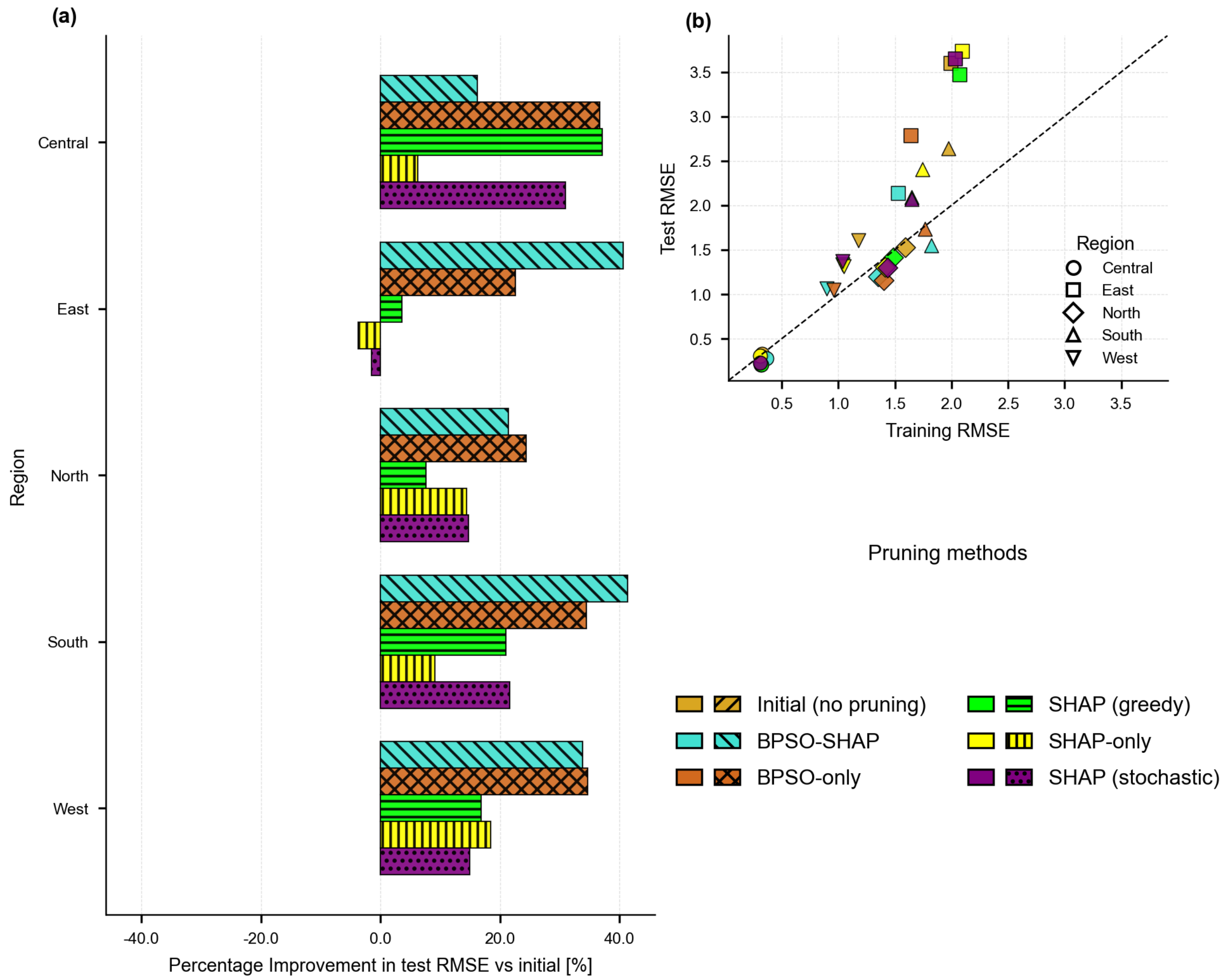
Supplementary Information

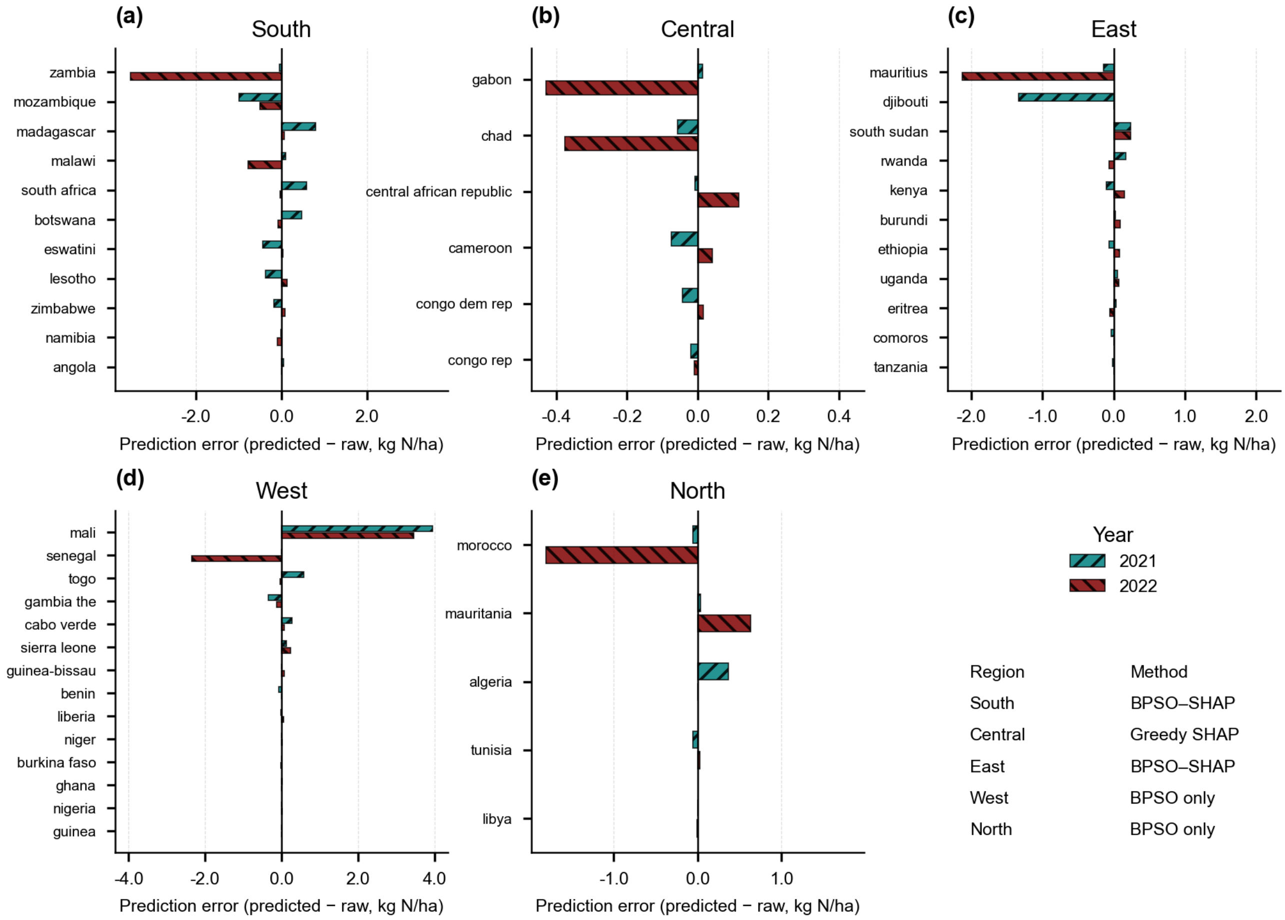
Projecting nitrogen fertilizer use in Africa using empirically bounded historical drivers

---

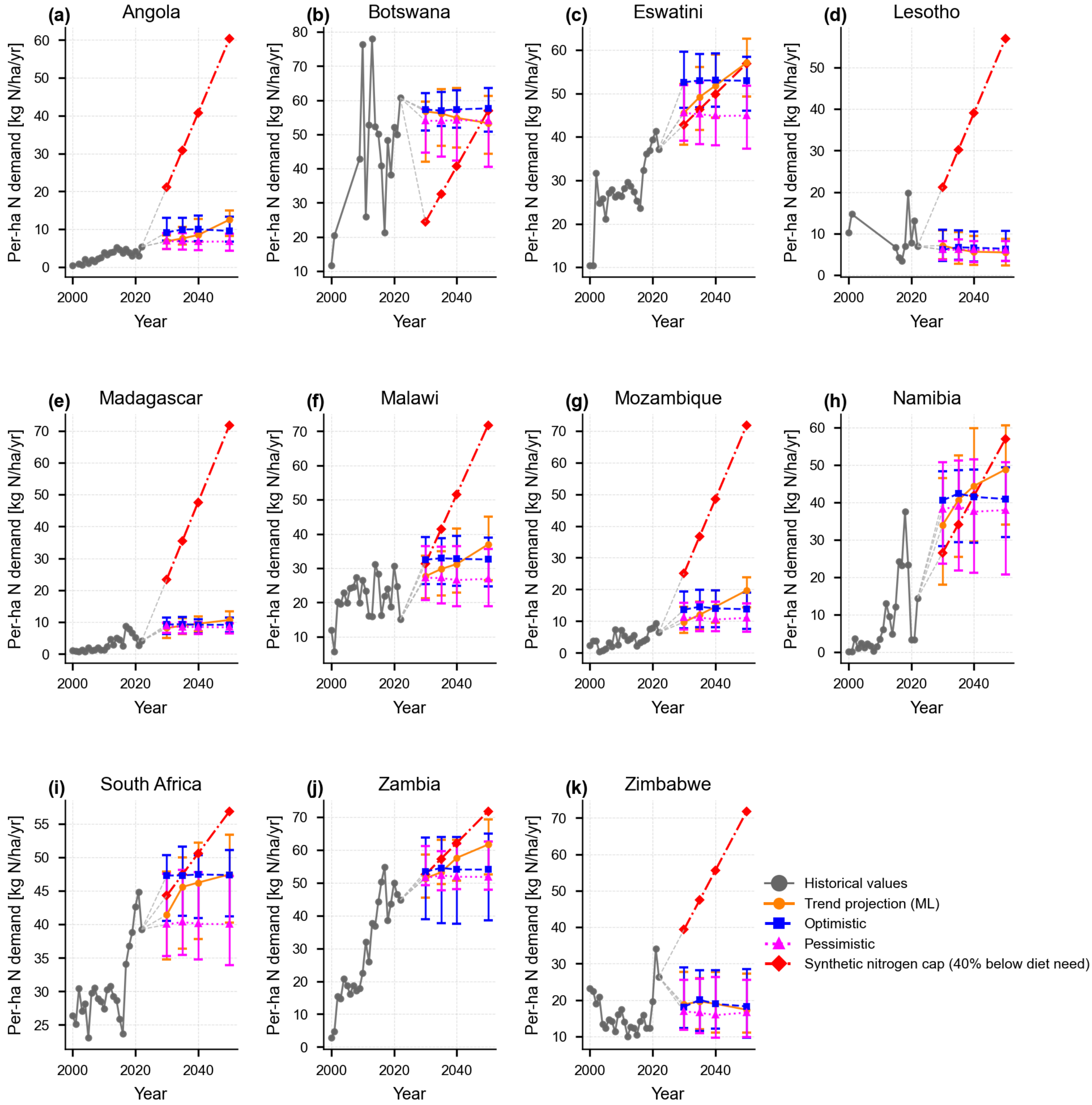
Supplementary Figures



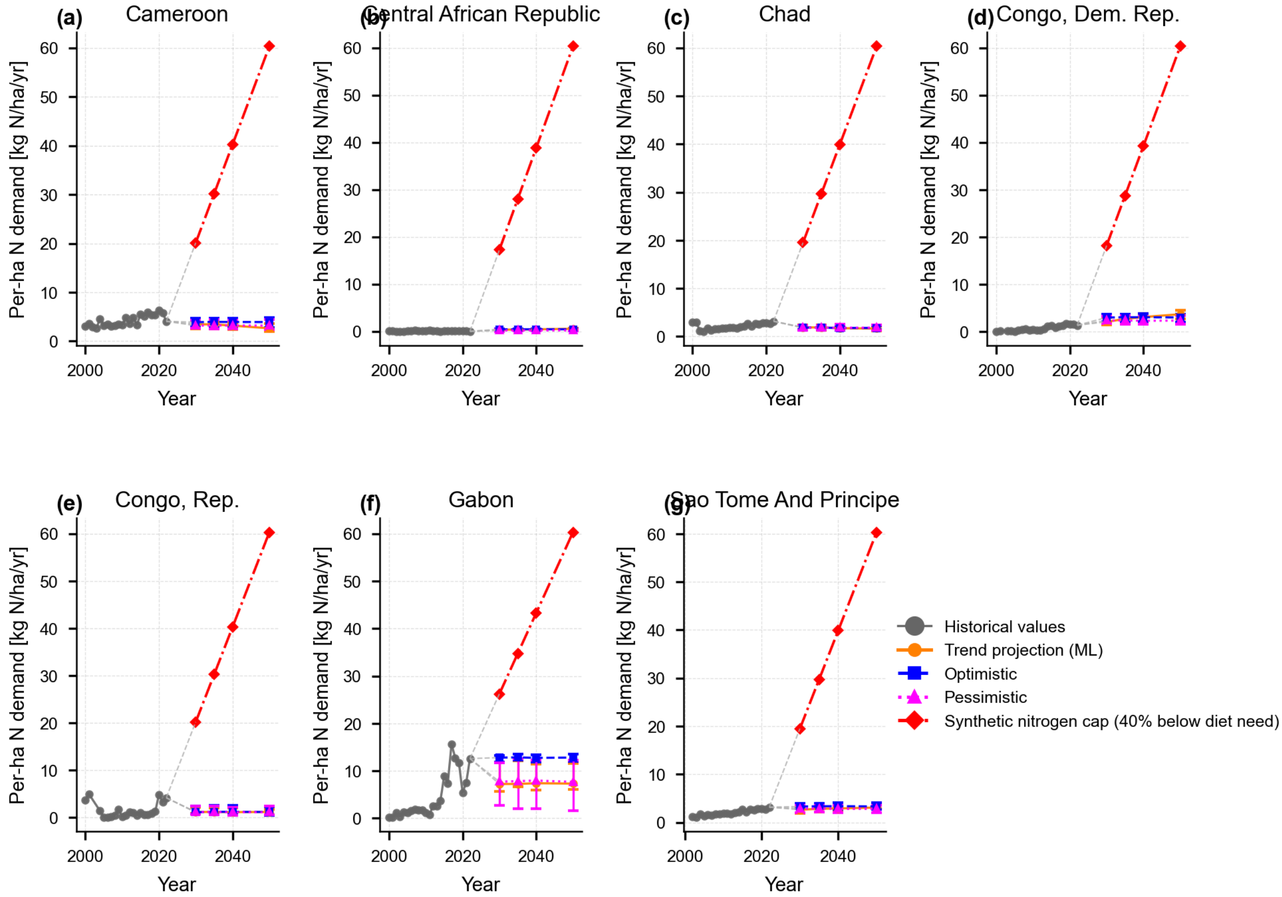
Supplementary Fig. S1 | Feature pruning improves accuracy and yields compact regional fertilizer-intensity models. (a) Percentage change in test root mean squared error (RMSE) relative to the unpruned initial model, shown for each region and pruning strategy. Positive values indicate reduced error (improved performance), whereas negative values indicate degradation. (b) Training versus test RMSE for each region–method combination. The dashed line indicates equal training and test error; points near this line indicate good generalization, while points above it indicate overfitting. Optimization-based pruning (BPSO-only and SHAP-guided BPSO) yields consistent error reductions without widening the train–test gap. See Supplementary Tables 6–7.



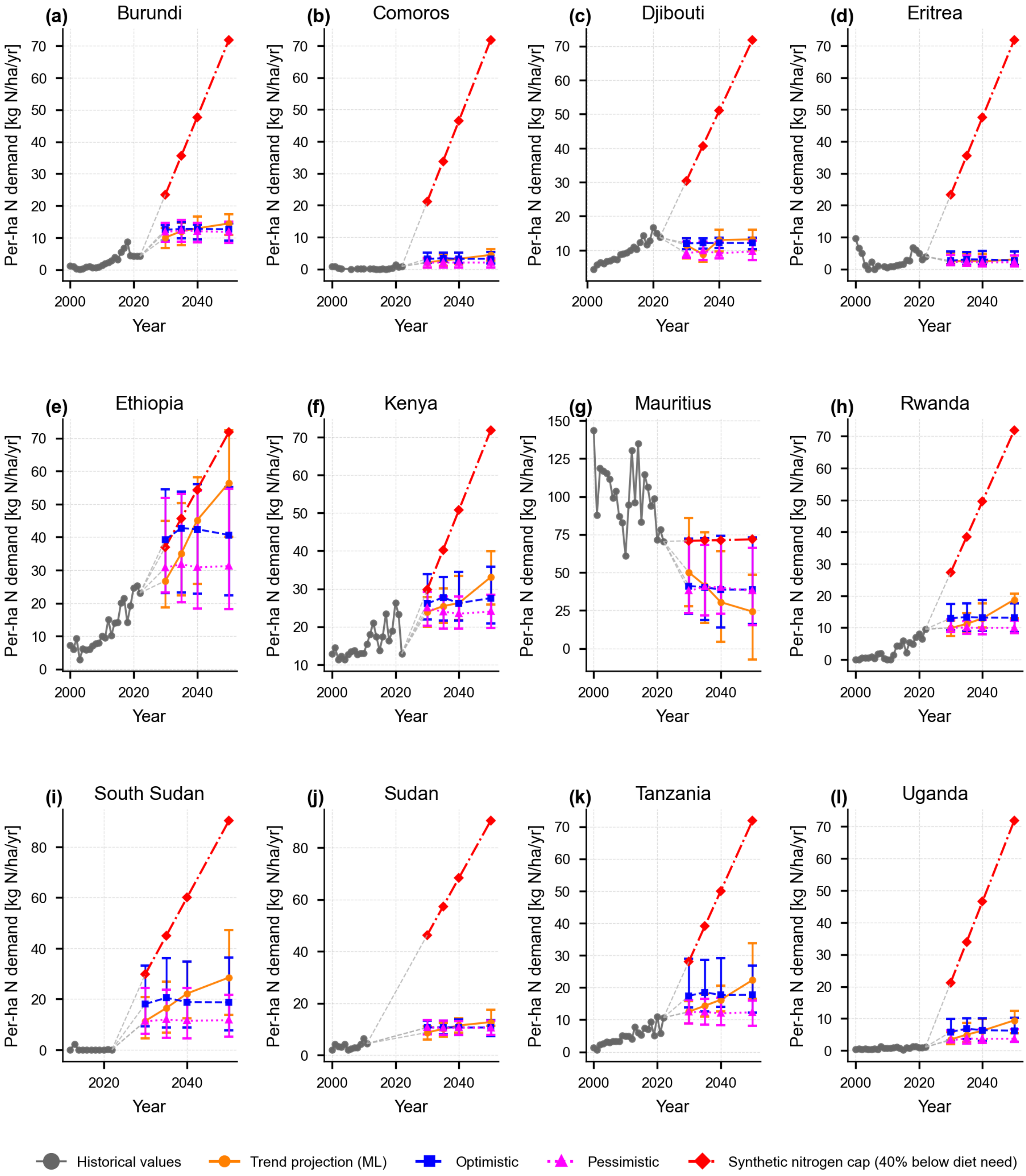
Supplementary Fig. S2 | Prediction-error tornado plots for regional fertilizer-intensity models. Tornado plots rank the largest country-level prediction errors for each regional model and SHAP-based Binary Particle Swarm Optimization pruning configuration on held-out data.



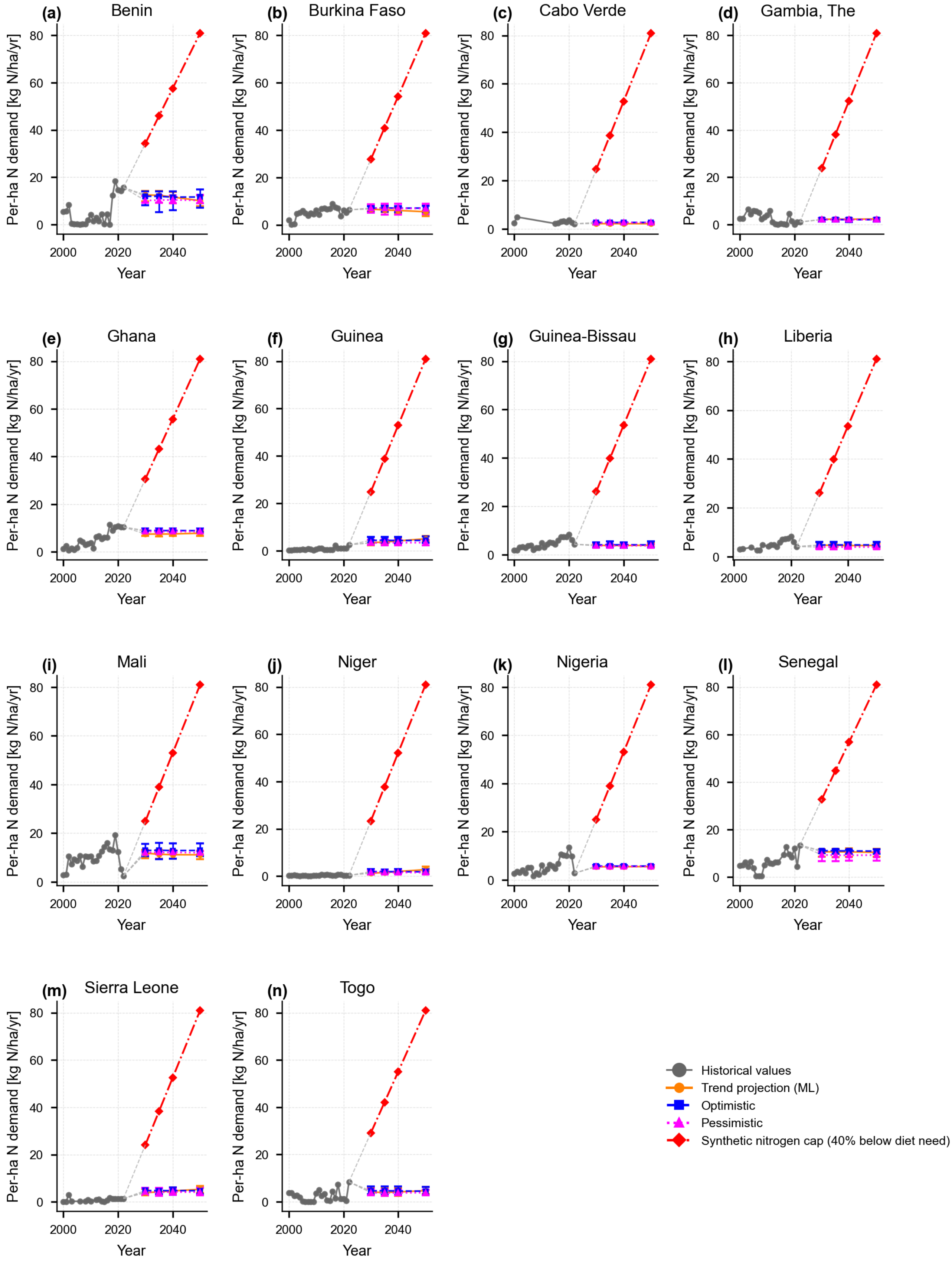
Supplementary Fig. S3 | Country-level fertilizer-intensity trajectories under alternative scenarios: Southern Africa. Country trajectories of nitrogen fertilizer intensity under business-as-usual (BAU), optimistic, and pessimistic scenarios. BAU is obtained by extrapolating driver variables forward in time and evaluating the extrapolated inputs using the trained regional model. Optimistic and pessimistic scenarios are obtained by applying SHAP-guided percentile shifts to a subset of influential, perturbable drivers. All shifts are constrained to empirically observed bounds and non-perturbable drivers remain at BAU. Uncertainty bands summarize uncertainty across bootstrap model ensembles. See Methods: Business-as-usual extrapolation, Percentile mapping and SHAP-guided constrained scenario optimisation, and Uncertainty quantification and scenario robustness.



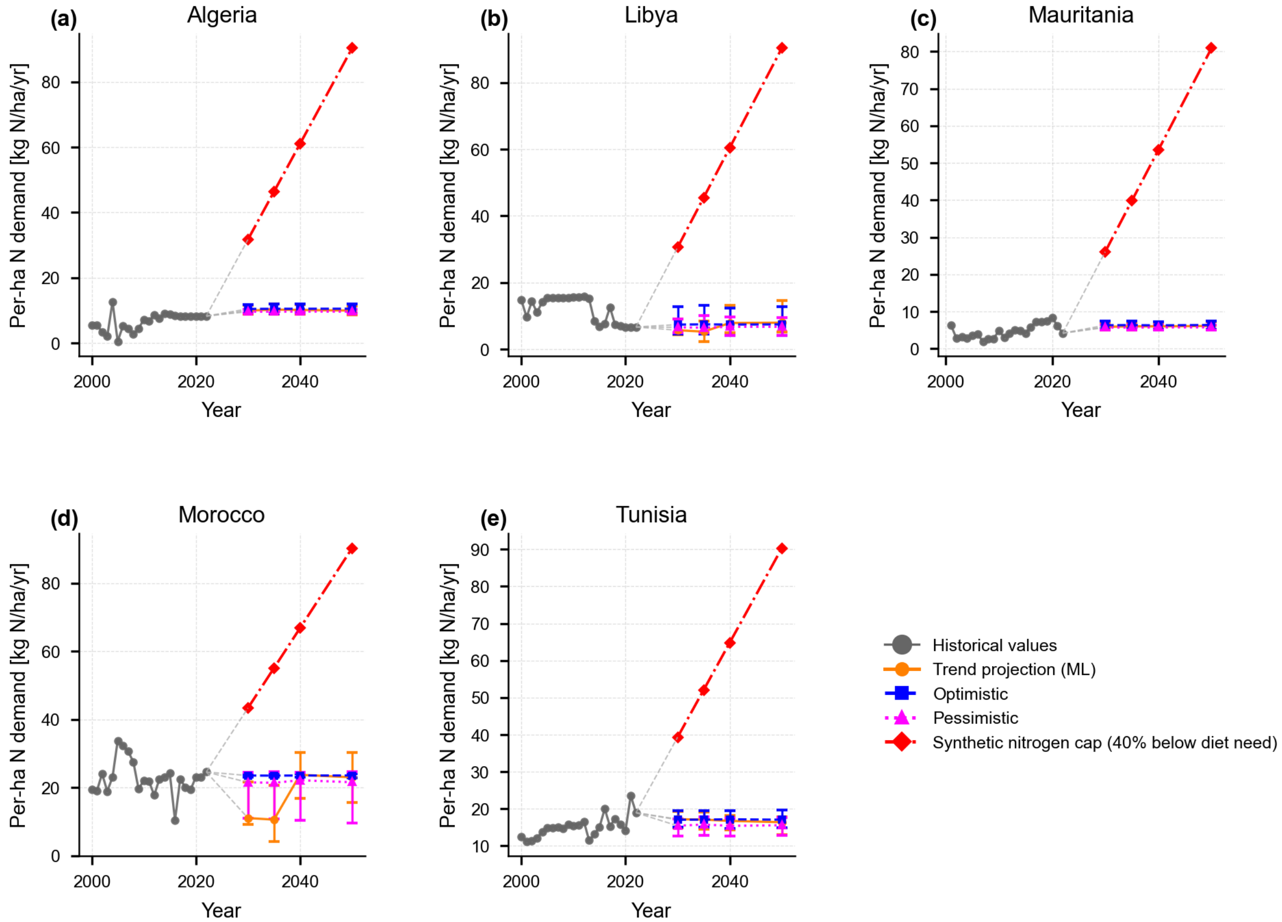
Supplementary Fig. S4 | Country-level fertilizer-intensity trajectories under alternative scenarios: Central Africa. Country trajectories of nitrogen fertilizer intensity under business-as-usual (BAU), optimistic, and pessimistic scenarios. BAU is obtained by extrapolating driver variables forward in time and evaluating the extrapolated inputs using the trained regional model. Optimistic and pessimistic scenarios are obtained by applying SHAP-guided percentile shifts to a subset of influential, perturbable drivers. All shifts are constrained to empirically observed bounds and non-perturbable drivers remain at BAU. Uncertainty bands summarize uncertainty across bootstrap model ensembles. See Methods: Business-as-usual extrapolation, Percentile mapping and SHAP-guided constrained scenario optimisation, and Uncertainty quantification and scenario robustness.



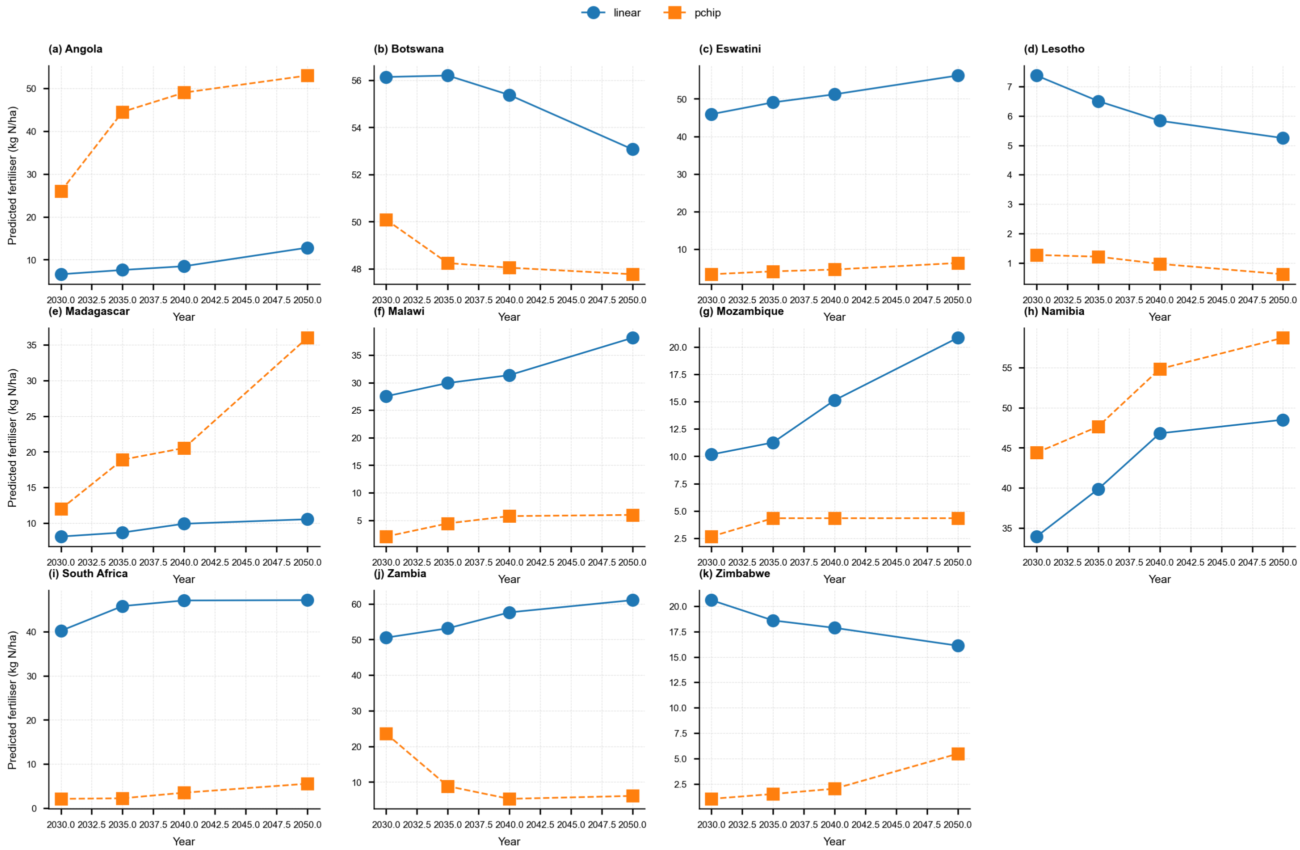
Supplementary Fig. S5 | Country-level fertilizer-intensity trajectories under alternative scenarios: East Africa. Country trajectories of nitrogen fertilizer intensity under business-as-usual (BAU), optimistic, and pessimistic scenarios. BAU is obtained by extrapolating driver variables forward in time and evaluating the extrapolated inputs using the trained regional model. Optimistic and pessimistic scenarios are obtained by applying SHAP-guided percentile shifts to a subset of influential, perturbable drivers. All shifts are constrained to empirically observed bounds and non-perturbable drivers remain at BAU. Uncertainty bands summarize uncertainty across bootstrap model ensembles. See Methods: Business-as-usual extrapolation, Percentile mapping and SHAP-guided constrained scenario optimisation, and Uncertainty quantification and scenario robustness.



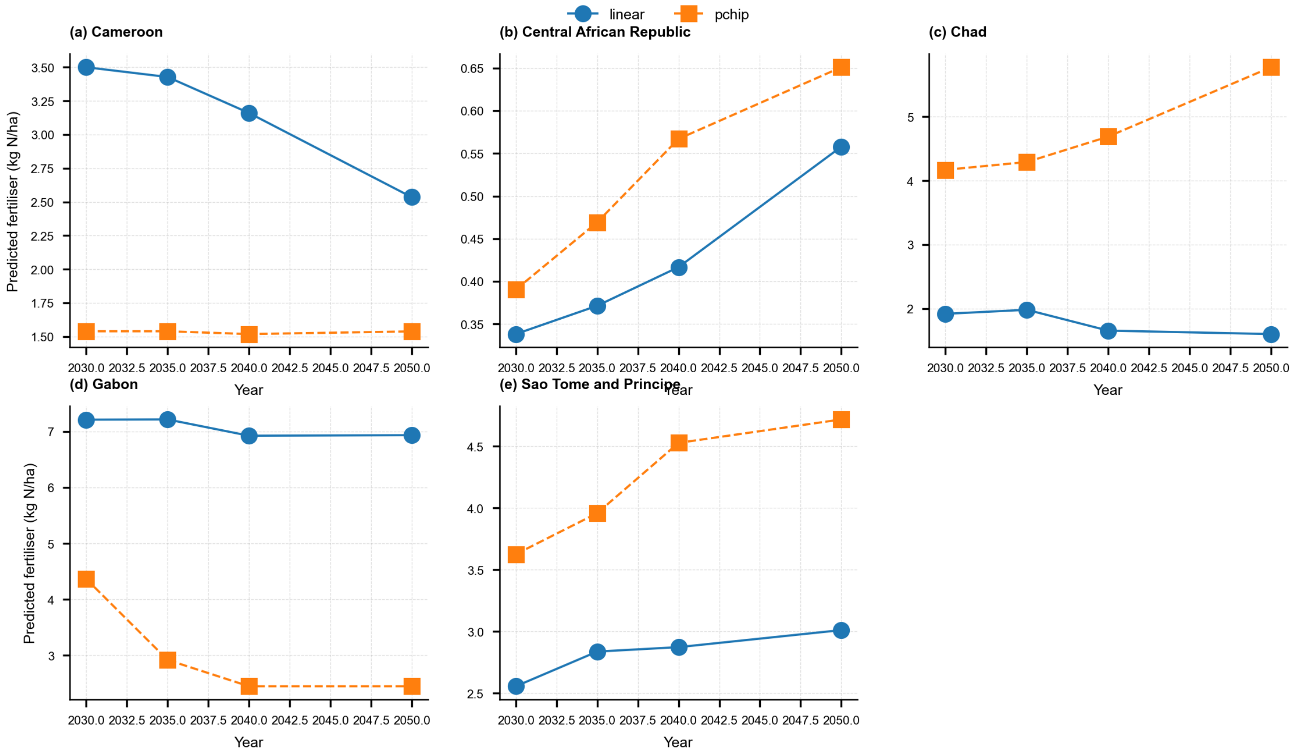
Supplementary Fig. S6 | Country-level fertilizer-intensity trajectories under alternative scenarios: West Africa. Country trajectories of nitrogen fertilizer intensity under business-as-usual (BAU), optimistic, and pessimistic scenarios. BAU is obtained by extrapolating driver variables forward in time and evaluating the extrapolated inputs using the trained regional model. Optimistic and pessimistic scenarios are obtained by applying SHAP-guided percentile shifts to a subset of influential, perturbable drivers. All shifts are constrained to empirically observed bounds and non-perturbable drivers remain at BAU. Uncertainty bands summarize uncertainty across bootstrap model ensembles. See Methods: Business-as-usual extrapolation, Percentile mapping and SHAP-guided constrained scenario optimisation, and Uncertainty quantification and scenario robustness.



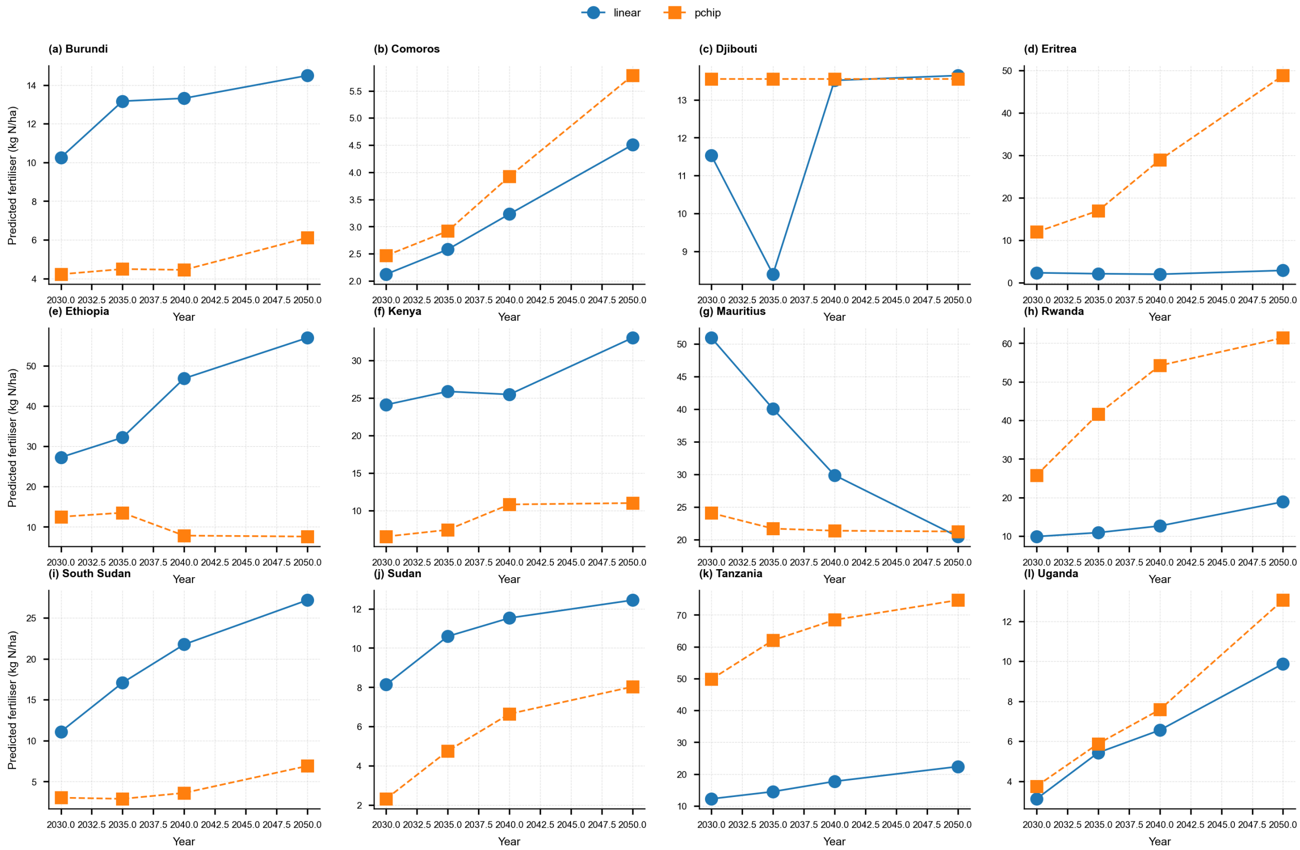
Supplementary Fig. S7 | Country-level fertilizer-intensity trajectories under alternative scenarios: North Africa. Country trajectories of nitrogen fertilizer intensity under business-as-usual (BAU), optimistic, and pessimistic scenarios. BAU is obtained by extrapolating driver variables forward in time and evaluating the extrapolated inputs using the trained regional model. Optimistic and pessimistic scenarios are obtained by applying SHAP-guided percentile shifts to a subset of influential, perturbable drivers. All shifts are constrained to empirically observed bounds and non-perturbable drivers remain at BAU. Uncertainty bands summarize uncertainty across bootstrap model ensembles. See Methods: Business-as-usual extrapolation, Percentile mapping and SHAP-guided constrained scenario optimization, and Uncertainty quantification and scenario robustness.



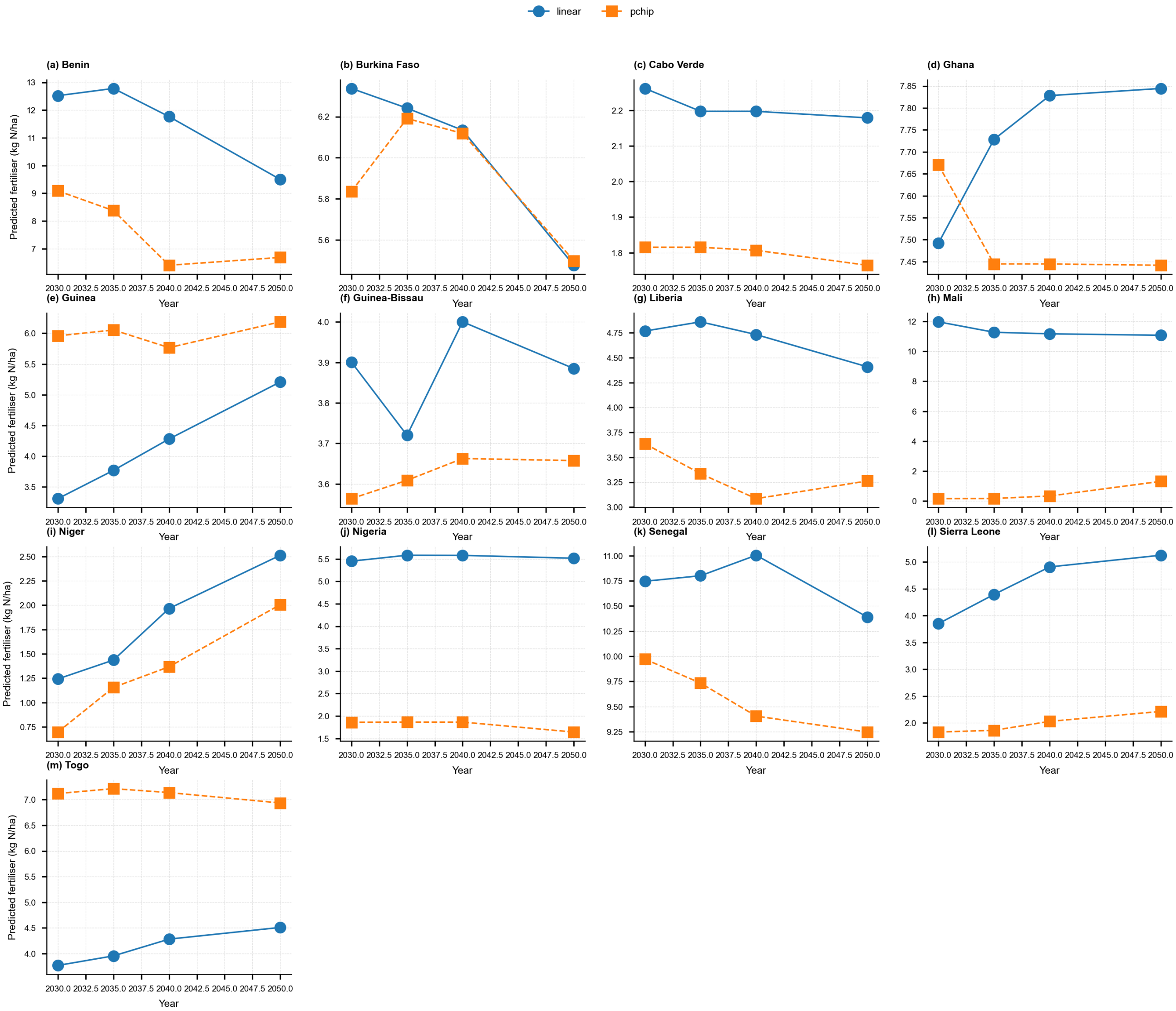
Supplementary Fig. S8 | Linear vs PCHIP driver extrapolation in BAU projections: Southern Africa. Comparison of BAU fertilizer-intensity trajectories generated by linear extrapolation of driver variables versus a shape-preserving piecewise-cubic Hermite interpolating polynomial (PCHIP) interpolation or extrapolation. In both cases, extrapolated drivers are evaluated through the fitted regional model to obtain BAU intensity trajectories. See Methods: Business-as-usual extrapolation.



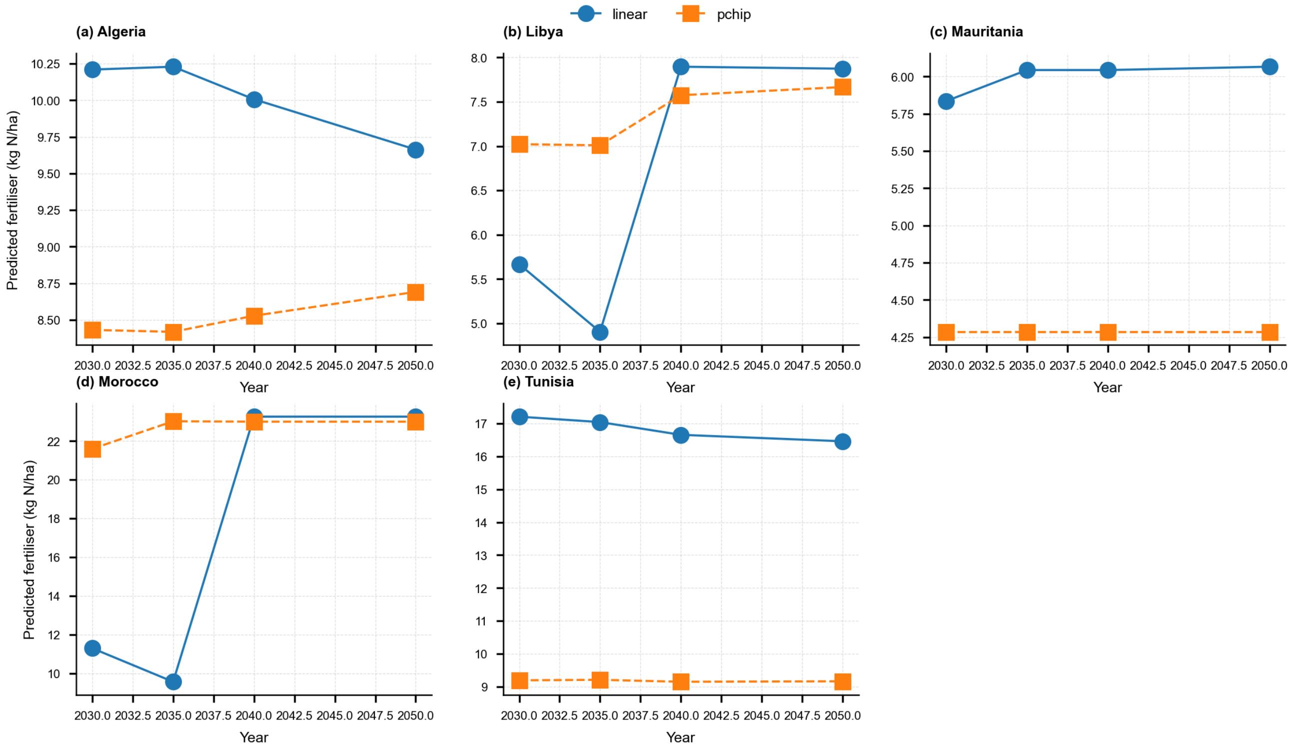
Supplementary Fig. S9 | Linear vs PCHIP driver extrapolation in BAU projections: Central Africa. Comparison of BAU fertilizer-intensity trajectories generated by linear extrapolation of driver variables versus a shape-preserving piecewise-cubic Hermite interpolating polynomial (PCHIP) interpolation or extrapolation. In both cases, extrapolated drivers are evaluated through the fitted regional model to obtain BAU intensity trajectories. See Methods: Business-as-usual extrapolation.



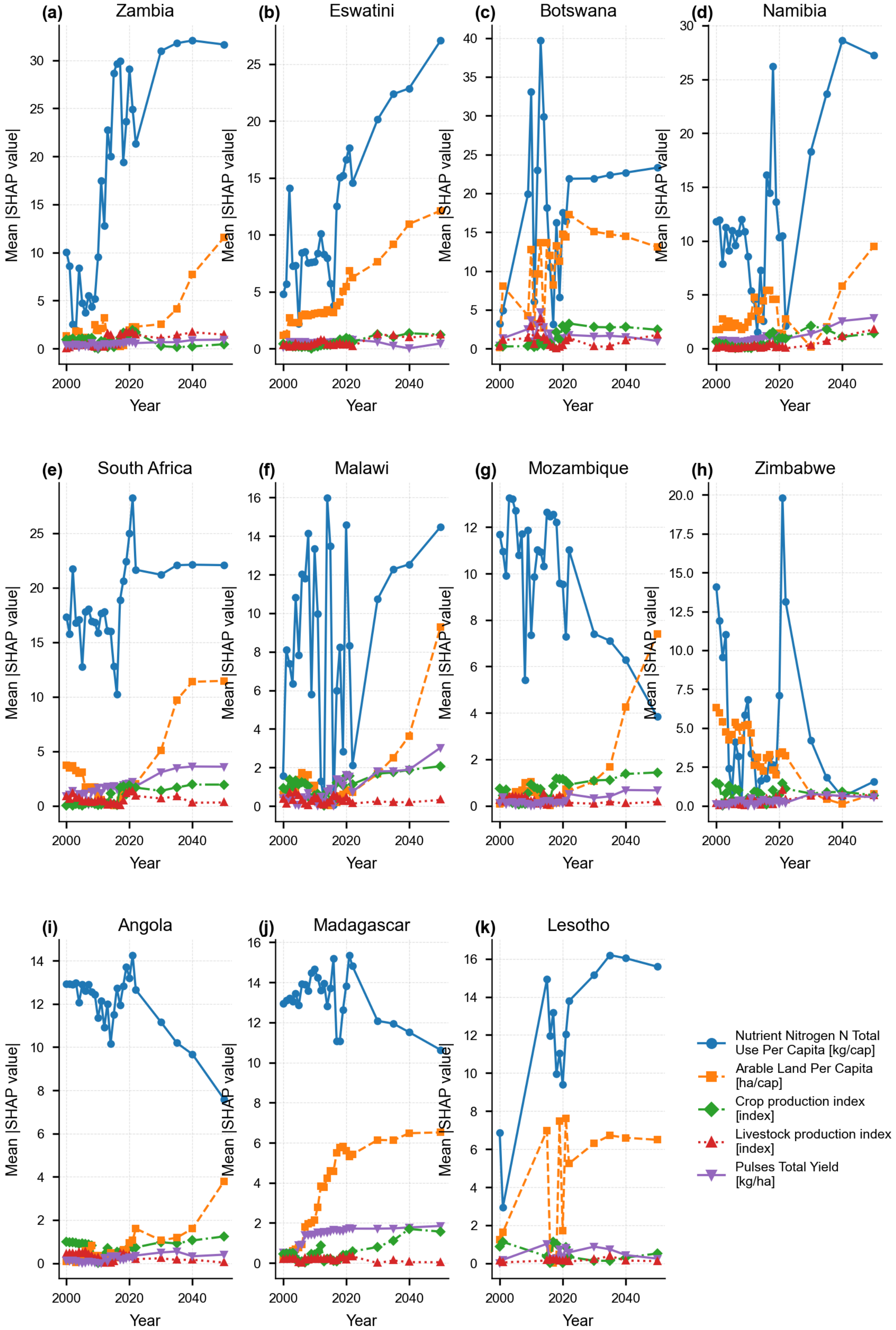
Supplementary Fig. S10 | Linear vs PCHIP driver extrapolation in BAU projections: East Africa. Comparison of BAU fertilizer-intensity trajectories generated by linear extrapolation of driver variables versus a shape-preserving piecewise-cubic Hermite interpolating polynomial (PCHIP) interpolation or extrapolation. In both cases, extrapolated drivers are evaluated through the fitted regional model to obtain BAU intensity trajectories. See Methods: Business-as-usual extrapolation.



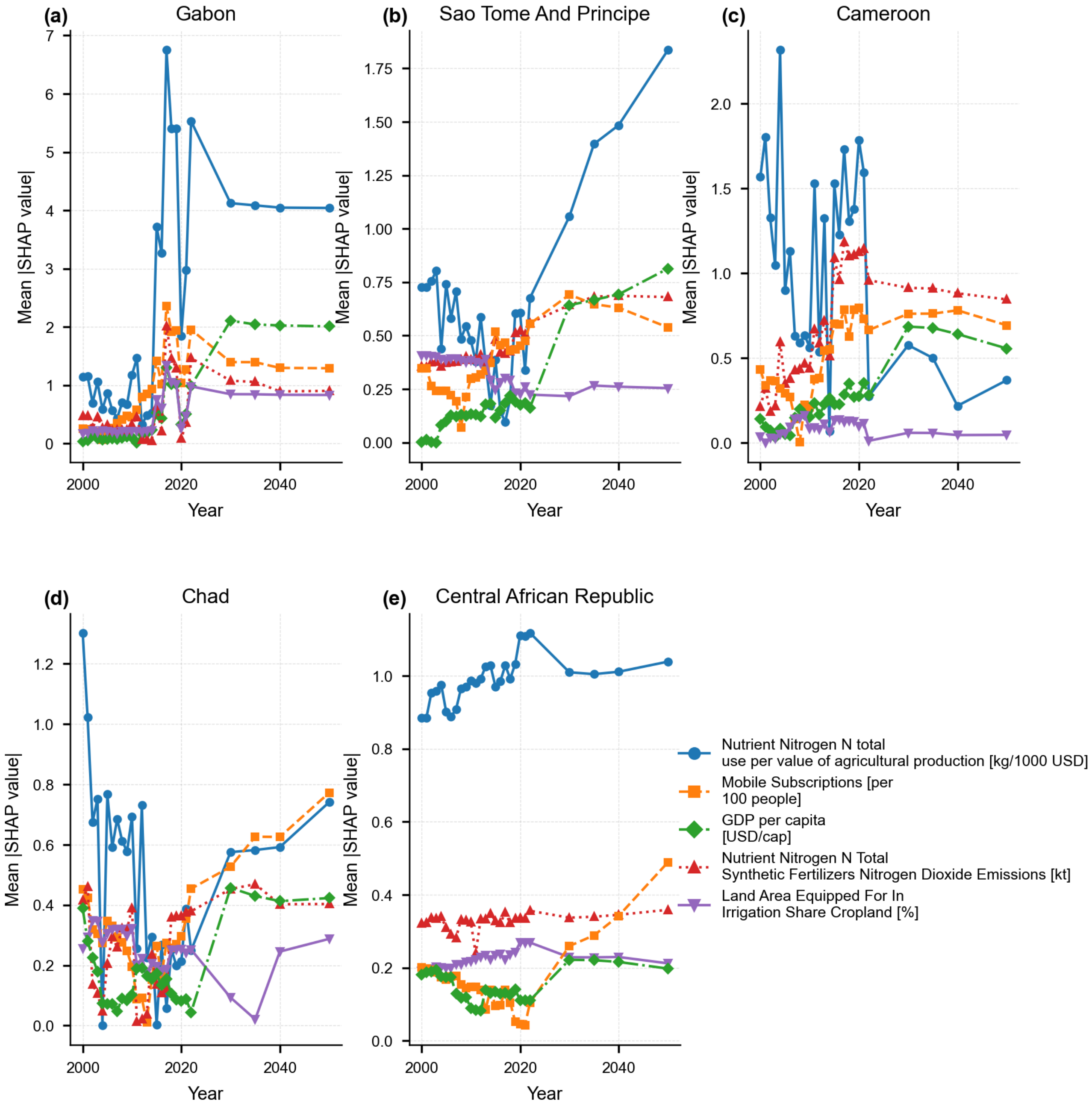
Supplementary Fig. S11 | Linear vs PCHIP driver extrapolation in BAU projections: West Africa. Comparison of BAU fertilizer-intensity trajectories generated by linear extrapolation of driver variables versus a shape-preserving piecewise-cubic Hermite interpolating polynomial (PCHIP) interpolation or extrapolation. In both cases, extrapolated drivers are evaluated through the fitted regional model to obtain BAU intensity trajectories. See Methods: Business-as-usual extrapolation.



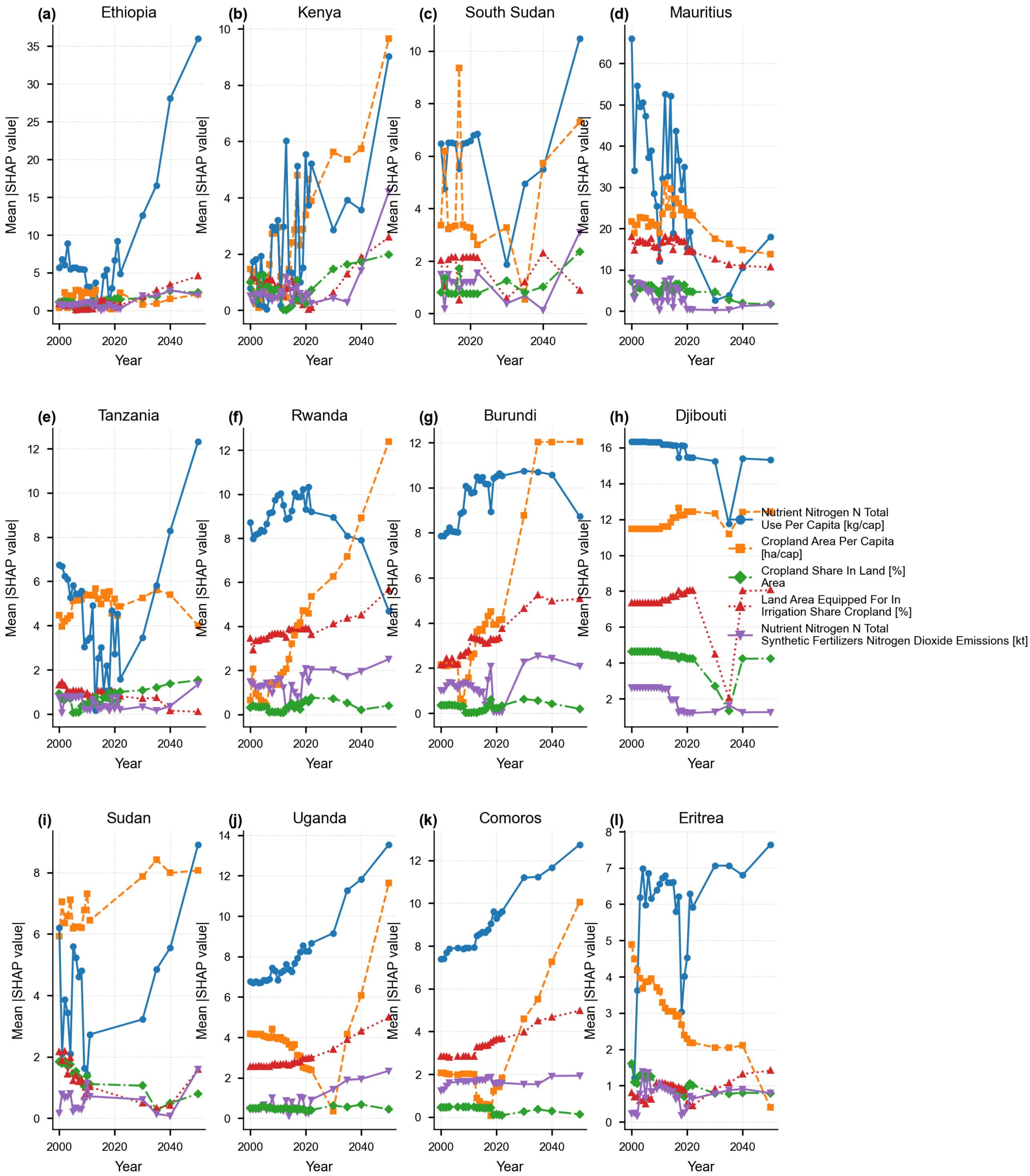
Supplementary Fig. S12 | Linear vs PCHIP driver extrapolation in BAU projections: North Africa. Comparison of BAU fertilizer-intensity trajectories generated by linear extrapolation of driver variables versus a shape-preserving piecewise-cubic Hermite interpolating polynomial (PCHIP) interpolation or extrapolation. In both cases, extrapolated drivers are evaluated through the fitted regional model to obtain BAU intensity trajectories. See Methods: Business-as-usual extrapolation.



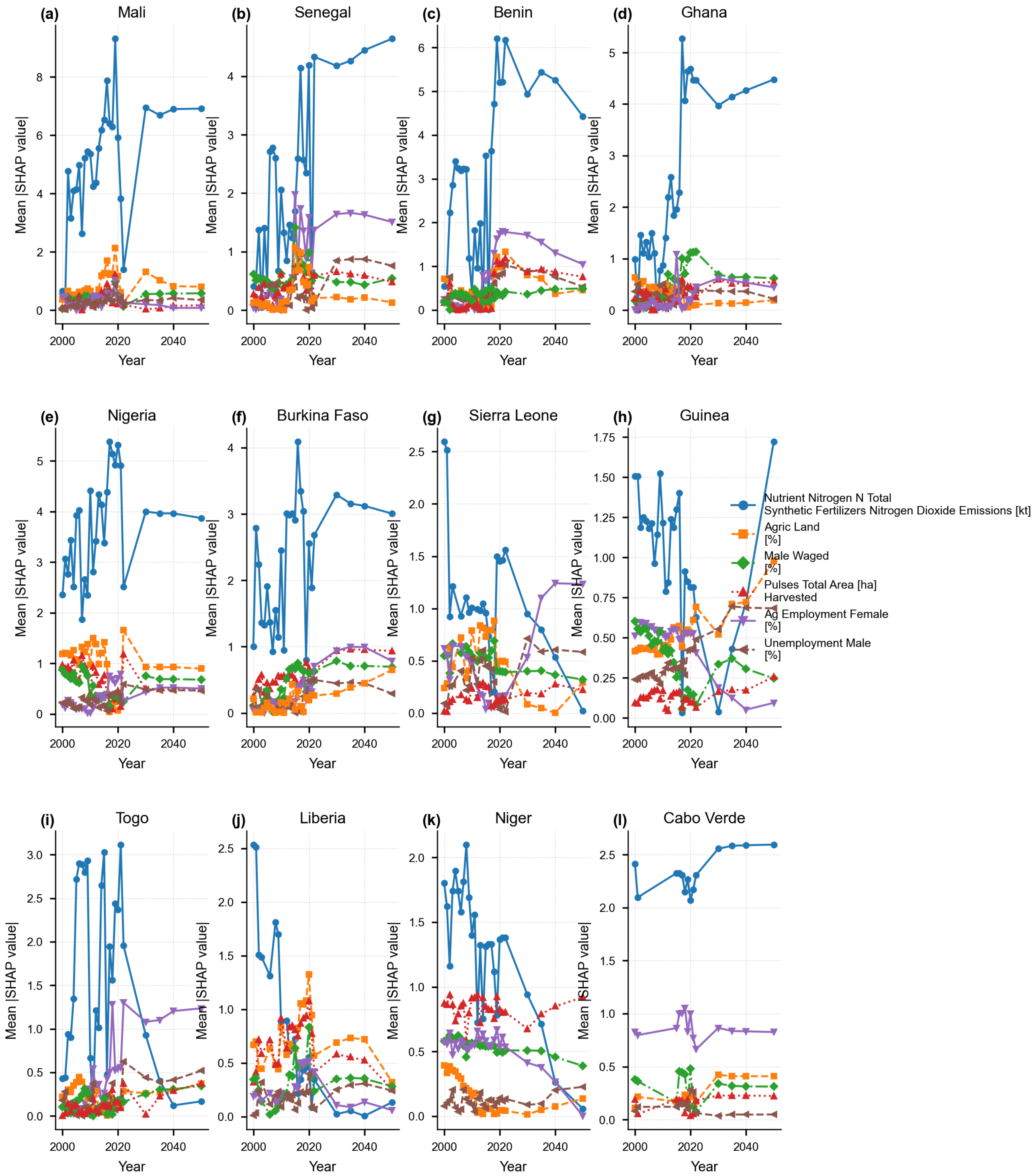
Supplementary Fig. S13 | SHAP driver contributions over time: Southern Africa. Mean absolute SHAP values by calendar year for the pruned regional model. Each series corresponds to one driver shown in the figure. Values summarize the magnitude of feature contributions to predicted nitrogen fertilizer intensity in that year. See Methods: Model attribution and global driver scores.



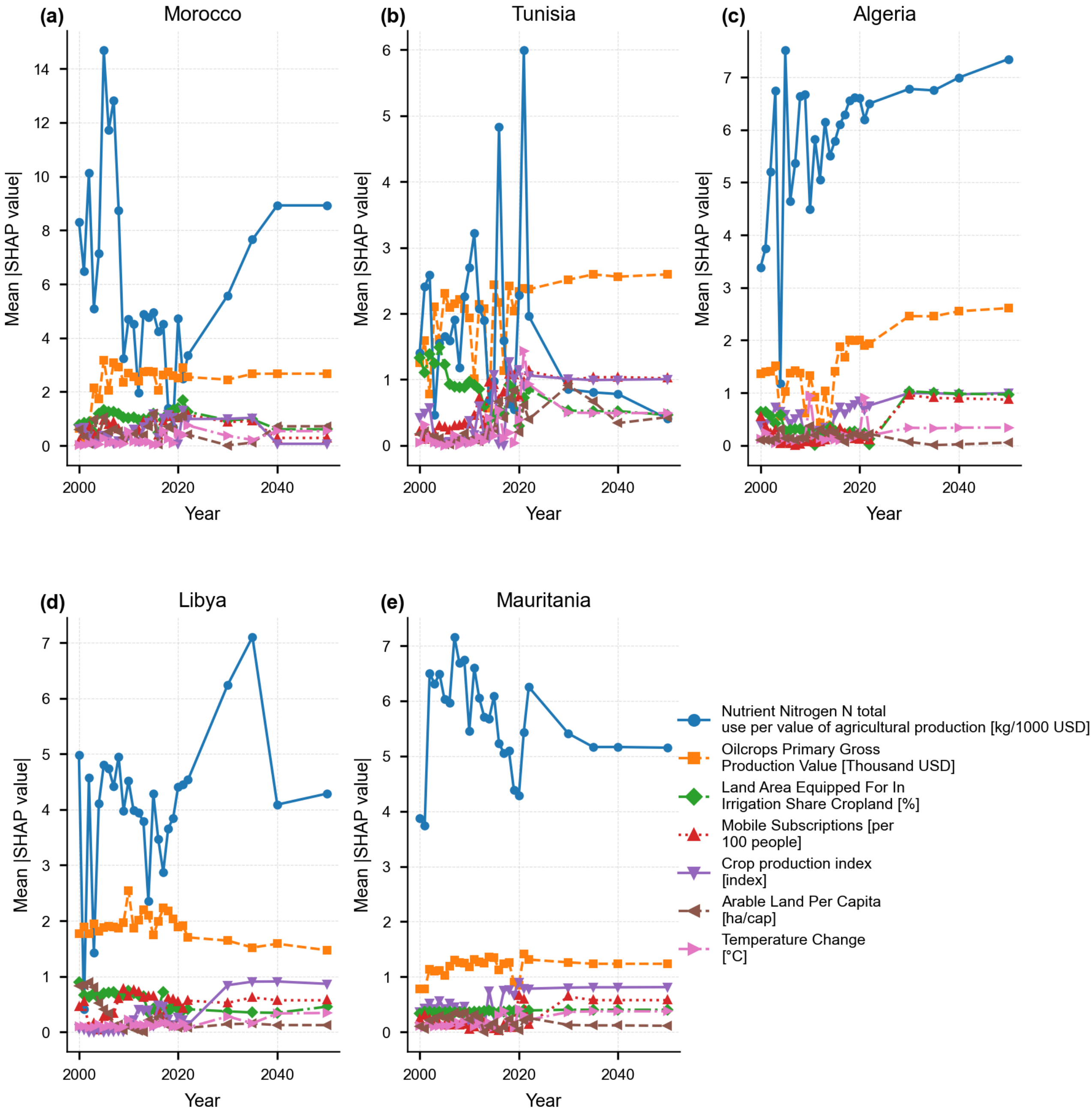
Supplementary Fig. S14 | SHAP driver contributions over time: Central Africa. Mean absolute SHAP values by calendar year for the pruned regional model. Each series corresponds to one driver shown in the figure. Values summarize the magnitude of feature contributions to predicted nitrogen fertilizer intensity in that year. See Methods: Model attribution and global driver scores.



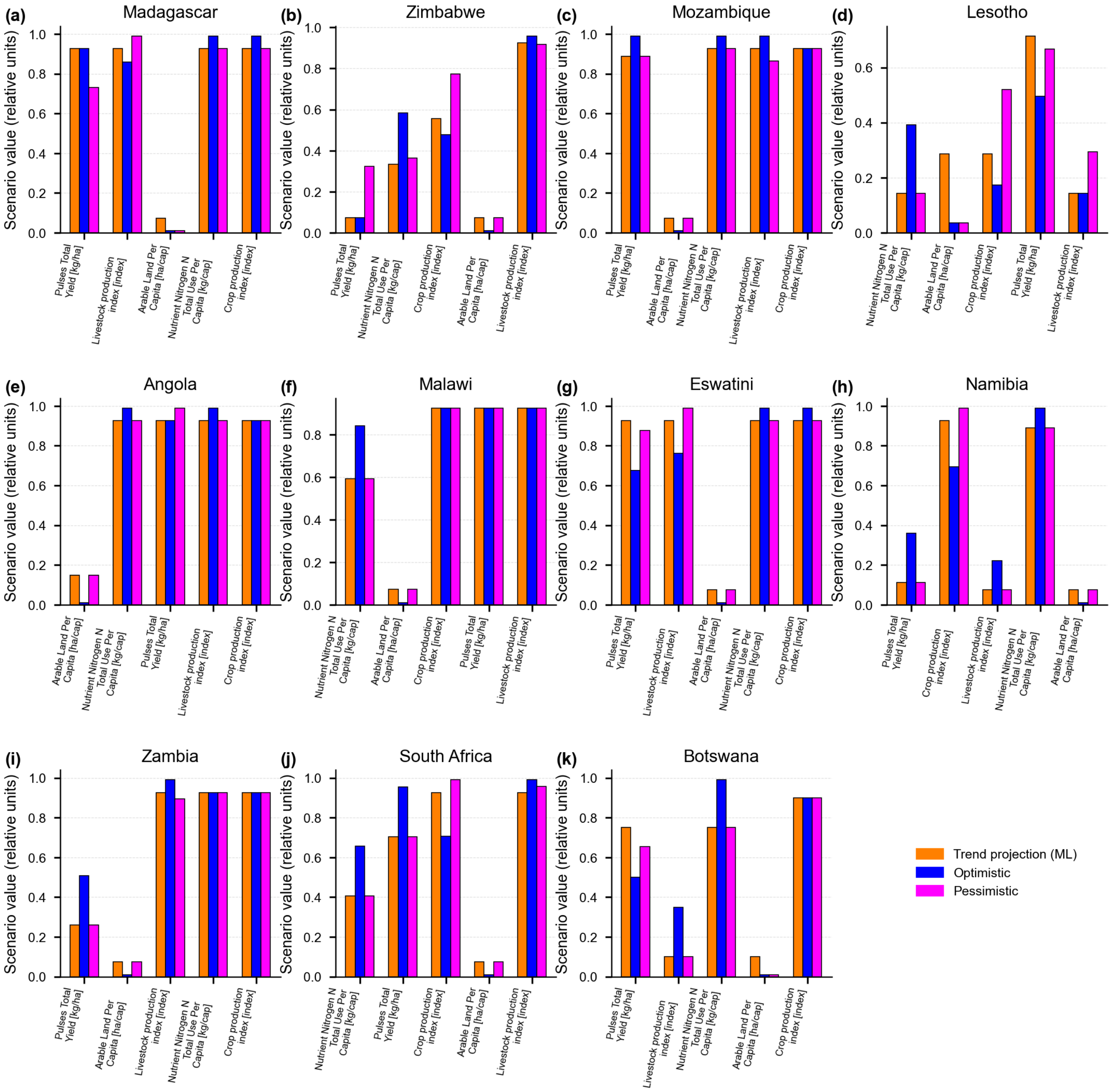
Supplementary Fig. S15 | SHAP driver contributions over time: East Africa. Mean absolute SHAP values by calendar year for the pruned regional model. Each series corresponds to one driver shown in the figure. Values summarize the magnitude of feature contributions to predicted nitrogen fertilizer intensity in that year. See Methods: Model attribution and global driver scores.



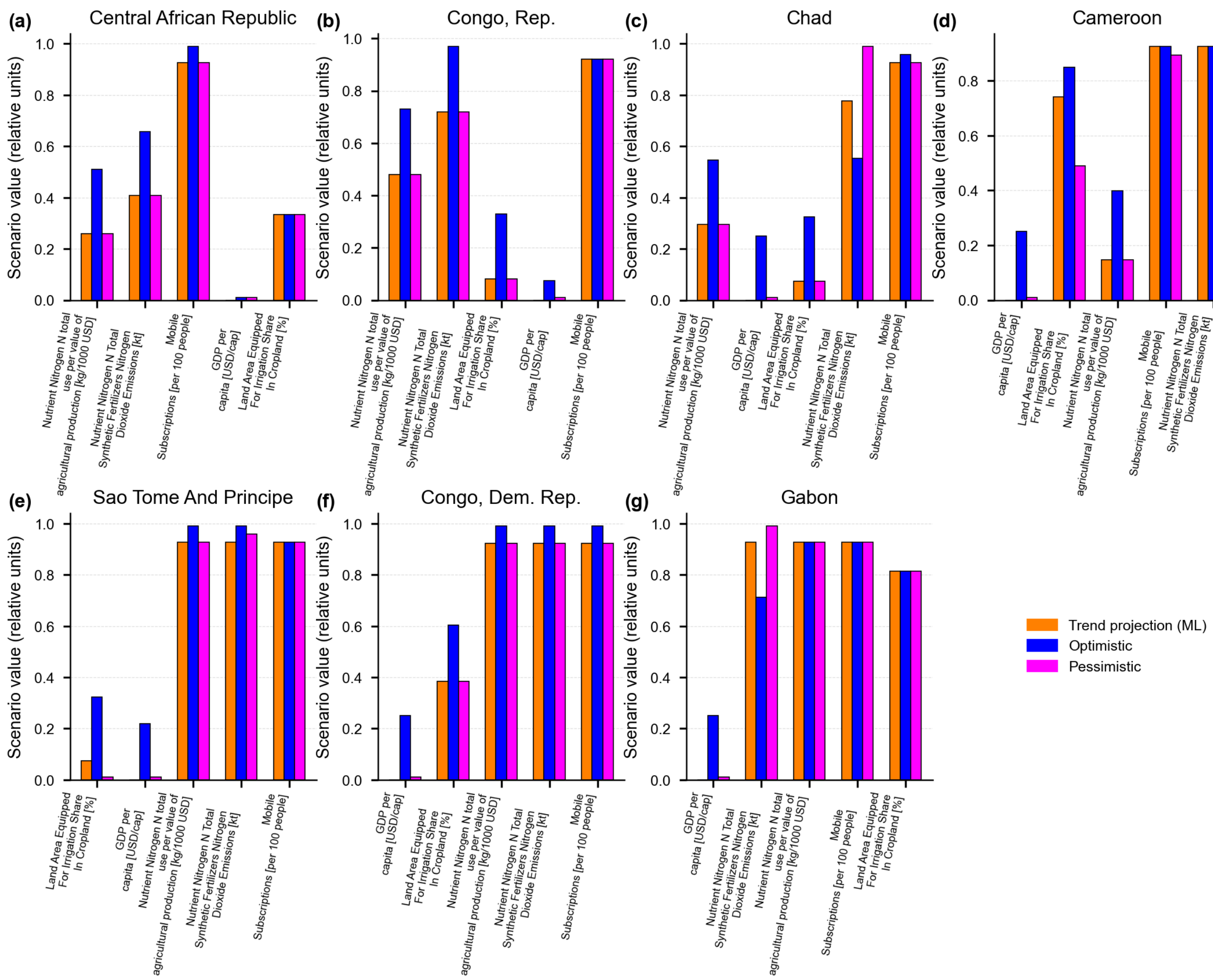
Supplementary Fig. S16 | SHAP driver contributions over time: West Africa. Mean absolute SHAP values by calendar year for the pruned regional model. Each series corresponds to one driver shown in the figure. Values summarize the magnitude of feature contributions to predicted nitrogen fertilizer intensity in that year. See Methods: Model attribution and global driver scores.



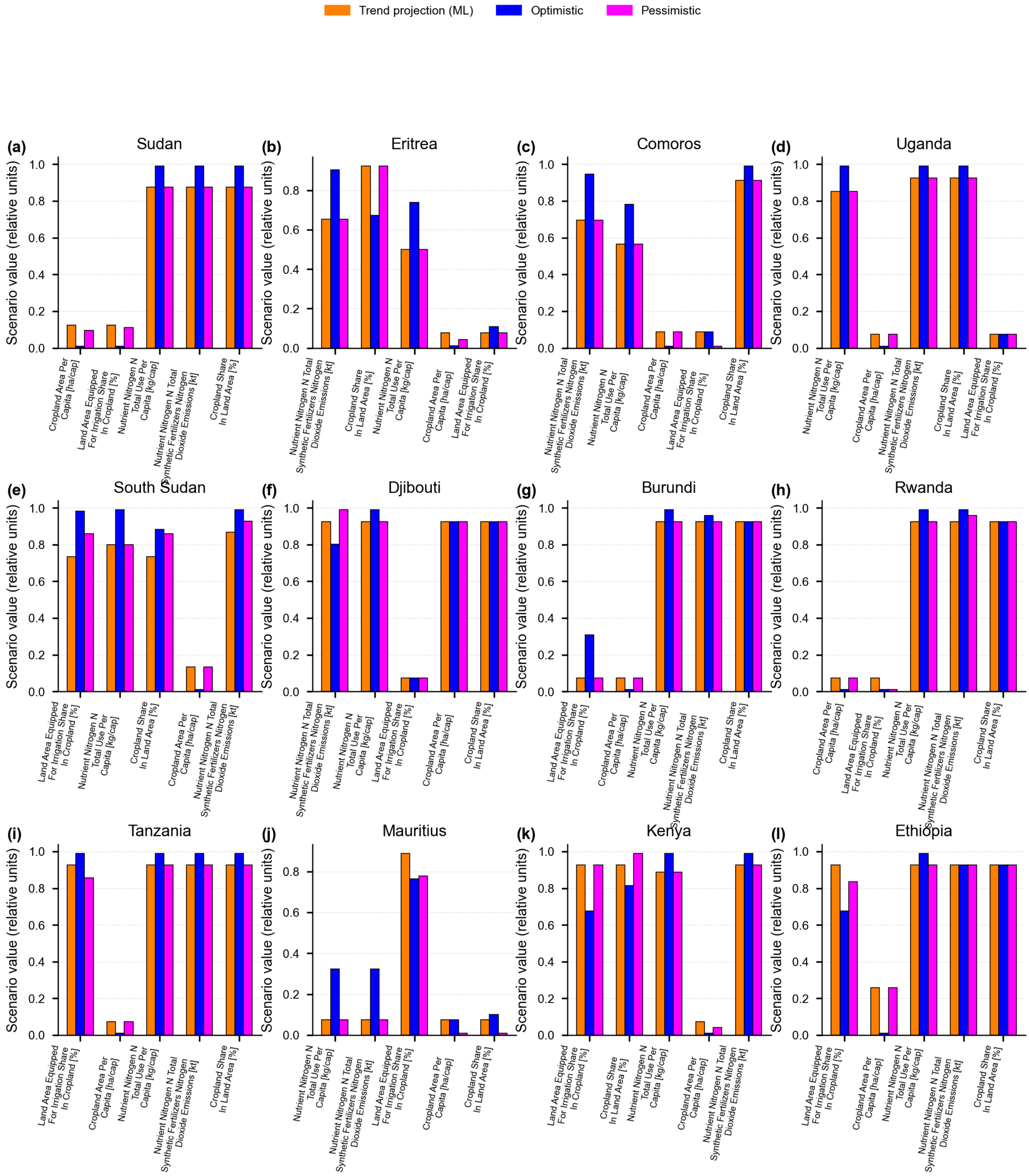
Supplementary Fig. S17 | SHAP driver contributions over time: North Africa. Mean absolute SHAP values by calendar year for the pruned regional model. Each series corresponds to one driver shown in the figure. Values summarize the magnitude of feature contributions to predicted nitrogen fertilizer intensity in that year. See Methods: Model attribution and global driver scores.



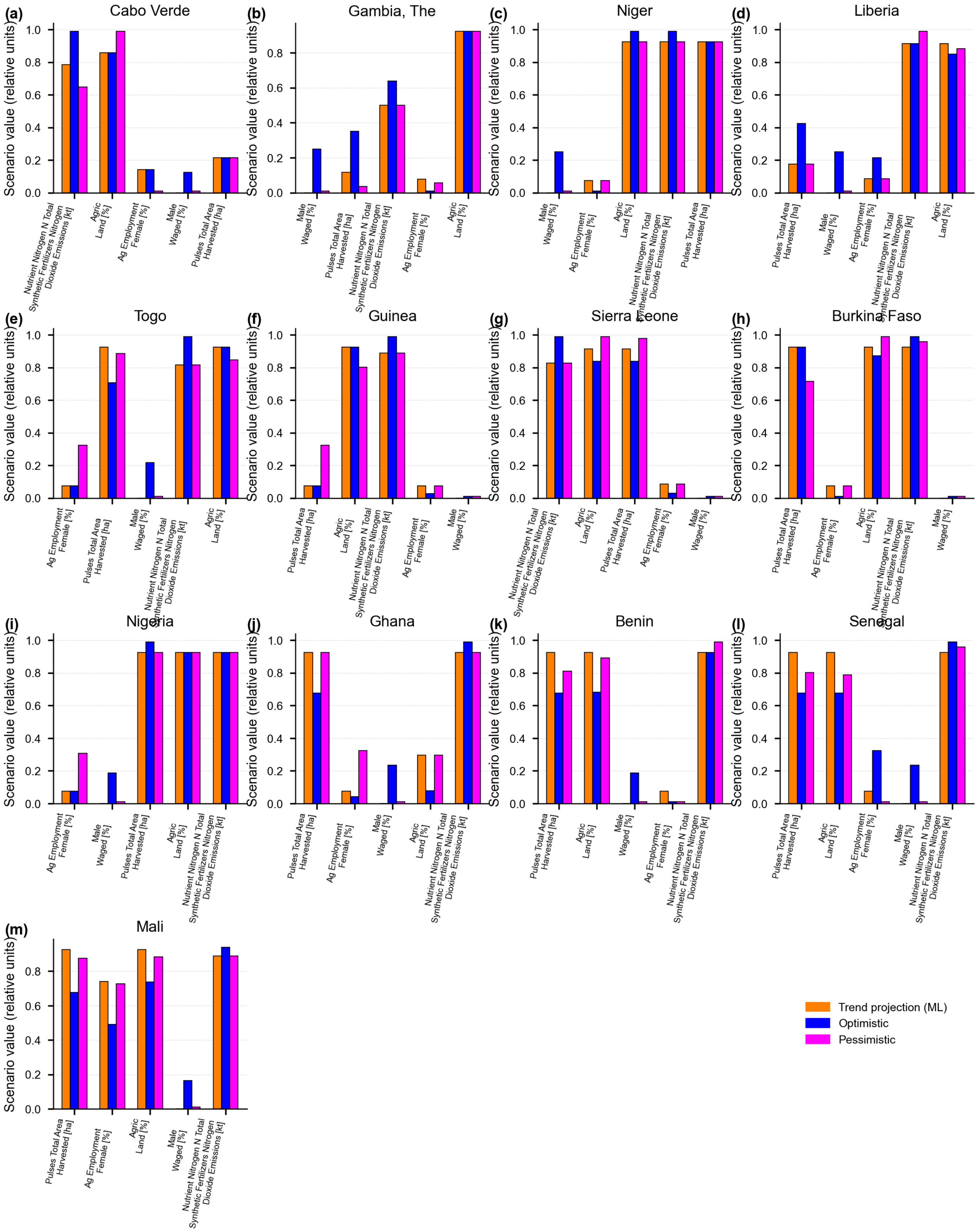
Supplementary Fig. S18 | Regional SHAP driver portfolio and perturbable levers: Southern Africa. SHAP-based driver portfolio for the pruned regional model. The figure summarizes the relative contribution of dominant drivers to predicted nitrogen fertilizer intensity and indicates which drivers were designated as perturbable in scenario construction. See Methods: Model attribution and global driver scores, Driver consolidation and perturbability rules, and Percentile mapping and SHAP-guided constrained scenario optimisation.



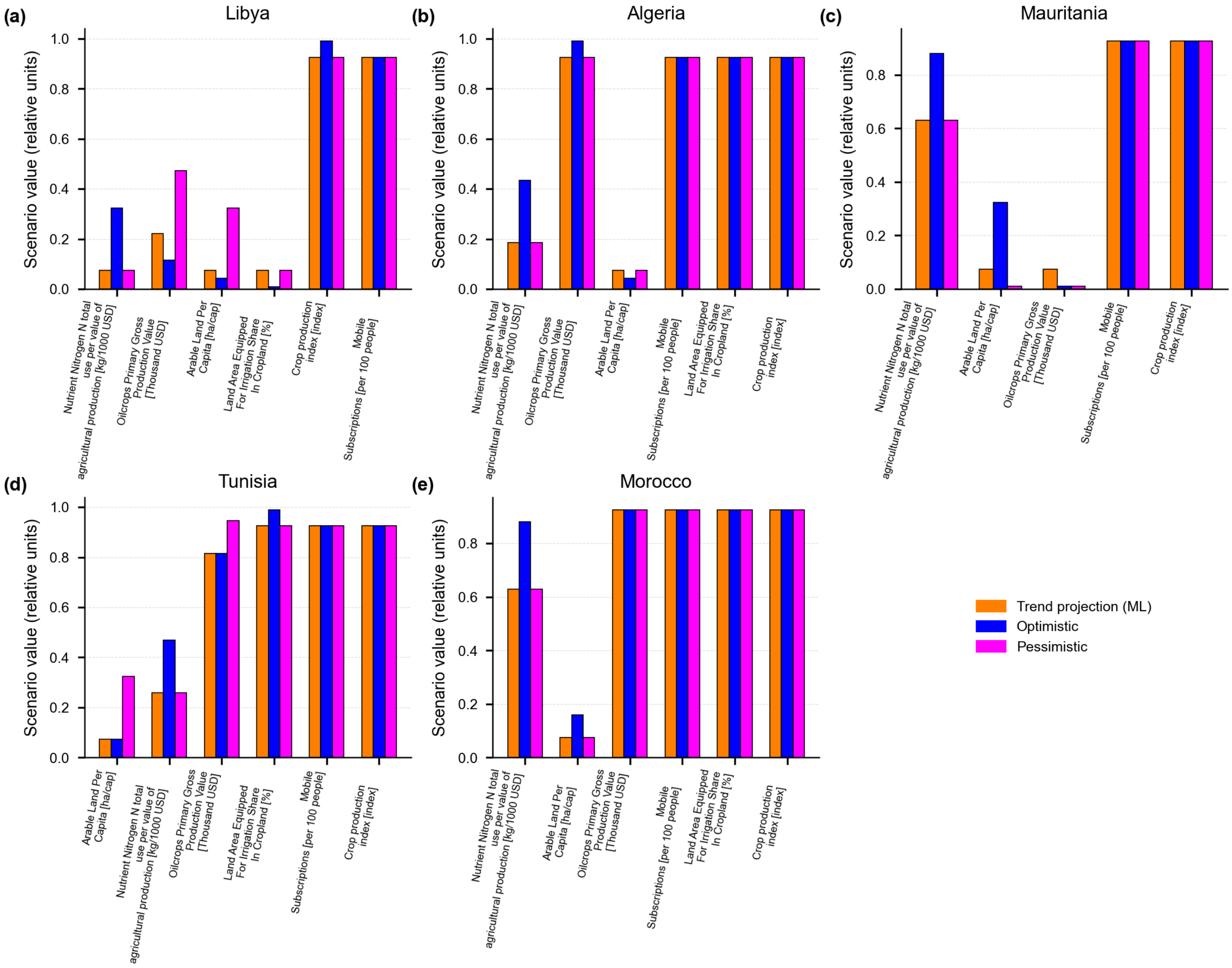
Supplementary Fig. S19 | Regional SHAP driver portfolio and perturbable levers: Central Africa. SHAP-based driver portfolio for the pruned regional model. The figure summarizes the relative contribution of dominant drivers to predicted nitrogen fertilizer intensity and indicates which drivers were designated as perturbable in scenario construction. See Methods: Model attribution and global driver scores, Driver consolidation and perturbability rules, and Percentile mapping and SHAP-guided constrained scenario optimisation.



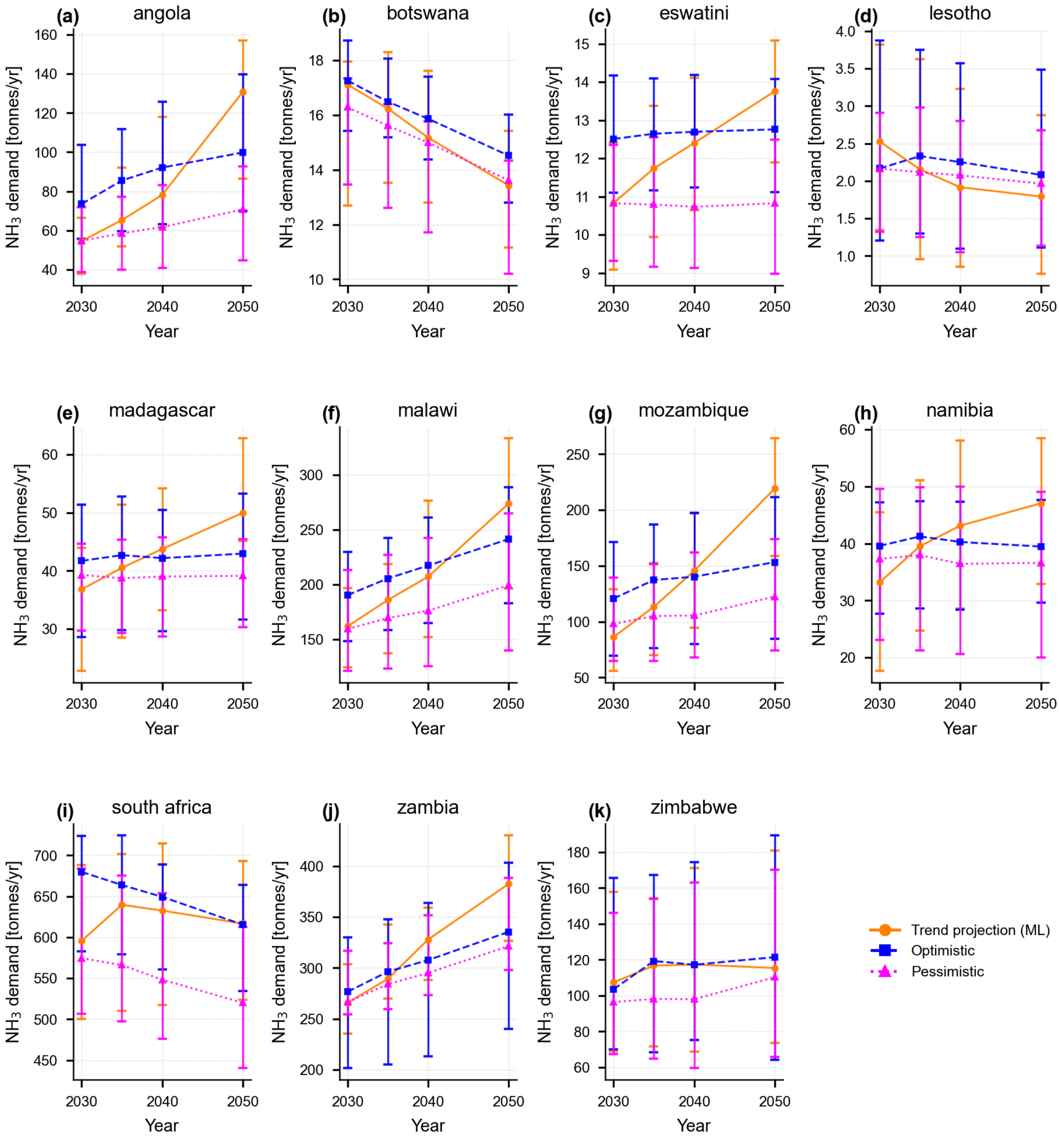
Supplementary Fig. S20 | Regional SHAP driver portfolio and perturbable levers: East Africa. SHAP-based driver portfolio for the pruned regional model. The figure summarizes the relative contribution of dominant drivers to predicted nitrogen fertilizer intensity and indicates which drivers were designated as perturbable in scenario construction. See Methods: Model attribution and global driver scores, Driver consolidation and perturbability rules, and Percentile mapping and SHAP-guided constrained scenario optimisation.



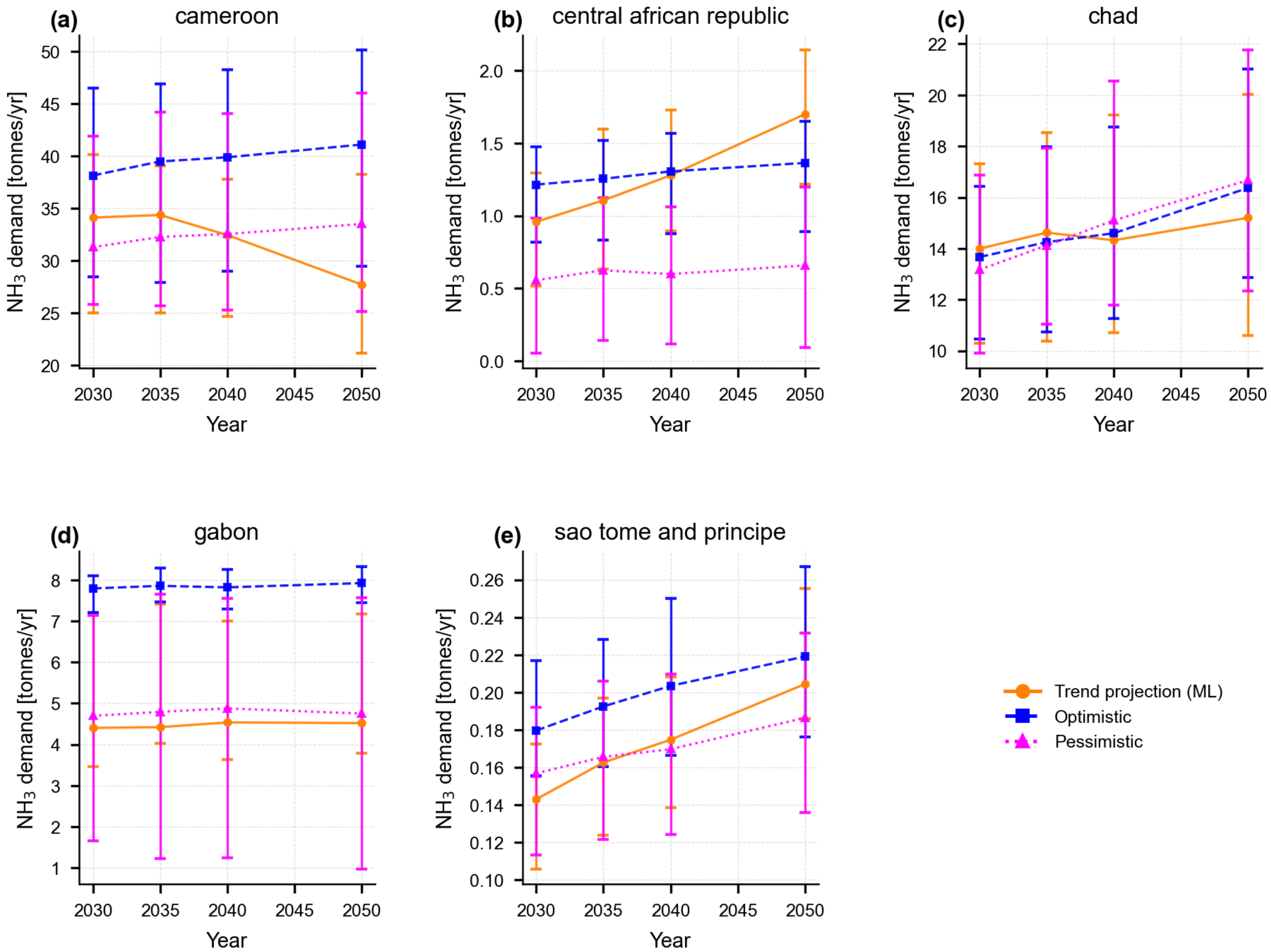
Supplementary Fig. S21 | Regional SHAP driver portfolio and perturbable levers: West Africa. SHAP-based driver portfolio for the pruned regional model. The figure summarizes the relative contribution of dominant drivers to predicted nitrogen fertilizer intensity and indicates which drivers were designated as perturbable in scenario construction. See Methods: Model attribution and global driver scores, Driver consolidation and perturbability rules, and Percentile mapping and SHAP-guided constrained scenario optimisation.



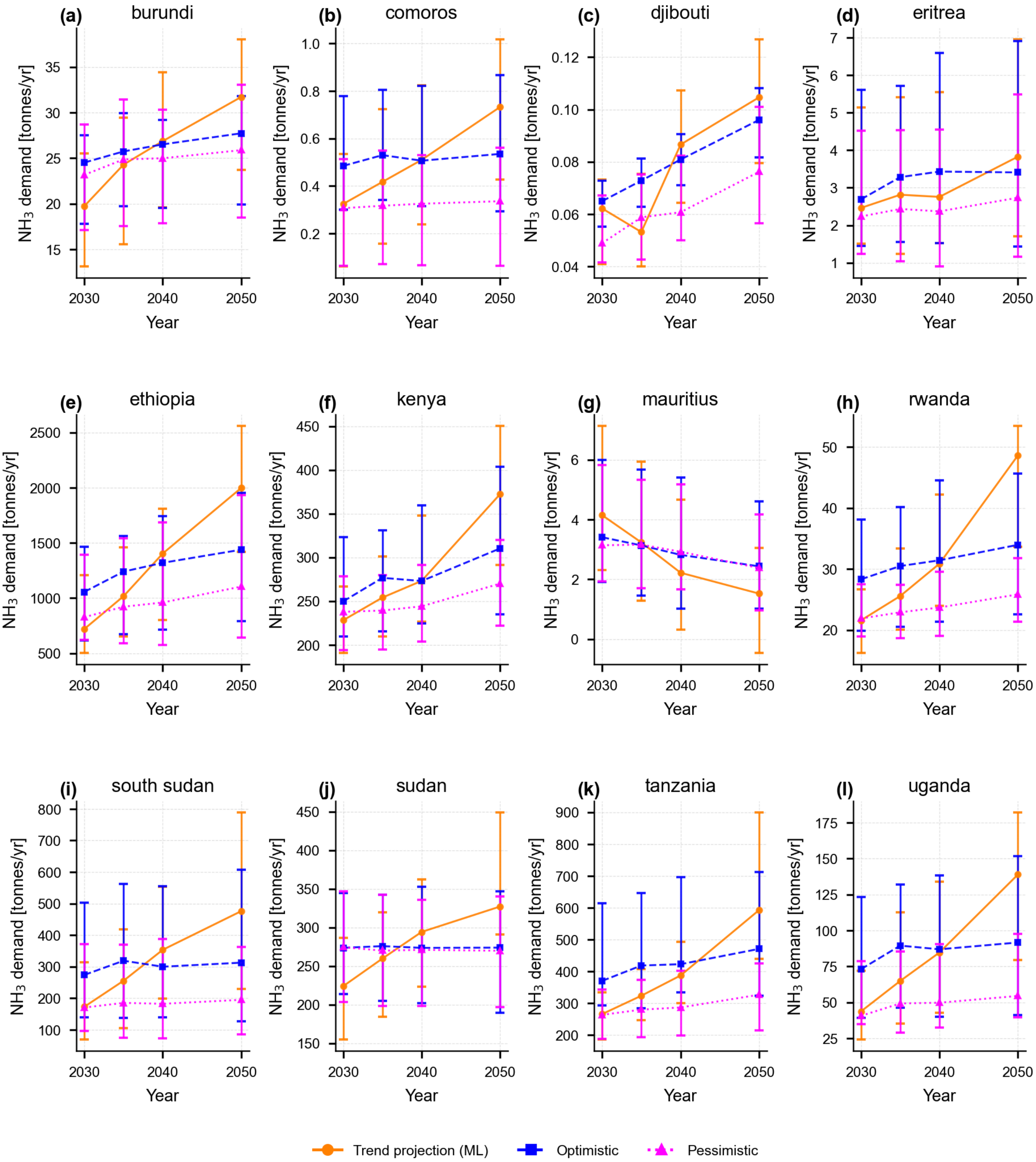
Supplementary Fig. S22 | Regional SHAP driver portfolio and perturbable levers: North Africa. SHAP-based driver portfolio for the pruned regional model. The figure summarizes the relative contribution of dominant drivers to predicted nitrogen fertilizer intensity and indicates which drivers were designated as perturbable in scenario construction. See Methods: Model attribution and global driver scores, Driver consolidation and perturbability rules, and Percentile mapping and SHAP-guided constrained scenario optimisation.



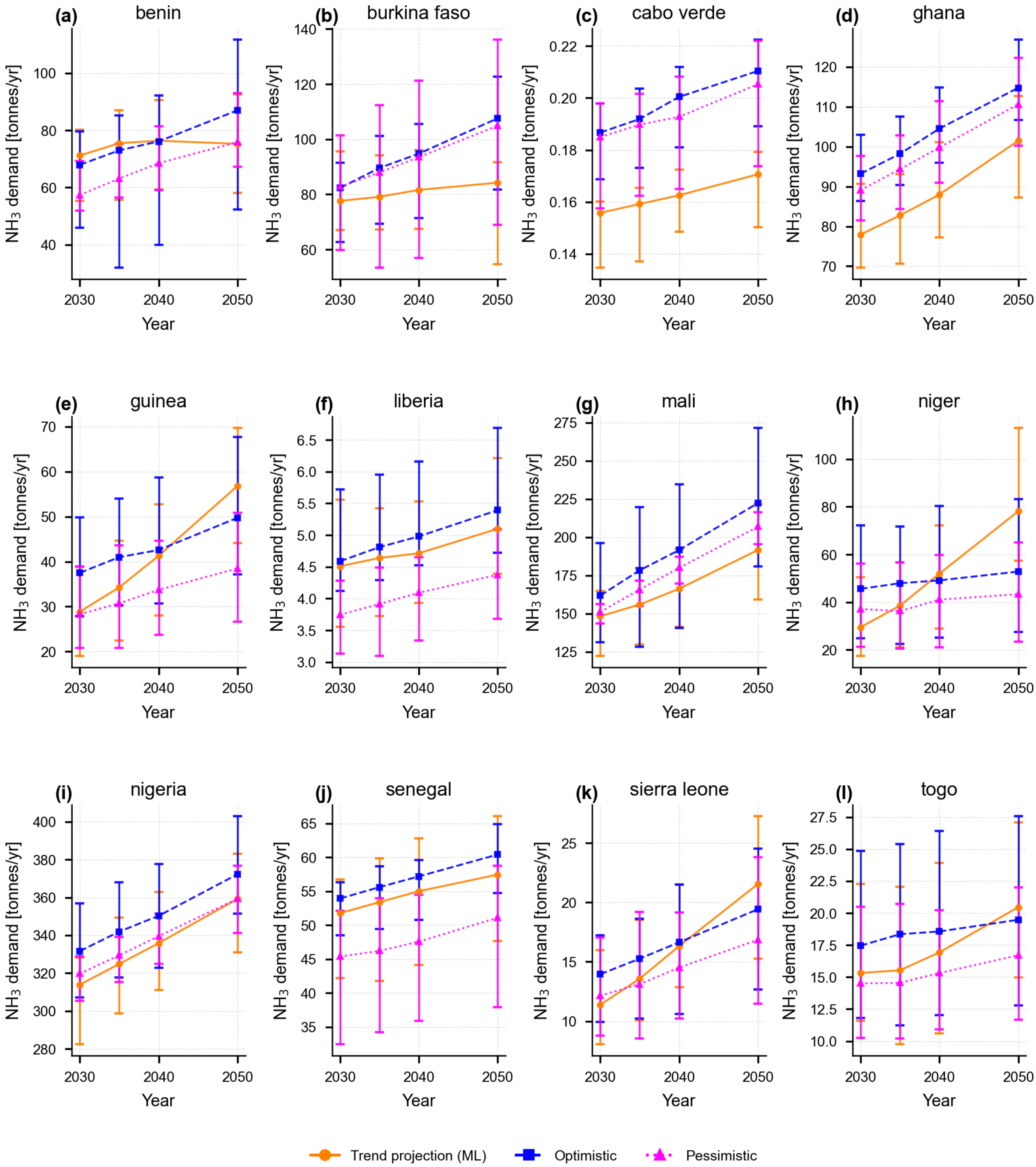
Supplementary Fig. S23 | Ammonia-equivalent demand implied by fertilizer-intensity scenarios: Southern Africa. Ammonia-equivalent totals (NH_3) implied by fertilizer-intensity scenario trajectories after scaling by national cropland area and applying the molecular-weight conversion from N to NH_3. Totals are shown for BAU, optimistic, and pessimistic scenarios. Uncertainty bands summarize uncertainty propagated from bootstrap model ensembles through the intensity-to-total conversion. See Methods: Equitable-diet cap, totals and ammonia translation and Uncertainty quantification and scenario robustness.



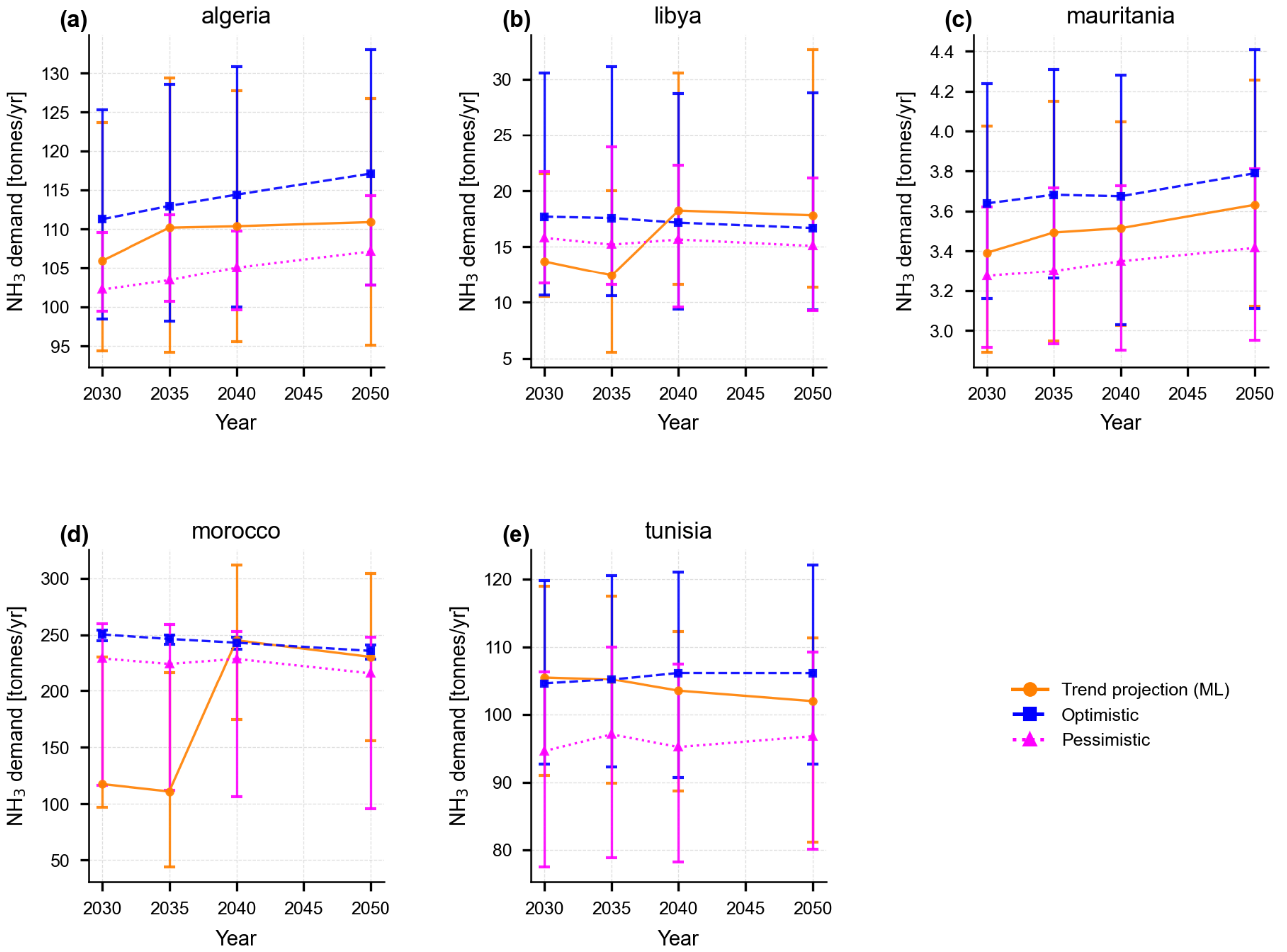
Supplementary Fig. S24 | Ammonia-equivalent demand implied by fertilizer-intensity scenarios: Central Africa. Ammonia-equivalent totals (NH_3) implied by fertilizer-intensity scenario trajectories after scaling by national cropland area and applying the molecular-weight conversion from N to NH_3. Totals are shown for BAU, optimistic, and pessimistic scenarios. Uncertainty bands summarize uncertainty propagated from bootstrap model ensembles through the intensity-to-total conversion. See Methods: Equitable-diet cap, totals and ammonia translation and Uncertainty quantification and scenario robustness.



Supplementary Fig. S25 | Ammonia-equivalent demand implied by fertilizer-intensity scenarios: East Africa. Ammonia-equivalent totals (NH_3) implied by fertilizer-intensity scenario trajectories after scaling by national cropland area and applying the molecular-weight conversion from N to NH_3. Totals are shown for BAU, optimistic, and pessimistic scenarios. Uncertainty bands summarize uncertainty propagated from bootstrap model ensembles through the intensity-to-total conversion. See Methods: Equitable-diet cap, totals and ammonia translation and Uncertainty quantification and scenario robustness.



Supplementary Fig. S26 | Ammonia-equivalent demand implied by fertilizer-intensity scenarios: West Africa. Ammonia-equivalent totals (NH_3) implied by fertilizer-intensity scenario trajectories after scaling by national cropland area and applying the molecular-weight conversion from N to NH_3. Totals are shown for BAU, optimistic, and pessimistic scenarios. Uncertainty bands summarize uncertainty propagated from bootstrap model ensembles through the intensity-to-total conversion. See Methods: Equitable-diet cap, totals and ammonia translation and Uncertainty quantification and scenario robustness.



Supplementary Fig. S27 | Ammonia-equivalent demand implied by fertilizer-intensity scenarios: North Africa. Ammonia-equivalent totals (NH_3) implied by fertilizer-intensity scenario trajectories after scaling by national cropland area and applying the molecular-weight conversion from N to NH_3. Totals are shown for BAU, optimistic, and pessimistic scenarios. Uncertainty bands summarize uncertainty propagated from bootstrap model ensembles through the intensity-to-total conversion. See Methods: Equitable-diet cap, totals and ammonia translation and Uncertainty quantification and scenario robustness.

Supplementary Tables

Supplementary Table 1 | Data harmonization, inclusion, and preprocessing rules. Rules applied prior to model training.

Rule	Implementation setting
Yearwindowretained	YEAR_MIN=1990,YEAR_MAX=2022
Coveragefilter	COVERAGE_THRESHOLD=50
Minimumcompletenessforcountry-series	MIN_COMPLETE_RATIO=0.7
Interpolationofnumericfeatures	INTERP_METHOD=linear
Varianceandcollinearitypruning	LOW_VARIANCE_TOPFRACTION=0.98;HIGH_CORR_THRESHOLD=0.98;VIF_THRESHOLD=10.0
Train/testsplit	TRAIN_FRACTION=0.7;RANDOM_STATE=42

Supplementary Table 2 | Regional training metadata. For each region: number of countries, years included, observations, train/test split, and cross-validation design applied prior to training.

Region	Countries	Years	Obs.	Train/Test	Cross-validation (CV)
Central Africa	7	1990–2022	205	143/62	5-fold KFold; shuffle=True; random_state=42; train_fraction=0.7
East Africa	12	1990–2022	345	241/104	5-fold KFold; shuffle=True; random_state=42; train_fraction=0.7
North Africa	5	1990–2022	165	115/50	5-fold KFold; shuffle=True; random_state=42; train_fraction=0.7
Southern Africa	11	1990–2022	343	240/103	5-fold KFold; shuffle=True; random_state=42; train_fraction=0.7
West Africa	14	1990–2022	440	308/132	5-fold KFold; shuffle=True; random_state=42; train_fraction=0.7

Supplementary Table 3 | Countries by region. Countries that met data harmonization, inclusion, and preprocessing rules in Supplementary Table 1.

Region	Country	Inclusion	Avg decade cov (%)	Target non-null	Mean row completeness
Central Africa	Cameroon	Yes	95.83	1.00	0.97
Central Africa	Central African Republic	Yes	89.18	1.00	0.95
Central Africa	Chad	Yes	91.40	1.00	0.96
Central Africa	Congo, Dem. Rep.	Yes	88.50	0.97	0.96
Central Africa	Congo, Rep.	Yes	83.10	1.00	0.95
Central Africa	Gabon	Yes	68.35	1.00	0.92
Central Africa	Sao Tome and Principe	Yes	47.25	0.64	0.79
East Africa	Burundi	Yes	92.78	1.00	0.96
East Africa	Comoros	Yes	72.00	0.97	0.94
East Africa	Djibouti	Yes	50.35	0.64	0.85
East Africa	Eritrea	Yes	76.60	1.00	0.95
East Africa	Ethiopia	Yes	94.28	1.00	0.94
East Africa	Kenya	Yes	95.75	1.00	0.97
East Africa	Mauritius	Yes	74.78	1.00	0.87
East Africa	Rwanda	Yes	92.40	0.94	0.96
East Africa	South Sudan	Yes	37.95	1.00	0.90
East Africa	Sudan	Yes	61.17	1.00	0.84
East Africa	Tanzania	Yes	96.20	1.00	0.97
East Africa	Uganda	Yes	87.25	1.00	0.97
North Africa	Algeria	Yes	84.40	1.00	0.97
North Africa	Libya	Yes	65.30	1.00	0.97
North Africa	Mauritania	Yes	85.47	0.97	0.97
North Africa	Morocco	Yes	96.23	1.00	0.97
North Africa	Tunisia	Yes	68.08	1.00	0.90
Southern Africa	Angola	Yes	89.38	0.97	0.96
Southern Africa	Botswana	Yes	77.00	1.00	0.95
Southern Africa	Eswatini	Yes	83.38	1.00	0.95
Southern Africa	Lesotho	Yes	64.83	1.00	0.93
Southern Africa	Madagascar	Yes	89.72	1.00	0.97
Southern Africa	Malawi	Yes	90.30	1.00	0.94
Southern Africa	Mozambique	Yes	96.17	1.00	0.97
Southern Africa	Namibia	Yes	78.53	0.76	0.95
Southern Africa	South Africa	Yes	93.25	1.00	0.97
Southern Africa	Zambia	Yes	90.80	1.00	0.93
Southern Africa	Zimbabwe	Yes	94.85	1.00	0.97
West Africa	Benin	Yes	93.80	1.00	0.94
West Africa	Burkina Faso	Yes	95.05	1.00	0.95
West Africa	Cabo Verde	Yes	70.65	0.65	0.90
West Africa	Gambia, The	Yes	81.23	1.00	0.92
West Africa	Ghana	Yes	88.20	1.00	0.92
West Africa	Guinea	Yes	91.83	1.00	0.94
West Africa	Guinea-Bissau	Yes	79.93	1.00	0.91
West Africa	Liberia	Yes	57.03	0.62	0.79
West Africa	Mali	Yes	95.83	1.00	0.97
West Africa	Niger	Yes	96.25	1.00	0.97
West Africa	Nigeria	Yes	95.18	1.00	0.95
West Africa	Senegal	Yes	92.60	0.97	0.94
West Africa	Sierra Leone	Yes	87.12	1.00	0.92
West Africa	Togo	Yes	83.82	1.00	0.92

Supplementary Table 4 | Feature sets used for initial regional training splits (baseline, pre-pruning).

Region	Feature label	Unit	Data source
Central Africa	Agricultural Land Share In Land Area	%	[1]
Central Africa	Arable Land	%	[2]
Central Africa	Cereals Primary Area Harvested	ha	[1]
Central Africa	Crop production index	index	[2]
Central Africa	Cropland Area Per Capita	ha/cap	[1]
Central Africa	GDP per capita	USD/cap	[2]
Central Africa	Land Area Equipped For Irrigation Share In Cropland	%	[1]
Central Africa	Livestock production index	index	[2]
Central Africa	Maize Corn Yield	t/ha	[1]
Central Africa	Mobile Subscriptions	per 100 people	[2]
Central Africa	Nutrient Nitrogen N Total Synthetic Fertilizers Emissions N2O	kt	[1]
Central Africa	Nutrient Nitrogen N Total Use Per Area Of Cropland	kg/ha	[1]
Central Africa	Nutrient Nitrogen N total use per value of agricultural production	kg/1000 USD	[1]
Central Africa	Population Growth	%	[2]
Central Africa	Temperature Change	°C	[1]
Central Africa	Unemployment Male	%	[2]
East Africa	Agricultural Land Share In Land Area	%	[1]
East Africa	Cereals Primary Area Harvested	ha	[1]
East Africa	Cereals Primary Yield	kg/ha	[1]
East Africa	Cropland Area Per Capita	ha/cap	[1]
East Africa	Cropland Share In Land Area	%	[1]
East Africa	Forest Land Share In Land Area	%	[1]
East Africa	Land Area Equipped For Irrigation Share In Cropland	%	[1]
East Africa	Maize Corn Area Harvested	ha	[1]
East Africa	Nutrient Nitrogen N Total Synthetic Fertilizers Emissions N2O	kt	[1]
East Africa	Nutrient Nitrogen N Total Use Per Area Of Cropland	kg/ha	[1]
East Africa	Nutrient Nitrogen N Total Use Per Capita	kg/cap	[1]
East Africa	Planted Forest Area	1000 ha	[1]
North Africa	Arable Land Per Capita	ha/cap	[2]
North Africa	Cereal Yield	t/ha	[2]
North Africa	Crop production index	index	[2]
North Africa	GDP per capita	USD/cap	[2]
North Africa	Hh Consumption GDP	%	[2]
North Africa	Land Area Equipped For Irrigation Share In Cropland	%	[1]
North Africa	Mobile Subscriptions	per 100 people	[2]
North Africa	Nutrient Nitrogen N Total Use Per Area Of Cropland	kg/ha	[1]
North Africa	Nutrient Nitrogen N total use per value of agricultural production	kg/1000 USD	[1]
North Africa	Oilcrops Primary Gross Production Value	Thousand USD	[1]
North Africa	Population Growth	%	[2]
North Africa	Potatoes Area Harvested	ha	[1]
North Africa	Pulses Total Production	ton	[1]
North Africa	Pulses Total Yield	kg/ha	[1]
North Africa	Temperature Change	°C	[1]
North Africa	Unemployment Male	%	[2]
Southern Africa	Arable Land Per Capita	ha/cap	[2]
Southern Africa	Crop production index	index	[2]
Southern Africa	Cropland Share In Land Area	%	[1]
Southern Africa	Forest Land Area	1000 ha	[1]
Southern Africa	GDP per capita	USD/cap	[2]

Continued in SuppFig. S28

Supplementary Table 5 | CatBoost hyperparameters for each regional model (pre-pruning baseline). Best hyperparameters selected by cross-validated optimization.

Region	bootstrap_type	bag_temp.	subsample	depth	iter.	learn_rate	l2_reg	seed	early_stop
Central Africa	Bernoulli	—	0.726069	4	1000	0.0624687	9.59902	42	200
East Africa	Bayesian	1.45387	—	4	1000	0.0749453	9.4131	42	200
North Africa	Bayesian	0.250153	—	4	1000	0.0295899	3.11024	42	200
Southern Africa	Bernoulli	—	0.791718	4	1000	0.0216928	1.0022	42	200
West Africa	Bernoulli	—	0.980554	4	1000	0.0972164	7.7892	42	200

Supplementary Table 6 | CatBoost hyperparameters for each regional model (post-pruning). For each region, the pruning method with the smallest test RMSE was selected as the best method.

Region	Pruning	RMSE	bootstrap_type	subsample	depth	iter.	learn_rate	l2_reg	seed	early_stop
Central Africa	SHAPGreedy	0.2072	Bernoulli	0.928433	6	1000	0.0321968	1.00072	42	200
East Africa	BPSO+SHAP	2.1375	Bayesian	—	4	1000	0.0506059	2.45156	42	200
North Africa	BPSO	1.1549	MVS	—	4	1000	0.0602065	1.64213	42	200
Southern Africa	BPSO+SHAP	1.5499	Bernoulli	0.68094	4	1000	0.0397461	1.04262	42	200
West Africa	BPSO	1.0502	Bernoulli	0.91215	7	1000	0.0809735	2.15675	42	200

Supplementary Table 7 | Feature pruning and search configuration parameters. Parameters used for SHAP-based pruning techniques.

Pruning method	Parameter	Value
BPSO	generations	100
BPSO	min_features	5
BPSO	inertia	0.7
BPSO	c1	1.5
BPSO	c2	1.5
BPSO	mutation_rate	0.02
BPSO	random_state	None
BPSO+SHAP	generations	100
BPSO+SHAP	min_features	5
BPSO+SHAP	inertia	0.7
BPSO+SHAP	c1	1.5
BPSO+SHAP	c2	1.5
BPSO+SHAP	mutation_rate	0.02
BPSO+SHAP	use_shap	True
BPSO+SHAP	shap_min_prob	0.1
BPSO+SHAP	shap_max_prob	0.9
BPSO+SHAP	lambda_shap	0.1
SHAP	min_features	5
SHAPGreedy	max_iter	100
SHAPGreedy	drop_per_iter	1
SHAPGreedy	patience	20
SHAPGreedy	min_features	5
SHAPGreedy	pre_prune_top	30
SHAPStochastic	max_iter	100
SHAPStochastic	drop_per_iter	1
SHAPStochastic	drop_candidate_pool_size	3
SHAPStochastic	patience	20
SHAPStochastic	min_features	5
SHAPStochastic	pre_prune_top	30
SHAPStochastic	random_state	None

Supplementary Table 8 | Post-pruning features (all regions and methods).

Region	Pruning method	Feature label	Unit
Central Africa	BPSO	Cereals Primary Area Harvested	ha
Central Africa	BPSO	Crop production index	index
Central Africa	BPSO	Cropland Area Per Capita	ha/cap
Central Africa	BPSO	Land Area Equipped For Irrigation Share In Cropland	%
Central Africa	BPSO	Maize Corn Yield	t/ha
Central Africa	BPSO	Mobile Subscriptions	per 100 people
Central Africa	BPSO	Nutrient Nitrogen N Total Synthetic Fertilizers Nitrogen Dioxide Emissions	kt
Central Africa	BPSO	Nutrient Nitrogen N total use per value of agricultural production	kg/1000 USD
Central Africa	BPSO	Temperature Change	degC
Central Africa	BPSO	Unemployment Male	%
Central Africa	SHAPGreedy	GDP per capita	USD/cap
Central Africa	SHAPGreedy	Land Area Equipped For Irrigation Share In Cropland	%
Central Africa	SHAPGreedy	Nutrient Nitrogen N Total Use Per Capita	kg/cap
Central Africa	SHAPGreedy	Nutrient Nitrogen N total use per value of agricultural production	kg/1000 USD
Central Africa	SHAPGreedy	Temperature Change	degC
Central Africa	SHAP	Agricultural Land Share In Land Area	%
Central Africa	SHAP	Arable Land	%
Central Africa	SHAP	Arable Land Per Capita	ha/cap
Central Africa	SHAP	Cereals Primary Area Harvested	ha
Central Africa	SHAP	Crop production index	index
Central Africa	SHAP	Cropland Area Per Capita	ha/cap
Central Africa	SHAP	GDP per capita	USD/cap
Central Africa	SHAP	Land Area Equipped For Irrigation Share In Cropland	%
Central Africa	SHAP	Mobile Subscriptions	per 100 people
Central Africa	SHAP	Nutrient Nitrogen N Total Synthetic Fertilizers Nitrogen Dioxide Emissions	kt
Central Africa	SHAP	Nutrient Nitrogen N Total Use Per Capita	kg/cap
Central Africa	SHAP	Nutrient Nitrogen N total use per value of agricultural production	kg/1000 USD
Central Africa	SHAPStochastic	GDP per capita	USD/cap
Central Africa	SHAPStochastic	Land Area Equipped For Irrigation Share In Cropland	%
Central Africa	SHAPStochastic	Nutrient Nitrogen N Total Use Per Capita	kg/cap
Central Africa	SHAPStochastic	Nutrient Nitrogen N total use per value of agricultural production	kg/1000 USD
Central Africa	SHAPStochastic	Temperature Change	degC
Central Africa	BPSO+SHAP	Arable Land	%
Central Africa	BPSO+SHAP	Cereals Primary Area Harvested	ha
Central Africa	BPSO+SHAP	Crop production index	index
Central Africa	BPSO+SHAP	Cropland Area Per Capita	ha/cap
Central Africa	BPSO+SHAP	Land Area Equipped For Irrigation Share In Cropland	%
Central Africa	BPSO+SHAP	Maize Corn Yield	t/ha
Central Africa	BPSO+SHAP	Mobile Subscriptions	per 100 people
Central Africa	BPSO+SHAP	Nutrient Nitrogen N Total Synthetic Fertilizers Nitrogen Dioxide Emissions	kt
Central Africa	BPSO+SHAP	Nutrient Nitrogen N total use per value of agricultural production	kg/1000 USD
Central Africa	BPSO+SHAP	Population Growth	%
Central Africa	BPSO+SHAP	Temperature Change	degC
Central Africa	BPSO+SHAP	Unemployment Male	%
East Africa	BPSO	Agricultural Land Share In Land Area	%
East Africa	BPSO	Cropland Area Per Capita	ha/cap
East Africa	BPSO	Forest Land Share In Land Area	%
East Africa	BPSO	Land Area Equipped For Irrigation Share In Cropland	%
East Africa	BPSO	Nutrient Nitrogen N Total Synthetic Fertilizers Nitrogen Dioxide Emissions	kt

Continued in SuppFig. S29

Supplementary Table 9 | Scenario search levers and uncertainty-quantification settings. Scenario percentile bounds and perturbability rules, plus bootstrap and uncertainty safeguards.

Section	Parameter	Value
`Scenariosearch`	`q_low`	`0.01`
`Scenariosearch`	`q_high`	`0.99`
`Scenariosearch`	`max_pct_shift`	`0.25`
`Scenariosearch`	`n_grid`	`41`
`Scenariosearch`	`perturbability_rule`	Non-perturbable when name tokens indicate population-like variables (tokens including `population` or prefixes `pop*`) unemployment, or temperature. Tokens `urban` and `rural` are explicitly allowed.
`Bootstrap/uncertainty`	`n_bootstrap`	`200`
`Bootstrap/uncertainty`	`seed_base`	`42`

Supplementary Notes

Feature pruning yields compact and robust regional models

To support credible long-term projections and scenario analysis, regional models must generalize well. Across regions, feature pruning produced more compact models and reduced out-of-sample error by 22–41% (Supplementary Fig. S1a). The largest gains occurred in East and Southern Africa (41% reduction in RMSE), with additional improvements in Central (37%), West (35%) and North Africa (24%) (Supplementary Tables 6–7).

Absolute test RMSE values and the selected post-pruning configurations are reported in Supplementary Table 6 (for example, Central Africa improves from 0.33 to 0.21 kg ha^{-1} yr^{-1}; East Africa from 3.60 to 2.14 kg ha^{-1} yr^{-1}; Southern Africa from 2.64 to 1.55 kg ha^{-1} yr^{-1}; North Africa from 1.53 to 1.16 kg ha^{-1} yr^{-1}; and West Africa from 1.61 to 1.05 kg ha^{-1} yr^{-1}).

Training–test comparisons show that optimization-based pruning (BPSO-only and SHAP-guided BPSO) improves test performance without widening the train–test gap (Supplementary Fig. S1b), whereas SHAP rank-based pruning can modestly degrade generalization in some settings (e.g., East Africa; Supplementary Fig. S2). This pattern is consistent with rank-based removal discarding stabilizing predictors under distribution shift, while optimization-based pruning searches directly for feature subsets that retain predictive accuracy under the evaluation protocol (see Methods). Country-level error extremes for the selected regional models are shown in Supplementary Fig. S2.

Supplementary tables overview

Supplementary Tables S1–S9 provide implementation details for data harmonization, regional training configuration, pruning and scenario search settings, and uncertainty procedures referenced in the main-text Methods. Supplementary Figures provide additional diagnostics and regional decompositions.

Supplementary References

[1] Food and Agriculture Organization of the United Nations, *FAOSTAT Statistical Database*, 2025. [Online]. Available: https://www.fao.org/faostat/.

[2] The World Bank, *World Bank Development Indicators*, 2025. [Online]. Available: https://data.worldbank.org/indicator.