nature food

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s43016-020-0060-7

In the format provided by the authors and unedited.

Local food crop production can fulfil demand for less than one-third of the population

Pekka Kinnunen[1]✉, Joseph H. A. Guillaume[1,2], Maija Taka[1], Paolo D'Odorico[3], Stefan Siebert[4], Michael J. Puma[5,6], Mika Jalava[1] and Matti Kummu[1]✉

[1]Water and Development Research Group, Aalto University, Espoo, Finland. [2]Institute for Water Futures and The Fenner School of Environment & Society, The Australian National University, Canberra, Australian Capital Territory, Australia. [3]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA. [4]Department of Crop Sciences, University of Göttingen, Göttingen, Germany. [5]NASA Goddard Institute for Space Studies, New York, NY, USA. [6]Center for Climate Systems Research, Earth Institute, Columbia University, New York, NY, USA. ✉e-mail: pekka.kinnunen@aalto.fi; matti.kummu@aalto.fi

Supplementary Information

Local food crop production can fulfil demand for less than one-third of the population

P. Kinnunen, J. H. A. Guillaume, M. Taka, P. D'Odorico, S. Siebert, M. J. Puma, M. Jalava, M. Kummu

Note S1. Mathematical formulation of the optimisation framework

Net energy supply

For a given crop, the gridded total net energy supply N_i in each cell *i* is the difference between local energy supply S_i and demand D_i. The values for all different parameters vary depending on the location and the crop in question.

$$N_i = S_i - D_i \tag{1}$$

The food supply consists of total food production, energy content of the food and food loss factor.

$$S_i = C_i E_i \left(1 - L_i^{prod}\right) \tag{2}$$

Where,

1. C_i is total production in kilograms for a given crop
2. E_i energy content in kilocalories per kilogram
3. L_i^{prod} is the loss percentage within the food production for a given crop [39]

The food demand within a cell is defined as the product of population, crop specific energy demand and waste percentage

$$D_i = P_i E_i^{req} (1 + L_i^{cons}) \tag{3}$$

Where,

1. P_i is the population in the cell *i*
2. E_i^{req} is the crop specific energy demand according to FAOSTAT
3. L_i^{cons} is the loss and waste percentage within food distribution and consumption networks [39]

The linear optimisation problem is given by Equation 4 subject to constraints in Equations 5 and 6. The optimisation problem was solved numerically in MATLAB®, using the interior-point method.

$$\min f^T V \qquad \textbf{(4)}$$

$$s.t.\ AV \leq N \qquad \textbf{(5)}$$

$$and\ V_i \geq 0 \qquad \textbf{(6)}$$

where

4. $\boldsymbol{f^T}$ is a vector describing the friction between cells (distance or transport travel time cost)
5. ***V*** is a vector containing all the food flows between cells, to be optimised.
6. ***A*** is a sparse matrix denoting connections between adjacent cells, allowing summation of supplies across cells to meet net energy needs, within the limits of available supply
7. ***N*** is a vector of net energy supply for a given crop. Positive net supply denotes surplus production in a given cell and negative net supply denotes deficit in local production.

Average distance calculation

Consider flows between two adjacent cells marked by indices *i* and *j*. For a given cell *j*, the total distance can be divided into two parts: the flows between adjacent cells *i* and *j* and the mean travelled distance before reaching the cell *i*. Multiplying the distance with the volume of imported food yields food miles. Thus, the total imported food miles for cell *j* is equal to the sum of all of the flows from each of the adjacent cells (Equation 7).

$$M_j^{imp} = \sum_i V_i (d_i^{mean} + d_{i,j}) \qquad \textbf{(7)}$$

Where:

1. M_j^{imp} is total imported food miles (kcal km)
2. *i* is a source location for food
3. V_i is the total food volume imported from source location *i* (kcal)
4. d_i^{mean} is the average distance travelled for food to get to source *i* (km)
5. $d_{i,j}$ is the distance between cells *i* and *j* (km)

The mean average distance for each cell is total imported food miles divided by local production and the total imported food volume (Equation 8).

$$d_i^{mean} = \frac{M_i^{imp}}{l_i + \sum_k V_{i,k}} \quad \textbf{(8)}$$

Where:

1. l_i is the food volume produced locally in location *i* (kcal)
2. $\sum_k V_{i,k}$ is the total imported food volume to location *i* from all connected locations *k* (kcal)

Substituting Equation 8 into the Equation 7 gives us Equation 9 with only one unknown (vector) variable, M_j^{tot}. Now, we have a system of linear equations which allows us to solve simultaneously the average imported food miles for all the grid cells.

$$M_j^{imp} - \sum_i V_i \frac{M_i^{imp}}{l_i + \sum_k V_{i,k}} = \sum_i V_i d_{i,j} \quad \textbf{(9)}$$

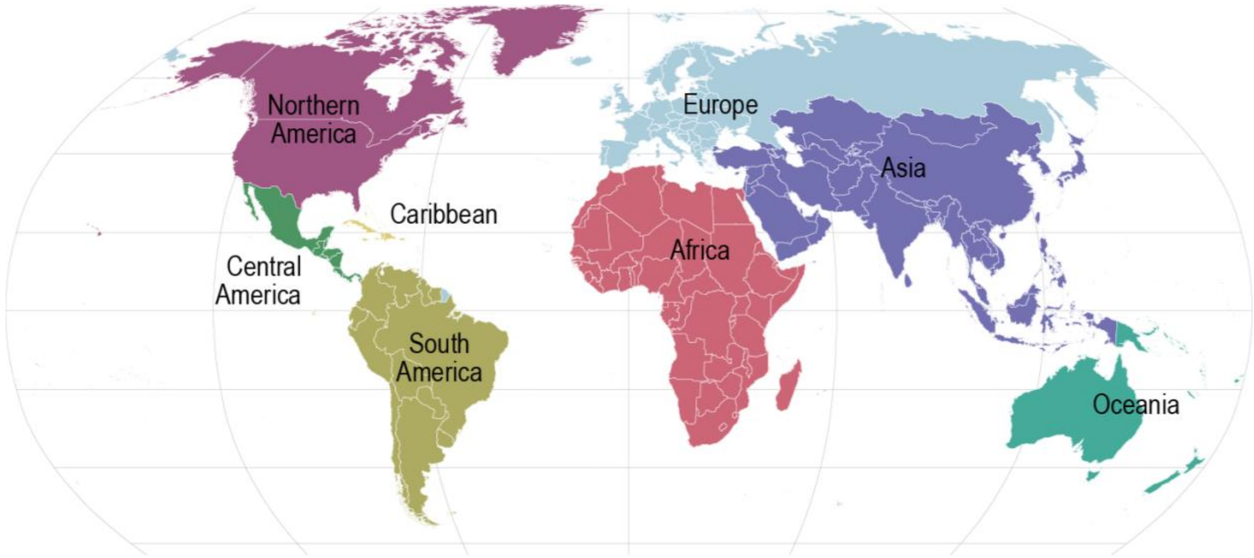
Lastly, we calculate the total food miles M_j^{tot} by subtracting the share of food flows continuing from cell *i* (exported or only passing through) from the imported food miles.

$$M_j^{tot} = \left(1 - \frac{\sum_j V_j^{exp}}{l_j + \sum_i V_i^{imp}}\right) M_j^{imp} \quad \textbf{(10)}$$

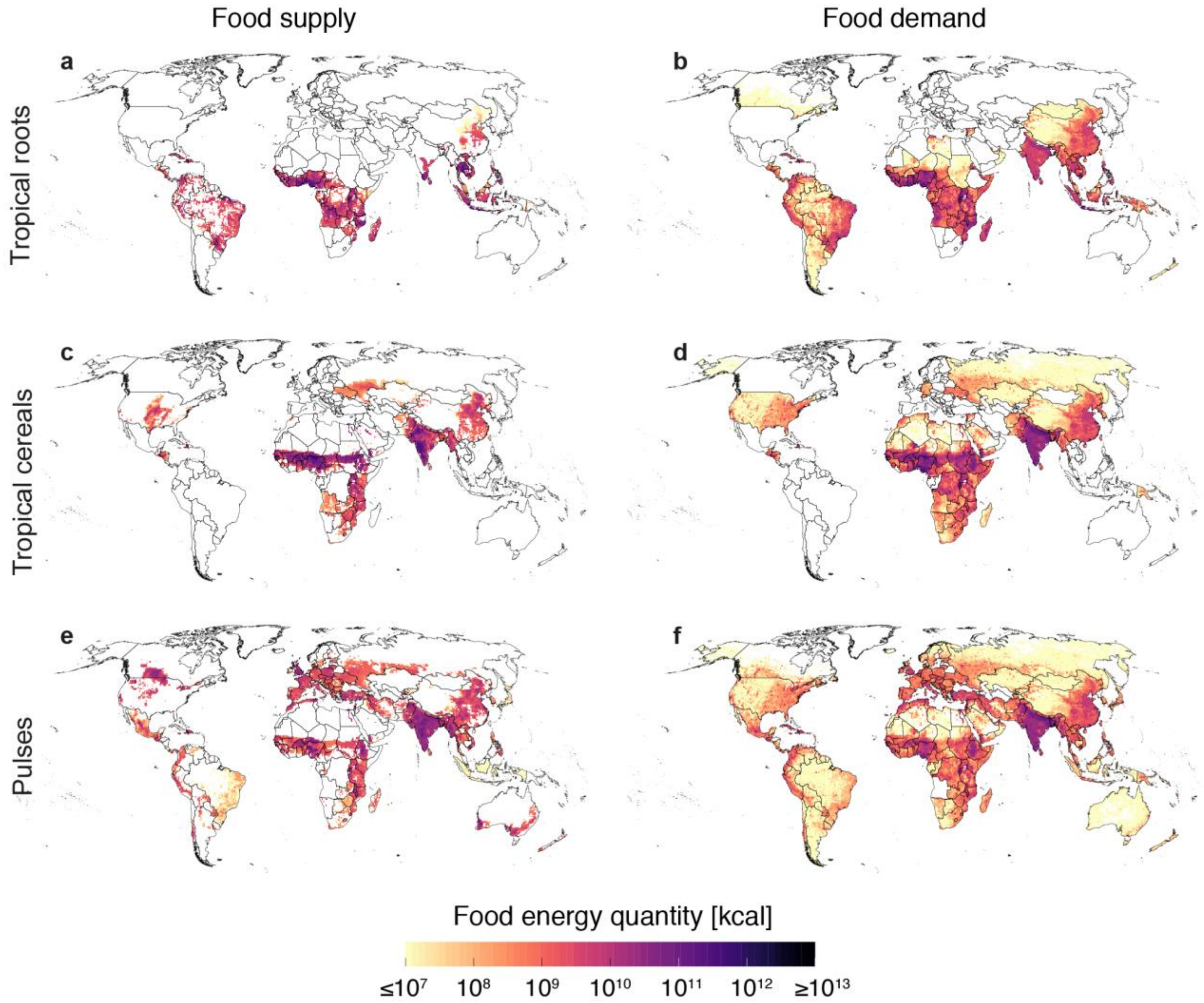
Where:

1. $l_j + \sum_i V_i^{imp}$ is the sum of local production and imports, i.e. the total food available in cell *j*,
2. $\sum_j V_j^{exp}$ is the sum of the outgoing flows from cell j

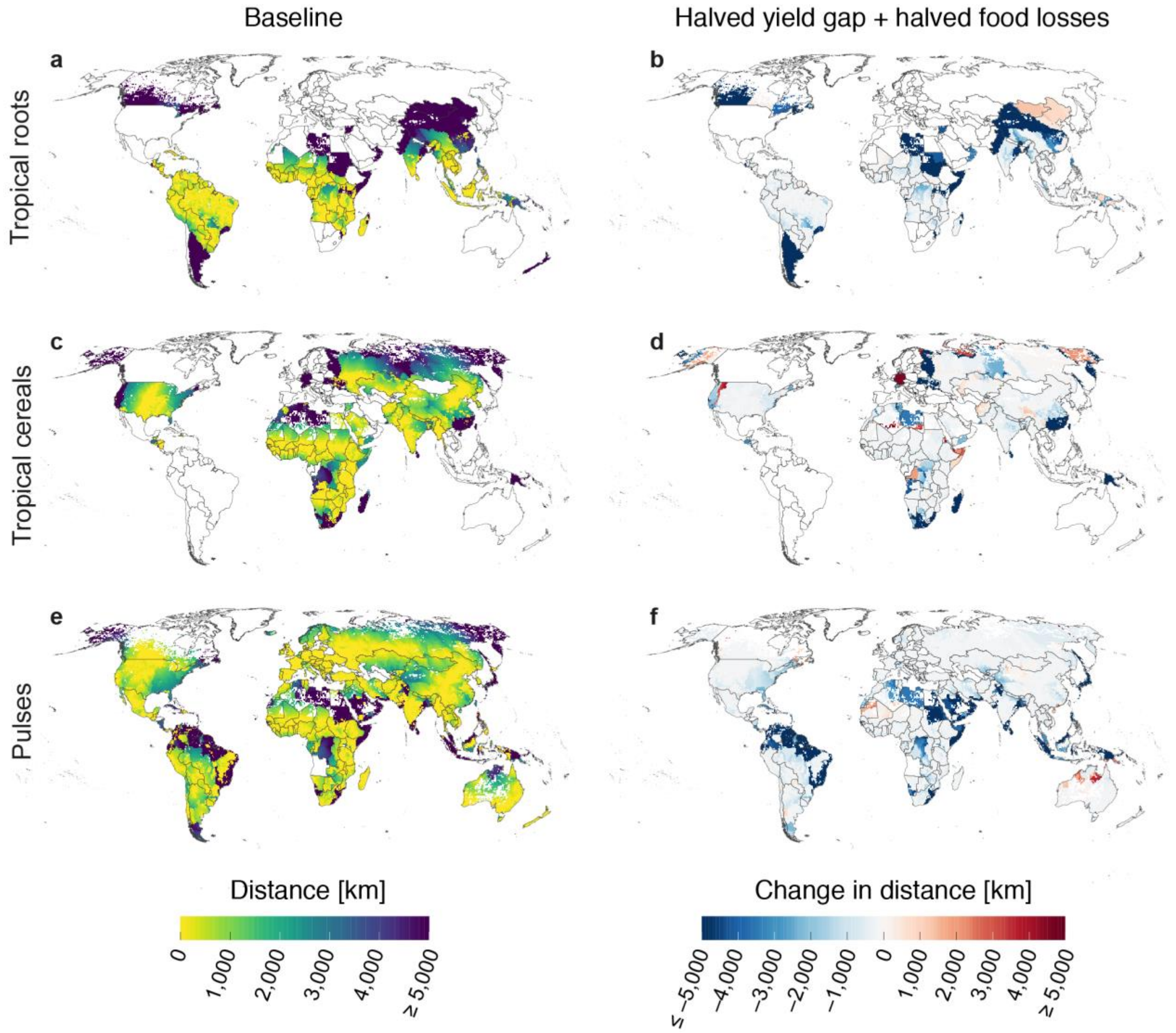
Supplementary figures



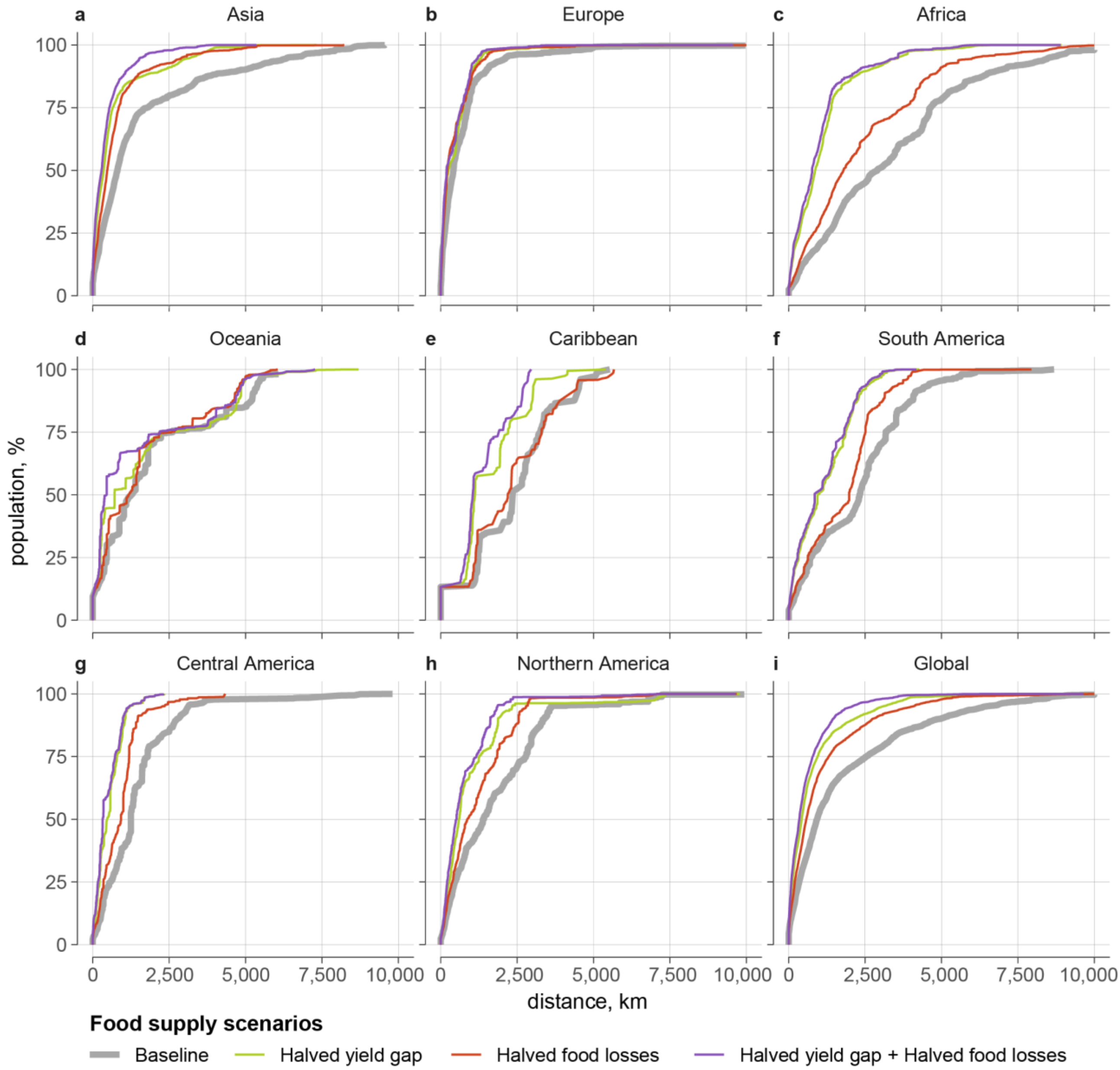
Supplementary Fig. 1: **FAOSTAT region classification.** Countries without classification were assigned the value of the closest country with a FAO region classification [31].



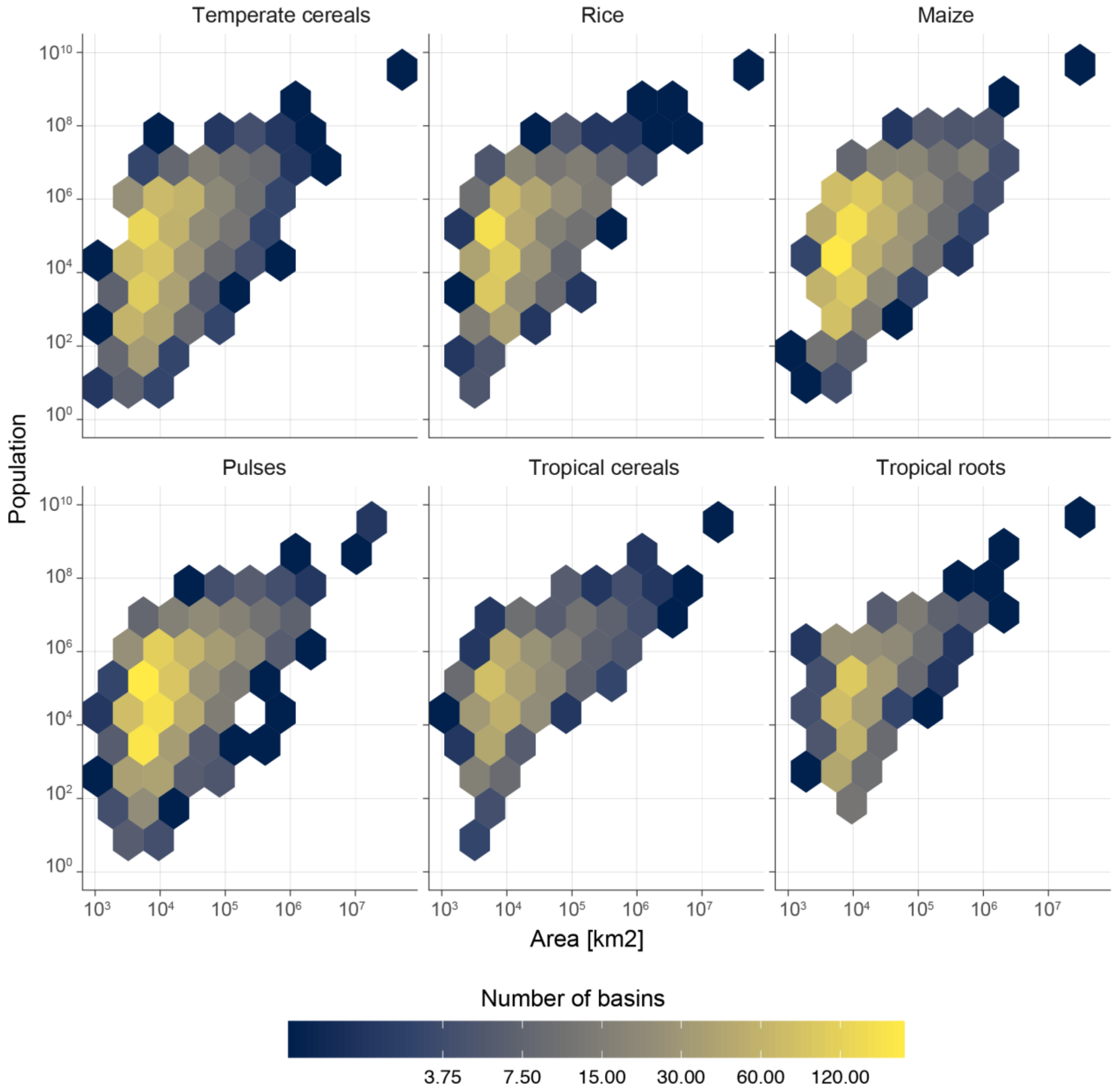
Supplementary Fig. 2: **Food supply and demand for baseline scenario**. Food supply in calories (left panels) for tropical roots (**a**), tropical cereals (**c**), and pulses (**e**). Food demand (right panels) for tropical roots (**b**), tropical cereals (**d**), and pulses (**f**). Both on a 30 arc-min grid (~50 km x 50 km at the equator).



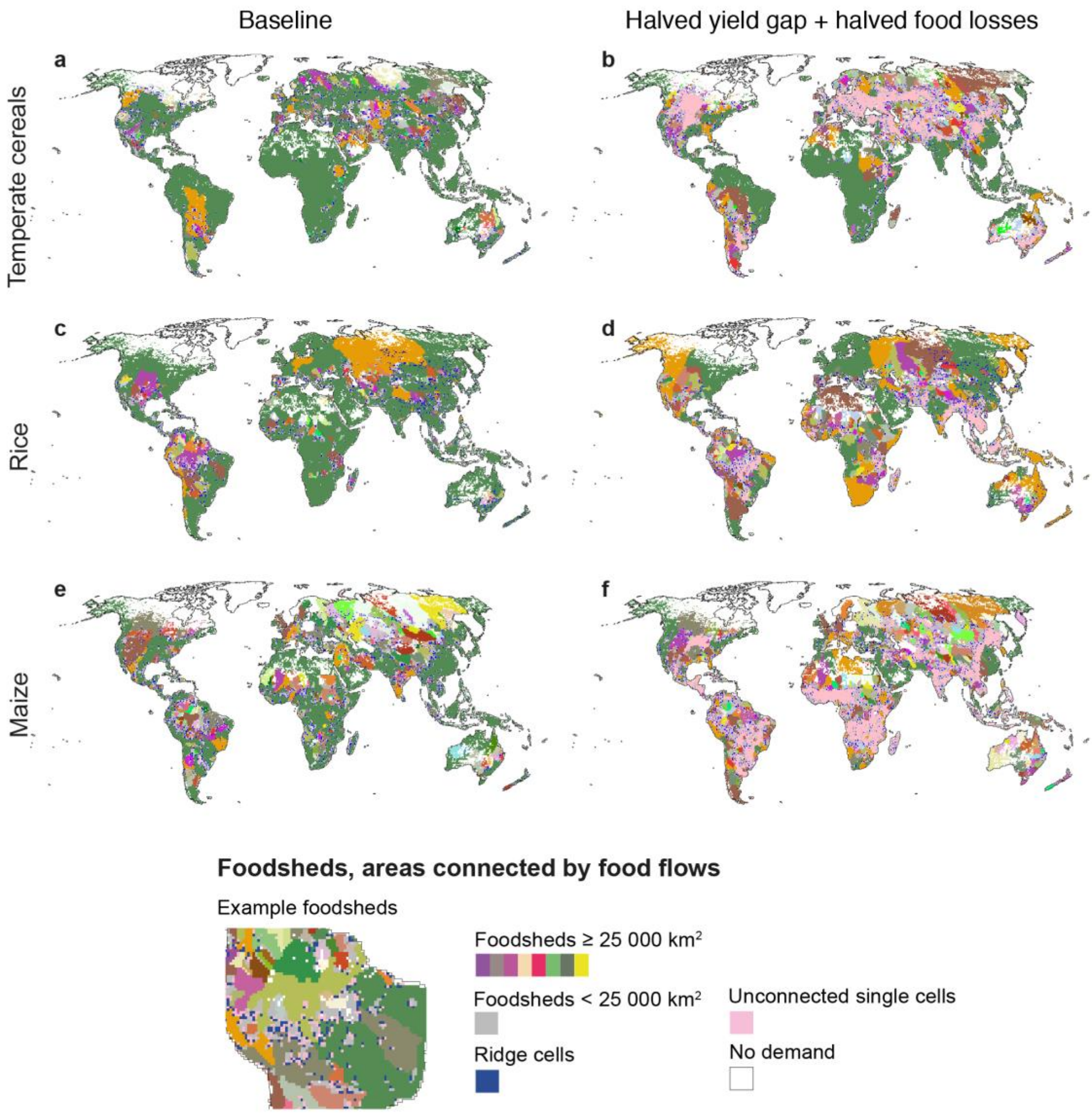
Supplementary Fig. 3: **Optimised simulated distance from food production to consumption.** Panels on the left show the distance under baseline conditions and panels on the right show the change in distance when the HalfLoss + HalfYieldGap scenario is considered. Distances needed to satisfy food demand for tropical roots (**a**), tropical cereals (**c**), and pulses (**e**). Change in distance relative to the baseline scenario for tropical roots (**b**), tropical cereals (**d**), and pulses (**f**). Food flows are determined by minimizing a friction surface capturing transport travel time costs.



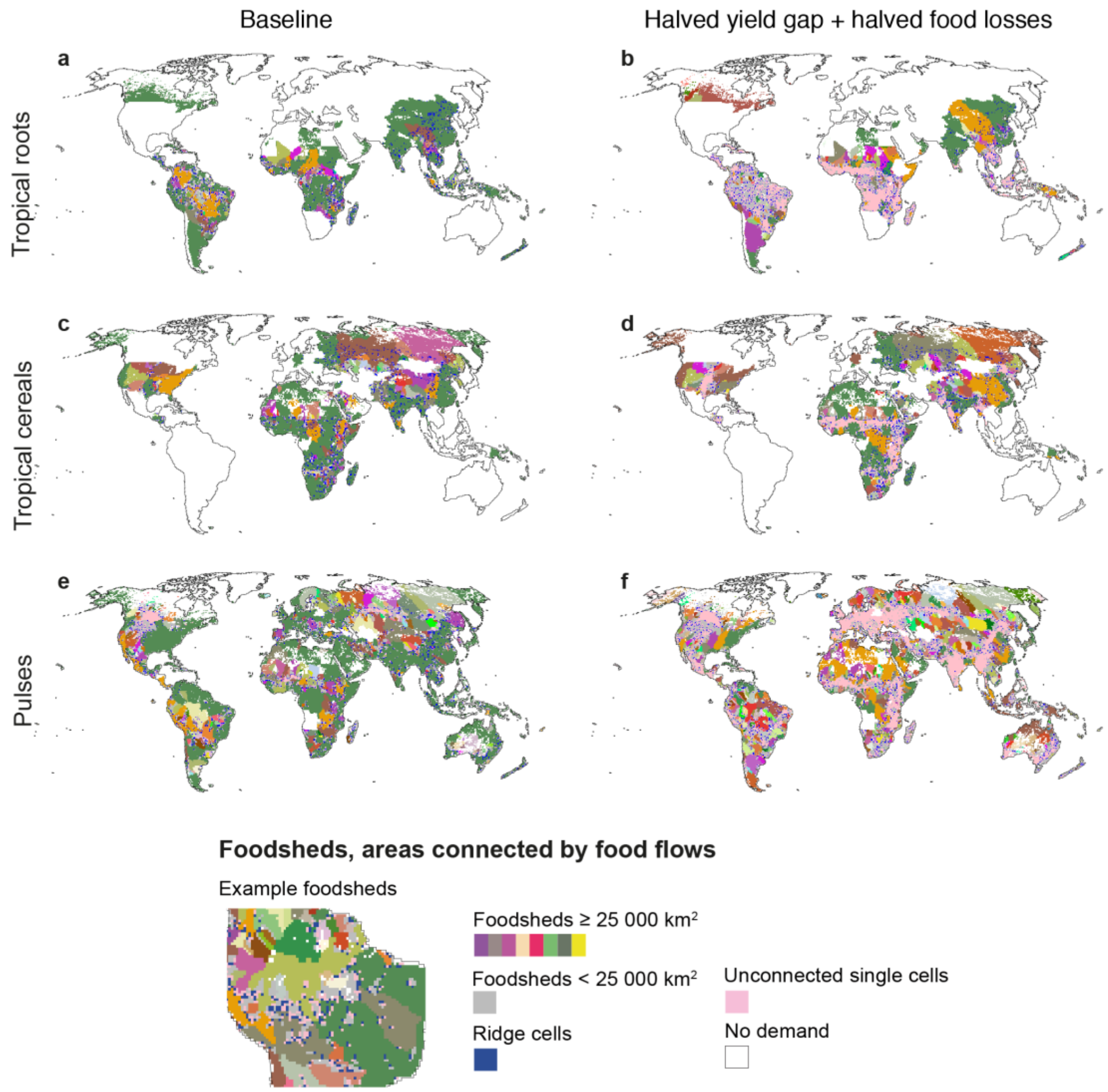
Supplementary Fig. 4: **Cumulative population distributions for the weighted mean distance calculated across six crops for baseline and three scenarios.** The distributions are aggregated globally and over eight regions. See region delineation in Supplementary Fig. 1.



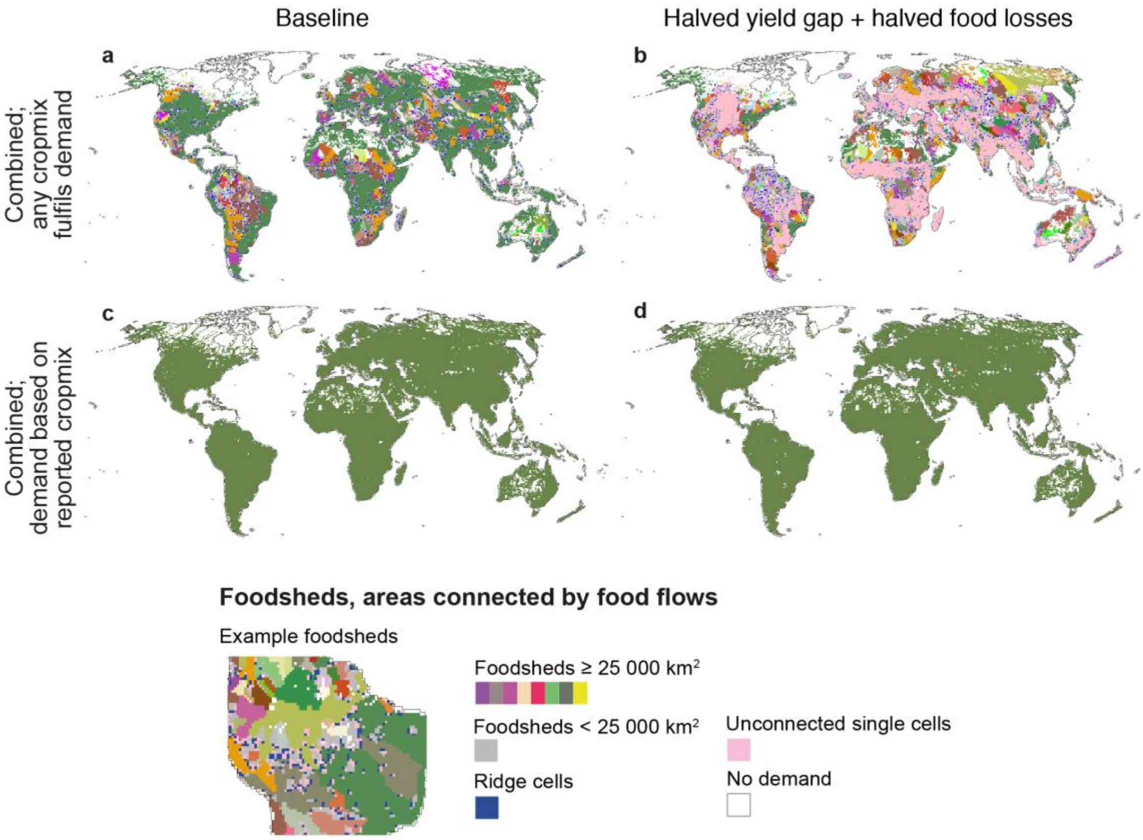
Supplementary Fig. 5: **Number of foodsheds and population for temperate cereals (a), rice (b), maize (c), pulses (d), tropical cereals (e) and tropical roots (f)**. Note the logarithmic scale in both area and population axes.



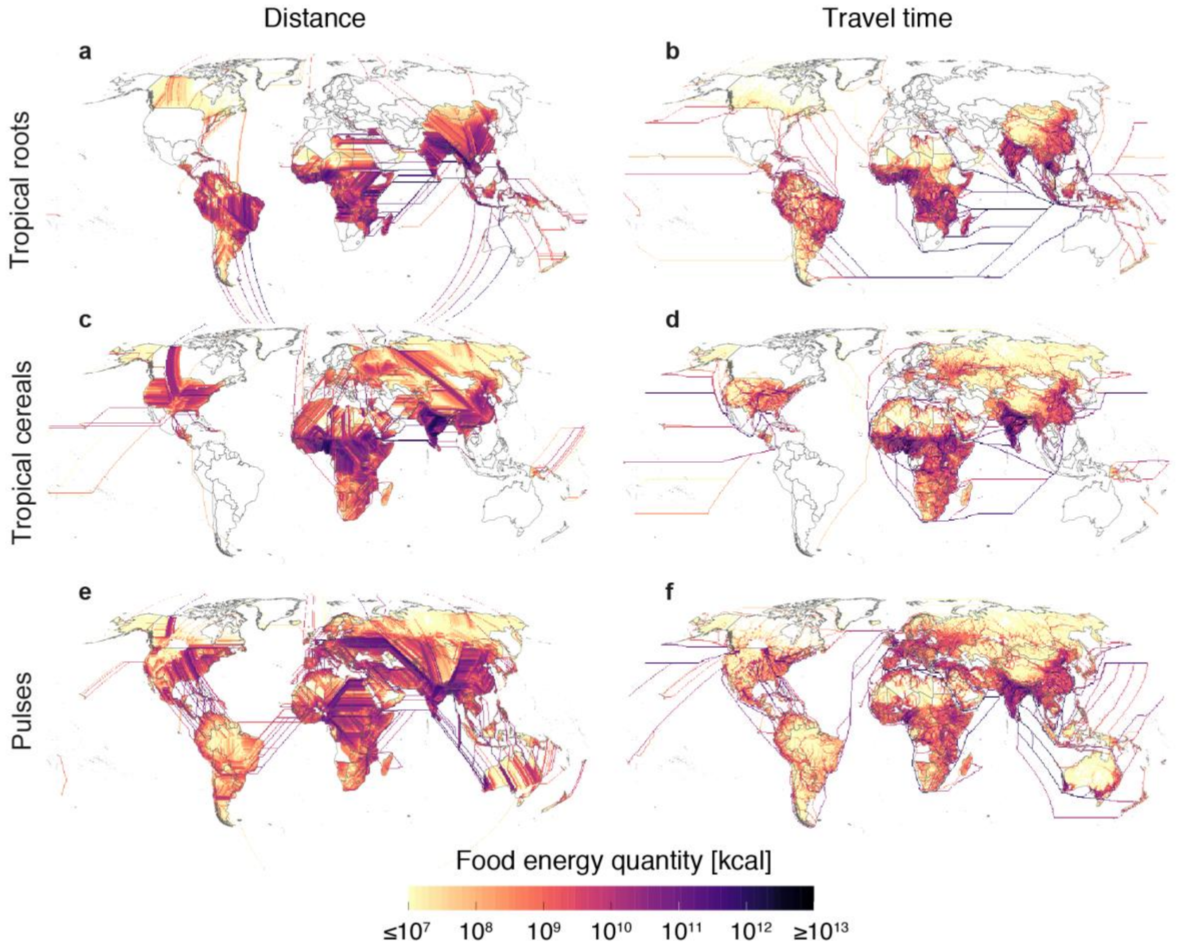
Supplementary Fig. 6: **Foodsheds for baseline (left column) and HalfLoss + HalfYieldGap scenarios (right column), temperate cereals (a, b), rice (c,d), and maize (e,f).** For each of the crops, all areas with the same colour belong to the same foodsheds. All foodsheds with area less than 25,000 km_2 are marked with grey colour. Ridge cells are surplus production cells that have no incoming flows and connect to two or more foodsheds. Unconnected single cells (pink) are cells that are able to fulfil demand within the cell but are not connected to any other cell. Foodsheds also include non-production demand-only areas. Food flows were determined by minimizing a friction surface capturing transport travel time costs.



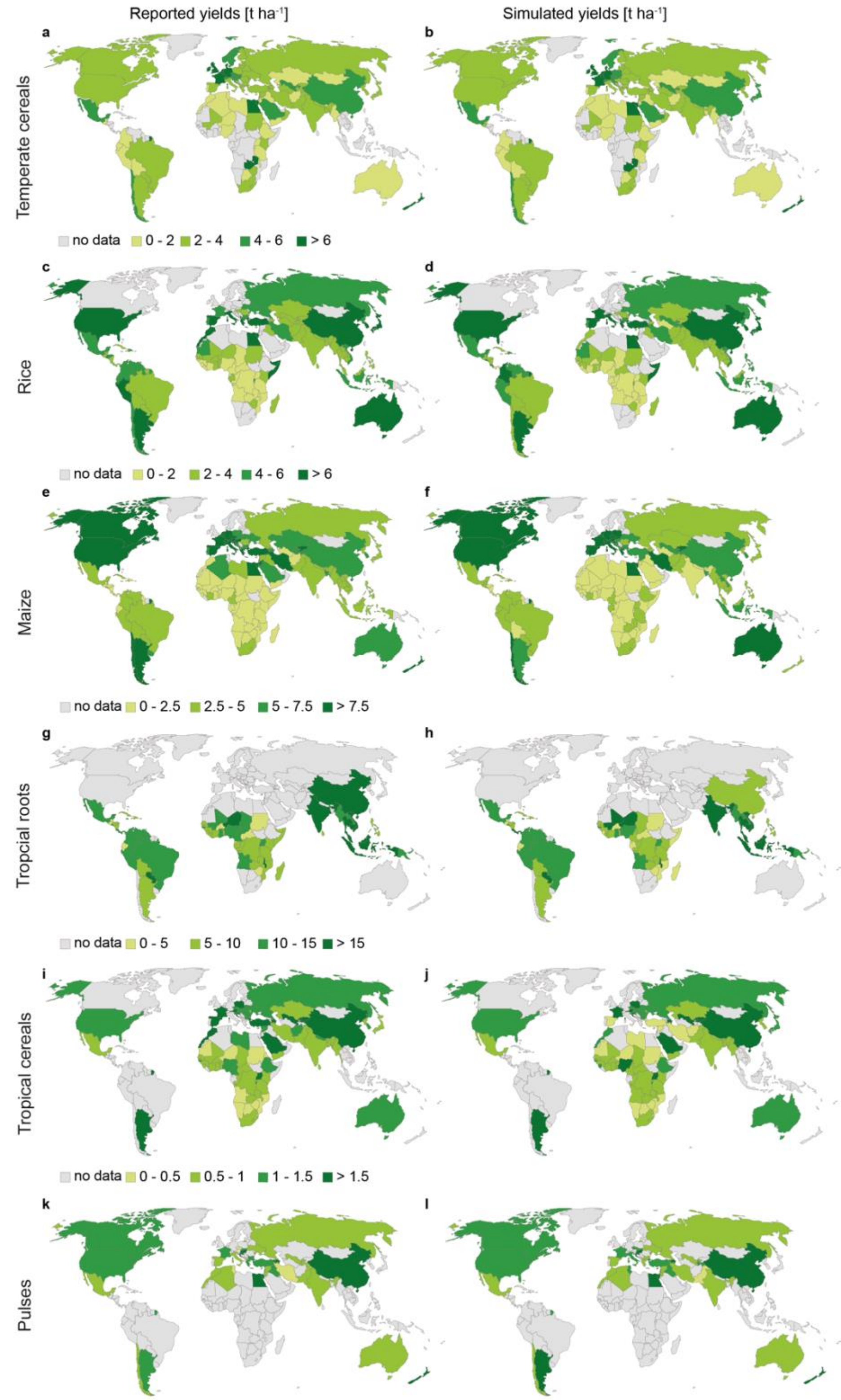
Supplementary Fig. 7: **Foodsheds for baseline (left column) and HalfLoss + HalfYieldGap scenarios (right column), tropical roots (a, b), tropical cereals (c,d), and pulses (e,f).** For each of the crops, all areas with the same colour belong to the same foodsheds. All foodsheds with area less than 25,000 km^2 are marked with grey colour. Ridge cells are surplus production cells that have no incoming flows and connect to two or more foodsheds. Unconnected single cells (pink) are cells with sufficient food production to satisfy demand in the cell but with no connection to any other cells. Foodsheds also include non-production demand-only areas. Food flows were determined by minimizing a friction surface capturing transport travel time costs.



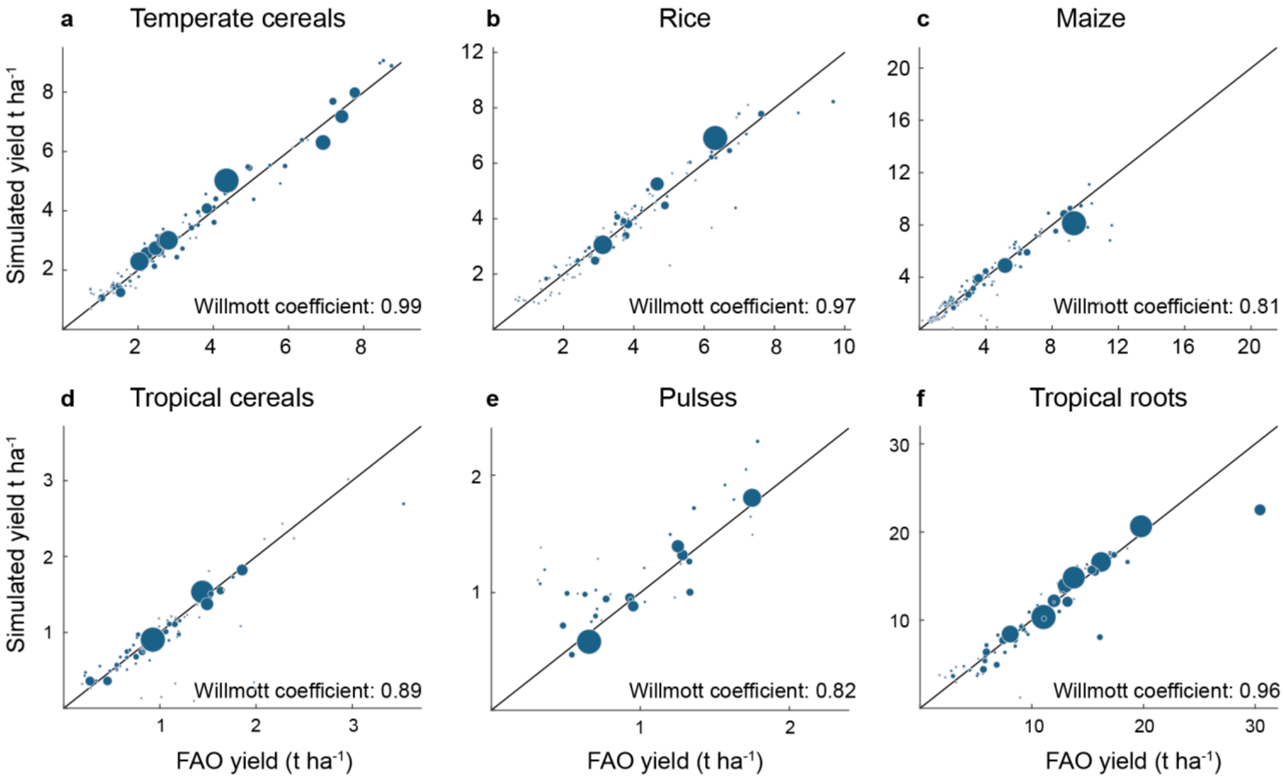
Supplementary Fig. 8: **Foodsheds for baseline (left column) and HalfLoss + HalfYieldGap scenarios (right column), combined when any crop mix can fulfil the reported kcal demand of the six assessed crops (a, b), combined when demand is based on reported crop mix (c,d).** For each of the crops, all areas with the same colour belong to the same foodsheds. All foodsheds with area less than 25,000 km^2 are marked with grey colour. Ridge cells are surplus production cells that have no incoming flows and connect to two or more foodsheds. Unconnected single cells (pink) are cells with sufficient food production to satisfy demand in the cell but with no connection to any other cells. Foodsheds also include non-production demand-only areas. Food flows were determined by minimizing a friction surface capturing transport travel time costs.



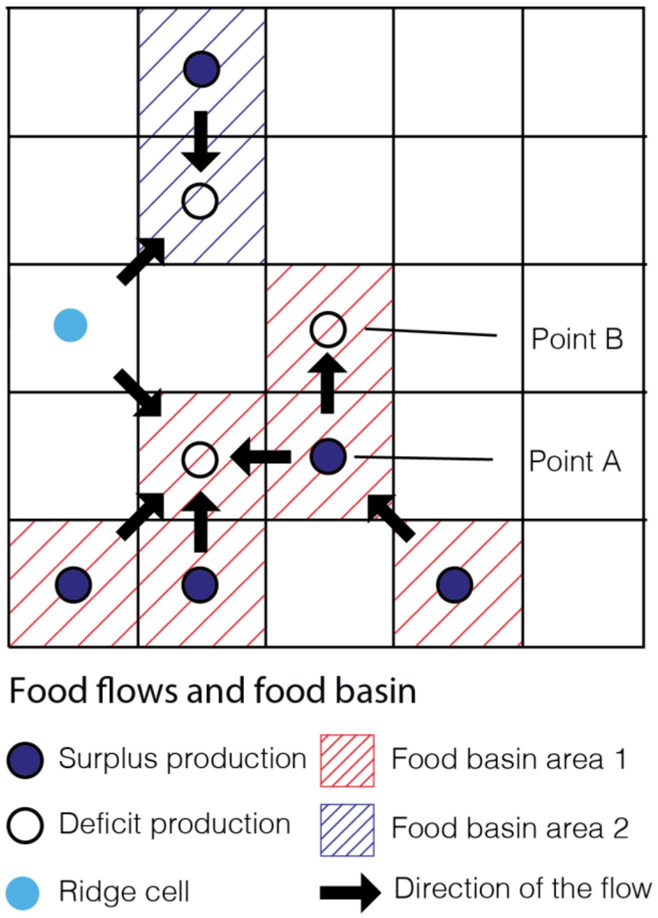
Supplementary Fig. 9: **The impact of friction surfaces on optimised food flows.** Flow routes using cell centroid distance as friction surface in the optimisation (left panels) for tropical roots (**a**), tropical cereals (**c**), and pulses (**e**). Flow routes when using transport travel time cost (see Methods) as a friction surface in the optimisation (right panels) for temperate cereals (**b**), rice (**d**), and maize (**f**). The optimisation is done on a 30 arc-min grid (~50 km x 50 km at the equator).



Supplementary Fig. 10: **Reported[37] and LPJmL[45] simulated (calibrated) yield surface for temperate cereals (a,b), rice (c,d), maize (e,f), tropical roots (g,h), tropical cereals (i,j), and pulses (k,l), averaged over the years 2001–2010.** The calibration and evaluation are done by Heino et al.[46].



Supplementary Fig. 11: **Country level scatter plots for calibration results[46] for temperate cereals (a), rice (b), maize (c), tropical cereals (d), pulses (e), and tropical roots (f) – averaged over the years 2001–2010.** The x-axis shows the yield reported in FAOSTAT and the y-axis shows LPJmL simulated yields. The Willmott coefficient[59] shows the goodness of fit of the calibration, similarly to Fader et al.[60]. The point sizes signify the average production for the country within the time period.



Supplementary Fig. 12: **Conceptual illustration of the optimisation framework, in which deficits are satisfied by flows from cells with surplus production.** Ridge cells are surplus production cells without incoming flows that connect two or more foodsheds. Food flows are routed from surplus production cells to cells without sufficient production. These flows, however, can also traverse several surplus cells, if that particular route provides the smallest friction to satisfy demand in the deficit cells, e.g. due to transport infrastructure.