Supplementary Information for:
"How much meat should we eat? Improved estimates accounting for food system dynamics influencing water use"

Figures

Figure S1: Change in land use from current diets to 20% reduced livestock using the objectives a) MinET and b) MinDiff. Land use for MinStress is not shown, as the objective is not sensitive to changes in rainfed agriculture and therefore optimized land use patterns are not well-defined. Changes in livestock production are shown in Figure S3.

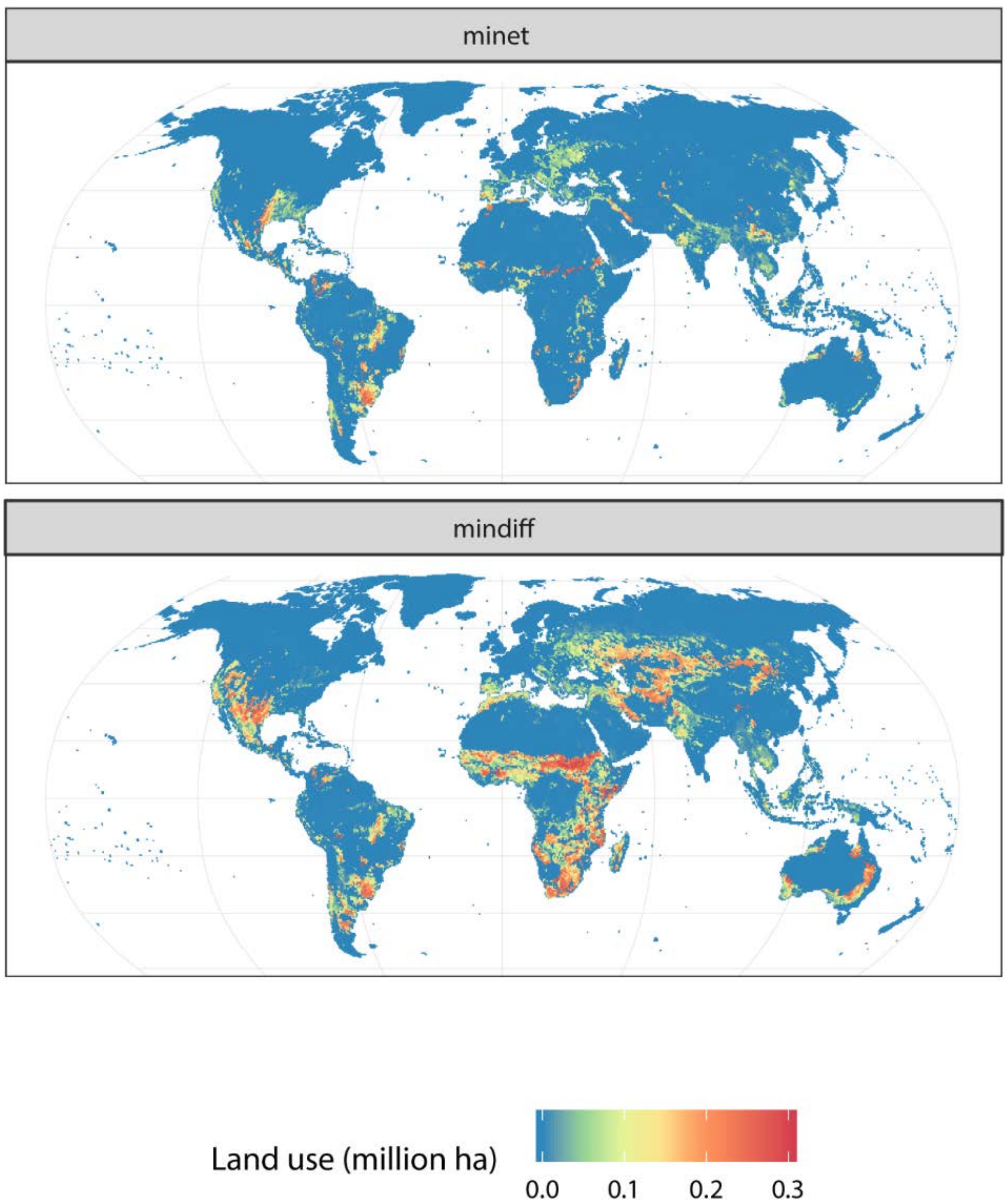


Figure S2: Calibration of LPJmL modelled vs. FAO recorded yields as described by Fader et al[1] (reproduced with permission). Data on yields for Aalto OptoFood was taken from the LPJmL model. Scatterplot of LPJmL-simulated yields (averages over 1999–2003) versus reported yields (t DM ha^{-1}) ((a) wheat and (b) maize). The bubbles indicate the relative size of the harvested area in the respective country according to FAO.)

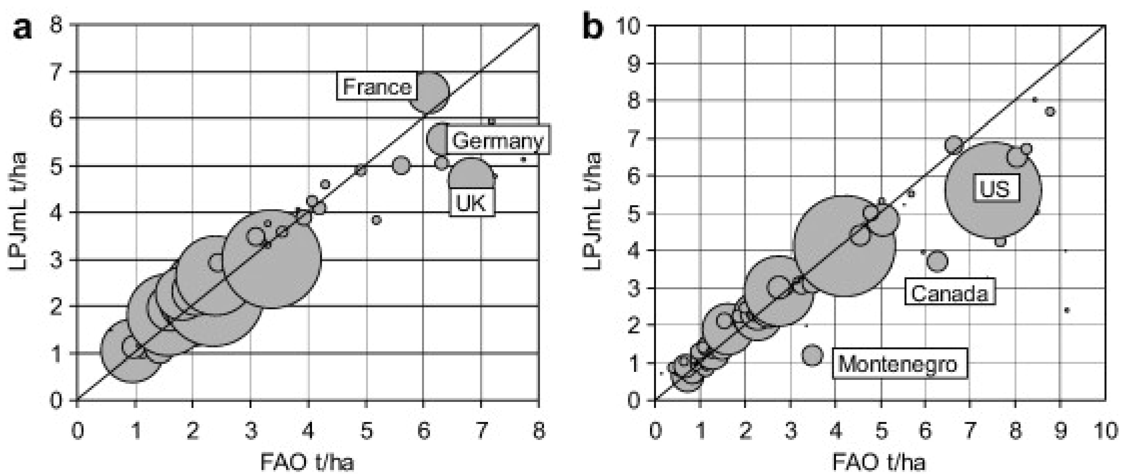
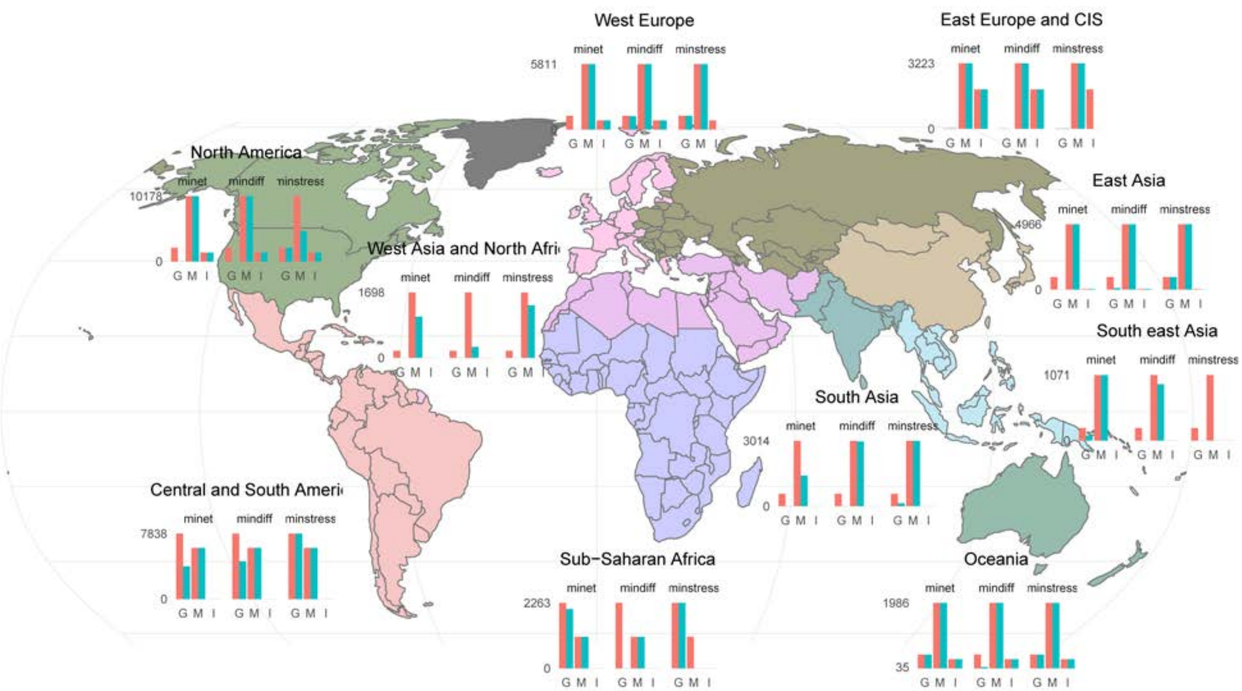
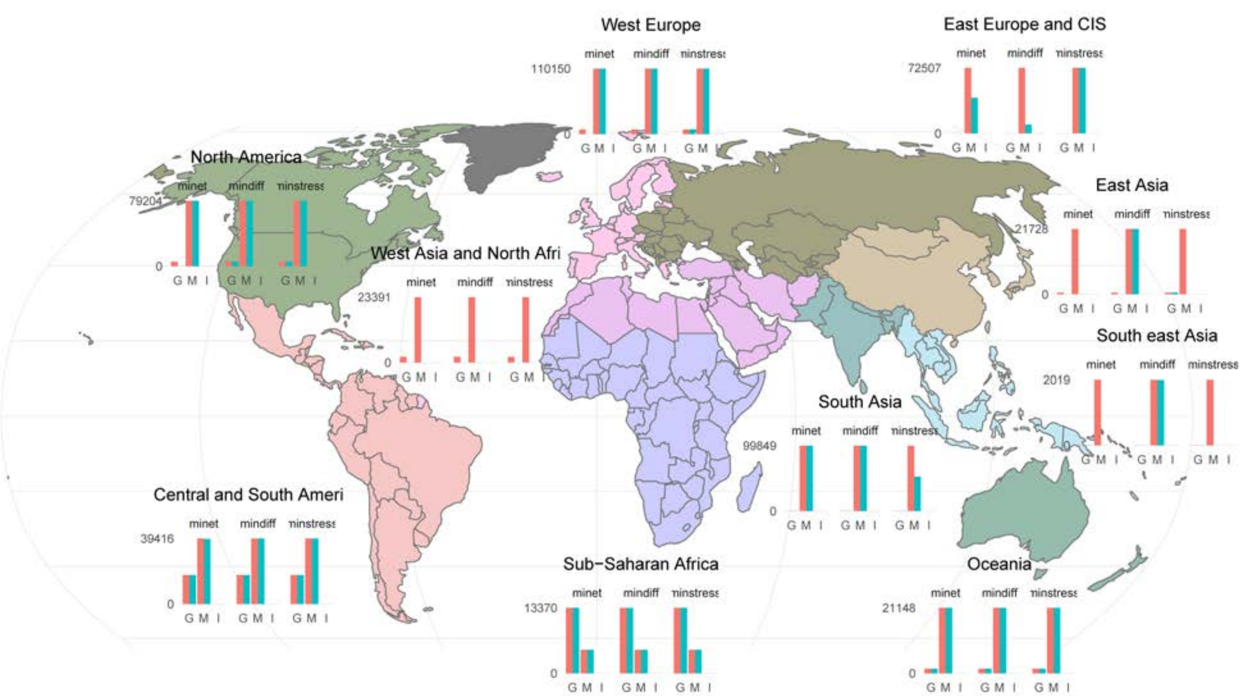


Figure S3 (S3A-F) Maps of livestock production regions and regional livestock production for current and 20% reduced livestock protein optimized with MinET, MinDiff and MinStress, for A) beef, B) milk, C) chicken, D) eggs, E) pork, F) sheep and goats. G, M and I correspond to grazing, mixed and industrial production systems respectively. Changes in crop land use are shown in Figure S1.

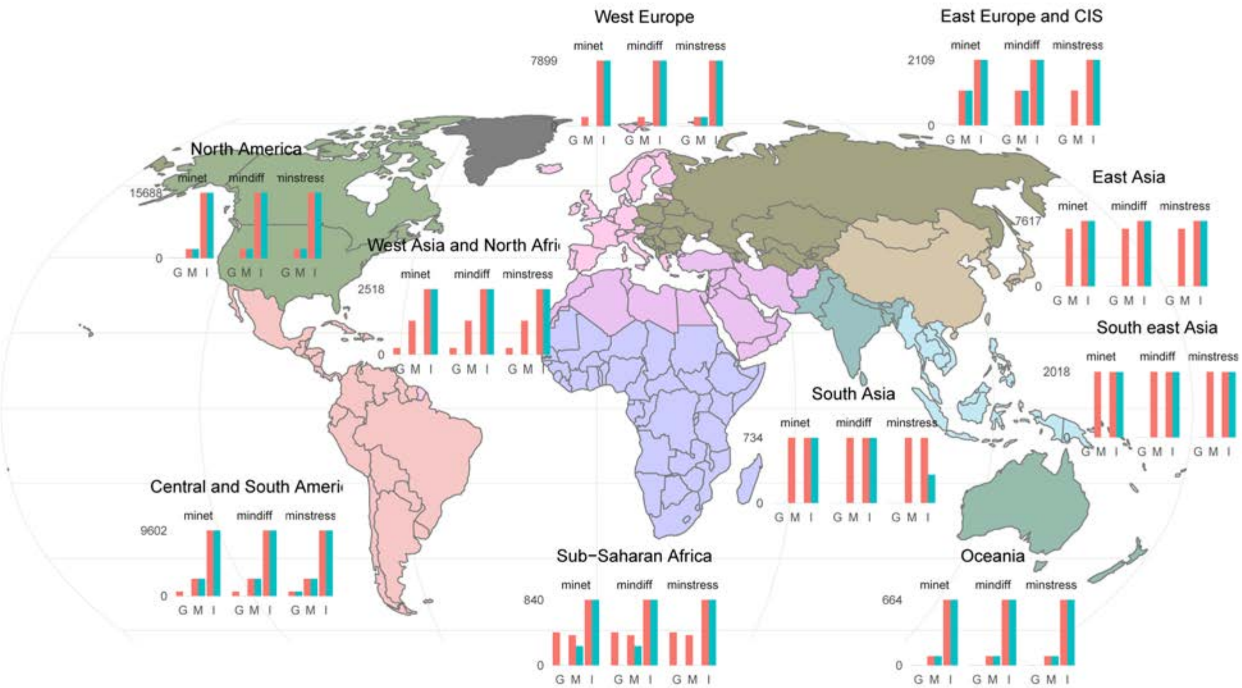
S3A Beef



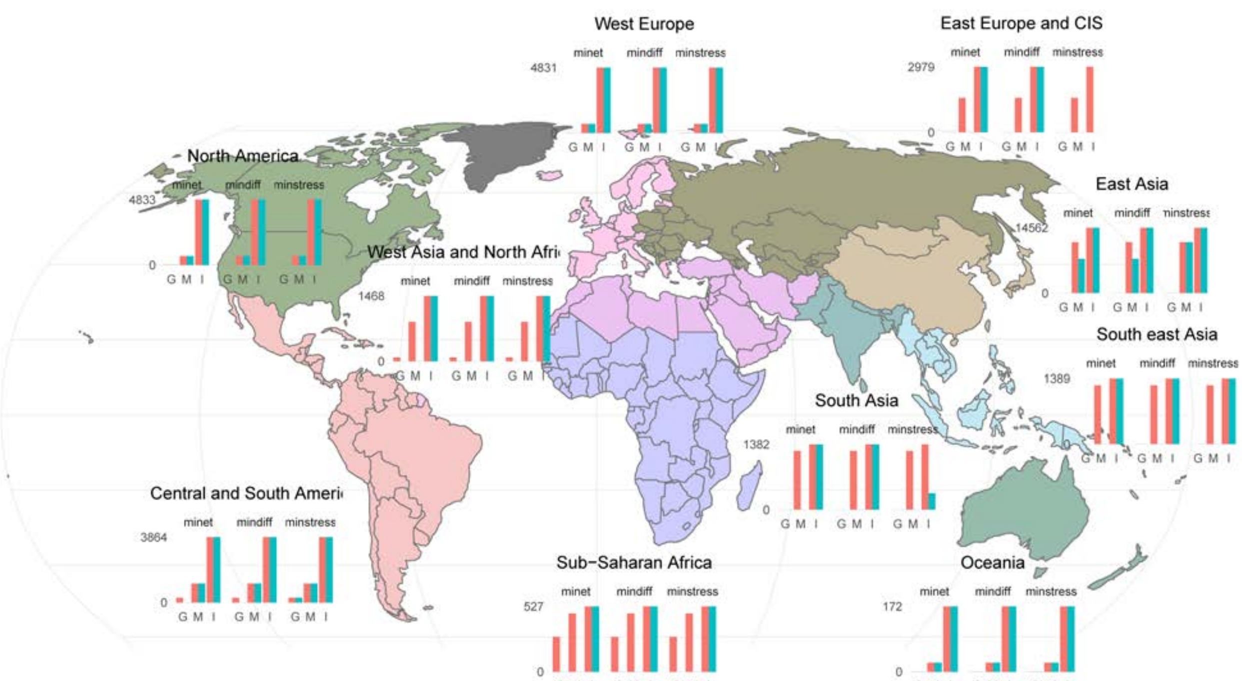
S3B Milk



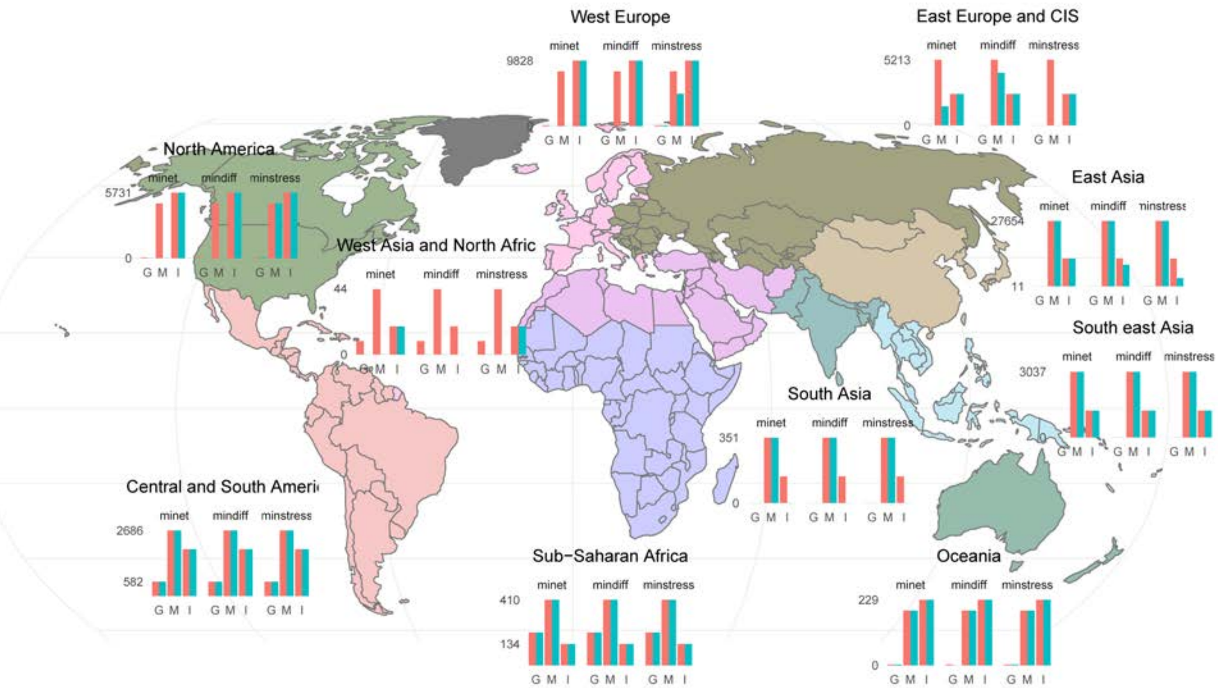
S3C Chicken



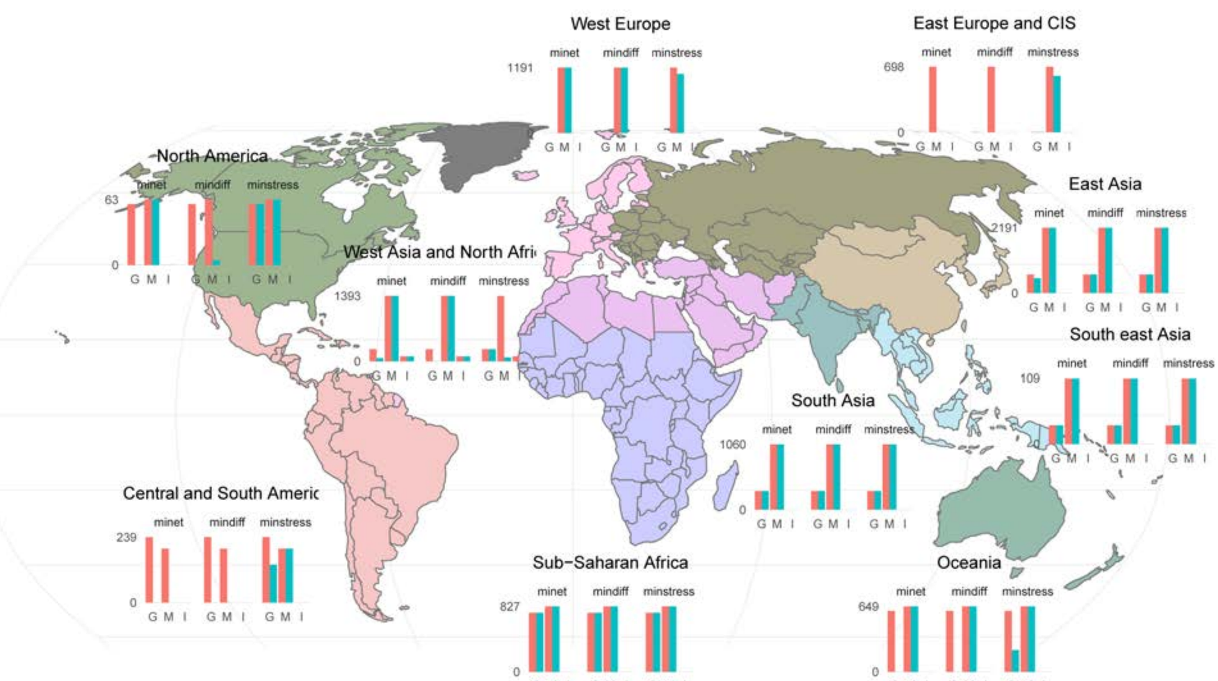
S3D Eggs



S3E Pork



S3F Sheep and goats



Tables

Table S1 Cake percentage: percentage of oilseed crop production yielding press cake, rather than oil (see "Model calibration" below)

Crop Functional Type	Cake percentage
1-7 (temperate cereals, rice, maize, tropical cereals, pulses, temperate roots, tropical roots)	0%
8 sunflower	42%
9 soybean	67%
10 groundnuts	28%
11 rapeseed	51%
12 sugarcane	0%
13 others	6%

Table S2: Properties of the scenarios explored – key parameters and optimization constraints.

Parameter/constraint	Effect
Dietary food demand constraints	Production goal of the optimisation. Current diet determines crop and livestock demand. Livestock products are replaced by pulses and soy with approximately equivalent calories and equal protein content.
Food losses	Current (no loss reduction)
Non-food crop production	Crop production used for other than food or feed is kept constant, based on FAO FBS utilisation
Harvestable area constraints	Within a grid cell, total land use for all crops must be less than or equal to the current agricultural land use. Additionally, irrigation is limited to areas currently irrigated.
Cropland reallocation constraints "stickiness"	Changes from current cropland distribution are restricted. Only changing production can be moved. Crops in increasing demand cannot be removed from any cell, those in decreasing demand cannot be added to any cell.
Irrigation water constraints	LPJmL simulation with current irrigation – excluding nonlocal and non-renewable water ("ILIM" scenario[2]) Total water abstraction for irrigation upstream for a cell cannot be increased. Water saved upstream may increase use downstream.
Water availability for MinStress	Maximum discharge within a basin, shared amongst all cells
Feed demand constraints	Determined dynamically within the optimisation. Roughage feed must be produced within the same region it is used.
Livestock "stickiness" constraints	Production cannot increase in any region when global demand decreases.
Current feed from non-food grade sources	70%
Contribution of crop residues to roughage	Max 40%
Grassland considered non-arable	Varies between scenarios, 75% (50%, 0%) of the cells closest to evapotranspiration by natural vegetation
Trade limitations	Food demand is to be satisfied globally. It is not known what trade restrictions would stay in place in the face of large demand changes. Roughage is used within its production region due to its low value.

Table S3 Quantity of livestock, pulses and soy in diet at each diet change level, and corresponding total protein and kilocalories in the average global diet. Note these quantities are influenced by calibration to LPJmL production (see Model Calibration below).

Diet change level		**Livestock**			**Pulses**	**Soy**	**Average diet**	
% of live-stock protein replaced	**% of protein from livestock**	g/cap/d	protein g/cap/d	kcal/cap/d	g/cap/d	g/cap/d	protein g/cap/d	kcal/cap/d
0	**34.2**	344	22	380	13	11	64	2157
10	**30.8**	309	20	342	17	14	64	2148
20	**27.4**	275	18	304	21	18	64	2138
30	**24.0**	241	15	266	25	21	64	2129
40	**20.5**	206	13	228	29	24	64	2120
50	**17.1**	172	11	190	33	28	64	2111
60	**13.7**	137	9	152	37	31	64	2102
70	**10.3**	103	7	114	41	34	64	2092
80	**6.8**	69	4	76	45	38	64	2083
90	**3.4**	34	2	38	49	41	64	2074
100	**0.0**	0	0	0	53	44	64	2065

Aalto OptoFood model equations

The Aalto OptoFood model is a linear programming optimization model. A less technical description is provided in the Method. Here, a mathematical description is given for decision variables, objective functions and constraints, and additionally the data calibration of the model.

Table S4: Summary of optimisation problem and data sources

Model element	**Approach**
Solver	Linear programming, Interior Point method in MATLAB[3]
Decision variables	1: Harvested area of crops and grass (LPJmL CFTs, spatial land use allocation at 0.5 degree resolution) 2: Livestock production (categories from ref[4], mass units, divided by geographical region, animal species and production system)
Optimisation objectives	MinET: Minimise evapotranspiration (ET) MinStress: Minimise average use-to-availability across basins MinDiff: Minimise absolute deviation from natural ET
Constraints	Demand, Harvestable area, Spatial re-allocation “stickiness”, Water availability
Scenario variation	**Description**
Diet change	Varied from current diet by replacing livestock products with pulses and soy down towards exclusively vegetarian diet *Current* refers specifically to the current diet
Data	**Source**
Crop properties	Yield, evapotranspiration and water consumption: LPJmL
Cropland and irrigation infrastructure	LPJmL standard land use model[5]
Cropland suitability	FAO GAEZ Suitability Indices[6]
Livestock production (baseline)	FAOSTAT Food and Commodity Balance Sheets
Livestock feed consumption (baseline)	FAOSTAT Food and Commodity Balance Sheets
Livestock feed conversion ratios	UNESCO Value of Water report 48[4]
Livestock regional production system distribution	World livestock production systems[7]
Population data	HYDE database[8]
Water availability	LPJmL ILIM[2] simulation (irrigation limited to local renewable water resources)
River routing network	STN-30p[9], as used by LPJmL ILIM simulation
Food losses and waste	FAO report: Global food losses and food waste: Extent, causes and prevention[10]

Common symbols

Symbol	Definition	Unit
$i \in I$	Cell index	
$c \in C$	Crop index, member of Crop Functional Types	
$u \in U$	Irrigation type, member of Irrigation Types	
$a_{i,c,u}$	Harvested area decision variable (cell, crop)	1000 ha
$a_{i,c,u}^{0}$	Current harvested area for each cell and crop	1000 ha
d_c	Demand for a crop product (food and other use)	1000 t
d_c^F	Food demand for crop	1000 t
$l \in L$	Livestock product	
$s \in S$	Livestock production system	
$r \in R$	Livestock production region	
d_l	Demand for a livestock product	1000 t
$p_{l,s,r}$	Livestock production	1000 t
$E_{i,c,u}$	Evapotranspiration	Mm
$Y_{i,c,u}$	Crop Yield	t/ha
$Y_{i,c,u}^{resid}$	Yield of Crop residue	t/ha (dry weight)
Y_i^{rough}	Yield of roughage crops (grass, CFT 14)	t/ha (dry weight)
$F_{l,s,r}$	Feed conversion ratio	kg/kg product
$W_{i,c}$	Water withdrawal for irrigation	mm
$R_{i,c}$	Return flow from irrigation	mm
α	Share of yield component currently used as food (oil for oilcrops)	0≤ α ≤1
φ	Share of crop yield satisfying food grade requirements	0≤ φ ≤1

Decision variables

1. the harvested area $a_{i,c,u}$ of each of the LPJ crop functional types (CFT) *c* in the each raster cell *c*, using a given irrigation type *u* (none, flood, sprinkler, drip) in thousands of hectares
2. the production amount $p_{l,s,r}$ of each of the included animal product groups in thousands of tons

Objective functions

MinET

The objective is to minimise total evapotranspiration E_{tot} over the area *I* covered by the raster cells. The coefficient for each area decision variable is the modelled evapotranspiration of the CFT in mm for that cell. The multiplication results in units of 10^4 m^3. For all other decision variables, the coefficients are zero.

$$Min(E_{tot}) = \sum_{i \epsilon I, c \in C, u \in U} E_{i,c,u} \cdot a_{i,c,u}$$

MinDiff

The objective is to minimise global difference between the evapotranspiration of the crops and the natural vegetation E_{tot}^{d}, over the area covered by the raster cells. The coefficient for each area decision variable is the absolute difference between modelled evapotranspiration of the CFT ($E_{i,c,u}$) and the natural vegetation (E_{i}^{nat}) for that cell, in mm.

$$Min\left(E_{tot}^{d}\right) = \sum_{i\epsilon I, c\in C, u\in U} \left|E_{i,c,u} - E_{i}^{nat}\right| \cdot a_{i,c,u}$$

MinStress

The objective is to minimise average water stress over all basins globally ($WSI^{average}$). The coefficient for each area decision variable is the blue water evapotranspiration modelled by LPJmL ($E_{i,c,u}^{blue}$) over the sum of water available within basin (A_B).

$$Min(WSI^{average}) = \frac{1}{n_B}\sum_{B}\sum_{i\epsilon B, c\in C, u\in U} \frac{(E_{i,c,u}^{blue} \cdot a_{i,c,u})}{A_B}$$

Constraints

Harvestable area

The harvestable area constraints only allow crops to be allocated to current agricultural land. For each cell *i*, the total harvested area across crops is limited to the current total harvested area ($a_{i,c}^{0}$).

$$\sum_{c\in C, u\in U} a_{i,c,u} \leq \sum_{c\in C, u\in U} a_{i,c,u}^{0}$$

Crop allocation is also limited by the availability of irrigation for each cell *i* and irrigation type *u*. The area irrigated using each one of three irrigation systems (flood, sprinkler, drip) cannot exceed the area equipped for that system.

$$\sum_{c\in C} a_{i,c,u} \leq \sum_{c\in C} a_{i,c,u}^{0}$$

Water constraints

Water availability for irrigation is limited by the current water use modelled using the ILIM[2] simulation in LPJmL for irrigation, but water saved by irrigation reduction can be made available downstream. For each cell *i*, total withdrawals therefore include local ($W_{i,c}$, $W_{i(ILIM)}$) and upstream withdrawals (W_{i}^{U}, $W_{i(ILIM)}^{U}$), after accounting for return flows (R_{i}^{U}, $R_{i(ILIM)}^{U}$).

$$\sum_{c\in C} W_{i,c} + W_{i,c}^{U} - R_{i,c}^{U} \leq W_{i(ILIM)} + W_{i(ILIM)}^{U} - R_{i(ILIM)}^{U}$$

Livestock roughage feed constraints

Roughage feed demand is calculated dynamically based on the livestock production allocated to a particular region and production system ($p_{l,s,r}$), and it needs to be met by production of roughage (CFT14) and crop residues within that region.

For each livestock region r,

$$\sum_{\substack{l\in L \\ s\in S \\ i\in I^r}} -Y_{i,c}^{resid} a_{i,c,u} - Y^{rough} a_{i,c=14,u} + F_{l,s,r}^{rough} \cdot p_{l,s,r} \leq 0,$$

To ensure a maximum amount of crop residues and minimum amount of grass, the shared denoted by grass_share of the roughage needs to be met by grass alone.

$$\sum_{\substack{l\in L \\ s\in S \\ i\in I^r}} -Y^{rough} a_{i,c=14,u} + \ grass\_share \cdot F_r^{rough} \cdot p_{l,s,r} \leq 0$$

Demand constraints

Human demand (food and other uses) is provided externally for each of the food crops and livestock products. Concentrate feed demand is generated endogenously by animal production decision variables and feed conversion ratios and feed composition data (FCR).

1. Production of crops must fulfil both human demand and feed demand globally, divided into two components (α for oil, β for cake) for oilcrops, all other crops are only α (see Model calibration and Table S1).

$$\sum_{i,c,u,l,s,r} Y_{i,c} \cdot \alpha_c a_{i,c,u} - F_{l,s,r}^{conc,\alpha} p_{l,s,r} \geq d_{c,\alpha}$$

$$\sum_{i,c,u,l,s,r} Y_{i,c} \cdot (1-\alpha_c) a_{i,c,u} - \ F_{l,s,r}^{conc,\beta} p_{l,s,r} \geq d_{c,\beta}$$

2. Food demand must be fulfilled by production of food grade crops (with share of production φ).

$$\sum_{i,c} Y_{i,c} \cdot \alpha_c \cdot \varphi \geq d_{c,\alpha}^F$$

$$\sum_{i,c} Y_{i,c} \cdot (1-\alpha_c) \cdot \varphi \geq d_{c,\beta}^F$$

3. Livestock demand constraints enforce that the production of each livestock product *l* satisfies human demand.

$$\sum_{\substack{r\in R \\ s\in S}} p_{l,s,r} \geq d_l$$

When diet change reduces livestock product demand (new d_l), human crop demand for pulses and soy is increased accordingly (see Table S3), replacing initial food demand (d^{F0}) with new (d^{Fn}). For soybean, the full crop is then also used as food, so the β component is also increased (d_β^{Fn}).

$$d_\alpha = d_\alpha^0 - d^{F0} + d^{Fn}$$

$$d_\beta = d_\beta^{Fn}$$

Stickiness constraints

Cropland cannot be removed in any grid cell from crops with increasing demand, and no additional land can be allocated to crops with decreasing demand. These constraints are implemented as upper and lower bounds for the decision variables:

$$a_{i,c}^{inc} \geq a_{i,c}^0$$

for crops with increasing demand, and

$$a_{i,c}^{dec} \leq a_{i,c}^0$$

for crops with decreasing demand.

Similarly, livestock species production within a region and production system can only decrease as diet change progresses (no scenarios with increasing livestock production are explored), so that no region/system can replace production removed elsewhere:

$$p_{l,s,r} \leq p_{l,s,r}^0$$

Model calibration

Aalto OptoFood combines data from three major datasets that need to be made compatible: the LPJmL model, the FAO Food Balance Sheets (FBS), and feed conversion ratios from Mekonnen & Hoekstra[4]. Given uncertainties in all three datasets, the priority is to create a coherent scenario where current consumption of food, feed and other users can be met by production with the current land use model – this provides a solid basis for then modifying the food system through optimisation.

LPJmL yields are calibrated to FAO yields (see Figure S2), but residual discrepancies and differences in land use cause a discrepancy in production. As the Aalto OptoFood model relies on yields and water use from LPJmL, production in LPJmL is used as a ground truth (i.e. we operate in an "LPJmL world"). The LPJmL production, OrigProd_LPJ, is used as a baseline, determined by the modelled yields and the current land use pattern.

OrigProd_LPJ = sum(yields * areas)

To ensure production can meet consumption demands (from FAO FBS), production is then allocated to different consumption uses, i.e. all current production is assumed to be required to fulfil the current demand. The calibration first calculates livestock feed consumption, and then estimates direct human demand to fit the difference.

Livestock calibration to LPJ feed availability

LPJmL provides no information about consumption or use for feed vs food. Given LPJmL production is used as the baseline, rather than using FAO utilisation of food and feed, we use FAO Commodity balances[11] to determine the proportion of each CFT used for feed, i.e. the *feed share*. This is applied to LPJmL production to obtain baseline feed availability, LPJFeedAmount.

LPJFeedAmount = OrigProd_LPJ * FeedShare

Livestock production then needs to be calibrated to fit this feed amount. Using FAO livestock production, feed requirements (MH_Feed_required) are calculated using estimated production system distribution[7] and feed conversion ratios.

MH_Feed_required = FAOLSProd * ProdSysPerc * FCR * FeedComposition

Where

FAOLSProd = Global FAO livestock production

ProdSysPerc = percentage of production in given production system and regions according to Sere and Steinfeld[7]

FCR = Feed conversion ratio[4]

FeedComposition = Percentage of (concentrate) feed for given CFT

Note here that feed conversion ratios are taken as ground truth as FAO data on livestock production and feed consumption are not split by production system. As a result of this decision, FAO feed consumption will not exactly match feed consumption in the model. To make the ProdSysPerc from Sere and Steinfeld compatible with Mekonnen and Hoekstra FCRs, sheep and goat production has also been moved to a mixed production system for Western Europe.

Feed requirements and production need to match. Livestock demand is scaled to meet production available. Specifically the baseline livestock production for the optimisation (OrigLSProd, before diet changes) is obtained by scaling FAO LS production:

OrigLSProd = FAOLSProd * (LPJFeedAmount / MH_Feed_required)

Human demand calibration:

The human consumption consists of the part of production used for food and all "other" uses in the FAO FBS. In simulation, food is calculated from diet and population (pop). The baseline diet (FSQ, food supply quantity per capita for each country) is calculated based on the FBS, and used to identify the food component of LPJmL production. A similar approach is used to identify "other" uses, and for feed, above:

OrigFood_FAO = FSQ_FAO * pop

OrigFood_LPJ = OrigProd_LPJ * (OrigFood_FAO / OrigProd_FAO)

OrigOtherUse_LPJ = OrigProd_LPJ * (OrigOtherUse_FAO / OrigProd_FAO)

Given that we operate in an LPJ world, the baseline average diet composition at global scale therefore differs from that calculated directly from the FAO FBS.

Decomposition of production using the FAO FBS guarantees that production and demand match in the baseline case.

OrigProd_LPJ = OrigFood_LPJ + OrigOtherUse_LPJ + LPJFeedAmount

For oilseed crops, production is additionally split to identify the part resulting in press cake rather than oil, which is therefore currently used mostly as feed. This is calculated from the FAO Commodity Balances as the mass percentage of production of oilseeds in the CFT (0% for non-oilseed crops, 100% for oilseed crops, in between for CFT 13 - "other" crops), and the percentage of cake supply within that production (see Table S1).

CakePerc = oilseedshare * (cakesupply_of_cft/oilseedsupply_of_cft)

Total demand (cake and non-cake) needs to be reduced by the amount used for feed (LPJFeedAmount in the baseline scenario). Feed is preferentially taken from the cake component, with any remaining feed requirements met from the non-cake component (when FeedShare > CakePerc).

NonCakeFeedPerc = max(0,(FeedShare - CakePerc) / FeedShare)

OrigDemandNoncake = (1-CakePerc) * OrigProd_LPJ – NonCakeFeedPerc * LPJFeedAmount

OrigDemandCake = CakePerc * OrigProd_LPJ – (1 – NonCakeFeedPerc) * LPJFeedAmount

In the baseline, demand and production are then also equal based on this decomposition.

OrigProd_LPJ = OrigDemandNonCake + OrigDemandCake + LPJFeedAmount

In the baseline case, production for food is taken from the non-cake component, so we have

OrigDemandOtherUseNonCake = OrigDemandNonCake – OrigFood_LPJ

Such that we know

OrigDemandOtherUseNonCake + OrigDemandCake = OrigOtherUse_LPJ

Combining the food/other/feed, and noncake/cake/feed decompositions, we have:

OrigProd_LPJ = OrigFood_LPJ + LPJFeedAmount + OrigDemandOtherUseNonCake + OrigDemandCake

References

1. Fader, M., Rost, S., Müller, C., Bondeau, A. & Gerten, D. Virtual water content of temperate cereals and maize: Present and potential future patterns. *J. Hydrol.* **384**, 218–231 (2010).
2. Rost, S. *et al.* Agricultural green and blue water consumption and its influence on the global water system. *Water Resour. Res.* **44**, n/a-n/a (2008).
3. Linear Programming Algorithms - MATLAB & Simulink - MathWorks Nordic. Available at: https://se.mathworks.com/help/optim/ug/linear-programming-algorithms.html. (Accessed: 21st March 2019)
4. Mekonnen, M. M. & Hoekstra, A. Y. *The green, blue and grey water footprint of farm animals and animal products*. (Unesco-IHE Institute for Water Education, 2011).
5. Portmann, F. T., Siebert, S. & Döll, P. MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Glob. Biogeochem. Cycles* **24**, (2010).
6. Fischer, G. *et al. Global Agro-Ecological Zones (GAEZ v3.0)*. (IIASA, 2016).
7. Sere, C. & Steinfeld, H. LEAD Digital library - World livestock production systems: current status, issues and trends. (1995). Available at: http://www.fao.org/WAIRDOCS/LEAD/X6101E/X6101E00.HTM. (Accessed: 7th May 2018)
8. Goldewijk, K. Anthropogenic land-use estimates for the Holocene; HYDE 3.2. (2017). doi:10.17026/dans-25g-gez3
9. Vörösmarty, C. J., Fekete, B. M., Meybeck, M. & Lammers, R. B. Global system of rivers: Its role in organizing continental land mass and defining land-to-ocean linkages. *Glob. Biogeochem. Cycles* **14**, 599–621 (2000).

10. Gustavsson, J., Cederberg, C., Sonesson, U., van Otterdijk, R. & Meybeck, A. *Global food losses and food waste: Extent, causes and prevention*. (Food and agriculture organization of the United Nations (FAO) Rome, 2011).

11. FAO. FAOSTAT. *FAOSTAT* (2019). Available at: http://www.fao.org/faostat/en/#data/. (Accessed: 5th January 2019)