 % by 2050 [Springmann et al., 2018]. Unfortunately, the complexity and interdisciplinarity of international trade barriers, cultural differences, geographic variations in productivity, nutritional requirements, economic forces, and other components of the food system inhibit a unified understanding of its properties [Dalgaard et al., 2003, Jones and Ejeta, 2015, Picchioni et al., 2017]. While much is known about these aspects in isolation, a collection of partitioned global food system analyses might not be sufficient to address the transdisciplinary nature of some of the more difficult problems facing humanity, including carbon emissions, world hunger, chemical pollution, water usage, animal cruelty, and other widespread challenges [Conijn et al., 2018, Gerten et al., 2020].

This thesis is an effort to present a new perspective on addressing these problems, by accounting for and studying the physical properties of the global food system, namely the rates of production, consumption, and spatial transportation of food energy, and to explore the system's underlying organizational dynamics. Our goal is not to accurately predict these food flow quantities, but rather to understand the topological properties[1] that govern the network of gridded food flows. In this way we simultaneously account for the physical consumption and production of most of the food consumed and produced on earth, and constrain the nature of how that food goes from production to consumption, while favoring both parsimony and physical reality in our representation of the food system. We aim to supplement models like MagPie [Dietrich et al., 2019] and GTAP [Corong et al., 2017], which predict global food trade data with fits based on a number of economic parameters, by furthering our understanding of how food flows physically with minimal parameters. While the simplifying assumptions we make mean our approach might lack some short-term predictive power when compared to a model with more parameters, its simple physical nature is by design accessible to a plethora of different fields, and may be more amenable to multi-decadal timescales.

1.1 Why study the Food System

The global food system accounts for nearly half the global workforce [Fajzel et al., 2023], nearly a fifth of carbon emissions [Ritchie et al., 2020], and 32% of final energy demand [Sims, 2011]. 38% of land surface area is used for agriculture [FAO, 2020]. 10% of the world lives in hunger, and 13% is obese [WHO, 2021]. Food production growth has outpaced population growth by a factor of 1.5 in recent years [OECD, 2018]. Humans spend on average 54 minutes per day working on food growth, collection, and processing, which accounts for over 6% of all human waking hours, or over 11.6 billion human hours per day [Fajzel et al., 2023]. As society continues to globalize and we face international crises, a mechanistic understanding of the food system can be invaluable.

Many recent works relating to the food system have emphasized its transdisciplinary nature. Research has shown that health and food security are innately connected [Hawkesworth

[1] Some examples of topological properties of a network include density (how sparsely connected the nodes are), degree distribution (how many nodes there are with 10, or 100, or with 1000 connections), how do these properties differ between in and out going connections, etc.

et al., 2010]. Others have examined the link between food value chains and malnutrition [Gómez and Ricketts, 2013]. The connection between water food production and water consumption is also well known [Hoekstra and Mekonnen, 2012, Hoff et al., 2014, Mekonnen and Gerbens-Leenes, 2020]. Some have pioneered assessment of how a sustainable future can and must address malnutrition and malnourishment [Willett et al., 2019], and others still have investigated the link between increasing meat production, health, and the environment [Godfray et al., 2018].

Furthermore, the environmental impacts of the food industry are well established [Roser et al., 2013, Kriewald et al., 2019]. In order to transform our food system to exist within the planetary boundaries, we must think holistically about the system, from initial production to final consumption [Conijn et al., 2018, Gerten et al., 2020], and develop models of how the food system could fit into these boundaries [Gerten et al., 2023]. In particular, the food system has to undergo dramatic change quickly, because it is a large component of green house gas emissions [Ritchie et al., 2022]. The food science community has established that both production and consumption patterns must change in order to synergistically give us the best chance at climate change mitigation [Poore and Nemecek, 2018].

A uniform, gridded, global account of the food system can potentially provide a basis for reform at the scale necessary to avert crisis. Current shortcomings of gridded data products include a lack of hybridization and accessibility [Kim et al., 2021]. Additionally, recent work by the OECD has identified many knowledge gaps with regard to the study of the global food system, including that "Existing evidence is often not detailed enough to be useful" and "Inconsistent methodologies may prevent comparison of available evidence" [Deconinck et al., 2021]. With the present work we seek to address these issues.

1.2 What is the Food System

This question has been asked by many prior authors, each landing on a wide variety of points of emphasis. The global food system includes an economic industry, a component of human health, a set of international trade agreements, and more. Food enters the system through production, which includes growth and collection of crops and livestock, and then leaves the food system through a variety of consumption paths. Food is consumed by humans, by animals and livestock, used for bio-fuels and other uses, wasted, processed into other foods. In table 1.1 we propose a set of definitions of fluxes that comprise the components of the food system. For each definition, we reference the data source for our analyses, which we attempt to reconcile with our definitions. Here we take a wide definition of food, which might be more aptly labeled as agricultural crops and livestock. In this thesis, "food" generally refers to anything that could be metabolized by humans, but doesn't necessarily have to be (eg. biofuels are considered food under this definition).

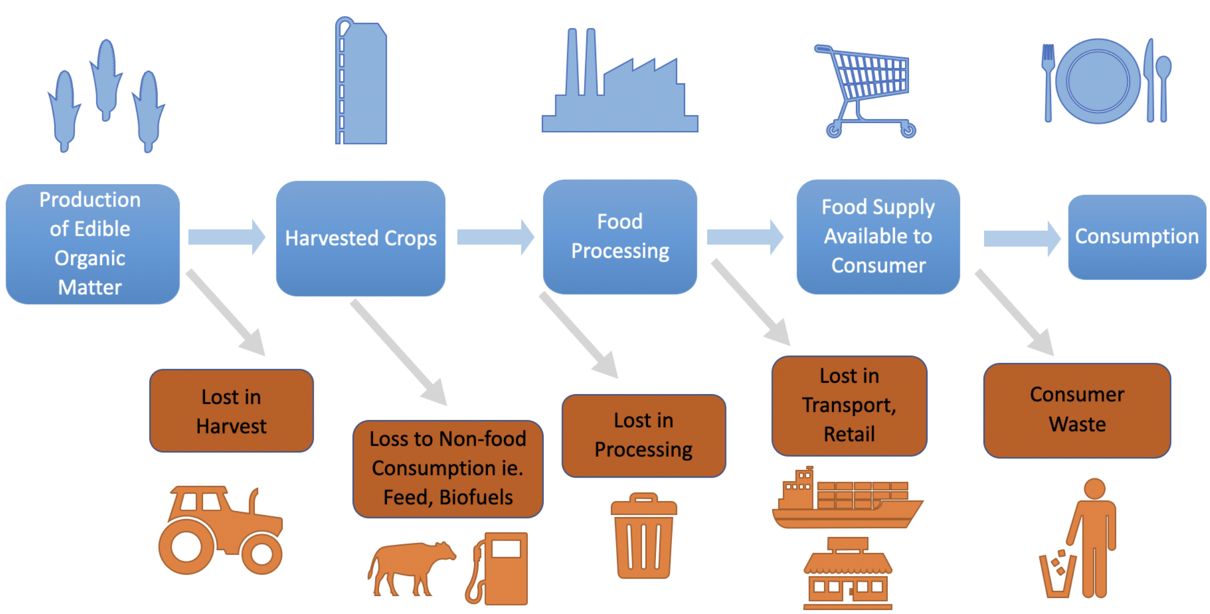


Figure 1.1: Loss and Waste Flow Diagram

Production to consumption account of food use, loss and waste fluxes, as defined in table 1.1.
In this thesis we do not directly account for different types of loss (orange bubbles), but rather balance the food system to be closed so that total production equals total consumption.

Food System Flux	**Definition**	**Data Source**
Production	Growth of edible animal or plant organic matter at the point of harvest.	GAEZ+ 2015 [Grogan et al., 2022] and FAO GLW [Gilbert et al., 2018]
Consumption and Food Supply	We use these terms interchangeably to refer to food available to be eaten by humans.	FAO FBS [Food and Agriculture Organization of the United Nations, 2020b] and WB Population [WorldBank, 2020]
Inflow/ Imports	Food flowing in to a food network node (either country or grid cell).	FAO DTM [Food and Agriculture Organization of the United Nations, 2020a] (for national trade)
Outflow/ Exports	Food flowing out of a food network node (either country or grid cell).	FAO DTM [Food and Agriculture Organization of the United Nations, 2020a] (for national trade)
Waste	Food that is made available to consumers but is instead discarded.	Estimated from Alexander et. al. [Alexander et al., 2017]
Loss	Edible organic matter that could have been consumed by humans, but was never made available to consumers as Food Supply. Includes food that is lost before it reaches market due to improper use, lack of adequate storage, processing or transportation capabilities [Gustavsson et al., 2011], as well as feed and other non-human food uses.	FAO FBS [Food and Agriculture Organization of the United Nations, 2020b]

Table 1.1: Glossary of Food Fluxes

Definitions of fluxes in the food system as used in this thesis, each of which can be measured in either mass or energy per unit time.

1.3 Food Flow Units

Food flows can be measured in a few different ways. The most common are mass (tons), value (USD), water volume (km^3) and energy (calories), all per unit time. It is worth considering environmental impact metrics as well, including green house gas emissions and water usage. Prior work has made significant headway in comparing the different types of food quantifying metrics - namely value, calories, land-use, and water-consumption [Hoff et al., 2014, MacDonald et al., 2015, Mekonnen and Gerbens-Leenes, 2020]. While we consider input data from masses, energies, and population counts, we attempt to compare food quantities primarily through energy rates, across every region of the world and through nearly every type of food consumed.

Examining caloric energy flows within the food system can offer valuable insights, as it is a standardized method for quantitatively comparing different foods based on how much energy they provide to humans. Furthermore there is the added benefit that energy is a natural quantification for both production data and consumption fluxes, providing a bridge that links the introduction of food to the system with its final destination. Prior research has quantified caloric flows [Cassidy et al., 2013], by exploring the distribution of calories through the food system, helping to understand how different regions contribute to global food security in terms of energy provision. Studies also show that, given recent trends in global food crop production, it is unlikely that the UN's goal of eliminating under-nourishment of by 2030 will be achieved [Ray et al., 2022]. Various approaches have been proposed to address hunger by the redistribution of food energy, including reducing food waste and loss [Alexander et al., 2017], or by replacing consumption of animal products with more efficient foods (the conversion of animal feed to human food has been estimate to have a 7%–8% efficiency [Shepon et al., 2016]). We seek to further existing research on food systems through this energetic lens by comparing food production and consumption energy rates in between global grid cells.

Some researchers have investigated the geospatial distribution of food supply through the concept of the "food shed", which was introduced in the 20th century as an analogue to watersheds, to conceptualize a region around a population center that can provide all necessary food [Kloppenburg et al., 2017]. Food shed research tends to focus on the capacity of lands surrounding cities to support the diet of the population, and how much food must flow in order to sustain populations [Schreiber et al., 2021]. The investigation of how each city or population center might be sustained through only local food production, and thus significantly reducing the environmental costs of the global food supply chain, is alluring. However, it has been shown that local food crop production can only fulfil demand for less than one-third the world population [Kinnunen et al., 2020]. Bearing this in mind, we seek to study the geospatial properties of energy flows through the food system without the stipulation that food must necessarily be supplied locally.

1.4 Food System Networks

Networks are a useful framework for understanding the complex interconnections and flows within global food systems [Tu et al., 2019, Karakoc and Konar, 2021]. When we observe a system that has billions of actors with complex motivations, it becomes useful to conceptualize the system as a network and study it on a large scale, particularly through the lens of a network flow model [Borgatti and Halgin, 2011]. By modeling locations as nodes and production-consumption flows between them as edges, some of the essential properties of how food flows can be easily revealed, namely as the dynamics of food production, distribution, and consumption [Aboah et al., 2019]. Generally, food system network nodes are either countries, sub-national regions, firms or population centers (cities). In this section, we will explore some notable contributions to the research of the global food system networks.

1.4.1 International vs Gridded Food Networks

Numerous works have studied the network of international food flows [Ercsey-Ravasz et al., 2012, Wang and Dai, 2021, Tu et al., 2019], including the topological properties of food trade [Konar et al., 2018], the network of meat trade between countries [Chung et al., 2020], and how food trade contributes to green house gas emissions [Li et al., 2023]. The increase of national dependence on international trade can be quantified using network trade metrics [Kummu et al., 2020]. These analyses have the advantage that food consumption and production data are often collected at the national scale. When it comes to the global food system, it has been pointed out that, while international trade flows [Tu et al., 2019] are theoretically interesting and certainly useful to study, ”more granular analyses are need to capture the real-world nuances of the actual global food system” [Puma, 2019].

Some have studied components of the gridded food system such as crop land [Tubiello et al., 2023], water footprints [Hoekstra and Mekonnen, 2012, Hoff et al., 2014], and food production [Grogan et al., 2022, Yu et al., 2020]. Unfortunately, there is a lack of bilateral trade data between geographic grid cells, which has inhibited scientists from using a network approach to study trade between arbitrary geographical locations, rather than jurisdictions. As a result, trade data is typically reported between jurisdictional areas, making it challenging to trace the precise flows of food commodities within and between regions. However, advancements in high-resolution global information systems have opened up new possibilities for overcoming this limitation. These systems provide comprehensive data on agricultural production, trade, and consumption, enabling the construction and study of more detailed food system networks.

1.4.2 Networks of Firms

An alternative perspective on the global food system is the economic study of the food system network as a collection of supply chain nodes, or firms. Some recent foundational work on ”Global Value Chains” [Gereffi, 2019] has provided quantitative and qualitative theoretical insights into the interconnectedness of global supply chains and international trade. The GVC method provides an excellent framework for understanding commodity flow between firms. However using firms as nodes has the shortcoming that it is not physical, and therefore is difficult to integrate with physical measurements such as geographic landcover, populations, etc. Furthermore, firm data is often proprietary and inaccessible to the scientific community for research on the global scale, leading to analyses focused on specific food sub-industries.

While the global food system can be studied with conventional economic metrics such as GDP and intensification/extensification [Coomes et al., 2019, Manzini and Accorsi, 2013], the value of a commodity varies from region to region, month to month, and transaction to transaction. Hence the abstraction process from a region's physical properties to financial metrics results in a loss of information. The methods that follow develop a model for the flow of food between different regions in the world based on physical parameters.

1.4.3 Food System Network Properties

There has been some recent effort to study food system networks in the abstract, in order to glean topological and systematic properties that distinguish it from other supply chain networks. In a series of papers, researchers have introduced the application of complex network theory to food systems across scales [Konar et al., 2018], and used properties such as the degree distribution, strength degree relationship, graph sparsity, and some linear programming methods to construct a food flow model between United States counties [Lin et al., 2019], which they later improved upon and developed [Karakoc et al., 2022]. These works were foundational for beginning to systematically understand food system networks, and have motivated the present study.

1.5 Our Approach

This thesis aims to study a novel food system network that complements prior approaches. We move away from jurisdictions and non-energetic flows, toward an energetic, gridded food flow data set, which enables study of the fundamental grid network topology[2], and exploration of whether the underlying organizational dynamics of the global food system might be simpler than the sum of their parts. While any given supply chain of a particular food product can be exceedingly complex, accounting for all of them simultaneously brings the added advantage of being able to apply global production and consumption constraints in order to study emergent systematic behavior.

Our global food system network uses a grid of 1 degree resolution. We chose this resolution since it is fine enough to capture trade between cities and large population centers as well as large scale farming regions, and yet coarse enough for network analysis on the grid cells as nodes. At this resolution, each grid cell covers approximately $10,000$ km^2 and there are roughly $16,000$ grid cell nodes contributing to the global human food system. Our method emphasizes food mass and energy flows as a way to measure the physical parameters of the food flow network, as opposed to the more arbitrary monetary value of food trade, which can be distorted by a variety of nonphysical factors including local supply and demand, financial markets, and political entities.

The unavailability of intra-grid trade data poses a considerable challenge when inferring food flows for a uniform global grid network of food energy flows. The lack of detailed information on trade between smaller geographical units makes it difficult to accurately map the movement of food within a network. Overcoming this data scarcity requires innovative methods and approaches to infer food flows[3].

[2] We use the term *topology* to denote the general idea of the shape of a network.

[3] This thesis formulates and analyzes idealized energetic systems in an effort to understand big picture trends and lay the groundwork for future research. It is not an attempt to represent realistic socio-political food systems with all there complexities, but rather to isolate and study the food system's physical properties.

The global food system operates through four main types of energy flows: production, consumption, inflow, and outflow. Food enters a given region either through production or inflow, and it leaves through consumption or outflow. Despite the lack of high-resolution inflow/outflow data, we seek to study the network properties of flows between grid cells, by inferring the in/out flows of each grid cell from available production and consumption data. In order to assign a flow strength to each grid cell, we introduce two simplified paradigms for determining in and out strength: *net flow*, which assigns flow strength based on the difference between production and consumption, and *maximum flow*, which assigns in strength as consumption and out strength as production. We then use these in and out flows to construct network models using two simplified edge assignment paradigms: the random assignment *configuration model*, and the minimum distance *linear programming model*. We present a notion of energy conservation to identify global production and consumption patterns, give approximate bounds on the distance between food production and final consumption, and observe some asymmetries in the food system network topology.

It is important to note that, while our approach provides an idealized energetic perspective on the food system, there are a range of socioeconomic factors which impact the nature of the food system, whose effects are not accounted for by our methods. We ignore differences in energy, value, price, demand, flavor, agro-ecological climate zones, and other factors between commodity types (both variations of individual food items and between different items that are aggregated together). We also treat all demand for consumption the same between regions, regardless of cultural relationships with food and local infrastructure for food trade. These assumptions allow us to develop simple models for food energy flow, however ensure that any of our idealized models cannot capture many important characteristics of the true food system.

Chapter 2

Methods

Here we construct four *food flow networks*, which are weighted directed graphs derived from food production and consumption flows. The starting nodes are grid cells that produce food, and the terminating nodes are grid cells that consume food[1]. This is a sparse network, and we assume some scale free properties based on prior work [Konar et al., 2018]. We utilize these facts to construct a global network of food flows between grid cells.

We first transform available raster data to into consistent units, before constructing idealized networks based on various limiting assumptions. We choose to convert all food flows into units of metabolic energy, in order to account for the stark difference between energy contents of different foods. For example, rice and milk, which account for 18.54% and 7.85% respectively of calories in the global food supply (table 5.1), differ in energy content by a factor of 6 (tables 5.6 and 5.5). We then apply the constraint that all food energy on earth must be conserved, in that an equal energy of food produced must be accounted for as loss, consumption, waste, or other uses[2]. That is, food can enter a given region either through production or inflow, and it can leave either through consump-

[1] Many nodes are both starting and terminating nodes.

[2] The food system does not conserve energy in the traditional physics sense. There is energy loss during transportation, energy gain from fertilizers, etc. However, the total metabolizable energy of all food produced is either consumed, lost, used, or wasted. In this sense the system is closed and ”food energy” is conserved.

tion (including non-food use and waste/loss) or outflow, and all food energy must be accounted for. We infer the net and maximum energy flows between nodes in the food system network. For all networks, the number of trading destinations for in each grid cell is proportional to a power law of the total flow of that grid cell. This means that as the number of trading destinations grows exponentially, so too does the flow strength out of the grid cell, which can be visualized as a linear plot on log-log axes. This relationship gives us a sequences of in-degrees an out-degrees for each grid cell. Finally, we construct networks given these degree sequences, by connecting edges between nodes and assign weights accordingly. With these basic principles we enable the study of conceptually simple but complete, idealized production-consumption flow networks.

Many decisions need to be made in order to approximate missing data. At each decision, we either refer to existing literature to proxy and assess the validity of our results, or we build the network with two different limiting assumptions. This will allow us to minimize arbitrary model parameters as we assess the space of possible 1 degree grid food system network models.

2.1 Raw Data Processing

We begin by populating a 1 degree resolution global grid with the following available data variables, all imported from publicly available data sets[3]:

- Crop Production Mass/Year
- Livestock Counts
- Human Population

We also collect the following *country-level* data, again obtained from publicly available data sets:

- Food Import Mass/Year

[3]Sources given in table 1.1

- Food Export Mass/Year
- Crop Production Mass/Year
- Livestock Production Mass/Year
- Food Waste, Loss, Feed, Non-food Use Mass/Year
- Average Per Capita Consumption Calories/Year

In order to obtain the desired grid cell variables, we often employ dasymetric mapping[4] to distribute nationwide data to grid cells proportionally to existing grid cell data. We then convert crop production mass to consumable energy production (Watts), and we convert population to consumption energy (Watts) based on national annual calorie consumption averages. An outline of this data processing step is displayed in figure 2.1. Once we have obtained food consumption energy and food production energy values for every grid cell, we consider the implications for grid cell energy inflow and outflow, and infer idealized network models for food energy flow.

[4]Dasymetric mapping is a method of distributing coarse data (in this case, national data) based on underlying characteristics (in this case, grid data), in order to enhance spatial precision.

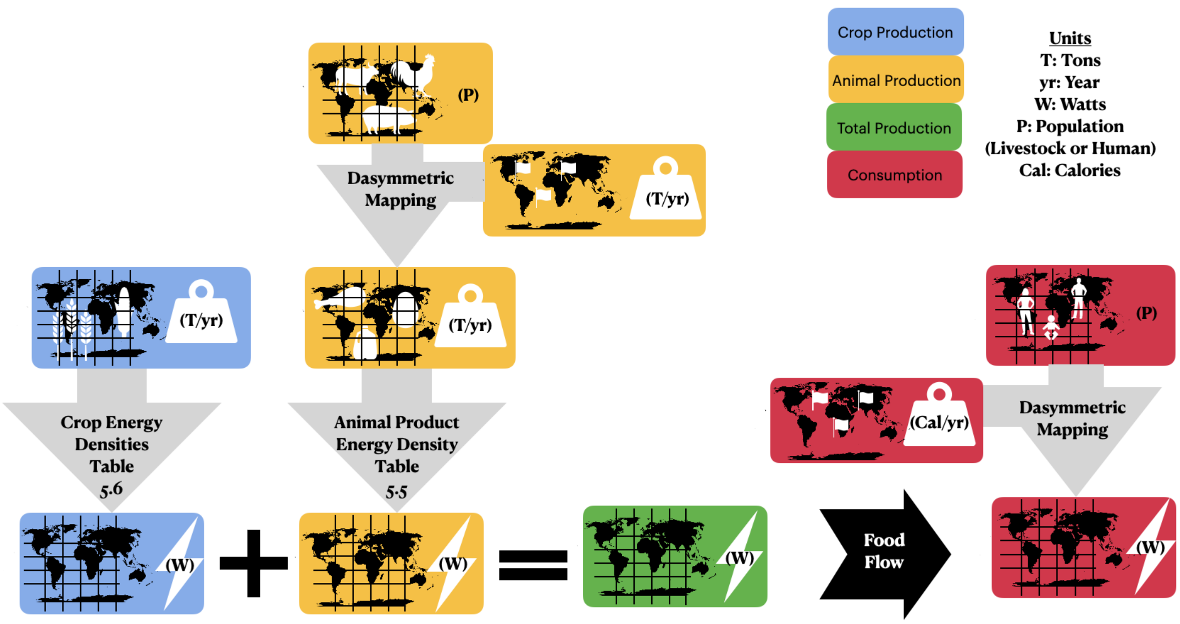


Figure 2.1: Data Processing Flow Chart

We reconcile three different data pipelines: crop production, animal product production and food supply, in order to assess global production and consumption at the grid level.

2.2 Food Energy Conservation

We now formalize the notion that, even though not all food produced is eaten, we can at least be sure that it all goes somewhere. According to the FAO food balance sheet "New Food Balances - Description of Utilization Variables" [Food and Agriculture Organization of the United Nations, 2020b], there is an existing idea analagous to our notion of food conservation, which is accounted for in their suggested methods for data collectors. This concept is referred to as Supply Utilization Accounts, or SUA. Their suggested approach accounts for the mass and stocks in the country being equal to the production plus the imports minus the exports plus an error term. Several problems have been pointed out with the SUA methodology, mainly due to a lack of available consistent data [Kabat, 2023]. With this in mind, we formulate a more parsimonious yet less specific concept of food energy conservation. We specifically account for non-food[5] consumption using processing, feed and loss data, and collect other SUA variables such as Other Uses, Feed, Seed, Waste, Processing, Stocks, into a single scale factor used to match production energy with consumption energy. This choice is motivated by figure 2.1, and allows us to compare the food balance sheet available food supply with the production, imports, and exports.

Hence, we assume the following "Food Energy Conservation" equations to be true[6].

$$\sum I_i = \sum E_i \tag{2.1}$$

$$\sum P_i = \sum C_i \tag{2.2}$$

$$I_i + P_i = E_i + C_i \tag{2.3}$$

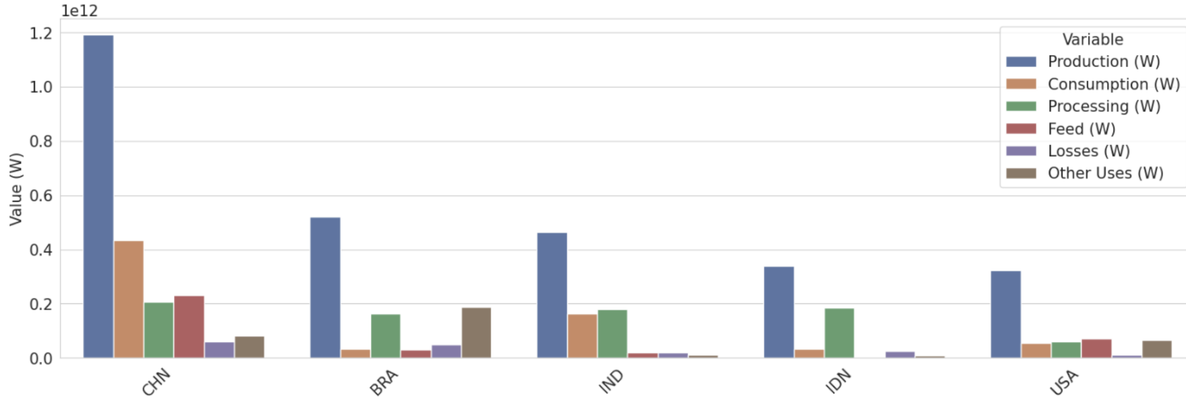
Equation 2.1 follows from the notion that the system is closed; all food imported by a node (country or grid cell), must have been exported by another node. Equation 2.2 follows from the notion that the accumulation of food in the system at each time step is negligible, hence the total production is equal to the total consumption. Equation 2.3 follows from the notion that food can enter a node only through production or imports, and it can leave a node only through consumption[7] or exports.

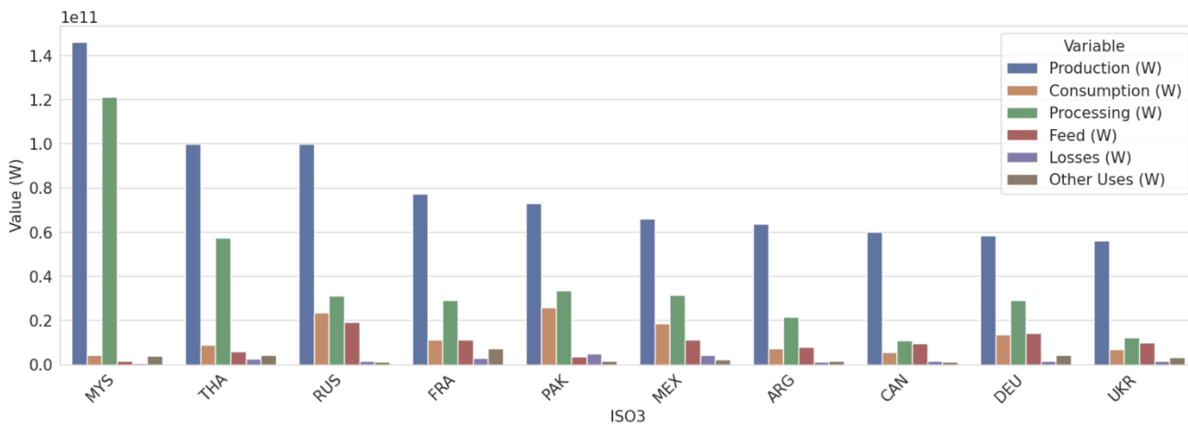
Equations 2.1, 2.2, and 2.3 describe the conservation of energy of any closed food system network at any scale, whether international, inter-grid, or any other food system network, so long as food only enters the system through P_i and only leaves through C_i, and trade (I_i and E_i) is closed between nodes in the system. While there will be heterogeneities among locations (for example some locations are primarily producers while others are primarily consumers), all food energy in a closed food system must be accounted for.

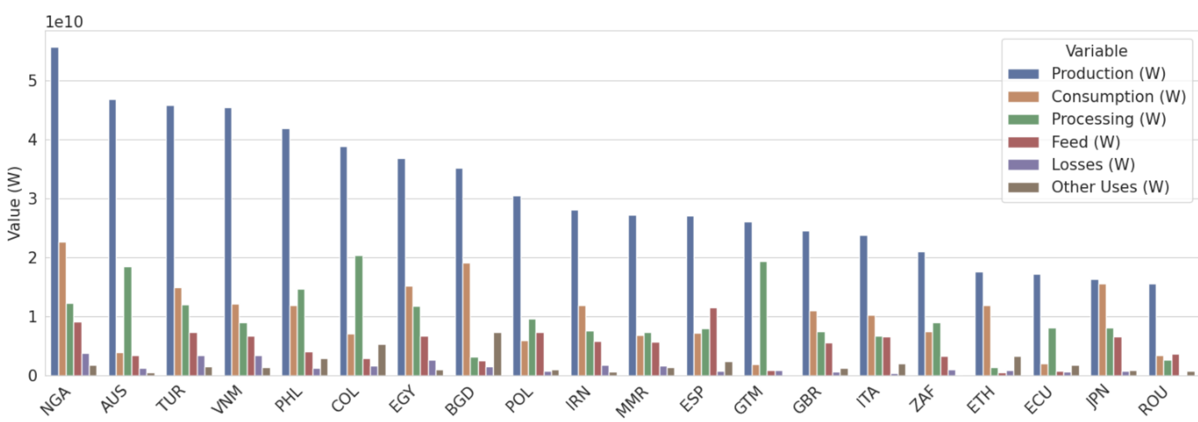
[5] *Food and non-food consumption* includes both human consumption and other non-food uses including feed, waste, loss, biofuels, etc.

[6] In the case of country nodes, I_i, P_i, E_i, and C_i denote imports, production, exports, and consumption, respectively. Production does not include loss (processing loss, use as feed or biofuels, etc.) as defined in table 1.1, and consumption does include waste.

[7] Here *consumption* refers to all food metabolized as well as wasted or put to any other end use.


1e12
Value (W)
Variable
Production (W)
Consumption (W)
Processing (W)
Feed (W)
Losses (W)
Other Uses (W)
CHN
BRA
IND
IDN
USA


1e11
Value (W)
Variable
Production (W)
Consumption (W)
Processing (W)
Feed (W)
Losses (W)
Other Uses (W)
MYS
THA
RUS
FRA
PAK
MEX
ARG
CAN
DEU
UKR
ISO3


1e10
Value (W)
Variable
Production (W)
Consumption (W)
Processing (W)
Feed (W)
Losses (W)
Other Uses (W)
NGA
AUS
TUR
VNM
PHL
COL
EGY
BGD
POL
IRN
MMR
ESP
GTM
GBR
ITA
ZAF
ETH
ECU
JPN
ROU

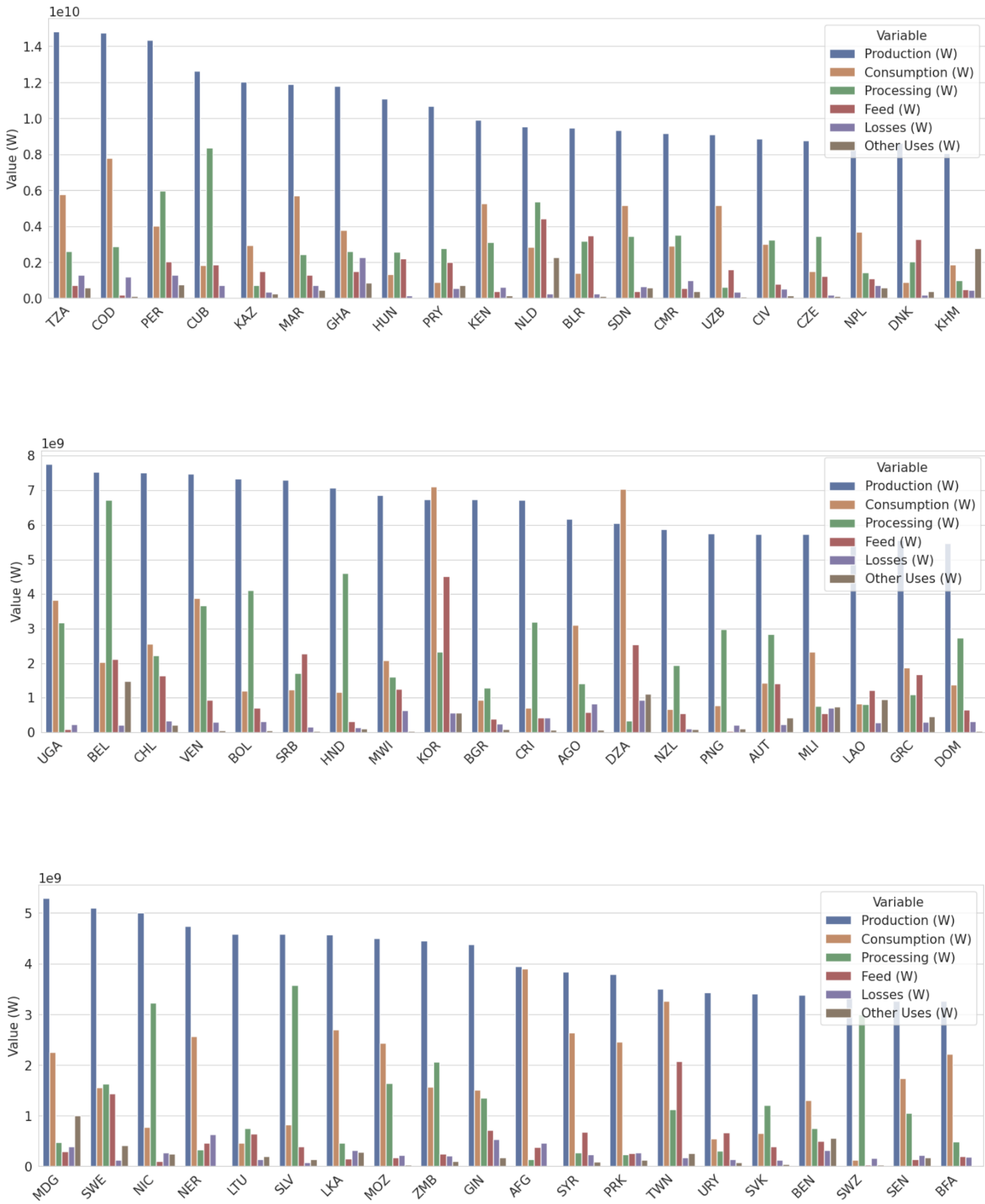


Figure 2.1: National Food Use Types (6)

Aggregate production and food end use energies for 95 countries, calculated from FBS data, and sorted by production. Definitions of the FBS variables are given in the the appendix table 5.2.

2.3 National Data Processing

Before we can study the food flow network at a more uniform, gridded resolution, we must first process national data. The national imports and exports from the FAO food balance sheet define a coarse food flow network, where the nodes are countries (section 1.4.1). The national food consumption and production data provides a window into the inputs and outputs of the global grid food system. We formulate a food energy flow network between nations, before dasymetrically mapping national data to a uniform grid, and comparing the predictions from various network models to the national trade data provided in the FAO Detailed Trade Matrix [Food and Agriculture Organization of the United Nations, 2020a].

We get a shape file labeling each country with its ISO3 code and population estimates from the World Bank database [WorldBank, 2020]. We obtain national average production, consumption, import, and export quantities from the 2020 FAO Food Balance Sheets [Food and Agriculture Organization of the United Nations, 2020b], which capture the total amount of all major food products in each country, accounting for production, consumption, waste, loss, feed, other end uses, and more types of food use. We also use the detailed trade matrix to obtain all food imports and exports between countries [Food and Agriculture Organization of the United Nations, 2020a]. The food items mentioned in the FBS and there correspondence with crop and livestock production are listed in tables 5.3 and 5.4. For a sense of the total quantity of each food globally and how it compares as a percentage of the total food supply, see table 5.1. The foods included in this study (in black in table 5.1) account for about 94% of all consumption in the food balance sheets.

In order to obtain gross energy and mass values, we eliminate all items from the food balance sheet that are not included in tables 5.3 and 5.4. This leaves us with only food products and not their categories as well[8], in order to avoid double counting. Similarly, we eliminate data associated with the groups of countries to not double count based on

[8] For example, rather than including the cereal crops category, we keep only rice, wheat, maize, barley, etc.

overlapping regions. We group each value in the table by Area (country name), Item (food product), and element (ie exports, production, etc.), and sum the groups.

We estimate that the total global food supply, including all foods in table 5.1, is approximtely $1.00 * 10^{16}$ kcal/year or 1.33 TW. The crops and animal products which we have data gridded for (black in table 5.1) include $9.40 * 10^{15}$ kcal per year, or 1.25 TW of food available. We remove some items (marked in red in table 5.1) which are comprised mostly of animal products that we could not easily dasymetrically map onto existing livestock data, or beverages that were difficult to account for from crop production data. About 2% of remaining food energy comes form beer and other alcoholic beverages, and 3% is from excluded aquatic food products.

2.3.1 Unit Conversion

The Food Balance Sheets give production, imports and exports in units of tons, whereas the calories available for consumption are given in units of calories (kcal). In order to fairly compare different types of food flow quantities[9], we convert all dimensions to both mass and energy.

For each food flow quantity starting in mass units (imports, exports, production), we obtain energy by converting it from tons to kilograms, and then use table 5.6 for crops and table 5.5 for animal products to convert from kilograms to calories (kcal). For each food in these tables, we searched the USDA food central database [Haytowitz et al., 2022] for a reasonable choice of energy density for that food. Notable uncertainty arose from the fact that some foods contain a large amount of non-edible mass, and it is sometimes unclear whether the mass of the crop in reference is solely the edible mass, or if other parts are included. For example, there is a potential discrepancy between the mass of crop production of corn, if it were to include cob, husk and kernel, and the mass of food

[9] *Food flow quantities* refer to imports, exports, production, and consumption. The reader should note that, as described in table 1.1, the consumption figures represent food energy available to consumers (food supply), and the production figures are reduced by subtracting loss and feed usage, which are the primary non-food form of consumption.

production of corn, which only includes the edible kernel. We then convert these annual energy flow figures into SI units Joules/Second (Watts). Conversely we do the same in reverse for consumption data in order to obtain consumption masses[10].

We present the mass fluxes in terms of Tons per year, and the energy fluxes in terms of Joules per second (Watts). 1 Watt is approximately equivalent to 1 Calorie per hour (1 W = $\frac{9}{10.46}$ kcal/hour), however we choose to analyze results in terms of Watts, the SI unit for power, to emphasize the general notion of energy conservation which we apply to food flux energy rates.

2.4 Grid Data Processing

Next we use the GAEZ+ 2015 gridded data set of crop production masses, which are then converted to gridded crop production energies, to populate the grid cells with food production energy. Food energy content from the USDA Food Data Central database [Haytowitz et al., 2022] corresponding to each crop in the GAEZ+ 2015 database [Grogan et al., 2022] was used to convert the food production masses to food production energy.

The food balance sheet production data differs from the GAEZ crop production data in that the former refers to food[11] production and the latter refers to crop[12] production. To address this mismatch, we compare quantities of each crop category with foods that fit into that category in figure 2.2. The food categories from the FAO Food Balance Sheets that are mapped to these crop categories are listed in table 5.4. The GAEZ data includes 8.6 total Gt of crop production, which is significantly lower than the FAO FBS crop production total of 10.8 Gt. We assume that this is attributable to missing crops or under-

[10]When converting from gross energy values to gross mass values, we use a mean energy density for crop consumption energy conversion, and a weighted mean energy density for animal product consumption energy conversion, where the mean is weighted by total energy contribution to the system. This was mostly to account for the skew that milk adds to the data, as it is the animal product with the largest annual mass produced and consumed by nearly an order of magnitude, and yet has the smallest energy density of all the animal products by nearly a large margin (see table 5.5).

[11]For the purposes of this paper, we take *food* to mean any foodstuffs listed in the FAO food balance sheet which edible.

[12]*Crops* are defined as biomass grown that can be converted to human consumable food, as measured by the GAEZ+ 2015 database.

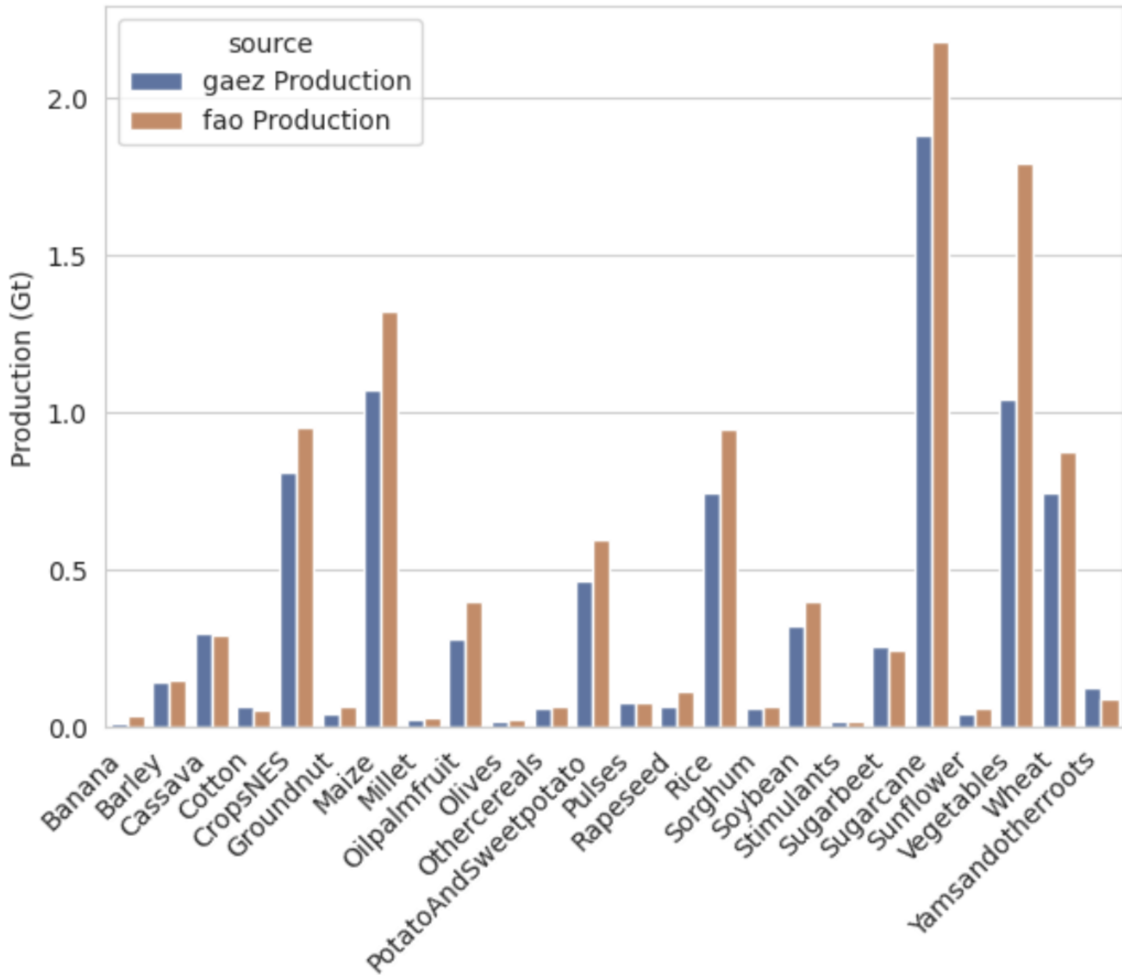


Figure 2.2: Food and Crop Production Comparison

GAEZ+ 2015 production mass shows reasonable consistency with 2015 FAO Food Balance sheet production mass.

estimates of GAEZ, and/or to over-reporting of FAO national production, as well as potential inconsistencies in definitions of foods, crops and edible organic matter.

In order to include energy content from animal products, we add to the gridded crop energy data gridded animal product food energy, which is obtained by dasymetrically mapping animal products from the FAO FBS (Food Balance Sheet) [Food and Agriculture Organization of the United Nations, 2020b] onto gridded livestock from the FAO GLW (Gridded Livestock of the World) database [Gilbert et al., 2018] using conversions given in tables 5.3 and 5.5 (which are analogous to tables 5.4 and 5.6 which we previously used for crops). We end up with 3.64 TW of crop production energy and 0.32 TW of animal product production energy.

Finally, we obtain 1 degree gridded consumption data by converting a human population raster map into a food energy consumption map by dasymetrically mapping national food supply energy [Food and Agriculture Organization of the United Nations, 2020b] onto gridded population data [WorldPop, 2020]. Note again that this ”consumption” figure refers to food available for consumption, which includes both consumption and waste, where waste refers not only to food waste but also overeating[13]. The FAO estimates that the global average daily energy supply is 2963 calories [FAO, 2022], which is approximately equivalent to a rate of 144W per person, or approximately 1.15 TW (8 billion * 144) for the global population. This population-based order of magnitude estimate of consumption agrees reasonably well with the calculated supply of 1.25 TW.

We aggregated all gridded food production and consumption across commodities in this idealized energetic model of the food system network, before inferring bilateral trade. There would be more precision if bilateral trade was estimated separately for desegregated food production and consumption, however this aggregate approach allows us to keep the idealized model more parsimonious, and also doesn't rely on local diet variations, for which data is not readily available. We discuss this desegregate approach as an avenue for future work further in section 3.5.

2.4.1 Distributing Losses, Waste, and Non-Food Uses

Now we come to the question of how to distribute losses from the national FAO data to each grid cell. Non-food use measurements are generally reported in terms of percentage loss [Alexander et al., 2017] [Gustavsson et al., 2011], which makes sense considering the nature of loss tends to scale with the amount of available commodity to be lost. This has the natural implication that, even if production side and consumption side losses were

[13]We use the term *overeating* in the sense of Alexander et. al. [Alexander et al., 2017], which accounted that 1.66 Gt of food is consumed annually whereas only 1.49 Gt are needed to sustain the population. This 170 MT deficit is ”overeating”.

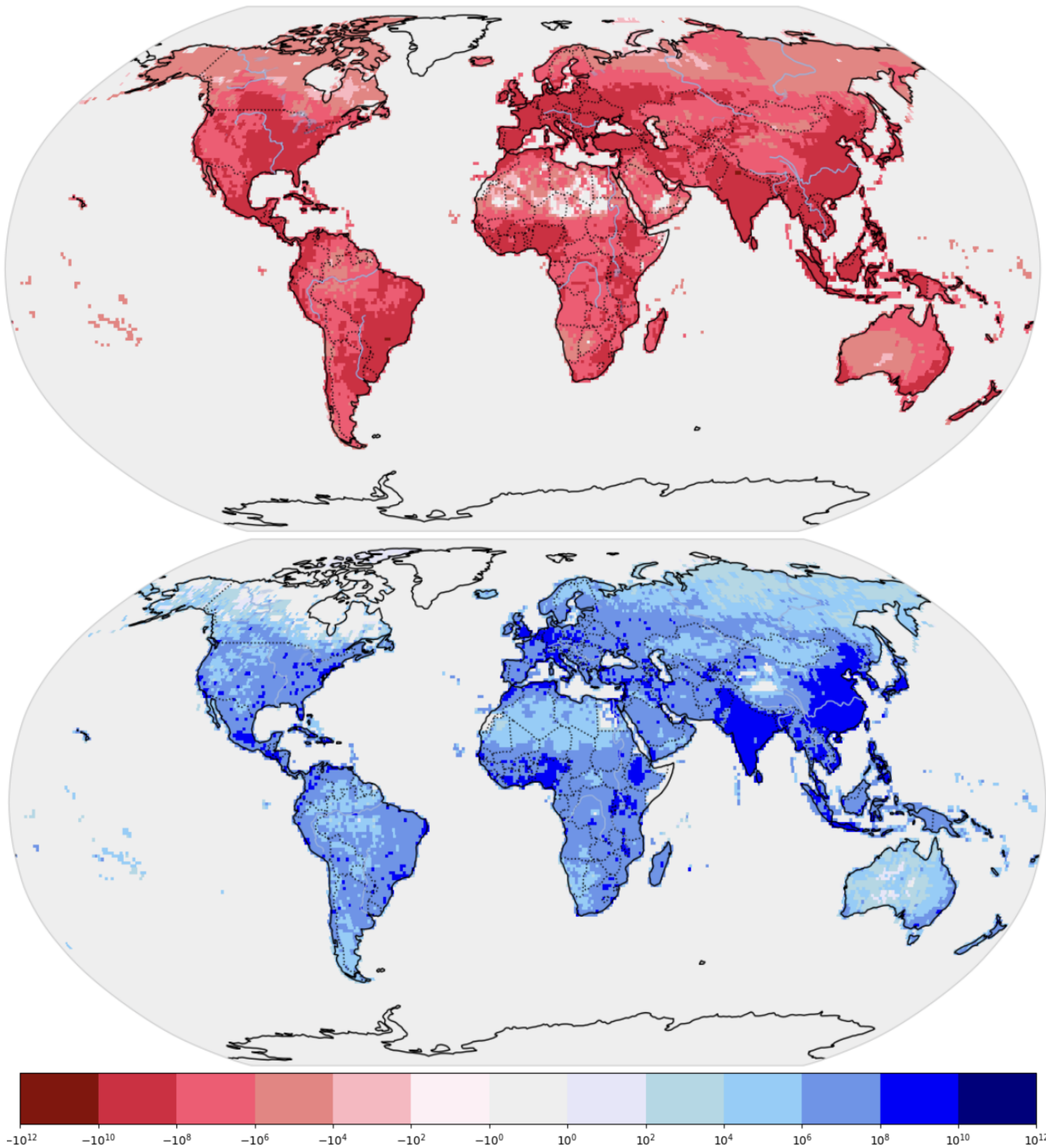


Figure 2.3: Production and Consumption (W)

Estimated food production energy (red) and consumption energy (blue). Grid cells are shaded to represent the rate of energy of production or consumption in Watts. Production includes net production after loss, feed, processing, and other non-feed uses are accounted for. Consumption refers to the food supply available for consumption, including food that is eaten as well as wasted.

the same, the majority of loss occurs before food reaches consumer markets, considering that food always flows through a unidirectional supply chain from initial production to end-use consumption, and loss is proportional to the amount of food present at any given level.

Some estimates of food loss during production come to around 30% [Gustavsson et al., 2011], others coming to 24 % [Kummu et al., 2012]. More recent studies have argued that these works are problematic based on a wide range of assumed and estimated loss rages [Alexander et al., 2017], and so estimated substantial loss at various stages of the supply chain including crops harvested, processed commodities, and animal products, all at significant percentages. Still though, consumption end waste of food available to consumers is still significantly lower, at approximately 10% of food available to consumption, which is yet a smaller fraction of originally produced food [Alexander et al., 2017].

The sum totals for all food energy produced and consumed are given in table 2.1. We start by calculating the total production side loss by comparing total production to total food available to the consumer (both with dimensions of mass and energy). We then reduce the production of all grid cells proportionally to the global percent loss to obtain a rough approximation of the amount of food produced in that grid cell that will eventually be available to a consumer.

We verify that these gross production and loss figures are in line with the literature by comparing to values from Table 1 of Alexander et. al. [Alexander et al., 2017]. Indeed 8.6 Gt of crop production falls close to the estimate of 9.76 Gt wet mass. Animal product mass is higher at 2.3 Gt, compared 1.14 Gt wet mass, for reasons that are not completely clear. Gross loss percentage falls in with existing estimates that between percent 19.2–31.9% food production biomass is consumed by humans [Alexander et al., 2017].

For our model we focus only on the food balance sheet categories of consumption, processing, loss, and feed as an effort to maximize parsimony, considering that the vast

	Production	Loss	Percent Loss	Food Available
Crop Mass (Gt)	8.6	6.4	74.4%	2.3
Animal Product Mass (Gt)	2.3	1.2	49.4%	1.1
Total Mass (Gt)	10.9	7.5	69.1%	3.4
Crop (TW) Energy	3.63	2.65	73.0%	0.98
Animal Product Energy (TW)	0.32	0.05	14.6%	0.27
Total (TW) Energy	3.94	2.69	68.4%	1.25

Table 2.1: Consumption Data Accounting

Sum total of grid cell production and consumption data

majority of national food uses can be accounted for with these categories. Stock variation accounts for less than 1% of production [Alexander et al., 2017].

2.4.2 Net Flow and Maximum Flow

With production and consumption data estimated in every grid cell, we define the node strength in two different ways, each representing extremes in terms of subsistence agriculture and the distribution of trade patterns. The "net flow" model for node strength equates the in-strength to the deficit between production and consumption within a grid cell, and the out-strength is the surplus of production and consumption. This provides a lower bound for how much food energy must flow in or out of each grid cell to sustain the population, or trade the surplus respectively. The strength of each grid cell node is mapped in figure 2.4.

We call the second model for node strength "maximum flow," which takes the opposite extreme of possible node strength. In this model, the entire quantity of consumed

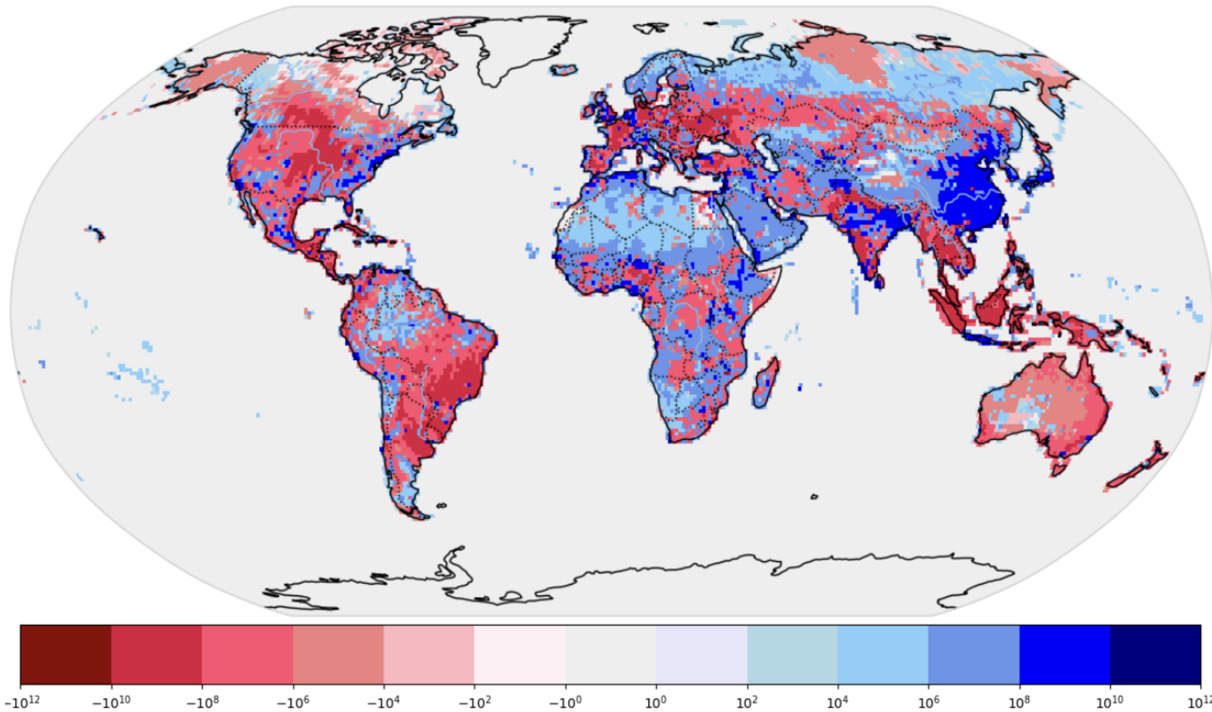


Figure 2.4: Net Energy Flow (W)

Estimated Net Flow (food production minus consumption energy) per 1 degree grid cell

food is assumed to be imported to each grid cell, and the entire quantity of produced food is exported (no local consumption).

From the food energy conservation equations, particularly equation 2.3, along with our production and consumption data, the net flow of food energy is determined for every grid cell.

Equation 2.3 allows us to rewrite the food conservation equations as follows with some algebra:

$$\begin{gathered} I_{ij} - O_{ij} = C_{ij} - P_{ij} \\ \sum_{i,j} (I_{ij} - O_{ij}) = 0 \end{gathered} \tag{2.4}$$

We define netflow as the difference between production and consumption.

$$
\begin{aligned}
I_{ij} &= \max(P_{ij} - C_{ij}, 0) \\
O_{ij} &= \max(C_{ij} - P_{ij}, 0)
\end{aligned}
\tag{2.5}
$$

Net flow is the first type of flow we study, as it is in a sense the minimum possible flow that could enter or leave a grid cell. It has some useful properties, including that any two nodes have at most one direction of food flow, as well as the fact that all local consumption demand is addressed from local production. Any deficit is imported and any surplus is exported.

We define maxflow as in equation 2.6:

$$
\begin{aligned}
I_{ij} &= C_{ij} \\
O_{ij} &= P_{ij}
\end{aligned}
\tag{2.6}
$$

Maximum flow is also quite simple to calculate, and is a reasonable upper bound for food flows. It is worth noting that it is possible that the food flow in reality is greater than max flow for any given grid cell, but that would imply that a grid cell is importing and re-exporting the same food on a scale that outweighs their consumption.

We believe that these two flow models, net flow and maximum flow, are ideal choices for grid inflow and outflow strength proxies. While reported flow data between grid cells does not exist, net flow provides a clear lower bound, and maximum flow provides a compelling yet not strict upper bound on how much food a grid cell reasonably would transport.

2.4.3 From Flow Strength to Node Degree

Next our goal is to approximate the number of trading partners[14] in each grid cell using flow strengths, so that we can feasibly construct a network between grid cell nodes. It has been shown that, for food flow networks, whether at the village, sub-national, or international scale, there is a power law relationship between the degree distribution (number of trading partners) and mass strength distribution (strength of trade) [Konar et al., 2018]. We assume this to be true for our grid network as well, as the "universal signature of human behavior" and "similar governing mechanisms" that drive this power-law relationship between food flows should hold for food flow networks at any resolution, and use mass flow strength to calculate node degree. As the sub-national scale is closest to our 1-degree resolution, we take the power law coefficient and exponent for our initial parameters. We perform sensitivity analysis with respect to these parameters to demonstrate that choice of value does not affect our results (section 3.4. Reference [Konar et al., 2018] presents a super-linear power law relationship between strength as a function of degree. This implies a sub-linear relationship in the other direction, with strength on the x axis and degree on the y, so we convert the exponent to its reciprocal and the coefficient to its respective value accordingly as our starting point is node strength,

$$d = as^b \tag{2.7}$$

where d = node degree and s = node strength. We let a = .0071 and b = 1.53 as initial parameters. These values are the same parameters[15] of the power law fit between food flows between Commodity Flow Survey Areas in the United States, which is the nearest scale to a 1 degree resolution of the three scales studied [Konar et al., 2018]. We explore sensitivity to this choice of parameters in section 3.4. Having defined node strengths

[14] We use the term trading partners here in a loose sense. The edges we will construct are not truly direct trade relationships, because we are not including intermediate steps in the trade network, just a start-to-finish approximation based on production and consumption data.

[15] We scaled them to match our mass flow units of tons per year instead of billion kg per year, and inverted the power law to represent degree as a function of strength rather than the other way around.

using these two methods, we infer from the relationship between mass flow strength and degree the number of in and out edges for each node in the network. The food mass production data comes directly from our raw data sources, whereas the food consumption mass is determined by dividing consumption energy by the average calories per kilogram of each food in table 5.6, which is sourced from the USDA Food Data Central database [Haytowitz et al., 2022].

2.5 Network Food Flow Models

At this stage, there are approximately 13,000 grid cell nodes containing some amount of production and/or consumption of human food energy. For each node, there are two different models for flow strength, and corresponding degree sequences for each flow model. We next connect the nodes in both the maximum flow and net flow models with two different methods for edge assignment: the configuration model and the linear optimization edge assignment algorithm.

2.5.1 The Random Connection Model

The configuration model randomly connects nodes to preserve the degree sequence. This model has the dual benefits of parsimony and computational simplicity. It can be thought of as bearing similarity to a free market trade system, which is agnostic to international trade barriers and long distance trade concerns. The maps of the net-flow configuration and max-flow configuration models are shown in figure 3.5. We visualize the top 1000 edges by flow strength.

2.5.2 The Linear Programming Model

Our other method of edge connection is only slightly more complicated than random assignment. We attempt to find the optimal configuration of edges to minimize both distance between initial production and final consumption, and international trade. The linear optimization edge assignment algorithm finds the optimal edges given a degree sequence constraint that minimizes the total cost between each pair of connected nodes. The cost function is defined as the sum of the geodesic distance between the nodes on earth plus a constant that indicates whether they were in the same country or not (representing international trade barriers). The linear programming algorithm minimizes the total edge cost by optimizing the objective function 2.8 with respect to the constraints 2.9 provided by the degree sequences.

$$f(e_{ij}) = \sum_{ij} c_{ij} e_{ij} \tag{2.8}$$

$$i_{out} = \sum_{j} e_{ij} \qquad\qquad j_{in} = \sum_{i} e_{ij} \tag{2.9}$$

$$c_{ij} = d(i,j) + b\delta^{i.ctry}_{j.ctry} \tag{2.10}$$

Where f is the objective function, c is the cost function, e is the food flow from node i to node j, i_{out} and j_{in} are the out and in strenghts of nodes i an j, respectively, b is the local country incentive constant (we use a value of 0 by default), and $\delta^{i.ctry}_{j.ctry}$ is the kronecker delta with respect to the countries containing nodes i and j.

While the linear optimization algorithm is theoretically possible with 10,000 nodes, we chose to enforce a node strength threshold of 100MW to reduce the complexity of the network[16]. This has the effect of decreasing the total degree of the nodes in the network, which allows us to use the CPLEX solver [Nickel et al., 2022] to find the optimal solution. This threshold also has the effect of ignoring nodes which are only small contributors to the food system, allowing us to focus on the few thousand nodes which contribute the

[16]This threshold was obtained through trial and error. Our Linux system has approximately 200GB of free CPU RAM memory and this is insufficient for large mixed integer optimization problems. Even with just a 500MW threshold, the linear programming problem becomes infeasible. While this threshold is relatively small in terms of the total energy dropped from the system, it does drops a significant number of nodes (see table 2.2)

largest in and out strengths. While this threshold does impact the statistical distributions of strength and degree which were used to construct the network, the qualitative shape of the distributions remain consistent.

2.5.3 Edge Weight Assignment

The final step to build the complete network model is to distribute the node flow strengths to the edges, making our graph weighted. In order to do this, we employ another linear programming algorithm. We simply maximize the total weight of all edges, subject to the constraint that for every node, the sum of its out edge strengths must be less than or equal to its out strength, and the same for in coming nodes and in strength. For each network some amount of flow strength is lost in this process, as it is not guaranteed that the edge connections will preserve the total strength of nodes. They can only be forced to preserve the total degree of nodes, which is a coarse approximation of each node's strength. Net flow configuration loses 12.6% of its flow strength, maximum flow configuration loses 8.3%, net flow linear programming loses 12.6%, and maximum flow linear programming loses 22.4%. We are optimistic that this loss of data can be substantially reduced in future iterations of this research (described in section 3.5, but we make due with some information loss in order to take the network study to its conclusion.

Having assigned edge weights, we are left with 4 directed weighted food flow networks. While these network models aren't necessarily indicative of the real world food flow network, they have the power to demonstrate what the food system *would* look like if grid cells erred more to the side of net flow subsistence or max flow globalization, and if the trade system erred more to the side or random global trade or preferential local flows. In reality, the actual food network is likely somewhere in between these four extreme cases.

	Net Flow	**Maximum Flow**
# nodes initial	16690	16690
Total Initial Flow Strength (GW)	658	1247
# nodes of degree 0 (removed)	5670	5026
# nodes below energy threshold (removed)	8448	7390
Total strength removed (GW)	173	179
# stubs removed to normalize degrees	151	4922
% stubs removed to normalize degrees	0.27	3.41
# nodes final	2572	4274
# edges final	27779	69791

Table 2.2: Flow Model Creation Statistics

Some statistics describing the state of each network model at each step of methods and how much information is lost at each step.

Chapter 3

Results and Discussion

In this chapter we present the results of the thesis in three sections. First we explore the implications of accounting for spatially resolved production and consumption energy of food within a country. Next we examine the network of international food energy trade between countries, rather than grid cells. We validate the accuracy of our transformation from FAO food balance sheet masses [Food and Agriculture Organization of the United Nations, 2020b] to energy by asserting that the food energy conservation laws 2.1, 2.2, and 2.3 do in fact hold true for international food energy flows.

Finally, we presents the results of the network analysis. The configuration models lead to food transport of upwards of 10x more distance than the linear programming models. We interpret these models not as reflections on reality, but rather as constraints on what range of network topologies might be possible in the real world, without changes to production and consumption energy patterns. We find that, as expected, none of the simplifying assumptions individually explains the nature of international trade flows reported by the FAO [Food and Agriculture Organization of the United Nations, 2020a] However, by comparison we attempt to glean some information about the underlying network.

Specifically, we compare 3 pairs of opposing types of food flow: production and consumption, maximum flow and net flow, imports and exports. There is an observed asymmetry between each of them. This manifests in a diversity of indicators. To name a few,

- Consumption is more concentrated among nodes than production.
- Maximum flow is more coupled to exports than imports.
- Imports are more coupled with consumption and exports are more coupled with production.

We posit that these asymmetrical relationships are mechanistically connected. We find that these overarching asymmetries seems to be robust to preliminary perturbations of the power law model parameters, hinting that the properties of our models might indeed reflect the asymmetrical topology of the underlying real global food system network.

3.1 Sub-National Production and Consumption

In figure 3.1 we plot each grid cell within a selection of large countries, as well as every grid cell in the food system. The R^2 value for the strength of correlation between production and consumption energy of grid cells within each country is mapped in figure 3.2. We observe that production and consumption tend to be correlated with each other, meaning that grid cells with large populations (and hence more consumption) tend to also produce a lot of food. Interestingly, the strength of this correlation varies between countries and regions. We suggest that the strength of this relationship could be used as a metric for the co-location of production and consumption. While a correlation between consumption and production does not necessarily mean that the produced food is being consumed locally, the inverse is true: if consumption is not co-located with production, it requires food trade and transport.

While there are some exceptions including Greece, Switzerland, and England, many countries in Europe have very little correlation between production and consumption within their grid cells, and therefore most of the food (both produced and consumed) in their grid cells cannot be consumed locally. While non-local consumption does not necessarily mean international import and export, as it could just mean grid cells within a country supplying each other, it seems that many European countries might to be more integrated with of the globalized food system, as evidenced also by their under-prediction even for the maximum flow configuration model (see figures 3.10 and 3.9).

China and India both have of huge populations and huge rates of production. For these reasons, they are in a class of their own with respect to the food system network. It is intriguing then, that their grid cells tend to be more correlated by production and consumption than most other countries. We speculate that this is because subsistence farming plays a greater roll in these nations.

3.2 Country Network Food Conservation

Just as in the network of food producing and consuming grid cells, the actual food energy consumed, produced, imported, and exported by each *country* (per unit time) must also satisfy the food conservation laws (equations 2.1, 2.2, 2.3). In other words, all food that is produced must be lost, wasted, or consumed through food or non-food usage, regardless of how we partition the food system into jurisdictions.

In figure 3.3 we show the relationship between production plus imports versus consumption plus exports. In essence, this is the relationship between the ways food can enter a country, and the ways food can leave the country. While these two axes are highly correlated, there is certainly some discrepancy, which can be accounted for by including the other forms of non-food consumption, including processing, loss, feed, and other uses. These are elements in the food balance sheet for each country. If we include these forms of non-food consumption, we see a much stronger correlation, as shown in figure 3.4.

This result is reassuring for two reasons. First of all, it indicates that the food balance sheets are robust and consistent with the SUA methodology [Kabat, 2023]; the food balance sheets are constructed to ensure a balance of food *masses*, preserving conservation of food mass in the international trade system, as described in section 2.2. Second of all,

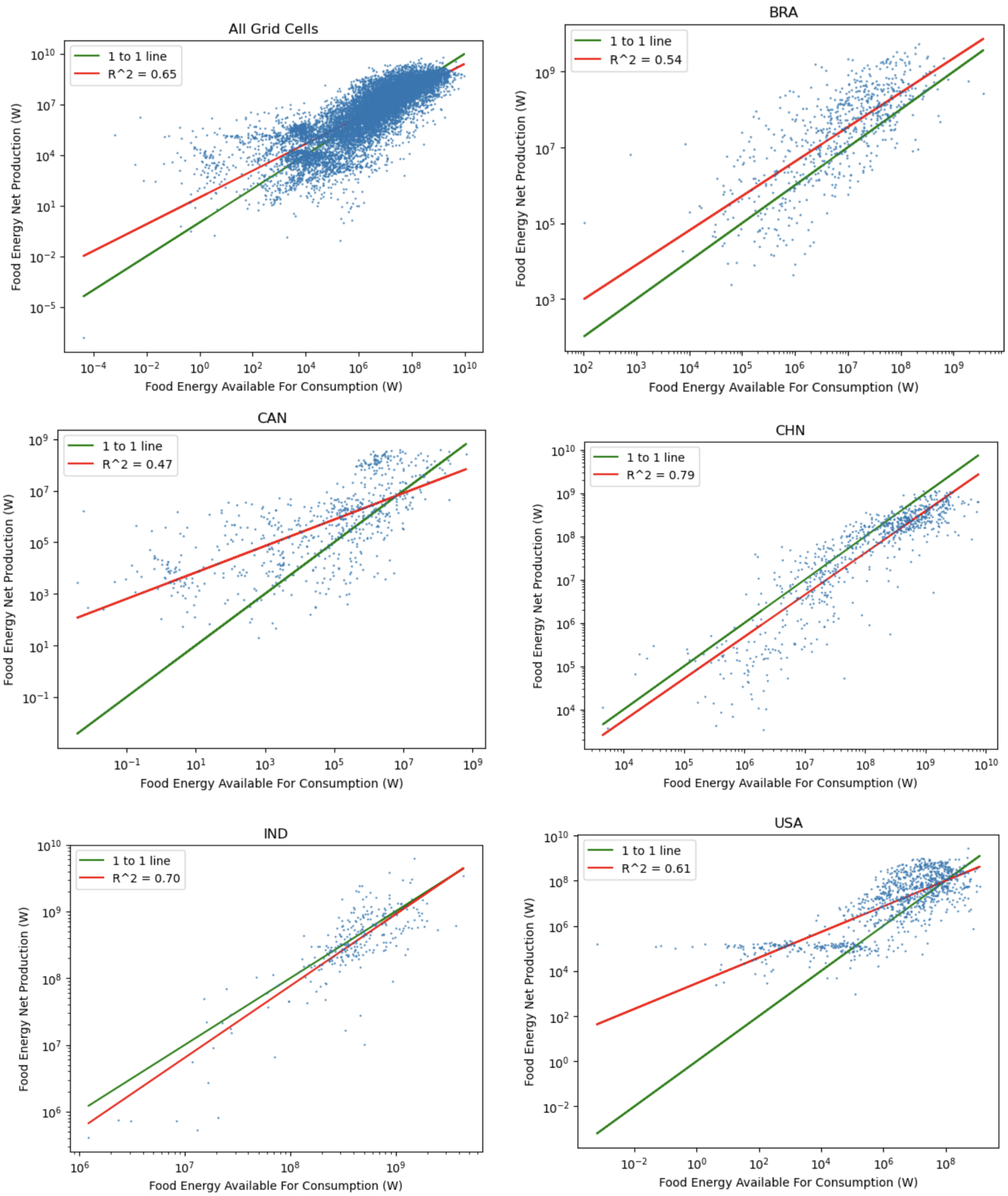


Figure 3.1: Grid Cell Production vs Consumption

Observed correlation between production and consumption in grid cells in different regions and in the world.

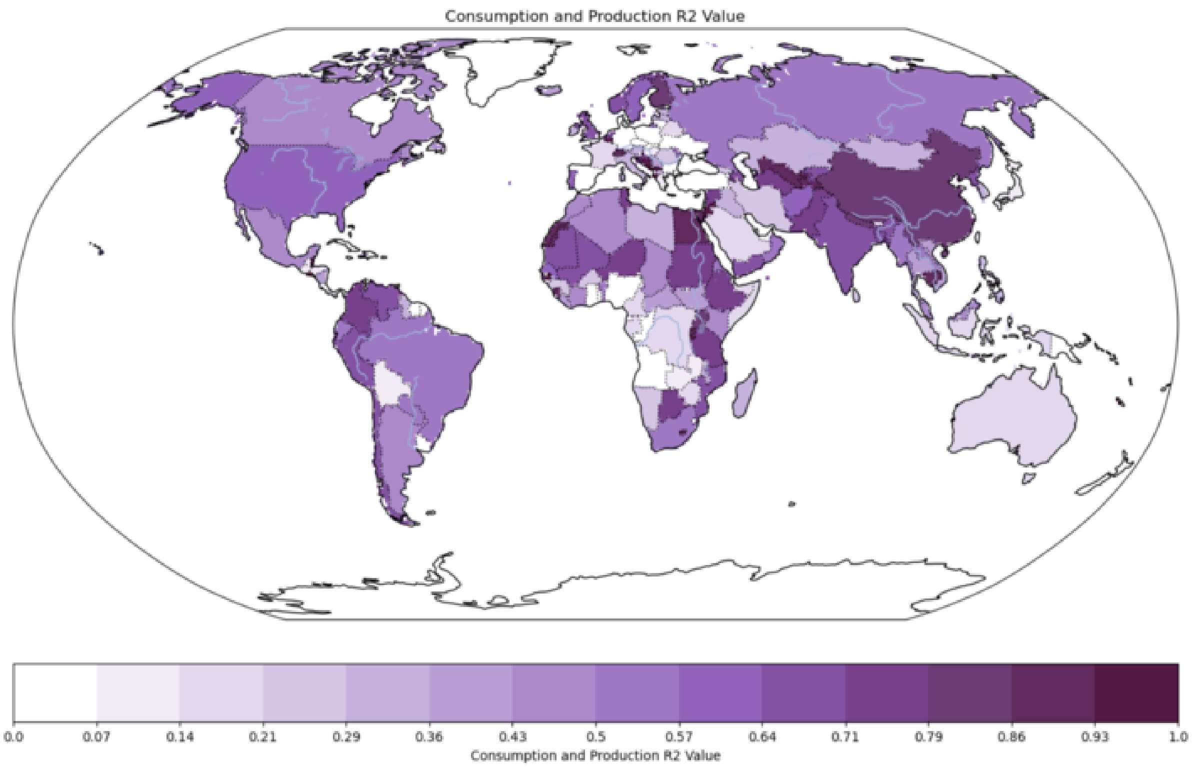


Figure 3.2: R^2 Production Consumption Correlation Map

This is a map of the strength of correlation between production and consumption within country grid cells.

we observe that *energy* inflow is also nearly equal to *energy* outflow. That is, import plus production energy is nearly identical to export plus consumption energy, even though it was mass that was forced to be equivalent by SUA [Kabat, 2023]. The preservation of this result despite transformation from mass to energy indicates that our unit transformation using the USDA food central caloric dense energy densities did not disturb the food conservation of the international food trade network. This indicates that we chose reasonable energy density values to reflect the average energy density of most foods in the food system.

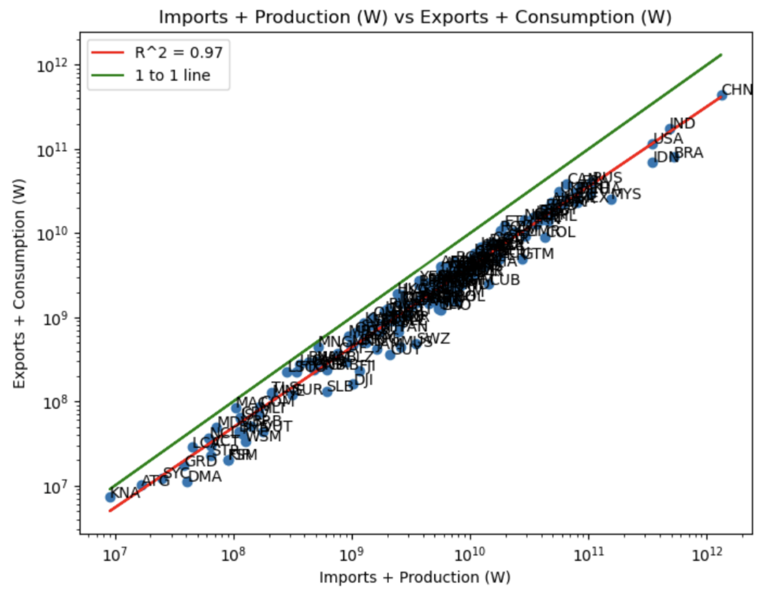


Figure 3.3: International Food Energy Conservation (1)

For each country, we compare the influx from imports and production to the outflux from exports and consumption. Consumption includes only data from the FAO food balance sheet column "food supply".

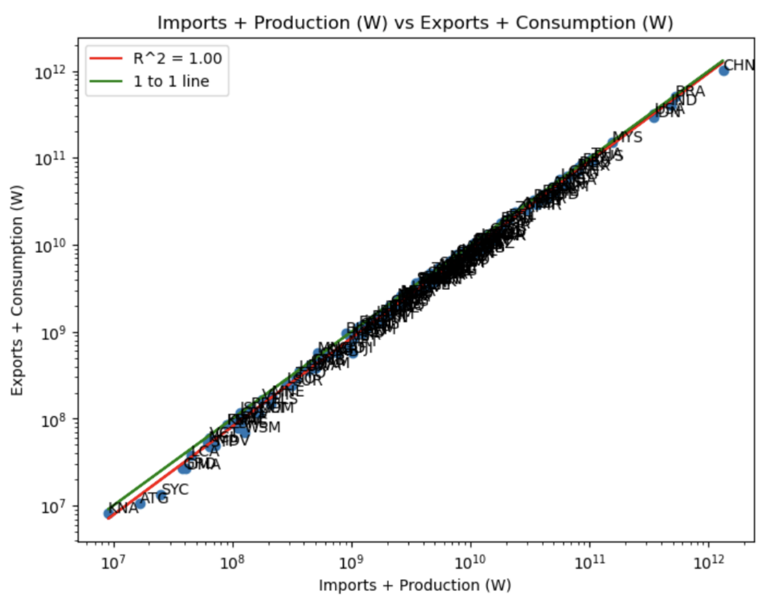


Figure 3.4: International Food Energy Conservation (2)

For each country, we compare the influx from imports and production to the outflux from exports and consumption. Consumption includes FAO food balance sheet columns "food supply", "processing", "loss", "feed", "other uses".

3.3 Food Flow Network Analysis

Now let us compare and contrast the four network models, mapped in figure 3.5. An overview of the final network statistics, after dropping nodes below the 100 MW threshold, constructing edges, and distributing weights, is given in table 3.1. As is to be expected, the maximum flow networks have greater total strength than the net flow networks. Before network normalization and weight reduction, the total strength of the net flow model is 657 GW, whereas the total strength of the maximum flow model is 1247 GW (table 2.2). We interpret the net flow value of 657 GW as a lower bound on possible food energy flows to sustain global consumption, ie the minimum required food flow rate to balance production and consumption nodes, and the 1247 GW as a conjectured upper bound. By comparison, we estimate the total energy of food trade between nations to be approximately 552 GW[1]. It has been estimated that international and intra-national (domestic) food trade are of about the same order of magnitude [Konar et al., 2018]. We can therefore make the rough order of magnitude conjecture that the actual total flow between grid cells is on the order of 1 TW.

A comparison of the two edge construction models reveals that the random networks have a significantly greater total path length. This is reasonable, as nodes were not paired to minimize distance as they were in the linear programming shortest path model, but rather randomly connected even across the entire globe. Theoretically the path length would be $\frac{1}{4}$ the circumference of the earth if production and consumption were evenly distributed (as is the the average distance between two random points on a sphere), or approximately 10 thousand kilometers. In fact it is not far off, as shown in table 3.1, at 8 to 9 thousand kilometers.

While it might be tempting to interpret the total path length of the net flow linear programming and maximum flow configuration models as minimum and maximum global food trade path lengths respectively, this energetic account gives no consideration to nu-

[1]We obtain this value of 552 GW by summing all imports and exports energies we converted from food balance sheet import and export masses and dividing by 2.

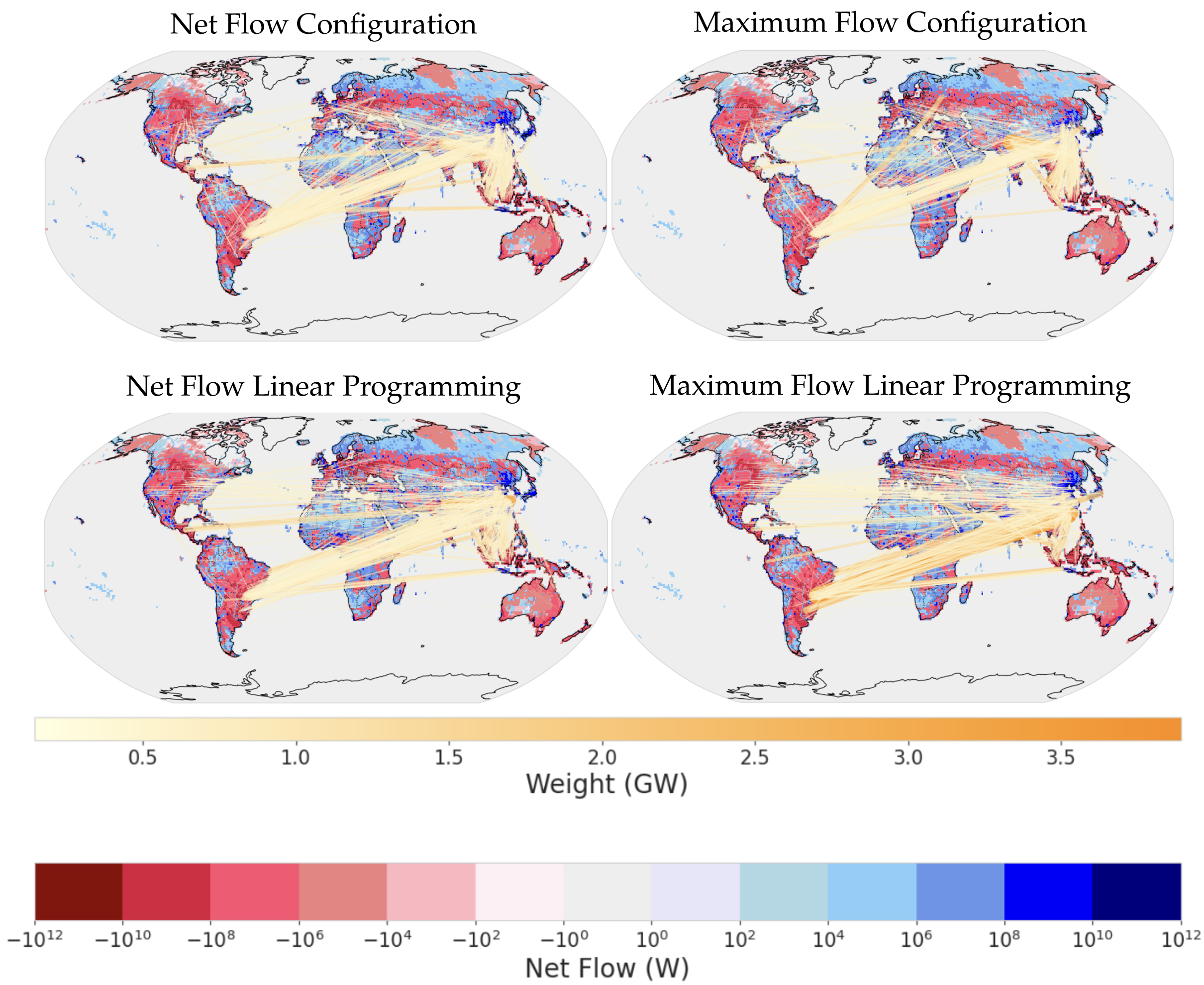


Figure 3.5: Network Model Maps

The colors in each grid cell represent its net food energy (production - consumption). The orange lines are network edges representing modeled food energy flow from initial production to final consumption. These are maps of net flow (left) maximum flow (right) configuration (top) and linear programming (bottom) models.

tritionional requirements. While the random configuration model, by the law of large numbers, likely distributes diets relatively evenly throughout the world, the linear programming model is designed to distribute food energy nearby to satisfy energy demand, regardless of a population's requirement of dietary balance. This is why the linear programming model is at best a lower bound on the distance food must travel between grid cells to sustain the global demand for food. It is likely significantly lower than what is possible in reality, as there is a need for dietary balance on the consumption side as well as the existence of crop-mono-cultures on the production side which inhibit a diverse local food supply.

The similarity of the four network shapes mapped in figure 3.5 despite the fact that they were obtained with very different assumptions, suggests that indeed the production and consumption in each grid cell is a useful constraint on the topology of flows between grid cells. We note that one of the defining features of all four networks seems to be heavy trade between Brazil and Asia, at least among the strongest node pair fluxes. While China and Brazil are indeed two of the three countries that trade the most food with each other, along with the USA [Food and Agriculture Organization of the United Nations, 2020a], we suspect sugarcane production might be an additional bias favoring Brazil-China trade in the models. Recall from figure 2.2 that sugarcane is the single most produced crop by mass, and also note that Brazil accounts for 45% of global sugar exports [OECD, 2018][2]. This might account for some of the nodes with the strongest flow strength between these two countries.

Next, we compare the results with national food trade reports. For each country, we sum the strength of all edges flowing in and out of that country to obtain the predicted in and out strengths. These values are compared with the reported import and export quantities from the UN food balance sheets. All data is converted to Watts.

[2]China is also the single greatest sugar cane importer [Food and Agriculture Organization of the United Nations, 2020a]. Still though, the magnitude of this relationship is likely an inaccuracy in the model, since the model is agnostic about crop type, and just treats all energy equivalently. In reality sugarcane and other food flows are probably dispersed more widely.

	Net Flow Random Configuration	Maximum Flow Random Configuration	Net Flow Linear Programming	Maximum Flow Linear Programming
Number of Nodes	2572	4274	2572	4274
Number of Edges	27294	69402	27779	69791
Total Flow Strength (GW)	556	603	556	603
Network Density[3]	0.41%	0.38%	0.42%	0.38%
Total Path Length (million km)	229	626	50	45
Average Path Length (thousand km)	8.39	9.02	1.80	0.65

Table 3.1: Idealized Network Model Statistics

The comparison between the configuration models and the linear programming models reveals striking similarity. While close examination of the location of some countries on the plot reveals slight differences, generally the flow strength (either net flow or maximum flow) has far more implication on the network topology than the manner in which we connect edges, which suggests that the degree sequence constraint is more relevant, at least to understanding the import and export patterns of nations, than the distance between grid cells. We are hopeful on this basis that, if the actual in/out flow or degrees of grid cells, the accuracy of model import-export predictions would be greatly improved.

3.3.1 Asymmetry in Food Flows

Now let us compare and contrast the nature of food energy imports and exports. They have characteristic differences, as demonstrated by the differences between figures 3.7 and 3.6. While neither the net flow nor the maximum flow model produce stellar pre-

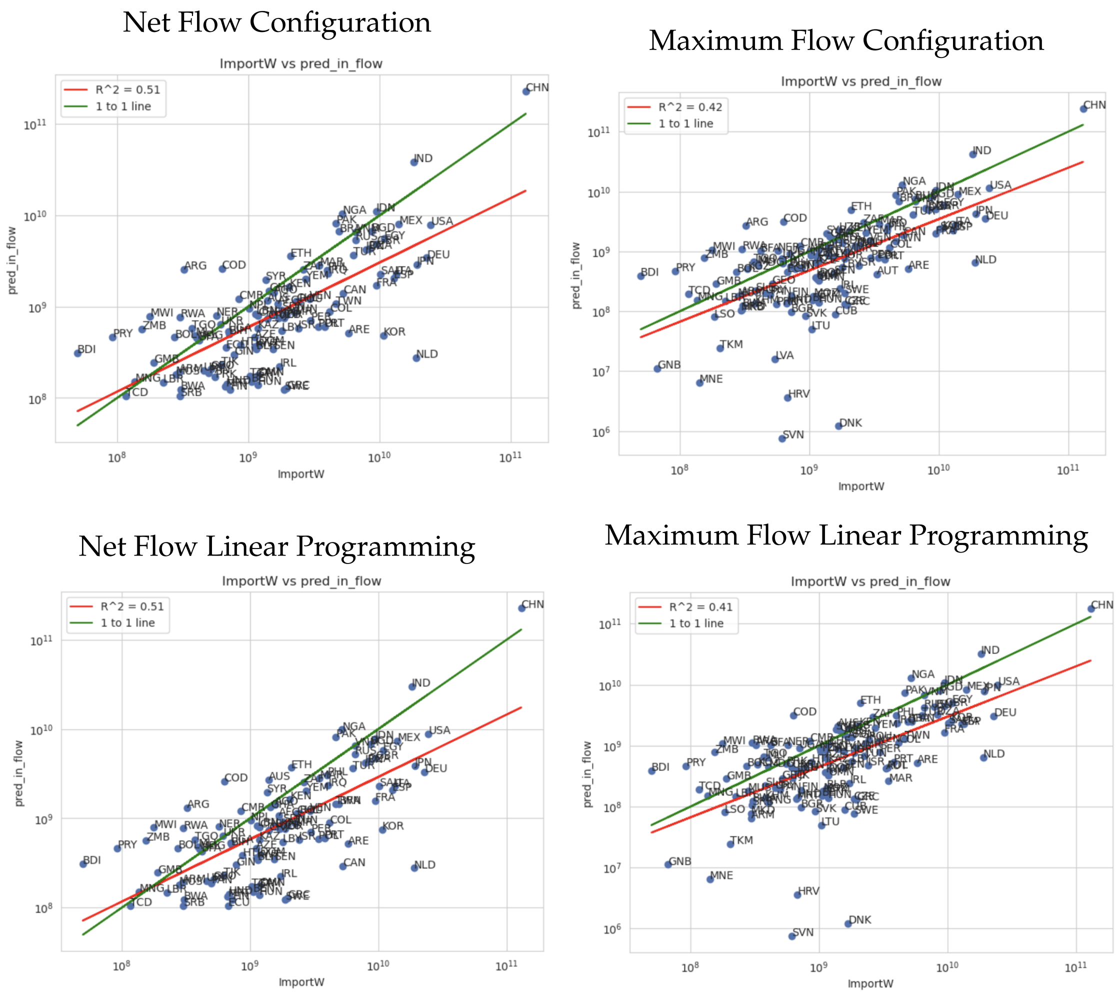


Figure 3.6: Import Predictions

Predicted imports by the net flow (left) and maximum flow (right) configuration (top) and linear programming (bottom) models.

dictions of either inputs or exports, the import predictions are more correlated with net flow than maximum flow. The reverse is true for exports, which are more correlated with maximum flow than maximum flow.

Both net flow and maximum flow are symmetric with respect to direction, in the sense that consumption and production are treated equally in their calculations. This means that if we were to call consumption ”production” and production ”consumption”, then our inflow and outflow quantities of each grid cell would be reversed, and our degree sequences would be swapped, which would generate the same networks that we generated,

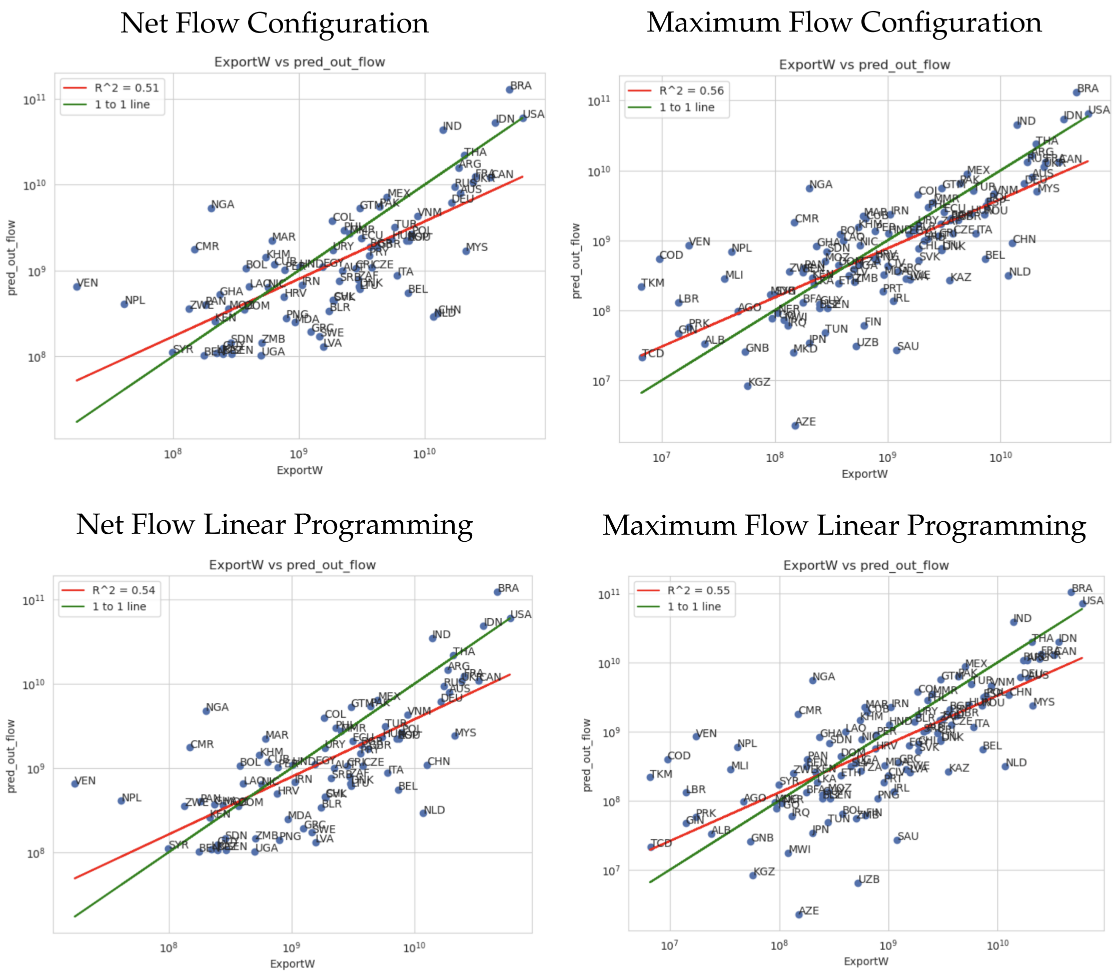


Figure 3.7: Export Predictions

Predicted exports by the net flow (left), maximum flow (right), configuration (top), and linear programming (bottom) models.

except with the direction of the arrows reversed. Therefore, the observed asymmetry in the models must be derived from the production and consumption input data. In fact, we can observe this asymmetry in figure 3.8, in which production is distributed among more nodes with a lower rate of production energy, while consumption is concentrated among fewer nodes with a higher rate of consumption energy. This is a reflection of the physical reality that cities are supplied by large areas of farmland.

Exports are more correlated to maximum flow. We posit that this is a result of the uneven distribution of production and consumption among grid cells, as shown in figure

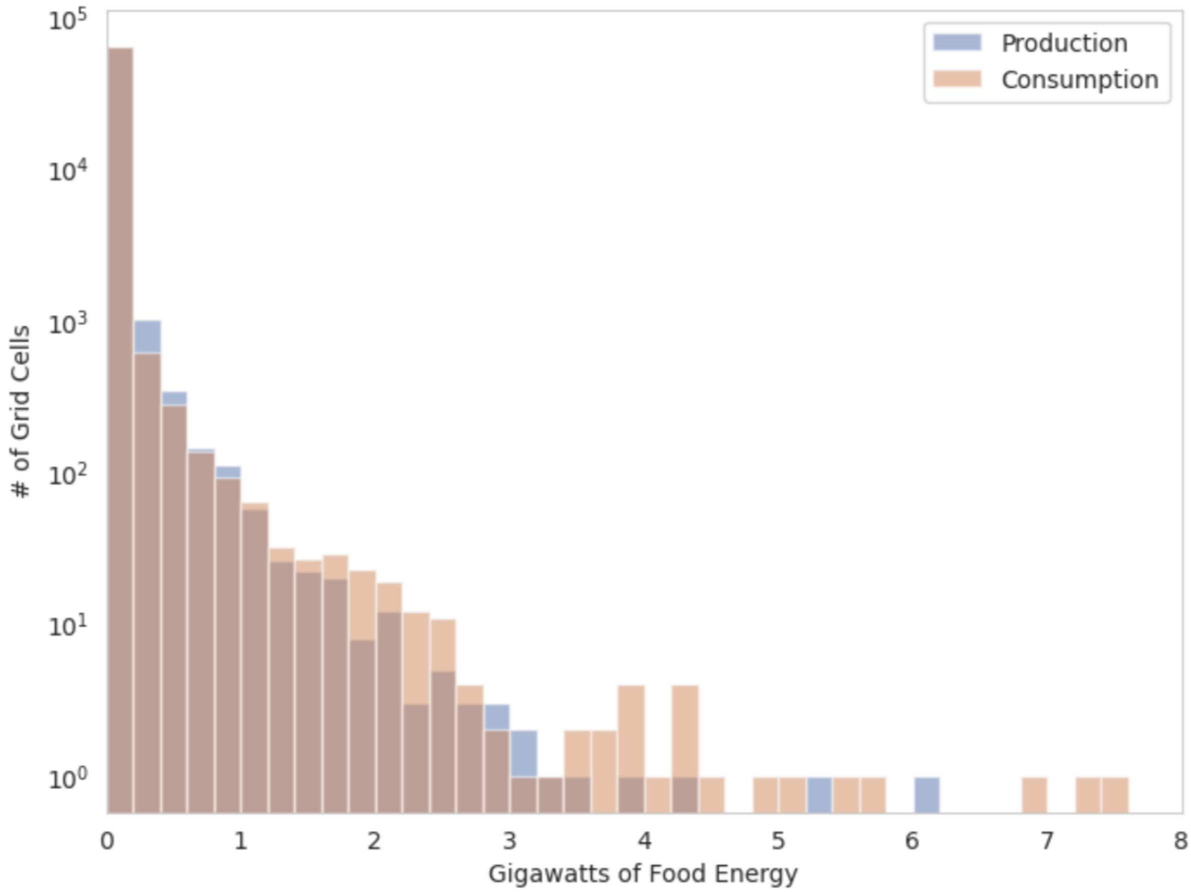


Figure 3.8: Histogram of Production and Consumption in Grid Cells

Due to the logarithmic scale, which is necessary to visualize this exponential distribution, the difference in height of the bins that contain the most grid cells are misleadingly small. Among grid cells in the bins at a rate of less than 1 GW, hundreds more of them are producing.

3.8. Since production is more dispersed among grid cells and consumption is more concentrated, more grid cells in the network must export nearly all of the food they produce in order to sustain the population. By contrast, since consumption is more concentrated among fewer grid cells (cities), the number of consuming grid cells that must import nearly all of their food (and hence follow the maximum flow model) is proportionally lower.

Additionally, it makes sense that production and consumption asymmetries lead directly to asymmetries in flow strength of nodes from a theoretical perspective, since equations 2.5 and 2.6 are functions of production and consumption. We posit that the character

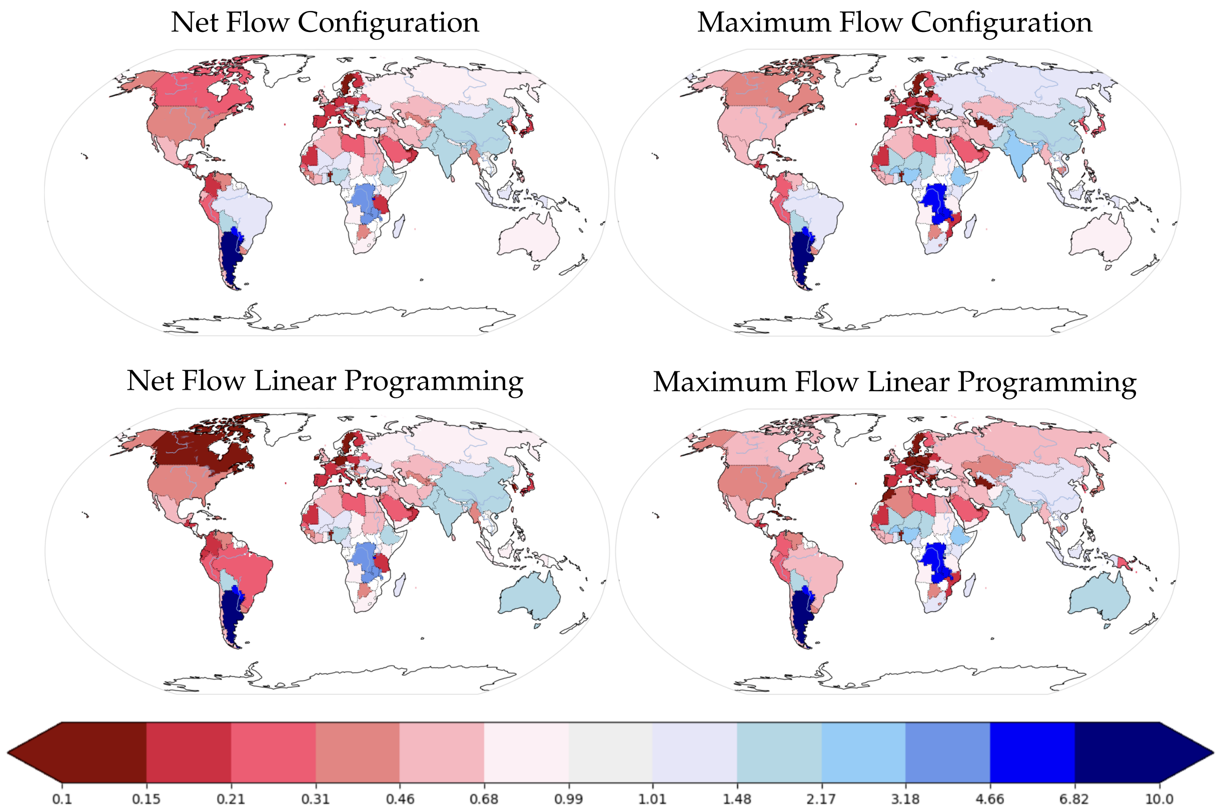


Figure 3.9: Network Model Import Accuracy Maps

Ratio of model predicted imports I_{pred} to each country's FAO reported imports I_{fao} for the each model. $R = I_{pred}/I_{fao}$.

of this asymmetry likely holds for the real food system network, since it holds for both the net flow and the maximum flow models, and is derived directly from production and consumption data that necessarily impacts real flow strength.

While the magnitudes of national exports reported by the FAO seem to be predicted to about the same degree of correlation as imports, maximum flow does a much better job of predicted exports. That maximum flow predicts exports well might be an indicator that the actual outflow strength of each grid cell is closer to maximum flow (ie the total production) than to net flow (ie the production minus consumption). However further research is necessary to establish if this is the case, as we are not observing grid cell flows directly, but rather only national imports and exports.

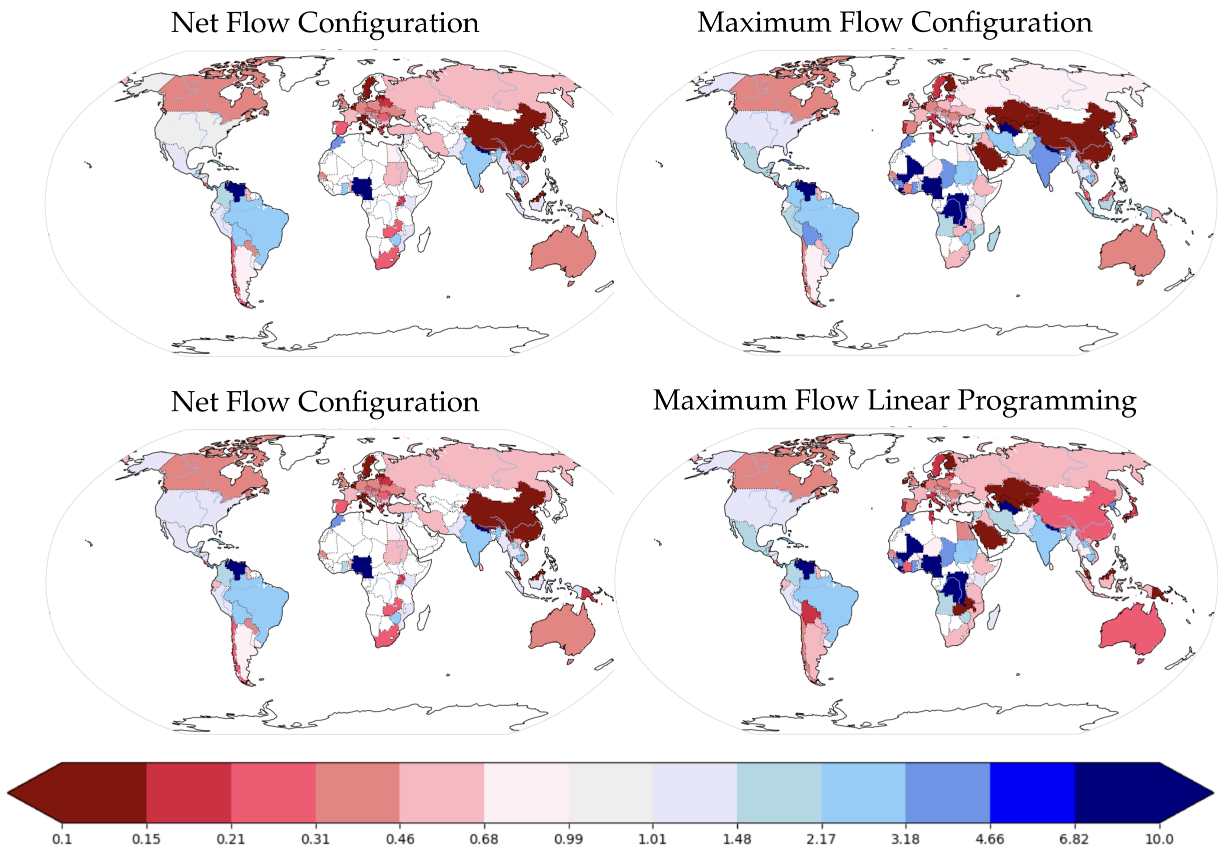


Figure 3.10: Network Model Export Accuracy Maps

Ratio of model predicted exports E_{pred} to each country's FAO reported exports E_{fao} for the each model. $R = E_{pred}/E_{fao}$.

Notice that, in figures 3.6 and 3.7, the slope of the line of best fit on log-log axes is less than the 1-1 line, implying that the predicted data is a sublinear power law of the actual data. For countries with lower imports and exports, the model tends to over-predict their values. Similarly, imports/exports of many of the countries with large reported trade fluxes are under predicted by our energetic flow model.

While the linear programming model searches for an optimal solution, the configuration model randomly connects outgoing food flow stubs with incoming food flow stubs to make network flow edges. A reasonable hypothesis would be that the configuration model would over-predict country in/out flows, as there is no incentive to trade locally in this model. We compare the actual imports and exports of each country with the pre-

	Imports			Exports		
	1.3	**1.53**	**1.7**	**1.3**	**1.53**	**1.7**
0.005	.57	.58	.57	.54	.54	.50
0.0072	.58	.57	.54	.55	.58	.55
0.01	.56	.58	.58	.56	.49	.56

Table 3.2: Net Flow Configuration Sensitivity

R^2 correlation of log(imports) vs log(predicted imports) as shown in figure 3.6, and of log(exports) vs log(predicted exports) as shown in figure 3.7

dicted imports and exports according to the net flow and maximum flow configurations model in figures 3.6 and 3.7 respectively, and find striking similarity in goodness of fit and overall shape to their configuration model counterpart, by comparing the top lots to the bottom plots.

3.4 Sensitivity Analysis

Here we display R^2 coefficients of deviation for polots analogous to 3.7 and 3.6, for nine combinations of choices of power law coefficients (0.005, 0.0072, 0.01) and power law exponents (1.3, 1.53, 1.7). While the choice of power law exponent and coefficient do seem to have an impact on the nature of the fits, the underlying sublinear power law relationship holds regardless of parameter choice. We suspect that this is a result of the asymmetric distribution of production and consumption. A greater coefficient can be interpreted as allowing more food to flow through each edge. A greater exponent can be interpreted as a greater increase of number of trading partners for each unit of strength increased in a node.

In tables 3.2 and 3.3, we present the R^2 values for the fits shown in figures 3.7 and 3.6 for the net flow configuration model and the maximum flow configuration model separately. These regressions, while each subtly different, retain the sublinear power law fit, and there are not significant gains or reductions in accuracy across different power law parameters.

	Imports			Exports		
	1.3	**1.53**	**1.7**	**1.3**	**1.53**	**1.7**
0.005	.52	.52	.52	.63	.62	.61
0.0072	.52	.52	.52	.63	.63	.63
0.01	.52	.52	.52	.63	.63	.63

Table 3.3: Maximum Flow Configuration Sensitivity

R^2 correlation of log(imports) vs log(predicted imports) as shown in figure 3.6, and of log(exports) vs log(predicted exports) as shown in figure 3.7

These perturbations of parameters indicate that our results are not dependant on the specific values in the power law relationship between strength and degree, which was instrumental in converting node strength into a degree sequence in order to construct the network models (section 2.4.3). Along with figure 3.4, which shows that our energy conversions maintained international food conservation reasonably well, we gain some confidence in our results.

3.5 Limitations and Future Work

This thesis contributes to a rich academic literature that seeks to study and model the nature of the food system. By establishing a proof of concept of an energetic production-consumption food network, we prepare future work to further constrain food flow within and between grid cell, both in terms of precision but also magnitude and direction.

While this is a step towards more granular and systematic approach to the global food system, there remains much to be done. Incorporation of fishery data and approximation of the average cereal crop components of alcoholic beverages would allow for complete coverage of essentially all food in the FAO food balance sheets. With more computing resources, these analyses can be run on a finer grid with more nodes included. Not only would this provide more versatile data, but would eventually allow researchers to make stronger assumptions about the nature of each grid cell as a node in the network, since they will on average tend to be more homogeneous the smaller they get (meaning there won't be grid cells that contain both hundreds of hectares of cropland and large population centers in the same grid cell). Furthermore, this research can not only be used to better understand the underlying mechanisms of global food flows, but also as an indicator for the effect of perturbations to global consumption and production on the rest of the food system.

Our model might be also made more comprehensive by including feed as inputs to the animal product consumption gridded data. While we account for feed as loss to the food system, in reality feed is a flow from one type of food energy (crop) to another (livestock). We will use feed energy conversion efficiencies [Shepon et al., 2016, Alexander et al., 2017] to constrain and assess this essential component of the food system.

Furthermore, the models would be much improved without the node energy threshold. This was necessary in order to obtain results for most computationally intensive maximum flow linear programming model, which was otherwise infeasible without some amount of coarsening. However this meant that the threshold needed to be applied to each model to enable a fair comparison. Since we are looking at only nodes with greater than 100 MW of total flow strength, our models are a representation of the characteristic topology of flows between the largest contributors to the food system, which is not all that different in magnitude between the two flow models. However in order to include the entire system within this framework, more powerful computational resources are required.

Another possible improvement to the methodology might be to combine edge construction and edge weight distribution in one step. There are novel statistical network methodologies which allow for generating weighted edges given directed strength and degree sequences [Vallarano et al., 2021], which might reduce the loss of strength we observed in section 2.5.3.

We took an aggregate approach in this project, as it enabled us to use the simple dasymetric equation of food supply energy per capita to determine consumption. However, in principle a disaggregated food system network could be constructed and analyzed with a similar methodology, if we were able to first assign more detailed dietary information to each grid cell. An ideal next step would be to assign to each grid cell a fraction of caloric consumption from each food item in the food balance sheet. At a fine enough resolution, this might enable policy makers to establish quotas or disincentives for certain types of unsustainable, high carbon footprint consumption such as beef and dairy products, and incentivize more sustainable alternatives.

A feasible first approximation of this could be reasonably attained using our methodology of dasymetric mapping. Rather than simply mapping the gross energy content onto human population as we did in this thesis, we will treat each food in the food balance sheet separately. Just as food energy conservation must be conserved, so too must rice energy, and wheat energy, and milk energy, and every food that is produced and consumed on planet earth (or at least every food that is listed in table 5.1). Using the national food supply quantities would naturally neglect important differences between consumption patterns within a country, however as more global information systems data becomes available, this sub-national heterogeneity can also be addressed through dasymetric mapping of smaller jurisdictional areas.

In addition to providing an entirely new dimension of data by specifying each food separately, this approach will have the added benefit of reducing the linear programming problem. We anticipate that we won't need to use the node energy threshold to reduce the complexity of the edge assignment problem, as the greatest energy we will need to map for an individual food item is less than 20% of the total food supply we mapped here.

Chapter 4

Conclusion

This thesis has taken what appears to be a novel approach to examining the global food system network. Using a combination of dasymmetric mapping and unit conversion we constructed a global grid of $10,000km^2$ grid cells with associated producing and consuming food energy rates.

We identified the conservation of food energy as a defining feature of this network. By accounting for and aggregating nearly all global food energy, we were able to constrain the properties of food fluxes between both countries (trade) and grid cells (food flows).

We also revealed major differences in collocation of production and consumption in different parts of the world.

We used the gridded production and consumption values along with food conservation constraints to generate two simplified and opposing models for how food can flow between grid cells as a function of production and consumption (net and maximum flow). We suggest that they impose reasonable lower and upper bounds on the actual food flow through each node.

Next we created degree sequences from the node strengths and used these to create two simplified and opposing network models (the minimum distance linear programming model and the random configuration model). These models were surprisingly similar, emphasizing the deep structural implications of the food flow network strength and degree sequences on its network topology.

We conclude that these four idealized network models display patterns of asymmetry between the different types of food flow quantities. We speculate that the network properties such as total path length and density can be used as first guess estimations of the actual network, and we hope that this research can help set the stage for future analyses of the global grid food flow network.

Chapter 5

Appendix

Table 5.1: Food Balance Sheet Food Products Food Supply

Red food products are not included. Included food products account for about 94% of all consumption in the food balance sheet.

Food Item	Food Supply (Total kcal)	% of Total Food Supply
Rice and products	1.86×10^9	18.54%
Wheat and products	1.72×10^9	17.14%
Milk - Excluding Butter	7.88×10^8	7.84%
Sugar (Raw Equivalent)	5.75×10^8	5.72%
Pigmeat	5.12×10^8	5.10%
Maize and products	4.44×10^8	4.42%
Vegetables, other	3.07×10^8	3.05%
Eggs	2.76×10^8	2.75%
Soyabean Oil	2.56×10^8	2.55%
Potatoes and products	2.15×10^8	2.14%
Poultry Meat	1.94×10^8	1.93%

Continued on next page

Table 5.1 – Continued from previous page

Crop	Food Supply (kcal)	% of Total Food Supply
Palm Oil	1.84×10^8	1.83%
Cassava and products	1.48×10^8	1.47%
Bovine Meat	1.16×10^8	1.16%
Beer	1.11×10^8	1.11%
Sorghum and products	1.11×10^8	1.11%
Rape and Mustard Oil	1.05×10^8	1.05%
Fruits, other	1.03×10^8	1.03%
Fats, Animals, Raw	1.02×10^8	1.02%
Groundnuts	1.02×10^8	1.02%
Pulses, Other and products	9.93×10^7	0.99%
Sunflowerseed Oil	9.41×10^7	0.94%
Sweet potatoes	8.47×10^7	0.84%
Beverages, Alcoholic	8.33×10^7	0.83%
Butter, Ghee	8.28×10^7	0.82%
Millet and products	7.01×10^7	0.70%
Sweeteners, Other	6.71×10^7	0.67%
Bananas	6.38×10^7	0.63%
Beans	6.18×10^7	0.61%
Freshwater Fish	5.55×10^7	0.55%
Nuts and products	5.04×10^7	0.50%
Soyabeans	4.95×10^7	0.49%
Apples and products	4.29×10^7	0.43%
Groundnut Oil	4.28×10^7	0.43%
Oilcrops Oil, Other	4.18×10^7	0.42%

Continued on next page

Table 5.1 – Continued from previous page

Crop	Food Supply (kcal)	% of Total Food Supply
Cottonseed Oil	4.07×10^7	0.41%
Mutton & Goat Meat	4.01×10^7	0.40%
Onions	3.81×10^7	0.38%
Tomatoes and products	3.76×10^7	0.37%
Yams	3.51×10^7	0.35%
Oranges, Mandarines	3.20×10^7	0.32%
Coconuts - Incl Copra	2.83×10^7	0.28%
Plantains	2.70×10^7	0.27%
Olive Oil	2.69×10^7	0.27%
Offals, Edible	2.45×10^7	0.24%
Cereals, Other	2.43×10^7	0.24%
Sugar non-centrifugal	2.38×10^7	0.24%
Barley and products	2.33×10^7	0.23%
Pelagic Fish	2.29×10^7	0.23%
Grapes and products (excl wine)	2.19×10^7	0.22%
Maize Germ Oil	21421878.01	0.213156
Peas	20593005.29	0.204908
Roots, Other	17944520.97	0.178555
Wine	17664437.95	0.175768
Spices, Other	17394955.13	0.173086
Demersal Fish	17335642.68	0.172496
Coconut Oil	16543001.02	0.164609
Palmkernel Oil	13904032.30	0.138350
Rye and products	13338548.92	0.132724

Continued on next page

Table 5.1 – Continued from previous page

Crop	Food Supply (kcal)	% of Total Food Supply
Cocoa Beans and products	12708468.22	0.126454
Sesame seed	12561463.70	0.124991
Sugar cane	12203537.42	0.121430
Beverages, Fermented	12177530.20	0.121171
Aquatic Products, Other	11969469.52	0.119101
Aquatic Plants	11969469.52	0.119101
Dates	11526745.44	0.114695
Pimento	10878394.96	0.108244
Meat, Other	9332297.92	0.092860
Crustaceans	9332282.41	0.092860
Oats	9127733.56	0.090824
Sesameseed Oil	8858915.58	0.088150
Miscellaneous	7921837.56	0.078825
Marine Fish, Other	7763578.72	0.077250
Ricebran Oil	7596460.78	0.075588
Pineapples and products	6947879.12	0.069134
Cream	6379067.55	0.063474
Honey	5815163.25	0.057863
Molluscs, Other	5543636.92	0.055161
Oilcrops, Other	4854636.11	0.048305
Citrus, Other	4360585.74	0.043389
Cephalopods	3795762.34	0.037769
Coffee and products	3557619.06	0.035400
Olives (including preserved)	3346375.66	0.033298

Continued on next page

Table 5.1 – Continued from previous page

Crop	Food Supply (kcal)	% of Total Food Supply
Lemons, Limes and products	3285501.00	0.032692
Infant food	3116974.86	0.031015
Tea (including mate)	3049735.99	0.030346
Grapefruit and products	2763744.21	0.027500
Rape and Mustardseed	1401774.98	0.013948
Pepper	1171604.82	0.011658
Sunflower seed	1114167.26	0.011086
Aquatic Animals, Others	686379.45	0.006830
Fish, Body Oil	635882.94	0.006327
Cloves	138722.80	0.001380
Fish, Liver Oil	40075.86	0.000399
Sugar beet	36963.21	0.000368
Palm kernels	29915.00	0.000298
Cottonseed	6005.65	0.000060
Meat, Aquatic Mammals	0.00	0.000000

Table 5.2: Verbatim Food Balance Sheet (FBS) Column Definitions [Food and Agriculture Organization of the United Nations, 2020b]

Element	Description
Production	Figures relate to the total domestic production whether inside or outside the agricultural sector, i.e. it includes non-commercial production and production from kitchen gardens. Unless otherwise indicated, production is reported at the farm level for crop and livestock products (i.e. in the case of crops, excluding harvesting losses) and in terms of live weight for fish items (i.e. the actual ex-water weight at the time of the catch). All data shown relate to total meat production from both commercial and farm slaughter. Data are expressed in terms of dressed carcass weight, excluding offal and slaughter fats. Production of beef and buffalo meat includes veal; mutton and goat meat includes meat from lambs and kids; pig meat includes bacon and ham in fresh equivalent. Poultry meat includes meat from all domestic birds and refers, wherever possible, to ready-to-cook weight. Source: FAO Statistics Division

Continued on next page

Table 5.2 – Continued from previous page

Element	Description
Stock Variation	Comprises changes in stocks occurring during the reference period at all levels between the production and the retail levels, i.e. it comprises changes in government stocks, in stocks with manufacturers, importers, exporters, other wholesale and retail merchants, transport and storage enterprises and in stocks on farms. In actual fact, however, the information available often relates only to stocks held by governments and even these are not available for a number of countries and important commodities. In the absence of information on opening and closing stocks changes in stocks are also used for shifting production from the calendar year in which it is harvested to the year in which it is consumed. Net increases in stocks (add to stock) are generally indicated by the sign "-". No sign denotes net decreases (from stock). Source: FAO Statistics Division
Feed	Data refer to the quantity of the commodity in question available for feeding to the livestock and poultry during the reference period, whether domestically produced or imported. Source: FAO. 1986. The ICS users' manual. Interlinked computer storage and processing system of food and agricultural commodity data. Rome.

Continued on next page

Table 5.2 – Continued from previous page

Element	Description
Losses	Amount of the commodity in question lost through wastage (waste) during the year at all stages between the level at which production is recorded and the household, i.e. storage and transportation. Losses occurring before and during harvest are excluded. Waste from both edible and inedible parts of the commodity occurring in the household is also excluded. Quantities lost during the transformation of primary commodities into processed products are taken into account in the assessment of respective extraction/conversion rates. Distribution wastes tend to be considerable in countries with hot humid climate, difficult transportation and inadequate storage or processing facilities. This applies to the more perishable foodstuffs, and especially to those which have to be transported or stored for a long time in a tropical climate. Waste is often estimated as a fixed percentage of availability, the latter being defined as production plus imports plus stock withdrawals. Source: FAO. 1986. The ICS users' manual. Interlinked computer storage and processing system of food and agricultural commodity data. Rome.
Processing	*Description is omitted from FAOSTAT. Assumed to be the quantity of food lost in food processing.*

Continued on next page

Table 5.2 – Continued from previous page

Element	Description
Other uses (non-food)	Data refer to quantities of commodities used for non-food purposes, e.g. oil for soap. In order not to distort the picture of the national food pattern quantities of the commodity in question consumed mainly by tourists are included here (see also ”Per capita supply”). In addition, this variable covers pet food. Source: FAO. 1986. The ICS users' manual. Interlinked computer storage and processing system of food and agricultural commodity data. Rome.
Food supply (kcal)	Refers to the total amount of food available for human consumption expressed in kilocalories (kcal). Caloric content is derived by applying the appropriate food composition factors to the quantities of the commodities and shown in million units. Source: FAO Statistics Division

Table 5.3: FAO Food Balance Sheet Animal Product and Corresponding FAO Gridded Livestock

FAO Food Balance Sheet Animal Product	**FAO Gridded Livestock**
Bovine Meat	Cattle
Milk - Excluding Butter	Buffalo, Cattle, Goat, Sheep
Mutton & Goat Meat	Goat
Poultry Meat	Chicken, Duck
Pigmeat	Pig
Eggs	Chicken
Butter, Ghee	Buffalo, Cattle, Goat, Sheep

Table 5.4: GAEZ+ 2015 Crop Categories and Corresponding Foods from the FAO Food Balance Sheet

GAEZ+ 2015 Crop Category	**FAO Food Balance Sheet Food Category**
Barley	Barley and products
Banana	Plantains, Bananas
CropsNES	Apples and products, Bananas, Citrus, Other, Cloves, Coconut Oil, Coconuts - Incl Copra, Dates, Fruits, other, Grapefruit and products, Grapes and products (excl wine), Lemons, Limes and products, Nuts and products, Oilcrops Oil, Other, Oilcrops, Other, Oranges, Mandarines, Pineapples and products, Sesame seed, Sesameseed Oil, Spices, Other
Cassava	Cassava and products
Groundnut	Groundnuts, Groundnut Oil
Maize	Maize and products, Maize Germ Oil
Millet	Millet and products
Oilpalmfruit	Palm Oil, Palm kernels, Palmkernel Oil
Olives	Olive Oil, Olives (including preserved)
Othercereals	Cereals, Other, Oats, Rye and products
PotatoAndSweetpotato	Potatoes and products, Sweet potatoes
Pulses	Beans, Peas, Pulses, Other and products
Rapeseed	Rape and Mustard Oil, Rape and Mustardseed
Rice	Rice and products, Ricebran Oil
Stimulants	Cocoa Beans and products, Coffee and products, Tea (including mate)
Soybean	Soyabean Oil, Soyabeans
Sorghum	Sorghum and products
Sugarbeet	Sugar beet
Sugarcane	Sugar cane, Sugar non-centrifugal, Sugar (Raw Equivalent)
Sunflower	Sunflower seed, Sunflowerseed Oil
Vegetables	Onions, Tomatoes and products, Vegetables, other
Wheat	Wheat and products
Yamsandotherroots	Yams, Roots, Other

Table 5.5: Energy Content of FAO Food Balance Sheet Animal Products

FAO Food Balance Sheet Animal Product	**kcal per kg**	**USDA Source (SR LEGACY ID)**
Bovine Meat	2400	BEEF, GROUND, UNSPECIFIED FAT CONTENT, COOKED (172161)
Milk - Excluding Butter	610	MILK, WHOLE, 3.25% MILKFAT, WITHOUT ADDED VITAMIN A AND VITAMIN D (172217)
Mutton & Goat Meat	1430	GAME MEAT, GOAT, COOKED, ROASTED (175304)
Poultry Meat	2430	POULTRY, MECHANICALLY DEBONED, FROM MATURE HENS, RAW (171106)
Pigmeat	2180	PORK, GROUND, 84% LEAN / 16% FAT, RAW (168372)
Eggs	1430	EGG, WHOLE, RAW, FRESH (171287)
Butter, Ghee	7170	BUTTER, WITHOUT SALT (173430)

Table 5.6: Energy Content of GAEZ+ 2015 Production Crops

GAEZ+ 2015 Crop Category	**kcal per kg**	**USDA Source (SR LEGACY ID)**
Barley	3540	BARLEY, HULLED (170283)
Banana	890	BANANAS, RAW (173944)
Cassava	1600	CASSAVA, RAW (169985)
Groundnut	5700	PEANUTS, SPANISH, RAW (174263)
Corn grain	3650	CORN GRAIN, YELLOW (170288)
CropsNES	3388	Average of FRUIT SALAD, CANNED (174670) and MIXED NUTS, OIL ROASTED (169428)
Millet	3780	MILLET, RAW (169702)
Oil palm fruit	8840	OIL, PALM (171015)
Olives	1450	OLIVES, PICKLED, CANNED OR BOTTLED, GREEN (169096)
Other cereals	3664	Average of RYE GRAIN (168884) and OATS (169705)
Potato and sweet potato	770	POTATOES, FLESH AND SKIN, RAW (170026)
Pulses	3410	BEANS, BLACK, MATURE SEEDS, RAW (173734)
Rapeseed	8840	OIL, CANOLA (172336)
Rice	3700	RICE, WHITE, GLUTINOUS, UNENRICHED, UNCOOKED (168883)
Stimulants	10	BEVERAGES, TEA, GREEN, BREWED, REGULAR (171917)
Soybean	1470	SOYBEANS, GREEN, RAW (169282)
Sorghum	3290	SORGHUM GRAIN (169716)
Sugarbeet	3870	SUGARS, GRANULATED (169655)
Sugarcane	3870	SUGARS, GRANULATED (169655)
Sunflower	5840	SEEDS, SUNFLOWER SEED KERNELS, DRIED (170562)
Vegetables	720	VEGETABLES, MIXED, FROZEN, UNPREPARED (170471)
Wheat	3390	WHEAT, DURUM (169721)
Yamsand-otherroots	1180	YAM, RAW (170071)

Bibliography

[Aboah et al., 2019] Aboah, J., Wilson, M. M., Bicknell, K., and Rich, K. M. (2019). Identifying the precursors of vulnerability in agricultural value chains: A system dynamics approach. *International Journal of Production Research*, 59(3):683–701.

[Alexander et al., 2017] Alexander, P., Brown, C., Arneth, A., Finnigan, J., Moran, D., and Rounsevell, M. D. (2017). Losses, inefficiencies and waste in the global food system. *Agricultural Systems*, 153:190–200.

[Borgatti and Halgin, 2011] Borgatti, S. and Halgin, D. (2011). On network theory. *SSRN Electronic Journal*.

[Cassidy et al., 2013] Cassidy, E. S., West, P. C., Gerber, J. S., and Foley, J. A. (2013). Redefining agricultural yields: From tonnes to people nourished per hectare. *Environmental Research Letters*, 8(3):034015.

[Chung et al., 2020] Chung, M. G., Kapsar, K., Frank, K. A., and Liu, J. (2020). The spatial and temporal dynamics of global meat trade networks. *Scientific Reports*, 10(1).

[Conijn et al., 2018] Conijn, J., Bindraban, P., Schröder, J., and Jongschaap, R. (2018). Can our global food system meet food demand within planetary boundaries? *Agriculture, Ecosystems amp;amp; Environment*, 251:244–256.

[Coomes et al., 2019] Coomes, O. T., Barham, B. L., MacDonald, G. K., Ramankutty, N., and Chavas, J.-P. (2019). Leveraging total factor productivity growth for sustainable and resilient farming. *Nature Sustainability*, 2(1):22–28.

[Corong et al., 2017] Corong, E., Thomas, H., Robert, M., Tsigas, M., and van der Mensbrugghe, D. (2017). The standard gtap model, version 7. *Journal of Global Economic Analysis*, 2(1):1–119.

[Dalgaard et al., 2003] Dalgaard, T., Hutchings, N. J., and Porter, J. R. (2003). Agroecology, scaling and interdisciplinarity. *Agriculture, Ecosystems amp;amp; Environment*, 100(1):39–51.

[Deconinck et al., 2021] Deconinck, K., Giner, C., Jackson, L. A., and Toyama, L. (2021). Overcoming evidence gaps on food systems. *OECD Food, Agriculture and Fisheries Papers*.

[Dietrich et al., 2019] Dietrich, J. P., Bodirsky, B. L., Humpenöder, F., Weindl, I., Stevanović, M., Karstens, K., Kreidenweis, U., Wang, X., Mishra, A., Klein, D., and et al. (2019). Magpie 4 – anbsp;modular open-source framework for modeling global land systems. *Geoscientific Model Development*, 12(4):1299–1317.

[Ercsey-Ravasz et al., 2012] Ercsey-Ravasz, M., Toroczkai, Z., Lakner, Z., and Baranyi, J. (2012). Complexity of the international agro-food trade network and its impact on food safety. *PLoS ONE*, 7(5).

[Fajzel et al., 2023] Fajzel, W., Galbraith, E. D., Barrington-Leigh, C., Charmes, J., Frie, E., Hatton, I., Le Mézo, P., Milo, R., Minor, K., Wan, X., and et al. (2023). The global human day. *Proceedings of the National Academy of Sciences*, 120(25).

[FAO, 2020] FAO (2020). Sustainable food and agriculture.

[FAO, 2022] FAO (2022). World food and agriculture statistical pocketbook 2022. https://doi.org/10.4060/cc2212en.

[Food and Agriculture Organization of the United Nations, 2020a] Food and Agriculture Organization of the United Nations (2020a). FAOSTAT - Trade. Accessed: March 25, 2023.

[Food and Agriculture Organization of the United Nations, 2020b] Food and Agriculture Organization of the United Nations (2020b). Food Balance Sheets. Accessed: March 25, 2023.

[Gereffi, 2019] Gereffi, G. (2019). *Global Value Chains and development: Redefining the contours of 21st century capitalism*. Cambridge University Press.

[Gerten et al., 2023] Gerten, D., Braun, J., Breier, J., Schaphoff, S., Stenzel, F., and Werner, C. (2023). Feeding 10 billion people and achieving negative emissions within planetary boundaries – insights from global modelling. *Potsdam Institute for Climate Impact Research*.

[Gerten et al., 2020] Gerten, D., Heck, V., Jägermeyr, J., Bodirsky, B. L., Fetzer, I., Jalava, M., Kummu, M., Lucht, W., Rockström, J., Schaphoff, S., and et al. (2020). Feeding ten billion people is possible within four terrestrial planetary boundaries. *Nature Sustainability*, 3(3):200–208.

[Gilbert et al., 2018] Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T. P., Vanwambeke, S. O., Wint, G. R., and Robinson, T. P. (2018). Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data*, 5(1).

[Godfray et al., 2018] Godfray, H. C., Aveyard, P., Garnett, T., Hall, J. W., Key, T. J., Lorimer, J., Pierrehumbert, R. T., Scarborough, P., Springmann, M., and Jebb, S. A. (2018). Meat consumption, health, and the environment. *Science*, 361(6399).

[Grogan et al., 2022] Grogan, D., Frolking, S., Wisser, D., Prusevich, A., and Glidden, S. (2022). Global gridded crop harvested area, production, yield, and monthly physical area data circa 2015. *Scientific Data*, 9(1).

[Gustavsson et al., 2011] Gustavsson, J., Cederberg, C., Sonesson, U., van Otterdijk, R., and Meybeck, A. (2011). Global food losses and food waste – extent, causes and prevention.

[Gómez and Ricketts, 2013] Gómez, M. I. and Ricketts, K. D. (2013). Food value chain transformations in developing countries: Selected hypotheses on nutritional implications. *Food Policy*, 42:139–150.

[Hawkesworth et al., 2010] Hawkesworth, S., Dangour, A. D., Johnston, D., Lock, K., Poole, N., Rushton, J., Uauy, R., and Waage, J. (2010). Feeding the world healthily: The challenge of measuring the effects of agriculture on health. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554):3083–3097.

[Haytowitz et al., 2022] Haytowitz, Ahuja, Wu, Somanchi, Nickle, Nguyen, Roseland, Williams, Patterson, Li, and et al. (2022). Usda national nutrient database for standard reference, legacy release: Ag data commons.

[Hoekstra and Mekonnen, 2012] Hoekstra, A. Y. and Mekonnen, M. M. (2012). The water footprint of humanity. *Proceedings of the National Academy of Sciences*, 109(9):3232–3237.

[Hoff et al., 2014] Hoff, H., Döll, P., Fader, M., Gerten, D., Hauser, S., and Siebert, S. (2014). Water footprints of cities – indicators for sustainable consumption and production. *Hydrology and Earth System Sciences*, 18(1):213–226.

[Jones and Ejeta, 2015] Jones, A. D. and Ejeta, G. (2015). A new global agenda for nutrition and health: The importance of agriculture and food systems. *Bulletin of the World Health Organization*, 94(3):228–229.

[Kabat, 2023] Kabat, L. (2023). Ess: Supply utilization accounts and food balance sheets - background information for your better understanding.

[Karakoc and Konar, 2021] Karakoc, D. B. and Konar, M. (2021). A complex network framework for the efficiency and resilience trade-off in global food trade. *Environmental Research Letters*, 16(10):105003.

[Karakoc et al., 2022] Karakoc, D. B., Wang, J., and Konar, M. (2022). Food flows between counties in the united states from 2007 to 2017. *Environmental Research Letters*, 17(3):034035.

[Kim et al., 2021] Kim, K.-H., Doi, Y., Ramankutty, N., and Iizumi, T. (2021). A review of global gridded cropping system data products. *Environmental Research Letters*, 16(9):093005.

[Kinnunen et al., 2020] Kinnunen, P., Guillaume, J. H., Taka, M., D'Odorico, P., Siebert, S., Puma, M. J., Jalava, M., and Kummu, M. (2020). Local food crop production can fulfil demand for less than one-third of the population. *Nature Food*, 1(4):229–237.

[Kloppenburg et al., 2017] Kloppenburg, J., Hendrickson, J., and Stevenson, G. W. (2017). Coming in to the foodshed. *Rooted in the Land*, page 113–123.

[Konar et al., 2018] Konar, M., Lin, X., Ruddell, B., and Sivapalan, M. (2018). Scaling properties of food flow networks. *PLOS ONE*, 13(7).

[Kriewald et al., 2019] Kriewald, S., Pradhan, P., Costa, L., Ros, A. G., and Kropp, J. P. (2019). Hungry cities: How local food self-sufficiency relates to climate change, diets, and urbanisation. *Environmental Research Letters*, 14(9):094007.

[Kummu et al., 2012] Kummu, M., de Moel, H., Porkka, M., Siebert, S., Varis, O., and Ward, P. (2012). Lost food, wasted resources: Global food supply chain losses and their impacts on freshwater, cropland, and fertiliser use. *Science of The Total Environment*, 438:477–489.

[Kummu et al., 2020] Kummu, M., Kinnunen, P., Lehikoinen, E., Porkka, M., Queiroz, C., Röös, E., Troell, M., and Weil, C. (2020). Interplay of trade and food system resilience: Gains on supply diversity over time at the cost of trade independency. *Global Food Security*, 24:100360.

[Li et al., 2023] Li, Y., Zhong, H., Shan, Y., Hang, Y., Wang, D., Zhou, Y., and Hubacek, K. (2023). Changes in global food consumption increase ghg emissions despite efficiency gains along global supply chains. *Nature Food*, 4(6):483–495.

[Lin et al., 2019] Lin, X., Ruess, P. J., Marston, L., and Konar, M. (2019). Food flows between counties in the united states. *Environmental Research Letters*, 14(8):084011.

[MacDonald et al., 2015] MacDonald, G. K., Brauman, K. A., Sun, S., Carlson, K. M., Cassidy, E. S., Gerber, J. S., and West, P. C. (2015). Rethinking agricultural trade relationships in an era of globalization. *BioScience*, 65(3):275–289.

[Manzini and Accorsi, 2013] Manzini, R. and Accorsi, R. (2013). The new conceptual framework for food supply chain assessment. *Journal of Food Engineering*, 115(2):251–263.

[Mekonnen and Gerbens-Leenes, 2020] Mekonnen, M. M. and Gerbens-Leenes, W. (2020). The water footprint of global food production. *Water*, 12(10):2696.

[Nickel et al., 2022] Nickel, S., Steinhardt, C., Schlenker, H., and Burkart, W. (2022). Decision optimization with ibm ilog cplex optimization studio: A hands-on introduction to modeling with the optimization programming language (opl).

[OECD, 2018] OECD (2018). Understanding the global food system - how we feed the world today.

[Picchioni et al., 2017] Picchioni, F., Aurino, E., Aleksandrowicz, L., Bruce, M., Chesterman, S., Dominguez-Salas, P., Gersten, Z., Kalamatianou, S., Turner, C., and Yates, J. (2017). Erratum to: Roads to interdisciplinarity – working at the nexus among food systems, nutrition and health. *Food Security*, 10(4):1141–1141.

[Poore and Nemecek, 2018] Poore, J. and Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science*, 360(6392):987–992.

[Puma, 2019] Puma, M. J. (2019). Resilience of the global food system. *Nature Sustainability*, 2(4):260–261.

[Ray et al., 2022] Ray, D. K., Sloat, L. L., Garcia, A. S., Davis, K. F., Ali, T., and Xie, W. (2022). Crop harvests for direct food use insufficient to meet the un's food security goal. *Nature Food*, 3(5):367–374.

[Ritchie et al., 2022] Ritchie, H., Rosado, P., and Roser, M. (2022). Environmental impacts of food production. *Our World in Data*. https://ourworldindata.org/environmental-impacts-of-food.

[Ritchie et al., 2020] Ritchie, H., Roser, M., and Rosado, P. (2020). Co and greenhouse gas emissions. *Our World in Data*. https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions.

[Roser et al., 2013] Roser, M., Ritchie, H., and Rosado, P. (2013). Food supply. *Our World in Data*. https://ourworldindata.org/food-supply.

[Schreiber et al., 2021] Schreiber, K., Hickey, G. M., Metson, G. S., Robinson, B. E., and MacDonald, G. K. (2021). Quantifying the foodshed: A systematic review of urban food flow and local food self-sufficiency research. *Environmental Research Letters*, 16(2):023003.

[Shepon et al., 2016] Shepon, A., Eshel, G., Noor, E., and Milo, R. (2016). Energy and protein feed-to-food conversion efficiencies in the us and potential food security gains from dietary changes. *Environmental Research Letters*, 11(10):105002.

[Sims, 2011] Sims, R. E. H. (2011). "energy-smart" food for people and climate issue paper.

[Springmann et al., 2018] Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., de Vries, W., Vermeulen, S. J., Herrero, M., Carlson,

K. M., and et al. (2018). Options for keeping the food system within environmental limits. *Nature*, 562(7728):519–525.

[Tu et al., 2019] Tu, C., Suweis, S., and D'Odorico, P. (2019). Impact of globalization on the resilience and sustainability of natural resources. *Nature Sustainability*, 2(4):283–289.

[Tubiello et al., 2023] Tubiello, F. N., Conchedda, G., Casse, L., Pengyu, H., Zhongxin, C., De Santis, G., Fritz, S., and Muchoney, D. (2023). Measuring the world's cropland area. *Nature Food*, 4(1):30–32.

[Vallarano et al., 2021] Vallarano, N., Bruno, M., Marchese, E., Trapani, G., Saracco, F., Cimini, G., Zanon, M., and Squartini, T. (2021). Fast and scalable likelihood maximization for exponential random graph models with local constraints. *Scientific Reports*, 11(1).

[Wang and Dai, 2021] Wang, J. and Dai, C. (2021). Evolution of global food trade patterns and its implications for food security based on complex network analysis. *Foods*, 10(11):2657.

[WHO, 2021] WHO (2021). Obesity and overweight.

[Willett et al., 2019] Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., and et al. (2019). Food in the anthropocene: The eat–lancet commission on healthy diets from sustainable food systems. *The Lancet*, 393(10170):447–492.

[WorldBank, 2020] WorldBank (2020). Data catalog.

[WorldPop, 2020] WorldPop (2020). Global 1km population total grid - 2000-2020.

[Yu et al., 2020] Yu, Q., You, L., Wood-Sichra, U., Ru, Y., Joglekar, A. K., Fritz, S., Xiong, W., Lu, M., Wu, W., Yang, P., and et al. (2020). A cultivated planet in 2010 – part 2: The global gridded agricultural-production maps. *Earth System Science Data*, 12(4):3545–3572.