ISSN 2521-7240

Food and Agriculture Organization of the United Nations

14

The resilience of domestic transport networks in the context of food security

A multi-country analysis

Background paper for
The State of Food and Agriculture 2021

FAO AGRICULTURAL DEVELOPMENT ECONOMICS TECHNICAL STUDY

14

The resilience of domestic transport networks in the context of food security

A multi-country analysis

Background paper for
The State of Food and Agriculture 2021

By

Andy Nelson
Professor, Faculty of Geo-Information Science and Earth Observation, University of Twente

Rolf de By
Associate professor, Faculty of Geo-Information Science and Earth Observation, University of Twente

Tom Thomas
Dr, Faculty of Engineering Technology, University of Twente

Serkan Girgin
Assistant Professor, Faculty of Geo-Information Science and Earth Observation, University of Twente

Mark Brussel
Senior lecturer and researcher, Faculty of Geo-Information Science and Earth Observation, University of Twente

Valentijn Venus
Lecturer and researcher, Faculty of Geo-Information Science and Earth Observation, University of Twente

Robert Ohuru
MSc, Faculty of Geo-Information Science and Earth Observation, University of Twente

Food and Agriculture Organization of the United Nations
Rome, 2021

Required citation:

Nelson, A., de By, R., Thomas, T., Girgin, S., Brussel, M., Venus, V. & Ohuru, R. 2021. *The resilience of domestic transport networks in the context of food security – A multi-country analysis.* Background paper for *The State of Food and Agriculture 2021.* FAO Agricultural Development Economics Technical Study No. 14. Rome, FAO. https://doi.org/10.4060/cb7757en

ISSN 2521-7240 [Print]

ISSN 2521-7259 [Online]

ISBN 978-92-5-135363-9

Cover photo: ©FAO/Giuseppe Bizzarri

Contents

Figures

Tables

Preface

This study aims to help bring the domestic food transport network into focus for *The State of Food and Agriculture 2021 – Making agrifood systems more resilient to shocks and stresses*. Transport infrastructure and logistics, not least domestic food transport networks, are an integral part of agrifood systems, and play a fundamental role in ensuring physical access to food at the local level, as well as in producing non-food agricultural output. The flow of food from farm to fork is vulnerable to various shocks; however, the resilience of this flow has rarely been studied. This study aims to fill that gap; in doing so, it develops a spatial analysis framework that realistically characterizes the physical transport network, and uses this framework to then analyse the network's ability to transport enough food to meet demand.

The analysis builds on a preliminary spatial workflow and on evaluated resilience metrics to analyse the structure of transport networks in the context of national food transport network resilience. For a total of 90 countries, it considers road, river and rail transport infrastructure, along with trade ports, border crossings and their respective import and export quantities. It then measures food transport network resilience for each country through three main indicators: proximity-based resilience, relative detour cost and alternative route availability.

Findings show that where food is transported more locally and where the network is denser, disturbances have a much lower impact. This is mostly the case for high-income countries, as well as for densely populated countries like China, India, Nigeria and Pakistan. Conversely, low-income countries have much lower levels of transport network resilience, although some exceptions exist.

A simulation of the impact of localized 1-in-10-year flooding events in Mozambique, Nigeria and Pakistan is also used to capture the effect of potential disruptions to food transport networks for crops in the three countries. The simulation illustrates the loss of network connectivity that results when links become impassable, potentially affecting millions of people.

Acknowledgements

We thank Andrea Cattaneo and Theresa McMenomy in the Agrifood Economics Division (ESA) of the Food and Agriculture Organization of the United Nations (FAO), as well as other members of *The State of Food and Agriculture 2021* Writing Team for their technical contributions and guidance. We thank Andi Shiraz for copy-editing and Daniela Verona (ESA, FAO) for the design and publishing coordination.

As part of the review process of *The State of Food and Agriculture 2021* report, this paper was revised by *The State of Food and Agriculture* team, the FAO specialists and external experts that reviewed *The State of Food and Agriculture* report, and the management team of FAO's Economic and Social Development Stream.

Acronyms

AI	artificial intelligence
CHEETAH	Chains of Human Intelligence towards Efficiency and Equity in Agro-Food Trade along the trans-Africa Highway
FAO	Food and Agriculture Organization of the United Nations
GADM	Global Administrative Areas
GDP	gross domestic product
GHG	greenhouse gas
GHSL	Global Human Settlement Layer
GIS	Geographic Information System
GSHHG	Global Self-consistent, Hierarchical, High-resolution Geography Database
IQR	interquartile range
KML	Keyhole Markup Language
OD	origin–destination
OSM	OpenStreetMap
PBF	Protocolbuffer Binary Format
PPP	purchasing power parity
QCL	Crops and Livestock Products (FAOSTAT)
SME	small and medium-sized enterprise
SPAM	Spatial Production Allocation Model
SWF	simplified workflow
UN DESA	United Nations Department of Economic and Social Affairs
WPI	World Port Index
WPS	World Port Source

Executive summary

Transport infrastructure and logistics – and domestic food transport networks in particular – are an integral part of agrifood systems, and play a fundamental role not only in ensuring physical access to food but also in producing non-food agricultural output. The flow of food from farm to fork is vulnerable to various shocks; however, the resilience of this flow has rarely been studied. In this study, we develop a spatial analysis framework that realistically characterizes the physical transport network, and use this framework to analyse the network's ability to transport enough food to meet demand. This approach supports easy visualization of the network and movement through it, and permits aggregation of network characteristics to subnational and national levels for tabulation and intercountry comparison.

This analysis builds on a preliminary spatial workflow and on evaluated resilience metrics to analyse transport network structure in the context of national food transport network resilience. For a total of 90 countries, it considers road, river and rail transport infrastructure, along with trade ports, border crossings, and import and export quantities.

It then measures food transport network resilience for each country through three main indicators: proximity-based resilience, relative detour cost and alternative route availability. Proximity-based resilience captures the way food is distributed between production and supply area. When average travel distance is higher than the optimum, systems are not very resilient. Relative detour cost evaluates the extra tonnage-minutes generated when a critical transport link is closed. Finally, alternative route availability measures the availability of alternative transport links when a given critical link is closed.

Findings show that where food is transported more locally and where the network is denser, disturbances have a much lower impact. This is mostly the case for high-income countries, as well as for densely populated countries like China, India, Nigeria and Pakistan. Conversely, low-income countries have much lower levels of transport network resilience, although some exceptions exist.

A simulation of the impact of localized 1-in-10-year flooding events in Mozambique, Nigeria and Pakistan is also used to capture the effect of potential disruptions to food transport networks for crops in the three countries. The simulation reveals the loss of network connectivity when links become impassable (see Section 5.3), either because the link is damaged (for example, when a bridge is washed away) or because access to that link is reduced (for example, when an access road to a main road or bridge is damaged or submerged). Such events have a major impact on the local population served by the affected transport of goods. In addition, they can also have an impact at the national level, as they compromise the ability of the transport network to efficiently transport goods throughout the country.

Overall, this work has provided a first geospatial framework to represent and model national food transport network resilience at a global scale. It has developed a unique and internally consistent database and plausible representations of complex transport networks, and has generated network and resilience metrics to characterize the network and its ability to transport food to meet demand. This work has established a new toolkit for measuring resilience, which promises further use and applications beyond this study. At the same time however, its conceptualization and development involved many decisions that deserve critical reflection, for which more work is needed.

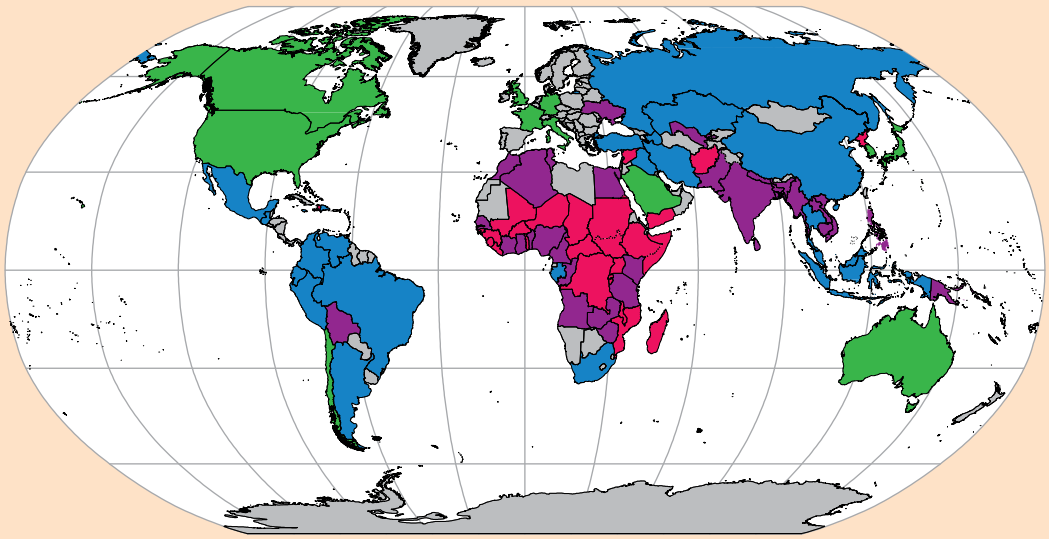
1 Introduction

The domestic food transport network is an integral part of agrifood systems. The flow of food from producer to consumer is vulnerable to shocks at many stages of its journey, but the resilience of this flow to shocks has rarely been studied, especially in comparison to the number of studies that have looked at production shocks (Davis, Downs and Gephart, 2021). Our aim is two-fold: Firstly, to develop a spatial analysis framework that realistically characterizes the physical transport network. Secondly, to use this framework to analyse the network's ability to transport sufficient food to meet demand, and its resilience to do so under shocks.

This analysis builds on a preliminary spatial workflow and on evaluated resilience metrics to analyse transport network structure in the context of national food transport network resilience. It considers road, river and rail transport infrastructure, along with trade ports, border crossings, and import and export quantities for a total of 90 countries. Figure 1 indicates the countries included in the analysis.

Our approach supports easy visualization of the network and of movement through it, and permits aggregation of network characteristics to subnational and national levels for tabulation and intercountry comparison. These data can be useful for future multinational or global studies.

FIGURE 1 Representation of the 90 countries by income group



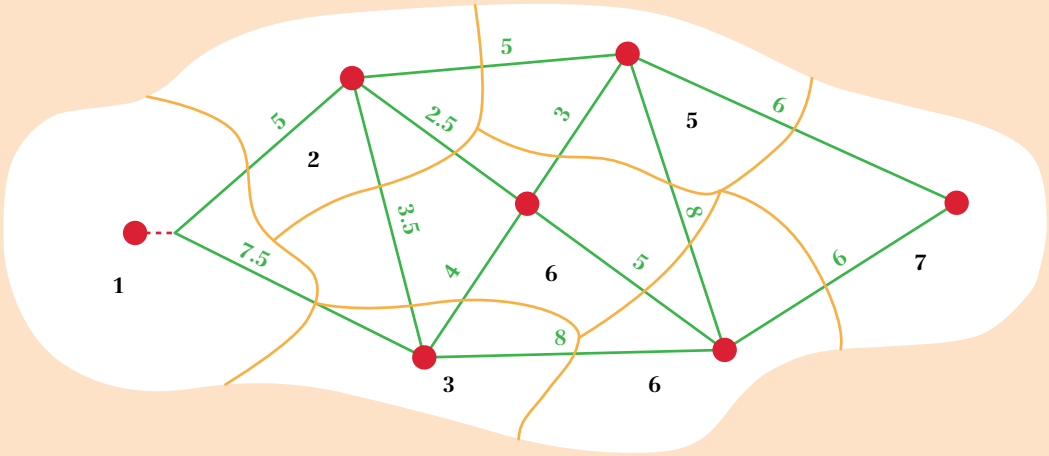
Notes: Dotted line represents approximately the Line of Control in Jammu and Kashmir agreed upon by India and Pakistan. The final status of Jammu and Kashmir has not yet been agreed upon by the parties. Final boundary between the Sudan and South Sudan has not yet been determined. Final status of the Abyei area is not yet determined. A dispute exists between the Governments of Argentina and the United Kingdom of Great Britain and Northern Ireland concerning sovereignty over the Falkland Islands (Malvinas).

Source: World Bank (2019) modified by the authors.

1.1 Inland transport network

Figure 2 is a simple diagram of the elements of the transport network within a country. Cities (red) are nodes on the network graph. While they are conceptually represented as points on the network, cities have spatial footprints that are more accurately represented as polygons.

FIGURE 2 Elements of the inland transport network



Notes: Cities (red), roads (green) and their distances, and catchments around each city (yellow).

Source: Authors' own elaboration based on Banks (2002).

Cities are connected to one another by transport infrastructure segments (as indicated by the green lines in Figure 2). These include roads, railroads, or rivers that support the transport of food commodities via truck, train, barge or ship. The analysis does not consider air transport infrastructure, and sea transport is considered for international transport only. Each transport infrastructure segment is defined by type (road, rail or river) and class (major road, secondary road, etc.), and has a defined speed of travel which varies by type, class and country. Two cities may therefore be connected by multiple segments, each of a different type.

Cities act as centres of attraction for services and employment for a surrounding catchment area; these are delineated by the yellow lines in Figure 2. The analysis considers all cities of 50 000 people or more, and associates each one with a single catchment; each catchment consists of a population that consumes food and a land area that is used to produce crops or food. When computing the demand for food, we therefore account for the total population of the catchment as a whole (and not just the population of the city within that catchment).

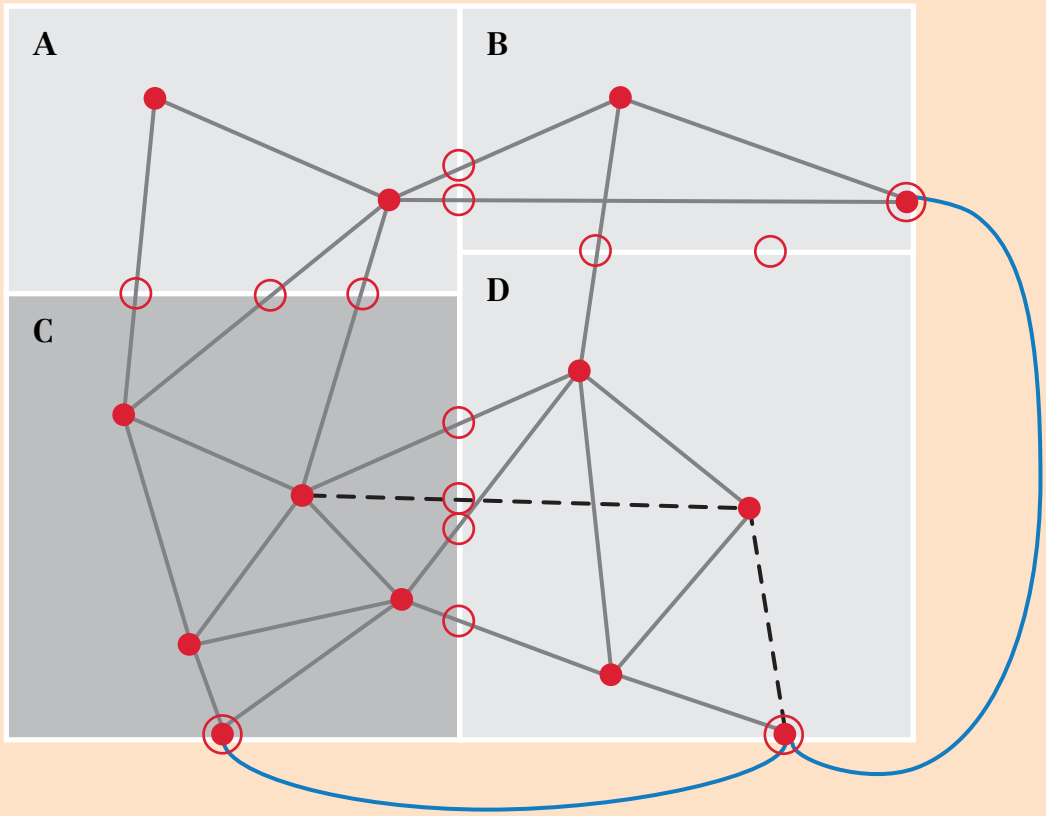
Trade, through import and export, takes place at sea and land border crossings as explained in the next section. These *trade stations* are properly embedded in the transport network as nodes, to play their role in the transport of food.

On this basis, we define the spatial and tabular data requirements to represent each of these elements, and the means to represent disruptions to movement and production. Six key requirements for the analysis were that all data must: (i) be open-source to permit redistribution and reuse; (ii) have global coverage to allow for analysis of any country and for comparisons across counties; (iii) represent the state of the art (in terms of recentness and quality of data); (iv) be spatially detailed enough to permit plausible subnational representations; (v) preferably be updated on a regular basis; and (vi) include certain layers that allow for adjustments in their spatial or attribute information so as to represent disruptions.

1.2 International transport network

Figure 3 is a diagram of the elements of the transport network between countries. Trade stations (red circles) are nodes on the network graph, and act as gateways to trade stations in other countries. To model long-haul transport through intermediate countries, land-based trade stations (i.e. border crossings) are also connected to other trade stations in the same country (via the grey and black lines). The network currently includes road, rail and sea transport at the global level.

FIGURE 3 Elements of the international transport network



Notes: Red circles represent trade stations that are connected to other trade stations by road (grey lines), rail (dashed lines) or sea (blue lines).

Source: Authors' own elaboration.

1.3 Production, import, export and food available for consumption

We follow the assumptions and definitions used in the Food Balances of the Food and Agricultural Organization of the United Nations (FAO) (FAO, 2017). We consider the top ten supplied commodities per country based on national supply quantities, excluding beverages, livestock, poultry, dairy, fish and seafood. Globally, this subset accounts for 66 percent of all food supply, 60 percent of calories, and 58 percent of protein.

Each food crop considered is grown in specific locations, such that its production (in tonnes) is spatially allocated to production zones in each country. Food available for consumption is spatially allocated based on the spatial distribution of population. We assume food available

for consumption per capita is constant within a country, although different food commodities have different per capita consumption. Import and export quantities are allocated to import and export locations based on the most likely mode of transport: ports for trade via sea or river routes, and land border crossings for trade via road or rail. We do not account for any informal international trade.

The quantity of food available domestically (*domestic food supply*) is equal to the amount of production and stocks (either positive or negative) plus imports, minus exports. As per FAOSTAT, the quantity of food available for consumption (*domestic food demand*) equals the domestic food supply minus any food that is used for feed, seed, or manufacturing processes that produce a derived food product, any food used for any other non-food products, and any food lost between production and supply (FAO, 2020).

1.4 Movement of food through the transport network

We make the following assumptions regarding the movement of food through the transport network:

- Each catchment produces a certain amount of food, and that production is associated with the city (node) at the centre of that catchment. Centrality is based on travel time.
- We do not model the initial transport from the farm to the nearest city, nor from the city to consumers within the same catchment. We ignore that food flows, in part, to processing factories and storehouses, instead of directly to consumers. We are assuming that processing plants are evenly distributed over cities. Since we concentrate food production of a catchment in the city, a substantial amount of food processing may already occur in that city (as it is the closest) before it is transported into the network.
- Each catchment has a certain population. Some – or all – of the locally produced food may be consumed by the population of the catchment.
- Food that is produced in a catchment but which is excess to the consumption needs of that catchment is distributed via the transport networks to all other cities. Cities that are closer receive a higher share if they have demand for that food.
- International trade (import/export) is modelled separately by a dedicated international transport network and trade distribution model. The results are transferred to the inland model using trade stations. While the model also quantifies trade volumes that transit through third countries, these numbers are not used in our analysis.
- Trade stations are additional nodes in the network that are not associated with a city or catchment, and their import or export values are added/subtracted to the total country production.
- We do not impose any constraints on the quantity of food that can be transported along a transport network segment in a given amount of time.
- We assume that globally, road infrastructure is relevant to transport only on roads of tertiary class or higher. In high-income countries we ignore the tertiary road class. In three low-income countries, we include unclassified roads and tracks, for cases where they are the only infrastructure connecting a city to the national network. In a few other countries, we include such roads at the local level only, to connect otherwise disconnected cities to the network.
- We assume that travel is possible bidirectionally, and direction does not influence cost.

2 Methods

KEY MESSAGES

- Using data from OpenStreetMap (OSM), we developed a complete data set for road infrastructure at country level, complemented with data on navigable rivers and associated railroad stations and ferry routes.
- We spatially assigned food production using 2010 and 2017 Spatial Production Allocation Model (SPAM) data, and food supply using 2015 Global Human Settlement Layer (GHSL) data.
- We calculated production, supply, imports and exports for each country, averaged over the years 2015–2017 using FAOSTAT data; where these data were missing, we used remedial tactics to address the data gaps.

2.1 Spatial base data

Global fundamental base data were collected to allow quantification and spatialization of characteristic functions of national food transport systems (Davis, Downs and Gephart, 2021). As food transport uses the national transport infrastructure, we collected complete, national-level road data from OSM (OpenStreetMap, 2021), and complemented this with OSM data on railroads, ferry routes, associated railroad stations and ferry terminals. Data on navigable rivers were retrieved from HydroSHEDS (Lehner, 2013), and on ports from World Port Index (WPI) (National Geospatial-Intelligence Agency, 2017) and World Port Source (WPS) (World Port Source, 2021). All these data were embedded in a base layer that provides countries sourced from Global Administrative Areas (GADM) 3.6 (Global Administrative Areas, 2021). For the international transport network, administrative boundaries (level 0–2) and road data from OSM were collected and utilized at the global level to identify border crossings, whereas the Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) was used as input for sea route identification (Wessel and Walter, 2017). Ferry routes considered elementary for connecting some regions were treated as a special, slower road type. Due to data constraints, we did not include domestic sea freight/shipping routes, nor did we include air freight, whether domestic or international. (The latter is also considered a minor form of food transport.)

We aimed to spatially assign places of food production using 2010 and 2017 SPAM data (Yu *et al.*, 2020). Food supply was spatially assigned using 2015 data from GHSL (JRC, 2015). As larger cities typically act as local hubs for regional food trade, we sourced GHSL for cities with a population above 50 000. These cities then became the centres of regional food trade catchments, with travel time as a metric (Cattaneo, Nelson and McMenomy, 2021).

We characterized and spatially assigned sea trade ports and land border crossings as key food transport locations between countries. While we modelled international food transport for transit countries, we did not analyse the transit of food volumes through these countries for the resilience metrics.

2.2 Food volume equations

In order to understand national food volume equations, we collected tabular data for all countries (Table 1) and for all crops (Table 2). We focused on the years 2015 to 2017. Crucial for our work on food volumes is a pivot table that (dis)aggregates and translates between the crop codes and the names of the different sources we used: MapSPAM, FAOSTAT Crops and Livestock Products (QCL), FAOSTAT Food Balances and FAOSTAT Detailed Trade Matrix (FAO, 2020). Most of our data derive from FAOSTAT, and where data were missing, we used remedial tactics to address the gaps.

TABLE 1 Spatial data layers, sources and data filters used

Layer	Source	URL	Filter
Countries	GADM	gadm.org	Level 0
Cities	GHSL	ghsl.jrc.ec.europa.eu	> 50k population
Roads	OSM	openstreetmap.org	On road classes; network simplification workflow
Railroads and stations	OSM		
Ferry routes and terminals	OSM		
Administrative boundaries	OSM		Level 0–2
Trade ports	WPI	msi.nga.mil/Publications/WPI	
	WPS	worldportsource.com	Coincidence with WPI
Navigable rivers		hydrosheds.org	River class and average annual discharge
Border crossings	OSM and derived		Border crossing generation workflow
Coastal lines	GSHHG	soest.hawaii.edu/pwessel/gshhg	
Sea routes	Derived		Sea route generation workflow
Human population	GHSL	ghsl.jrc.ec.europa.eu	
Crop production	MapSPAM	mapspam.info	Food crops

Source: Authors' own elaboration.

TABLE 2 Sources of food production, supply, imports and exports

Variable	Source	Missing	Remedy
Production	FAOSTAT's Food Balances (85 countries)	5 countries	Used FAOSTAT QCL
Imports	FAOSTAT Detailed Trade Matrix (74 countries)	16 countries	Used the mirrored reports of partner countries from which imported
Food available for consumption	FAOSTAT's Food Balances (85 countries)	5 countries	Used regional food supply averages (FAOSTAT) and UN World Population Prospects for national population figures
Exports	FAOSTAT Detailed Trade Matrix (74 countries)	16 countries	Used the mirrored reports of partner countries to which exported

Source: Authors' own elaboration.

For Burundi, the Democratic Republic of the Congo, Somalia, South Sudan and the Syrian Arab Republic, there are no production and supply data available. Annual production figures per crop for these were gleaned from FAOSTAT QCL; annual food supply figures were imputed from regional averages,[1] with annual population figures obtained from the United Nations Department of Economic and Social Affairs (UN DESA). Likewise, there are 16 countries for which we have no import and export data: Angola, Chad, the Democratic People's Republic of Korea, the Dominican Republic, Guinea-Bissau, Haiti, Iraq, the Lao People's Democratic Republic, Liberia, Mozambique, Myanmar, Somalia, South Sudan, Sudan, Uzbekistan and Viet Nam. For these, we used the reversed data as reported by their trade partners. While the imports reported by one country will not usually be identical to the exports of its partner country (and vice versa), we believe such adjustments should not in principle result in significant errors at the system level.

2.3 Selection of production crops and commodities available for consumption

Our selection of (groups of) food crops was constrained by the level of disaggregation in the spatial and tabular data on crop production/consumption and on the congruence between these data sources. Our starting point was the SPAM data set, which provides crop production estimates in tonnes at a five-minute resolution (approximately 10 km at the equator) for 42 different crops or crop aggregates, of which 27 are food crops. SPAM data are available globally for the year 2010, and for sub-Saharan Africa for 2017.

We compared the 27 food crops in SPAM to the crops listed in the FAOSTAT QCL and Detailed Trade Matrix, and to the crops and crop products in FAOSTAT's Food Balances. The online documentation for MapSPAM (MapSPAM, 2019) provides a list of FAOSTAT item codes to relate MapSPAM to the crops and item codes in the FAOSTAT QCL and Detailed Trade Matrix. The resulting 21 crops are listed in Table 3. For each of the 90 countries, we considered only the top ten supplied crops / crop groups, based on quantity (tonnes).

[1] For Burundi and the Democratic Republic of the Congo, the regional average of Middle Africa was used; that of Northern Africa was used for Somalia and South Sudan; and that of Western Asia for the Syrian Arab Republic.

2.4 Spatial assignation of production, supply, imports and exports

The 27 MapSPAM layers were aggregated into 21 layers; for each country, the production value (in tonnes) per pixel was adjusted so that the national sum matched the national production in FAOSTAT, as averaged over 2015–2017. Production was assigned to the catchment around each city. Catchments were defined based on a 30-arc-second cost-allocation model, which allocates each pixel to a city based on shortest travel time to the nearest city. The catchment population was used to compute the food available for consumption (in tonnes) per catchment, using national per capita food available for consumption (FAO, 2020).

TABLE 3 Relation between MapSPAM, FAOSTAT and our 21 crops/crop groups

MapSPAM crop	Codes in QCL and Detailed Trade Matrix tables	Codes in Food Balances	Our code	Our name
Wheat	15	2511	1	Wheat
Rice	27	2805	2	Rice
Maize	56	2514	3	Maize
Barley	44	2513	4	Barley
Pearl millet	79	2517	5	Millets
Small millet				
Sorghum	83	2518	6	Sorghum
Other cereals	68, 71, 75, 89, 92, 94, 97, 101, 103, 108	2515, 2516, 2520	7	Other cereals
Potato	116	2531	8	Potatoes
Sweet potato	122	2533	9	Sweet potatoes
Yams	137	2535	10	Yams
Cassava	125 (128)*	2532	11	Cassava
Other roots	135, 136, 149	2534	12	Other roots and tubers
Bean	176	2546	13	Beans
Chickpea	191	2547, 2549	14	Other pulses
Cowpea	195			
Pigeonpea	197			
Lentil	201			
Other pulses	181, 187, 203, 205, 210, 211			
Soybean	236	2555	15	Soyabeans
Groundnut	242 (243)**	2556	16	Groundnuts
Coconut	249	2560	17	Coconuts
Banana	486	2615	18	Bananas

TABLE 3 (cont.) Relation between MapSPAM, FAOSTAT and our 21 crops/ crop groups

MapSPAM crop	Codes in QCL and Detailed Trade Matrix tables	Codes in Food Balances	Our code	Our name
Plantain	489	2616	19	Plantains
Tropical fruit	490, 495, 497, 507, 512, 567, 568, 569, 571, 572, 574, 577, 587, 591, 600, 603	2611, 2612, 2613, 2614, 2617, 2618, 2620, 2625	20	Fruits
Temperate fruit	515, 521, 523, 526, 530, 531, 534, 536, 541, 542, 544, 547, 549, 550, 552, 554, 558, 560, 592, 619			
Vegetables	358, 366, 367, 372, 373, 388, 393, 394, 397, 399, 401, 402, 403, 406, 407, 414, 417, 420, 423, 426, 430, 446, 449, 459, 461, 463	2601, 2602, 2605	21	Vegetables

Notes: * Dried cassava (code 128) is traded; fresh weight conversion is 0.33. ** Groundnut is code 242 in QCL and 243 (shelled) in the Detailed Trade Matrix.

Source: Authors' own elaboration.

Total national import and export trade volumes were sourced from FAOSTAT's Food Balances and used as the leading numbers. We determined an annual average based on maximally three annual totals for the years 2015–2017. The split of imports/exports to partner countries was based on what the country itself reports in the FAOSTAT Detailed Trade Matrix, and not on what the partner country reported. An exception to this was made for 16 countries that do not report import/export volume by partner; here we used the mirrored export/import numbers from the partner countries. Among these, the Dominican Republic and Haiti, and Chad and Sudan are neighbour countries, which has likely led to some unknowns. The Detailed Trade Matrix numbers were used to assess the ratios by which to spatially assign trade volumes to partner countries.

Due to the data processing methods used, the FAOSTAT Detailed Trade Matrix includes entries with very small trade volumes between countries. In order to properly model food transport, the study requires the distribution of country-level trade volumes to multiple specific locations in each country (i.e. trade stations). However, further division of such small trade volumes was not acceptable; a filtering method was therefore applied to eliminate trade volumes that were insignificant for both importing and exporting countries (in other words, less than 0.1 percent of country-specific import and export volume). This resulted in a 40 percent reduction in data rows, while retaining 99.77 percent of the global trade.

The four-step transport model requires that supply equals demand. In this case, this requires that production plus imports equals supply plus exports, wherein "supply" represents food available for consumption after losses and other uses have been removed. We account for this by scaling up the supply such that the equation balances out. This has no impact on the resilience metrics that we compute, but ensures that the four-step transport model (Section 5) will generate a result.

3 Geospatial workflow and four-step transport model

KEY MESSAGES

- We developed five workflows to determine where food is produced and supplied, where it enters or exits a country, and where it travels to meet demand.
- One workflow took a highly detailed, open-source road data set and simplified it to represent food transport at the national level.
- Other workflows derived sea routes and estimated international trade volumes for these routes.
- We then used a four-step transport model to predict the amount of travel in terms of quantity (tonnes), of food kilometres and of food hours travelled over a network.

3.1 Workflows overview

The geospatial workflows produce the geospatially explicit data sets required to model national food transport systems for a given country. The central question is, where? Where is food produced and supplied, where does it enter or exit the country, and where does it travel to meet demand? To answer these questions and prepare the data required for a spatially explicit transport model, we developed the following five workflows, each of which is described in the following sections:

- spatial production and food available for consumption per catchment workflow;
- border crossing identification workflow;
- sea trade route workflow;
- trade volume assignation to trade stations workflow; and
- transport network simplification workflow.

Spatial production and food available for consumption per catchment

As described in Section 2.4, we base the spatial production data set on MapSPAM. For each country and for each of its top ten crops available for consumption, we spatially intersect the respective MapSPAM crop raster with the country's territory. We sum the crop production (in tonnes) for the country and determine a factor by which that production must be scaled to equal the FAOSTAT production annual average as reported for the crop. The crop-specific factor is then applied to the crop grid, which allows us to estimate crop production by grid cell aggregation per catchment.

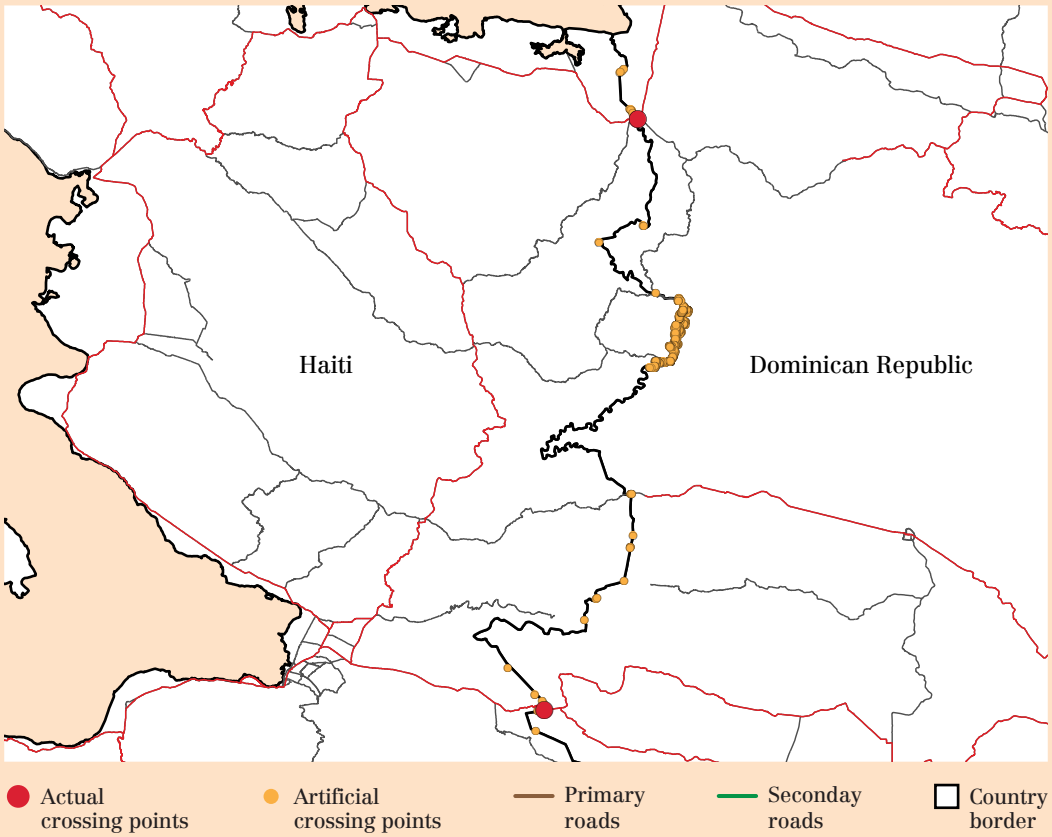
Spatially explicit figures for food available for consumption combine population data (i.e. on the spatial grid, as obtained from GHSL) with food demand data (i.e. annual kg per capita, as obtained from FAOSTAT's Food Balances). Next, we determine national population as an aggregation over the GHSL grid and compute a factor to scale against FAOSTAT's national population figure. This scale factor is applied to the grid at catchment level, in order to render catchment population in FAOSTAT's terms. That number is simply multiplied by the per capita food available for consumption for any crop that we are interested in, to give us the quantity of crop available for consumption per catchment.

Border crossing identification

Trade between countries only takes place at designated points where trade stations are located, thus at ports and major land border crossings. Our food transport model includes import and export volumes and associates them with the national network at these designated locations. We used a global ports data set for the analysis, but because no global data set of road transport trade stations exists, we derived these from the OSM road network and OSM administrative boundaries. To identify the important border crossings, a global roads data set comprising the highest three classes (motorway, trunk and primary) was intersected with global level-zero administrative boundaries (which correspond with country boundaries).

This part of our analysis uses OSM boundaries (instead of GADM, which is used elsewhere) to reduce the risk of non-conformity between roads and borders. Nevertheless, many artificial (non-existing) border crossings were found due to misaligned roads along borders (see Figure 4). To address this problem, a simplification method was developed to automatically remove points belonging to roads that zigzag along borders, and to consolidate proximate points in an iterative way that considers road types, prioritizing higher-level roads over others (for example, motorways over trunks). The method was validated by comparing its results to manually identified crossing points in Africa, the United States of America and Viet Nam, and was found to identify actual crossing points in a satisfactory manner. The global border crossing data set for roads is shown in Figure 5.

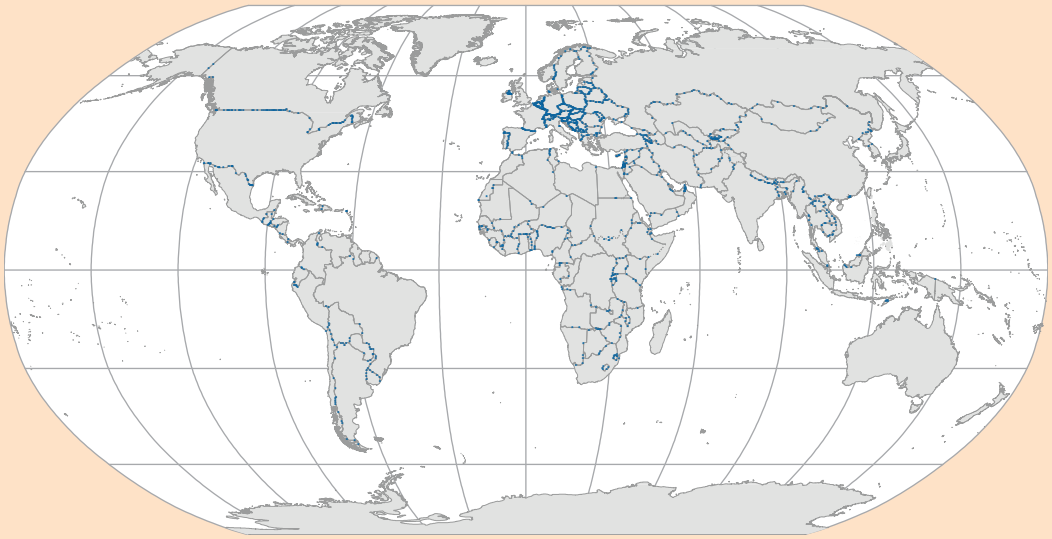
FIGURE 4 Example of border crossing simplification for the border between the Dominican Republic and Haiti



Notes: Actual crossing points (red) are separated from artificial (or zigzagging) points.

Source: UN (2020) modified by the authors.

FIGURE 5 Global distribution of border crossings generated in the study



Notes: For special country/region-specific disclaimers see notes of Figure 1.

Source: UN (2020) modified by the authors.

Sea trade route

A significant portion of global food transport takes place by sea. Port facilities play an important role in the food trade and act as gateways to the national transport network for gathering and distribution purposes. Likewise, sea routes that connect port facilities are significant elements of the global transport network, and may also be quite vulnerable – as demonstrated by the blockage of the Suez Canal in 2021.

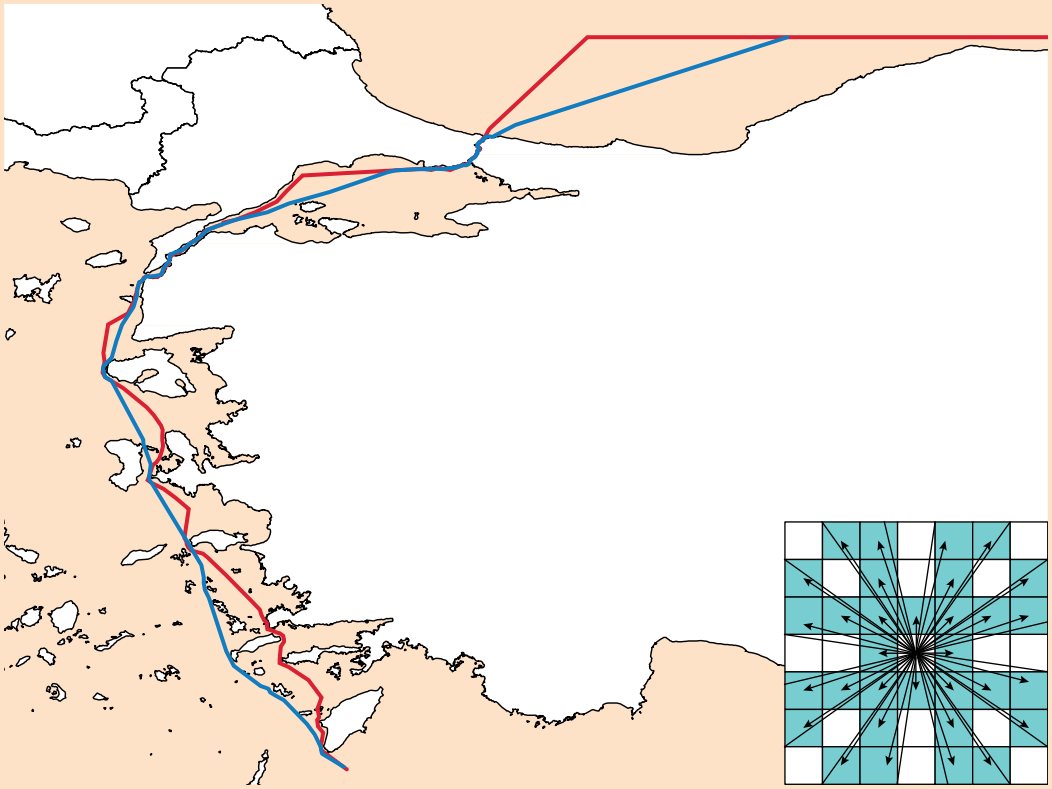
We consider more than 1 400 major (i.e. large and medium-sized) ports as part of the international food transport network. Although tabulated distances for sea routes between countries and selected major ports are available (National Geospatial-Intelligence Agency, 2001; Bertoli, Goujon and Santoni, 2016), modelling international trade requires a public data set providing sea routes between all major ports, which does not exist. Long-haul sea routes are determined by long-term experience and practices, which are difficult to reproduce with the limited data available. Given the scope of the study however, it is sufficient to have a good approximation of the routes, so as to allow for the elimination of routes that are infeasible in the context of global food trade distribution. For this purpose, a grid-based minimum distance and routing method was developed. The method, which is based on Dijkstra's (1959) shortest-path algorithm, specializes in sea routing and allows for proper circumnavigation of landmasses and islands along the route.

Although high-resolution vector coastline data sets are available, using vector data for pathfinding is complicated and computationally intensive. The common practice is to convert vector data into a regular raster for grid-based routing, which starts from an origin cell and reaches the destination cell by following a single neighbour (N) cell at a time. Several neighbourhood visiting patterns can be used for this purpose, the most common being N4 and N8, which visit four and eight cardinal directions respectively. Although N8 visits all direct neighbours, due to restricted directionality the resulting routes are not smooth and natural. To allow a higher degree of freedom in potential moves, the custom sea routing method includes an implementation of N32, allowing 32 (instead of 8) cells to be considered at each move (see Figure 6).

Accurate pathfinding requires a grid with a small cell size, as this allows for the identification of narrow passages or obstacles which are not uncommon at sea (for example, channels and small islands). However, small cell sizes result in very large grid sizes at global scale; these in turn are either difficult to process or require very long computation times. To retain essential connectivity information, the custom sea routing method includes a special data preparation step, which starts with a high-resolution grid and creates grids of progressively lower resolution by resampling gradually, until the target resolution is reached. Distribution of sea and land cells are considered at each sampling step to ensure transfer of connectivity information from the higher to the lower resolution.

Figure 7 shows an example of a set of sea routes from a single port (Mariupol, Ukraine) to all other ports, using the N32 neighbour visiting pattern with a connectivity-aware sea raster grid.

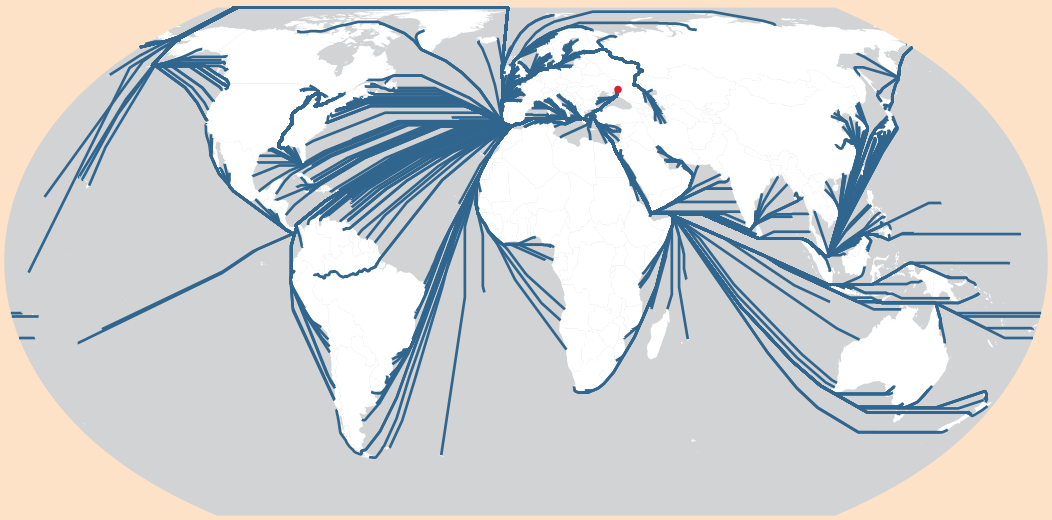
FIGURE 6 Comparison of sea routes obtained by N8 and N32 visiting patterns along the Aegean Sea, Dardanelles and Bosporus



Notes: The route in red was obtained using the N8 visiting pattern. The route in blue was obtained using the N32 visiting pattern. N32 moves are shown at bottom right.

Source: UN (2020) modified by the authors.

FIGURE 7 Sea routes from the port of Mariupol, Ukraine



Notes: For special country/region-specific disclaimers see notes of Figure 1.

Source: UN (2020) modified by the authors.

Trade volume assignation to trade stations

Each national transport model considers the spatial distribution of production and food available for consumption of each relevant crop by using areal catchments. The movement of crop products between catchments is also properly taken into consideration by using a country-specific transport network. To complement this and to have a complete picture of food transport, international trade (import/export) should also be integrated with the model considering spatial aspects. Unfortunately, the FAOSTAT Detailed Trade Matrix – which is the fundamental data source on international trade – provides information only at (partner) country level. Hence, there is no information available on where products enter or exit a country, or on the trade volumes at those locations. Border crossing and sea route data sets, as explained in the previous section, aim to fill the gap for the former ("where?"), whereas the international trade distribution model developed in the study aims to provide missing information for the latter ("how much?").

The international trade model is a sophisticated model that assigns crop-specific trade volumes to trade stations based on: total trade volumes between countries; available trade stations; direct as well as indirect connections between trade stations (the latter referring to transit freight through a transport chain); and transport costs, including factors such as type of transport, distance, number of intermittent stations and number of instances in which the transport mode needs to be changed (for example, from sea to road or road to sea). Our model automatically generates multi-segment transport routes between all trade stations, evaluates possible routes by comparing estimated transport cost through the use of representative metrics, and eliminates infeasible routes, thereby allowing for the logical distribution of trade volumes to the remaining routes according to the rules or preferences of the transport mode. Because international trade is complex and has many exceptions, the model also allows for the customization and correction of automatically generated data sets to yield better results. Considering the need to keep track of the manual interventions required, and to guarantee the reproducibility of results, the model features a specialized manual data entry method based on revision tables. To support the manual data entry method, the model uses a dedicated database and a custom-developed modelling framework; it is also well integrated with the existing study database and can easily import up-to-date data.

To distribute the FAOSTAT Detailed Trade Matrix country-level data on trade through trade stations, the model first simplifies available crop-specific trade data between pairs of countries; it does this by removing any entries for a given crop that are less than 0.1 percent of the total import or export volumes between the related countries for that crop. For the existing data set, the simplification resulted in a 37.1 percent reduction in pair-wise trade data, while retaining 99.8 percent of the global trade volume, thereby allowing more targeted distribution of trade to major trade stations. Links between trade stations can then be identified. A link is a direct connection between two trade stations located in two different countries. For border crossings, a link is created automatically for each pair of stations located opposite each other in neighbouring countries. For ports, sea routes are generated as explained previously, and are utilized accordingly. As with the links connecting different countries, inner links connecting trade stations located in the same country are similarly generated. The model currently assumes that all trade stations in a country are connected to each other directly, independent of the national transportation network, and the distances between the stations are calculated as great-circle distances. For each link and inner link, a transport cost is calculated based on weighted link length and on the types of stations at each end. Link length weighting factors of 1.0 and 2.0 are utilized for sea and land links respectively, which results in the preference of sea routes over land routes whenever possible. For border crossing links, the distance is set to zero. A fixed cost of 1 000 km is assigned to links having stations with the same transport mode (such as sea or road),

whereas for differing stations this cost includes a penalty for the change of transport mode, which is set as 2 500 km.

Next, the model automatically creates routes between trade stations. A route is a continuous path of links and inner links between two trade stations located in two different, non-neighbouring countries. The routes are generated by connecting links and inner links in the appropriate order and then summing their costs. Alternative routes between two stations are automatically handled by keeping the shortest route only: first by route cost, then by number of stages. For this purpose, the route is first compared to the lowest-cost route identified between the countries (if any). If the difference in cost is more than 10 percent, then the route is eliminated. If not, the route is then compared with the lowest-cost route identified between the stations (if any), and again, if it is more costly it is eliminated. In short, an existing route is replaced with a new route only if the new route is less costly by a difference of more than 10 percent, or if it has at least two fewer intermediate stations (in other words, if it crosses fewer countries). The model allows selected links or inner links to be disabled in order to customize routing when necessary, for example to consider disruptions or sociopolitical conditions more accurately. Once all transport routes connecting global trade stations are identified, the routes of the same transport mode connecting pairs of countries are prioritized.

The model allows a ratio to be assigned to each route so as to indicate the weight of the route as compared with routes of the same transport mode between related countries. For example, if a↔b is a road-to-road route between countries A and B, and if there are two other road-to-road routes between A and B, which are c↔d and e↔f, then assigning a ratio of 0.5 to a↔b results in a ratio of 0.25 for c↔d and a ratio of 0.25 for e↔f (whereas normally, each of the three would have a ratio of 0.33). Default route ratios are automatically calculated by using a ratio lookup table for station types and by considering all alternative routes between countries. Preference is given to major trade stations, such as major ports and border crossings located at major roads. The model then introduces directionality into the equation by converting routes into trade routes that consider the direction of trade – in other words, from exporting countries to importing countries (for example, a→b, if A→B). For each crop traded over a trade route, a trade route ratio is assigned, which is by default equal to the route ratio (for example, a→b: 0.5 for rice and a→b: 0.3 for wheat) (r_s in the equation below). Next, transport mode ratios between pairs of export–import countries are calculated by using a lookup table that is specific to transport mode (for example, A→B by road: 0.7 and A→B by sea: 0.3), and then customized for each crop (for example, A→B rice by road: 0.3 and A→B rice by sea: 0.7) (r_m in the equation below). Finally, crop-specific trade volumes between trade stations are calculated as follows:

$$a_c = A_c \cdot r_m \cdot r_s$$

where:

a_c = amount of crop c exported from station Si in country C to station Sj in country D

A_c = amount of crop c exported from country C to country D

r_m = trade ratio between countries C and D for crop c and transport mode of station Si

r_s = trade ratio between stations Si and Sj for crop c

Consistency between country-specific (r_m) and route-specific (r_s) ratios is automatically handled and guaranteed by the model, which also automatically handles dependencies between different data tables.

Transport network simplification

This workflow generates a connected graph that represents a country's transport network reasonably and simply. The nodes in the graph reflect cities associated with a food consumption/production catchment, as well as principal trade stations through which international trade takes place. Additional nodes represent major road junctions. The edges of the graph are attributed with a transport mode class, which comprises various higher-priority road classes, as well as rail, ferry and river transport modes. Another attribute is a cost factor, which expresses the travel time in minutes along the edge for the attributed transport mode. Edges are assumed to be bidirectional and transport cost is assumed to be independent of direction.

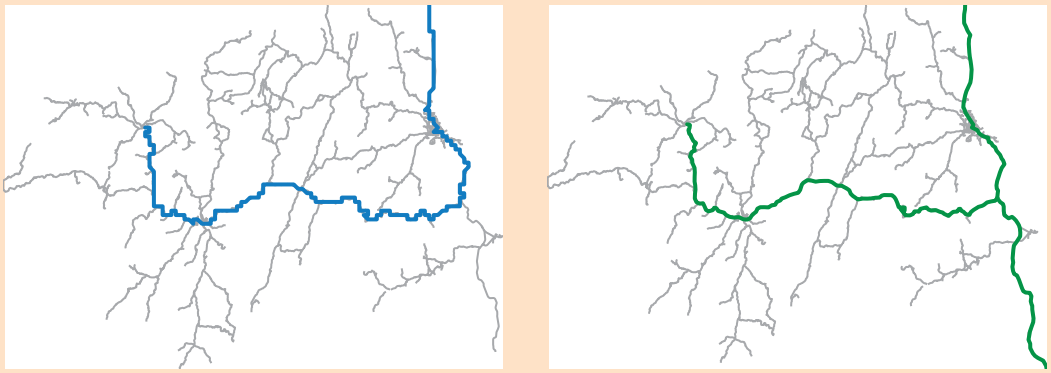
A connected graph lends itself well to transport modelling if and only if it is sufficiently downsized. As our source geospatial data is mostly taken from OSM, this results in some challenges. The prefiltered road data sets for China, India and the United States of America consist of 3.8, 0.5 and 2.8 million road segments respectively. The cities number 1 841, 3 247 and 321, and trade stations number 89, 57 and 258 respectively. Graphs of such size cannot be practically analysed for the purposes of resilience metrics unless they are simplified. The OSM data lends itself to such simplification because of considerable redundancy in the context of its transport models. Single motorway lanes, feeder roads and link roads are all separate road segments, and complex junctions may consist of tens of edges. While OSM road segments are generally well-behaved in how they connect with each other, the points of connection appear to be the arbitrary result of the acquisition and maintenance history, and a single stretch of road may be represented by multiple segments in line. These can in principle be combined into one segment, if the contributing segments are of the same transport mode. In addition, urban (residential) roads make up a large proportion of the total, but are largely irrelevant for national transport studies. We exclude these roads from our data and replace them with an artificial within-city road network that restores road connectivity within the urban footprint.

The principle of our simplification workflow (SWF) is to collapse the coordinate space to a grid at a fixed distance (by "snapping" or repositioning road segments to a fixed grid). The distance we chose is 0.004 degrees in the native coordinate system (445 m at the equator), as early experimentation indicated it was the best compromise between sufficient collapse and acceptable coordinate precision loss. Figure 8a illustrates the snap-to-grid procedure, with original road data in red and snapped road data in blue. The snapped road data will put multiple road segments (such as road lanes running in parallel) on top of each other, after which a geospatial union operator will allow them to be considered as one. We register the highest transport mode available for each segment for later use. The collapse/snap operations do cause unwanted artefacts as well, and an array of techniques is applied to remove these from the graph, each technique simplifying the graph further. Eventually, the snapped graph edges are unsnapped again, vertex-by-vertex, to arrive at an original coordinate pair that is true in the source data for a road in the same road class. This restores coordinate accuracy to some extent. The overall effect is illustrated in Figure 8b, with the original data shown in red as before, and the resulting unsnapped road data that is used in the graph shown in green. This simplification method leads to a systematic shortening of roads (in both metric length and travel time), which is likely less severe in flat areas as compared to hilly or mountainous areas.

FIGURE 8 Original OSM roads for an area around Mangabe, Madagascar

A. ORIGINAL OSM ROADS AND THEIR SNAPPED-TO-GRID COUNTERPARTS

B. ORIGINAL OSM ROADS AND THEIR UNSNAPPED COUNTERPARTS



Notes: The village of Mangabe is located on the northern shores of Lake Itasy, approximately 80 km west of Antananarivo. In both figures, original road data is shown in grey. In Figure 8a, snapped road data is shown in blue, and in Figure 8b, unsnapped road data is shown in green.

Source: UN (2020) modified by the authors.

All of the above is done for roads only, as we keep railroads, long ferry stretches, and river networks separate. Rivers are only considered if we have river ports for the country, which is the case for 12 countries in our target list (Argentina, Brazil, Canada, China, Germany, France, the Islamic Republic of Iran, Iraq, Kazakhstan, Peru, the Russian Federation and the United States of America). Where ferry routes connect decently with the road network, they are simply added. Relevant ferry routes that do not connect well are added manually at a later stage of the SWF. Both railroads and rivers are handled as graphs in isolation that only connect with the main graph through railroad stations and river ports respectively. These in turn are then connected with the main graph through an artificial, lower-class road. We use all river ports and railroad stations available to us. This appears reasonable for ports but is overkill for railroad stations, which may be numerous in larger cities. At this point in time, we unfortunately have no data with which to filter for stations that are relevant to food freight. The added connector segments do not exist in real life as we create them, but are assumed to be realistic placeholders in terms of food transports and related costs.

The SWF consists of 16 steps, described as follows:

1. **Data entry and preparation:** The transport network uses OSM data, which are downloaded from the Geofabrik geoportal (Geofabrik, 2021) in a compressed file format (Protocolbuffer Binary Format or PBF). Using the polygon boundary of a particular country, we cut the PBF file and extract all the relevant data, which include roads, railroads and stations, ferry routes and their terminals, and waterways. The extracted data are then inserted into the database, which is used in the subsequent steps.
2. **Precondition testing:** In this step, we identify data that could cause the workflow to break at later stages. We specifically look for problematic geometries, and label them either as *non-included* or as *to-be-repaired*. Such repairs are also conducted in this step. Of specific interest are geometries that have unexpected intersections (specifically overlaps) with each other.
3. **Road segment snapping:** In this step, we pursue the snap of each constituent vertex of each geometry onto a regular grid, with snap points regularly spaced out at 0.004 degrees'

distance in both longitude and latitude. This means that at the equator, snap points are spaced at 445 m, and this number gradually decreases with growing latitude. Such a collapse of coordinates may bring multiple vertices onto the same snap point, and thus render resulting geometries invalid. After repairs for this, resulting geometries are collected and their mutual intersection points are identified, working towards a *fully noded network*. We use a tiling approach to battle computing time complexity, which requires a merge phase for the per-tile results. The approach aggregates over edges of multiple road classes, so assigning a proper road class to each resulting segment is also required after the merge. This is one of the two most time-consuming SWF steps.

4. **Road segment reconciliation:** While the previous step regularized graph segments in terms of coordinates, it also created geometric artefacts (such as two-vertex edges with zero length), which need to be removed from the graph. In this step, we find several anomalies and simplify the graph further, while leaving intact its topology. Two fundamental techniques are used: The first is *topological simplification*, and consists of a search-and-replace of edge combinations that form unwanted structures in the graph, such as small triangles and semi-parallel edges with shared end nodes. These structures are traced and removed, while ensuring a correct graph remains. The second technique, which we call *long-edging*, consists of finding maximal stretches of edges of the same road class by combining two-degree nodes. Such stretches can be replaced by a single edge that has a union geometry. These two techniques are also time-consuming; hence this step, together with the preceding, is the most performance-intensive. We made significant effort in the implementation to also allow for the handling of the biggest countries in terms of data (China, India and the United States of America).
5. **Road segment unsnapping:** In this step, we take each edge and map all its vertices back to a nearby known position in the original road data. We respect road class and map back only onto an original road of the same class. To some extent, this repairs the coordinate accuracy of constituent vertices. The unsnapping method is based on a large set of (snap/unsnap) point pairs that is built up in advance. Guarantees are built in to ensure that valid geometries are delivered.
6. **Graph repairs I:** While the previous steps depended in large part on geometric computations, the resulting graph is an abstraction of reality in which geometry does not play a significant role. This step therefore aims to construct a graph where nodes have numeric identifiers, and where edges are primarily characterized by the two identifiers of their end nodes, along with attributes such as edge length and edge cost. We devised a mechanism to derive a unique (big integer) identifier from a location's coordinate pair; this allows us to construct a bijection (one-to-one correspondence) between geometric points and node identifiers. This step creates the graph as a combination of a node set, an edge set and a graph set.
7. **Urban footprinting:** This step aims to further simplify the graph by removing the high number of road edges that are usually generated within the urban footprint, and replacing these with a simple, artificial, skeleton road system that allows for the connection of roads entering the urban footprint. The base data for this step is the GHSL urban footprint data set, which is raster-derived. We generalize the polygons with a line simplification mechanism and a subtle tolerance factor that is just high enough to remove the raster cell effects that are present in the original. Next, any overlaps between road segments and the urban footprint are removed. In each case, the overlapping segment removed is replaced with a straight segment which connects to the nearest vertex of the artificial, skeleton road network as previously created for the city, using an approximate medial (skeleton) axis. While this step mostly only affects residential roads (if included), it may reduce the number of urban road segments by a factor of ten or more.

8. **Graph repairs II:** This step involves revisiting the graph once more and performing a series of minor clean-up actions.
9. **Ferry routing:** The OSM data includes ferry routes, but these display rather diverse modes of connecting to other network elements. Here, we include only the ferry routes that directly connect with road segments at both ends of the route. Additional ferry routes are covered in the step involving manual segment addition. For the present, ferry terminals are ignored.
10. **Railroad routing:** As OSM data represents railroad segments as single-track entities, railroads require a similar approach as that used for roads, to arrive at a more simplified network topology. In principle, we follow a snap/unsnap cycle here, as we did for roads with similar simplification techniques (as described previously). The railroads are kept separate from roads because points where they spatially overlap are not points where a change of transport mode is possible. Thus, we also model railroad stations as "mode change points", enabling this by adding a connector road segment to each station. Unfortunately, current OSM railroad station data do not give us enough information to decently filter between passenger-only stations and others; this means our networks include more stations than are actually available for food transport.
11. **Trade station inclusion:** These entities need to become first-class nodes in the network graph. We treat trade stations within an urban footprint differently from those outside of it. For the latter, we essentially snap the trade station to the nearest node and split up the edge if needed. For urban trade stations, we leave the position unchanged, but create a connector road to the urban road network.
12. **River routing:** Inclusion of rivers is similar to that of railroads in the sense that no connector nodes should be constructed in the graph such that they overlap spatially with roads or railroads. The only nodes that allow a change of transport mode will be those of river ports; though we exclude a river port if very close to a seaport, as they are likely the same. River ports are considered first-class nodes that connect with river edges, and are also connected to the main network by an added connector road. The implementation of this step leaves some room for improvement; for example we are quite certain that for some countries our set of river ports does not adequately represent the ports used for food transport. Moreover there is often a high number of unused river edges – namely those that do not connect to any river port – which should be removed. Lastly, we should ensure that end-of-river river ports become fully connected to nearby seaports.
13. **Segment lengths and costs:** This step revisits the graph's edges again, ensuring that end node identifiers are proper and that the length of edges and the cost in travel time are appropriately computed. The scheme for node identifiers is as discussed previously. Edge length in metres is the length of the associated geometry. Edge cost is determined in minutes of travel time, based on edge length and on the cruising speed for the transport mode in the country at hand. Where such values are unknown for a given country, default global values are in place. Edge costs are thus minimal estimates, and are likely lower than travel time over the corresponding edges in real life.
14. **Manual segment addition:** The last data generation step is executed manually, based on a visual inspection of the emerging network and on reports of unconnected cities that are fed back from initial transport model runs. We have found this scrutiny crucial to obtaining reasonably correct and complete networks. While some countries' networks appear to need little interference of this type, there are other countries for which we encountered surprising network omissions. For instance, exclaves may present problems of overall connectivity, as is the case for Alaska (United States of America), Kaliningrad (Russian Federation) and the Nakhchivan Autonomous Republic (Azerbaijan). For these,

main roads through other countries (Canada, the Baltic states and Belarus, and Armenia) were included manually, so as to allow for transport between the network components. Islands and even island states also presented connectivity issues when insufficient ferry routes were included, with Indonesia overall and Hawaii (United States of America) being points in case. Where ferry routes were known to exist, we included them; however, in cases where these were unknown, we may have left islands disconnected from the main network. In addition, we did not attempt to connect islands that have no city as defined for this project. Lastly, some countries show cities which are off the network, sometimes apparently by design. This is the case in Ethiopia for example, where many rural cities are not connected by a major road, even when one is nearby.

15. **Data delivery:** The data delivery step simply constructs four data sets from the results obtained: the country's cities, the country's trade stations, the country's transport graph, and an adjusted version of that graph for city node centrality analysis. They are created as materialized views for direct software consumption (such as the transport model), but can also be exported into various formats. These include GeoJson or Keyhole Markup Language (KML) for the web; PostGIS, SQLLite or SpatialLite for databases; and KML, Shapefile or Geopackages for GIS (Geographic Information System).
16. **Data clean-up:** This is a simple step in which intermediate data results created during the workflow are removed from the database.

3.2 The four-step transport model

Our transport model is based on the well-known four-step travel demand model (see for example, Ortúzar and Willumsen, 2011), which was developed in the United States of America in the 1950s, and which is still the main transport planning tool currently used at different spatial scales. The theoretical basis of this model lies in travel demand theory and random utility theory. Its main purpose is to predict the amount of travel (normally in terms of people and vehicles over a network). It is used for the analysis of the current situation and for forecasting.

The model is based on the execution of four consecutive steps: trip generation, trip distribution, mode choice and trip assignment. Trip generation involves an estimate of the number of trips made from a given zone, often based on land use and socio-economic information; and the number of trips attracted to a zone, often determined from land use characteristics/functions such as employment, shops, etc.

In our model, we use agricultural production in tonnes for each zone (catchment) to represent production, and food available for consumption in each zone (catchment with one city) to represent attraction. We therefore work with goods rather than with trips. A skim matrix is created to hold the origins (O), the destinations (D) and the impedance (travel time) between each origin and destination.

This matrix is useful to calculate distances and travel times between the nodes, but it does not tell us how much of the production from a particular catchment goes to any other catchment(s), or specify which catchments it goes to. That is calculated in the second step (goods distribution) where we use a doubly constrained gravity model, analogous to Newton's gravity model. This model assumes that large production areas and large cities are more important poles of attraction than smaller ones. It also assumes that locations that are close to each other attract more goods travel than locations that are spaced further apart.

We use the crop production total in each catchment as one constraint, and the total quantity of crop available for consumption in each catchment as the other constraint. To distribute the production across catchments, we estimate a so-called distribution function.

This is usually of the negative power function or negative exponential type, whichever is more fitting to the data. As a result, we get a matrix with discounted impedances. This matrix needs to be balanced iteratively using what is known as the Furness method, to ensure that the production and consumption figures for O and D are correct. Once that is done, we know how much product a catchment receives, where the goods come from, and where the production in each catchment is going; in other words, we know all the flows between production and consumption catchments. In transport terms this would be the trip matrix; we call it the *goods matrix*.

The third step involves the distribution of trips over different modes. As we focus on the flow of goods in tonnes and not on which vehicles (with their specific capacities) transport the goods, we ignore this step. Nevertheless, we introduced railroad and river networks in the model, next to the road-based networks. These networks are provided with their own impedances. We do not really analyse the specific proportions in which goods are transported by different modes; rather we look at the total that is transported.

In the fourth step, we use this balanced matrix to model the flows over the network. We apply a so-called *all-or-nothing* procedure, which uses the calculated shortest route between every origin–destination (OD) pair, selects the links involved, and assigns the flow from the matrix to these links. The links could be road links, railroad links, river links, or a combination – depending on which sequence of links provides the shortest route between the OD pair. More complicated methods use road / railroad / river capacity in relation to flows, but in view of data limitations, this was not feasible here.

4 Metrics

KEY MESSAGES

- We measure food transport network resilience with three main indicators: proximity-based resilience, relative detour cost and alternative route availability.
- Proximity-based resilience captures the way food is distributed between production and consumption area. When average travel distance is higher than the optimum, systems are not very resilient.
- Relative detour cost evaluates the amount of extra tonnage-minutes generated when a critical transport link is closed.
- Alternative route availability measures the availability of alternative transport links when a given critical link is closed.

In this section, we outline the types of static graph metrics and characteristics of the transport network, as well as our metrics for food transport network resilience. Nigeria is used as example.

4.1 Static graph metrics for the transport network

We refer to Nelson *et al.* (2020) for a discussion of theoretical background on resilience in transport networks, and for a categorization of related metrics. In this section, we focus on metrics that can be derived directly from the country graphs that result from the SWF described in the previous section, or from the same graphs but with additional semantic attributes. Such annotations may be labels for graph nodes (for example to identify cities) or graph edges (for example to provide transport mode and class), and may involve positioning information that allows for geospatial inference.

In addition, static graph metrics may be scale-dependent (thus indicating the scale or size of the graph) or scale-independent (thus capturing a characteristic that appears not to correlate with graph size). Both types of metric have their uses: scale-dependent metrics help to quantify graph size, and can be used in the interpretation of resilience metrics that are derived later (some of these display a correlation with size, which for certain purposes may need to be factored out). The scale-independent metrics allow for direct graph-to-graph (and thus country-to-country) comparisons. In the explanations that follow, the "#" symbol represents a count of entities.

Our SWF generates a view of a country's transport network that is not entirely realistic in every way. Our treatment of the network within the urban footprints is of specific note here, as it is highly artificial (yet usefully simple). As all urban footprints are treated in this way, we judge this to be an acceptable compromise. One exception however, relates specifically to the degree centrality of graph nodes that represent cities, as the simplification workflow disrupts city node degrees significantly. For this specific metric, we therefore work with a

graph reconstruction in which city nodes display a more natural edge connectivity that is worthy of reporting and analysis.

The role of static graph metrics is one of early warning: their computation is mostly simple and straightforward, and thus allows for the rapid detection of resilience concerns. Future work will elaborate on how spatial scale affects such analysis, and whether our tools can work at the subnational level as well. The preceding discussion on the different types of metric allows our static graph metrics to be categorized, and we discuss the metrics that we derive for each category in turn.

The first split-out separates topological metrics from traffic-related metrics, which require notions of distance or travel time.

Scale-dependent topology-based metrics characterize the graph size in some way. The direct base metrics are #*components*, #*nodes* and #*edges*. The more involved metrics are #*articulation nodes* and #*maximal independent sets*. An articulation node is a node whose removal would increase the number of graph components. An independent set consists of nodes that are not pairwise direct neighbours of each other in the graph; it is maximal if more nodes cannot be added to it while keeping it independent. Both metrics will generally correlate positively with graph size, and as such are relatively less useful.

The **scale-independent topology-based metrics** that we derive are *average degree*; *alpha, beta, and gamma indices*; *assortativity*; and *articulation ratio*. All of these are measures of connectivity, and the first four are strongly intercorrelated as they express some ratio between #nodes and #edges. Assortativity is a metric that expresses to what extent degrees of neighbouring nodes correlate; it is a Pearson correlation coefficient. Articulation ratio is the ratio between #*articulation nodes* and #*nodes*, and can thus be interpreted as a node vulnerability index.

There are also two **scale-independent centrality metrics** that we consider: *degree centrality* focuses on nodes and their degree as a measure of importance, and *average absolute degree centrality* is therefore a proxy for connectivity. *Betweenness centrality* also focuses on nodes and determines how often the node is on the shortest route between two other nodes. It can be seen as a proxy for sensitivity to the risk of node failure. We determine *average absolute betweenness centrality* over all nodes in the graph. This is the most computationally intensive metric that we apply.

The **scale-dependent semantic metrics** are #*cities* and #*trade stations*; we may include more when looking at different transport modes (which we have not included here).

Various **scale-independent semantic metrics** can be identified. One such metric is *local sustenance percentage*, which derives from the food balance situation per catchment and crop that we take as the starting point for the transport modelling: we know local production and food available for consumption; we can therefore determine how much food (per crop in tonnes) can maximally stay and will minimally need to enter or leave the catchment. In this context, trade stations can be seen as catchments where no food stays local. A summation over catchments and crops can provide the total volumes for food tonnage that stays local and for that which is transported, leading to the country's local sustenance percentage. Other metrics can be derived from such base data.

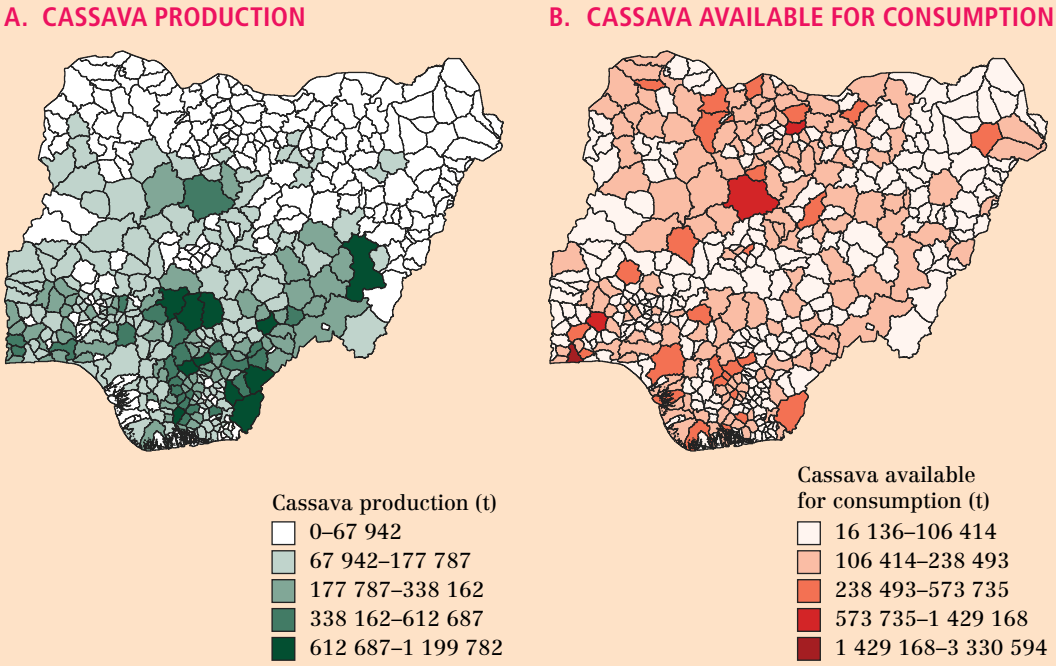
The **scale-dependent traffic-related metrics** are *diameter*, *area size* and *pi index*. Diameter is defined as the maximum of the shortest travel time required to reach one city from another. It is thus a travel time, and a one-dimensional size characteristic. Area size is the spatial size of a concave hull around all the city nodes of a given country, which uses a fixed tightness setting. It is expressed in km2 and is thus a two-dimensional size characteristic. Pi index is the sum of travel times over all edges divided by diameter; it is thus a one-dimensional size characteristic, and can be used as a proxy for network density.

The **scale-independent traffic-related metrics** are the *eta-index* and *detour index*. The first is best interpreted as the average edge length over all edges, measured in travel time. The second relates to routes between city pairs and is the ratio between actual travel time and fictitious travel time over a straight primary road edge. This means that 1.0 is as fast as such a virtual road would be, and below 1.0 is even faster.

4.2 Proximity-based resilience of the food transport network

The first type of food transport network resilience we consider is related to the food matrix – that is, the way food is expected to be distributed between the areas where it is produced and the areas where it is available for consumption. We call this proximity-based resilience. The OD food matrix estimates the tonnage of crops between origin zones (where crops are produced) and destination zones (where crops are available for consumption). The zones, which are catchment areas around all cities, are shown in Figure 9, using the example of Nigeria.

FIGURE 9 Cassava production and available for consumption, per catchment, in Nigeria



Source: UN (2020) modified by the authors.

The total attraction in the model is based on the total available for consumption and export of crops (in tonnage). The total available for consumption is equal to the total production plus import minus export, and is distributed proportionally to the population of each zone. The import and export are distributed over the trade stations according to the trade model. We distinguish between specific crops, as each crop may have different transport characteristics. First, the distribution of the production can vary widely for different crops. Second, the distribution functions can also be different. For example, highly perishable foods such as fruits may be transported over shorter distances than non-perishable foods like rice. The distribution function describes the propensity to transport food between

A and B, irrespective of production and attraction. We use an exponential decay function as a distribution function, as follows:

$$f(c) = e^{-\beta c}$$

with c as the travel cost between origin and destination, and β as the slope of the distribution function. We use travel time as the best available option to express cost of transport (given that "time is money"). The coefficient β describes how fast the attractiveness drops with travel time (cost). In principle, it is most economic to consume food where it is produced. However, market mechanisms, consumer preferences and various other factors can all contribute to a shallower distribution function (with a relatively low value for β), in which longer travel times are still relatively attractive even when there is attraction nearby (that is, people may travel further away even when there are closer locations). We use an internal travel time per zone of 20 minutes, which implies that all food transported in that time period is consumed within the zone.

Irrespective of the value of β, systems will be more resilient when there is a balance between production and attraction. If for each zone, production and demand are the same, then crops do not need to travel for long periods (even if that will happen due to factors that lead to a low value for β). In that case, systems are optimally resilient against network disturbances, because the possibility to consume the crops locally is always present (in theory). In contrast, if crops are produced in one part of the country and are consumed in another part, they will always need to be transported. This will result in longer trip durations, even if β is high (and even if there is no propensity to transport the crops on journeys with such long durations). In that case, systems are less resilient and instead more vulnerable, for example if for some reason the network connections between production and consumption areas break down.

In other words, when the average travel time is much longer than in the optimal situation (that is, when production and attraction are perfectly in balance), systems are not very resilient. The indicator for proximity-based resilience is the ratio between the average trip duration for the optimal situation and the actual situation. Table 4 shows the results for ten crops in Nigeria.

TABLE 4 Results for 10 crops in Nigeria using an internal transport time of 20 minutes

Crop	Transport time (min.)
Cassava	167
Fruits	189
Maize	112
Plantains	253
Pulses	163
Rice	111
Sorghum	219
Vegetables	113
Wheat	518
Yams	147

TABLE 4 (cont.) Results for 10 crops in Nigeria using an internal transport time of 20 minutes

Crop	Transport time (min.)
Average	155
Best case	28
Proximity-based resilience	0.18

Notes: The table presents the results after 50 Furness iterations, using an internal transport time (within zone) of 20 minutes. For the optimal situation, production is equal to the attraction for each zone.

Source: Authors' own elaboration.

4.3 Route-based resilience of the food transport network

Relative detour cost

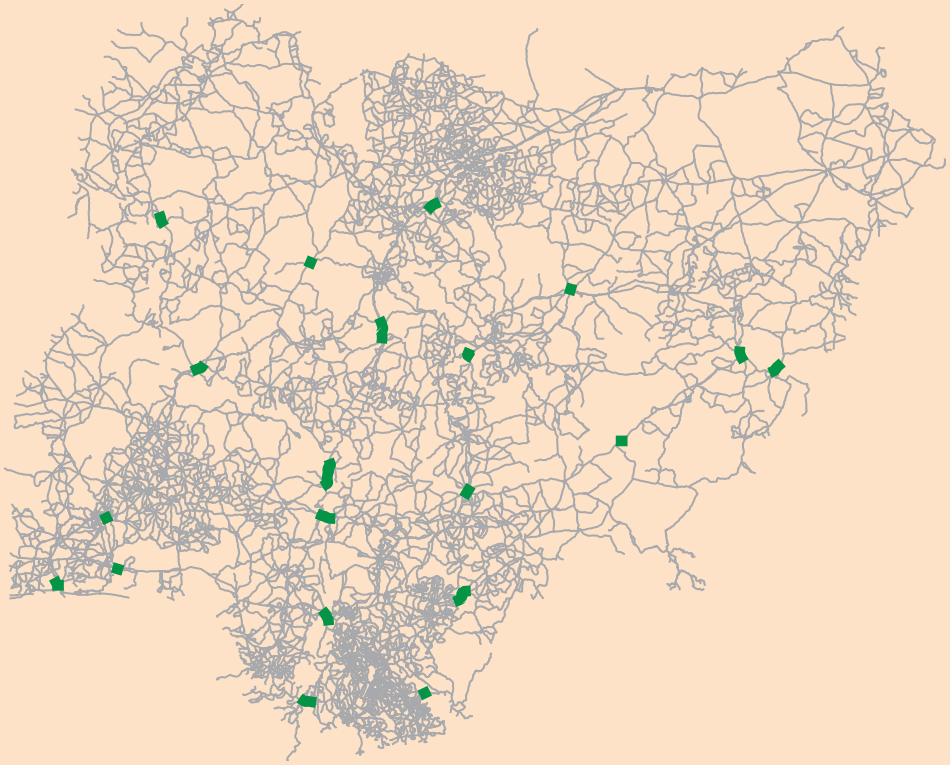
The second type of food transport network resilience is related to the effect of disruptions in the transport network on the movement of goods. Often, sections of transport networks are closed off, for example due to natural disasters, accidents or political instability. To evaluate the effect of such types of closure, we introduce disruptions on particular transport links and evaluate their effect. We outline *three route-based resilience metrics* below.

After the OD food matrix is estimated for a given β, the food volumes can be assigned to the network links. For the assignment, we assume that food is transported over the paths of shortest travel time. This assignation to links gives us *link intensity*: the total food volume [tonnes] that is/will be transported over the link, as contributed by the various shortest paths that make use of the link. To estimate the resilience, we select the 20 links of highest intensity, and we calculate the extra tonnage-minutes resulting from the closure of one such link. To select the high-intensity links, we perform the following procedure:

1. For each link with traffic, we close the link and then calculate the travel time (cost) required on the alternative route between start and end nodes of the closed link.
2. We multiply this travel time with the link's intensity. The higher the resulting value, the higher the impact of the link removal. We then select the links with the highest impact values. We cannot simply select the 20 links with highest intensity, because this may result in selecting a number of links along the most important route, which would simulate a similar closure multiple times. Therefore, we select all OD relations over the link with the highest impact value.
3. We assign again, but disregard transports over selected OD relations from the previous iteration(s).
4. We iterate back to the first step and repeat this procedure until 20 links are selected.

This route-based resilience metric, which we call the **relative detour cost**, is the relative difference (in percent) between the tonnage-minutes of the normal situation and the situation with the 20 high-intensity links removed. Figure 10 shows the high-intensity links for Nigeria. As the figure illustrates, these links correspond quite well to expectations, in that they are predominantly located in the middle of the country, serving the north–south routes that are important for food transport. The relative detour cost is 23.1 percent, or 1.8 percent on a total country basis.

FIGURE 10 Selected critical links based on link intensity times link distance, Nigeria



Notes: The selected critical links are shown in green.

Source: UN (2020) modified by the authors.

Alternative route availability

When estimating network-based resilience with relative detour cost, there is one important issue regarding step 1 of the procedure as described previously. We can only select a link for which there is an alternative route between its start and end node, and consequently between the origins and destinations of the OD relations using that link. If there is no alternative route, the link cannot be considered. For most countries this is not really an issue, because such links are rare and only occur in the periphery of the network. However, for some countries with a sparse network, this is an issue. A good example is Somalia, as shown in Figure 11. The red link in the figure is a critical link between the north and south of the country, and serves as the only connection in our simplified network. If this link is closed, there are no alternative routes between the north and south. This link is therefore not considered in the resilience of the network based on relative detour cost, because there is no detour or alternative route for it.

To include this type of resilience, we calculate the total tonnage (T) for all links, as well as the tonnage for any links for which there is an alternative route when the given link is closed. The ratio between the two gives the second route-based resilience metric, called **alternative route availability**.[2] For most countries this value lies above 0.80, indicating

[2] Due to network size and complexity, our alternative route availability calculation for India used a different workflow, based on multiple subsamples of the transport network.

that alternative routes are readily available. For Somalia however, it is around 0.31. In other words, 69 percent of the country's tonnage is transported over links for which there are no alternative routes. This is an indication that this network is not resilient at all.

FIGURE 11 Sparse network for Somalia

Notes: There are no alternative routes available for the critical link, as shown in red.

Source: UN (2020) modified by the authors.

People affected

The closure of high-intensity links (in the relative detour cost metric) will have an impact on the population that consumes that travel-delayed food. We estimate the impact on the population purely in terms of **people affected** by delay, by considering each closed link in turn and determining which (OD) pairs of nodes (or catchments) are involved when the link is removed, and whether they are O or D catchments (or both, if goods are moving both ways across the link). Then, without double counting, we compute the average number of people in the D catchments (in other words, the consumers) affected across all link closures, based on the known population per catchment. This metric does not consider the quantity of food that is delayed: in other words the count of people affected in a given catchment is the same regardless of the quantity.

4.4 Correction for country size

In the next section, we compare the resilience metrics of different countries. When we do this for proximity-based resilience and relative detour cost, we find a clear correlation between resilience metrics and country size. While this correlation can have various causes, it can be divided into the following two components:

1. "Artificial scale effects", for which a correction is justified. For example, the average trip length, and hence the proximity-based resilience, of a small country is constrained by the physical size of the country. Similarly, detours will on average be shorter for small countries, but relative detours will on average be longer. This is simply because in general, there tend to be more alternatives for longer trips.
2. "Real effects", which should not be corrected for. Good examples include population and network density. In general, when the population density increases, people will live closer to the production locations of crops, and when the network density increases, detours will be shorter. They therefore have a real positive effect on proximity-based and route-based resilience respectively.

The most straightforward way to correct is to use the physical size – for example, the square root of the country area – as a scale factor. By multiplying the proximity-based resilience and relative detour cost metric with this scale factor, artificial differences in both these metrics may be levelled out between countries of different sizes. However, hidden factors such as population density may also correlate with country size. For example, small countries typically have higher population densities, which have a real positive effect on the proximity-based resilience metric. We therefore used the square root of the total tonnage of transported crops as a scale factor in our correction. We multiplied the **proximity-based resilience** metric and the **relative detour cost** metric with this scale factor (and divided by the average scale factor to keep the same order of magnitude in both metrics). In doing so, we tried to correct for the artificial effects (for the same population density, you actually correct for country size), while at the same time avoiding overcorrection due to other hidden factors in the correlation between our resilience metrics and country size. Only when we compare all countries in the following section will we use the corrected **proximity-based resilience** metric and corrected **relative detour cost** metric to take the artificial scaling effects into account. This correction is not applicable to the **route redundancy** and **people affected** metrics.

5 Results and discussion

KEY MESSAGES

- Where food is transported more locally, the network will not be overly affected by disturbances, independent of its quality. In countries where food is moved over longer distances but with a dense network (for instance, the United States of America) there will also be relatively little effect. But in countries with networks that are both low in food transport efficiency and density (mostly low-income countries), there will be a significant effect.
- High food transport efficiency is found in high-income countries (for example in Europe), as well as in densely populated countries like China, India, Nigeria and Pakistan.
- In sub-Saharan Africa, some countries produce food relatively close to where it is consumed, and thus have higher proximity-based resilience.
- A simulation of the impact of localized 1-in-10-year flooding events in Mozambique, Nigeria and Pakistan reveals the loss of network connectivity when links become impassable, potentially affecting millions of people.
- A vast amount of spatial and statistical data needs to be brought together and harmonized before such models of food transport network resilience can provide meaningful results.

We follow the same order as in the previous section, first presenting the static graph metrics/characteristics of the transport network, followed by the metrics for food transport network resilience.

5.1 Static graph metrics

Primarily from OSM base data, the simplification workflow derives simplified graphs per country, thereby allowing various comparisons between countries. We present here an analysis of the characteristics of graphs by country, and subsequently share some early findings. This work comprises approximately 90 countries. These statistics allow for research into various correspondences, such as how many people in the population must be served per network node.

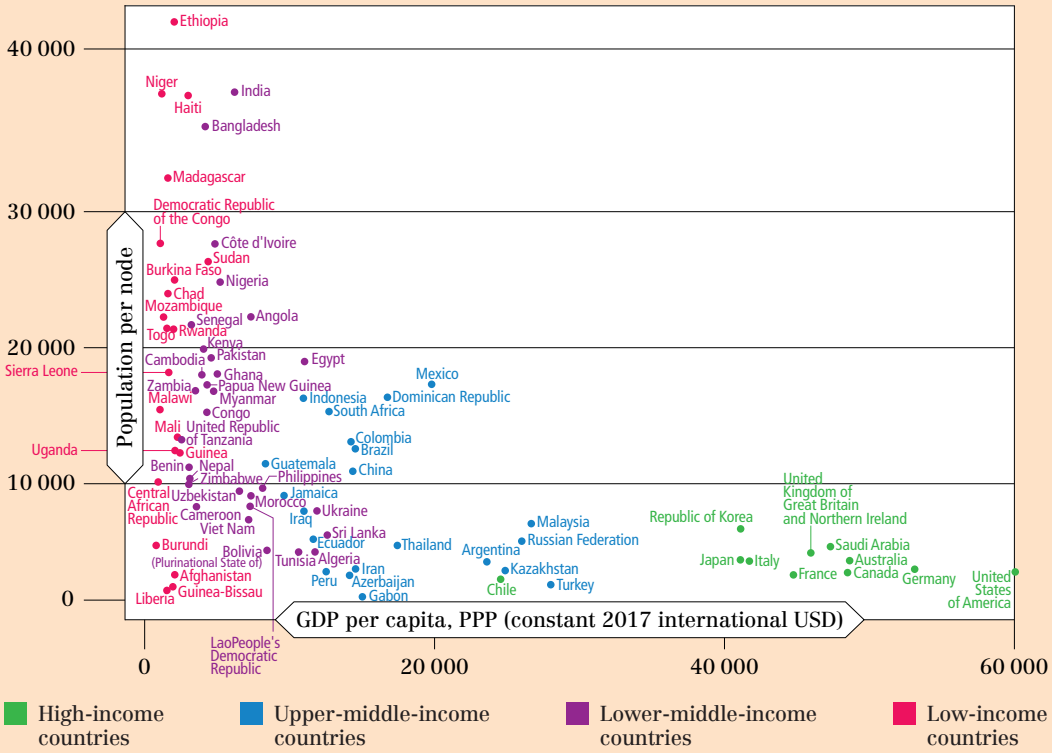
The direct *scale-dependent topology-based metrics* typically consist of basic counts of the graph's components, nodes and edges, as well as more involved counts of articulation points and maximal independent sets. An articulation node is a node whose removal will break the graph; it can thus be used to assess *node vulnerability*. A maximal node set is a set in which no two nodes are direct neighbours – the lower the maximal node set count, the denser the graph is – it can thus be used to assess *connection density*. These numbers do not seem telling in themselves, but when normalized by some appropriate size gauge for the graph or country, they may help to display more useful patterns. Some of these

relative metrics are classical graph metrics (such as average degree), as discussed below. The articulation node ratio simply determines how many nodes there are per articulation point. A high ratio means a low node vulnerability.

A word of warning in interpreting results: Due in large part to data complexity, we did not include tertiary roads in the graphs of 16 high-income countries: Australia, Brazil, Canada, China, Germany, France, Italy, Japan, the Republic of Korea, Mexico, the Russian Federation, Saudi Arabia, South Africa, Ukraine, the United Kingdom of Great Britain and Northern Ireland, and the United States of America. This has affected various graph counts and derived metrics, and means that results cannot readily be compared with other countries. At the other end, for three low-income countries (Gabon, Guinea-Bissau and Liberia), we included unclassified roads and tracks to ensure node connectivity and alternative routing.

Any of the above metrics can be set off against population size or level of economy. In Figure 12, we display gross domestic product (GDP) per capita against the number of people served per network node. As the figure illustrates, while some low-income countries such as Burundi do well, it is typically high-income countries that display low headcounts per node. (Incidentally, many of these would have had more nodes if all their roads had been included, leading to even lower scores.)

FIGURE 12 Population per network node vs gross domestic product



Notes: GDP per capita expressed in international dollars at purchasing power parity (PPP). There is no GDP data for the Democratic People's Republic of Korea, South Sudan, the Syrian Arab Republic, the Bolivarian Republic of Venezuela and Yemen.

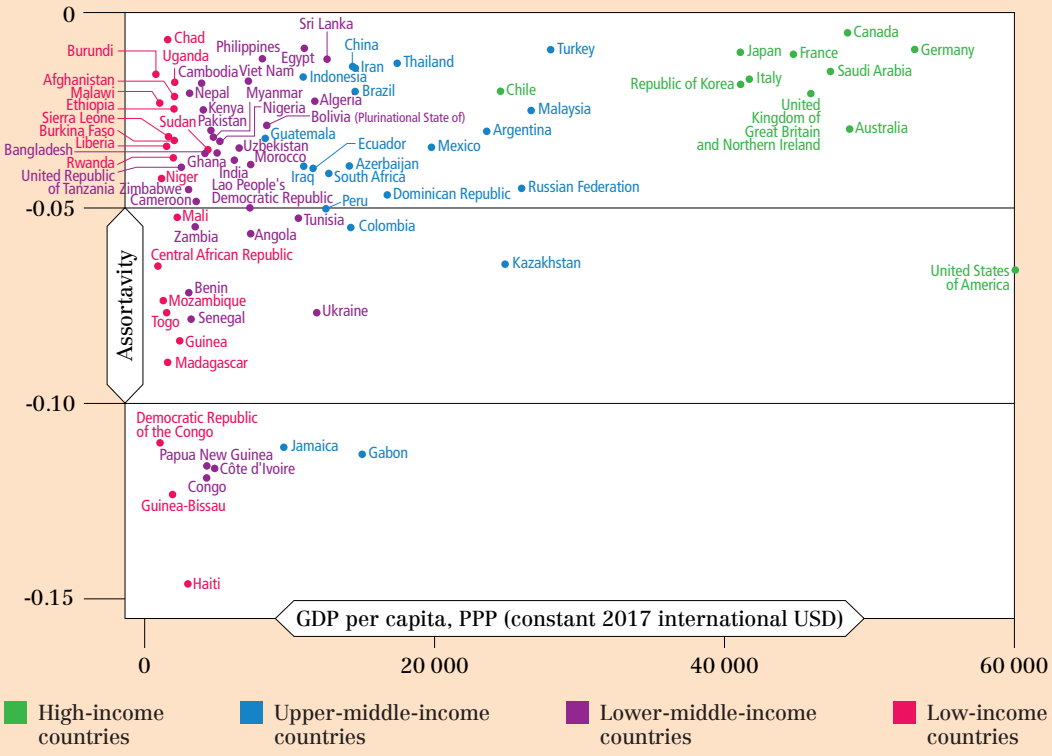
Source: Authors' own elaboration.

The *scale-independent topology-based metrics* are expressions of connectivity, and many are highly correlated. High connectivity is an advantage, as it offers alternative routing when parts of the network infrastructure collapse. For assortativity (see Figure 13), we must first

understand that our graph simplification removes as many two-degree nodes as possible, thereby pushing down the scores. Removal of such nodes is appropriate because a high percentage of these nodes carry no intrinsic meaning for the network: they do not represent a feature of some sort. Removal of tertiary roads has likely had a positive impact on some of the assortativity scores for high-income countries. As seen in Figure 13, it appears that scores between -0.025 and 0.0 characterize mature transport networks in our study, and that values below -0.025 are indicative of less well-developed networks. Lower assortativity means there are more instances in which nodes of different degree values are neighbours; this happens especially when higher priority roads more frequently intersect with lower priority roads, as this causes nodes with more alternating degrees. This may be characteristic of a transport network's "development stage". Figure 13 bears a strong resemblance to similar figures for connectivity metrics (which we do not include here).

Among the static graph metrics, our two *scale-independent centrality metrics* were the hardest to compute. *Absolute betweenness*, in particular, presented performance issues for the largest of countries, and we have not yet produced scores for China, India and the United States of America. Regardless, the charts for GDP versus degree centrality and log(*betweenness centrality*) show an almost identical pattern to Figure 13, with high variability in scores between low- and lower-middle-income countries, and more congruence among higher-income countries. Average absolute degree varies between 1 and 9 for the former group, and between 7 and 9.5 for the latter. Average absolute betweenness varies between 5 000 and 10 000 000 for the former, and between 100 000 and 50 000 000 for the latter.

FIGURE 13 Assortativity per network vs gross domestic product

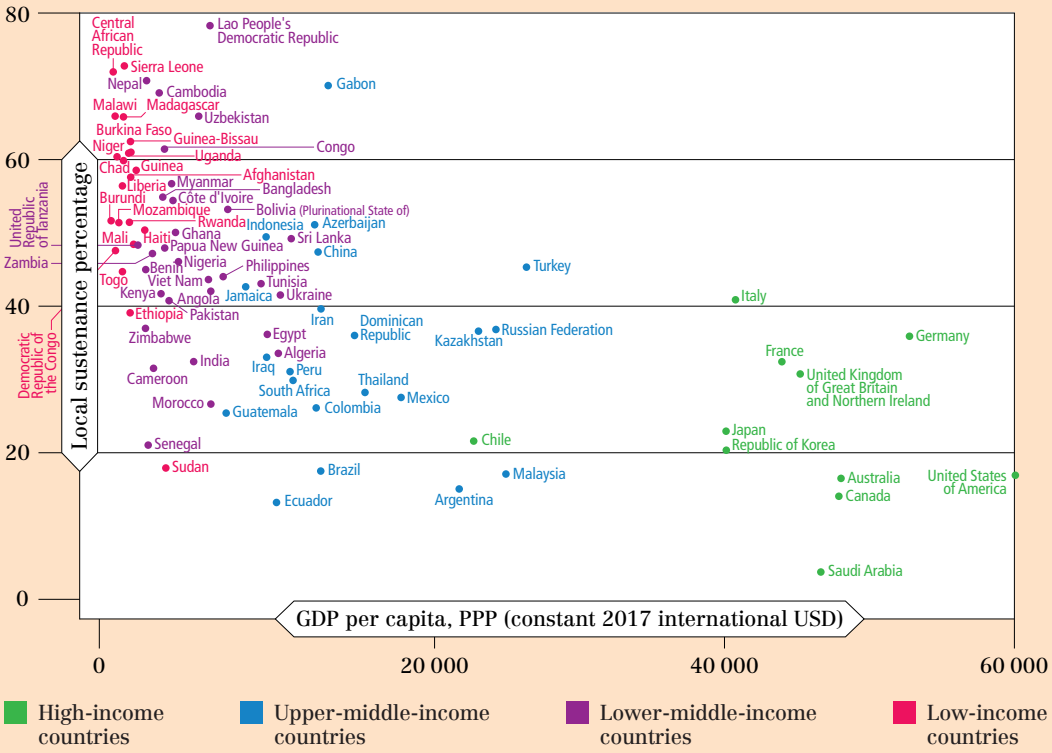


Notes: GDP per capita expressed in international dollars at purchasing power parity (PPP). There is no GDP data for the Democratic People's Republic of Korea, South Sudan, the Syrian Arab Republic, the Bolivarian Republic of Venezuela and Yemen.

Source: Authors' own elaboration.

Scale-dependent semantic metrics have proven to be a determining data complexity factor in some of our graph computations. One *scale-independent semantic metric* that we believe is fundamental is local food sustenance – the maximum percentage of food volume that is produced and consumed in the same catchment. In Figure 14, local food sustenance is inversely correlated with GDP per capita, as agriculture in many countries is not a high-opportunity industry. There is high variability between lower-income countries however: those with reasonably functioning food production systems show high local sustenance, while those with failing systems show lower scores. High local sustenance is one way to lower vulnerability against infrastructural failures. Higher-income countries can more easily afford to bring more food into a city.

FIGURE 14 Local food sustenance percentage vs gross domestic product

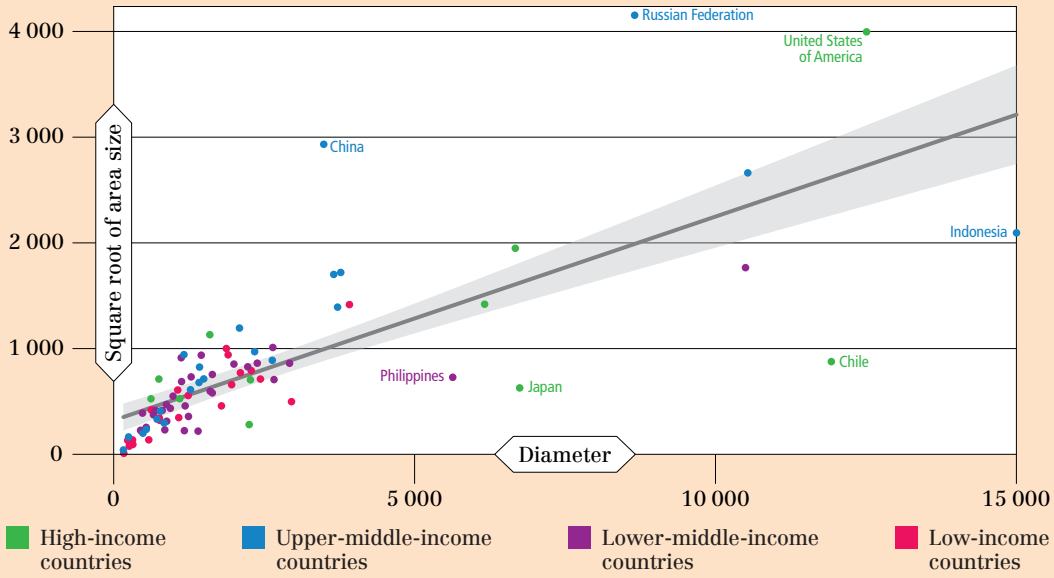


Notes: GDP per capita expressed in international dollars at purchasing power parity (PPP).
There is no GDP data for the Democratic People's Republic of Korea, South Sudan, the Syrian Arab Republic, the Bolivarian Republic of Venezuela and Yemen.

Source: Authors' own elaboration.

Our *scale-dependent traffic-related metrics* provide a notion of graph or country size, against which other metrics that reflect size dependency can be normalized; countries freed from size dependency thus become more comparable. We conduct such comparisons in the next section. Here we display the relationship between a country's diameter and the square root of its area size. Both can be used to gauge size. Figure 15 shows how they are related. A linear fitted trendline provides an R2 of 0.56; countries above the line are rounded in shape, whereas those below are elongated. China, the Russian Federation and the United States of America are the most prominent examples of the former, while Chile, Indonesia, Japan, and the Philippines are examples of the latter. Some of our other metrics often feature eccentric scores for such countries.

FIGURE 15 Country area size (square rooted) vs diameter

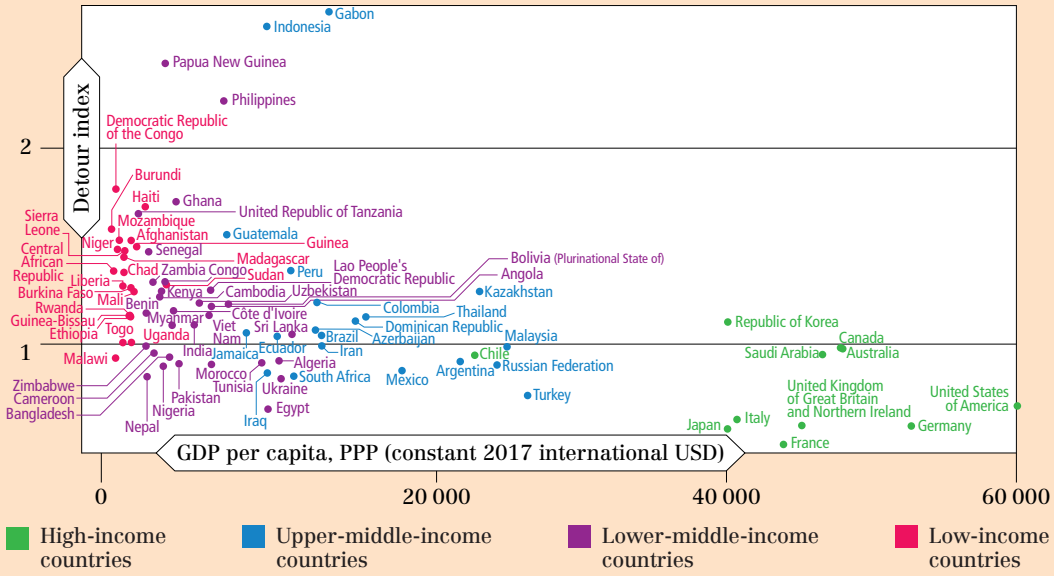


Notes: Only countries which are significantly distant from the trendline were identified.

Source: Authors' own elaboration.

Of the *scale-independent traffic-related metrics*, we want to draw attention to the detour index. It is a simple metric that characterizes the speed with which one can travel between cities in the country, and thus it is a transport quality characteristic. Where its value is below 1.0, travel is fast; where it is above 1.5, travel is slow. One can expect higher-income countries to have lower detour indices, and this is corroborated in Figure 16.

FIGURE 16 Detour index vs gross domestic product



Notes: GDP per capita expressed in international dollars at purchasing power parity (PPP). There is no GDP data for the Democratic People's Republic of Korea, South Sudan, the Syrian Arab Republic, the Bolivarian Republic of Venezuela and Yemen. Detour index could not be computed for China due to network size and complexity.

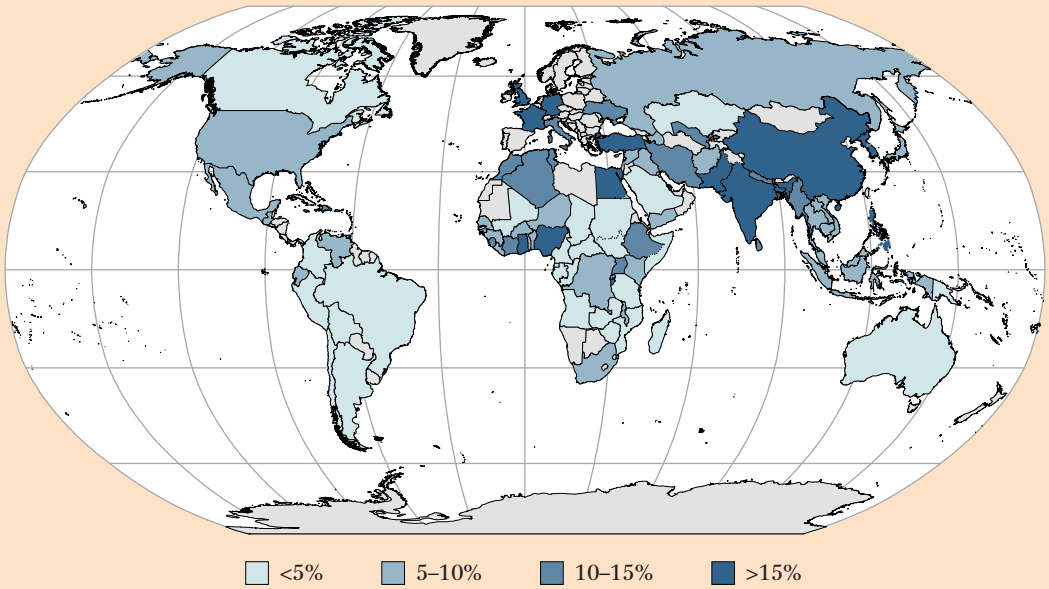
Source: Authors' own elaboration.

5.2 Resilience metrics

Proximity-based resilience

The proximity-based resilience metric is the ratio between the average transport time in the ideal situation (when all food that is produced in a catchment is also available to be consumed in the same catchment) and the average transport time in the actual situation (when food produced in a catchment is distributed to all cities). If the average transport time for the actual situation in a country is relatively short, this indicates that food is available for consumption relatively close to production locations. (The ideal situation being one in which transport times are at their lowest because all food is supplied close to where it is produced.) The results are also a good indication of the food transport systems' overall sustainability for the crops involved. The metric is mapped in Figure 17, and plotted against average transport time in Figure 18.

FIGURE 17 Proximity-based resilience

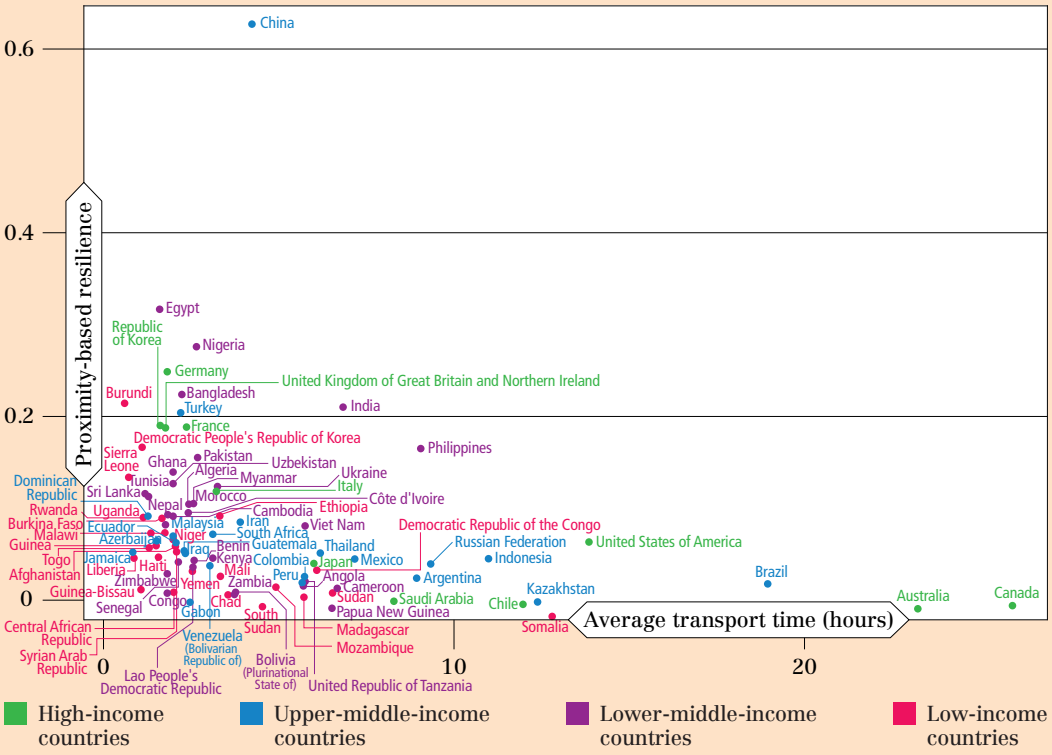


Notes: Darker colours reflect higher scores; i.e. greater resilience. For special country/region-specific disclaimers see notes of Figure 1.

Source: UN (2020) modified by the authors.

Figure 17 shows that high food transport efficiencies can be found in the high-income European countries, but also in densely populated lower-income countries like Bangladesh, China, India, Nigeria and Pakistan. The ease of transporting food from production locations to consumers not only depends on the quality of the network, but also on "real" (or Euclidean) distances between production and consumption. For example, Egypt scores well due to high population density in agricultural areas that are situated along narrow, well-connected corridors by the coast and the Nile River. And despite their lower quality networks, the relatively high resilience scores of some sub-Saharan countries reflect that in the current situation, food is produced relatively close to where there is demand. It is also noteworthy that these distances increase in large countries with relatively low population (density) numbers, such as Brazil, Canada, and Australia.

FIGURE 18 Proximity-based resilience vs average transport time



Source: Authors' own elaboration.

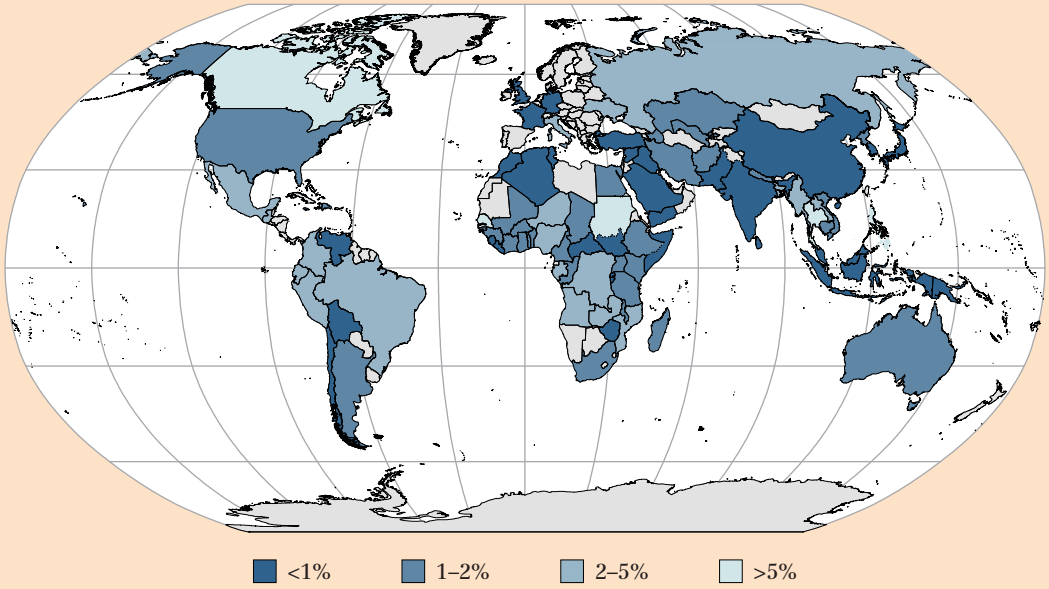
Figure 18 shows a clear (negative exponential) pattern, where countries that have shorter average transport times have higher proximity-based resilience. This can be explained as follows: Proximity-based resilience represents the ratio between the average transport time in the ideal situation (when all food that is produced in a catchment is also supplied in the same catchment) and average transport time in the actual situation (when food produced in a catchment is distributed to all cities). If the average transport time for the actual situation in a country is relatively short (as in the examples of France, Nepal, Tunisia, Ukraine, and the United Kingdom of Great Britain and Northern Ireland), this indicates that food is available for consumption relatively close to production locations. This situation is more comparable to the ideal situation, in which transport times are at their lowest since all food is supplied close to where it is produced. The results are also a good indication of the food transport systems' overall sustainability for the crops involved. In Tunisia and the United Kingdom of Great Britain and Northern Ireland for example, food transport systems are organized such that food travels over relatively short distances, resulting in much lower costs and negative externalities such as energy consumption and pollution. Unsurprisingly, this is not the case for countries like Brazil, Canada and the United States of America. These countries are of course much larger than those previously mentioned, but they have also organized their food production and distribution systems differently. Other large countries such as China and Nigeria have organized their food production and distribution in a much more resilient and sustainable manner. China's position here as an outlier is also influenced by our scaling against $\sqrt{tonnage}$ which has amplified its score. We return to this in the following section, with respect to the challenge of developing such metrics across diverse geographies, economies, transport network topologies and qualities.

Network-based resilience

For relative detour cost, and as shown in Figure 19, higher values mean that a longer detour time is incurred if routes are removed from the network due to disruption. The alternative route availability in Figure 20 is a measure of route alternatives or route options. A score of 100 percent means there are always alternative routes between OD pairs. Figure 21 shows people affected by the closure of high-intensity links (in the relative detour cost metric). For all three figures, and as with Figure 17, darker colours in the map indicate better scores on the resilience metric.

As may be expected, small, high-income countries have more resilient networks than lower-income countries, but some lower-income countries also score consistently well, such as Algeria, the Plurinational State of Bolivia, Liberia and Sri Lanka. Relative detour costs are low where networks are highly developed, generally indicating that there are alternative routes available. However, some large and sparsely populated countries like Canada score low on the relative detour metric but quite well on the alternative route availability metric. Canada's network is in general less dense, implying relatively large detours when an important link is closed. On the other hand, some elongated countries such as Chile score relatively well in network detour costs but poorly in alternative route availability. In such countries, important routes may have limited alternatives, especially when confined by mountains. If such routes have no alternatives, they contribute negatively to alternative route availability, but they do not contribute to the detour indicator (because there is no detour possible). These two resilience indicators can be seen as complementary. If a country scores poorly on one of them, its network resilience is considered to be low. The relation between both network resilience measures is further illustrated in the scatter plot in Figure 22. For clarity, we invert the alternative route availability percentage on the y-axis by subtracting it from 100, so it shows "no alternative routes (%)". To clearly distinguish between countries, we use a log10 scale for both axes. Countries in the top right score poorly in both, and those in the bottom left score well in both. Most countries are between these two extremes.

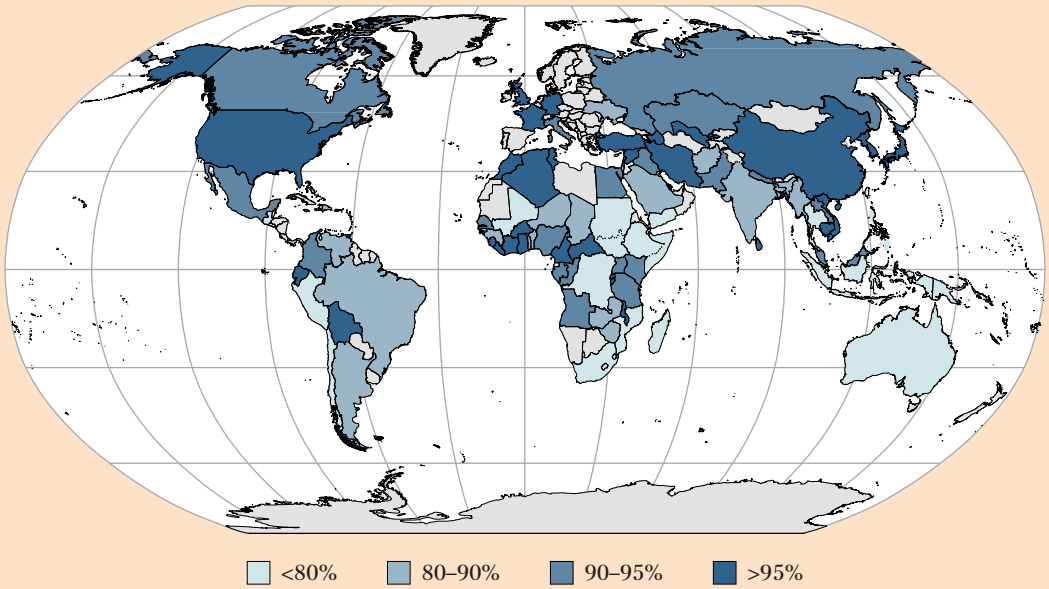
FIGURE 19 Relative detour cost



Notes: Darker colours reflect lower scores; i.e. greater resilience. For special country/region-specific disclaimers see notes of Figure 1.

Source: UN (2020) modified by the authors.

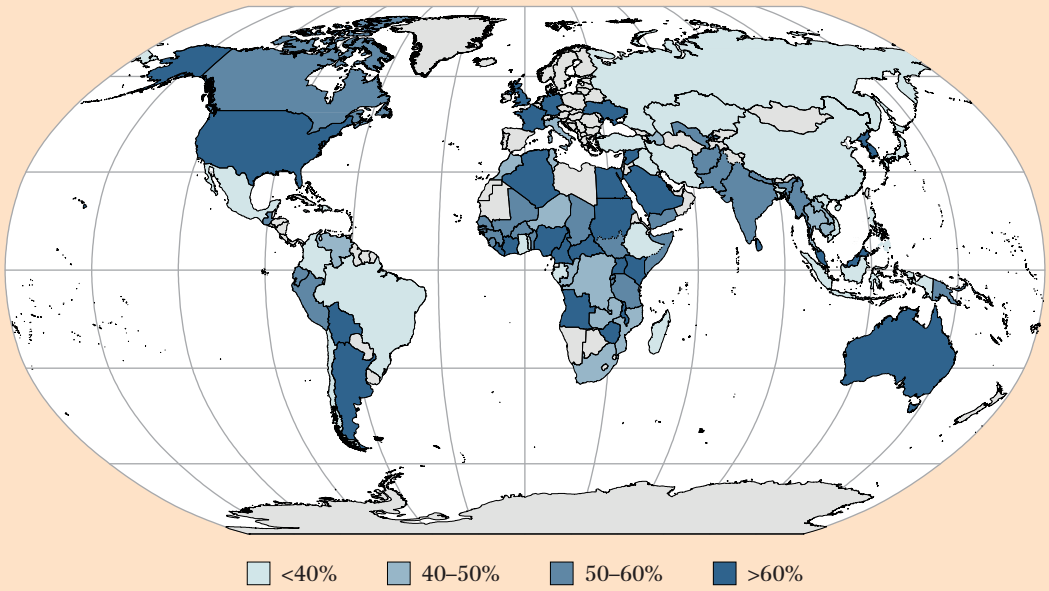
FIGURE 20 Alternative route availability



Notes: Darker colours reflect higher scores; i.e. greater resilience. For special country/region-specific disclaimers see notes of Figure 1.

Source: UN (2020) modified by the authors.

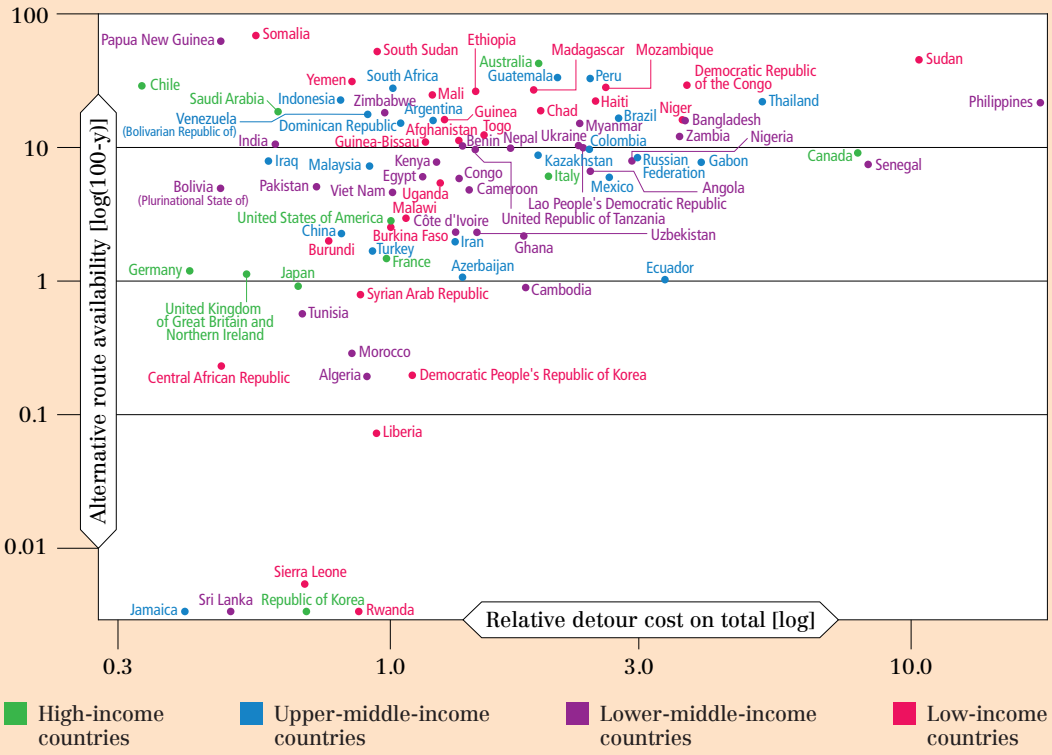
FIGURE 21 People affected



Notes: Darker colours reflect lower scores; i.e. greater resilience. For special country/region-specific disclaimers see notes of Figure 1.

Source: UN (2020) modified by the authors.

FIGURE 22 Alternative route availability vs relative detour cost



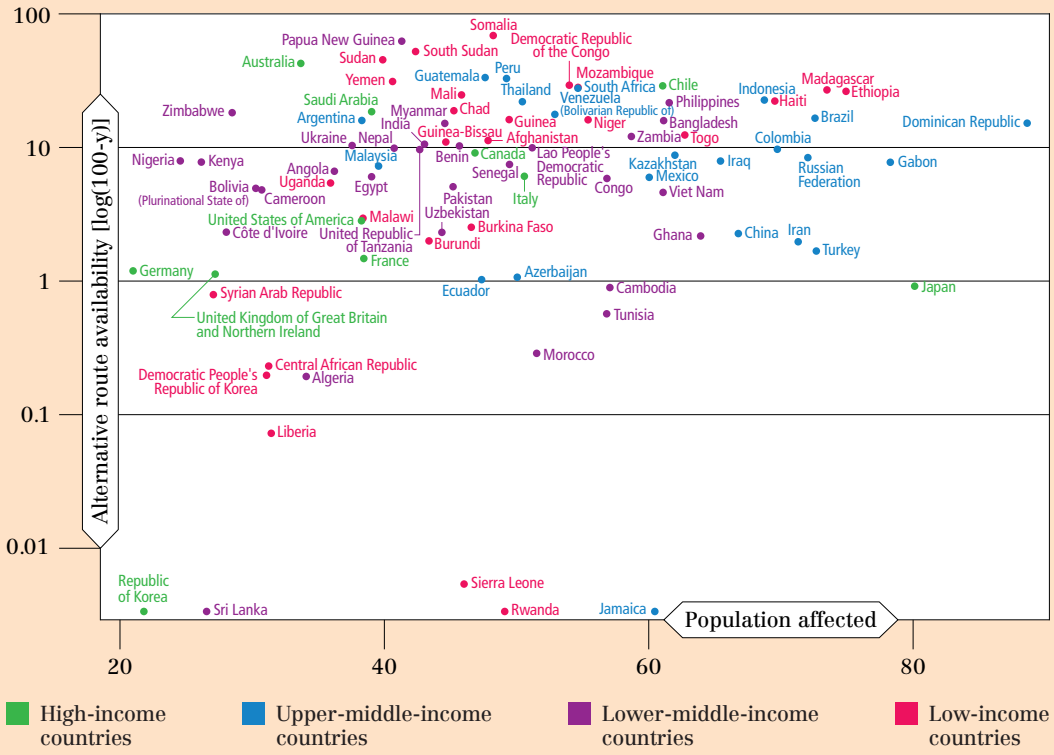
Notes: Darker colours reflect lower scores; i.e. greater resilience.

Source: Authors' own elaboration.

People affected adds another dimension to network resilience by considering the proportion of the population that is potentially impacted when high-intensity links become unavailable. This means that a country may have a resilient network – for example in the case of China – but any detours or delays can still impact a significant proportion of the population served by food transport routes that include those links. Again, small, dense, high-income countries generally score well on this metric. In Figure 23 we see that upper-middle-income countries generally have a higher percentage of people affected by the loss of high-intensity links than any other income group. This may suggest a difficult transition in emerging economies, where the food transport network struggles to meet demands in the event of disruptions, resulting in higher impacts. There seems to be a trend of lower percentages on either side of this upper-middle-income group: In lower-income countries, impacts are generally low because of higher local food sufficiency, while in higher-income countries, a well-developed network is able to continue providing sufficient food to most people despite disruptions.

Overall, food transport efficiency is not strongly related with the income level of a country. As mentioned before, population density and more generally, spatial distances between crop production sites and consumers are as important if not more important. Countries with low network resilience also tend to have low food transport efficiency. If a country has low food transport efficiency, food is transported not only locally, but also over longer distances, creating important corridor routes serving high-traffic loads. If these corridors are part of a sparse network (i.e. one that lacks many alternative routes), it will have a significant negative effect on network resilience indicators. However, if food is only transported locally, network disturbances will have a limited effect on the country as a whole, even if the network is sparse.

FIGURE 23 Alternative route availability vs people affected



Source: Authors' own elaboration.

Summary observations on results from the four resilience metrics

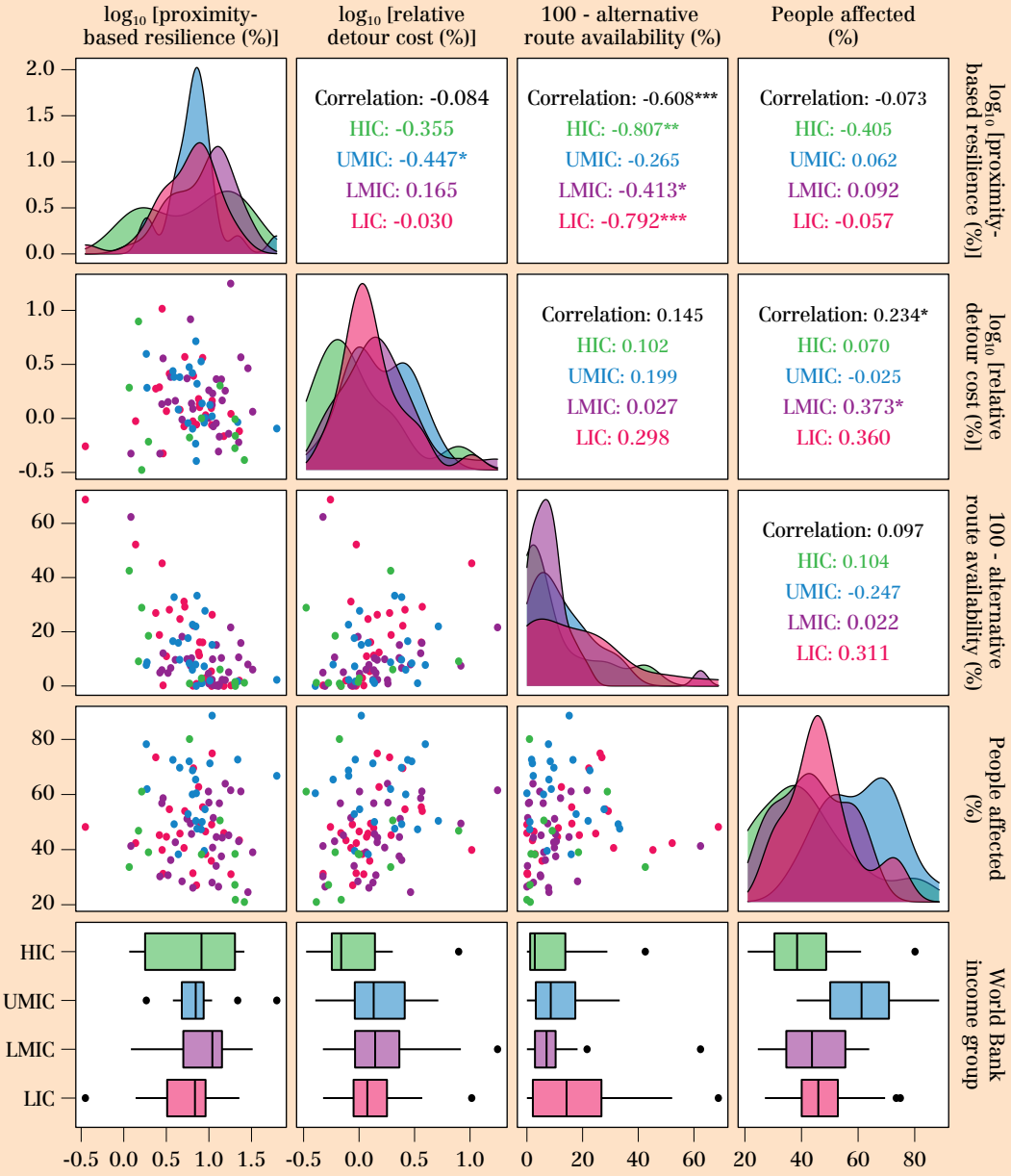
Figure 24 shows the relationships between the four food transport resilience metrics by country income group (World Bank, 2019). (Note the use of log10 transformations for some variables.)

There is little correlation between the four metrics (with the exception of a moderate negative correlation between proximity-based resilience and alternative route availability for high- and low-income countries). The four metrics capture different aspects of food transport network resilience, but do not preclude some interaction between characteristics of the network and the ways in which food is transported across it. High proximity-based resilience due to high local food sufficiency or short transport times can still lead to low levels of alternative route availability and long detours if there are few alternatives. The metrics also show no overall pattern between income group and food transport network resilience. This suggests a diversity of transport network "configurations" in terms of transport route densities, diversity in the spatial distributions of cities and trade stations, and diversity in the spatial patterns of food supply and demand. This diversity is more dependent on the area and shape of a country, as well as on the distribution of people and agricultural land and the ways in which they are connected via the transport network, than it is to country income level.

A preliminary typology of countries can be derived from a *k*-means clustering on the four (scaled) metrics. Using Gap Statistics (Tibshirani, Walther and Hastie, 2001), we determined that the optimum number of clusters in this case was three. The clusters are mapped in Figure 25; as the figure illustrates, there is some geographic grouping (for example for

Europe, North Africa and much of Asia) in cluster 1, and (in the Sahel, East and Southern Africa, Central Asia and much of Latin America) in cluster 2, whereas cluster 3 is smaller and shows no geographic pattern. The first two principal components of the clusters are visualized in Figure 26, and while these two components explain over 70 percent of the variation in the data set, the size and shape of the clusters suggest the need for caution in their interpretation. This is especially the case for cluster 1, though care must be taken with interpreting outliers on the edges of the convex hulls in all three clusters.

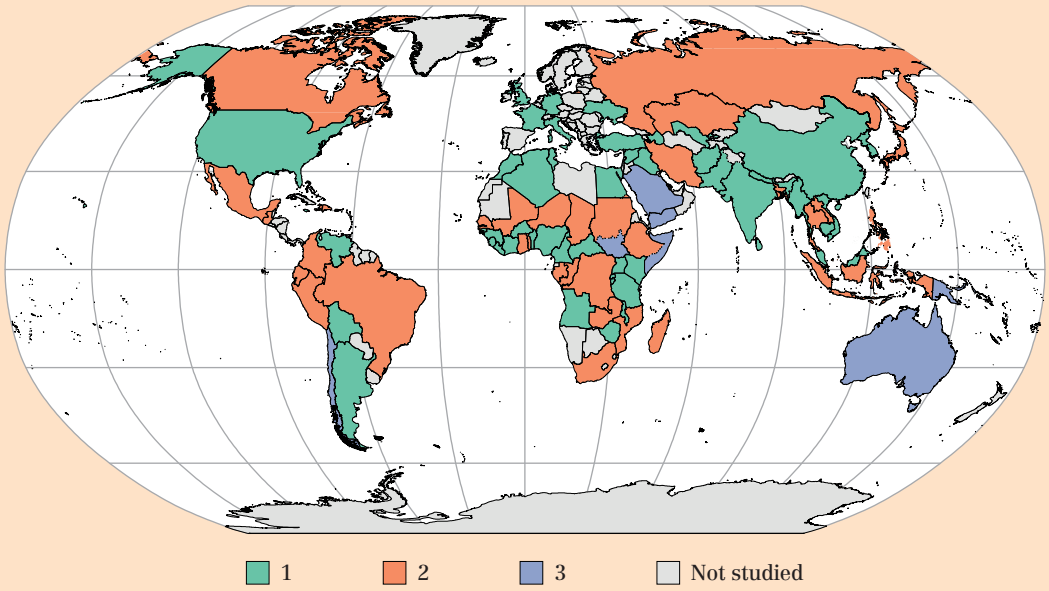
FIGURE 24 Correlation between resilience metrics (top right), scatter plots (bottom left), and their distributions (diagonal and bottom row)



Notes: * Statistically significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level. HIC = High-income countries; UMIC = Upper-middle-income countries; LMIC = Lower-middle-income countries; LIC = Low-income countries.

Source: Authors' own elaboration.

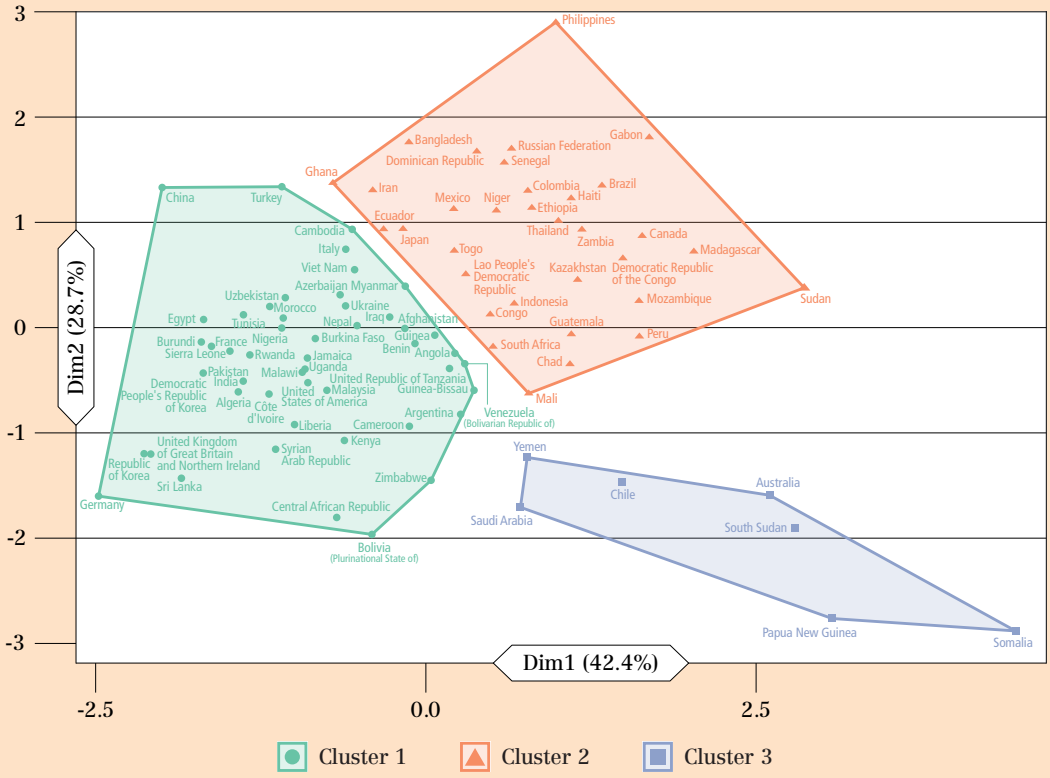
FIGURE 25 A *k*-means clustering on the four (scaled) resilience metrics



Notes: For special country/region-specific disclaimers see notes of Figure 1.

Source: UN (2020) modified by the authors.

FIGURE 26 First two principal components of the four (scaled) resilience metrics with a *k*-means clustering



Source: Authors' own elaboration.

Quartile statistics for each cluster are provided in Table 5. Countries in cluster 1 (n = 49) generally have the best performing networks in terms of proximal food availability. While there are alternative routes in case of disruptions, relative detours can be costly, with disruptions affecting 40 percent of the population (interquartile range or IQR: 31–49 percent). These countries have high food transport efficiency, and may or may not have a highly developed and dense network. As food is transported more locally, network disturbance will have a relatively limited effect, independent of the quality of the network.

Cluster 2 countries (n = 33) are where we see the largest potential impact on the population from disruptions at 60 percent (IQR: 50–70 percent), and where disruptions have a relatively high cost. That said, alternative route availability is generally good, and the proximity-based resilience lies between that of clusters 1 and 3. These countries have a lower food transport efficiency than cluster 1 and incur higher costs due to disruptions, but their networks are able to offer alternative routes to most disruptions, perhaps because they are dense enough.

A small number of diverse countries make up cluster 3 (n = 7), forming a long and narrow cluster that requires caution in interpretation. In general, this cluster represents countries with very poor proximity-based resilience, and possibly a lack of alternative routes (although the range from Somalia at 31 percent to Saudi Arabia at 81 percent should be noted). Where detours are possible, they are generally not expensive in terms of additional travel time, and disruptions affect 41 percent of the population (IQR: 40–45 percent). These countries have low food transport efficiency and have sparse transport networks that are vulnerable to any loss of critical links. The cluster contains both low- and high-income countries.

TABLE 5 Quartiles of resilience metrics per cluster and countries per cluster

Cluster	Statistics*	Proximity-based resilience	Relative detour cost	Alternative route availability	People affected	Countries per cluster, by World Bank country income group
1	min	0.03	0.40	81.8	21.0	n = 49 **LIC:** Afghanistan, Burundi, Burkina Faso, Central African Republic, Democratic People's Republic of Korea, Guinea, Guinea-Bissau, Liberia, Malawi, Rwanda, Sierra Leone, Syrian Arab Republic, Uganda; **LMIC:** Angola, Benin, Bolivia (Plurinational State of), Côte d'Ivoire, Cameroon, Algeria, Egypt, Kenya, Cambodia, India, Sri Lanka, Morocco, Myanmar, Nigeria, Nepal, Pakistan, Tunisia, United Republic of Tanzania, Ukraine, Uzbekistan, Viet Nam, Zimbabwe; **UMIC:** Argentina, Azerbaijan, China, Iraq, Jamaica, Malaysia, Turkey, Venezuela (Bolivarian Republic of); **HIC:** Germany, France, United Kingdom of Great Britain and Northern Ireland, Italy, Republic of Korea, United States of America
	lwq	0.07	0.76	92.1	31.2	
	med	**0.11**	**1.00**	**97.0**	**40.2**	
	upq	0.15	1.35	99.1	49.2	
	max	0.63	2.91	100.0	72.7	

TABLE 5 (cont.) Quartiles of resilience metrics per cluster and countries per cluster

Cluster	Statistics*	Proximity-based resilience	Relative detour cost	Alternative route availability	People affected	Countries per cluster, by World Bank country income group
2	min	0.01	0.67	54.7	39.9	n = 33 **LIC:** Democratic Republic of the Congo, Ethiopia, Haiti, Madagascar, Mali, Mozambique, Niger, Sudan, Chad, Togo; **LMIC:** Bangladesh, Congo, Ghana, Lao People's Democratic Republic, Philippines, Senegal, Zambia; **UMIC:** Brazil, Colombia, Dominican Republic, Ecuador, Gabon, Guatemala, Indonesia, Iran (Islamic Republic of), Kazakhstan, Mexico, Peru, Russian Federation, Thailand, South Africa; **HIC:** Canada, Japan
	lwq	0.03	1.51	75.3	50.4	
	med	**0.06**	**2.42**	**84.1**	**60.0**	
	upq	0.08	3.64	91.6	69.7	
	max	0.24	17.71	99.1	88.6	
3	min	0.00	0.33	31.2	33.7	n = 7 **LIC:** Somalia, South Sudan, Yemen; **LMIC:** Papua New Guinea; **HIC:** Australia, Chile, Saudi Arabia
	lwq	0.01	0.52	42.7	39.8	
	med	**0.01**	**0.61**	**57.5**	**41.3**	
	upq	0.02	0.89	70.0	45.3	
	max	0.05	1.93	81.5	61.1	

Notes: min = minimum; lwq = lower quartile; med = median; upq = upper quartile; max = maximum. HIC = high-income countries; UMIC = upper-middle-income countries; LMIC = lower-middle-income countries; LIC = low-income countries.

Source: Authors' own elaboration.

While these are interesting clusters and trends, we believe that this preliminary analysis requires more in-depth interpretation and exploration of alternative clustering methods, perhaps considering hierarchical models to acknowledge the wide range of metric scores within these preliminary clusters, and with more descriptive subgroupings. Additionally, other variables should be considered in the clustering. For example, we have noted that the four resilience metrics are influenced by the density and connectivity of the transport network, the number and spatial configuration of cities and trade stations, the spatial patterns of food supply and demand, and the area and shape of a country. All of these variables exist in our results and all are candidate variables for future characterization and grouping of countries to better classify the resilience of their domestic food transport networks.

5.3 Disruption scenarios

In this section we assess the impact of flood events on food transport network resilience in three countries: Nigeria, Mozambique and Pakistan. Flooding, whether from flash floods or from longer-term stagnant flooding , reduces the connectivity of any transport network, impacting the movement of people, goods and societal functioning in general (Pregnolato *et al.*, 2017). Direct and immediate impacts may occur when roads become impassable, when bridges become unusable (if they are blocked/damaged or washed away), or when lives are lost. If there is little material damage, then once the flood subsides the direct impacts on the infrastructure may be short-lived. But there can also be long-lasting effects if required repairs are slow to occur, which is often the case in lower-income countries. Indirect impacts affect a larger area for a longer period of time, for example when there are traffic delays and congestion on alternative routes, increased journey distances/durations, increased fuel consumption and associated greenhouse gas (GHG) emissions (Pyatkova *et al.*, 2019). Due to increasing urbanization, along with the expansion of intra-urban transport infrastructure, the built environment – which is in turn critical to economic development – is increasingly exposed to extreme weather events associated with climate change (Dawson *et al.*, 2018).

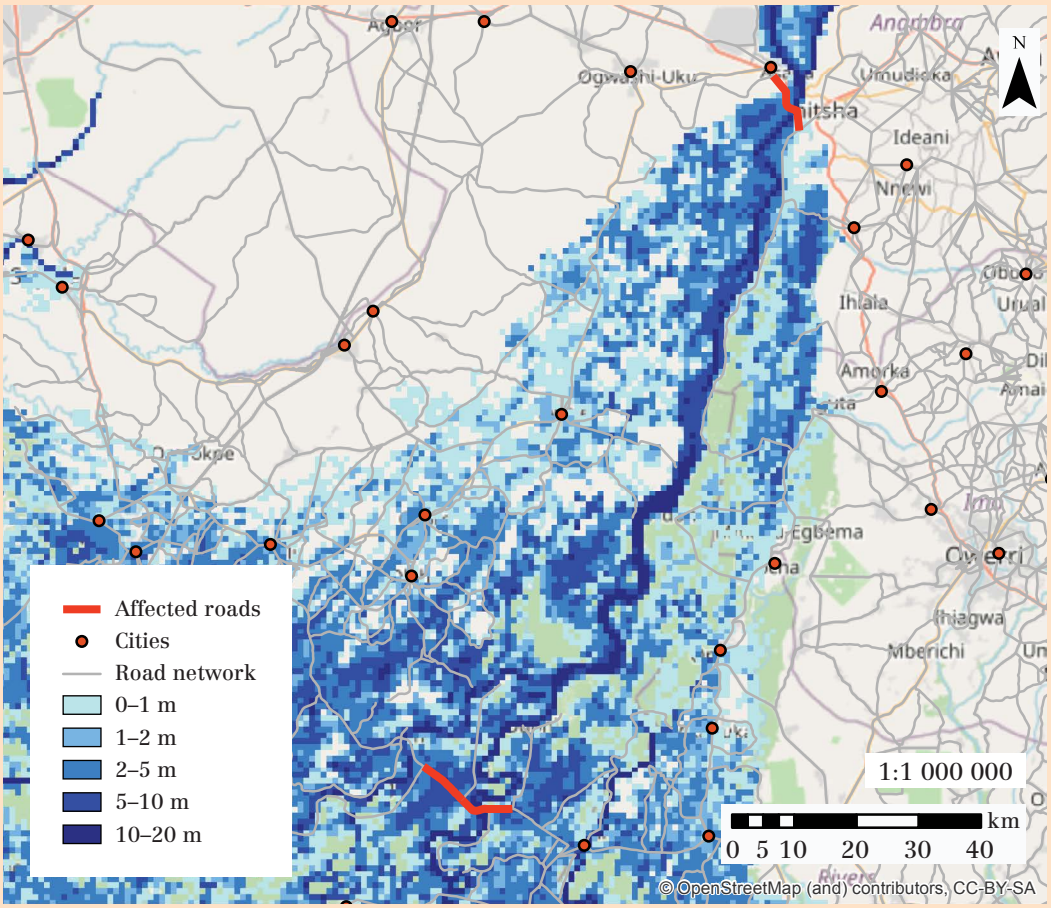
We assess the impact of localized flooding events in Nigeria, Mozambique and Pakistan, based on 1-in-10-year flood events. We consider plausible losses of network connectivity due to network links becoming impassable due to flooding, either because the link is damaged (for example, when a bridge is washed away), or because access to that link is reduced (for example, when an access road to a main road or bridge is damaged or submerged). These events have a major impact on the local population, not only in terms of reduced accessibility but also in terms of impact on all those who are served by the transport of goods through those damaged transport links. Finally, there is also a national impact, in that the ability of the transport network to efficiently transport goods may be significantly reduced. We illustrate the three scenarios below, with maps showing the flood risk, the transport links that we assume are affected, and the impact this has on our transport resilience metrics at both local and national level, based on the rerouting of origin–destination relations (which were previously routed over the disturbed links) over the remaining network.

Nigeria

In the Niger Delta region of Nigeria, where the Niger River empties into the Gulf of Guinea, there are large areas that flood regularly and make most roads impassable, particularly unpaved and unelevated roads. Most roads in the Niger Delta are local roads, but there are also two major roads that run from east to west (as shown by the red links in Figure 27). We have modelled a disruption on these two roads, based on a 1-in-10-year flood event.[3] The removal of these links amounted to an increase of 4.7 percent in the total travel time on the national network. Across the ten crop groups for which there is the highest demand in Nigeria, this scenario led to a 108 percent increase in travel time for trips that would have otherwise used the damaged links.

[3] While there is some doubt as to whether removing the northern link is realistic, as the bridge is elevated (according to Google Street View), it is plausible that access roads become impassable.

FIGURE 27 Flood scenario and affected roads in Nigeria



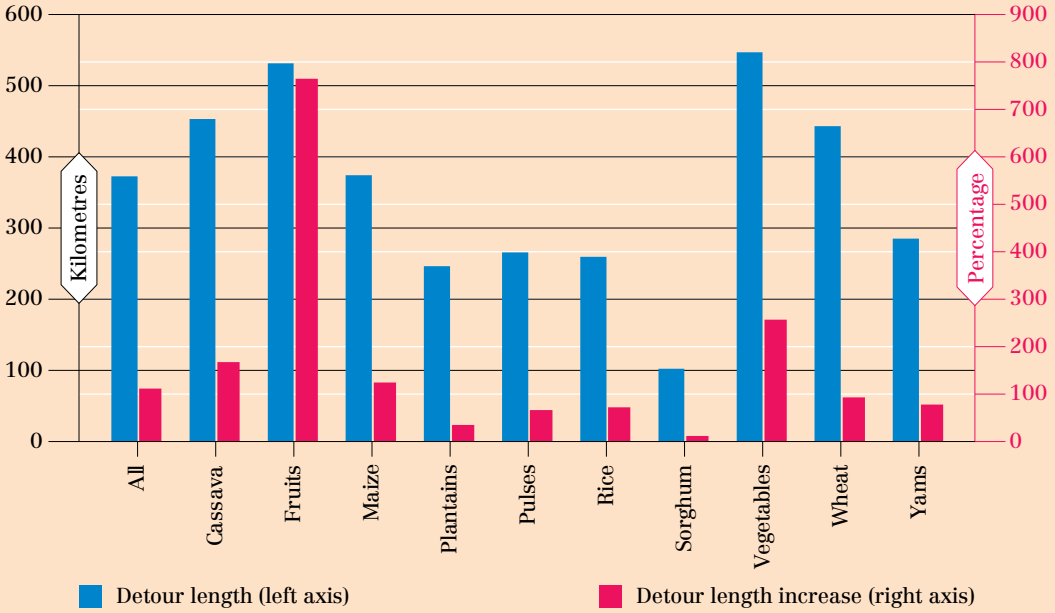
Source: UN (2020) modified by the authors.

Transport time detour by crop: Vegetables and fruits have the greatest detour lengths and times as percentages of their usual route journey. The distances in km (around 500 km for both) and the durations (around six hours for both) are also substantial. These increases can be detrimental to the quality of the crops, especially if they are exposed to high temperatures and/or humidity for extended periods, for example while being transported in non-refrigerated trucks. For all crop groups, the detour was longer than 100 km and more than 90 minutes (see Figure 28).

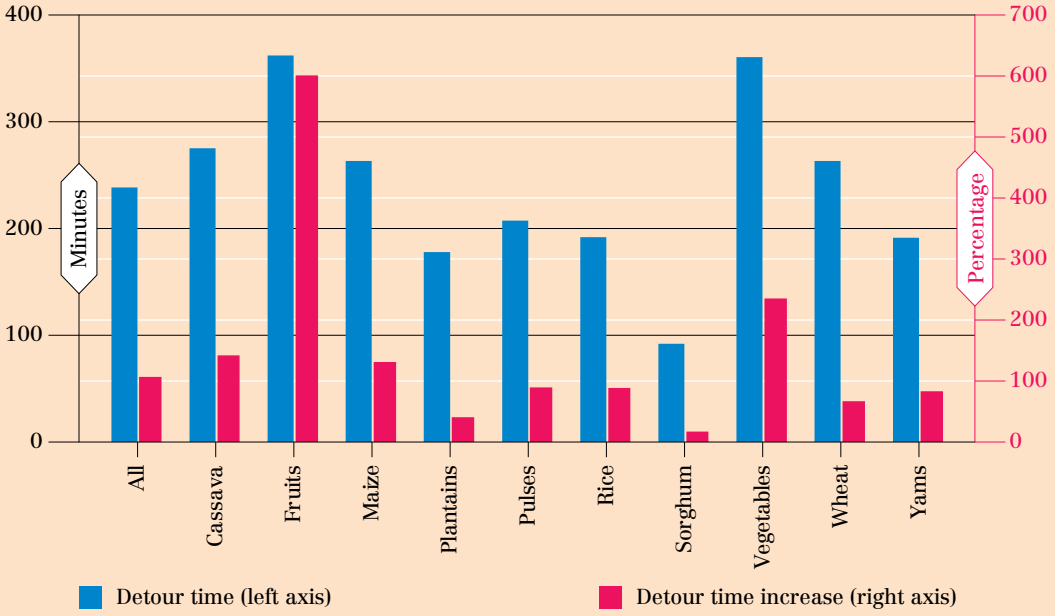
Tonnage diverted and people affected: We can estimate the tonnage diverted on its route between origin and destination locations in the regional network (see Figure 29). Because the duration of the flood's impact is difficult to assess, we report these estimates in terms of impact per month of closure across all crops, and for all trips through origins and destinations – with Lagos, Asaba, and Benin City having the greatest quantity of diverted crops in this scenario.

FIGURE 28 Detours by length and duration in the Nigeria scenario, by crop

A. TRIP LENGTH DETOURS

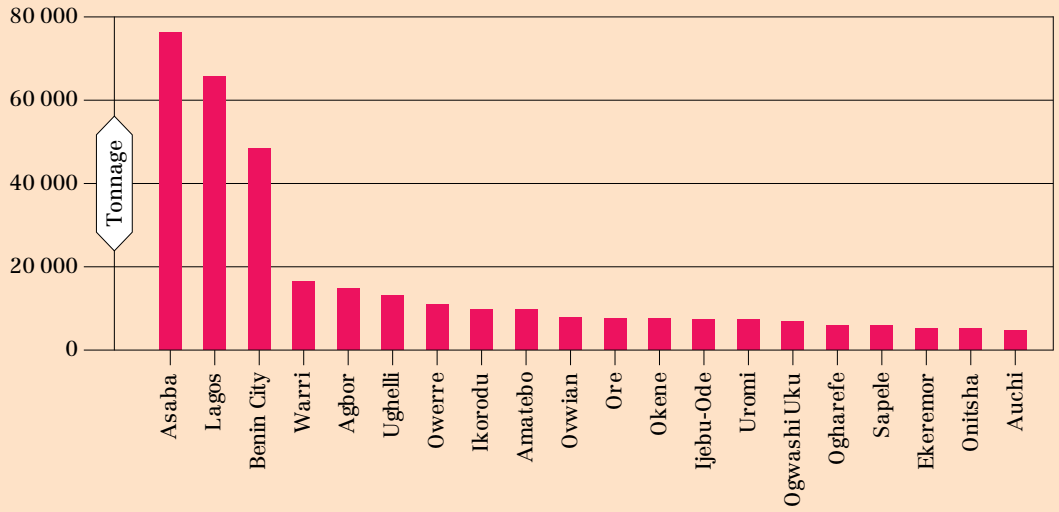


B. TRAVEL TIME DETOURS



Source: Authors' own elaboration.

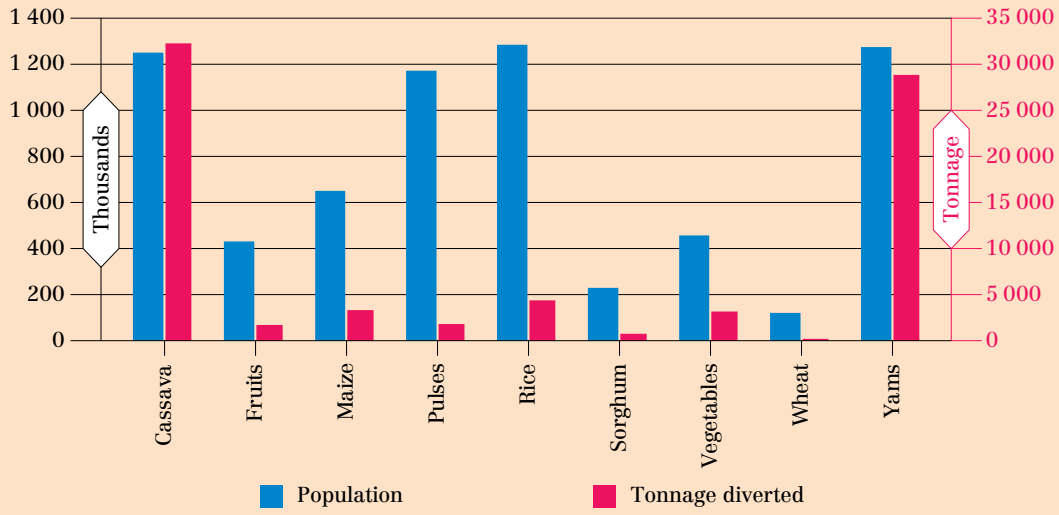
FIGURE 29 Tonnage diverted per month in the Nigeria scenario



Source: Authors' own elaboration.

We assess the potential impact more locally in the flood-affected area, by considering the people affected and the tonnage diverted in the Asaba region (see Figure 30). We look at the share of the original origin–destination volumes on the total volumes from those cities, to get an estimate of the people affected. This differs by crop; for example, 96 percent of all trips delivering cassava to Asaba use the two (now impassable) transport links, which means that 1.2 million people are directly affected. This number is lower for sorghum and wheat, which are produced elsewhere and routed differently, and are therefore minimally affected by the flooding. The figure also shows the tonnage per crop diverted, which is high for the two locally produced staple crops – cassava and yams. The rice crop is similar in that it is locally produced and affects over 1.2 million people; however the quantity is much lower.

FIGURE 30 People affected and tonnage diverted per month by crop in the Nigeria scenario

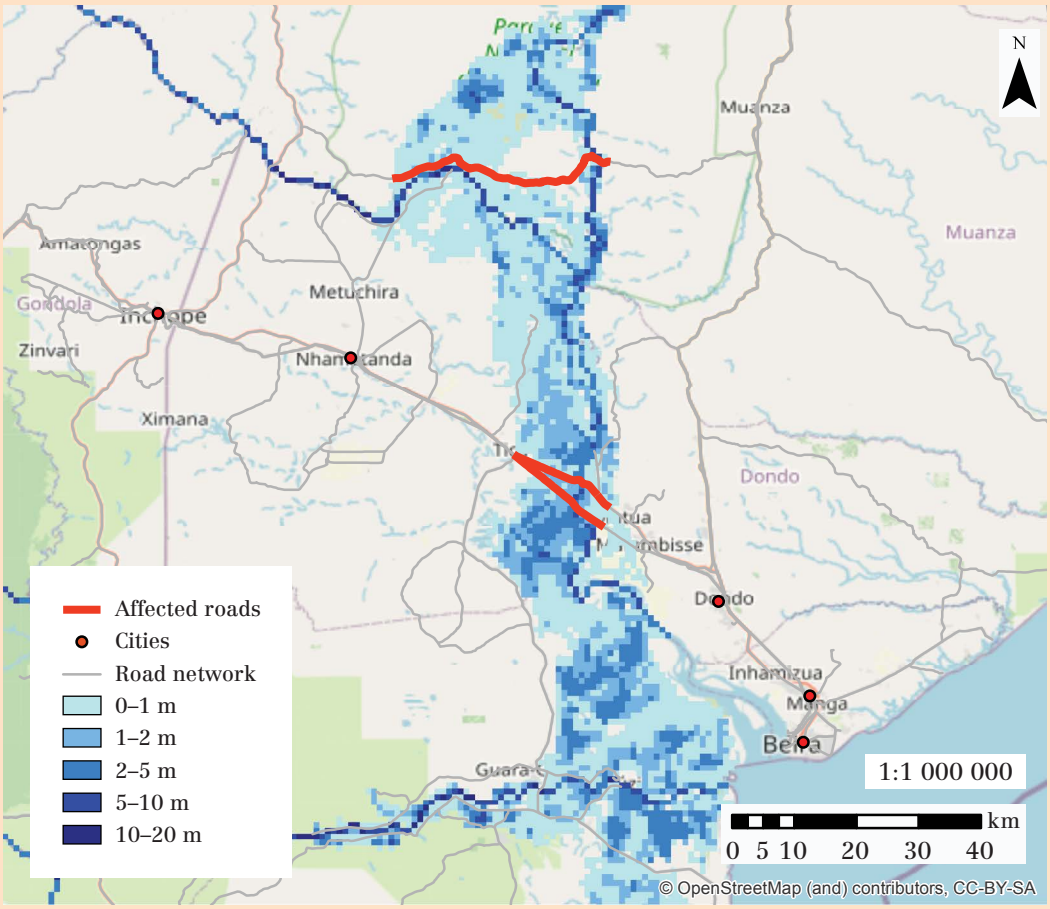


Source: Authors' own elaboration.

Mozambique

Mozambique experiences regular flooding associated with tropical cyclones that affect the low-lying coastal areas and deltas of the Zambezi, Pungwe and Limpopo rivers, among others. In this scenario, we model a 1-in-10-year inundation near the city of Beira (see Figure 31). As a result of this flood scenario, several roads become flooded, which disrupts part of the network around Beira and leads to longer transport times. For Mozambique, this scenario led to a 103 percent increase in travel time for those trips that would have otherwise used the damaged links. For the national network as a whole, the removal of these two links resulted in an overall increase in travel time of 1.25 percent.

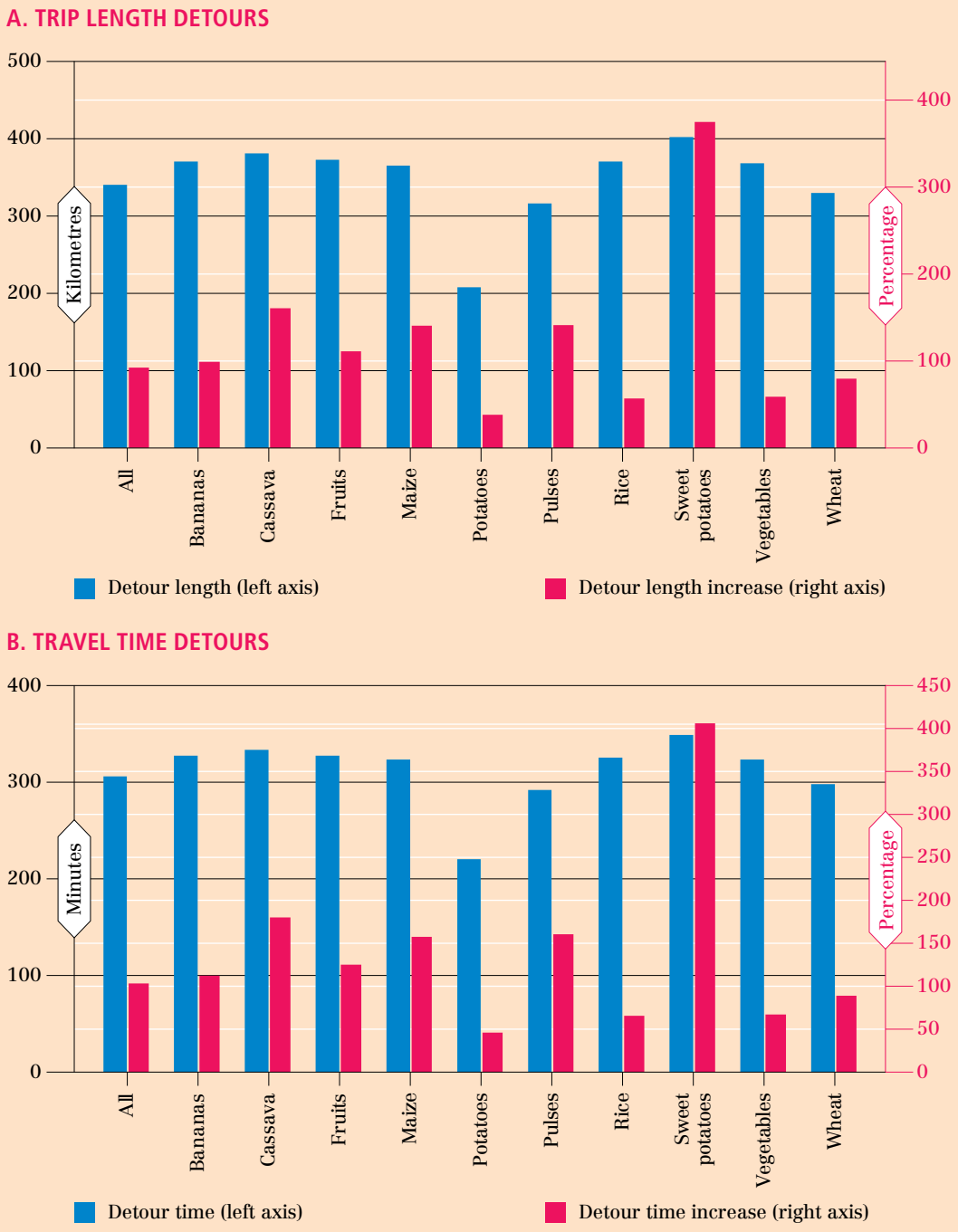
FIGURE 31 Flood scenario and affected roads in Mozambique



Source: UN (2020) modified by the authors.

Transport time detour by crop: All ten crops have substantially long detours (from 200 to 400 km) in their origin–destination trips. Potatoes have the lowest percentage increase (38 percent) while sweet potatoes have the highest (374 percent), suggesting that these routes are critical for sweet potatoes. In terms of travel time, the detours are approximately between 4 and 6 hours, and range from an increase of 46 percent to over 400 percent – again for potatoes and sweet potatoes (see Figure 32).

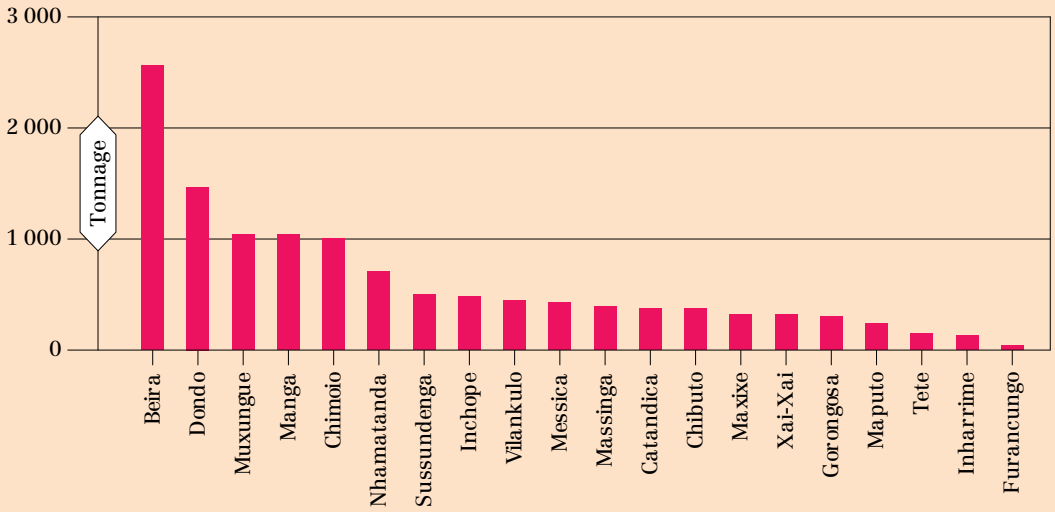
FIGURE 32 Detours by length and duration in the Mozambique scenario, by crop



Source: Authors' own elaboration.

Tonnage diverted and people affected: As may be expected, the greatest impact in terms of quantity of diverted crops per month is felt locally in Beira, but is also experienced in many other cities that are origins or destinations in the network (see Figure 33). The tonnage of diverted goods remains relatively high across several cities, whereas the impact tailed off quickly in Nigeria. This suggests that the impact was much more localized in Nigeria than in Mozambique.

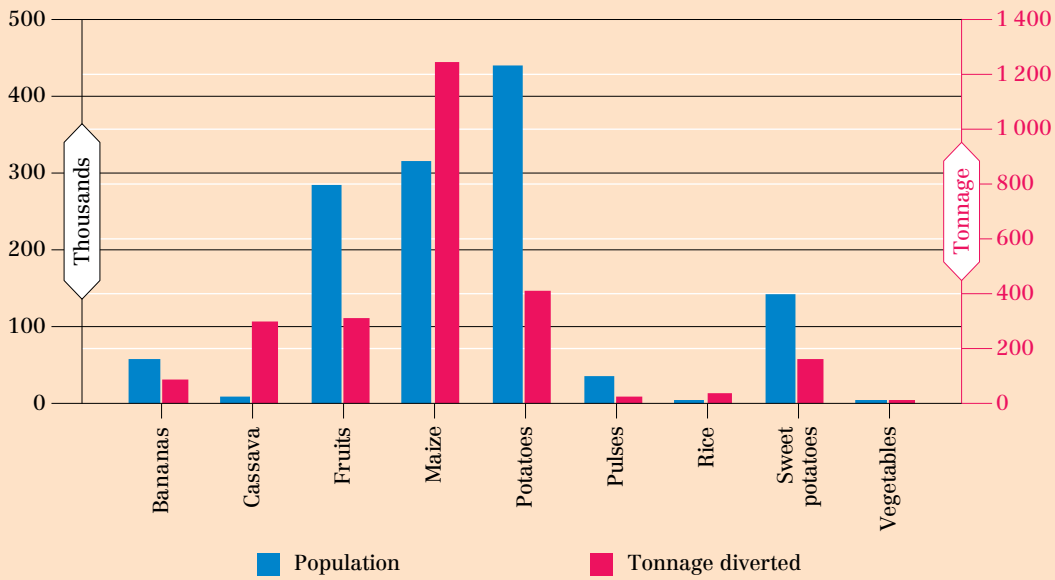
FIGURE 33 Tonnage diverted per month in the Mozambique scenario



Source: Authors' own elaboration.

We assess the potential impact more locally in the flood-affected area, by considering the people affected and the tonnage diverted in the Beira region (see Figure 34). We look at the share of the original origin–destination volumes on the total volumes from those cities, to get an estimate of the people affected. This differs by crop; for example, while potatoes were not affected much in terms of detour length or duration, the damaged network links are critical for its transport, and account for almost 100 percent of all trips delivering potatoes to Beira, thereby affecting 400 000 people. The numbers for people affected with regard to fruits and maize are similarly high.

FIGURE 34 People affected and tonnage diverted per month by crop in the Mozambique scenario

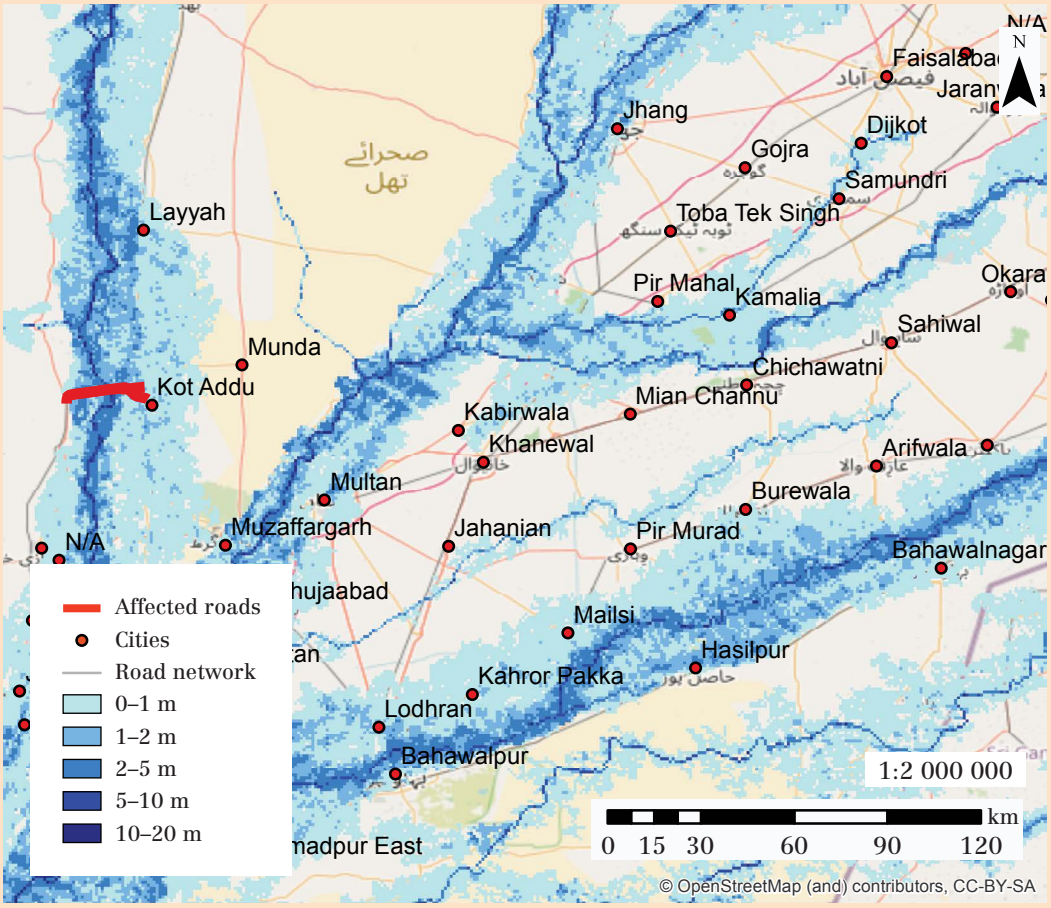


Source: Authors' own elaboration.

Pakistan

Parts of Pakistan are regularly flooded, impacting the distribution of food across the road network. In monsoon periods, local rainfall often occurs in combination with riverine flooding, leading to the flooding of large areas. In this scenario, we model a 1-in-1-year flooding event in the Indus River plain, in particular the area near Dera Ismail Khan (see Figure 35). This area has in fact undergone recent flooding. In this scenario, the connecting roads leading to three crossing bridges are temporarily taken out of the network. For Pakistan, this scenario led to a 32 percent increase of travel time for those trips that would have otherwise used the damaged links. The removal of these two links resulted in an overall increase in travel time of 1.19 percent for the national network.

FIGURE 35 Flood scenario and affected roads in Pakistan

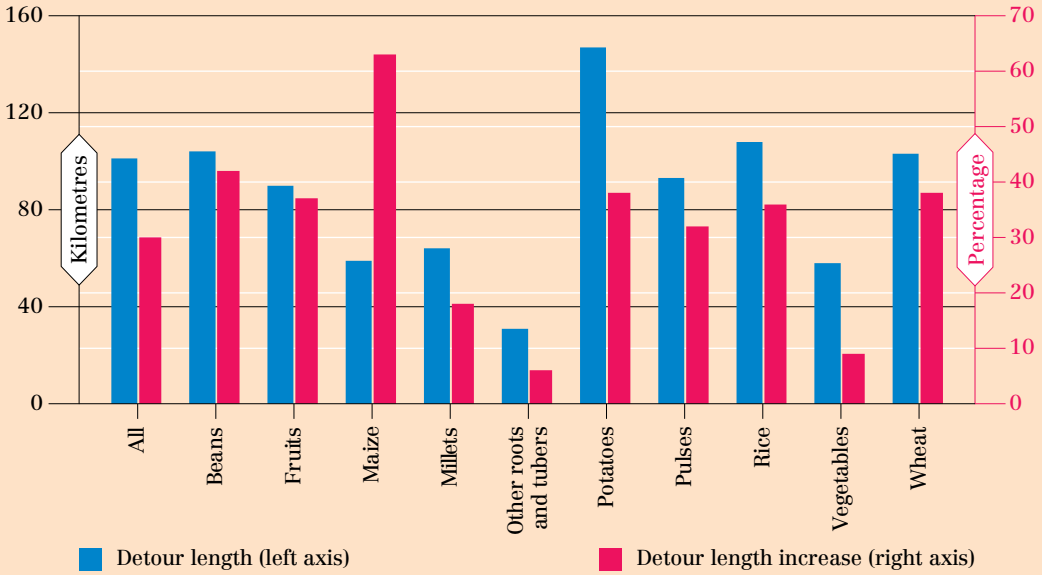


Source: UN (2020) modified by the authors.

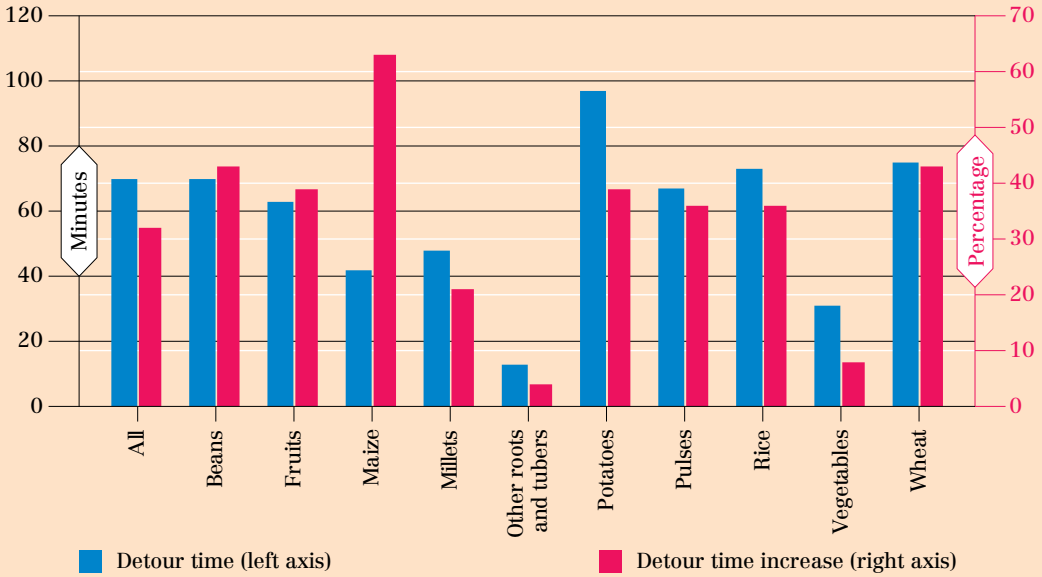
Transport time detour by crop: The overall impact is much lower in Pakistan, as the length and duration of detours are much lower than for Nigeria or Mozambique (less than 150 km and 100 minutes), with similarly small percentage increases (4 to 65 percent) relative to the other two countries (see Figure 36). The longest detours are for staples like rice, wheat and potatoes, though pulses, beans and fruit are also impacted. It is important to note the map scale here, as this is a much more localized flood event over a smaller geographic area, and therefore has a comparatively lower impact on the overall performance of the network in this region.

FIGURE 36 Detours by length and duration in the Pakistan scenario, by crop

A. TRIP LENGTH DETOURS



B. TRAVEL TIME DETOURS

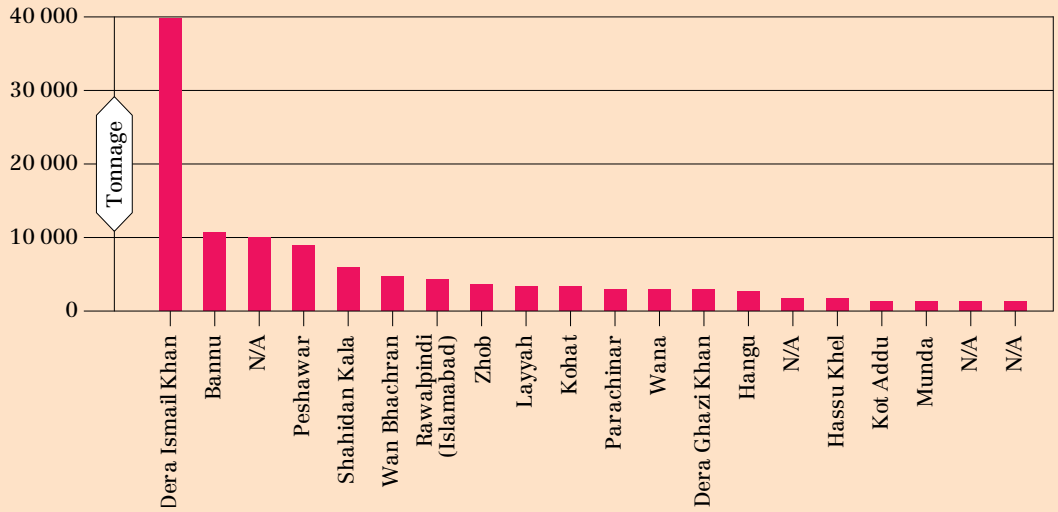


Source: Authors' own elaboration.

Tonnage diverted and people affected: The localized aspect of the flood is clearly seen in Figure 37: the impact felt is greatest by far in the city in the centre of the study area, and tails off quickly in the surrounding towns (where alternative routes outside the study area are more rapidly reached).

Even a very localized flood event can have a substantial impact. In Figure 38 we see that over 50 percent of all products delivered to Dera Ismail Khan are affected, albeit with significant variation in impact for each crop. Almost 100 percent of potato transport is affected, impacting over 70 000 people. The figures for transported tonnage of rice and wheat affected are 78 percent and 77 percent respectively, affecting over 50 000 people in each case.

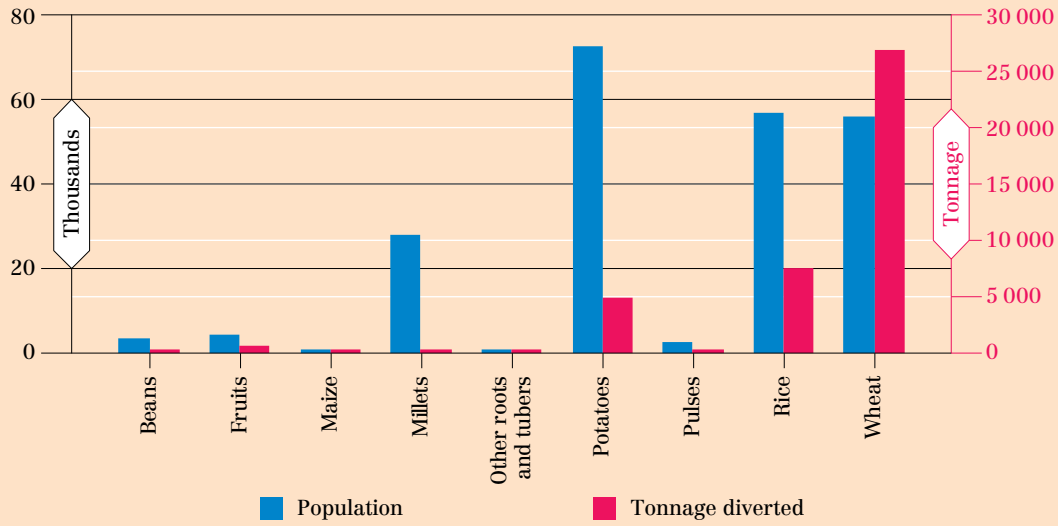
FIGURE 37 Tonnage diverted per month in the Pakistan scenario



Notes: N/A refers to unspecified cities in the country.

Source: Authors' own elaboration.

FIGURE 38 People affected and tonnage diverted per month by crop in the Pakistan scenario



Source: Authors' own elaboration.

Further comparisons across countries

The uncorrected proximity-based resilience metric scores are 0.075 (for Mozambique), 0.174 (for Pakistan) and 0.182 (for Nigeria); however there is very little change before and after the flooding, because the parts of the national network affected by these scenarios are relatively small. The relative detour cost resilience metric is more sensitive to local impacts, and for the cumulative links that have been closed off it changes by 0.75 percent for Pakistan, 2.7 percent for Mozambique and 3.3 percent for Nigeria.

We can also look at changes in travel times resulting from the flood's impact on the network. The percentage changes for all crops have already been given in each scenario described, but we expand them here for all crops (see Table 6).

Finally, we also consider the topology and connectivity of the transport network, independently through graph metrics. In many of our graph metrics, Nigeria and Pakistan are very close in their scores, and Mozambique's scores usually compare unfavourably against those of the other two. For instance, both Nigeria and Pakistan have a detour index of 0.90, while Mozambique has an index of 1.53. Mozambique also has the lowest *pi* index (edge density), by a factor of 10 against that of Pakistan and of 34 against Nigeria. A very similar comparison arises on cut points of the graph, once again with Mozambique's scores among the lowest of the three. These results indicate that there is a local disturbance effect only. This is comparable with the scenarios that were calculated in the network resilience metric, and are evidence of internal consistency in our modelling.

TABLE 6 Changes in travel time over the whole network, per crop and per country

Crop	Nigeria	Mozambique	Pakistan
All crops	4.7%	1.2%	1.2%
Bananas	N/A	1.0%	N/A
Beans	N/A	N/A	0.8%
Cassava	6.7%	0.2%	N/A
Fruits	1.1%	3.3%	0.2%
Maize	2.8%	3.7%	0.1%
Millets	N/A	N/A	1.1%
Other roots and tubers	N/A	N/A	0.0%
Plantains	4.5%	N/A	N/A
Potatoes	N/A	1.2%	1.6%
Pulses	1.5%	0.5%	0.1%
Rice	5.1%	0.3%	0.9%
Sorghum	N/A	N/A	N/A
Sweet potatoes	N/A	2.1%	N/A
Vegetables	4.8%	0.1%	0.6%
Wheat	1.1%	15.9%	2.5%
Yams	4.8%	N/A	N/A

Source: Authors' own elaboration.

6 Limitations and conclusions

This work has provided a first geospatial framework to represent and model national food transport network resilience at a global scale. It has developed a unique and internally consistent database along with plausible representations of complex transport networks, and it has generated network and resilience metrics to characterize the network and its ability to transport food to meet demand. While this work has established a new toolkit for measuring resilience – which in turn promises further use and applications beyond this study – its conceptualization and development involved many decisions that deserve critical reflection.

In this section, we therefore look back on the challenges of representing diverse transport networks across many countries, the inevitable choices that needed to be made to model their resilience, and opportunities to address some of the limitations of the current workflow. We also consider how the workflow can be applied to other uses.

6.1 Limitations and further work

Diverse data: Bringing together a wide variety of data sources, as in this project, is a challenge in itself. The data need to be internally consistent, to rest on a sound conceptual/semantic basis that is well-understood, and fit both temporally and spatially with the other sources. Among other things, this meant that appropriate methods had to be devised to address data gaps, and that default slot fillers had to be developed. Our workflows needed to have appropriate data interfaces as well; this was not fully understood at the start.

Thresholds: Our models take in data for many countries, and therefore had to depend on a wide variety of assumptions – often involving decisions on what to include and exclude, and at what level to define thresholds. We took great care in picking reasonable values for these decisions, but given the limited time available to experiment with such settings, some of the choices may be suboptimal. (An obvious example in this regard is the "snapping" grid resolution as used in the simplification workflow.) As such, various sensitivity tests still need to be conducted, in order to assess the optimality of chosen values and thresholds.

Computational intensity: While it was theoretically possible to generate resilience metrics for all countries, a small number of static graph metrics – such as local vertex connectivity for India and the centrality metrics for China, India and the United States of America – could not be computed. (This is despite many days of computation on dedicated high performance hardware.) These constraints could of course be addressed with more time for code and workflow optimization, but serve here as examples of the challenges involved in modelling a complex and detailed transport network.

Global trade model: Our global food trade model allows for international trade volumes available at country level to be distributed geographically to trade stations, while considering transport costs as characterized by type of transport and distance. The model takes transit routes passing from neighbouring countries and enables realistic estimates with minimum data input. However, better results may be obtained by studying major international transport routes in detail, and fine-tuning the model parameters accordingly.

National transport model: In view of available data and time, the transport model employed is based on several assumptions, three of the most important of which relate to:

(i) the all-or-nothing assignment; (ii) the treatment of multimodality; and (iii) the within-catchment travel time. These are discussed in more detail as follows:

(i) We assume that all freight between an origin and a destination gets assigned fully to one set of links that provide the overall shortest travel time. In doing so, we do not take into account the effects of network link capacity and/or congestion. This leads to over-assignment for certain links that are in reality congested, and under-assignment for others.

(ii) With respect to multimodal networks, we have assumed that food gets transported over the links that involve the lowest cost, irrespective of whether these are road, rail or waterway links. We use time as cost function, based on the free-flowing speeds of these networks. We have not been able to use any data on the distribution of food freight over the different modes. It may be that in one country, freight is predominantly transported by rail (for example, the United States of America), whereas in our model the road network receives the largest share as it is more expansive. Alternatively, it may be that a given country uses the rail network for passengers only, yet we assign it freight.

(iii) We assumed an average of 20 minutes for within-catchment travel time. Changing this average travel time would not change the results substantially (in other words, countries would keep the same ranking), but the proximity-based resilience values would change – for example an increase in average within-catchment travel time would result in higher proximity-based resilience metrics. It is therefore important to emphasize that this resilience metric should be used mainly as a comparative metric between countries.

Metric correction: To enable a fair comparison between countries with different sizes, we applied a metric correction factor based on the total tonnage transported. This is important, because without such a correction factor, small countries would score artificially high on the proximity-based resilience metric, for example. Our correction aimed to account for "artificial scale effects" (such as when the average trip time – and hence the proximity-based resilience – of a small country is constrained by the physical size of the country), without impacting "real effects" such as population and network density. There are different ways to correct for size, but the difficulty is that there are probably latent variables influencing the resilience metrics, and these can vary across countries. For example, in some countries when population density increases, people will tend to live closer to the production locations of crops, and when network density increases, detours will be shorter. In other countries, urban development may push out agriculture (especially large-scale agriculture), for reasons of land competition, thereby increasing the distance between producers and consumers. Another effect relates to areas undergoing urbanization; these areas are typically also undergoing economic development, which brings with it more buying power and changes in consumer purchasing (as when people allow themselves to buy more expensive or imported food items). Our models do not "punish" for international transport, and the cost of transport begins to accrue from the trade station only. This is an effect that appears to bring production closer to consumption, at least in our models. For possible future improvements to correction factors, it is important to better study the relationships between resilience metrics and these latent variables (such as population density and country size), and explore various typologies of countries based on size, network, city distribution and spatial production patterns.

Validation: Several methods and algorithms developed during the study allowed for the analysis of a massive amount of geospatial data at the global level, in order to provide food transport resilience metrics for more than 90 countries, as well as for the ten most important food products for each country based on food supply. Due to the extent of the study and as already mentioned, certain assumptions and simplifications were unavoidable. A comprehensive validation study could have provided better insights on the validity of

these assumptions and the accuracy of the results, but was not feasible due to the limited time and resources available. As a suggestion for follow-up, a dedicated validation study may therefore be considered – such a study could aim to collect detailed information on actual national and international food transport for selected countries and crops, and test the developed models by comparing them to multiple historical disruption case studies.

Temporal aspects of network performance and resilience: Neither the global trade model nor the national transport models consider the temporal dimension (i.e. seasonality) of food transport. Although some crops are traded and transported throughout the year, it is well-known that others are produced and traded seasonally. Moreover, seasons vary by location, and depending on their latitude, different countries can experience different seasons at a given time of year. This has direct implications on global food trade. Depending on data availability, the consideration of seasonal crop production and trade will therefore be useful for future applications, and models need to be developed for this purpose.

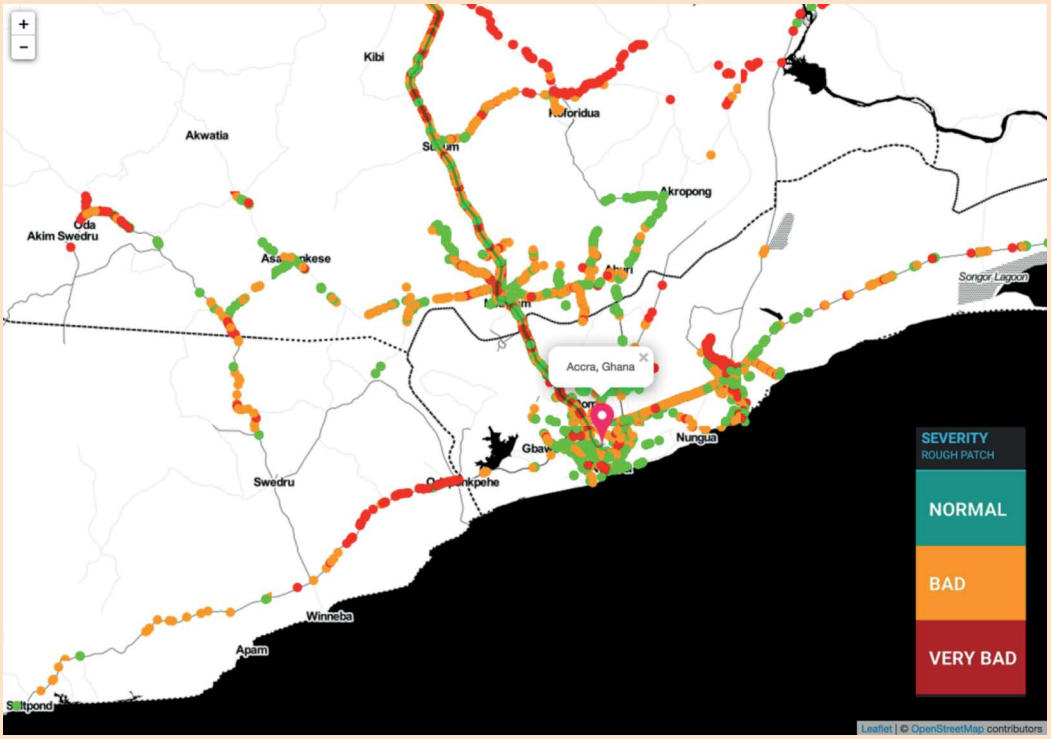
Multinational network analysis: At present, the network resilience model and related resilience metrics consider national transport networks only. Each country is considered individually, and any disruptions considered are also limited to the national level (they do not cross country boundaries). Given that international food trade is also vulnerable to similar disruptions, future work could involve creating and analysing multinational network resilience models by combining available national transport networks with a global trade model. This may allow for better identification of the cascading impacts of large-scale disruptions affecting multiple countries. That said, single-point disruptions can also have significant global impact, particularly if the point in question is located on a critical international trade route – as demonstrated by the blockage of the Suez Canal in 2021.

Positive disruptions: Given that the network resilience model and related resilience metrics are fine-tuned to the level of considering each road, waterway and railroad individually, it is not only the impact of disruptions but also the impact of possible improvements to the food transport network that can be modelled at the lowest applicable hierarchical level. This offers the opportunity to explore "what-if" scenarios and assess the cost–benefit of different transport infrastructure innovations, such as the treatment of dirt roads to improve their resilience to tropical rains and to wear-and-tear, or the introduction of transport terminal technology to prevent post-harvest losses. As an example of the former, dirt road surfaces can be sprayed with an environmentally friendly sealant such as Road Rapid (OSO Enschede, n.d.). This results in a permanent change to the molecular structure of the soil and an instant hardening of the road surface to the consistency of rock, thereby not only reducing maintenance but also increasing the load-bearing capacity of the road. This in turn creates more alternative routes, including for heavier classes of transport vehicles, greatly enhancing transport network resilience in the process. Additionally, and given the fine-grained OSM data used in our models, the sourcing of manpower and materials (such as soil and gravel) for labour-intensive public road works can now be optimized, for example by identifying small and medium-sized enterprises (SMEs) that are active in (road) engineering along trade corridors, and potentially creating youth employment opportunities that also benefit rural economies. As examples of the latter, there has been renewed interest in post-harvest terminals and cold stores such as those developed by ColdHubs (n.d.). As off-grid, renewable energy solutions become more generally accessible, they enable such innovations to become more broadly applicable.

Including non-traditional data sources: In a potential future evolution, we could consider crowdsourcing or citizen science techniques to capture relevant socio-economic aspects, as illustrated in the example of an award-winning app that was developed by a spin-off group at the University of Twente in the Netherlands (CHEETAH, n.d.). CHEETAH

(which is an acronym for Chains of Human Intelligence towards Efficiency and Equity in Agro-Food Trade along the trans-Africa Highway) uses a smartphone app to capture information on non-tariff barriers that impede transport operations; these include highway tolls, unapproved barriers, customs, traffic, illicit payments, police checkpoints, local tolls and road barriers. At the same time, Edge AI (artificial intelligence) that also runs on the smartphone is used to capture information on road pavement quality, including speed bumps, rough patches, potholes, bridge expansion joints, rumble strips, corrugated surfaces, road stops and sunken utility covers (see Figure 39).

FIGURE 39 Example of road patch quality in Ghana as crowdsourced from CHEETAH



Source: Ujuizi Laboratories, 2018.

CHEETAH can reduce costs for drivers, for example on tire damages and frequent mechanical maintenance, and in the process help public agencies in capturing road pavement quality to address some of the prioritization, planning and monitoring needs associated with improvements to transport resilience. Similarly, current cost–benefit analyses often ignore the ways in which (positive) changes and improvements to road infrastructure reduce post-harvest losses (Vursavuş and Özgüven, 2004; Idah, Ajisegiri and Yisa, 2007; Venus *et al.*, 2013); these can now also be (dis)aggregated, given the richness of the underlying data models.

6.2 Conclusions

This work has provided a first geospatial framework to represent and model national food transport network resilience at a global scale. It has developed a unique and internally consistent database along with plausible representations of complex transport networks, and has generated network and resilience metrics to characterize the network and its ability to transport food to meet demand. As discussed, the process of conceptualization and development involved many decisions that deserve critical reflection, and further work is required; nevertheless the work completed to date has established a new toolkit for measuring resilience, and offers a range of further uses and applications beyond this study.

References

Banks, J. 2002. *Introduction to Transportation Engineering*. 2nd edition. Boston, USA, McGraw-Hill.

Bertoli, S., Goujon, M. & Santoni, O. 2016. *The CERDI-seadistance database*. Clermont-Ferrand, France, Centre d'Études et de Recherches sur le Développement International (CERDI). (also available at https://halshs.archives-ouvertes.fr/halshs-01288748/document).

Cattaneo, A., Nelson, A. & McMenomy, T. 2021. Global mapping of urban–rural catchment areas reveals unequal access to services. *Proceedings of the National Academy of Sciences*, 118(2).

CHEETAH. (no date). *CHEETAH Data Products* [online]. [Cited May 2021]. http://cheetah.ujuizi.com

ColdHubs. (no date). *ColdHubs* [online]. [Cited May 2021]. www.coldhubs.com

Davis, K.F., Downs, S. & Gephart, J.A. 2021. Towards food supply chain resilience to environmental shocks. *Nature Food*, 2: 54–65.

Dawson, R.J., Thompson, D., Johns, D., Wood, R., Darch, G., Chapman, L., Hughes, P.N. *et al.* 2018. A systems framework for national assessment of climate risks to infrastructure. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2121): 20170298.

Dijkstra, E.W. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1): 269–271.

FAO. 2017. *New food balances: description of utilization variables*. (also available at http://fenixservices.fao.org/faostat/static/documents/FBS/New%20FBS%20methodology.pdf).

FAO. 2020. FAOSTAT. In: *FAO* [online]. [Cited October 2020]. http://faostat.fao.org

Geofabrik. 2021. *OpenStreetMap Data Extracts* [online]. [Cited March 2021]. http://download.geofabrik.de

Global Administrative Areas. 2021. GADM database of Global Administrative Areas, version 3.6. In: *GADM* [online]. [Cited October 2020]. https://gadm.org

Idah, P.A., Ajisegiri, E.S.A. & Yisa, M.G. 2007. Fruits and vegetables handling and transportation in Nigeria. *Assumption University Journal of Technology*, 10(3): 175–183.

Joint Research Centre – European Commission (JRC). 2015. *GHS-POP R2015A - GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015) - OBSOLETE RELEASE. [data set]. JRC*. (also available at https://data.jrc.ec.europa.eu/dataset/jrc-ghsl-ghs_pop_gpw4_globe_r2015a).

Lehner, B. 2013. *HydroSHEDS Technical Documentation*. Version 1.2. Washington, DC, World Wildlife Fund US. (also available at www.hydrosheds.org/images/inpages/HydroSHEDS_TechDoc_v1_2.pdf).

MapSPAM. 2019. Methodology: a look behind SPAM and what makes it run. In: *MapSPAM* [online]. [Cited August 2020]. http://mapspam.info/methodology

National Geospatial-Intelligence Agency. 2001. *Distance between ports*. Pub. 151. 11th edition. Bethesda, USA. (also available at https://msi.nga.mil/api/publications/download?key=16694076/SFH00000/Pub151bk.pdf).

National Geospatial-Intelligence Agency. 2017. *World Port Index*. 26th edition. Springfield, USA.

Nelson, A., de By, R., Girgin, S., Brussel, M. & Ohuru, R. 2020. *Analysing the resilience of domestic transport networks in the context of food security*. Background paper for *The State of Food and Agriculture 2021*. Rome, FAO (unpublished).

OpenStreetMap. 2021. *Planet OSM* [online]. [Cited March 2021]. https://planet.osm.org

Ortúzar, J. de D. & Willumsen, L.G. 2011. *Modelling Transport*. 4th edition. Chichester, UK, John Wiley & Sons.

OSO Enschede. (no date). *Intro Road Rapid* [online]. [Cited May 2021]. www.oso-enschede.com/?page_id=177

Pregnolato, M., Ford, A., Wilkinson, S.M. & Dawson, R.J. 2017. The impact of flooding on road transport: a depth-disruption function. *Transportation Research Part D: Transport and Environment*, 55: 67–81.

Pyatkova, K., Chen, A.S., Butler, D., Vojinović, Z. & Djordjević, S. 2019. Assessing the knock-on effects of flooding on road transportation. *Journal of Environmental Management*, 244: 48–60.

Tibshirani, R., Walther, G. & Hastie, T. 2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2): 411–423.

Ujuizi Laboratories. 2018. *CHEETAH: Chains of human intelligence; towards efficiency and equity in agro-food trade along the Trans-Africa Highway*. Pilot project report. Enschede, Netherlands. (also available at https://docs.google.com/document/d/1vTiCgnY3D-kd80nIdq5QbQMX-xDkw3jYJmm7BtrJpEA/edit).

United Nations (UN). 2020. *Map of the World* [online]. [Cited 1 January 2021]. https://unfao.sharepoint.com/sites/OCCP/UN%20Maps/Forms/AllItems.aspx

Venus, V., Asare-Kyei, D.K., Tijskens, L.M.M., Weir, M.J.C., de Bie, C.A.J.M., Ouedraogo, S., Nieuwenhuis, W. *et al.* 2013. Development and validation of a model to estimate postharvest losses during transport of tomatoes in West Africa. *Computers and Electronics in Agriculture*, 92: 32–47.

Vursavuş, K. & Özgüven, F. 2004. Determining the Effects of Vibration Parameters and Packaging Method on Mechanical Damage in Golden Delicious Apples. *Turkish Journal of Agriculture and Forestry*, 28: 311–320.

Wessel, P. & Walter, S. 2017. A Global Self-consistent, Hierarchical, High-resolution Shoreline Database. Version 2.3.7. In: *GSHHG* [online]. [Cited May 2021]. www.soest.hawaii.edu/pwessel/gshhg

World Bank. 2019. New country classifications by income level: 2019-2020. In: *World Bank* [online]. [Cited June 2020]. https://blogs.worldbank.org/opendata/new-country-classifications-income-level-2019-2020

World Port Source. 2021. *World Port Source* [online]. [Cited December 2020]. www.worldportsource.com

Yu, Q., You, L., Wood-Sichra, U., Ru, Y., Joglekar, A.K.B., Fritz, S., Xiong, W. ***et al.*** 2020. A cultivated planet in 2010 – Part 2: the global gridded agricultural-production maps. *Earth System Science Data*, 12(4): 3545–3572.

Annex

TABLE A1 List of countries analysed

Afghanistan	Democratic Republic of the Congo	Papua New Guinea
Algeria	Ghana	Peru
Angola	Guatemala	Philippines
Argentina	Guinea	Republic of Korea
Australia	Guinea-Bissau	Russian Federation
Azerbaijan	Haiti	Rwanda
Bangladesh	India	Saudi Arabia
Benin	Indonesia	Senegal
Bolivia (Plurinational State of)	Iran (Islamic Republic of)	Sierra Leone
Brazil	Iraq	Somalia
Burkina Faso	Italy	South Africa
Burundi	Jamaica	South Sudan
Cambodia	Japan	Sri Lanka
Cameroon	Kazakhstan	Sudan
Canada	Kenya	Syrian Arab Republic
Central African Republic	Lao People's Democratic Republic	Thailand
Chad	Liberia	Togo
Chile	Madagascar	Tunisia
China	Malawi	Turkey
Colombia	Pakistan	Uganda
Congo	Malaysia	Ukraine
Côte d'Ivoire	Mali	United Kingdom of Great Britain and Northern Ireland
Democratic People's Republic of Korea	Mauritius	United Republic of Tanzania
Dominican Republic	Mexico	United States of America
Ecuador	Morocco	Uzbekistan
Egypt	Mozambique	Venezuela (Bolivarian Republic of)
Ethiopia	Myanmar	Viet Nam
France	Nepal	Yemen
Gabon	Niger	Zambia
Germany	Nigeria	Zimbabwe

Source: Authors' own elaboration.

Transport infrastructure and logistics, not least domestic food transport networks, are an integral part of agrifood systems, and play a fundamental role in ensuring physical access to food. However, the resilience of these networks has rarely been studied. This study fills this gap and analyses the structure of food transport networks for a total of 90 countries, as well as their resilience through a set of indicators.

Findings show that where food is transported more locally and where the network is denser, systematic disturbances have a much lower impact. This is mostly the case for high-income countries, as well as for densely populated countries like China, India, Nigeria and Pakistan. Conversely, low-income countries have much lower levels of transport network resilience, although some exceptions exist.

The study further simulates the effect of potential disruptions – namely floods – to food transport networks in three countries. The simulation illustrates the loss of network connectivity that results when links become impassable, potentially affecting millions of people.

Overall, this study provides a first geospatial framework to represent and model national food transport network resilience at a global scale considering not only local production and consumption, but also international trade. It has established a new toolkit for measuring resilience, which promises further use and applications beyond this study.

14

FAO AGRICULTURAL DEVELOPMENT ECONOMICS TECHNICAL STUDY

The FAO Agricultural Development Economics Technical Study series collects technical papers addressing policy-oriented assessments of economic and social aspects of food security and nutrition, sustainable agriculture and rural development.

The series is available at www.fao.org/economic/esa/technical-studies

FOR FURTHER INFORMATION

Agrifood Economics - Economic and Social Development

- www.fao.org/economic/esa
- ESA-Director@fao.org

Food and Agriculture Organization
of the United Nations (FAO)

Rome, Italy

ISBN 978-92-5-135363-9 ISSN 2521-7240

9 789251 353639

CB7757EN/1/11.21