Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Economic Affairs,
Education and Research EAER

Agroscope

M2 statistical inference analysis on winter wheat variety trial

Nicolas Vuille-dit-Bille

Agroscope: swiss federal institute for agricultural research

Group:

- Cultivation Techniques and Varieties in Arable Farming

Team:

- 35 People

Domains of research:

- Agronomy
- Crop science

Main areas of work:

- Characterization of genotypes of field crops
- Improvement of management practices of field crops

Crop species:

- Winter cereals (**wheat**, barley, oat, …)
- Maize
- Potato and Sugar beet
- Oil crops and grain legumes
- Minor crops (e.g. Quinoa, Amaranth)

Variety trial context

- Winter wheat varieties are tested on different experimental fields across Switzerland with different modality (low vs high inputs)
- Each year, winter wheat varieties performance are reported and recommended to farmers according to a list
- Challenge: not easy to estimate and understand accurately specific varieties response to specific environment linked to abiotic and biotic stresses (uncontrolled factors on the field)

Méthodes de phénotypage dans le cadre de l'examen des variétés

Blé d'automne — Variétés

Schweizerische Eidgenossenschaft · Confédération suisse · Confederazione Svizzera · Confederaziun svizra
Département fédéral de l'économie, de la formation et de la recherche DEFR
Agroscope

swiss granum

Liste recommandée des variétés de blé d'automne pour la récolte 2023

		TOP										I						II				Fourrager				Biscuit
Variété		RUNAL	TITLIS	CH CLARO	CH NARA	PIZNAIR	AXEN[5]	DIAVEL[6]	MONTALBANO[L]	BONAVAU[5]	BARETTA	CADLIMO	ARINA	FOREL	SIMANO[4]	ALPVAL	HANSWIN	CAMPANILE	POSMEDA	LUDWIG	LEVIS	SPONTAN	PONCIONE	CAMPESINO	SAILOR	MULAN	DILAGO
Année d'inscription		1995	1996	2009 (dern. année)	2010	2020	2022 (prov.)	2020	2018	2022 (prov.)	2018	2020	1981	2008	2012 (dern. année)	2022 (prov.)	2015	2021	2019	2004	1997	2017	2019	2022 (prov.)	2015	2007 (dern. année)	2019
Rendement (Extenso)[1]		-	-	-	+	+	+++	+	++	++	+(+)	++	Ø	Ø	+	++	++	+++	+++	+++	+	++++	++++	++++	++++	+++(+)	++++
Rendement (PER)[2]		--	--	+	-(-)	Ø	+(+)	+	+	Ø	+	+	-	+	Ø	+(+)	+(+)	+(+)	++(+)	++	+(+)	+++(+)	++++	++++	++++	+++	+++(+)
Précocité à l'épiaison[1]		mt	t	mp	mt	mt	tp	tp	t	mt	t	t	t	mp	tp	t	mp	mt	mp	mt	mt	mt	t	mt	t	t	t
Hauteur des plantes[1]		m	ml	mc	to	m	l	tl	m	c	ml	m	tl	m	c	m	m	m	tl	tl	c	m	l	c	ml	m	ml
Verse[1,2]		++(+)	++	+++	+++	+	++	++(+)	+++	+++	++	+	Ø	++	+++	+++	++(+)	++	+	+	+++	+++	+	+++	++(+)	++(+)	++
Résistance	Oïdium[1]	+(+)	Ø	+	++(+)	+(+)	Ø	++(+)	++	+	+(+)	+(+)	Ø	+	+	+(+)	Ø	+	+	+	Ø	+(+)	+(+)	++	++	+	Ø
	Rouille jaune[1]	+	++	-	+++	++	Ø	++(+)	+++	+	+++	++	Ø	Ø	+	++	+	++	Ø	+	+	Ø	++	+++	+	-	+
	Rouille brune[1]	-	-	Ø	Ø	+	++	+++	++	Ø	+(+)	+(+)	--(-)	--(-)	+	Ø	Ø	++	Ø	-	-	Ø	Ø	++	Ø	-(-)	Ø
	Septoria nodorum feuilles[1,3]	-	Ø	--	+	-	Ø	Ø	Ø	Ø	Ø	Ø	-	-	Ø	+	Ø	Ø	++	Ø	Ø	++	++	++	Ø	+	++
	Septoria nodorum épi[1,3]	Ø	Ø	-	Ø	+	++	Ø	++	+	++	+	++	Ø	-	+	+	Ø	++	++	-	++	++	++	++	+	++
	Septoria tritici feuilles[1,3]	-	Ø	--	+	Ø	+	++	+	++	+	++	Ø	--	-	++	+	++	++	+	Ø	++	++	Ø	+	-	Ø
	Fusariose épi[1,3]	Ø	+	-	--	Ø	Ø	Ø	++	Ø	+	Ø	++	-	Ø	Ø	-	Ø	-	Ø	-	++	-	+	Ø	Ø	+
	Germination sur pied[1]	+(+)	+	-	+	+	Ø	+	Ø	-	Ø	Ø	Ø	+	+	+	+	+	+	Ø	Ø	Ø	+	+	Ø	--	Ø
Teneur en protéines[1]		+++	++(+)	+++	+++	+++	+++	++(+)	+++	+++	++(+)	++(+)	++(+)	+(+)	++(+)	++(+)	+(+)	++	Ø	Ø	++	Ø	Ø	-	-(-)	-(-)	Qualité biscuit
Zélény[7]		+(+)	++	++	+++	++(+)	++	++(+)	++	++(+)	++	++(+)	Ø	++	++	++	+	++	Ø	Ø	++	-	Qualité fourragère				
Poids à l'hectolitre[1]		++	++	Ø	+++	+(+)	++	++	+	++	Ø	+(+)	+++	+++	+(+)	++	+++	+++	+(+)	+(+)	++	+(+)	+(+)	+	++(+)	-	+++
Poids de mille grains[1]		moyen	grand	petit	petit	petit	moyen	petit	grand	moyen	petit	très petit	moyen	très petit	moyen	grand	grand	moyen	très grand	très grand	moyen	moyen	très grand	petit	très grand	moyen	grand

[1,2]Cette description variétale a été établie sur la base des résultats moyens de deux (trois) ans obtenus dans le réseau d'essais d'Agroscope et DSP[1], complétés par ceux du réseau d'essais culturaux swiss granum pour le rendement en mode PER[2] et quelques autres appréciations. Ces informations peuvent toutefois varier en fonction du lieu et des conditions climatiques de l'année.

L'appréciation du rendement s'entend entre variétés au sein d'un même type de production (lecture horizontale!). La description ci-dessus ne permet donc pas une comparaison entre types de production (Extenso ou PER)

[3]Échelle d'appréciation à cinq niveaux (++, +, Ø, - et --).

[4]Variété à épis barbus.

[L]Variété pouvant être semée au printemps, la qualité étant légèrement supérieure.

Toutes les variétés qui sont actuellement inscrites ou l'ont été dans le passé sur la liste recommandée de swiss granum peuvent être commercialisée sous Suisse Garantie.

Légende:

+++	= très bon
++	= bon
+	= moyen à bon
Ø	= moyen
-	= faible à moyen
--	= faible
---	= très faible
case vide	= aucune information

Précocité:

tp	= très précoce
p	= précoce
mp	= mi-précoce
mt	= mi-tardive
t	= tardive

Hauteur des plantes:

to	= très courte
c	= courte
mc	= moyenne à courte
m	= moyenne
ml	= moyenne à longue
l	= longue
tl	= très longue

www.agridea.ch | www.swissgranum.ch | www.agroscope.ch

agridea

Juin 2022

Agroscope

Case study: variety trial with 2 treatments over 4 years

- 36 winter wheat varieties

- 4 years: 2021 to 2024

- Case study context: study variety response to none vs conventional application

➢ Find variety that are performing well also under limited available N in order to reduce N fertilization (economic and environmental cost)

- 2 treatments: conventional application of nitrogen (N) and no application of nitrogen
- Measurement to estimate varieties performance:

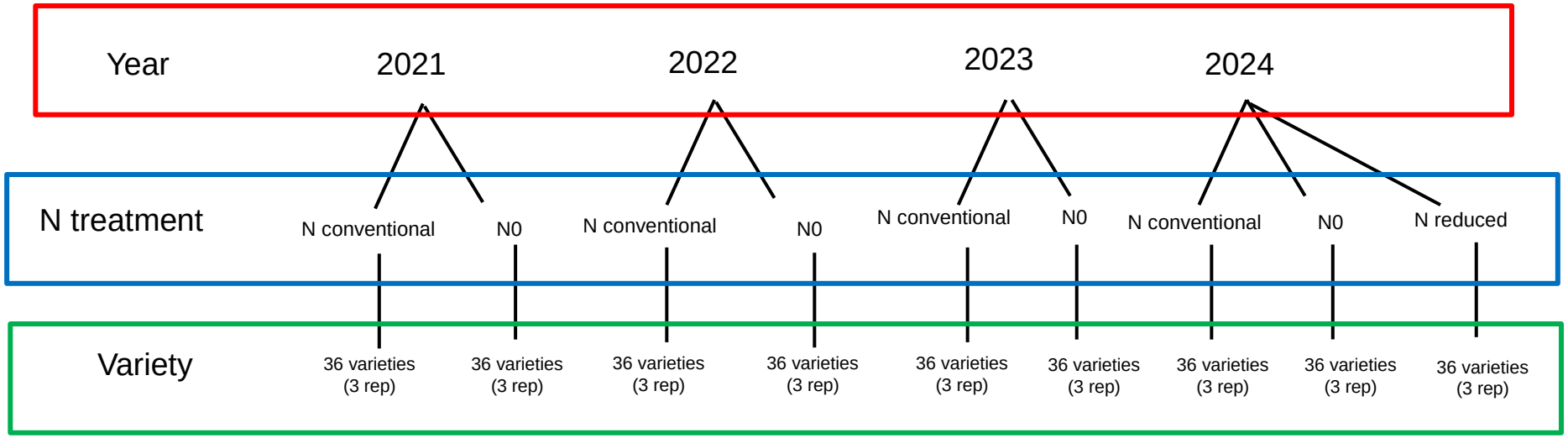
➢ Grain yield
➢ Straw yield
➢ Grain protein
➢ Other physiological parameters (harvest index, leaf area index, chlorophyll content, canopy cover…)

Agroscope

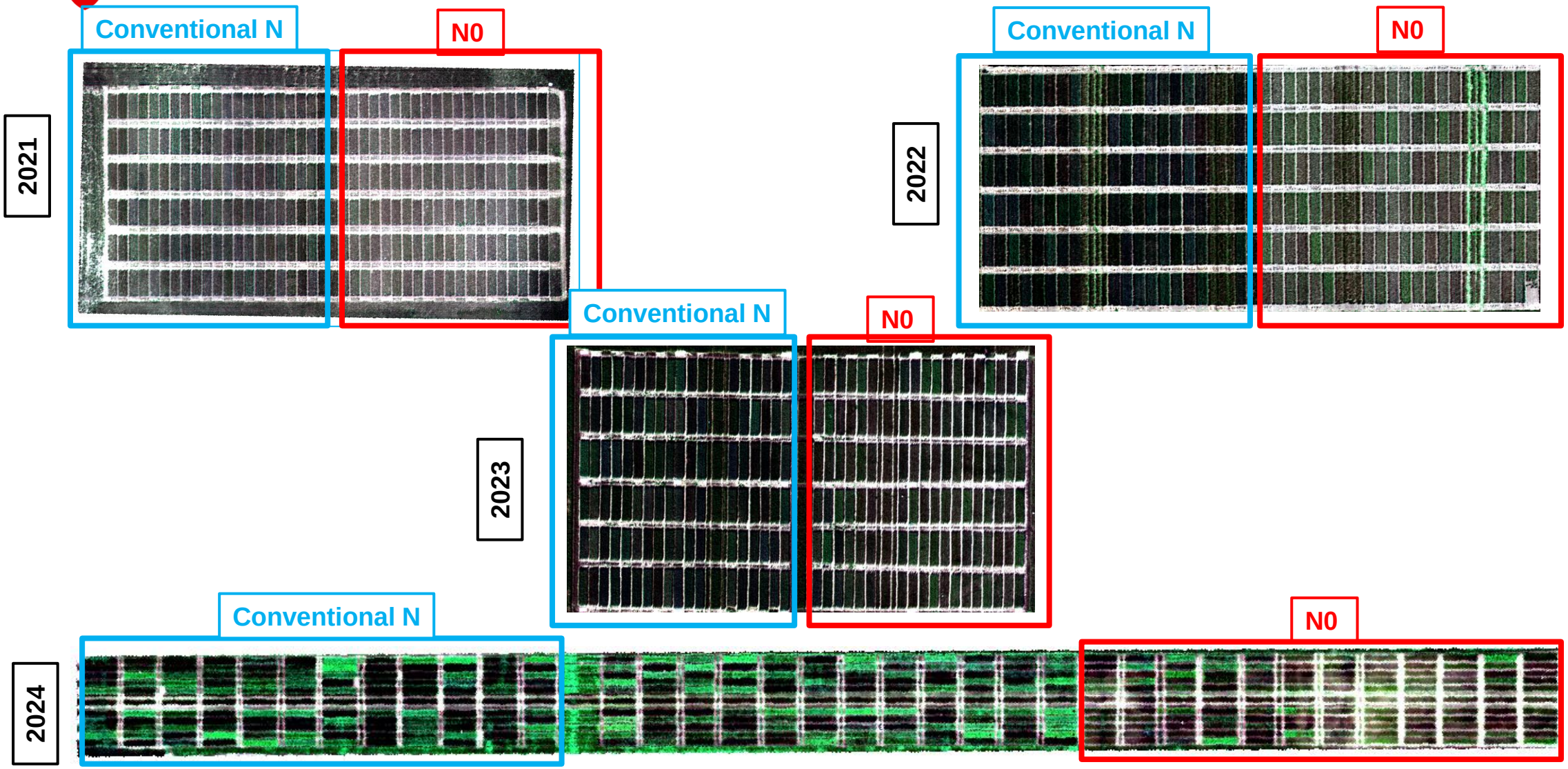
Case study: field experimental design

Case study: experimental design

- 4 contrasted climatic conditions over 4 years
- 2 blocks corresponding to 2 treatments (conventional and no N treatment) within field
- 36 wheat varieties (3 rep by block)



Case study: field trials over 4 years



Data preparation

- The varieties tested are not always the same over years

➢ Find common varieties tested over 4 years

- The measurements to evaluate the variety performance are not always the same over years

➢ Find common variables measured over 4 years

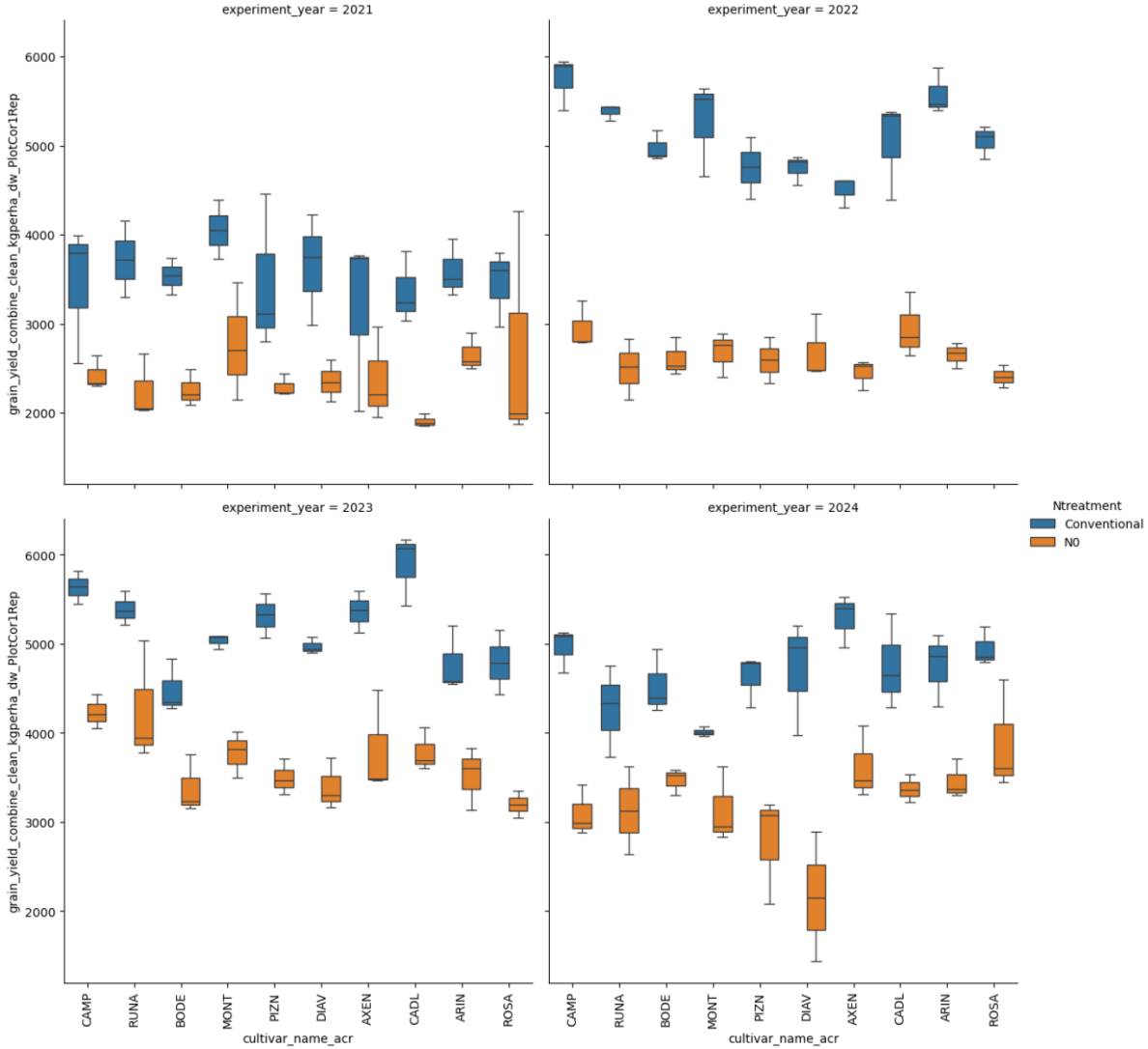
- Due to experimental issue some values are lost and one replicate less can imbalance the statistical analysis (especially with rep=3)

➢ Remove varieties with not sufficient replicates

- Datasets are stored by years on excel sheet

➢ Merge datasets for over years analysis and visualisation

Estimate statistical differences between varieties at N treatments and years level: grain yield



- Conventional N treatment: higher grain yield
  - ➢ high effect in 2022 (dry year)
- Varieties: grain yield performance depending on years an N treatment blocks

Model parameters

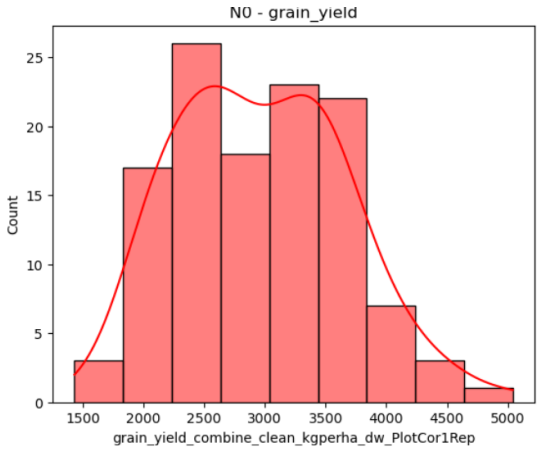
- Response variable: grain yield

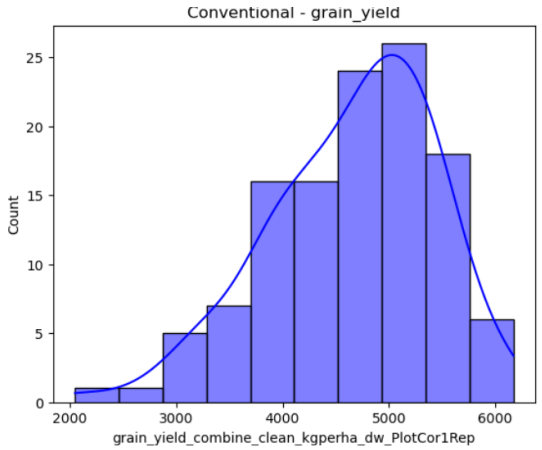
- 3 factors as explanatory variables:
  - Varieties : 10 levels
  - N treatment: 2 levels (N0, Conventional)
  - Years: 4 levels (2021-2024)

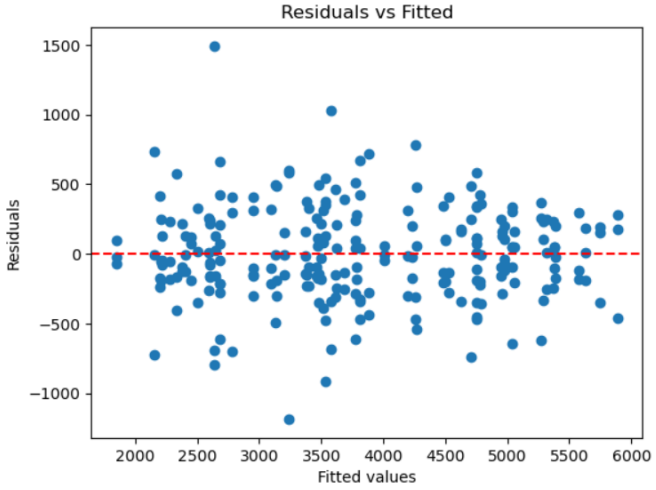
- H0: there is no difference of grain yield between varieties.

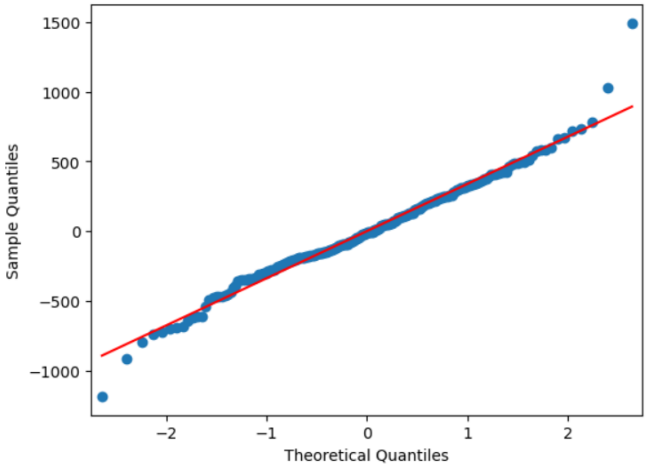
Statistical test option (> 2 groups): ANOVA

- Model assumptions:
  - ➢ Independence of observations ✖
  - ➢ Normality ✖
  - ➢ Homogeneity of variances (homoscedasticity)







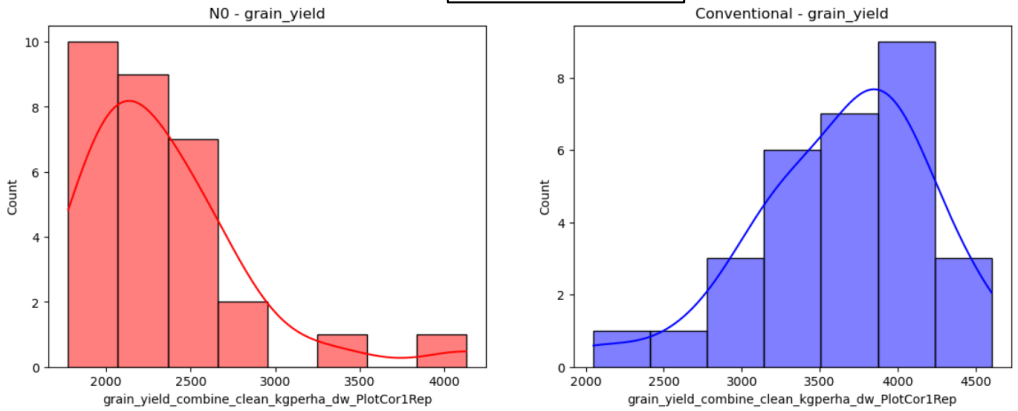


Shapiro-Wilk test p-value: 0.0067

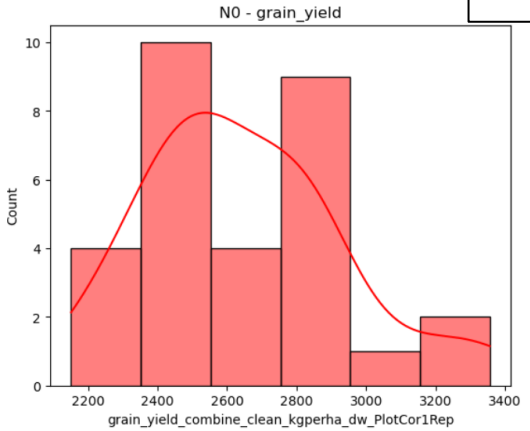
Agroscope

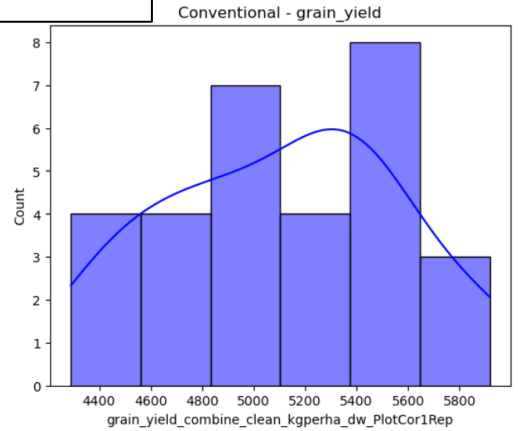
Data distribution by year and N treatment block

2021

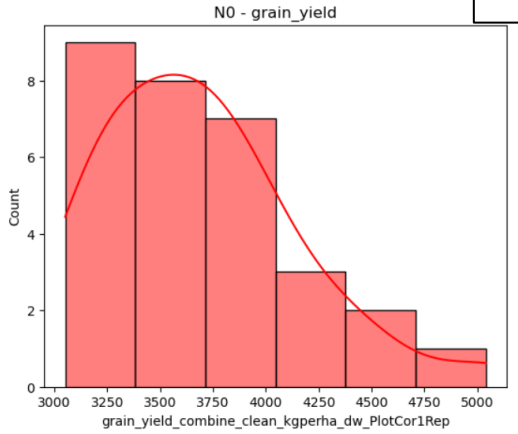


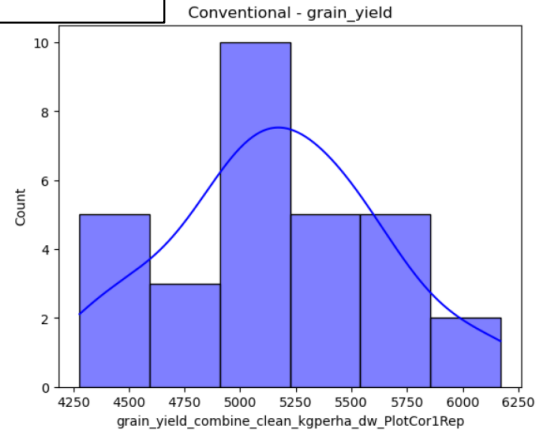
2022



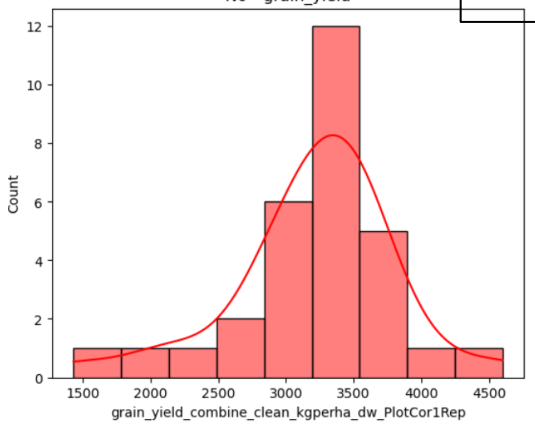


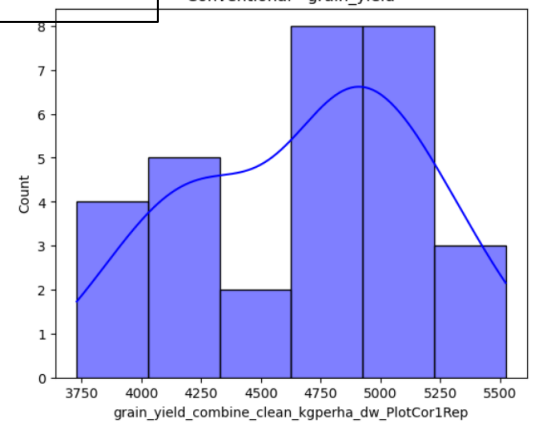
2023





2024





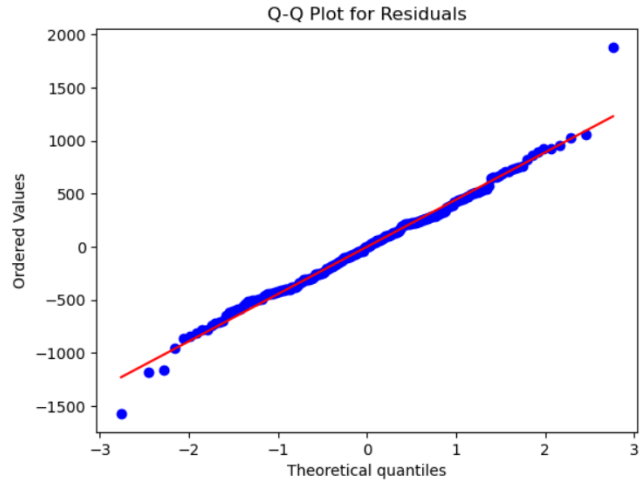
Linear mixed model

- Handles lack of independence of observations
  - ➢ Fixed effects: **varieties**
  - ➢ Random effects: **N treatments** nested in **years factor**

(account for correlation between observations within the same treatment block)

- Python formula:
  - ➢ Model = smf.mixedlm("grain_yield_combine_clean_kgperha_dw_PlotCor1Rep ~ **cultivar_name_acr**", data=merged_df, groups=merged_df["**experiment_year**"], re_formula="1 + **Ntreatmen**t")

- Normality of residuals ? ✖



Shapiro-Wilk Test: W=0.989, p-value=0.053

Mixed Linear Model Regression Results

Model:	MixedLM	Dependent Variable:	grain_yield_combine_clean_kgperha_dw_PlotCor1Rep
No. Observations:	240	Method:	REML
No. Groups:	4	Scale:	225366.5953
Min. group size:	60	Log-Likelihood:	-1777.0535
Max. group size:	60	Converged:	Yes
Mean group size:	60.0		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	3459.534	1118.269	3.094	0.002	1267.768	5651.301
cultivar_name_acr[T.AXEN]	-51.542	137.042	-0.376	0.707	-320.140	217.056
cultivar_name_acr[T.BODE]	-218.613	137.042	-1.595	0.111	-487.210	49.985
cultivar_name_acr[T.CADL]	10.955	137.042	0.080	0.936	-257.643	279.553
cultivar_name_acr[T.CAMP]	185.972	137.042	1.357	0.175	-82.626	454.570
cultivar_name_acr[T.DIAV]	-277.565	137.042	-2.025	0.043	-546.163	-8.968
cultivar_name_acr[T.MONT]	-39.152	137.042	-0.286	0.775	-307.749	229.446
cultivar_name_acr[T.PIZN]	-209.978	137.042	-1.532	0.125	-478.575	58.620
cultivar_name_acr[T.ROSA]	-71.461	137.042	-0.521	0.602	-340.059	197.137
cultivar_name_acr[T.RUNA]	-12.474	137.042	-0.091	0.927	-281.072	256.123
Group Var	2038305.579	6673.745				
Group x Ntreatment[T.N0] Cov	-2267175.585	5328.917				
Ntreatment[T.N0] Var	3019187.484	4559.434				

Fit linear mixed models for each N treatments

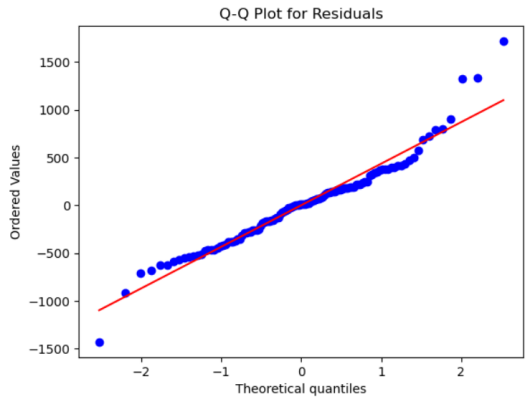
- A lot of variability seems not explained by the fixed effects (year effect).
- At this level, no significant differences between varieties are found
- Use N treatments blocks as separate experiment and analyse variety differences within year (random facror)

N0

Mixed Linear Model Regression Results

Model:	MixedLM	Dependent Variable:	grain_yield_combine_clean_kgperha_dw_PlotCor1Rep
No. Observations:	120	Method:	REML
No. Groups:	4	Scale:	216423.5738
Min. group size:	30	Log-Likelihood:	-850.0383
Max. group size:	30	Converged:	Yes
Mean group size:	30.0		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	3058.094	325.088	9.407	0.000	2420.932	3695.255
cultivar_name_acr[T.AXEN]	-5.775	189.923	-0.030	0.976	-378.016	366.467
cultivar_name_acr[T.BODE]	-142.335	189.923	-0.749	0.454	-514.577	229.906
cultivar_name_acr[T.CADL]	-68.870	189.923	-0.363	0.717	-441.111	303.372
cultivar_name_acr[T.CAMP]	107.142	189.923	0.564	0.573	-265.099	479.384
cultivar_name_acr[T.DIAV]	-428.236	189.923	-2.255	0.024	-800.477	-55.994
cultivar_name_acr[T.MONT]	9.907	189.923	0.052	0.958	-362.335	382.148
cultivar_name_acr[T.PIZN]	-286.609	189.923	-1.509	0.131	-658.850	85.633
cultivar_name_acr[T.ROSA]	-26.005	189.923	-0.137	0.891	-398.246	346.237
cultivar_name_acr[T.RUNA]	-37.297	189.923	-0.196	0.844	-409.539	334.944
Group Var	350588.660	636.544				



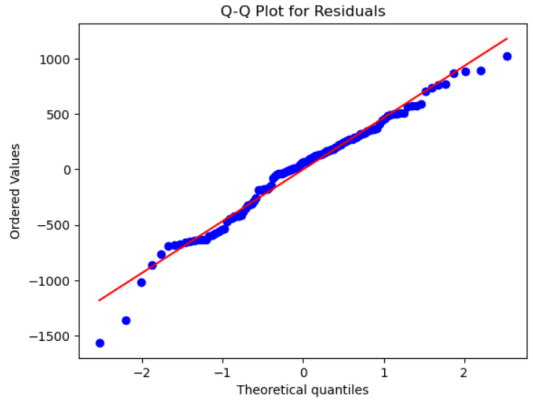
Shapiro-Wilk Test: W=0.956, p-value=0.001

Conventional N

Mixed Linear Model Regression Results

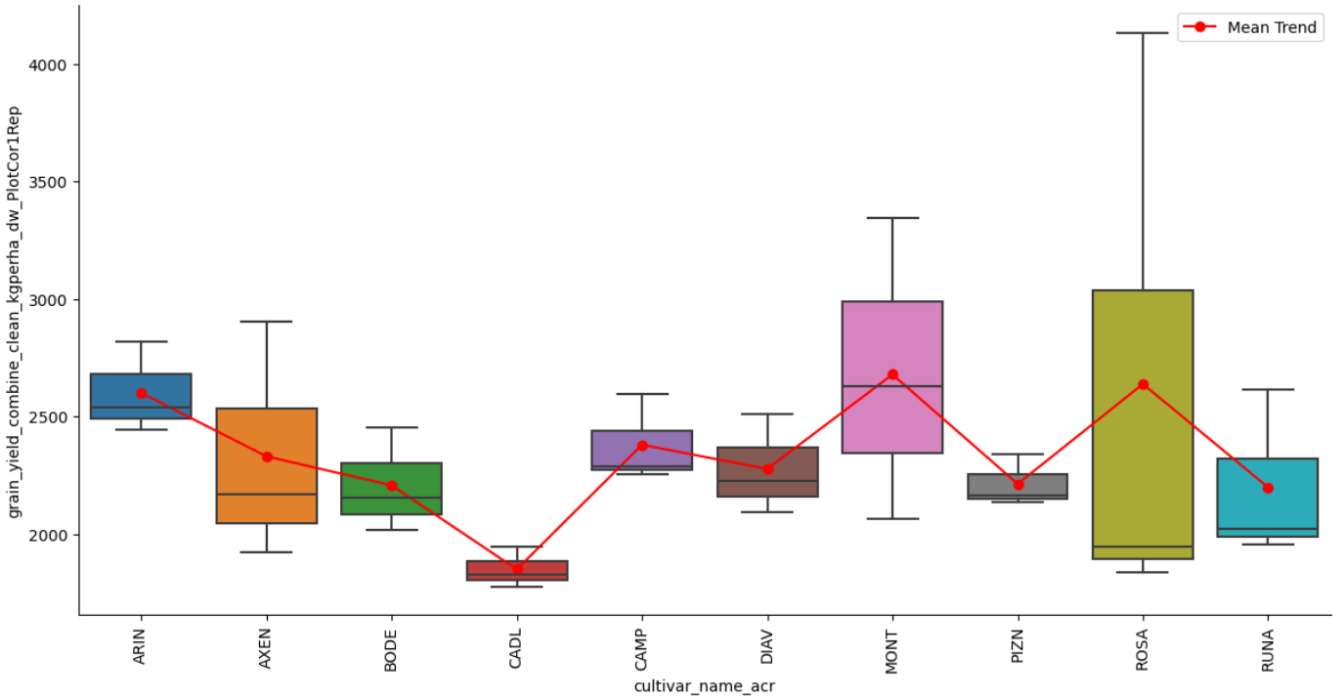
Model:	MixedLM	Dependent Variable:	grain_yield_combine_clean_kgperha_dw_PlotCor1Rep
No. Observations:	120	Method:	REML
No. Groups:	4	Scale:	241966.9754
Min. group size:	30	Log-Likelihood:	-856.4969
Max. group size:	30	Converged:	Yes
Mean group size:	30.0		

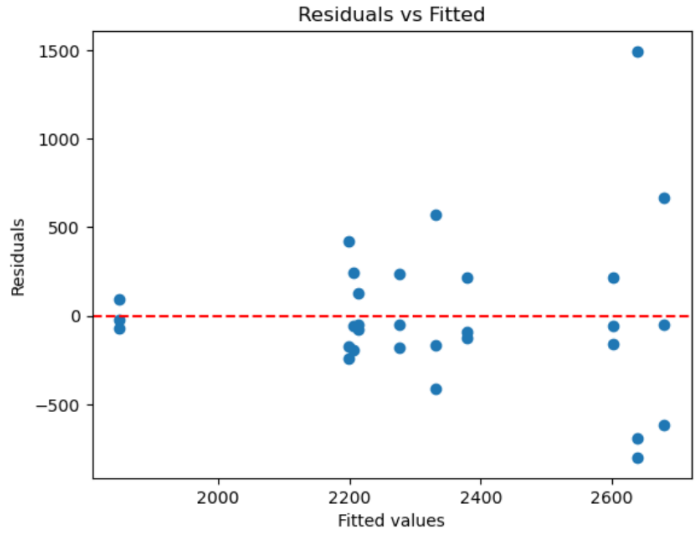
	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	4700.548	377.036	12.467	0.000	3961.571	5439.524
cultivar_name_acr[T.AXEN]	-97.309	200.818	-0.485	0.628	-490.905	296.287
cultivar_name_acr[T.BODE]	-294.890	200.818	-1.468	0.142	-688.486	98.706
cultivar_name_acr[T.CADL]	90.780	200.818	0.452	0.651	-302.816	484.376
cultivar_name_acr[T.CAMP]	264.802	200.818	1.319	0.187	-128.794	658.397
cultivar_name_acr[T.DIAV]	-126.895	200.818	-0.632	0.527	-520.490	266.701
cultivar_name_acr[T.MONT]	-88.210	200.818	-0.439	0.660	-481.806	305.386
cultivar_name_acr[T.PIZN]	-133.347	200.818	-0.664	0.507	-526.943	260.249
cultivar_name_acr[T.ROSA]	-116.917	200.818	-0.582	0.560	-510.513	276.679
cultivar_name_acr[T.RUNA]	12.348	200.818	0.061	0.951	-381.248	405.944
Group Var	487967.949	834.382				

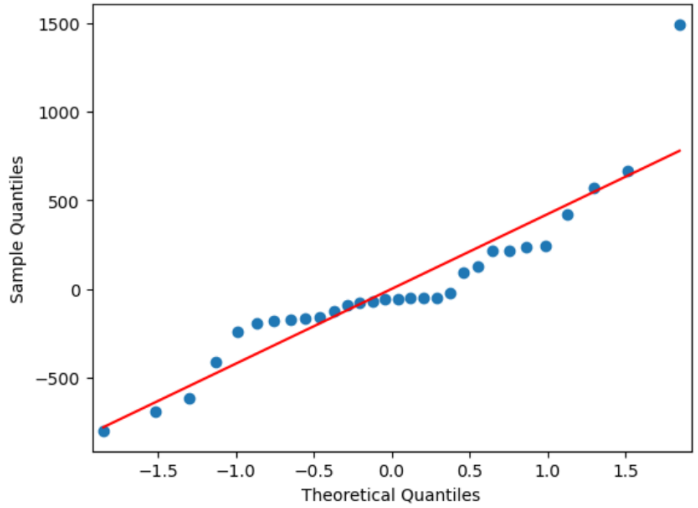


Shapiro-Wilk Test: W=0.979, p-value=0.054

N0 analysis: 2021





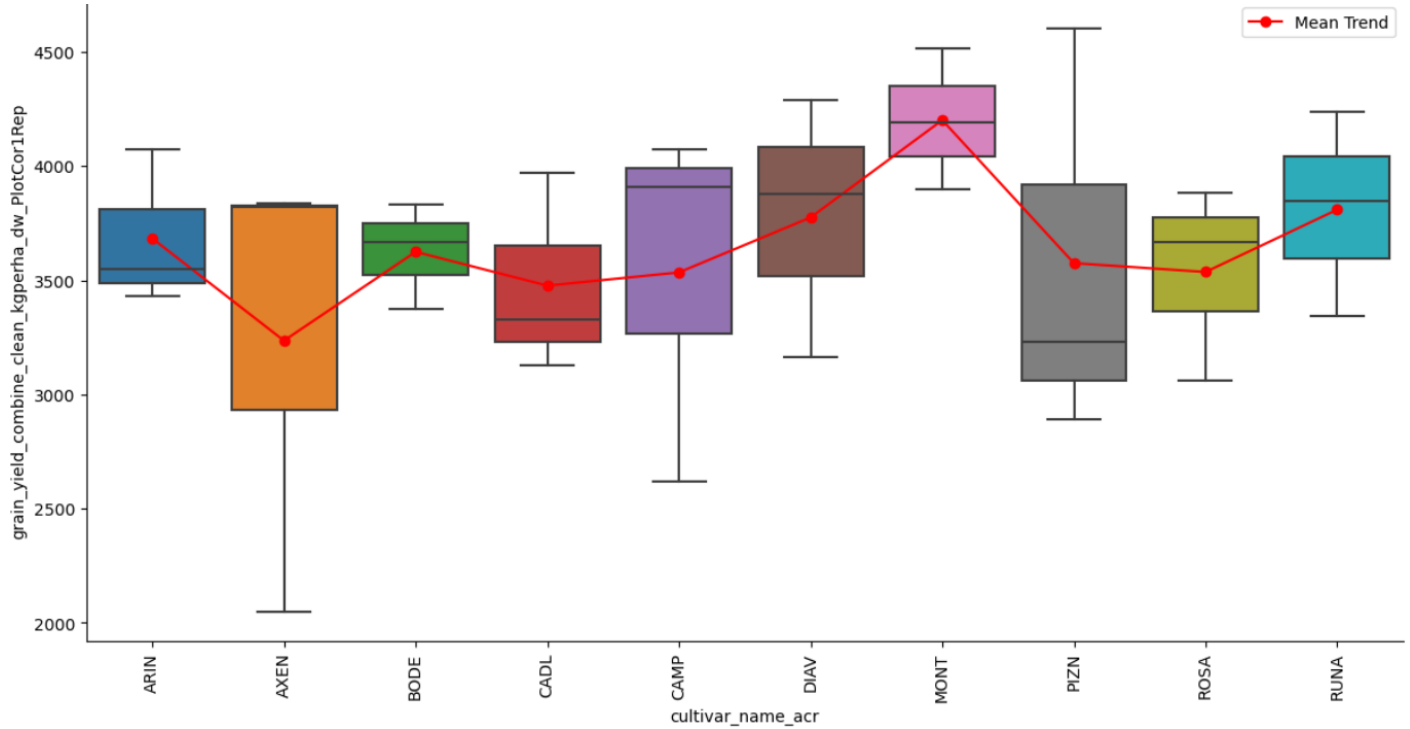



```
Bartlett's Test Statistic: 21.162204888309002
P-value: 0.011949357713574856
The variances are significantly different (homogeneity of variance violated).
```

```
Kruskal-Wallis Test Statistic: 11.39139784946235
P-value: 0.2498313845196752
There is no significant difference between the groups.
```

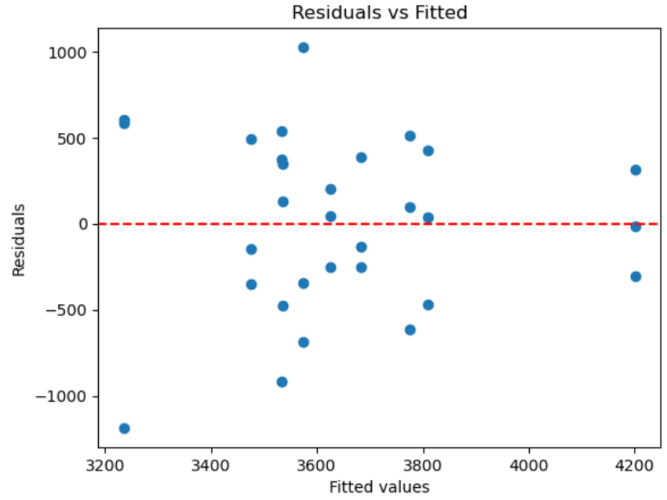
```
Shapiro-Wilk Test: W=0.883, p-value=0.003
```

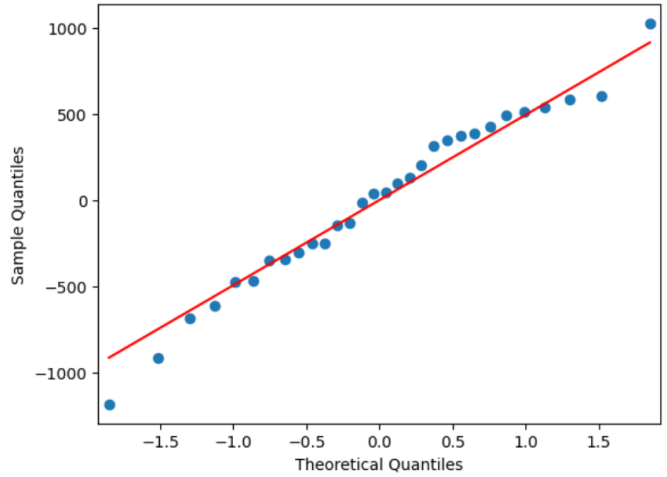
Conventional N analysis: 2021



Bartlett's Test Statistic: 8.809594737807885
P-value: 0.4550328502148302
The variances are not significantly different (homogeneity of variance met).

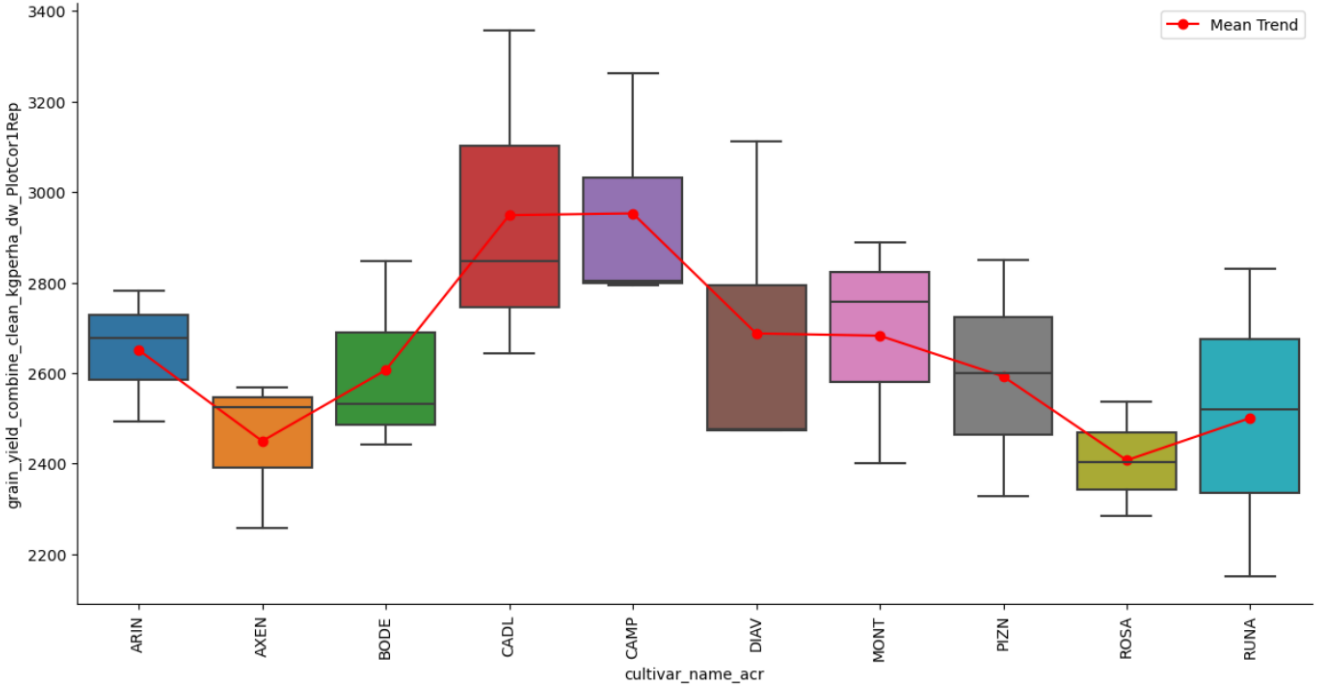
	df	sum_sq	mean_sq	F	PR(>F)
C(cultivar_name_acr)	9.0	1.740199e+06	193355.397277	0.525746	0.838819
Residual	20.0	7.355466e+06	367773.292242	NaN	NaN

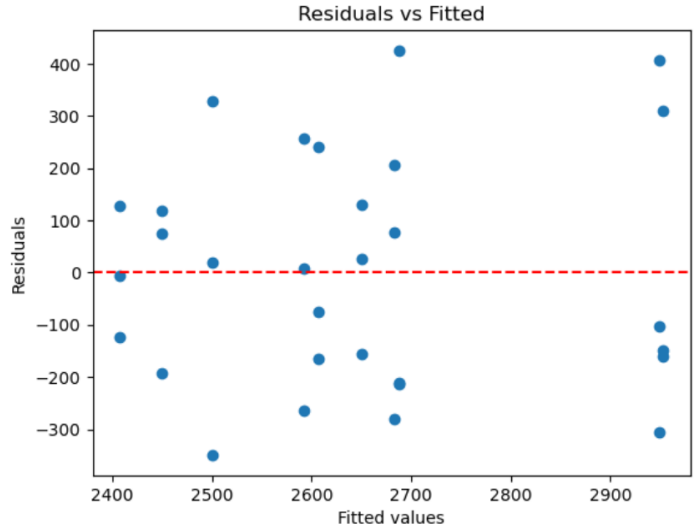


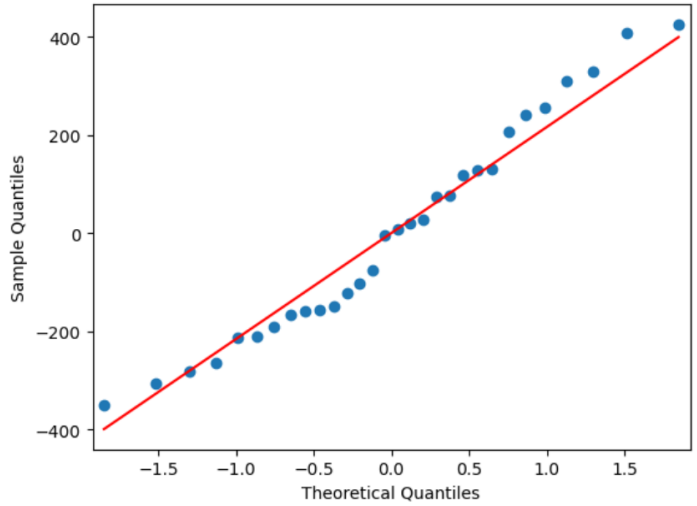


Shapiro-Wilk Test: W=0.980, p-value=0.833

N0 analysis: 2022







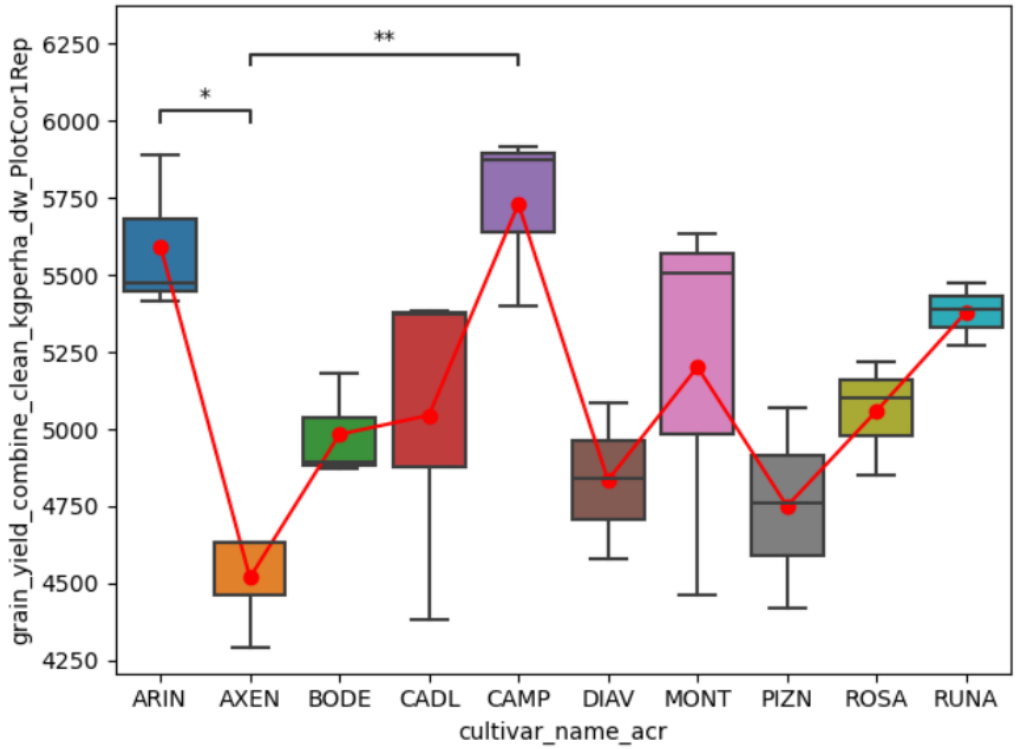
Shapiro-Wilk Test: W=0.958, p-value=0.278

```
Bartlett's Test Statistic: 3.888844197683194
P-value: 0.9185749111061812
The variances are not significantly different (homogeneity of variance met).
```

	df	sum_sq	mean_sq	F	PR(>F)
C(cultivar_name_acr)	9.0	9.290701e+05	103230.005615	1.475667	0.223524
Residual	20.0	1.399096e+06	69954.810694	NaN	NaN

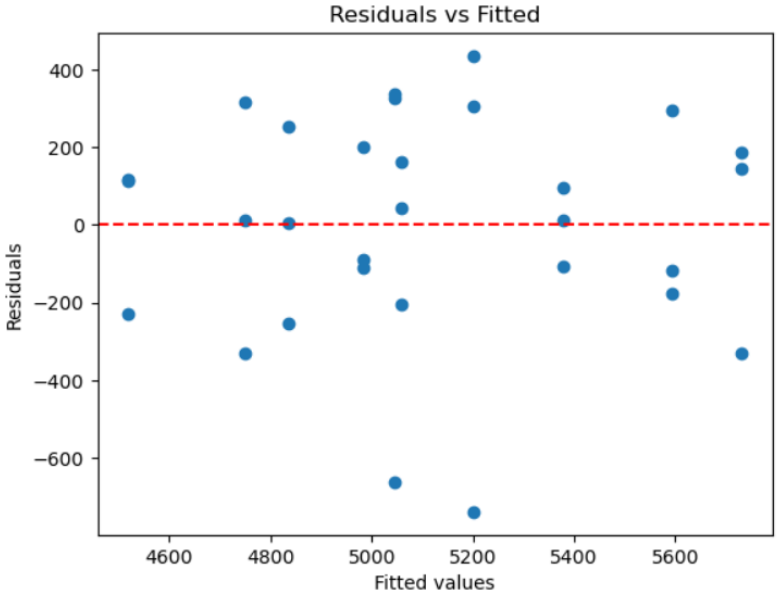
Conventional N analysis: 2022

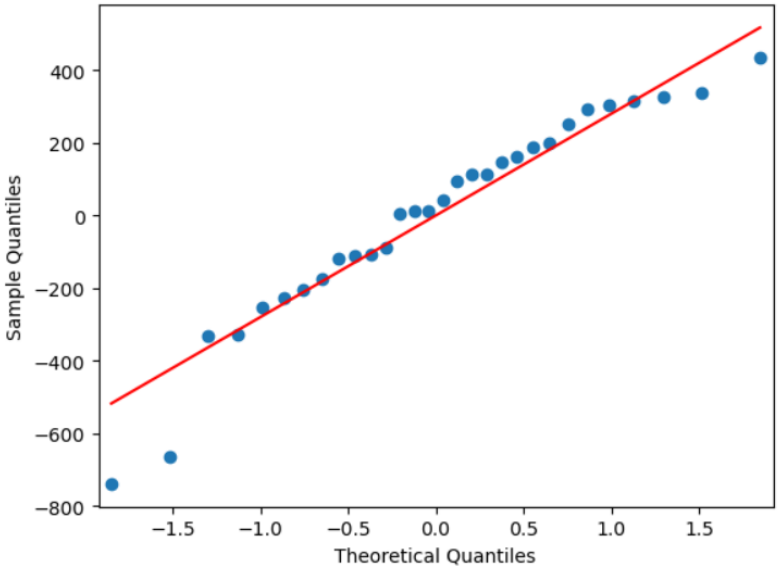
```
ARIN vs. AXEN: Custom statistical test, P_val:2.710e-02
AXEN vs. CAMP: Custom statistical test, P_val:9.517e-03
```




```
Bartlett's Test Statistic: 9.24690216708312
P-value: 0.4148023297396884
The variances are not significantly different (homogeneity of variance met).
```

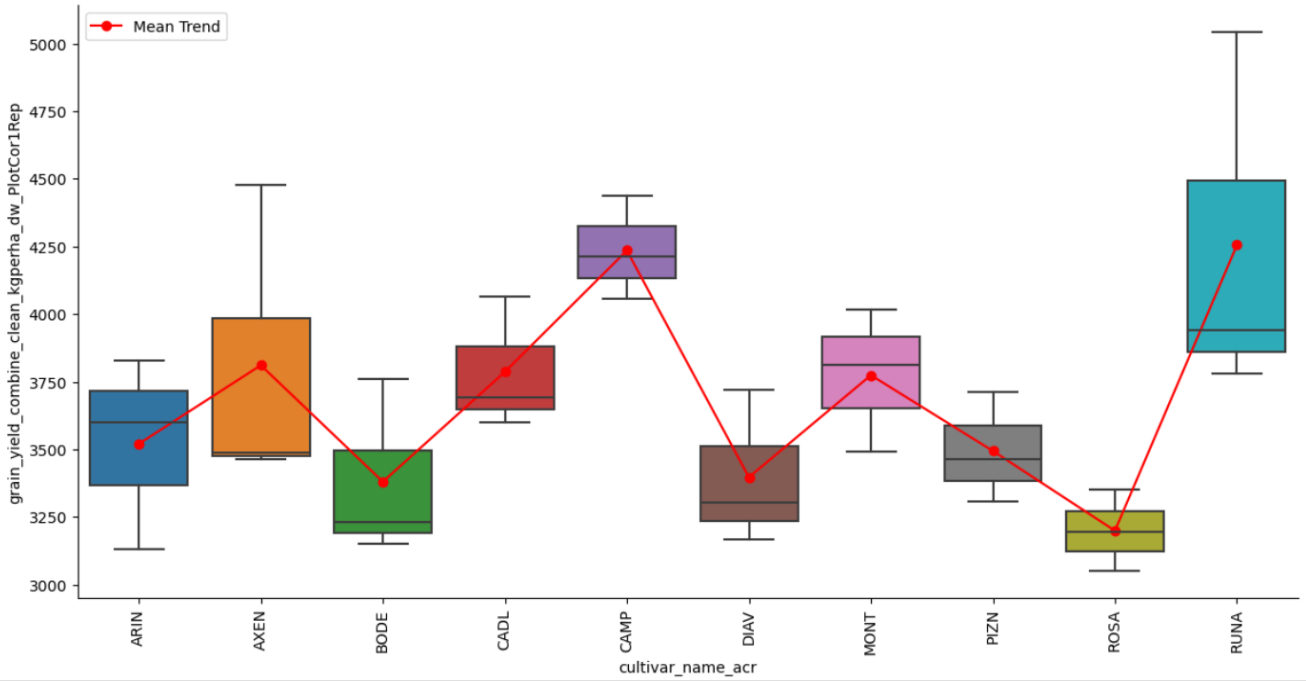
	df	sum_sq	mean_sq	F	PR(>F)
C(cultivar_name_acr)	9.0	3.832210e+06	425801.133291	3.618656	0.007953
Residual	20.0	2.353366e+06	117668.321377	NaN	NaN

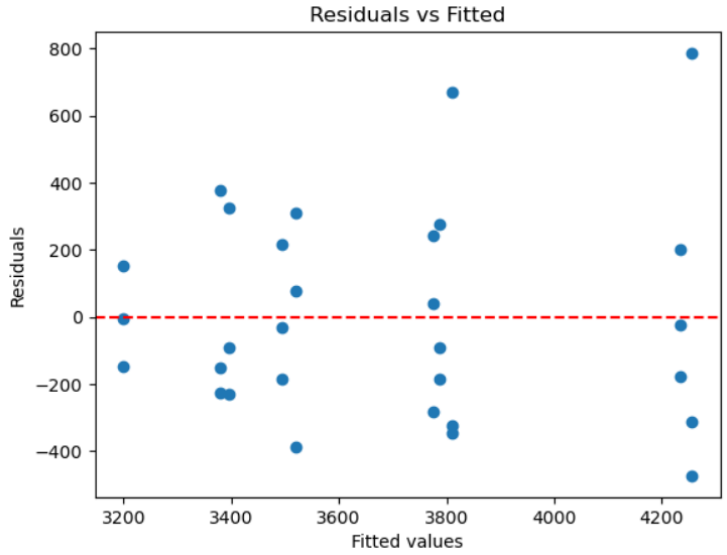


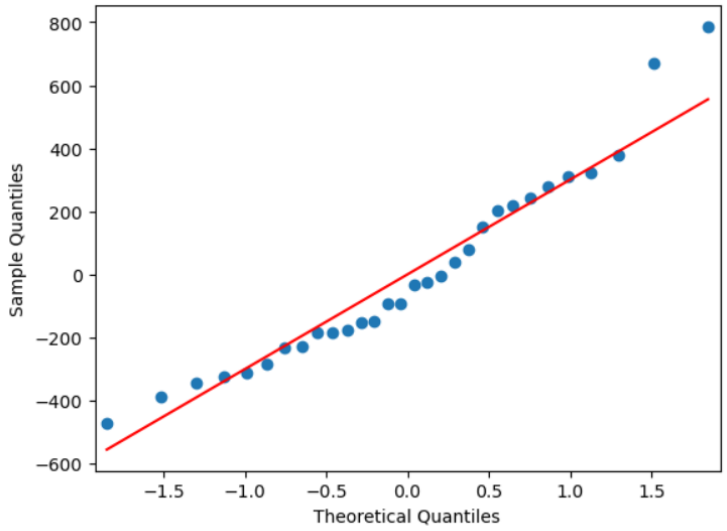



```
Shapiro-Wilk Test: W=0.941, p-value=0.098
```

N0 analysis: 2023







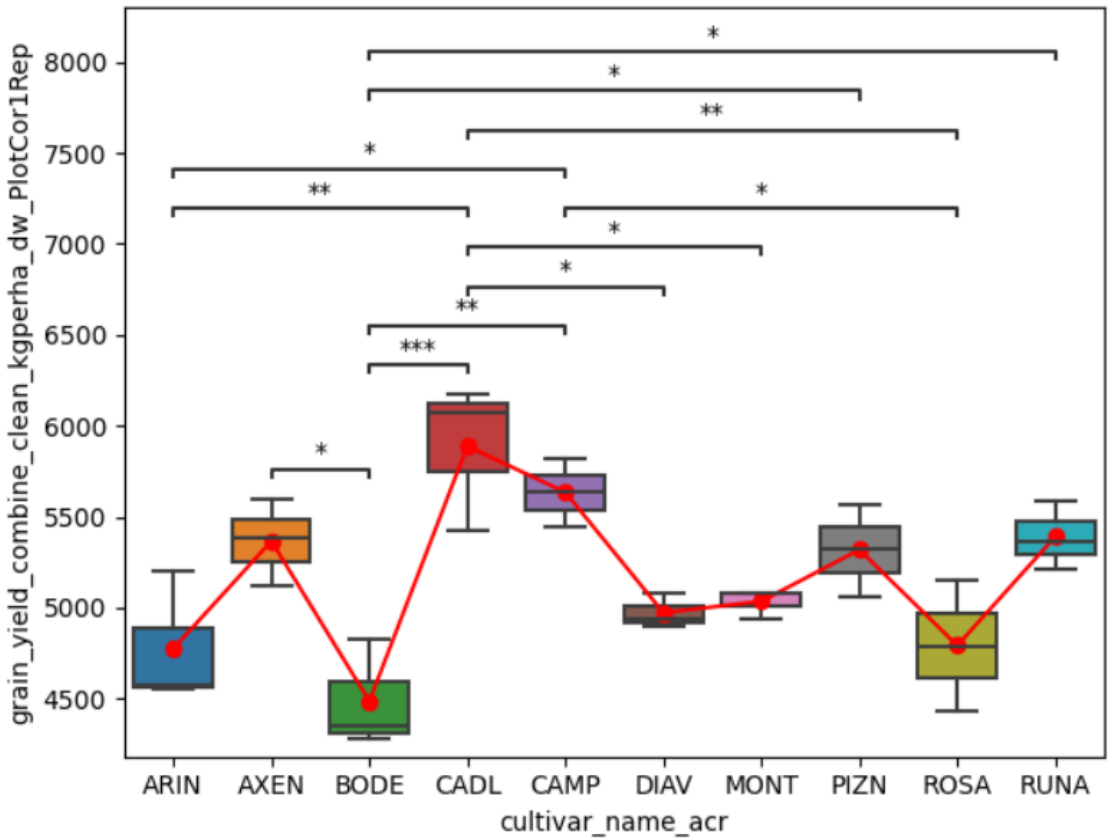
Shapiro-Wilk Test: W=0.946, p-value=0.136

```
Bartlett's Test Statistic: 7.406273016990933
P-value: 0.5948990035238204
The variances are not significantly different (homogeneity of variance met).
```

	df	sum_sq	mean_sq	F	PR(>F)
C(cultivar_name_acr)	9.0	3.410594e+06	378954.920347	2.797207	0.026552
Residual	20.0	2.709524e+06	135476.197805	NaN	NaN

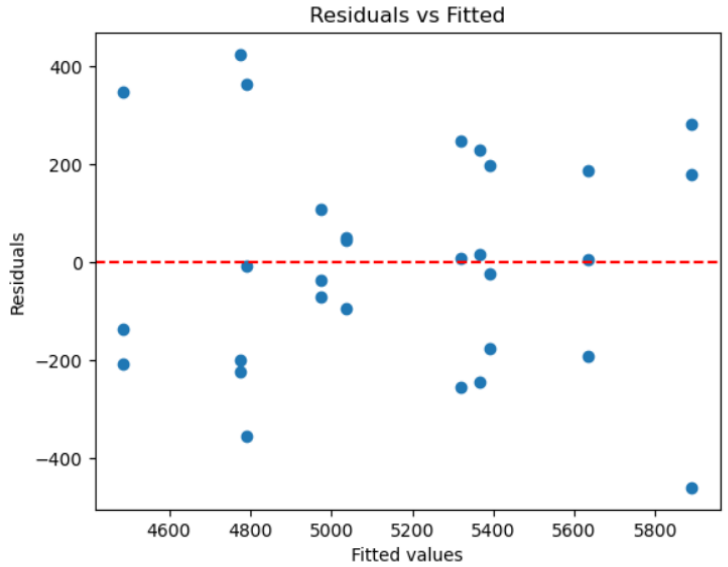
Agroscope

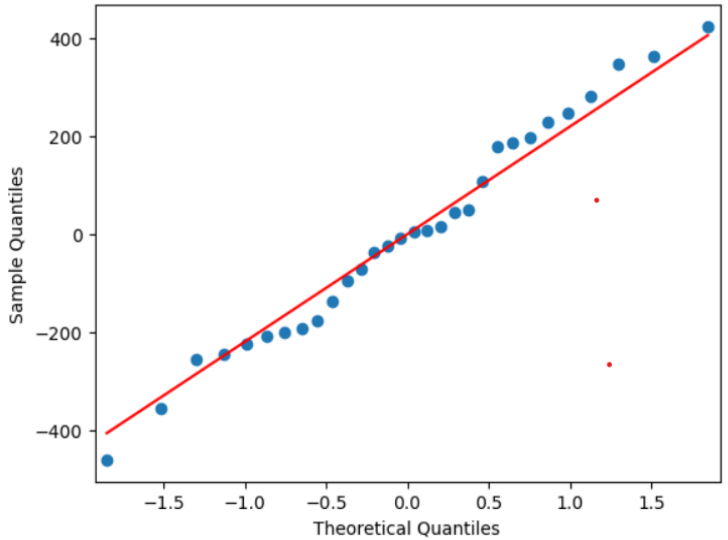
Conventional N analysis: 2023



	df	sum_sq	mean_sq	F	PR(>F)
C(cultivar_name_acr)	9.0	4.989126e+06	554347.311137	7.672642	0.00008
Residual	20.0	1.444997e+06	72249.860362	NaN	NaN

Bartlett's Test Statistic: 6.785088500241943
P-value: 0.6594834543853669
The variances are not significantly different (homogeneity of variance met).

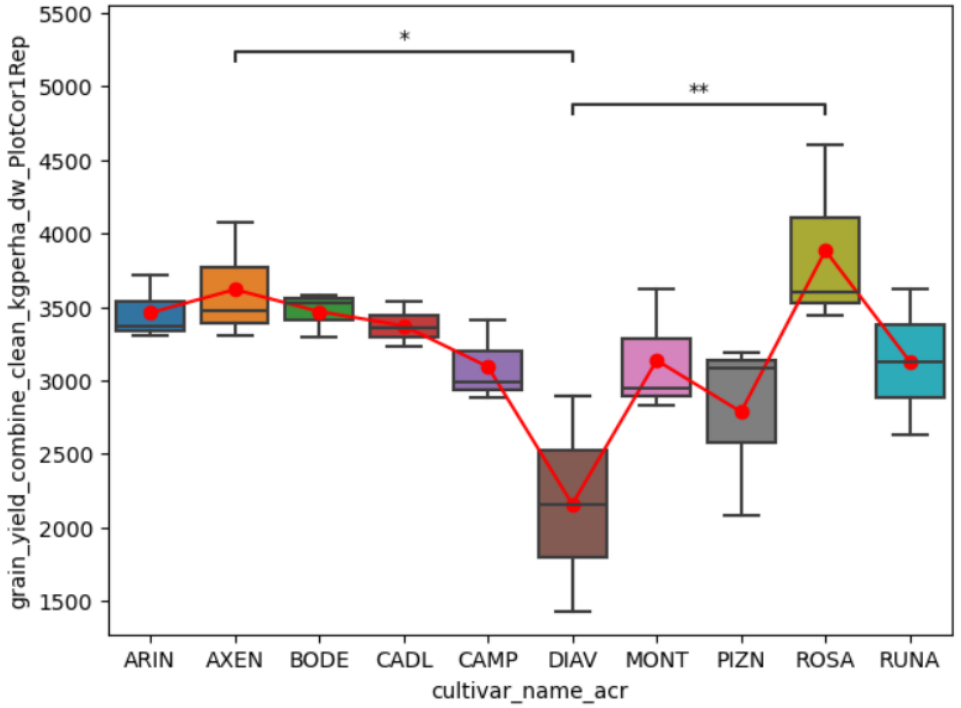




Shapiro-Wilk Test: W=0.980, p-value=0.838

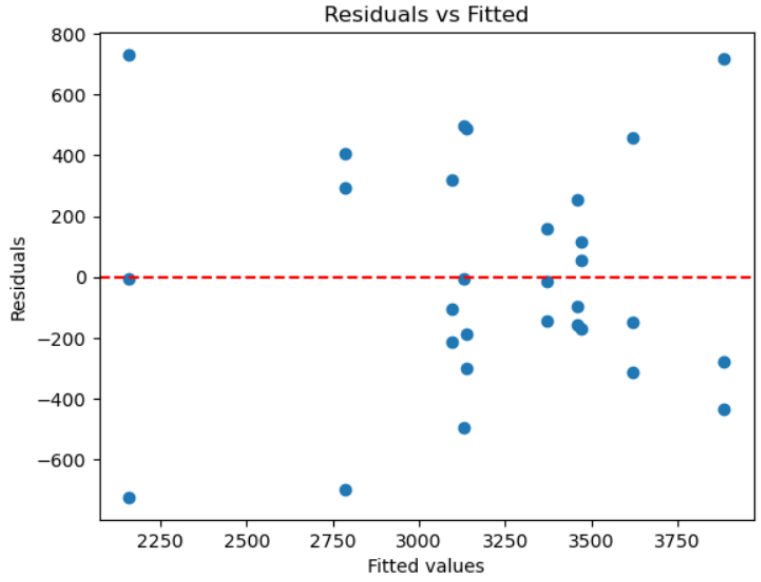
N0 analysis: 2024

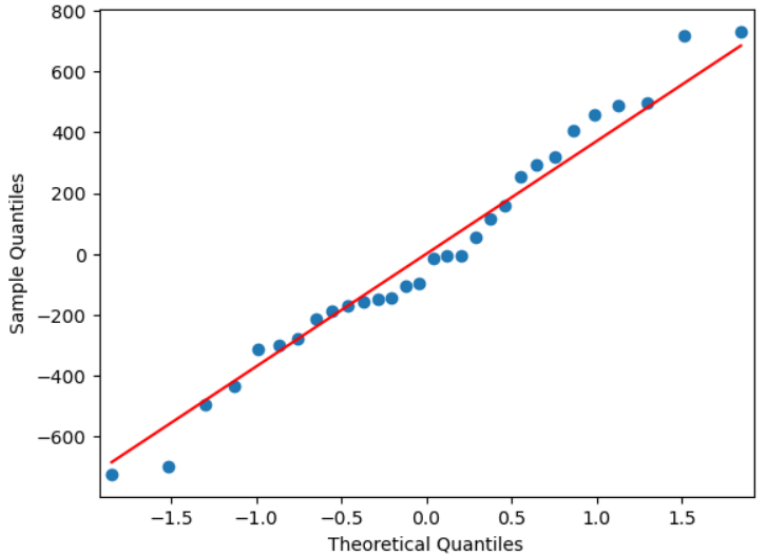
```
DIAV vs. ROSA: Custom statistical test, P_val:4.678e-03
AXEN vs. DIAV: Custom statistical test, P_val:2.198e-02
```



	df	sum_sq	mean_sq	F	PR(>F)
C(cultivar_name_acr)	9.0	6.261364e+06	695707.124299	3.373153	0.011272
Residual	20.0	4.124966e+06	206248.296859	NaN	NaN

```
Bartlett's Test Statistic: 8.024622816168675
P-value: 0.5316667461540587
The variances are not significantly different (homogeneity of variance met).
```

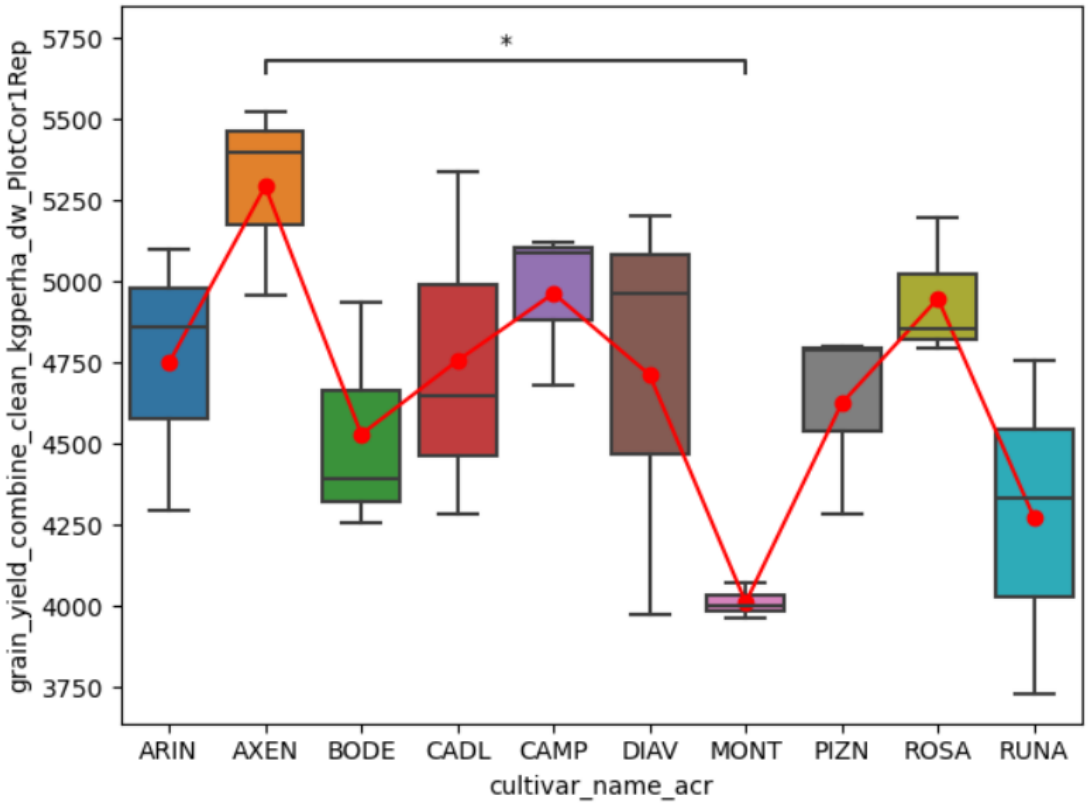





```
Shapiro-Wilk Test: W=0.975, p-value=0.680
```

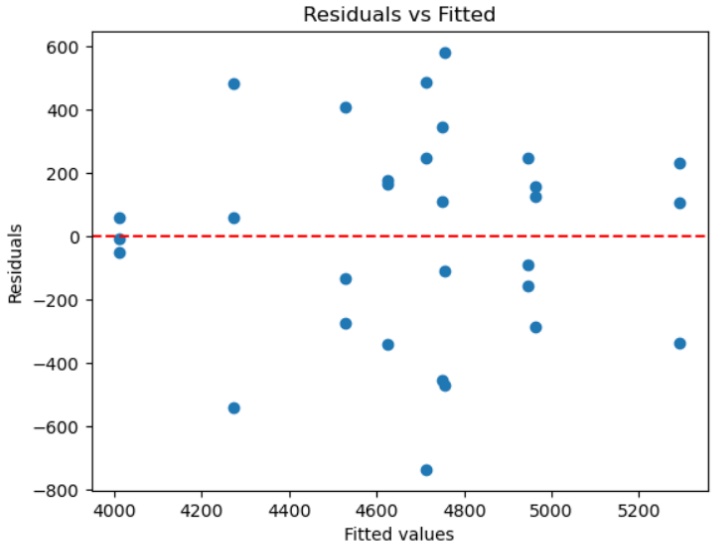
Conventional N analysis: 2024

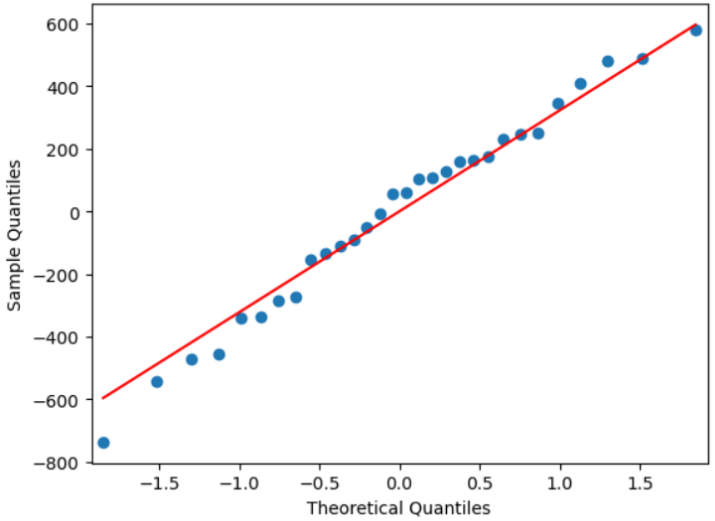
AXEN vs. MONT: Custom statistical test, P_val:2.038e-02



Bartlett's Test Statistic: 8.809594737807885
P-value: 0.4550328502148302
The variances are not significantly different (homogeneity of variance met)

	df	sum_sq	mean_sq	F	PR(>F)
C(cultivar_name_acr)	9.0	3.530605e+06	392289.441904	2.514883	0.041199
Residual	20.0	3.119743e+06	155987.146134	NaN	NaN

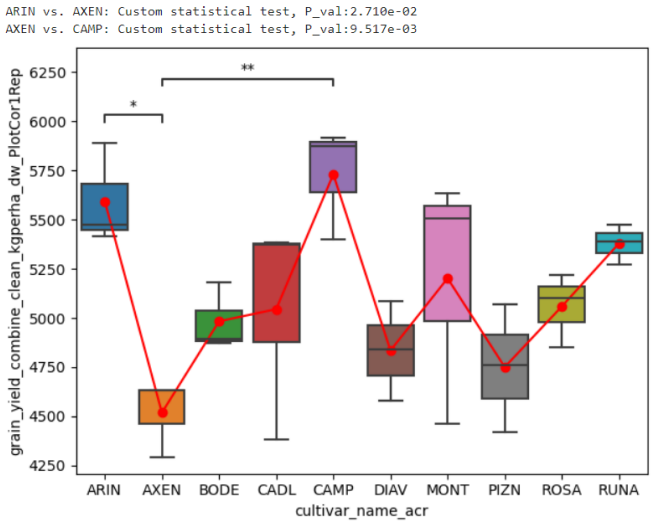




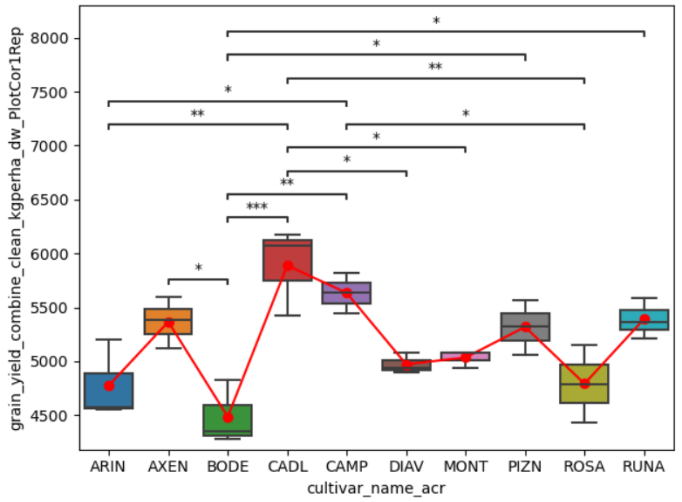
Shapiro-Wilk Test: W=0.980, p-value=0.836

Summary: statistical differences between varieties

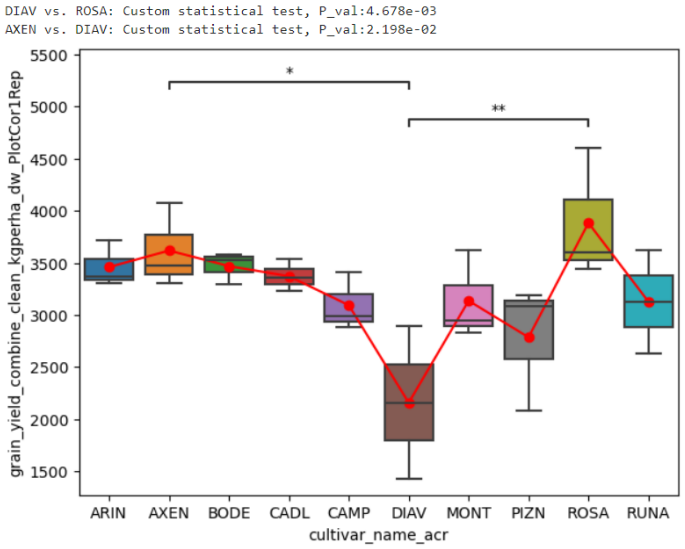
Conventional N - 2022



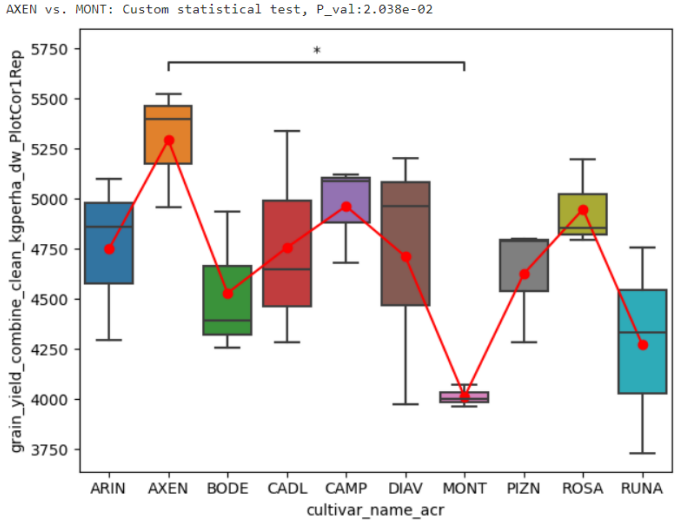
Conventional N - 2023



N0 - 2024



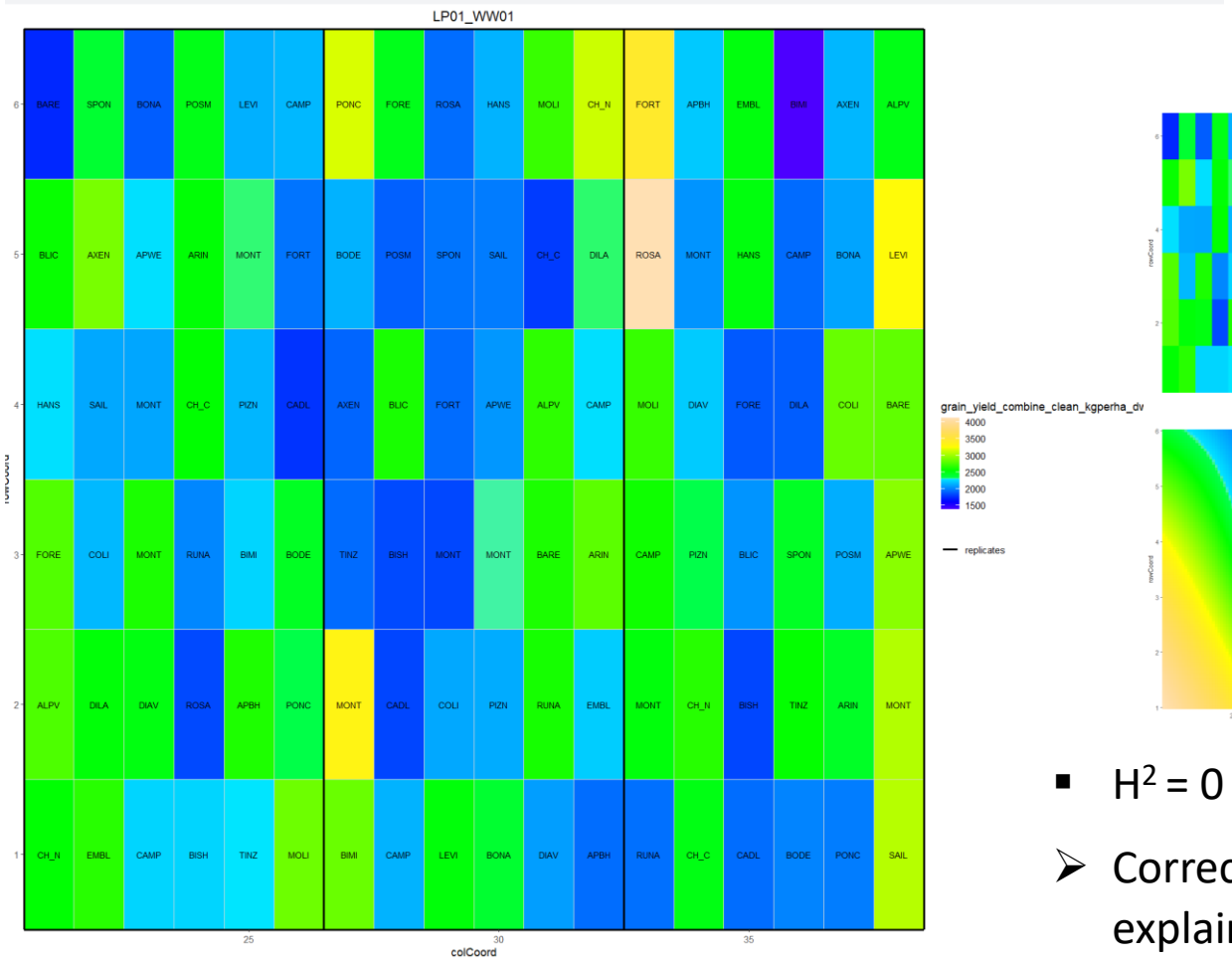
Conventional N - 2024

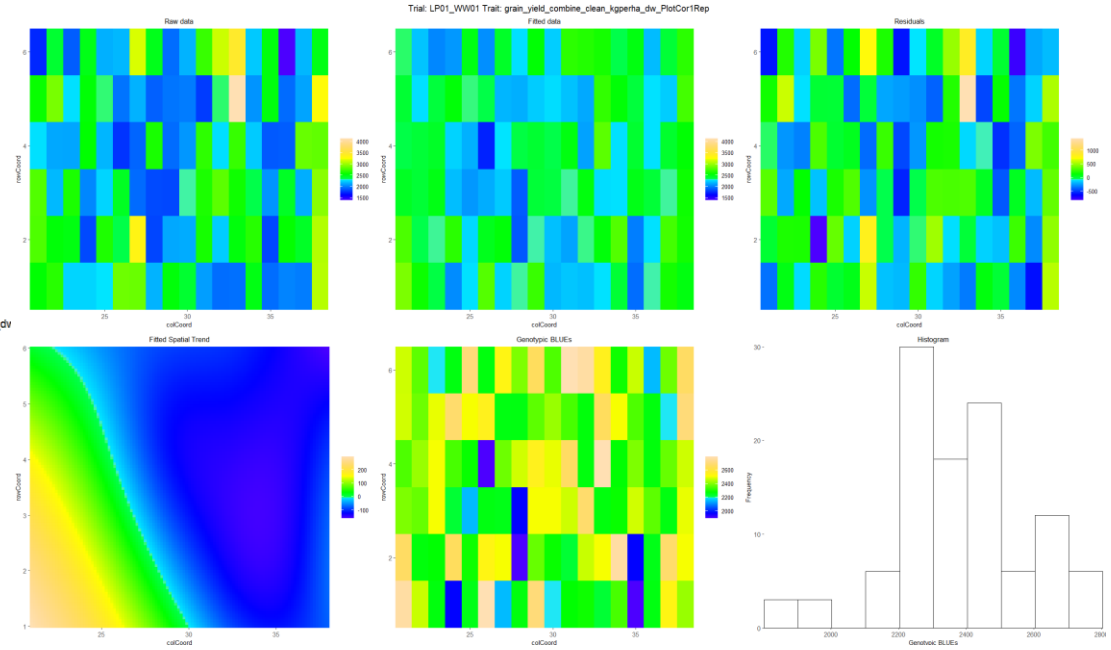


- 2021: grain yield too low – no significant differences found between varieties in both N treatment block.
- N0: no significant differences found, except in 2024.
- Campesino variety: consistent high grain yield over years and significantly different from other varieties.
- Axen variety: high grain yield (2023 and 2024) with significant differences from other varieties but seemed sensible to not optimal growth condition (2021 and 2022).

Agroscope

Spatial heterogeneity: source of variance: 2021 (N0)



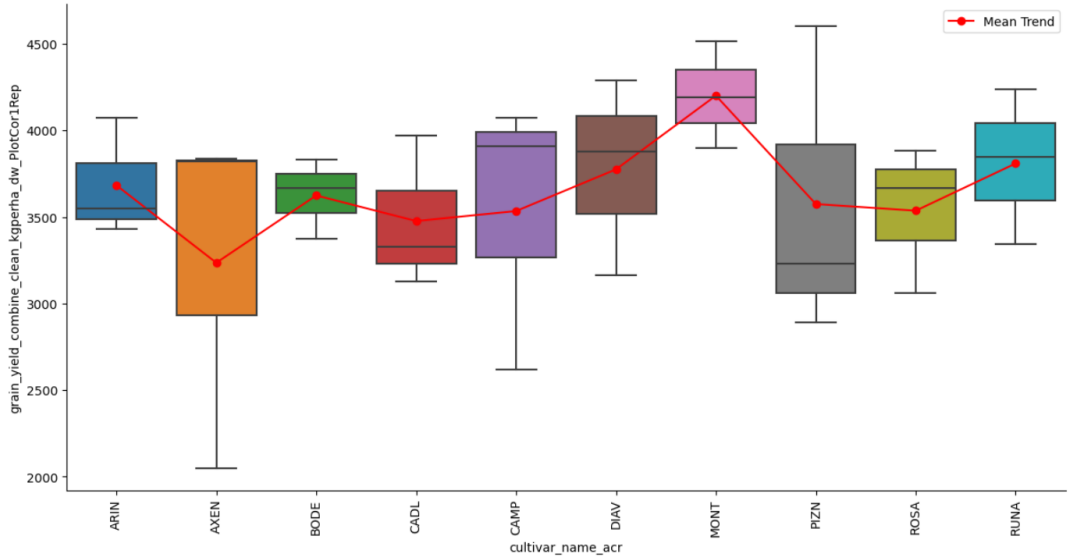


- $H^2 = 0$

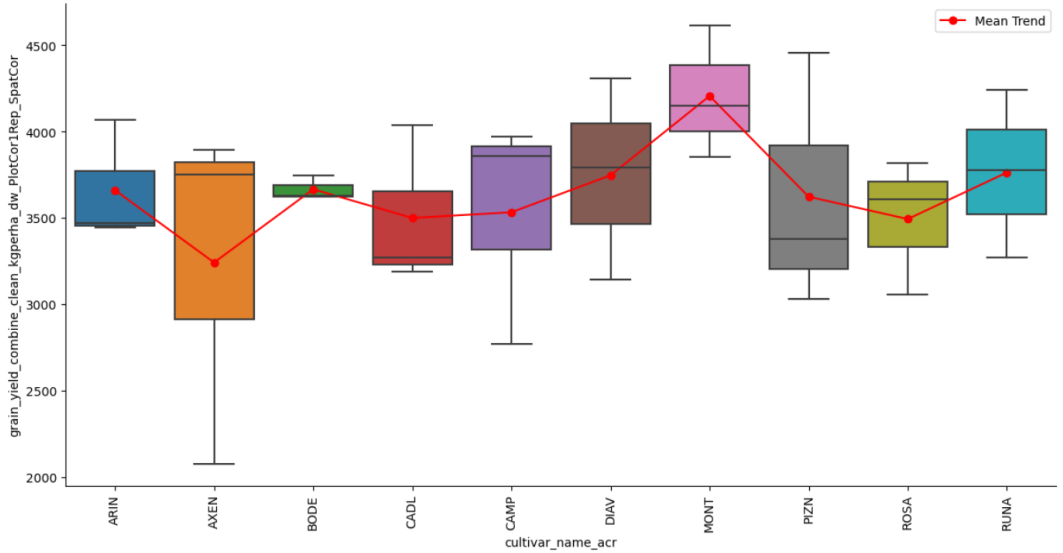
➢ Correcting spatial trend did not increase variance part explained by the variety

N0 analysis: spatial correction in 2021

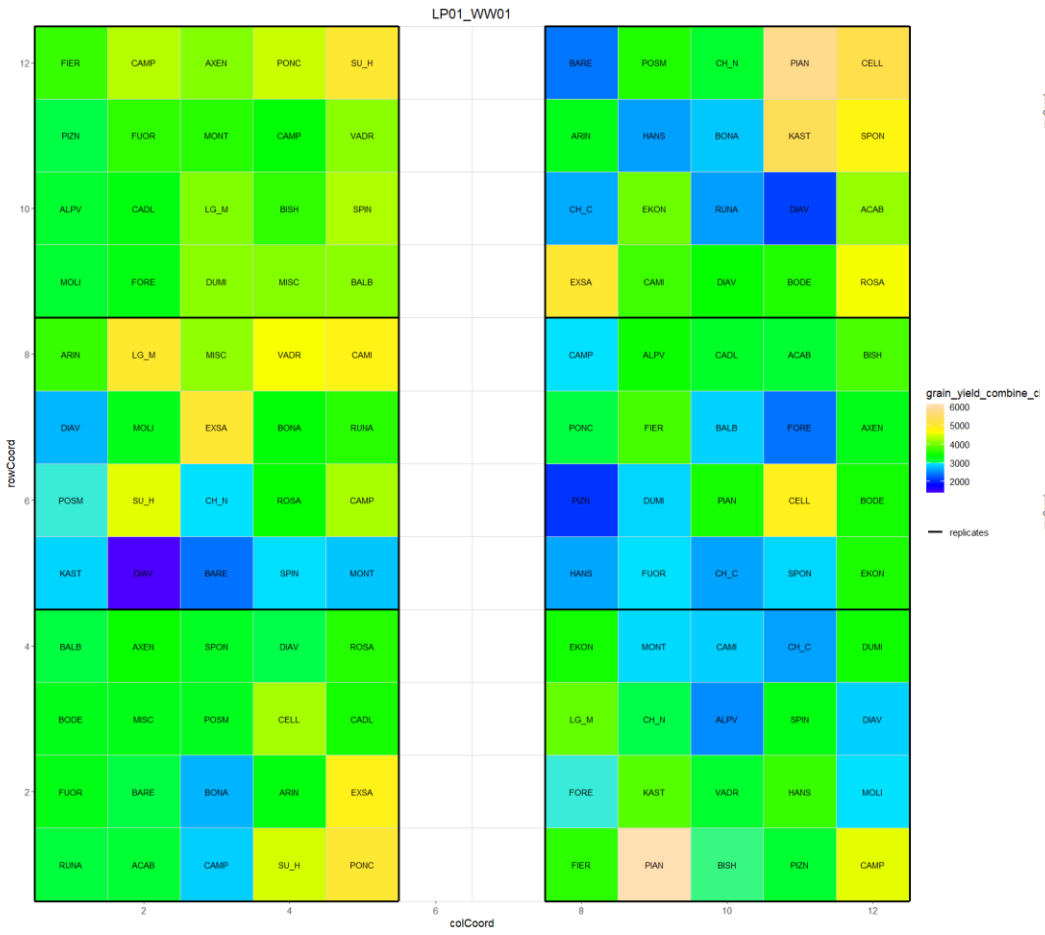
Before correction

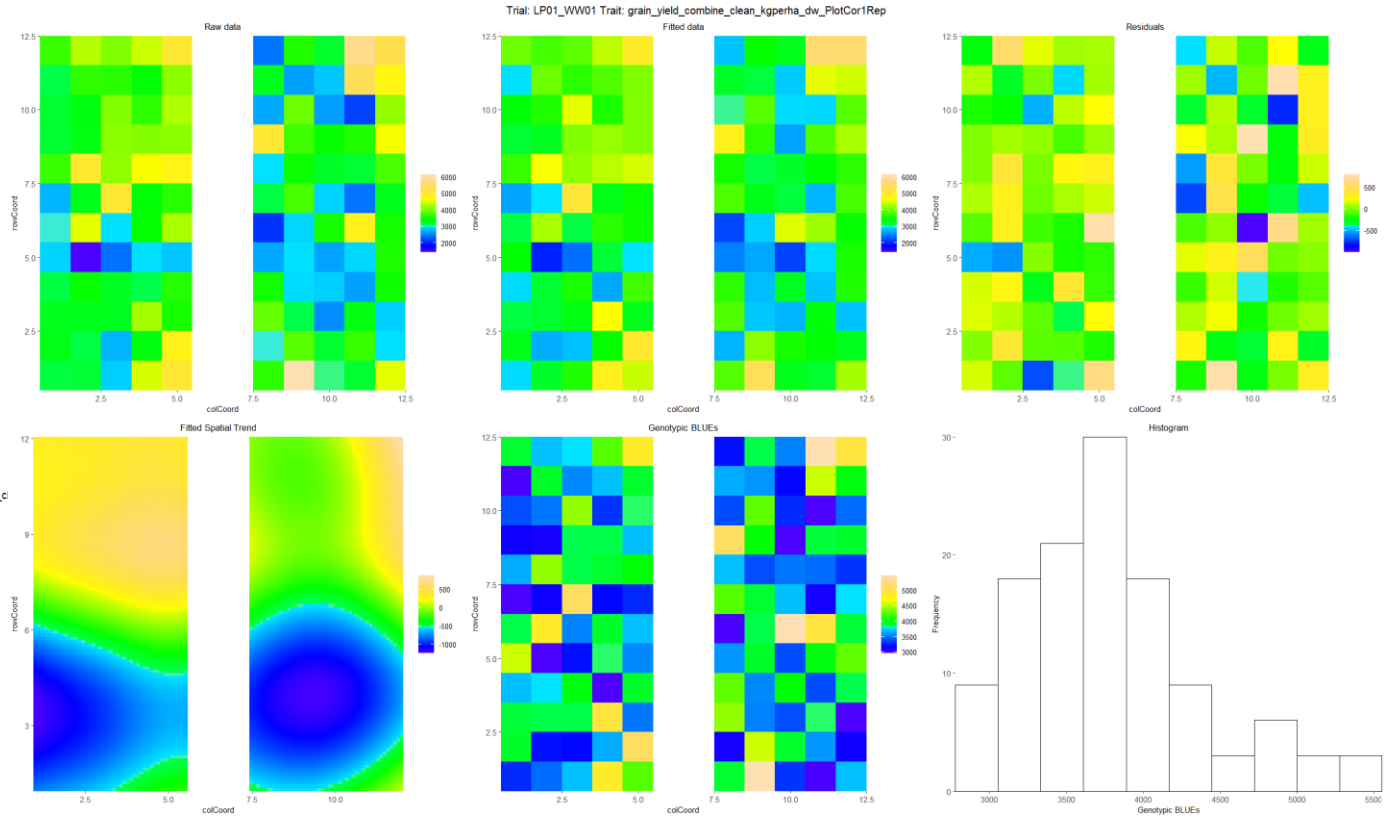


After correction



Spatial heterogeneity: source of variance: 2024 (N0)



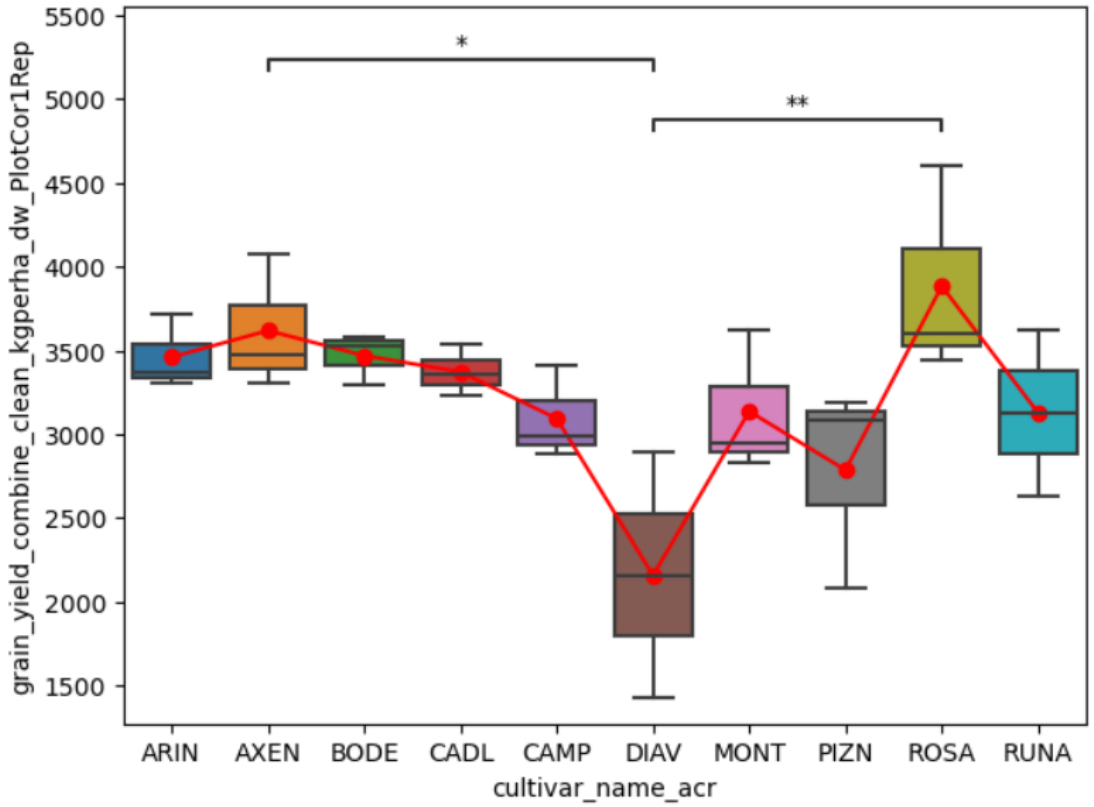


- $H^2 = 0.84$

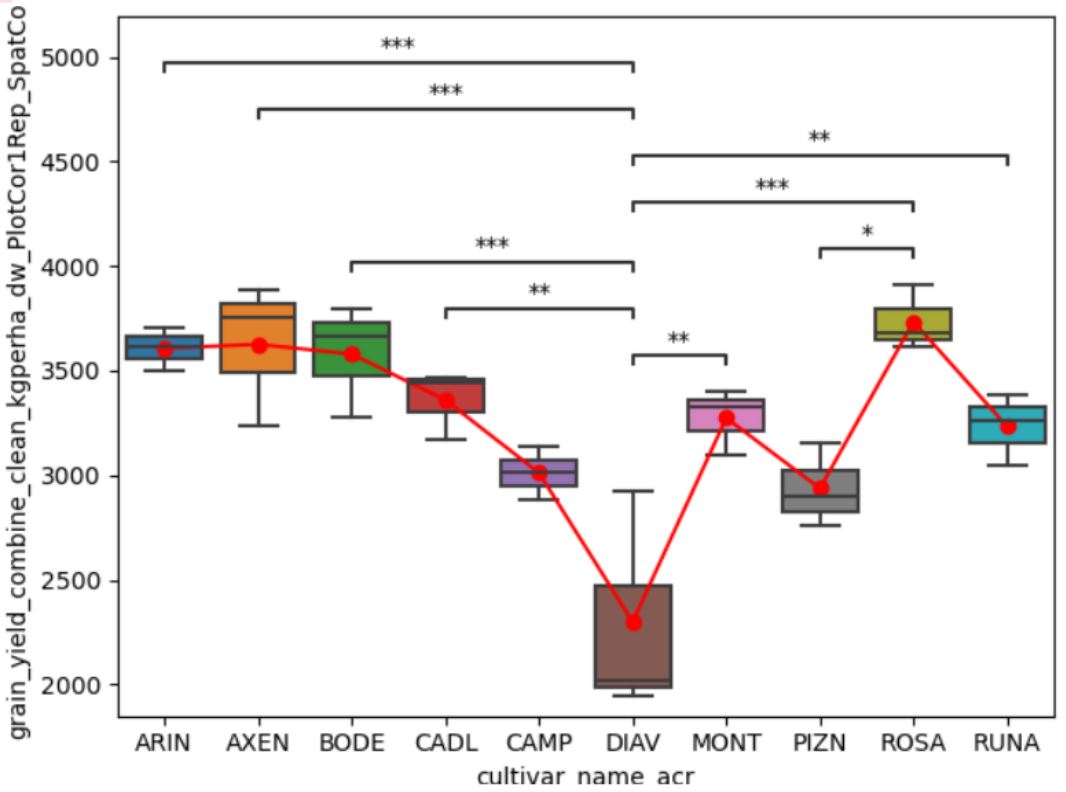
➢ Correcting spatial trend increased variance part explained by the variety

N0 analysis: spatial correction in 2024

Before correction



After correction



Conclusion

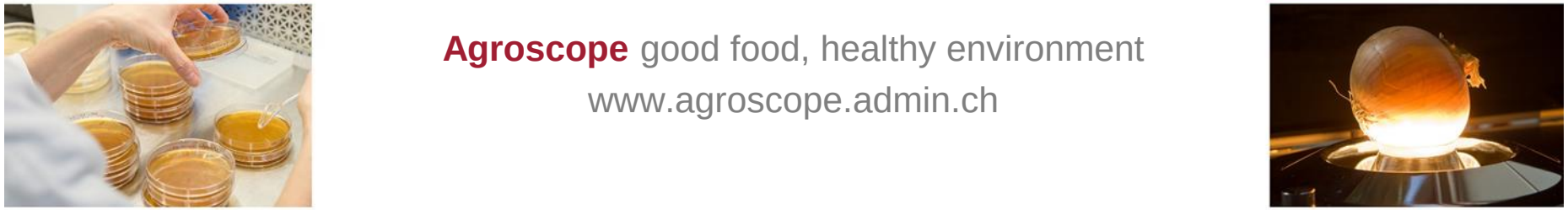
- **Field data: can be noisy:**
  - Difficult to highlight statistical differences between variety
  - Low statistical power (3 replicates)
  - Spatial correction could be used (Spat package from R) to reduce noise regarding field spatial heterogeneity

- **Field experimental design not optimal to study interaction between N treatment and variety:**
  - Limited to field practical resource
  - Dependence of observations within blocks
  - Other trials with less varieties are used to investigate this interaction with optimal field trial design

- **Difficult to find statistical differences when most of values are low (N0 block treatment):**
  - 2024: reduced N block (only one year of experiment)

- **Other agronomic parameter could be analysed to really understand how the variety response is interacting with specific environment and climatic conditions**
  - Straw yield (positively correlated with grain yield)
  - Harvest index (indicates the ratio of grain on total plant weight (grain + straw)
  - Protein in grain (negatively correlated with grain yield)
  - Plant height (positively correlated with grain and straw yield; lodging effect)
  - Disease (can explain unexpected low yield)



Thank you for your attention

Nicolas Vuille-dit-Bille

nicolas.vuille-dit-bille@agroscope.admin.ch

Agroscope good food, healthy environment

www.agroscope.admin.ch