AgWise

Rwanda Rice fertilizer recommendation data analytics report

EiA in collaboration with RAB and CIP

INITIATIVE ON
Excellence in
Agronomy

AgWise

Fertiliser response in rice for Rwanda

STEP 1:	assemble data
STEP 2:	extract BLUPs (structural variation)
STEP 3:	determine indigenous nutrient supply (reverse QUEFTS)
STEP 4:	aggregate INS at marshland level
STEP 4b:	calculate fertilizer recommendation

CGIAR INITIATIVE ON Excellence in Agronomy

STEP 1: assembling data

INITIATIVE ON
Excellence in
Agronomy

Assembling available rice fertilizer trial data

Data is available from two experiments:

- Africa Rice trials: wet season (2015), 4 treatments (NPK, NP. NK, PK)
- RwaSIS fertilizer response trials (RS-PFR): 2 seasons (2022A and 2022B), 11 treatments (Increased NPK, NPK_120N, NPK_100N, NPK_40K, NPK_60P, NPK_45P, NPK_30P, NPK_17_3, NPK_20K, NPK_60N, Control)

A reference treatment was set for the two data sources differently:

- Africa Rice : NPK (160-25-70)
- RwaSIS season 1: increased NPK (120-45-40)
- RwaSIS season 2 NPK_120N (120-15-28)

AgWise analysis rice field trial data Rwanda

2 experiments available, of which 2 repeated during two consecutive seasons:

RS-PFR-1 (2022A, 2022B)

Treat	N	P	K
NPK 17*3	80	15	28
NPK 60N	60	15	28
NPK 100N	100	15	28
NPK 120N	120	15	28
NPK 30P	80	30	28
NPK 45P	80	45	28
NPK 60P	80	60	28
NPK 20K	80	15	20
NPK 40K	80	15	40
Increase NPK	120	45	40
Control	0	0	0

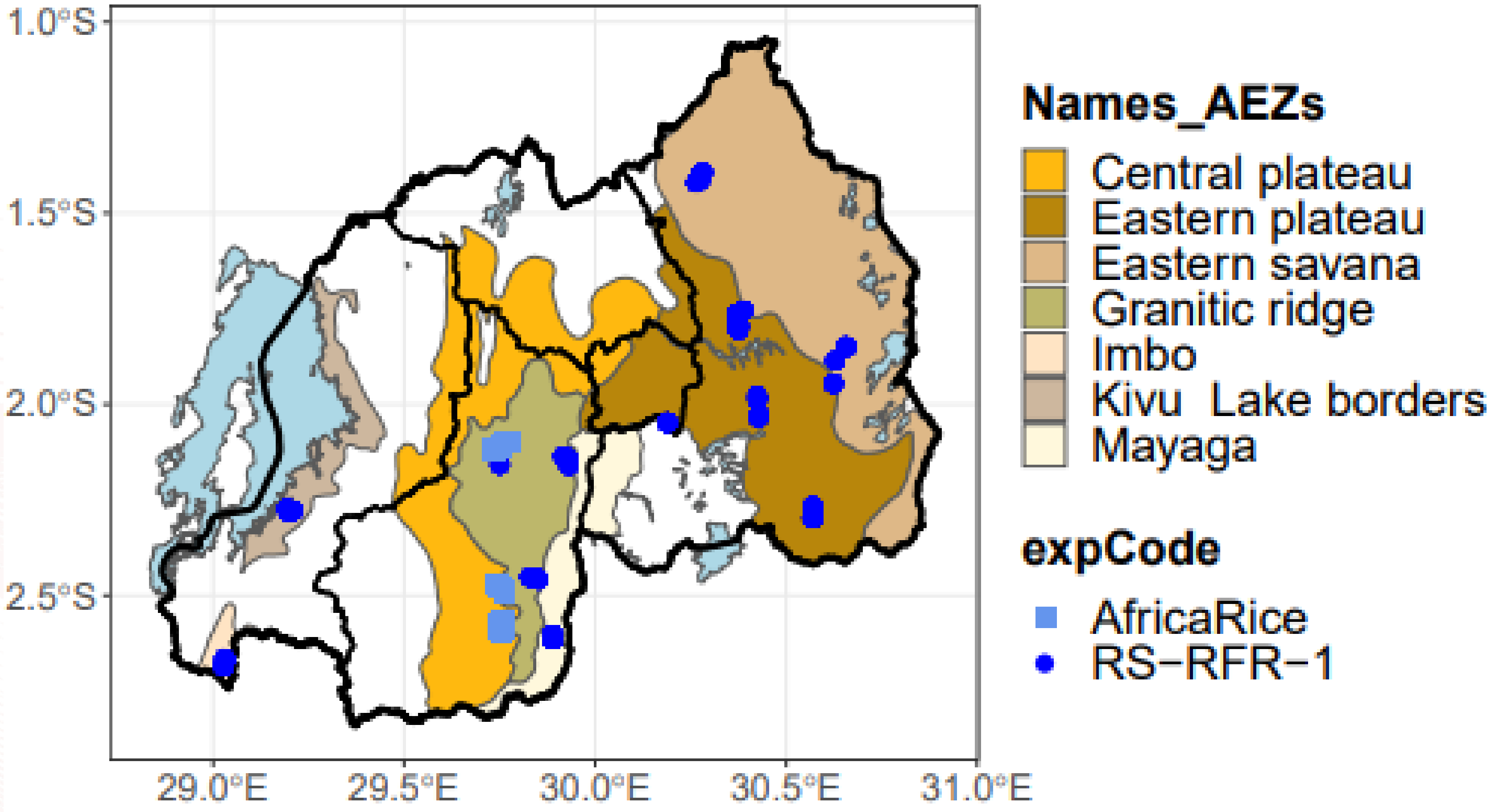
Reference treatment
(NPK increased,)

Africa Rice (2015 wet season)

Treat	N	P	K
NPK	160	25	70
NP	160	25	0
NK	160	0	70
PK	0	25	70

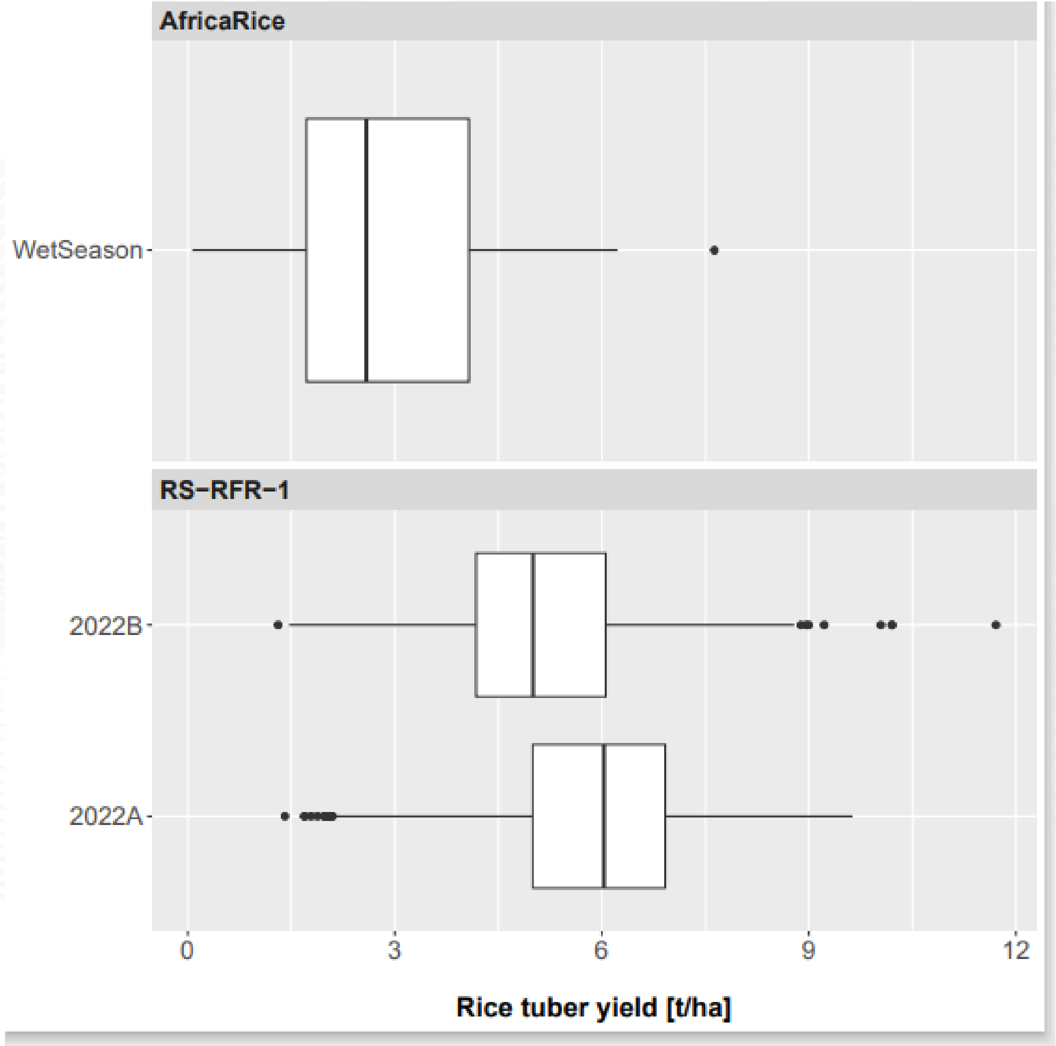
Reference treatment
(NPK)

AgWise analysis rice field trial data Rwanda

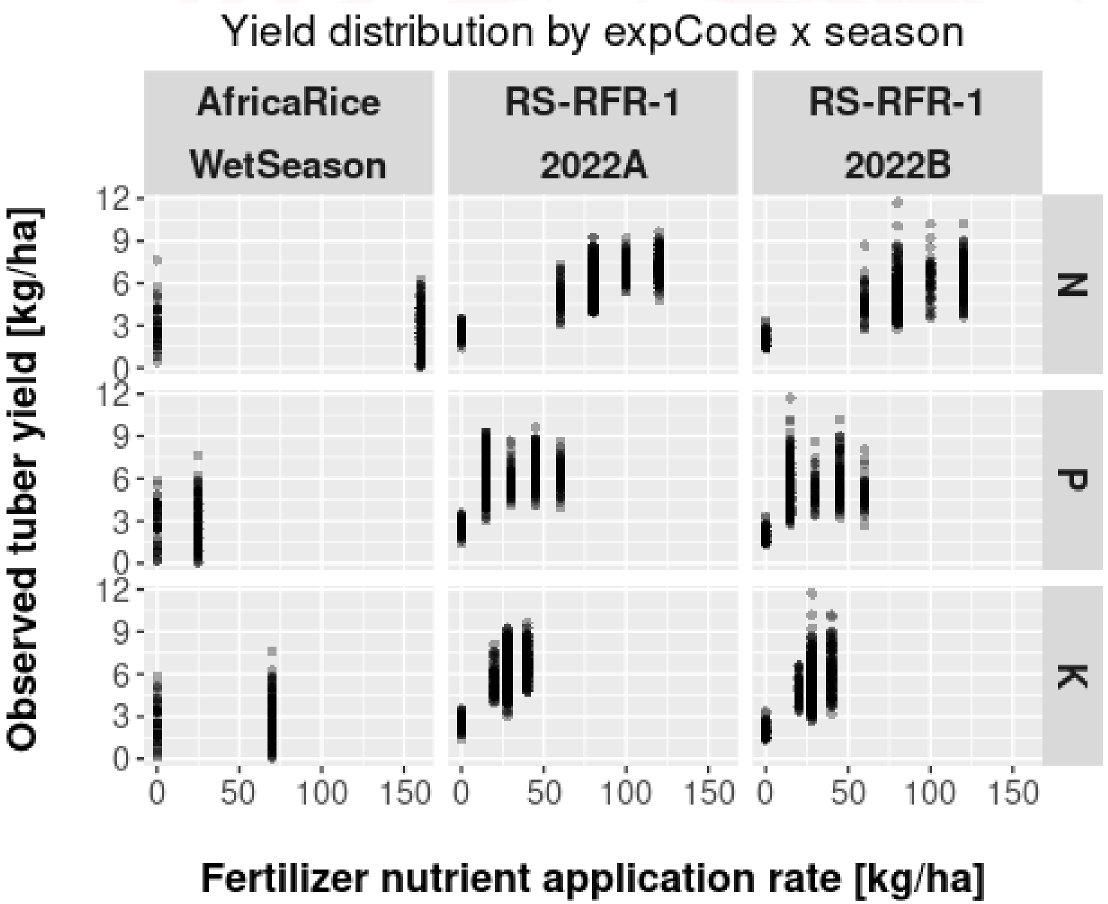


Overview yield distributions by expCode × season

Yield distributions are different in the two data, with Africa Rice data yields around 2.8 t/ha and in RwaSIS 2022A 5.9 t/ha and 2022B 5 t/ha:



Overview yield distributions by expCode × season



STEP 2: eliminate residual error through linear mixed effects modelling

Variation in the raw experimental data is composed of

- Overall treatment effect (application of N, P and K at varying rates)
- Season and experiment-specific effects (choice of variety, planting density, overall weather conditions,...)
- Trial-specific effects (location-specific differences in soil, weather and crop management conditions,...)
- Residual error (plot-specific effects, measurement error,...)

In order to evaluate the overall and location-specific effects of fertilizer application and relate this to digital soil map information and other predictors, the structural variation needs to be differentiated from the residual error. By eliminating the noise in the data and only carrying the meaningful signal forward, the performance of prediction models can be better evaluated.

This is done by fitting a linear mixed effects model:

```
sqrt(TY) ~ N + P + K + I(N^2) + I(P^2) + I(K^2) +
           N:P + N:K + P:K + N:P:K +
           season +
           (1|TLID) +
           (0 + N|TLID) + (0 + P|TLID) + (0 + K|TLID)
```

- Fixed effects of N, P and K (linear and quadratic terms in N, P and K)
- Fixed two- and three-way interactions between N, P and K
- Fixed seasonal effect (which allows for differences in yield between experiments and seasons)
- Random intercept for trial (allowing for differences in yield between individual trial locations)
- Random uncorrelated slopes for trial (allowing for different shapes in the response curve between trial locations)

A square root transformation is applied to ensure homoscedastic distribution of residuals.

The aim of this model is to describe the structural variation in yield, and eliminate the residual error. The model assumes a quadratic shape of the response curve, which is adequate if application rates are not too large to the extent that a plateau in the yield response curve is reached. In such situations a non-linear model with a logistic dose-response curve is required.

Performance of the linear mixed effects model

Linear mixed model fit by REML ['lmerModLmerTest']

```
Random effects:
 Groups   Name        Variance  Std.Dev.
 TLID     (Intercept) 0.02789   0.1670
 TLID.1   N100        0.02183   0.1477
 TLID.2   P100        0.02029   0.1424
 TLID.3   K100        0.08022   0.2832
 Residual             0.02137   0.1462
Number of obs: 1613, groups:  TLID, 171
```

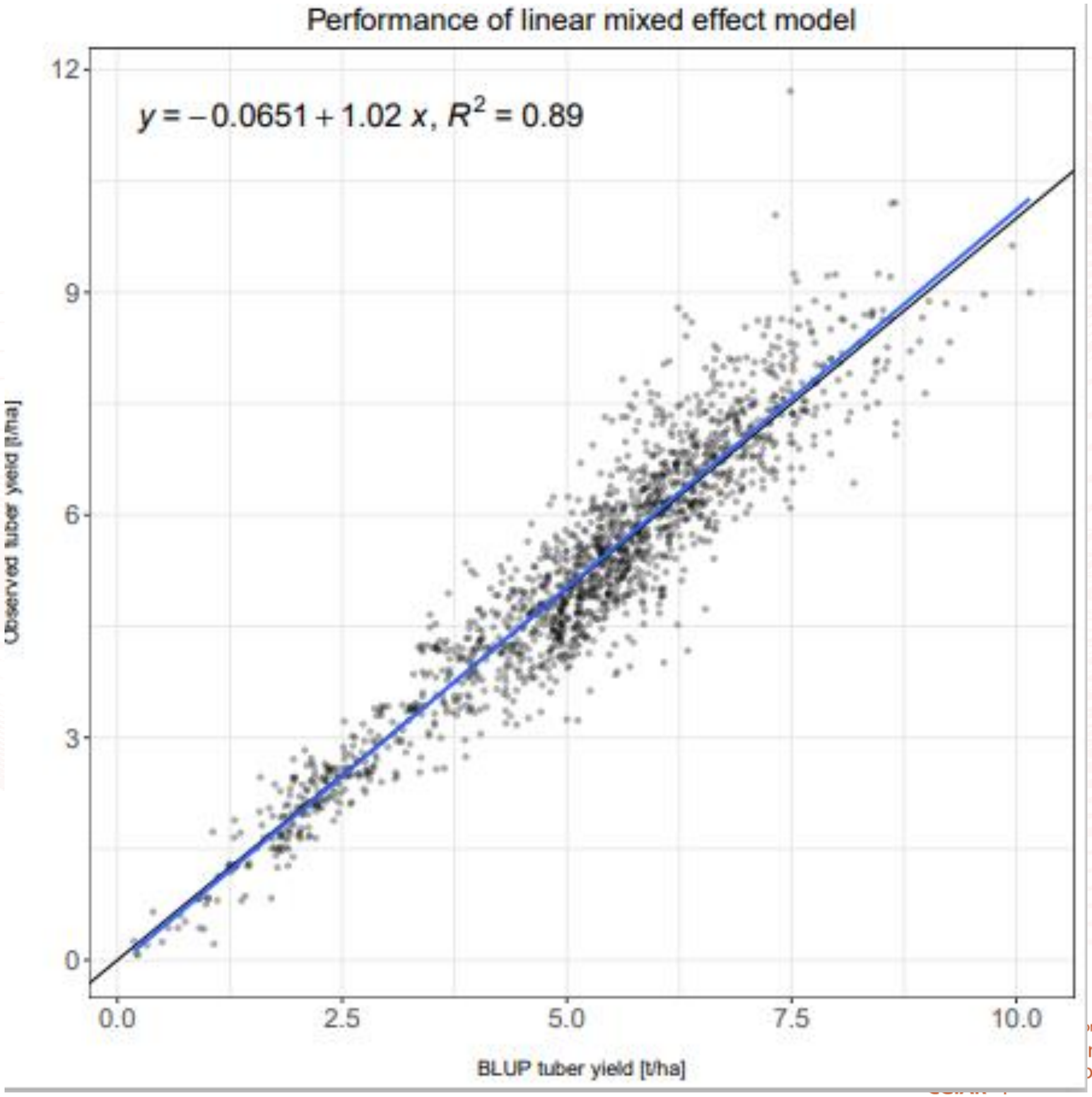
Type III Analysis of Variance Table with Satterthwaite's method

```
                Sum Sq Mean Sq  NumDFDenDF F value    Pr(>F)
N              0.95821 0.95821 1 1149.76 44.8411 3.335e-11 ***
P              0.05458 0.05458 1 1144.09 2.5543 0.1102688
K              0.11446 0.11446 1 1146.03 5.3565 0.0208218 *
I(N^2)         0.01010 0.01010 1 1111.49 0.4729 0.4918210
I(P^2)         0.05917 0.05917 1 1084.00 2.7691 0.0963908 .
I(K^2)         0.31942 0.31942 1 1091.22 14.9480 0.0001170 ***
season         2.44485 1.22242 2 188.44 57.2055 < 2.2e-16 ***
N:P            0.29033 0.29033 1 1163.03 13.5865 0.0002383 ***
N:K            0.37331 0.37331 1 1177.86 17.4698 3.135e-05 ***
P:K            0.00307 0.00307 1 1233.62 0.1437 0.7046532
N:P:K          0.00634 0.00634 1 1210.36 0.2966 0.5861084
```

Marginal R2 = 0.6176 (Fixed effects explain about 61% of the total variation)

Conditional R2 = 0.89 (Fixed + random effects explain 90% of variation, and 10% is residual error)

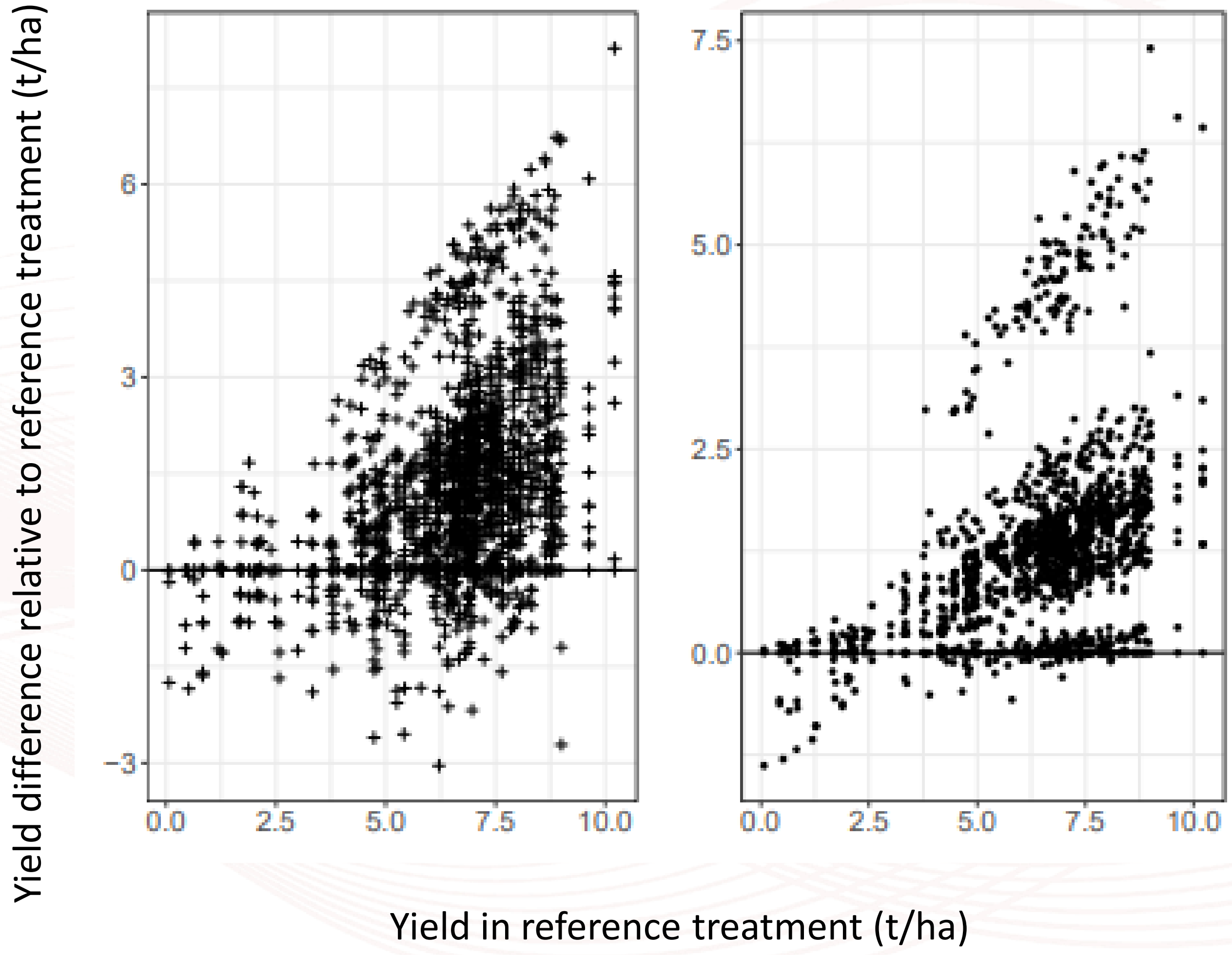
Performance of the linear mixed effects model



Why BLUPs?

BLUPs are Best Linear Unbiased Predictors. They “attenuate” the observed yield variation by distinguishing noise from signal, and represent the structural variation in the data. They assume that variation between locations, in terms of both mean yield (intercept) and yield response (slope) follow normal distributions. The procedure uses all data to make inferences about the structural (fixed + random terms) and residual variation.

BLUPs represent the structural variation in yield (and yield response)

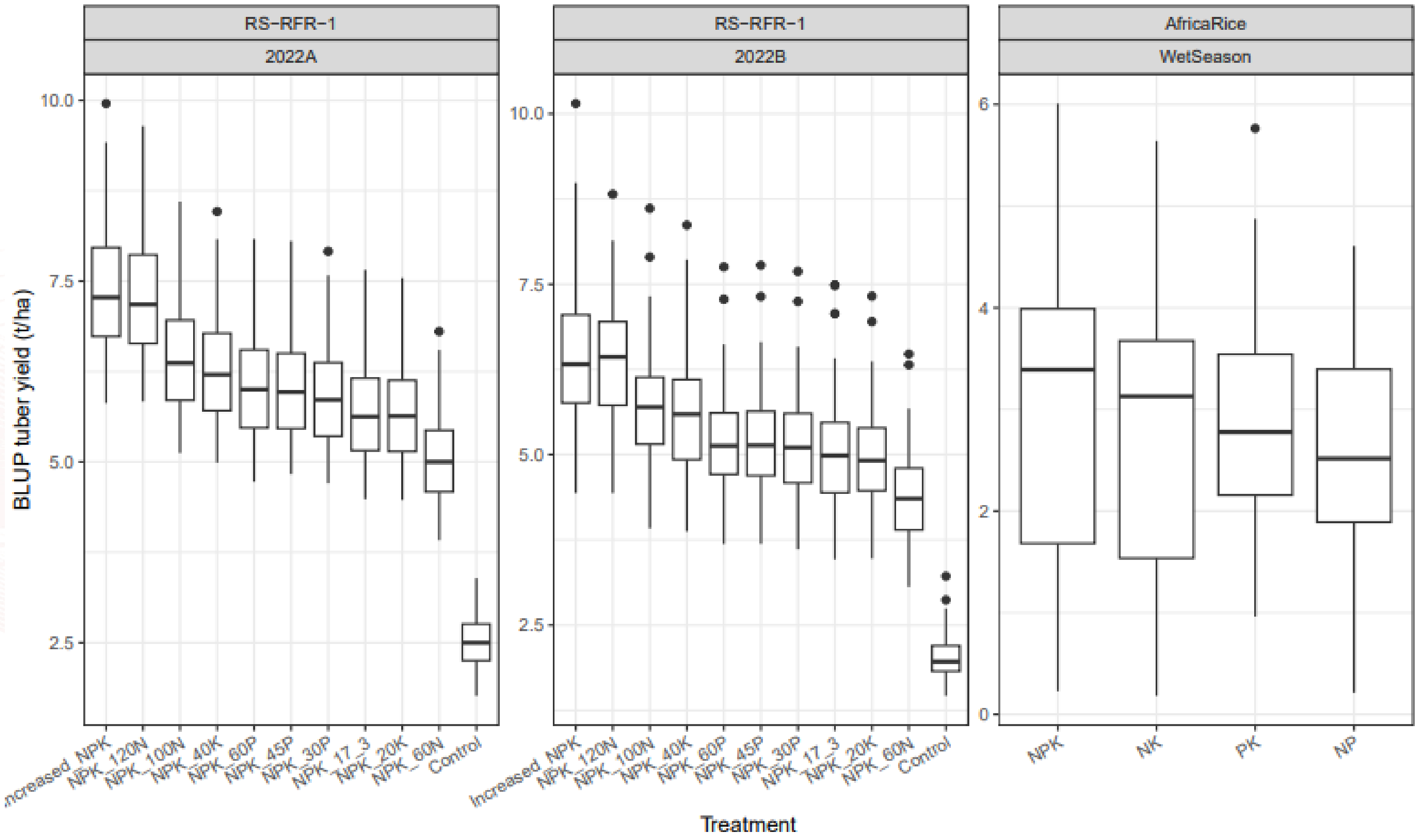


Comparing raw versus BLUP yield effects

We can calculate for each observation the yield difference to the reference treatment ($TY_{ref} - TY_{treat}$), and plot this yield difference against the yield in the reference treatment. We observe large variation in the raw observations, with frequent negative yield effects (meaning that yields in treatments with lower nutrient application have higher yields), which are not biologically meaningful. Variation in BLUP yield effects is lower, with fewer and smaller negative effects observed. It is therefore plausible that the mixed model approach has largely eliminated residual plot-level error (noise), and retained structural variation.

This can also be illustrated by looking at variation in yield response within individual trial locations (see next slide). BLUPs are closely aligned to raw observations in locations with consistent yield patterns. In locations with yield observations that deviate from overall patterns, BLUPs will attribute this as error and attenuate such values.

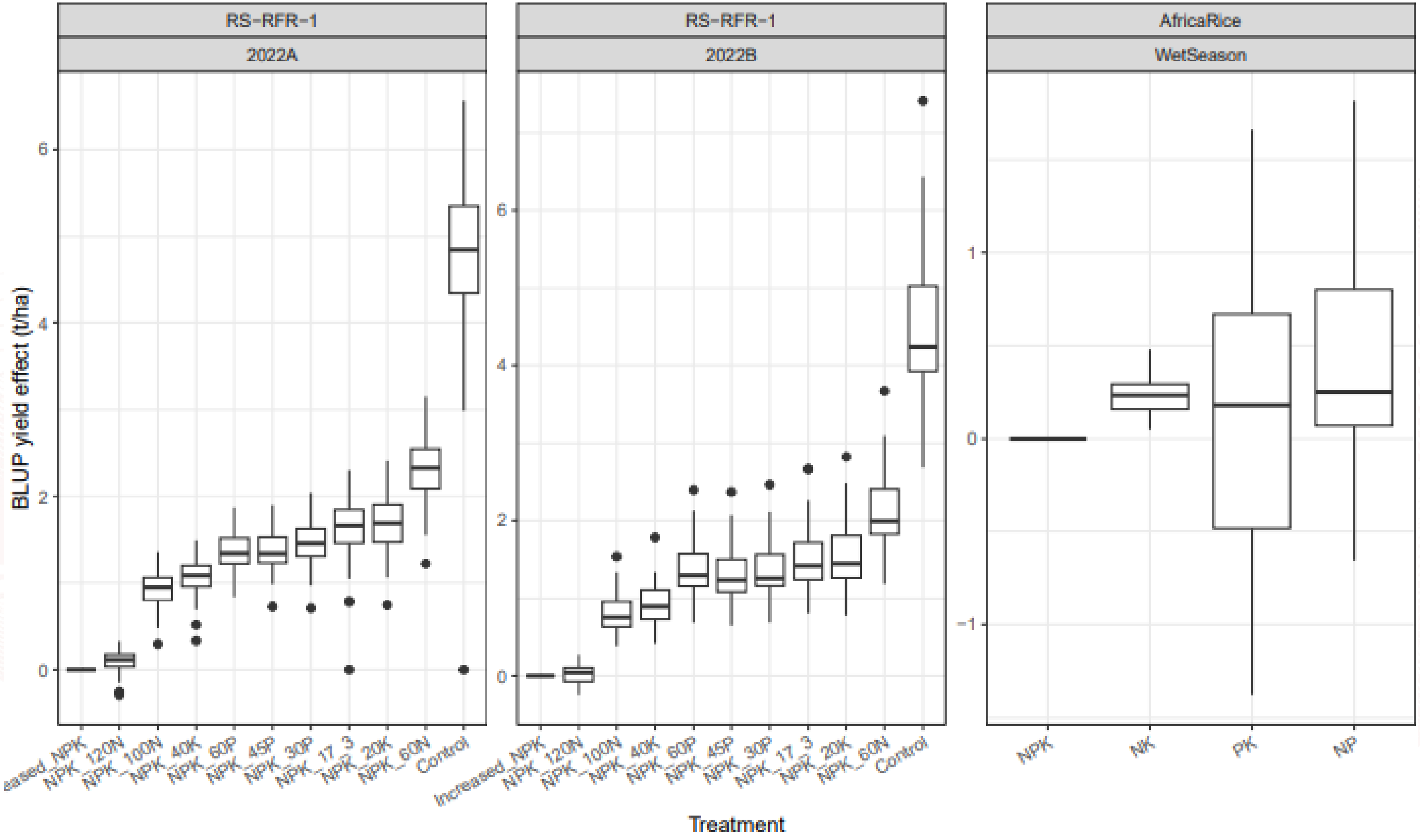
The performance of treatments, BLUPs



Treat	N	P	K
NPK 17*3	80	15	28
NPK 60N	60	15	28
NPK 100N	100	15	28
NPK 120N	120	15	28
NPK 30P	80	30	28
NPK 45P	80	45	28
NPK 60P	80	60	28
NPK 20K	80	15	20
NPK 40K	80	15	40
Increase NPK	120	45	40
Control	0	0	0

Treat	N	P	K
NPK	160	25	70
NP	160	25	0
NK	160	0	70
PK	0	25	70

The performance of treatments, BLUPs



Treat	N	P	K
NPK 17*3	80	15	28
NPK 60N	60	15	28
NPK 100N	100	15	28
NPK 120N	120	15	28
NPK 30P	80	30	28
NPK 45P	80	45	28
NPK 60P	80	60	28
NPK 20K	80	15	20
NPK 40K	80	15	40
Increase NPK	120	45	40
Control	0	0	0

Treat	N	P	K
NPK	160	25	70
NP	160	25	0
NK	160	0	70
PK	0	25	70

STEP 3: Determining indigenous nutrient supply

Reverse QUEFTS

The "reverse QUEFTS" is an optimisation algorithm that seeks the values for the soil indigenous N, P and K supply that best explains the observed yield responses to the fertilizer application rates in each individual field trial.

Data from different experiments with different treatment structures are standardised and each set of observations is described by a single set of INS, IPS and IKS values.

For every trial ID i:

TLID	Treat	N	P	K	blup
i	1	N1	P1	K1	Y1
i	2	N2	P2	K2	Y2
i	3	N3	P3	K3	Y3
i	...	...	...	...	...
i	n	Nn	Pn	Kn	Yn

n treatments with different N, P and K fertilizer application rates and observed yields

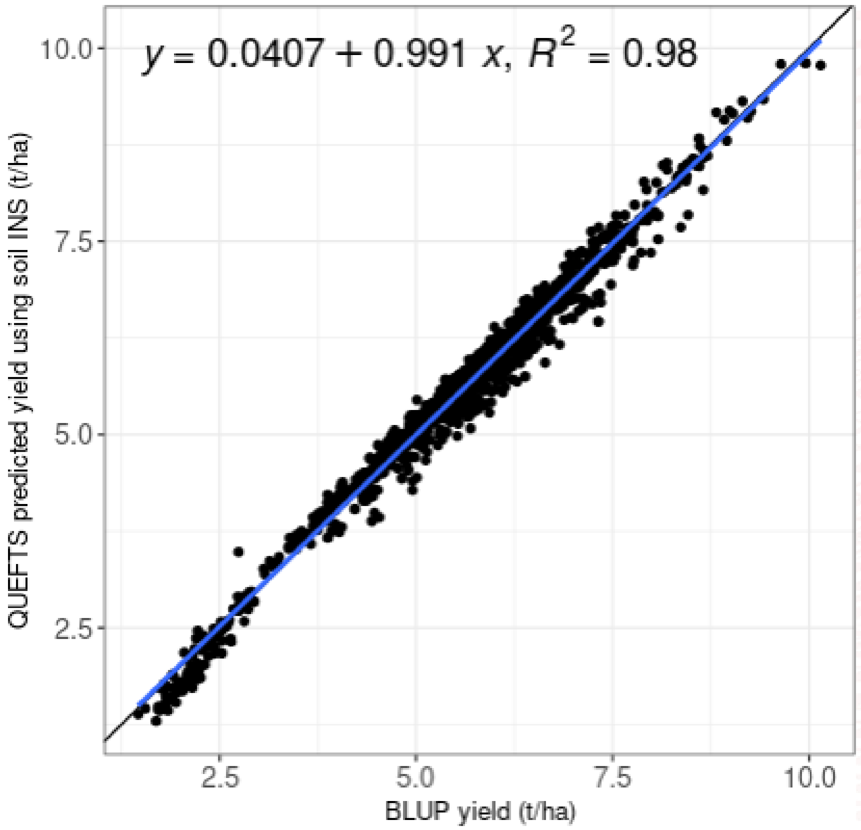
`revQUEFTS` →

TLID	INS	IPS	IKS	aY
i	INSi	IPSi	IKSi	aYi

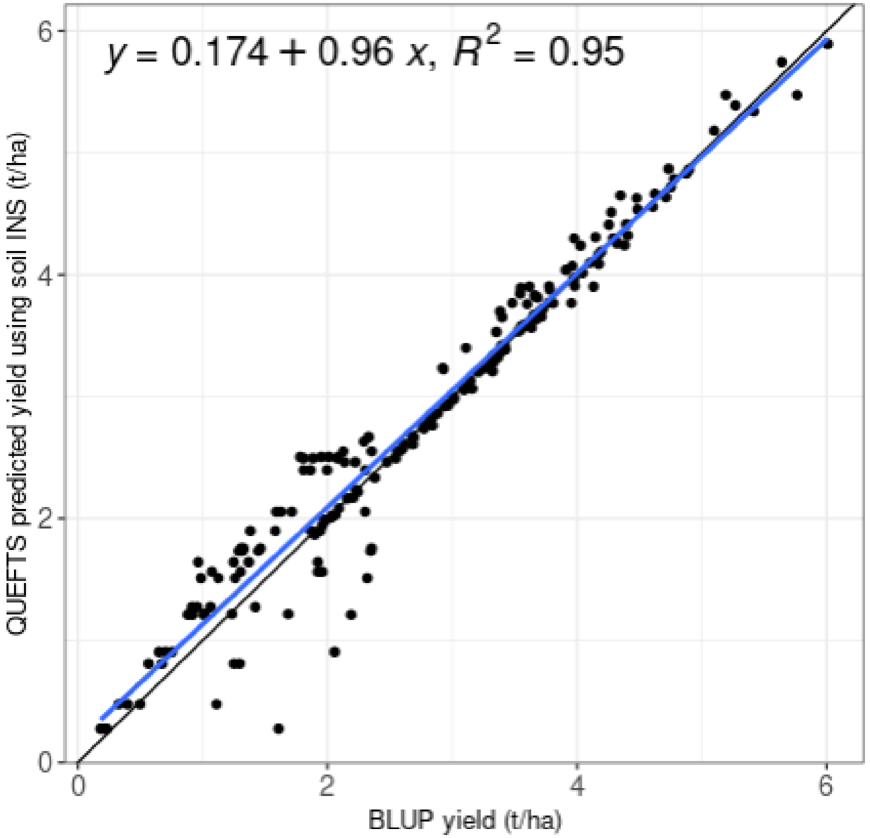
1 set of apparent soil indigenous N, P and K supply values and an attainable yield.
The attainable yield is set to 20% above the blup yield in the reference treatment.

How well does reverse QUEFTS describe the data?

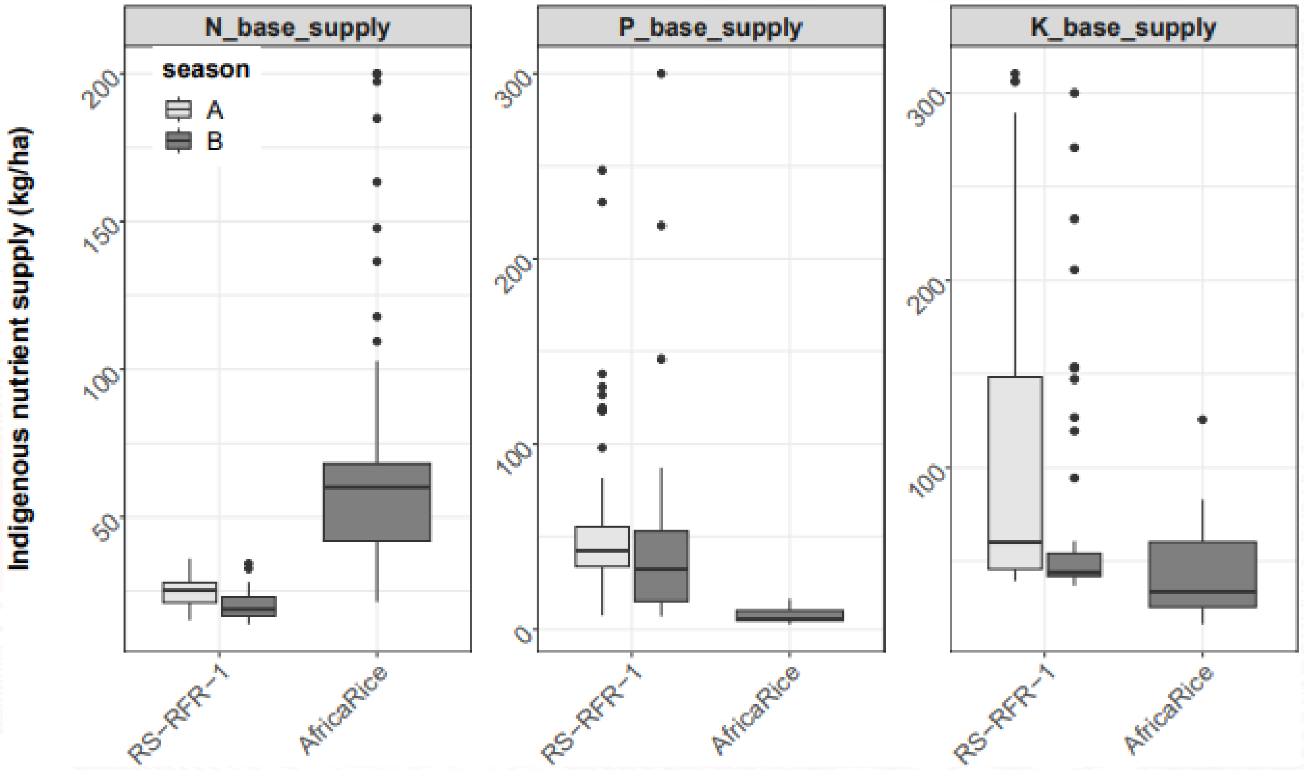
Using observed attainable yield, for RwaSIS data



Using observed attainable yield, Africa Rice data



Ranges in apparent indigenous nutrient supply



- All N , P and K supplies are different between RwaSIS and Africa Rice data

- Soil N:
  - In RwaSIS data, N supply is similar for the two seasons, and is rather low with very little variation.

  - The Africa Rice data predicts twice N supply compared to RwaSIS data (similar rate as we get for potato)
- Soil P
  - RwaSIS data predicts higher P supply than Africa Rice data
  - There is a bit more variation in P supply in RwaSIS data
- Soil K
  - Large difference between seasons for RwaSIS data.
- Response to P in RwaSIS
  - There is hardly any response to the different P levels
- Response to K in RwaSIS
  - There is a it more response to K if compared with response to P

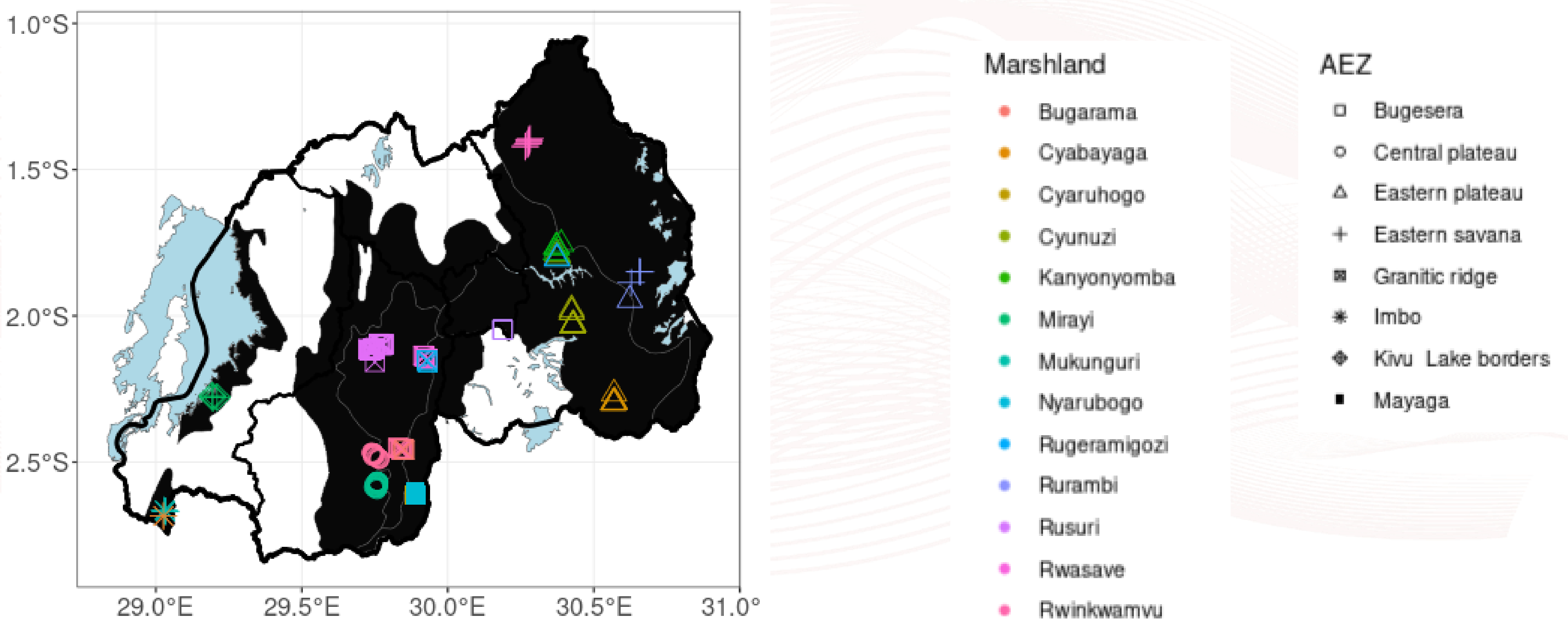
CGIAR INITIATIVE ON Excellence in Agronomy

STEP 4: Predicting yield and yield response

Aggregating the indigenous nutrient supply and the reference yield

Marshlands within AEZ:

Given the aggregated INS is going to be used to produce fertilizer advice and not the INS estimated using random forest models which would use different reference yield, in case of rice, the reference yield class is not implemented.



INITIATIVE ON
Excellence in
Agronomy

Aggregating the indigenous nutrient supply

What is the effect of Africa Rice data?

The soil NPK estimates for Africa Rice data is given in bracket
The two data sets are resulting in different soil INS, priority is given to get fertilizer advice using only RwaSIS data.

Marshland	soil_N	soil_P	soil_K
Bugarama	27	66	159
Bugarama	27	66	159
Cyabayaga	21	41	123
Cyaruhogo	24	42	74
Cyunuzi	21	33	78
Kanyonyomba	21	28	60
Mirayi	23	47	75
Mukunguri	27	94	120
Nyarubogo	26	132	132
Rugeramigozi	19 (73)	31 (7)	60 (44)
Rurambi	28	83	181
Rusuri	21 (72)	41 (6)	59 (47)
Rwasave	20 (69)	39 (7)	65 (39)
Rwinkwamvu	30	65	86
Rwinkwamvu	30	65	86

Using predicted indigenous nutrient supply to calculate yield using QUEFTS

How does NPK supply affect yield and yield response in the calculated ranges?

1. The soil INS obtained from QUEFTS is aggregated at level of Sector, marshland, AEZ and district.
2. The reference yield and yield with blanket recommendation (200 kg NPK 17*3 & 100 kg urea per ha) are also aggregated at the of Sector, marshland, AEZ and district
3. Setting target yield at blanket recommendation yield * (0%, 10%, 20%, 30%, 40%, 50%) yield increase, the lowest rates of fertilizer required was calculated with aggregation of soil INS, reference yield and target yield being aggregated at
    - Marshland
    - Sector
    - AEZ
    - District

4. Given the current price (see below) the cost of the advice is calculated

Fertilizer	N	P	K	price (RWF/kg)
DAP	18	20	0	722
Urea	46	0	0	640
NPK	17	7.4	14	654

Fertilizer requirements for two selected scenarios

Fertilizer requirements to achieve the same yield as compared to the yield obtained with the current blanket recommendation, aggregated at AEZ

By aggregating the INS and reference yield at AEZ,

Simplify to bags per hectare (1 bag = 50 kg)

AEZ	Blanket recommendation yield (t/ha)	Same yield as current recommendation (kg ha^{-1})			Same yield as current recommendation (bags)		
		DAP	NPK 17:17:17	Urea	DAP	NPK 17:17:17	Urea
Bugesera	6.4	29	64	115	0.5	1.5	2.5
Central plateau	5.2	33	70	127	0.5	1.5	2.5
Eastern plateau	5.3	32	69	124	0.5	1.5	2.5
Eastern savana	5.8	29	63	114	0.5	1.5	2.5
Granitic ridge	5.8	32	69	124	0.5	1.5	2.5
Imbo	6.7	33	71	128	0.5	1.5	2.5
Kivu Lake borders	6.0	32	70	125	0.5	1.5	2.5
Mayaga	5.9	33	72	129	0.5	1.5	2.5

The same yield level as compared to the yield obtained with the current blanket recommendation can be obtained with 15% less fertilizer

- The six bags of blanket recommendations (4 bags of NPK 17:17:17 and 2 bags of Urea) are replaced by 5 bags of fertilizers (1 bag DAP, 1.5 bags NPK 17:17:17 and 2.5 bags of urea)
- The result suggests to apply almost the same amount of N and P while reducing K
- This results should increase the agronomic efficiency and increase return on investment by reducing fertilizer investment by 16%, given current prices

Fertilizer requirements for two selected scenarios

Fertilizer requirements to achieve a 20% yield increase as compared to the yield obtained with the current blanket recommendation, aggregated at AEZ

By aggregating the INS and reference yield at AEZ,

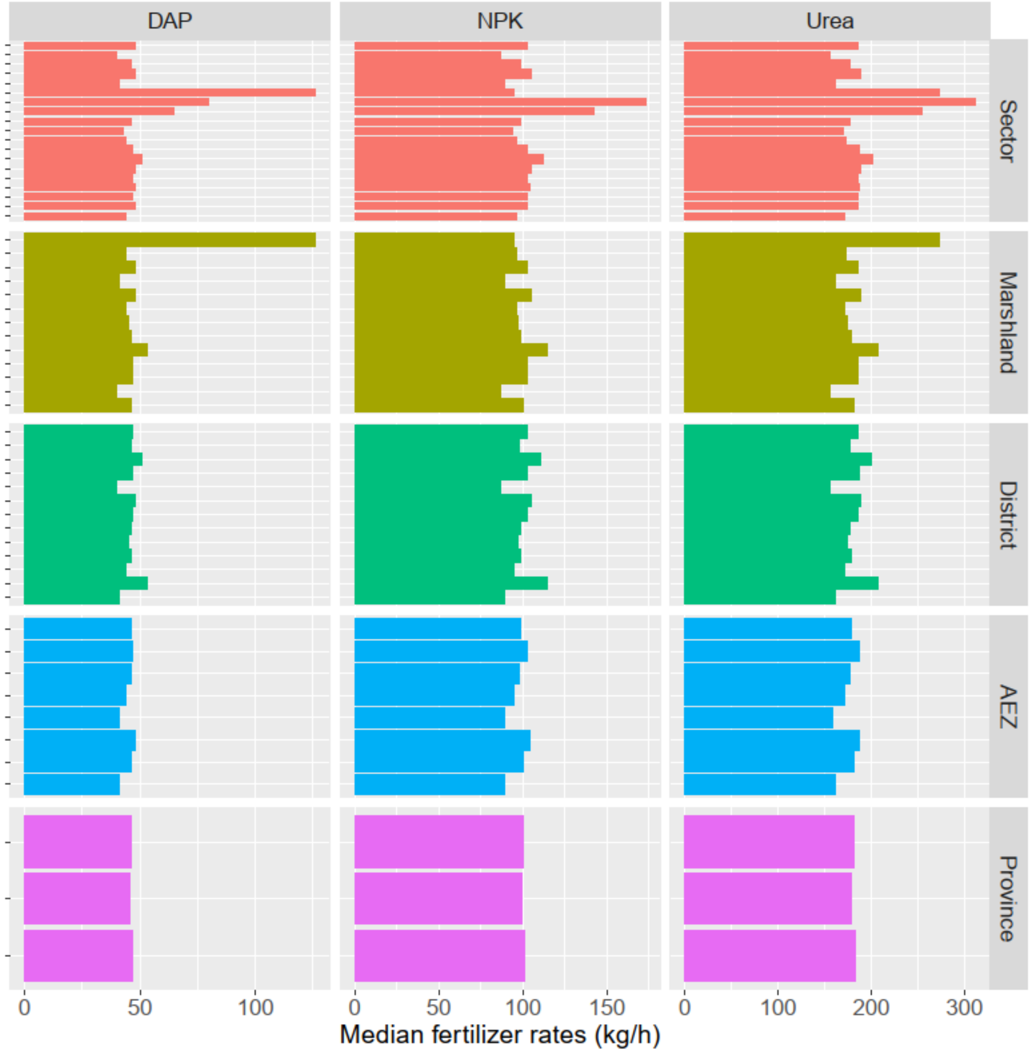
Simplify to bags per hectare (1 bag = 50 kg)

AEZ	Blanket recommendation yield (t/ha)	Same yield as current recommendation (kg ha^{-1})		
		DAP	NPK 17:17:17	Urea
Bugesera	6.4	41	89	161
Central plateau	5.2	46	100	181
Eastern plateau	5.3	48	104	187
Eastern savana	5.8	41	89	159
Granitic ridge	5.8	44	95	171
Imbo	6.7	46	98	177
Kivu Lake borders	6.0	47	103	187
Mayaga	5.9	46	99	197

Same yield as current recommendation (bags)		
DAP	NPK 17:17:17	Urea
1	2	3
1	2	4
1	2	4
1	2	3
1	2	4
1	2	4
1	2	4
1	2	4

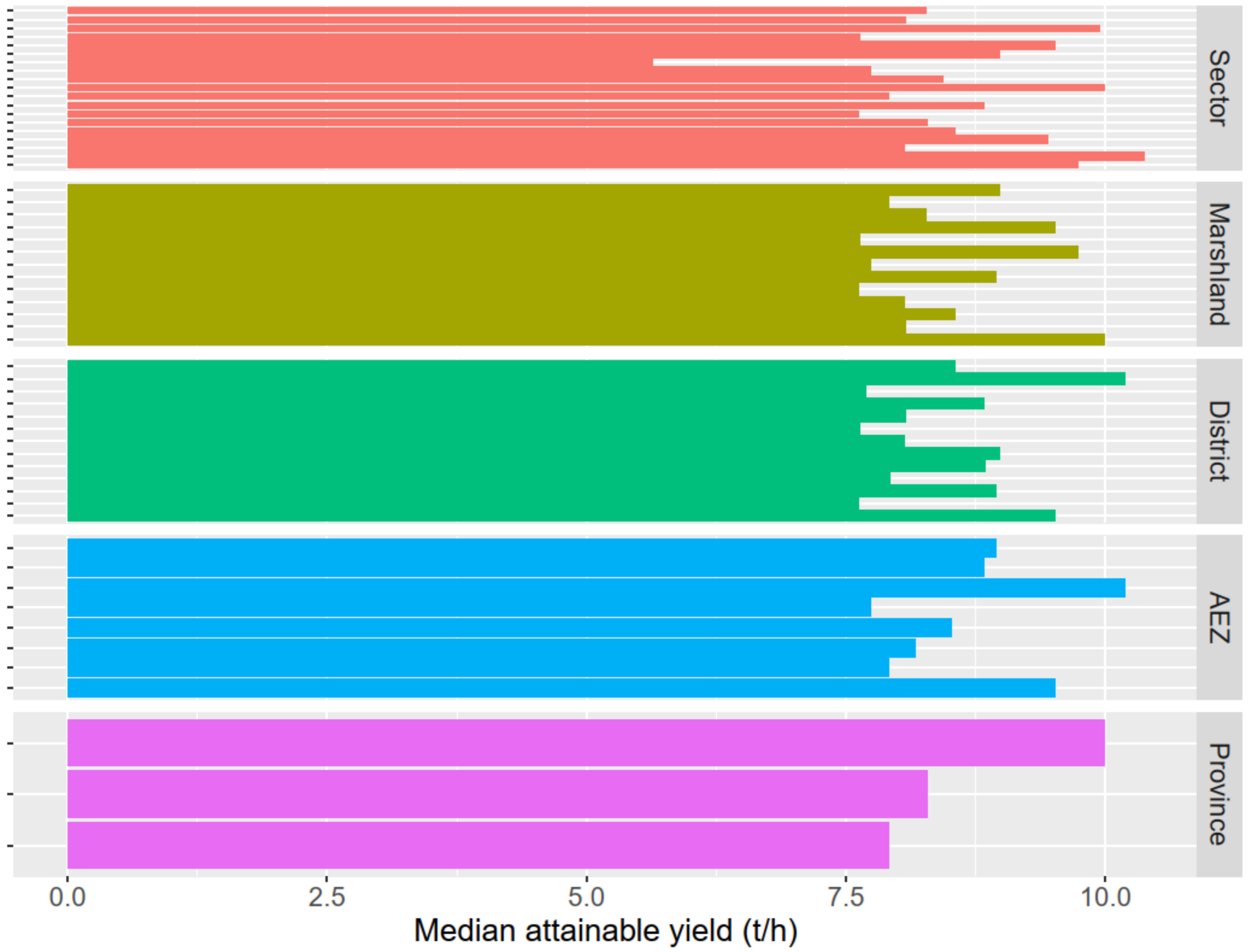
- 20% increase above the yield with the current blanket recommendation can be achieved by doubling the amount of urea applied and halving the NPK17:17:17 from 4 to 2 bags by supplementing it with 1 bag of DAP
  - The result suggests to increase the amount of N applied by about 48 % and P by 17% while halving K indicating limitation by N followed by P as a major yield limiting nutrients
  - Given the current fertilizer prices in the target area, around a 17% increase in fertilizer investment is expected to achieve a 20% yield increase.

The median of fertilizer rates at different aggregations level



Except for very few sectors, the median fertilizer rates does not vary sufficiently between marshlands, district, AEZ or provinces.
This is expected as the soil INS does not show much variation either.

The median of attainable yield at different aggregations level



The attainable yield variation is within the range of 1 t/ha.

Using predicted indigenous nutrient supply to calculate yield using QUEFTS

- The variation in fertilizer rates among AEZ and districts is minimal. It means we could probably simply replace the current blanket recommendation with a new (better) one. (However, in the validation trials, we could still test the disaggregated recommendations.)
- The same yield as the blanket recommendation can be achieved with an average 24% cost reduction
- More urea is required and less NPK. Some NPK is substituted by DAP.
- We can achieve a 20% yield increase at a cost that is 0 to 15% (average of 10% increase) higher than that of the blanket recommendation, depending on AEZ/district. This is a realistic scenario in terms of cost. Relatively high urea rates are required, which should be split-applied (159-187 kg/ha, depending on district/AEZ).
- We can achieve a 30% yield increase at a cost that is 12 to 79% higher than that of the blanket recommendation, depending on AEZ/district. This scenario has very high urea rates and high cost in some districts and seems less desirable.
- As variation between marshlands and AEZs is very minimal, we can choose only one level of aggregation instead of testing both in the validation trials. I propose we go for a two-plot validation trial (blanket versus new recommendation) and aggregate recommendation at the level of AEZ.
- Minor comment: There is something odd with the Rwinkamvu marshland in Eastern plateau district. It has very high fertilizer rates while its INS, INK and INP data and target yield are not so much different from other marshlands. Do you have any idea what causes this?

FertRecom_rice.xlsx