

Supporting Information for

Tiered maize and wheat nutrient removal coefficients estimated from available data.

C. I. Ludemann^{1*}, R. Hijbeek¹, M.P. van Loon¹, T.S. Murrell², A. Dobermann³, and

M.K. van Ittersum¹

¹Wageningen University & Research, Plant Production Systems group,

Wageningen, Netherlands.²African Plant Nutrition Institute and Mohammed VI

Polytechnic University, Ben Guerir, Morocco.³International Fertilizer Association, Paris,

France.

Contents of this file

Table S1: Countries assigned to world regions from Tier 2 data.

Text S1: Further methodology explanation on analysis of data to support Tier 3 estimates.

Table S2: Prediction accuracy for Tier 3 maize coefficients

Table S3: Prediction accuracy for Tier 3 wheat coefficients

Table S4: Tier 1 and 2 coefficients

Table S5: Mean and standard deviation for data to support Tier 3 estimates per region

Table S6: Harvest index models per region

Figure S1: Overview of experimental locations included in data to support Tier 3 estimates

References

Table S1. Countries assigned to world regions from Tier 2 data.

Count	World Region (as stated in original dataset)	Countries assigned to world region
1	Atlantic Zone of Costa Rica	Costa Rica
2	Australia	Australia
3	Australia, New South Wales South East	Australia
4	Australia, New South Wales South West	Australia
5	Australia, South Australia Lower EP	Australia
6	Australia, South Australia Mid North	Australia
7	Australia, South Australia Murray Mallee	Australia
8	Australia, South Australia South East	Australia
9	Australia, South Australia Upper EP	Australia
10	Australia, South Australia Yorke Penn.	Australia
11	Australia, Victoria Mallee	Australia

12	Australia, Victoria North Central	Australia
13	Australia, Victoria North East	Australia
14	Australia, Victoria Wimmera	Australia
15	Canada	Canada
16	Central America	Belize, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama
17	East Africa	Burundi, Chagos Archipelago, Comoros, Djibouti, Eritrea, Ethiopia, Ethiopia PDR, French Southern Territories, Kenya, Madagascar, Malawi, Mauritius, Mayotte, Réunion, Rwanda, Seychelles, Somalia, South Sudan, Uganda, United Republic of Tanzania, Zambia, Zimbabwe
18	East Asia	China, China, Hong Kong SAR, China, Taiwan Province of, Democratic People's Republic of Korea, Japan, Mongolia, Republic of Korea
19	Eastern Europe	Belarus, Bulgaria, Czechia, Czechoslovakia, Hungary, Poland, Republic of Moldova, Romania, Russian Federation, Slovakia, Ukraine, USSR
20	Ethiopia	Ethiopia
21	European Union and United Kingdom	Austria, Belgium, Belgium-Luxembourg, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, United Kingdom
22	Former USSR	Armenia, Azerbaijan, Belorussia, Belarus, Estonia, Georgia, Kazakhstan, Kirgiziya, Kyrgystan, Latvia, Lithuania, Moldova, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan
23	Ghana	Ghana
24	India	India
25	India, Odisha	India
26	Japan	Japan
27	Kenya	Kenya
28	Malawi	Malawi
29	Mali	Mali
30	Middle East	Akrotiri and Dhekelia, Bahrain, Cyprus, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Palestine, Qatar, Saudi Arabia, Syria, Turkey, United Arab Emirates, Yemen
31	North Africa	Algeria, Egypt, Libya, Morocco, Sudan, Sudan (former), Tunisia, Western Sahara

32	North America	Canada, United States, United States Minor Outlying Islands, Virgin Islands (US).
33	North Macedonia	North Macedonia
34	Oceania	American Samoa, Australia, Christmas Island, Cocos (Keeling) Islands, Cook Islands, Fiji, French Polynesia, Guam, Heard and McDonald Islands, Johnston Island, Kiribati, Marshall islands, Micronesia, Midway Island, Nauru, New Caledonia, New Zealand, Niue, Norfolk Island, Northern Mariana Islands, Pacific Islands Trust Territory, Palau, Papua New Guinea, Pitcairn, Polynesia, Samoa, Solomon Islands, Tokelau, Tonga, Tuvalu, United States Minor Outlying Islands, Vanawatu, Wake Island, Wallis and Futuna Islands
35	OECD Europe	Austria, Belgium, Belgium-Luxembourg, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom
36	South America	Argentina, Bolivia (Plurinational State of), Bouvet Island, Brazil, Chile, Colombia, Ecuador, Falkland Islands (Malvinas), French Guyana, Paraguay, Peru
37	South and South East Asia	Brunei Darussalam, Cambodia, China, China, Hong Kong SAR, China, Macao SAR, China, Mainland, China, Taiwan Province of, Democratic People's Republic of Korea, Indonesia, Japan, Lao People's Democratic Republic, Malaysia, Mongolia, Myanmar, Philippines, Republic of Korea, Singapore, Thailand, Timor-Leste, Viet Nam
38	South Asia	Afghanistan, Bangladesh, Bhutan, India, Iran Islamic Republic of, Maldives, Nepal, Pakistan, Sri Lanka
39	Southeast Asia	Brunei Darussalam, Cambodia, China, China, Hong Kong SAR, China, Macao SAR, China, Mainland, China, Taiwan Province of, Democratic People's Republic of Korea, Indonesia, Japan, Lao People's Democratic Republic, Malaysia, Mongolia, Myanmar, Philippines, Republic of Korea, Singapore, Thailand, Timor-Leste, Viet Nam
40	Southern Africa	Botswana, Eswatini, Swaziland, Lesotho, Namibia, South Africa
41	United States of America	United States of America
42	USA	United States of America
43	West Africa	Benin, Burkino Faso, Cabo Verde, Cote d'Ivoire, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Saint Helena, Ascension and Tristan da Cunha, Senegal, Sierra Leone, Togo

Text S1. *Further explanation of material and method section: Analysis of data to support Tier 3 estimates.*

Data collation and standardizing

Tier 3 estimates of maize and wheat crop harvest index and nutrient concentration of crop products and crop residues were based on the general methodology described by Ludemann et al. (2022). The Tier 3 data were checked for outliers, and standardized following the method described by Ludemann et al. (2022). Outliers were visually assessed using the “ggplot2” package in R version 4.1.0 (R Core Team, 2020) (herein referred to as ‘R’). Harvest index data points were deemed to be biologically infeasible and therefore outliers if they were above 0.65 based on maximum harvest index values in a review by Unkovich et al. (2010). Moreover, Tier 3 data from the two sources described in Table 1 were merged into one dataset based on weighting mean values by the replicates and years of data from each source as described by Ludemann et al. (2022). For example, if a mean value for harvest index was available over a 3-year period from a field experiment with four replicates, a value of $3 \times 4 = 12$ replicates was used to weight that mean.

Crop product yields on a fresh weight basis were converted to a dry matter basis using the fresh grain to grain dry matter conversion factor of 0.845 kg dry matter per kg (fresh) grain weight for maize (McKevith, 2004) and 0.88 kg dry matter per kg (fresh) grain weight for wheat. Some data sources had grain and/or residue yield and nutrient uptake of grain and/or residue available, but did not include grain and/or residue nutrient concentration. In these cases the nutrient concentration was interpolated by dividing the uptake of the applicable nutrient for that plant component by the dry matter yield of that plant component. If crop product protein concentration data were available, but crop product N concentration data were not available, the grain protein concentration of maize was divided by a factor of 6.25 (WHO and FAO, 2007) to obtain an estimate of the crop product N concentration.

Following the method described by Ludemann et al. (2022) non-water limited yield potential (Y_p) and water limited yield potential (Y_w) values (van Ittersum et al., 2013) were included in the analysis using data from the Global Yield Gap Atlas (www.yieldgap.org) where available. In cases where it was indicated that the site had been irrigated the Y_p was used to represent the Yield Potential (Y_{pot}), if this was not indicated, the Y_w was used to represent the Y_{pot} value.

Following Ludemann et al. (2022), both random forest and mixed-effects models were used to estimate Tier 3 estimates of harvest index and nutrient concentration of crop products and crop residues of maize.

The methodology in the present study did vary from that described by Ludemann et al. (2022) in two important respects. Firstly, more sources of data were used in the analysis with the inclusion of wheat data (with data openly available at www.croplandnutrientdata.net and Ludemann et al. (2024) as per Table 1). Secondly, the actual models used to explain variation in, and test predictive accuracy of, the crop coefficients were refined based on lessons from Ludemann et al. (2022). As a consequence there were fewer models tested, and new models were developed for explaining variation in, and testing prediction accuracy of the N, P and K nutrient concentration parameters. Further explanation for how this was performed is described in the next section.

Model selection for estimating Tier 3 crop coefficients

The best mixed-effects models for predicting harvest index, Grain N and Residue N - as described by Ludemann et al. (2022)- were used to predict Tier 3 harvest index, Grain N and Residue N coefficients for maize and wheat. Before use, the models were tested in their ability to explain variation in, and predict harvest index, Grain N and Residue N following the

methodology described by Ludemann et al. (2022). The random factor of all the mixed-effects models was region information, based on UN ‘region’ categories of countries (FAO, 2024).

The models were tested for their ability to explain variation in the dependent variable using Akaike Information Criteria (AIC) (Burnham and Anderson, 2002) and conditional Nakagawa R^2 (Nakagawa and Schielzeth, 2013) values following Ludemann et al. (2022). The explanatory power of the models was considered when choosing what models to use for testing predicting accuracy of crop nutrient concentrations. Models were then subjectively chosen based on those that were most parsimonious (most economical in terms of use of predictive variables), making biological sense (e.g. direction of coefficients should align with plant physiology) whilst still providing acceptable levels of prediction accuracy.

For brevity, only results from Scenario 1 for AIC, Nakagawa R^2 and prediction accuracies of models are presented for maize and wheat in Table S2 and S3 respectively. These show in bold the mixed-effects models that were eventually chosen for comparison in prediction accuracy compared with the random forest models.

Table S2: Linear mixed-effects models used to predict maize harvest index, nitrogen concentration of crop products (Grain N) and crop residues (Residue N), elemental phosphorus concentration of crop products (Grain P) and crop residues (Residue P), elemental potassium concentration of crop products (Grain K) and crop residues (Residue K) of maize using crop product yield (CPY), yield potential (Ypot) and/or fertilizer application rates of nitrogen (FN), elemental phosphorus (FP) and elemental potassium (FK) as predictor variables for Scenario 1. Models in bold were chosen for comparison with the prediction accuracy of random forest models.

Dependent variable	Linear mixed-effects model equation*	AIC**	Nakagawa R ² ***	Prediction accuracy R ² ****
Harvest index	~a×ln(CPY)+b	-	-	0.45
Grain N	~FN		0.13	0.10
	~FN - CPY		0.14	0.10
	~FN + Ypot	Best	0.16	0.12
Residue N	~FN		0.09	0.14
	~FN - CPY		0.13	0.16
	~FN + Ypot	Best	0.14	0.22
Grain P	~FP	Best	0.05	0.04
	~FP + CPY		0.06	0.03
	~FP - Ypot		0.23	0.06
	~FP + CPY + FN		0.06	0.03
	~ FP - Ypot - FN		0.24	0.07
Residue P	~FP		0.11	0.10
	~ FP - CPY		0.36	0.15
	~FP - Ypot		0.20	0.12
	~FP - CPY - FN		0.39	0.18
	~ FP - Ypot - FN	Best	0.32	0.20

Grain K	~ FK		0.15	0.23
	~FK - CPY		0.15	0.23
	~FK - Ypot	Best	0.17	0.25
	~FK - CPY - FN		0.15	0.23
	~FK - Ypot + FN		0.17	0.25
Residue K	~ FK		0.21	0.14
	~-FK +CPY		0.22	0.18
	~ FK - Ypot	Best	0.25	0.14
	~-FK + CPY + FN		0.22	0.18
	~ FK - Ypot + FN		0.24	0.14

*All linear mixed-effects models used the UN region as a random effect. Country categorization into regions aligned to those reported by FAO (2024). In the model for harvest index, the ‘a’ represents the factor to multiply against the natural logarithm (ln) of CPY and b represents a value to vary the plateau height of harvest index.

**Akaike Information Criterion (AIC) is a measure of the relative quality of statistical models for a set of data whereby a more negative value indicates a greater relative quality compared with another model (Burnham and Anderson, 2002). The AIC column indicates which linear mixed-effects model is best based on AIC for each sub-selection of the dataset, i.e. the sub-selection of the overall dataset used for the CPY, Ypot, FN, FP and FK models.

***The R^2 value for each model is a conditional coefficient of determination for a linear mixed-effects model based on variance of the fixed and random effects (a value between 0 and 1)(Nakagawa et al., 2017). AIC and Nakagawa R^2 values cannot be derived from non-linear models such as that used for harvest index.

**** The prediction accuracy R^2 was the coefficient of determination for the predicted values from each model compared with actual values from the ‘test’ dataset.

Table S3: Linear mixed-effects models used to predict wheat harvest index, nitrogen concentration of crop products (Grain N) and crop residues (Residue N), elemental phosphorus concentration of crop products (Grain P) and crop residues (Residue P), elemental potassium concentration of crop products (Grain K) and crop residues (Residue K) of maize using crop product yield (CPY), yield potential (Ypot) and/or fertilizer application rates of nitrogen (FN), elemental phosphorus (FP) and elemental potassium (FK) as predictor variables for Scenario 1. Models in bold were chosen for comparison with the prediction accuracy of random forest models.

Dependent variable	Linear mixed-effects model equation*	AIC**	Nakagawa R ² ***	Prediction accuracy R ² ****
Harvest index	~a×ln(CPY)+b			0.50
Grain N	~FN		0.08	0.03
	~FN - CPY		0.11	0.07
	~FN + Ypot	Best	0.08	0.03
Residue N	~FN		0.29	0.36
	~FN - CPY	Best	0.37	0.45
	~FN -Ypot		0.25	0.35
Grain P	~FP		0.41	0.15
	~FP - CPY		0.39	0.21
	~-FP + Ypot	Best	0.83	0.62
	~FP - CPY + FN		0.40	0.21
	~ -FP + Ypot - FN		0.83	0.62
Residue P	~-FP		0.36	0.33
	~ -FP + CPY		0.51	0.48
	~-FP + Ypot	Best	0.24	0.06
	~-FP + CPY - FN		0.52	0.48

	~FP + Ypot - FN		0.24	0.06
Grain K	~FK		0.07	0.01
	~ FK + CPY		0.49	0.02
	~FK - Ypot		0.40	0.09
	~FK + CPY - FN		0.50	0.03
	~FK - Ypot + FN	Best	0.41	0.08
Residue K	~ FK		0.27	0.08
	~FK -CPY		0.27	0.07
	~FK + Ypot		0.56	0.29
	~FK -CPY + FN		0.27	0.08
	~FK + Ypot + FN	Best	0.56	0.30

*All linear mixed-effects models used the UN region as a random effect. Country categorization into regions aligned to those reported by FAO (2024). In the model for harvest index, the 'a' represents the factor to multiply against the natural logarithm (ln) of CPY and b represents a value to vary the plateau height of harvest index.

**Akaike Information Criterion (AIC) is a measure of the relative quality of statistical models for a set of data whereby a more negative value indicates a greater relative quality compared with another model (Burnham and Anderson, 2002). The AIC column indicates which linear mixed-effects model is best based on AIC for each sub-selection of the dataset, i.e. the sub-selection of the overall dataset used for the CPY, Ypot, FN, FP and FK models.

***The R^2 value for each model is a conditional coefficient of determination for a linear mixed-effects model based on variance of the fixed and random effects (a value between 0 and 1)(Nakagawa et al., 2017). AIC and Nakagawa R^2 values cannot be derived from non-linear models such as that used for harvest index.

**** The prediction accuracy R^2 was the coefficient of determination for the predicted values from each model compared with actual values from the 'test' dataset.

Table S4. Tier 1 and Tier 2 coefficients of the seven nutrient removal coefficients for each original reported region for maize and wheat, whereby HI= harvest index, Grain N= nitrogen concentration of crop product, Residue N = nitrogen concentration of crop residue, Grain P= phosphorus concentration of crop product, Residue P = phosphorus concentration of crop residue, Grain K= potassium concentration of crop product, Residue K = potassium concentration of crop residue where units of concentration are % ($100 \times \text{kg elemental nutrient kg}^{-1} \text{ DM}$).

Crop	Region	Reference	Tier	HI	Grain N	Residue N	Grain P	Residue P	Grain K	Residue K
Maize	Atlantic Zone of Costa Rica	Bessembinder (1995)	2	0.34	1.54	0.83	0.34	0.15	0.46	1.49
Maize	Australia	Unkovich et al (2010)	2	0.52						
Maize	Canada	Unkovich et al (2010)	2	0.5						
Maize	United States of America	Unkovich et al (2010)	2	0.53						
Maize	World	Unkovich et al (2010)	1		1.84					
Maize	Ethiopia	Kumssa et al (2022)	2				0.26		0.36	
Maize	Malawi	Kumssa et al (2022)	2				0.29		0.4	
Maize	European Union and United Kingdom	Panagos et al (2022)	2	0.5			0.24			
Maize	India	Roy (2016)	2		2.99		0.59		2.72	
Maize	North America	Roy (2016)	2		1.5	0.99	0.24	0.16	0.76	1.2
Maize	World	Roy (2016)	1	0.4		0.59		0.14		1.09
Maize	India Odisha	Swain et al (2022)	2		1.95	1.26				
Maize	South and South East Asia	Koopmans & Koppejan (2008)	2	0.47, 0.7		1.04				
Maize	United States of America	USDA (2009)	2		1.25	1.11	0.26	0.2	0.49	1.34
Maize	World	Bouwman et al (2017)	1		1.4	1.1	0.29	0.2		
Maize	World	FAO (2004)	1		1.91	1.32	0.47	0.26	0.55	2.9
Maize	World	FAO (2020)	1		1.36	1.08	0.31	0.18	0.42	2.25

Maize	World	IPCC (2019)	1	0.5		0.6				
Maize	World	Lassaletta et al	1		1.26					
Maize	World	(2014)	1							
Maize	World	Zhang et al (2021)	1		1.55		0.3			
Maize	World	Mean (standard deviation) Tier 1	1	0.45 (0.05)	1.55 (0.25)	0.94 (0.29)	0.34 (0.07)	0.20 (0.04)	0.49 (0.07)	2.08 (0.75)
Maize	World	Mean (standard deviation) Tier 2	2	0.51 (0.10)	1.85 (0.61)	1.05 (0.14)	0.32 (0.12)	0.17 (0.02)	0.87 (0.84)	1.34 (0.12)
Wheat	Australia	Unkovich et al (2010)	2	0.36						
Wheat	Canada	Unkovich et al (2010)	2	0.4						
Wheat	United States of America	Unkovich et al (2010)	2	0.48						
Wheat	World	Unkovich et al (2010)	1		2.24					
Wheat	Australia New South Wales South East	Norton (2011)	2		3.13		0.36		0.51	
Wheat	Australia New South Wales South West	Norton (2011)	2		2.52		0.27		0.42	
Wheat	Australia South	Norton (2011)	2		2.17		0.31		0.45	
Wheat	Australia Lower EP	Norton (2011)	2		2.67		0.39		0.47	
Wheat	Australia South	Norton (2011)	2		2.79		0.35		0.45	
Wheat	Australia South East	Norton (2011)	2		2.45		0.36		0.49	
Wheat	Australia South	Norton (2011)	2		2.78		0.31		0.48	
Wheat	Australia Upper EP	Norton (2011)	2		2.4		0.31		0.44	
Wheat	Australia South	Norton (2011)	2		2.42		0.31		0.43	
Wheat	Australia Yorke Penn.	Norton (2011)	2							
Wheat	Australia Victoria	Norton (2011)	2							
Wheat	Mallee	Norton (2011)	2							

Wheat	Australia Victoria North Central	Norton (2011)	2		2.42		0.29		0.42	
Wheat	Australia Victoria North East	Norton (2011)	2		2.75		0.3		0.43	
Wheat	Australia Victoria Wimmera	Norton (2011)	2		2.6		0.41		0.5	
Wheat	Ethiopia	Kumssa et al (2022)	2				0.32		0.45	
Wheat	European Union and United Kingdom	Panagos et al (2022)	2	0.5			0.37			
Wheat	India	Roy (2016)	2		2.5		0.4		2.74	
Wheat	North America	Roy (2016)	2		2.08	0.6	0.46	0.07	0.52	0.86
Wheat	World	Roy (2016)	1	0.4		0.49		0.11		1.06
Wheat	South and South East Asia	Koopmans & Koppejan (2008)	2	0.41		0.98				
Wheat	United States of America	USDA (2009)	2		2.08	0.67	0.62	0.07	0.52	0.97
Wheat	World	Bouwman et al (2017)	1		1.9		0.34			
Wheat	World	FAO (2004)	1		2.55	0.5	0.49	0.21	0.66	3.1
Wheat	World	FAO (2020)	1		2.4	1.39	0.47	0.14	0.53	1.92
Wheat	World	IPCC (2019)	1	0.43		0.6				
Wheat	World	Lassaletta et al (2014)	1		2.23					
Wheat	World	Mean (standard deviation) Tier 1		0.42 (0.02)	2.26 (0.22)	0.75 (0.37)	0.43 (0.07)	0.15 (0.04)	0.60 (0.07)	2.03 (0.84)
Wheat	World	Mean (standard deviation) Tier 2		0.43 (0.05)	2.52 (0.28)	0.75 (0.17)	0.36 (0.08)	0.07 (-)	0.61 (0.55)	0.92 (0.06)

Table S5. Mean, standard deviation and number of data for the seven nutrient removal coefficients for each United Nations region for maize and wheat using data to support Tier 3 estimates, whereby HI= harvest index, Grain N= nitrogen concentration of crop product, Residue N = nitrogen concentration of crop residue, Grain P= phosphorus concentration of crop product, Residue P = phosphorus concentration of crop residue, Grain K= potassium concentration of crop product, Residue K = potassium concentration of crop residue where units of concentration are % ($100 \times \text{kg elemental nutrient kg}^{-1} \text{ DM}$). Standard deviations are provided in parentheses while the number of data points used for each mean calculation follow “n=”.

Crop	Region	HI	Grain N	Residue N	Grain P	Residue P	Grain K	Residue K
Maize	Africa	0.39 (0.1) n=1899	1.31 (0.31) n=2048	0.67 (0.32) n=1588	0.26 (0.12) n=1922	0.07 (0.06) n=1524	0.67 (0.48) n=1775	1.81 (1.09) n=1414
Maize	Americas	0.5 (0.07) n=1853	1.32 (0.21) n=1135	0.79 (0.39) n=1121	0.27 (0.08) n=1252	0.06 (0.06) n=949	0.36 (0.07) n=992	2.27 (0.6) n=1721
Maize	Asia	0.44 (0.09) n=1204	1.37 (0.4) n=652	0.68 (0.28) n=483	0.29 (0.11) n=493	0.13 (0.13) n=414	0.42 (0.19) n=427	1.23 (0.74) n=510
Maize	Europe	- n=0	1.24 (0.36) n=61	- n=0	0.26 (0.24) n=34	- n=0	0.33 (0.13) n=25	- n=0

Maize	Oceania	0.36 (0.07) n=30	1.13 (0.13) n=30	0.68 (0.14) n=31	- n=0	- n=0	- n=0	- n=0
Wheat	Americas	0.43 (0.07) n=56	2.3 (0.37) n=326	0.58 (0.18) n=52	0.52 (0.71) n=63	0.03 (0.01) n=17	0.56 (0.1) n=23	1.63 (0.18) n=4
Wheat	Asia	0.39 (0.07) n=960	2 (0.57) n=559	0.49 (0.22) n=442	0.35 (0.12) n=627	0.13 (0.16) n=431	0.67 (0.87) n=481	1.34 (0.63) n=439
Wheat	Europe	0.46 (0.08) n=170	2.22 (0.51) n=313	0.56 (0.25) n=128	0.43 (0.12) n=68	0.04 (0.02) n=20	0.46 (0.12) n=49	0.82 (0.34) n=30
Wheat	Oceania	0.37 (0.11) n=274	2.08 (0.37) n=672	1.09 (0.51) n=1180	0.13 (0.06) n=29	0.32 (0.05) n=24	0.33 (0.05) n=24	1.61 (0.26) n=24

Table S6. Coefficients for harvest index natural logarithm (ln) models differentiated by region for Scenario 1.

Crop	Region	ln 'a' value*	ln 'b' value*
Maize	Africa	0.09	0.53
Maize	Americas	0.08	0.50
Maize	Asia	0.06	0.50
Maize	Europe	NA	NA
Maize	Oceania	0.13	0.46
Maize	World	0.07	0.51
Wheat	Africa	0.04	0.43
Wheat	Americas	0.06	0.50
Wheat	Asia	0.01	0.43
Wheat	Europe	0.11	0.55
Wheat	Oceania	0.13	0.57
Wheat	World	0.05	0.46

*Where harvest index = $a \times \ln(\text{CPY}) + b$ and CPY=crop product yield in kilograms (kg) of dry matter (DM) per m².

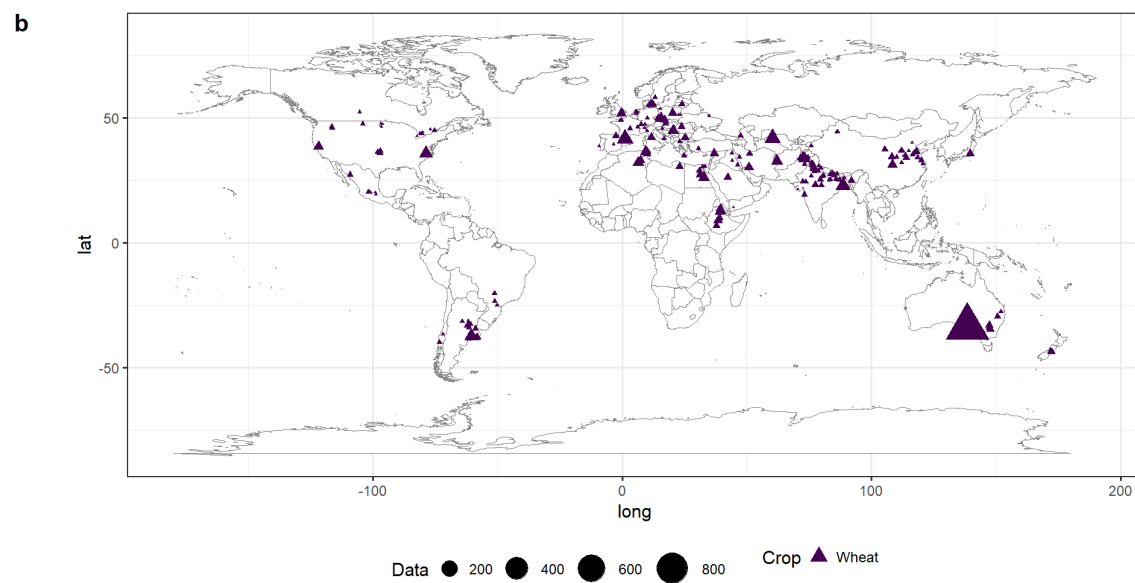
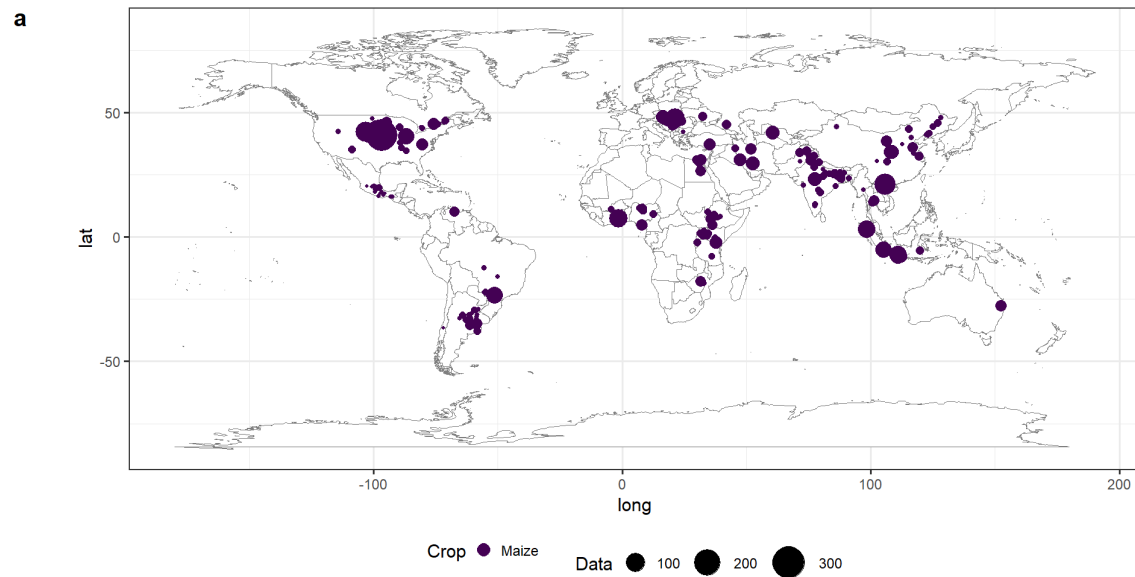


Figure S1: Locations of trials for data to support Tier 3 estimates of nutrient removal coefficients for maize (sub-plot a) and wheat (sub-plot b). The size of the symbol relates to the number of unique datum for each location. A unique datum represents a unique treatment-replicate combination.

References

- Burnham, K. P., and Anderson, D. R. (2002). "Model Selection and Multimodel Inference," 2nd edition/Ed. Springer, New York, USA.
- FAO (2024). FAOSTAT. Vol. 2020. FAO, Rome.
- Ludemann, C. I., Hijbeek, R., van Loon, M., Murrell, S. T., Dobermann, A., and van Ittersum, M. K. (2024). Published field experimental data of crop yields, nutrient concentrations, and harvest indices from around the world. <https://doi.org/10.5061/dryad.djh9w0w4m>. In "Dryad," Dryad [Dataset].
- Ludemann, C. I., Hijbeek, R., van Loon, M. P., Murrell, T. S., Dobermann, A., and van Ittersum, M. K. (2022). Estimating maize harvest index and nitrogen concentrations in grain and residue using globally available data. *Field Crops Research* **284**, 1-25.
- McKevith, B. (2004). Nutritional aspects of cereals. *Nutrition Bulletin* **29**, 111-142.
- Nakagawa, S., Johnson, P. C. D., and Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface* **14**, 20170213.
- Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* **4**, 133-142.
- R Core Team (2020). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Unkovich, M., Baldock, J., and Forbes, M. (2010). A review of biological yield and harvest index in Australian field crops. In "Advances in Agronomy", pp. 174-219. Elsevier Academic Press,.
- van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., and Hochman, Z. (2013). Yield gap analysis with local to global relevance—A review. *Field Crops Research* **143**, 4-17.
- WHO, and FAO (2007). "Codex Alimentarius," 1st edition/Ed. World Health Organisation, Food and Agriculture Organization of the United Nations,, Rome.