



***Vorstellung einer R-Funktion zur
Auswertung und Ergebnisdarstellung
einfacher Feldversuche***

Johannes Döhler

SKW Piesteritz, LAF Cunnersdorf



Piesteritz (Lutherstadt Wittenberg)
Produktionsstätte

Forschung und Entwicklung

- Analytik
- Chemische Forschung



Cunnersdorf (bei Leipzig)

Forschung und Entwicklung

- Landwirtsch. Anwendungsforschung
(24 Mitarbeiter)

Landwirtschaftliche Anwendungsforschung Cunnersdorf (LAF)

Freilandversuche



Klimakammer



Gewächshaus



Labor



Freilandversuchswesen LAF Cunnersdorf



Flüssigdüngung



**~18 ha Düngungsversuche
+ ca. 2 ha Auftragsvers.
pro Jahr**



Emissionsmessungen



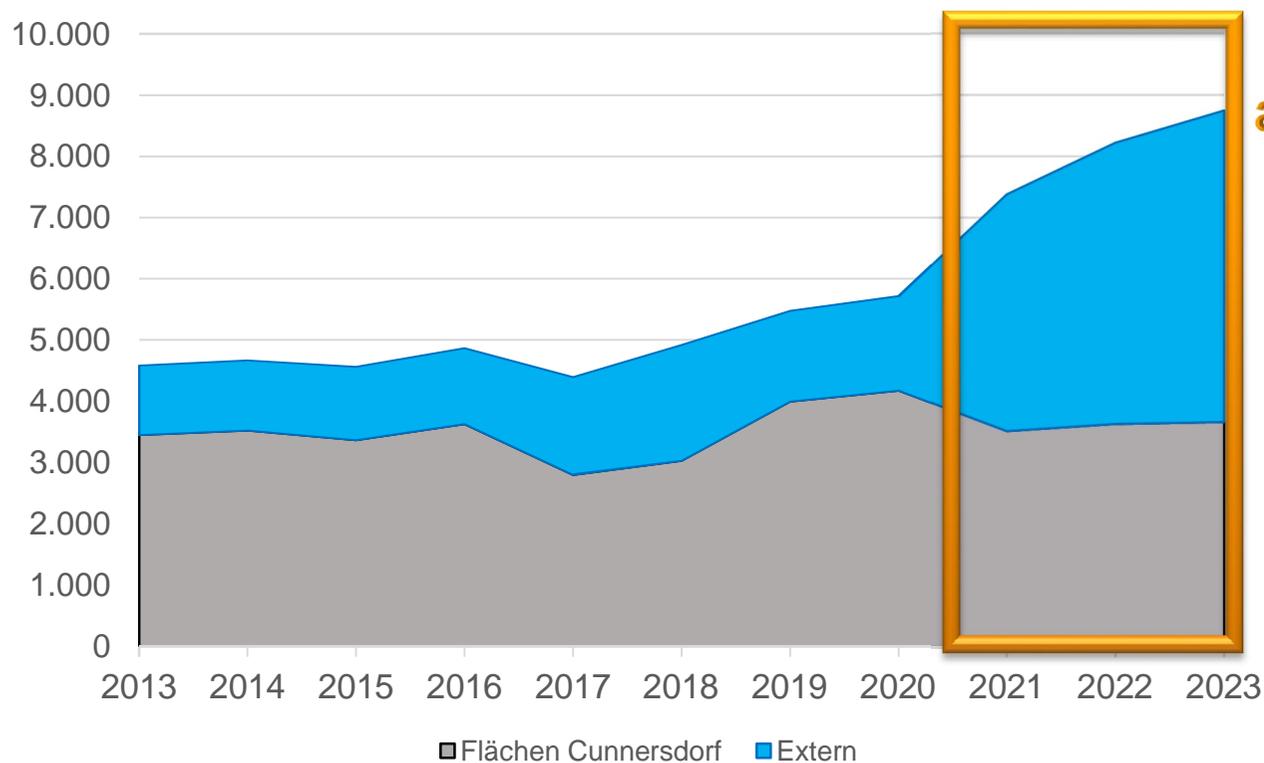
Feststoffdüngung



Gülle

Warum eine R-Funktion selbst erstellen?

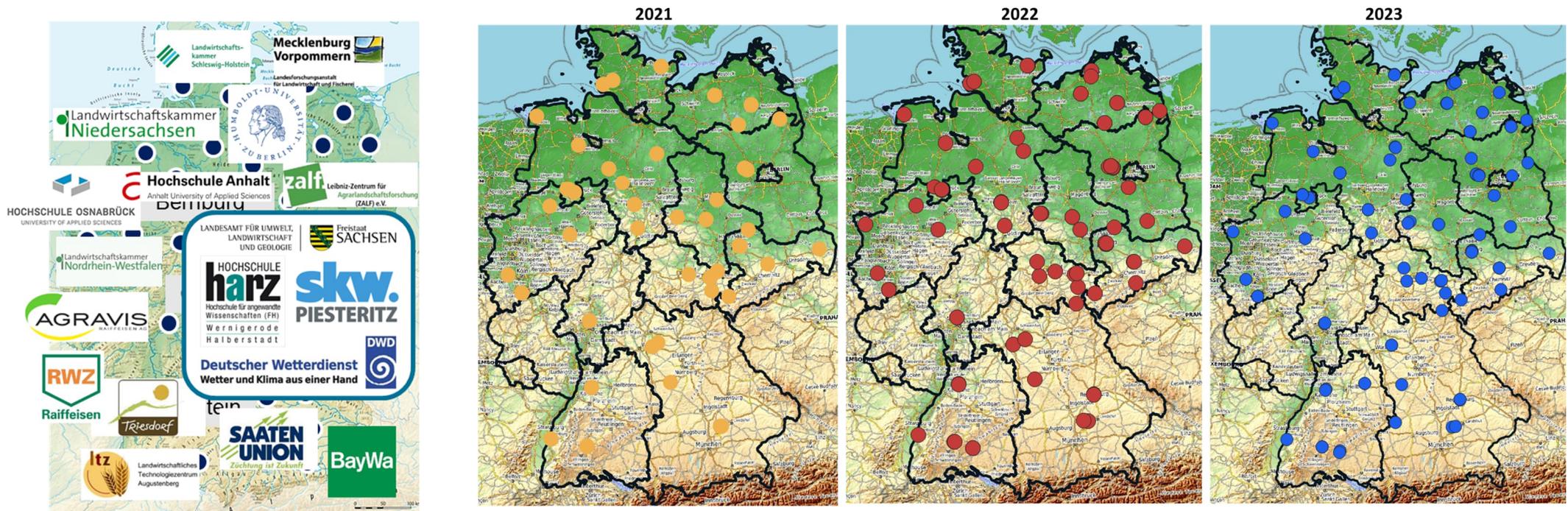
Entwicklung der Parzellenanzahl Düngungsversuche SKW



aktuell fast 9.000 Parzellen
pro Jahr

Warum eine R-Funktion selbst erstellen?

Das Projekt „StaPraxRegio“ – Regionalisierung von Düngestrategien nach bodenkundlichen und meteorologischen Parametern



Was meint der Ausdruck „einfacher Feldversuch“?

- Grundprinzipien des Exaktversuchs eingehalten (Wiederholung, Blockung, Randomisation)
- Einfaktorielle Anlage
- Säulen optional (lat. Rechteck bzw. Quadrat)
- lineares Modell ohne zufällige Effekte - Block und Säule sind fix

Historie der Versuchsauswertung in Cunnersdorf



bis 2012:
Excel-VAs,
Fehlertschätzung mit DAVEP

- ☹ Aufwand
- ☹ Flexibilität
- ☹ Fehlerte

Lageplan	Säulen									
Block	I		II		III		IV			
d	5	3	8	2	7	4	6	1		
c	4	7	1	6	8	2	3	5		
b	6	8	5	7	3	1	4	2		
a	1	2	3	4	5	6	7	8		
Korntrag (dt/ha 86 % TS) (nach Lageplan geordnet)										
Block	I		II		III		IV		B	
d	102,18	72,96	104,48	91,94	95,15	95,25	95,00	103,48	760,43	
c	95,12	89,79	99,09	89,77	102,32	91,41	72,53	100,53	740,55	
b	99,19	104,24	106,01	91,74	72,39	106,78	98,85	96,05	775,25	
a	105,45	93,66	71,38	95,59	104,89	96,03	94,63	103,89	765,52	
L	762,59		750,00		764,21		764,95		3041,74	
Korntrag (dt/ha 86 % TS) (nach Prüfgliedern geordnet)										
Block	1		2		3		4		B	
d	103,48	91,94	72,96	95,25	102,18	95,00	95,15	104,48	760,43	
c	99,09	91,41	72,53	95,12	100,53	89,77	89,79	102,32	740,55	
b	106,78	96,05	72,39	98,85	106,01	99,19	91,74	104,24	775,25	
a	105,45	93,66	71,38	95,59	104,89	96,03	94,63	103,89	765,52	
L	414,80	373,07	289,26	384,81	413,61	379,98	371,30	414,92	3041,74	
SG	289131,57		Ertrag - Wertetabelle							
Varianzanalyse				Pgl.Nr.	Mittelwert	Streuung	Var.Koeff.	Grenzdifferenzen		
Streuungsursache		SQ	FG	MQ	1	103,70	3,36	3,24	Ertrag (dt/ha)	
Gesamt		3129,44	31		2	93,27	2,09	2,24	GD10%	1,95
Block		80,08	3		3	72,31	0,67	0,92	GD 5%	2,37
Säule		18,52	3		4	96,20	1,77	1,84	GD 1%	3,25
Versuchsglied		2985,06	7	426,44	5	103,40	2,50	2,42	GD 0,1%	4,42
Rest		45,77	18	2,54	6	95,00	3,91	4,12		
sd	1,13				7	92,82	2,52	2,72		
sx	0,80				8	103,73	0,97	0,94		
					GD 5%	2,37				

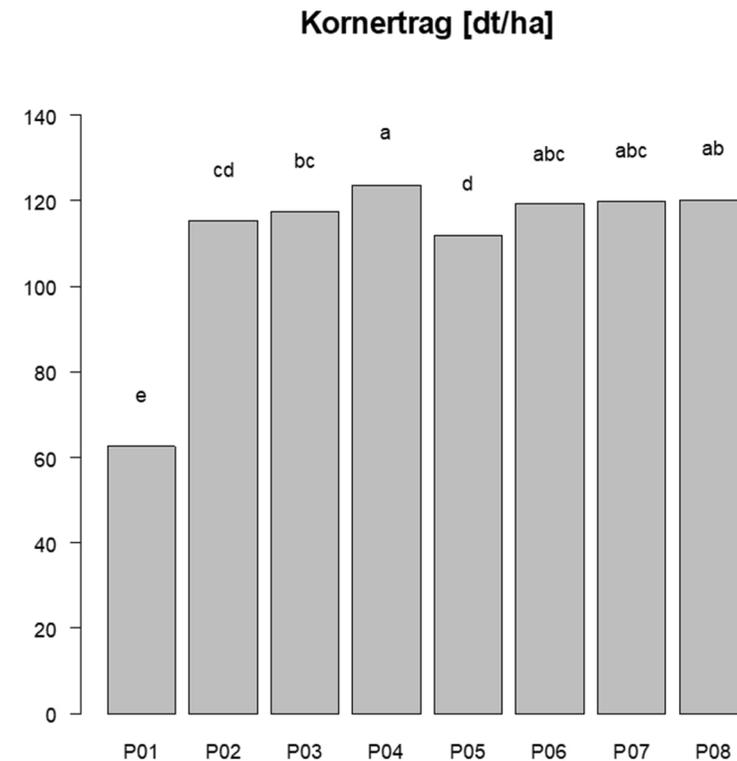
bis 2012:

Excel-VAs,
Fehlertschätzung mit DAVEP

2014 bis 2019:

Excel-VAs + R-Paket agricolae

- ☹ Vergleichs bei unbalancierten Daten
- ☹ Informationsgehalt



(t-Test, alpha = 0.05, LSD= 4.462)

bis 2012:

Excel-VAs,
Fehlertschätzung mit DAVEP

2014 bis 2019:

Excel-VAs + R-Paket agricolae

seit 2019:

Nutzung von PIAFStat (ohne PIAF)

seit 2020:

eigener R-Code, keine Excel-VAs mehr

Anforderungen an auszuwertende Daten:

- Eine Spalte je Merkmal oder Faktor
(wie bei allen gängigen Statistik-Programmen)
- Erforderliche Spalten:
 - *Prüfglied* oder *Pgl* [*Zahl oder Text*]
 - *Block* [*Zahl oder Text*]
 - Zielvariable, z. B. *Ertrag*

Pgl	Block	Ertrag
1	a	98,0
1	b	95,1
1	c	101,8
1	d	95,9
2	a	97,1
2	b	97,2
2	c	99,4
2	d	100,9
3	a	99,7
3	b	95,4
3	c	99,3
3	d	97,4

- R-Software (kostenfrei) mit den Paketen *stringr*, *emmeans*, *multcomp*
- Funktion einmalig einlesen:
`source("C: /R-Skri pte/FUN_Auswertung_FV. R")`
- Daten einlesen als data frame *.csv oder *.xlsx):
`Bei spi el daten <- read.csv2("C: /Düngung/2023/Versuch1.csv", header=TRUE)`



oder mit R-Paket *readxl*:

```
library(readxl)
```

```
setwd("C: /Düngung/2023")
```

```
Bei spi el daten <- as.data.frame(read_excel (Versuch1.xlsx, sheet="Daten"))
```

Vorschau des Datenrahmens prüfen:

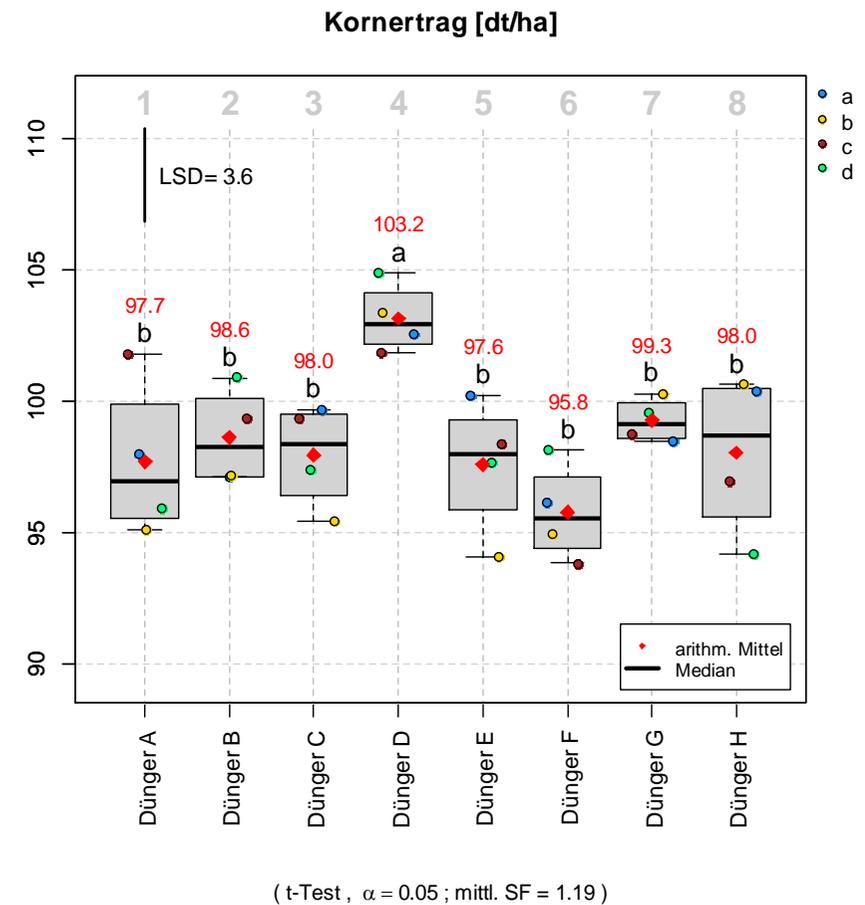
head(Beispiel Daten)

	Pgl	Bez	Block	Säule	Ertrag	RP	NEntz	HL	Sedi	TKM	ÄTH
1	1	Dünger A	a	I	97.96852	11.8	174.4183	80.5	38.7	43.27601	NA
2	1	Dünger A	b	III	95.11221	12.4	177.9433	79.4	43.0	43.05399	NA
3	1	Dünger A	c	II	101.80147	12.3	188.9221	80.3	42.0	43.23179	NA
4	1	Dünger A	d	IV	95.94778	12.1	175.1636	80.1	40.2	43.48026	NA
5	2	Dünger B	a	I	97.13390	12.3	180.2601	80.4	42.3	44.54757	NA
6	2	Dünger B	b	IV	97.16472	12.2	178.8513	80.6	41.4	44.03903	NA

Nutzung der Funktion

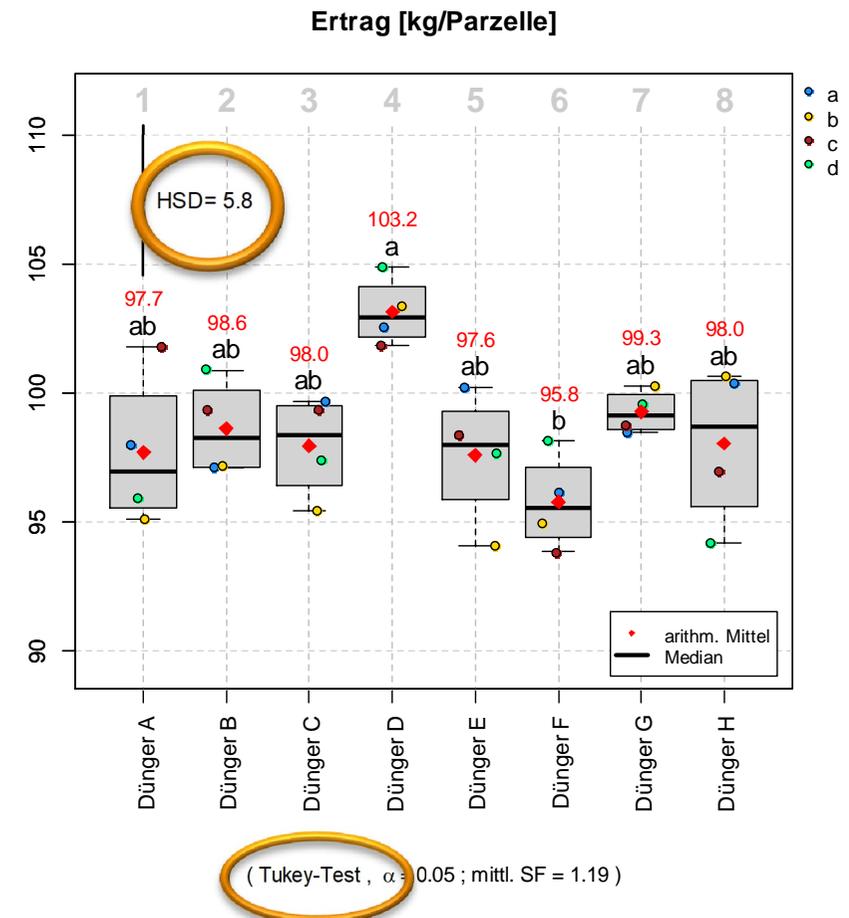
Funktion mit erforderlichen Angaben aufrufen:

`FUN_Auswertung_FV` (Daten. df = Beispiel daten,
Zielvariable = „Ertrag“)



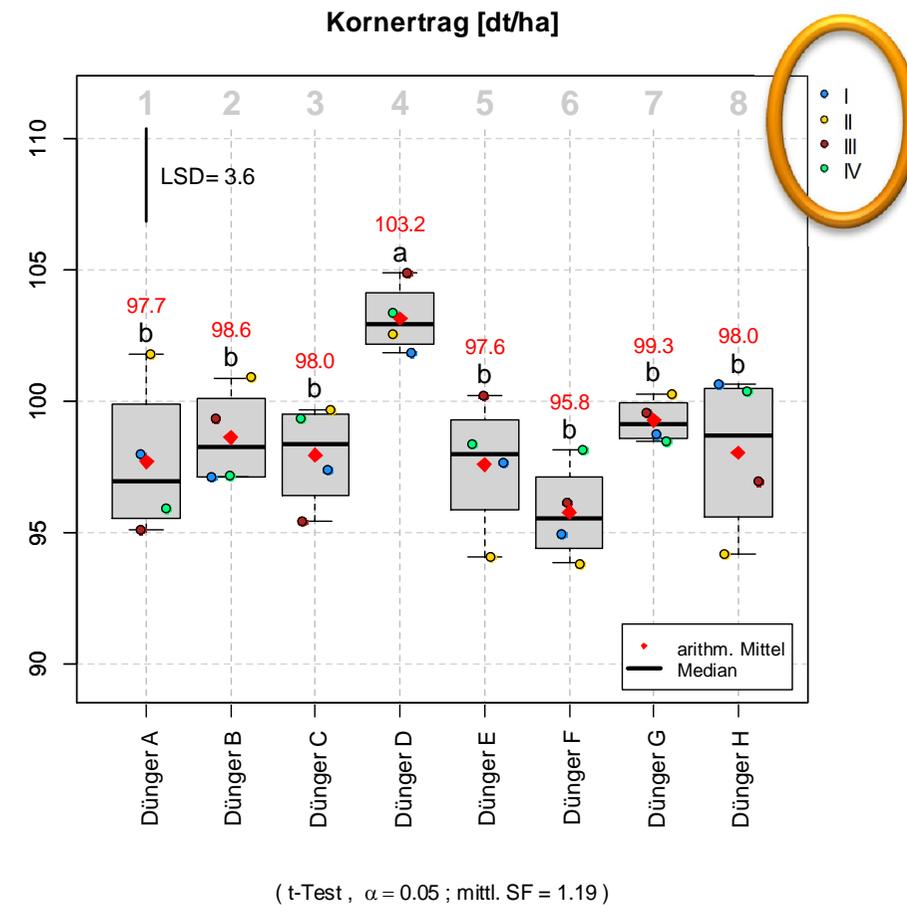
Optionale Anpassungen:

`FUN_Auswertung_FV` (Daten.df = Bei spi el daten,
Zielvariable = „Ertrag“,
Test_Typ=„Tukey-Test“)



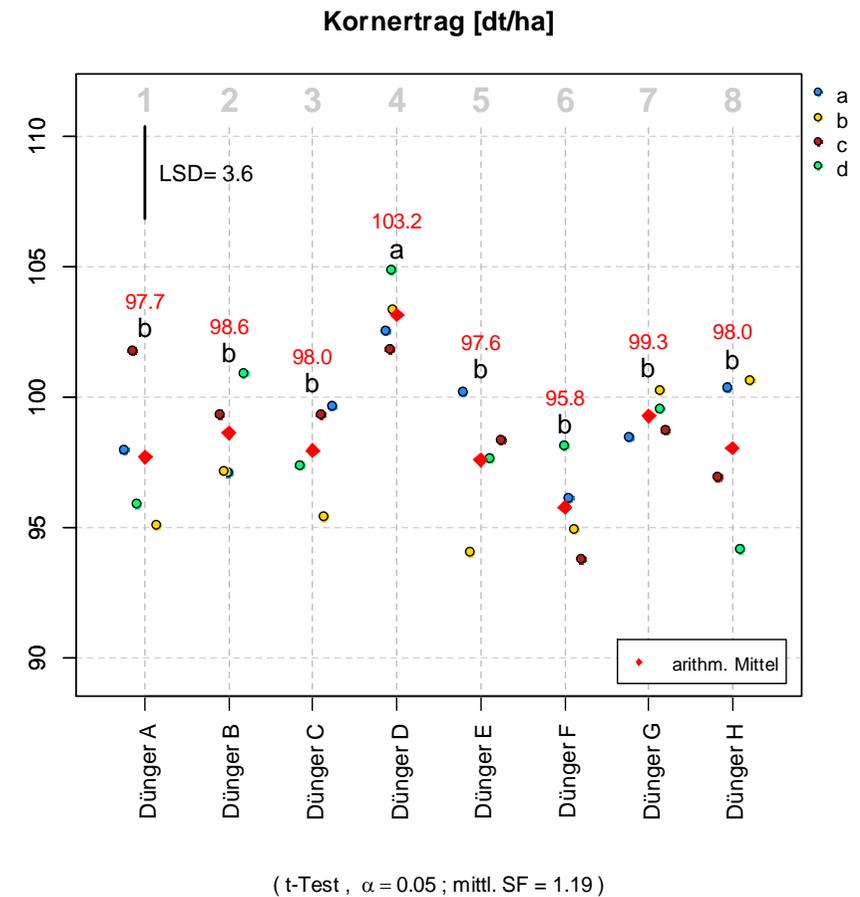
Optionale Anpassungen:

FUN_Auswertung_FV (Daten.df = Beispiel Daten,
Zielvariable = „Ertrag“,
Faktor.Farbe = „Säule“)



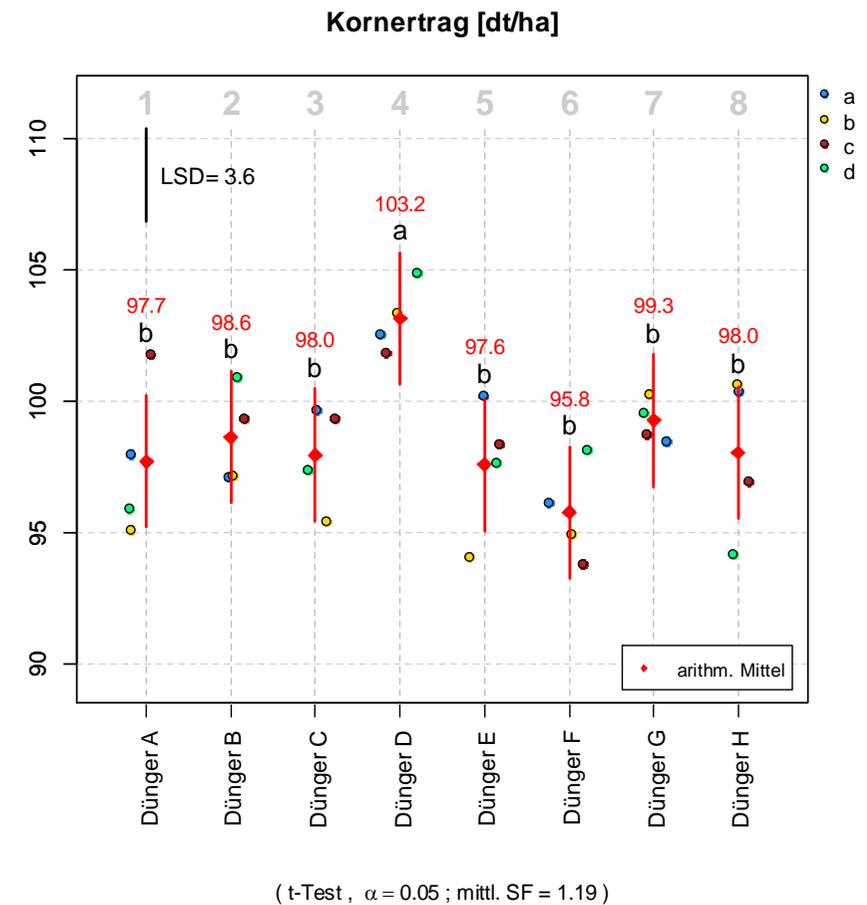
Optionale Anpassungen:

FUN_Auswertung_FV (Daten.df = Bei spi eldaten,
Zi elvari abel e = „Ertrag“,
Boxpl ot_zei gen = F)



Optionale Anpassungen:

FUN_Auswertung_FV (Daten.df = Bei spi el daten,
Zi el vari abel e = „Ertrag“ ,
Boxpl ot_zei gen = F,
VI_zei gen = T)

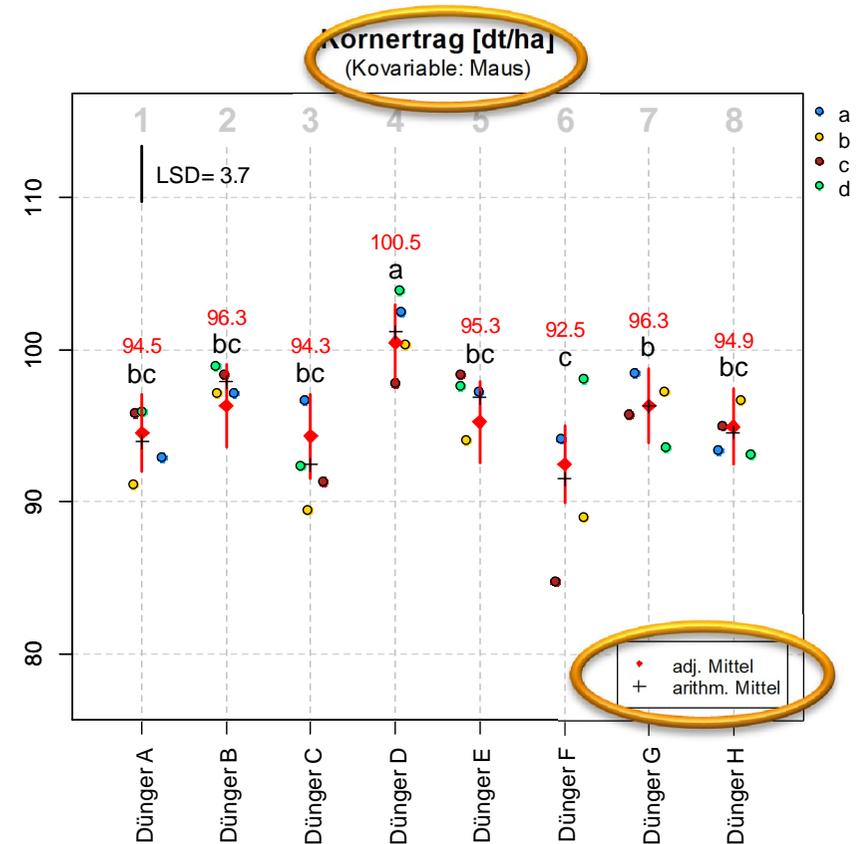


Nutzung der Funktion

Simulierte Daten: Mäusebefall

FUN_Auswertung_FV (Daten.df = Bei spi el daten,
 Zi el vari abl e = „Ertrag“ ,
 Kovari abl e = „Maus“ ,
 Boxpl ot_zei gen = F ,
 VI_zei gen = T)

Pgl	Block	Säule	Maus	Ertrag_alt	Ertrag
Dünger A	a	I	5	98,0	93,0
Dünger A	b	III	4	95,1	91,1
Dünger A	c	II	6	101,8	95,8
Dünger A	d	IV	0	95,9	95,9
Dünger B	a	I	0	97,1	97,1
Dünger B	b	IV	0	97,2	97,2



(t-Test , $\alpha = 0.05$; mittl. SF = 1.23)

Simulierte Daten: Mäusebefall

`FUN_Auswertung_FV` (Daten.df = Bei spi el daten,
Zielvariab le = „Ertrag_al t“,
Kovariabl e = „Maus“,
Boxplot_zei gen = F,
VI_zei gen = T)

Pgl	Block	Säule	Maus	Ertrag_alt
Dünger A	a	I	5	98,0
Dünger A	b	III	4	95,1
Dünger A	c	II	6	101,8
Dünger A	d	IV	0	95,9
Dünger B	a	I	0	97,1
Dünger B	b	IV	0	97,2

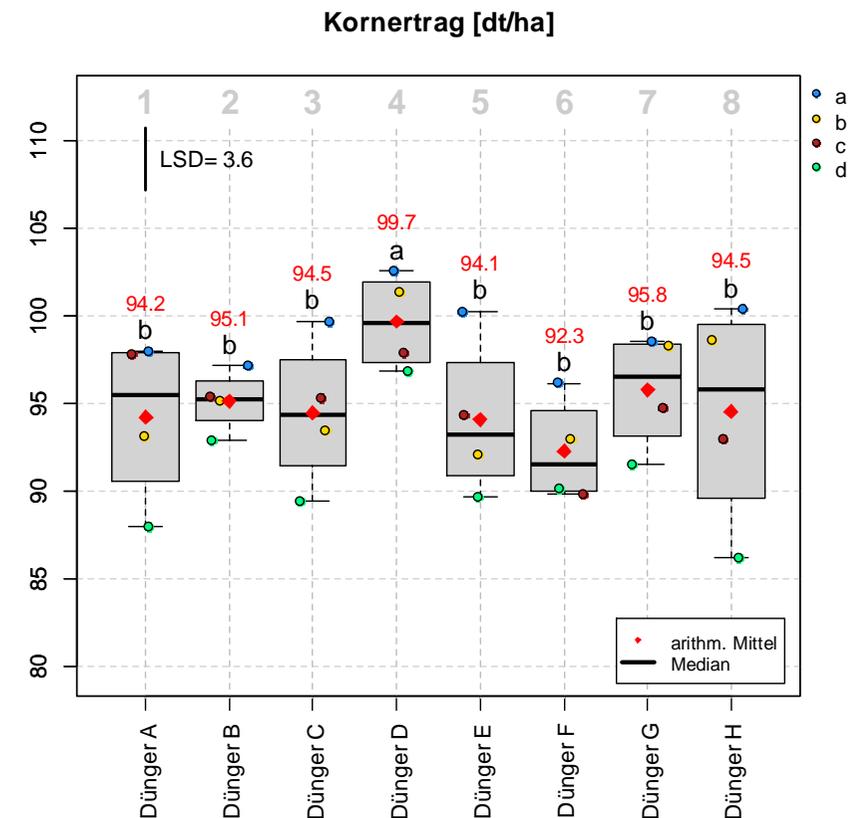
Fehler in `FUN_Auswertung_FV`(Daten.df = Bei spi el daten,
Zielvariab le = "Ertrag_al t", :

ACHTUNG: 'Maus' ist mit $p=0.2232$ als Kovariabl e nicht geei gnet!

Simulierte Daten: Blockeffekt

FUN_Auswertung_FV (Daten.df = Bei spi el daten,
Zi el vari abl e = „Ertrag“)

Pgl	Block	Säule	Blockeffekt	Ertrag_alt	Ertrag
Dünger A	a	I	0	98,0	98,0
Dünger A	b	III	-2	95,1	93,1
Dünger A	c	II	-4	101,8	97,8
Dünger A	d	IV	-8	95,9	87,9
Dünger B	a	I	0	97,1	97,1
Dünger B	b	IV	-2	97,2	95,2
Dünger B	c	III	-4	99,4	95,4
Dünger B	d	II	-8	100,9	92,9



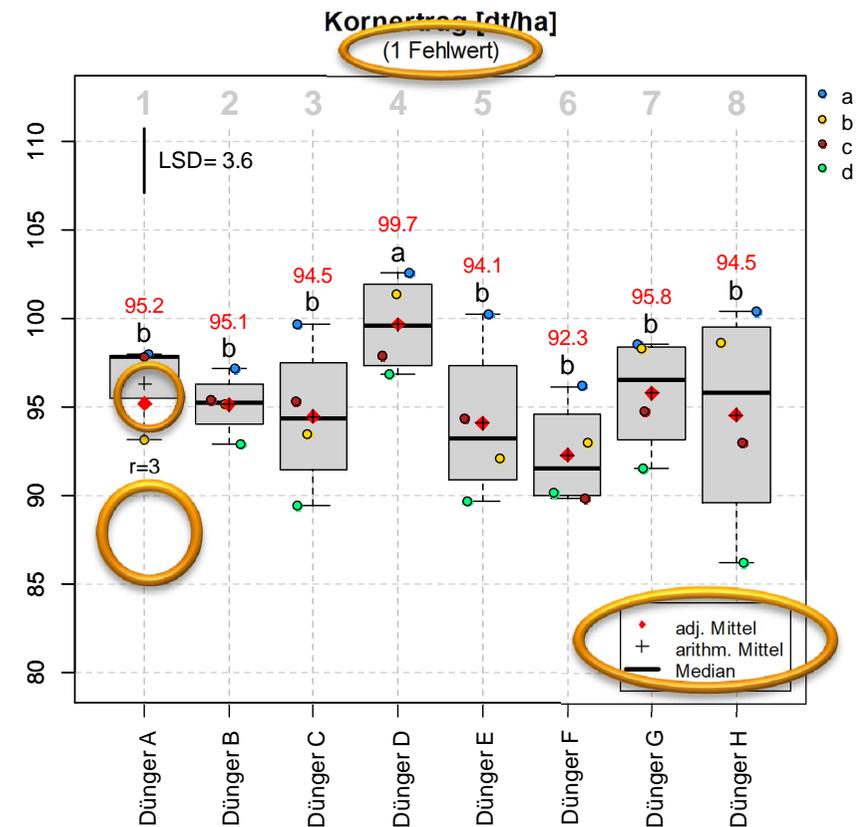
(t-Test , $\alpha = 0.05$; mittl. SF = 1.19)

Nutzung der Funktion

Simulierte Daten: Blockeffekt mit Fehlwert

FUN_Auswertung_FV (Daten.df = Bei spi eldaten,
Zielvariable = „Ertrag“)

Pgl	Block	Säule	Blockeffekt	Ertrag_alt	Ertrag
Dünger A	a	I	0	98,0	98,0
Dünger A	b	III	-2	95,1	93,1
Dünger A	c	II	-4	101,8	97,8
Dünger A	d	IV	-8	95,9	
Dünger B	a	I	0	97,1	97,1
Dünger B	b	IV	-2	97,2	95,2
Dünger B	c	III	-4	99,4	95,4
Dünger B	d	II	-8	100,9	92,9

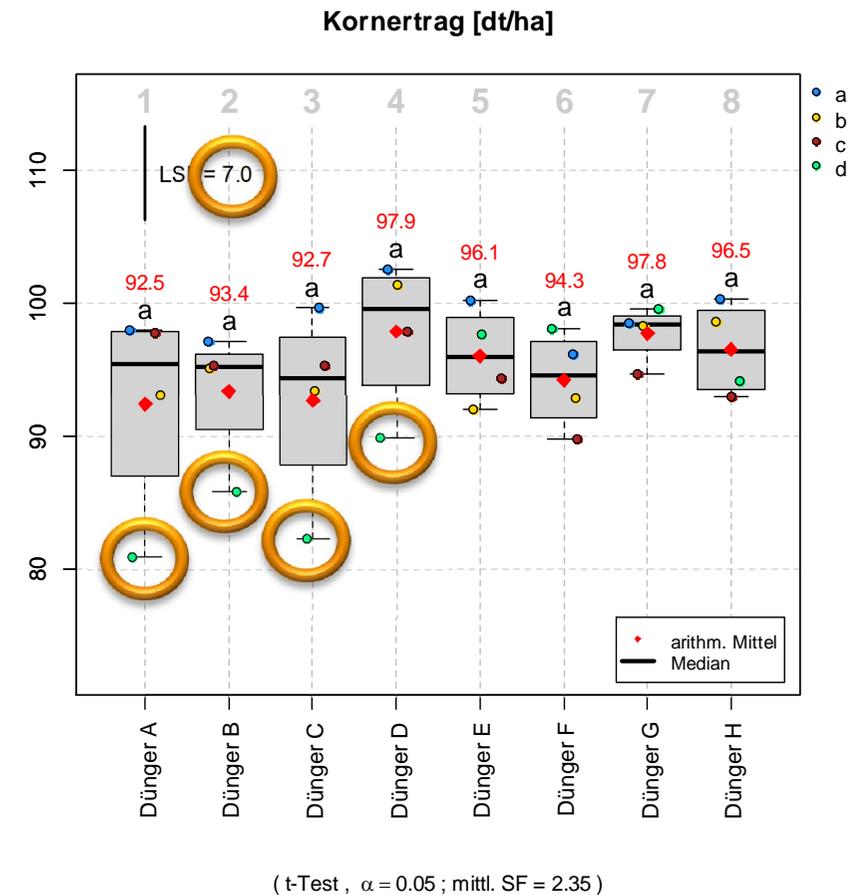


(t-Test , $\alpha = 0.05$; mittl. SF = 1.21)

Simulierte Daten: Störung in Block D

1d bis 4d : Störung
5d bis 8d : keine Störung

FUN_Auswertung_FV (Daten.df = Bei spi eldaten,
Zielvariable = „Ertrag“)

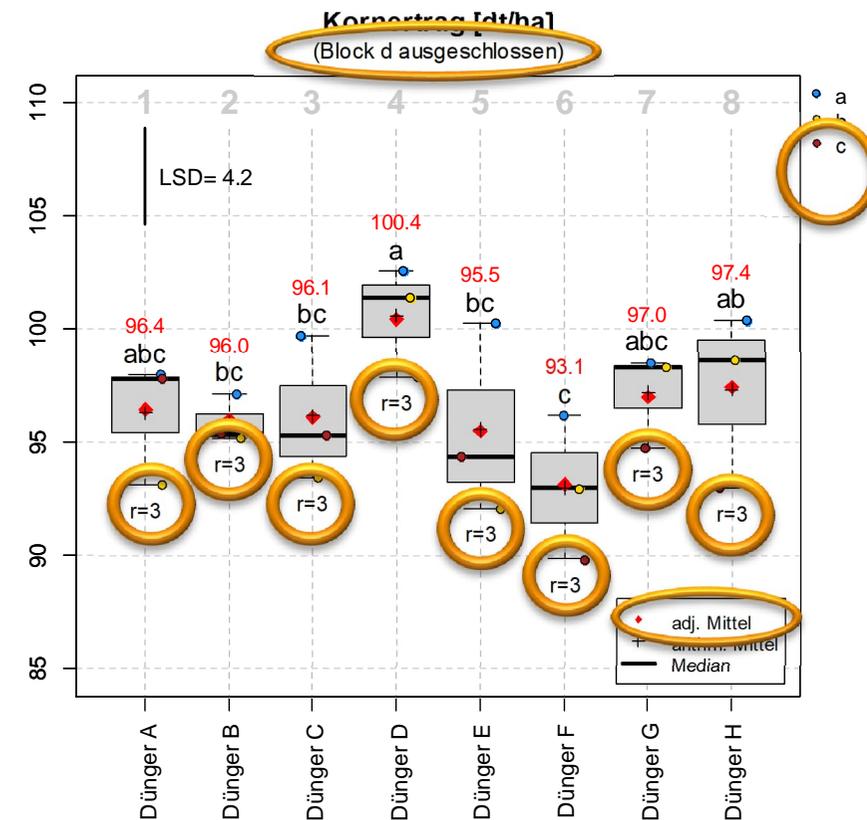


Simulierte Daten: Störung in Block D

1d bis 4d : Störung
5d bis 8d : keine Störung

FUN_Auswertung_FV (Daten.df = Bei spiel daten,
Zielvariable = „Ertrag“ ,
Ausschluss_Block=„d“)

Analog für Säule:
FUN_Auswertung_FV (Daten.df = Bei spiel daten,
Zielvariable = „Ertrag“ ,
Ausschluss_Säule=„I“)

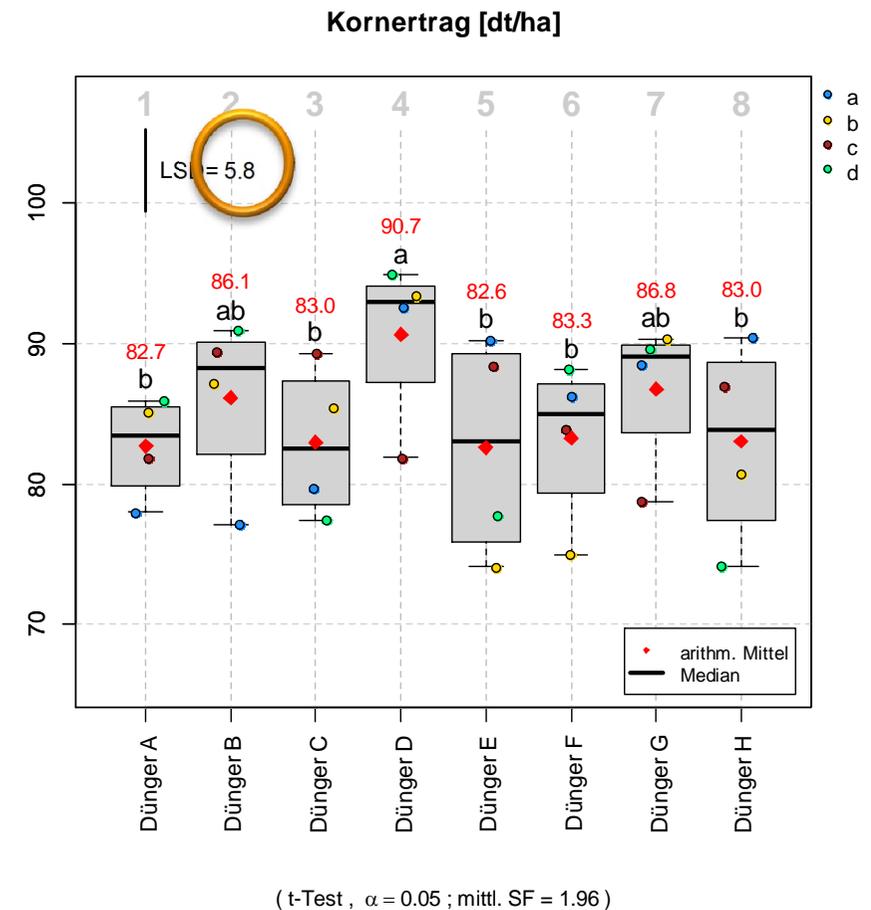


(t-Test , $\alpha = 0.05$; mittl. SF = 1.36)

Simulierte Daten: Unterschiedliches Erntedatum

FUN_Auswertung_FV (Daten. df = Bei spi el daten,
Zi el vari abl e = „Ertrag“)

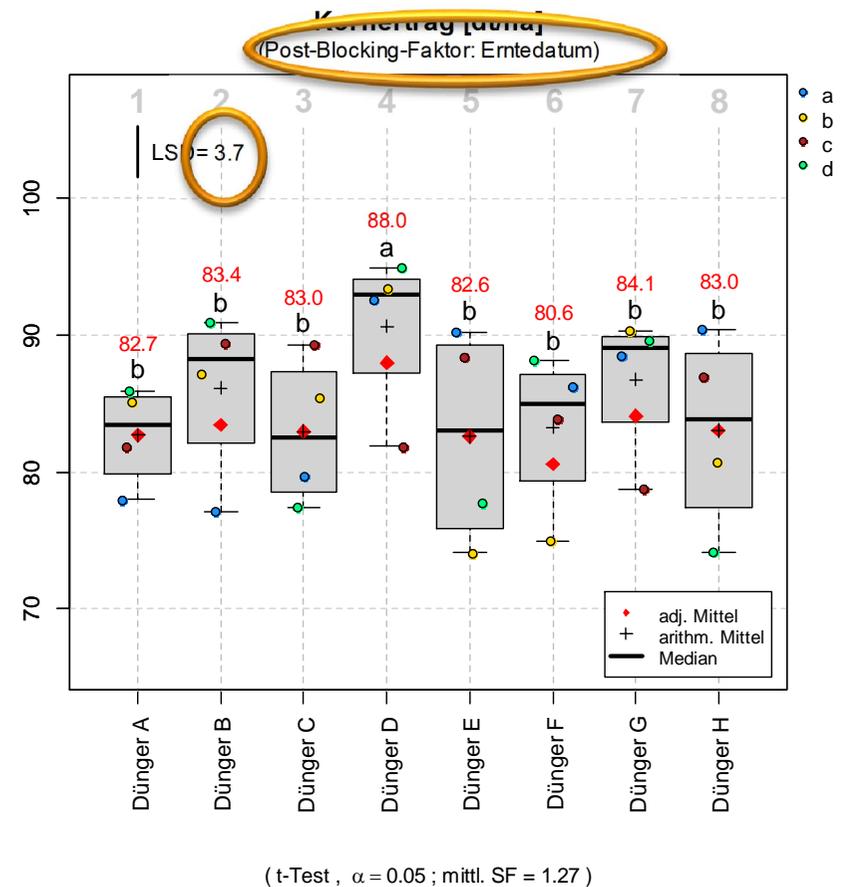
Pgl	Block	Säule	Erntedatum	Ertrag_alt	Ertrag
Dünger A	a	I	20	98,0	78,0
Dünger A	b	III	10	95,1	85,1
Dünger A	c	II	20	101,8	81,8
Dünger A	d	IV	10	95,9	85,9
Dünger B	a	I	20	97,1	77,1
Dünger B	b	IV	10	97,2	87,2
Dünger B	c	III	10	99,4	89,4
Dünger B	d	II	10	100,9	90,9
Dünger C	a	II	20	99,7	79,7



Simulierte Daten: Unterschiedliches Erntedatum

FUN_Auswertung_FV (Daten. df = Bei spi el daten,
 Zielvariab le = „Ertrag“,
 post_blocking = „Erntedatum“)

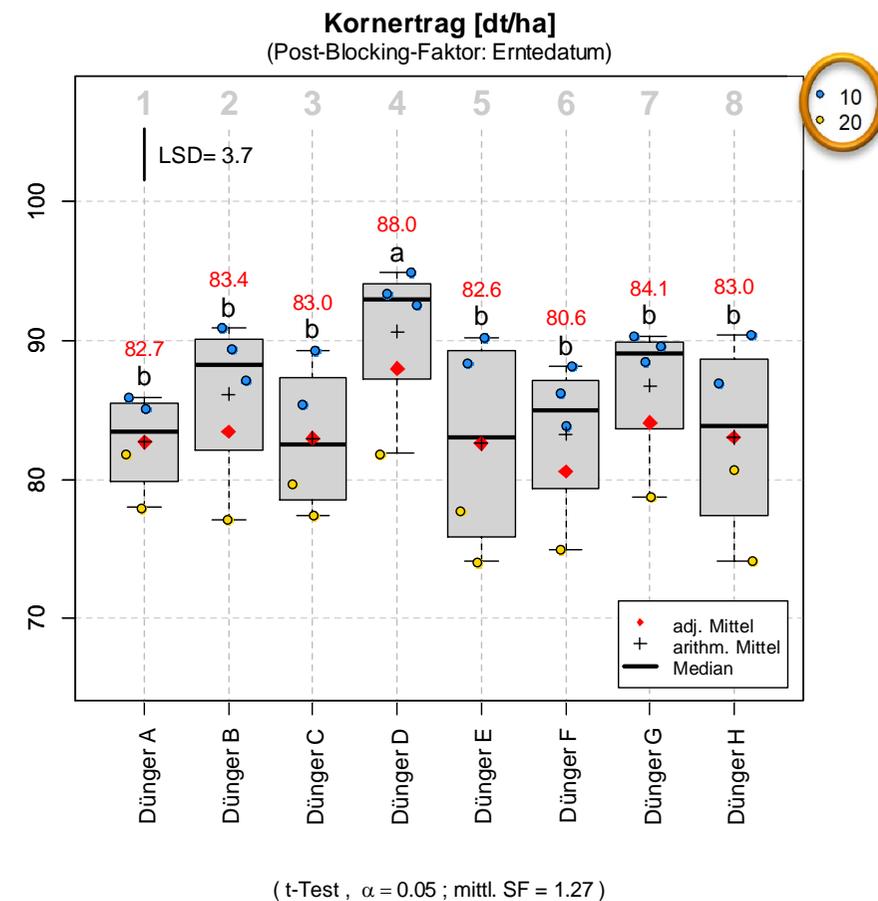
Pgl	Block	Säule	Erntedatum	Ertrag_alt	Ertrag
Dünger A	a	I	20	98,0	78,0
Dünger A	b	III	10	95,1	85,1
Dünger A	c	II	20	101,8	81,8
Dünger A	d	IV	10	95,9	85,9
Dünger B	a	I	20	97,1	77,1
Dünger B	b	IV	10	97,2	87,2
Dünger B	c	III	10	99,4	89,4
Dünger B	d	II	10	100,9	90,9
Dünger C	a	II	20	99,7	79,7



Simulierte Daten: Unterschiedliches Erntedatum

FUN_Auswertung_FV (Daten. df = Bei spi el daten,
 Ziel variable = „Ertrag“,
 post_blocking = „Erntedatum“,
 Faktor. Farbe= „Erntedatum“)

→ Wie ist die Verteilung der Residuen?

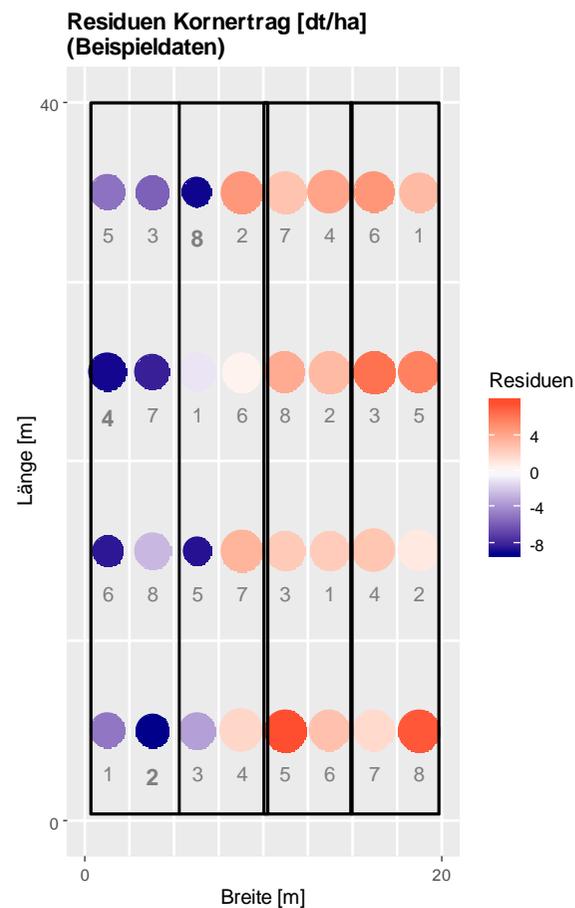


Datenstruktur für FUN_Residuen_FV

Pgl	Block	Säule	Erntedatum	Ertrag	Parzelle	Zeile	Spalte
Dünger A	a	I	20	78,0	1a	1	1
Dünger A	b	III	10	85,1	1b	2	6
Dünger A	c	II	20	81,8	1c	3	3
Dünger A	d	IV	10	85,9	1d	4	8
Dünger B	a	I	20	77,1	2a	1	2
Dünger B	b	IV	10	87,2	2b	2	8
Dünger B	c	III	10	89,4	2c	3	6
Dünger B	d	II	10	90,9	2d	4	4
Dünger C	a	II	20	79,7	3a	1	3
Dünger C	b	III	10	85,4	3b	2	5

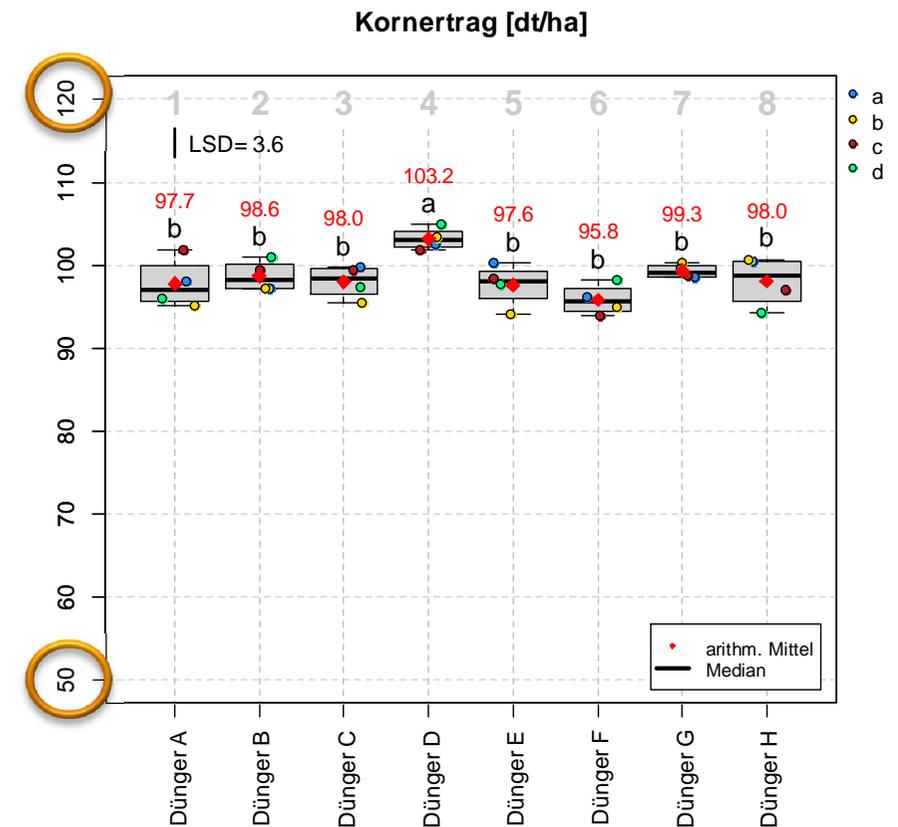
Simulierte Daten: Unterschiedliches Erntedatum

`FUN_Residuen_FV` (Daten. df = Beispiel daten,
Zielvariable = „Ertrag“,
Breite_Parz=2.5,
Länge_Parz=10)



Kosmetische Anpassungen:

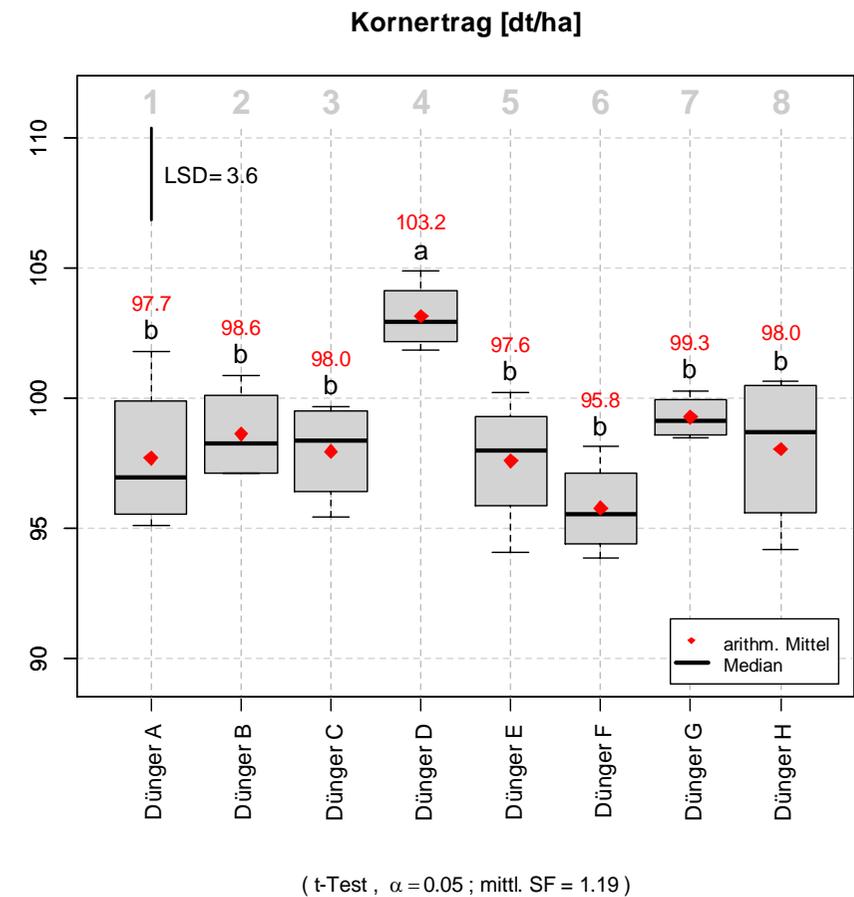
FUN_Auswertung_FV (Daten.df = Bei spi eldaten,
Zi elvari able = „Ertrag“,
ymi n= 50,
ymax= 120)



(t-Test , $\alpha = 0.05$; mittl. SF = 1.19)

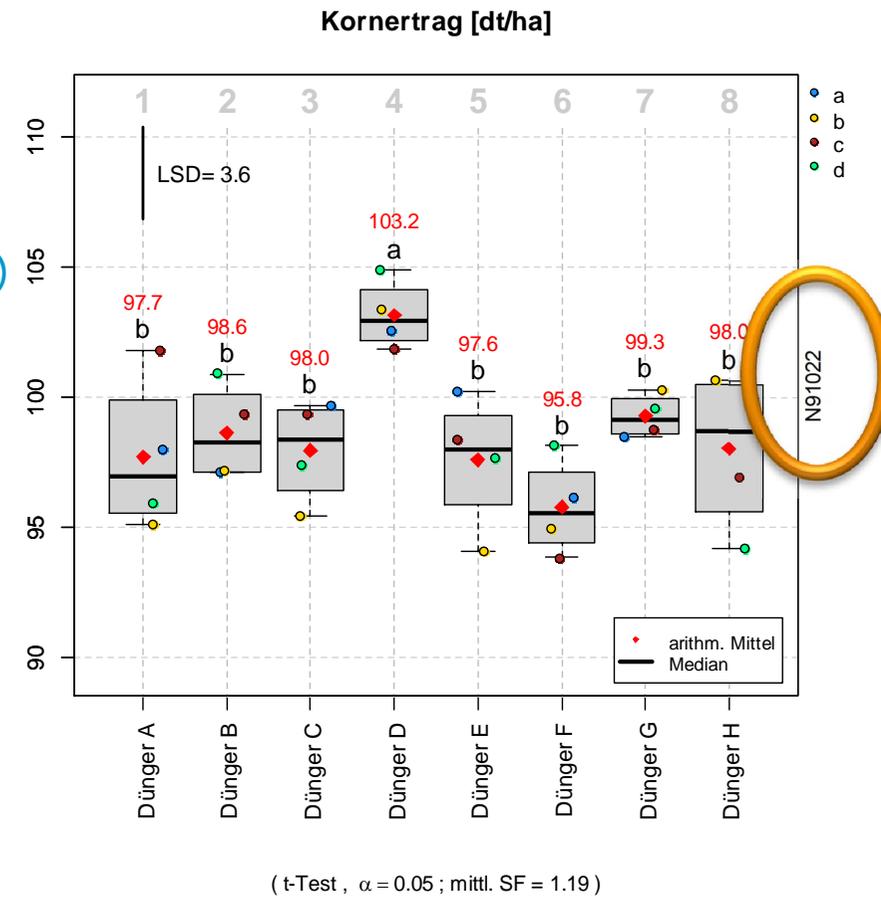
Kosmetische Anpassungen:

FUN_Auswertung_FV (Daten.df = Bei spi el daten,
Zi el vari abl e = „Ertrag“ ,
Ei nzel werte_zei gen=F)



Kosmetische Anpassungen:

FUN_Auswertung_FV (Daten.df = Bei spi el daten,
 Zi el variable = „Ertrag“,
 Versuchsbezei chnung=„N91022“)

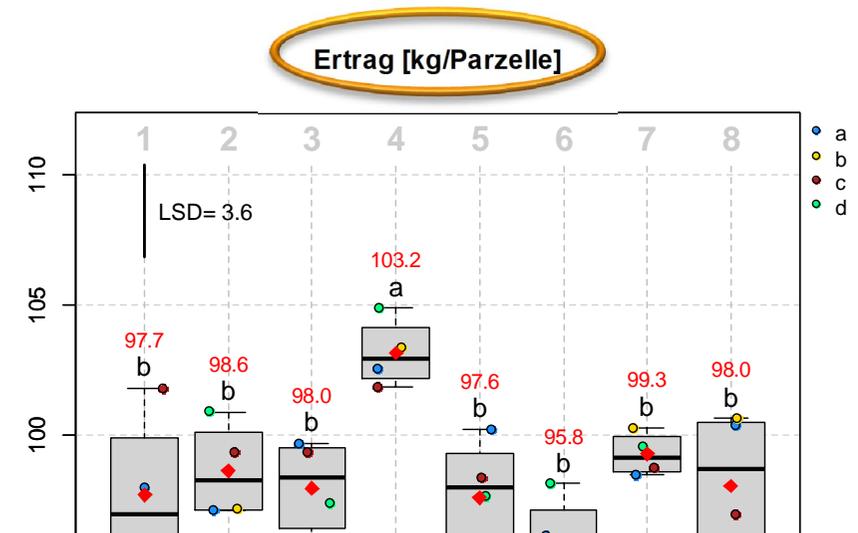


Automatische Erstellung der Überschrift

Bez_kurz	Bez_l ang
1 Ertrag	Kornertrag [dt/ha]
2 RP	Rohproteingehalt [%]
3 NEntz	N-Entzug [kg/ha]
4 SEntz	S-Entzug [kg/ha]
5 HL	Hektolitergewicht [kg/hl]
6 Sedi	Sedimentationswert [ccm]
7 ÄTH	Ährentragende Halme [1/m ²]
8 ÄTH_Hand	Ährentragende Halme [1/m ²] (Handzählung)
9 ÄTH_Drohne	Ährentragende Halme [1/m ²] (Drohne)
10 Pfl_m2	Anzahl Pflanzen pro m ²
11 TKM	Tausendkornmasse [g]
12 Ölgeh	Ölgehalt [%] (91% TS)
13 Öl ertr	Öl ertrag [dt/ha] (91% TS)
14 FM_Ertrag	Frischmasseertrag [dt/ha]
15 TM_Ertrag	Trockenmasseertrag [dt/ha]
16 ...	

Kosmetische Anpassungen:

FUN_Auswertung_FV (Daten. df = Beispiel daten,
Zielvariable = „Ertrag“,
Bezeichnung_Zielvariable = „Ertrag [kg/Parzelle]“)



- Schnelle und bequeme Vorauswertung von Versuchen, nur wenige Sek. pro Merkmal
- Grafiken ungestörter Versuche sind sofort „veröffentlichungsfähig“
- Visualisierung der Daten äußerst hilfreich zum Erkennen von Störungen
- Kombination mit räumlicher Residuenverteilung sinnvoll
- Verbesserungsoption: zufällige Effekte + automatische Modellreduktion?
- Versuche mit räumlichen Störungen, welche das einfache Modell nicht erfassen kann, müssen gesondert ausgewertet werden (PIAFStat).

Vielen Dank für die Aufmerksamkeit!

skw.
PIESTERITZ



Alle Optionen im Überblick

```
Ausschluss_Block="", # "a", "b", "c", "d" # Block bei Auswertung ausschließen
Ausschluss_Säule="", # "I", "II", "III", "IV" # Säule bei Auswertung ausschließen
Kovariablen="", # "[beliebiger Faktor aus Daten]" # bei Eignung des Faktors als Kovariable sie genutzt, bei Nichteignung Ausgabe Ursache
post_blockings="", # "[beliebiger Faktor aus Daten]" # bei Eignung des Faktors wird er genutzt, bei Nichteignung Ausgabe Ursache
Test_Typ="t-Test", # "t-Test", "Tukey-Test" # führt den entsprechenden Test durch und stellt die entspr. Grenzdifferenz dar
Beschriftung_Zielvariablen="auto", # "[beliebiger Text]", "auto" # "auto" erkennt die gängigen Zielvariablen und wählt die entsprechende Überschrift
Beschriftung="auto", # "gerade", "gedreht", "auto" # "auto" entscheidet automatisch, ob alles horizontal nebeneinander passt (klappt nicht immer)
Versuchsbeschriftung="", # "[beliebiger Text]" # fügt den Namen des Versuchs rechts an den Rand der Grafik
VI_zeigen=F, # T, F # Vertrauensintervalle einzeichnen
Boxplot_zeigen=T, # T, F # Boxplots zeichnen
Prüfgliednummern_zeigen=T, # T, F # Pgl -Nummern einfügen
Wdh_zeigen=F, # T, F # wenn r ungleich 4, wird "r=.." immer angezeigt; wenn r=4, nur dann, wenn Wdh_zeigen auf TRUE umgestellt ist
Einzelwerte_zeigen=T, # T, F # Einzelwerte einzeichnen
Einzelwerte_farbig=T, # T, F # wenn F, dann werden die Einzelpunkte zwar eingezeichnet, aber die Legende mit den Farben wird nicht erstellt
Faktor.Farbe="Block", # "[beliebiger Faktor aus dem Modell]" # in der Regel Block oder Säule, Einzelpunkte werden entsprechend eingefärbt
pos_Legende="bottomright", # "bottomright", "topright", "bottomleft", "topleft" # mögliche Positionen für Legende
ymax=100, # [beliebige Zahl] # obere Begrenzung der y-Achse selbst festlegen
ymin=50 # [beliebige Zahl] # untere Begrenzung der y-Achse selbst festlegen
```

Umgebung leeren (Verwendung „alter“ Daten ausschließen)

Daten einlesen

Daten prüfen:

- Spalten *Pgl* oder *Prüfglied* vorhanden?
- Spalten Block und Zielvariable vorhanden?
- Säule vorhanden (nicht erforderlich)?

Nutzereingabe prüfen:

Zu viele Einschränkungen (z. B. Ausschluss von Block und Säule)?

Spaltennamen der gewählten Modelle korrekt?

Datensubset entsprechend Nutzereingabe:

Block oder Säule ausschließen?

Anzahl Fehlwerte bestimmen

Auswertungsmodell wählen

- RCB
- Lat. Rechteck
- Post Blocking
- Kovarianz

- Durchführung der Auswertung mit emmeans entsprechend Modell
- Tabelle mit emmeans und Vertrauensintervallen
- Tabelle aller paarweisen Vergleiche → multcomp → Buchstabendarstellung
- Mittlere GD aller paarweisen Vergleiche
- Skalierung der Grafik
- Anpassung der Grafikränder entsprechend der Länge der Pgl-Bezeichnungen
- Erstellung Boxplot inkl. Überschrift
- Plot Vertrauensintervalle
- Plot Einzelwerte (mit Jitter)
- Plot Mittelwerte (arithmetisch und adjustiert)
- Beschriftung der Mittelwerte
- Signifikanzbuchstaben
- Einzeichnung GD
- Legende
- Dateninfo
- Bei Kovarianz und Post_Blocking: Test auf „Sinnhaftigkeit“
- Bei sign. WW oder nicht signifikanten Effekten → Schließen der Grafik, Fehlermeldung.

Hinweis zum Download



Die Funktion gibt es nicht online zum Download,
dennoch stelle ich sie gerne zur Verfügung.
Bei Interesse bitte ich um eine kurze Email an:

johannes.doehler@skwp.de