

RESI, SAS/AF-Anwendung zur Bewertung der partiellen Resistenz von Getreidesortimenten

Eckard Moll

Biologische Bundesanstalt für Land- und Forstwirtschaft
14532 Kleinmachnow
Tel. 033203-48-331
E-Mail: e.moll@bba.de

1. Inhalt und Struktur von RESI

RESI, Version 3.1, ist eine SAS/AF-Applikation unter SAS 6.12 zur Bewertung der partiellen Resistenz von Getreidesortimenten. Durch Anklicken der englischen Flagge kann vom deutschsprachigem Eröffnungsbildschirm (Abb. 1a) auf den englischsprachigen (Abb. 1b) umgeschaltet werden. Damit wird gleichzeitig ein Wechsel der Sprache für alle Ein- und Ausgaben vorgenommen.



Abb. 1a: Eröffnungsbildschirm (deutsch)



Abb. 1b: Eröffnungsbildschirm (englisch)

Das Programmsystem RESI ist in mehrere Teilkomplexe gegliedert (MOLL 2000, MOLL und FLATH 2001):



Schadbilder

Verschiedene Getreidekrankheiten sind visualisiert. Gezeigt werden auf der linken Seite Boniturhilfen, um den prozentualen Anteil befallener Blattfläche oder Ähre möglichst gut schätzen zu können, und rechts natürliche Befallssymptome. Am Beispiel vom Gelbrost (*Puccinia striiformis* Westend) sind in der Abb. 2 die Boniturhilfen und die natürlichen Befallssymptome vorgestellt.

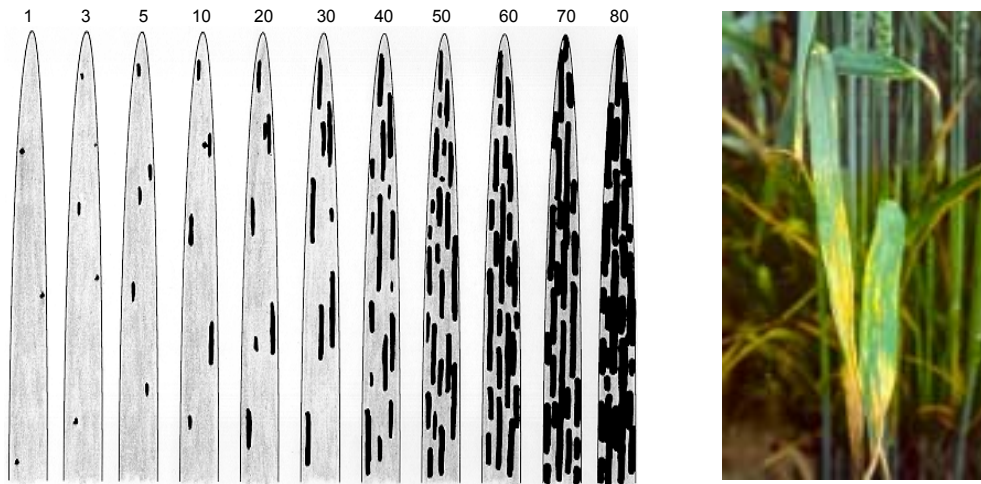


Abb. 2: Boniturhilfen und natürliche Befallssymptome für den Gelbrost



Konstruktion eines Lageplanes

Versuchsanlage ist eine einfaktorielle Blockanlage A-BI. Die zu prüfenden Getreidesorten liegen in einer Textdatei vor. Unter Hinzuziehung von Standards, resistenter, lokaler und/oder anfälliger Standard, wird ein randomisierter Lageplan konstruiert, wobei die Standards in jedem Block entsprechend der Vorgabe häufiger wiederholt werden als die zu prüfenden Sorten. Neben dem Lageplan, der auf Datei ausgegeben wird, wird eine weitere Datei angelegt. Diese enthält alle Prüfglieder und Wiederholungen. Werden zu diesen Informationen die geschätzten Befallswerte zu den einzelnen Boniturterminen hinzugefügt, entsteht die Datendatei.



Auswertung und Resistenzeinschätzung

- Auswertung eines Einzelversuches

Für jede Getreidesorte (einschließlich der Standards) wird für jedes Teilstück aus dem Befallsverlauf ein mittlerer Befall je Teilstück berechnet. Ausgegeben werden der mittlere Befall jeder Sorte und orientierend eine ganzzahlige Boniturnote. Die statistischen Analysen basieren auf den mittleren Befallswerten der Teilstücke. Diese berechnen sich aus in möglichst wöchentlichem Abstand aufeinander folgende Befallsschätzungen (Abb. 3, nach WALTHER u.a. 2000).

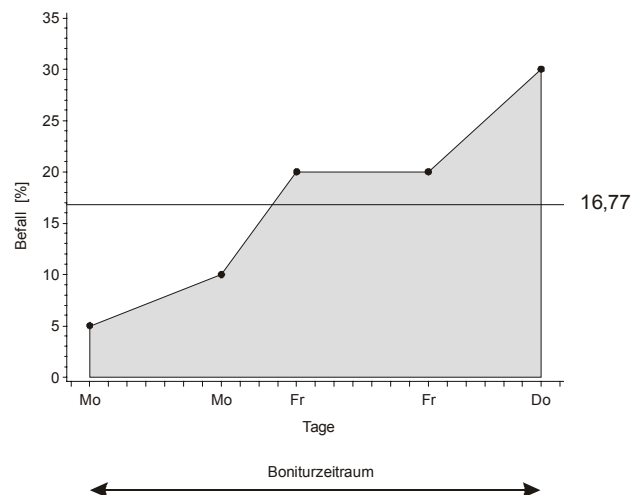


Abb. 3: Fläche unter der Befallsverlaufskurve

Die Fläche unter der Befallsverlaufskurve je Teilstück entspricht dem mittleren Befall je Teilstück (Abb. 3). Für das in der Abb. 3 dargestellte Beispiel ist

$$\text{mittlerer Befall}_{\text{Teilstück}} = \left[\frac{5+10}{2} * 7 + \frac{10+20}{2} * 4 + \frac{20+20}{2} * 7 + \frac{20+30}{2} * 6 \right] / (7+4+7+6) = 16,77$$

Mit diesen für jedes Teilstück vorliegenden mittleren Befallswerten wird die Varianzanalyse für eine einfaktorielle randomisierte Blockanlage A-BI unter Einbeziehung optionaler multipler Mittelwertvergleiche Tukey-Prozedur, t-Test gegen die Standards, Dunnett-Prozedur gegen die Standards und Maximum-Modulus-Prozedur gerechnet.

Alle Ergebnisse werden in Textdateien gespeichert. Des weiteren wird eine Textdatei angelegt, die die mittleren Befallswerte je Teilstück, also die Einzelwerte, enthält. Dieses ist die Daten-Datei für eine Serienanalyse.

- Auswertung einer Versuchsserie

Mehrere solcher Daten-Dateien, die als Ergebnis der varianzanalytischen Auswertung eines Einzelversuches gebildet wurden, werden zur Serienauswertung von Versuchen an verschiedenen Versuchsstandorten und in verschiedenen Versuchsjahren herangezogen. Modellabhängig, d. h. Orte fix oder zufällig und Jahre fix oder zufällig, werden die Einzeldaten der Versuche varianzanalytisch ausgewertet, wobei insbesondere bei den optionalen multiplen Tests zum Vergleich der zu prüfenden Getreidesorten die Balanciertheit der Versuchsserie und ihre Orthogonalität in den Versuchsorten und -jahren berücksichtigt wird. Bei der Durchführung der multiplen Tests und der Interpretation der Ergebnisse wird auch berücksichtigt, ob und welche Wechselwirkungen der fixen Effekte signifikant sind. Für Informationen über den biometrischen Hintergrund, der hier nicht betrachtet werden kann und soll, wird auf MOLL u.a. (2000) verwiesen.

Für den Nutzer von RESI ist es wichtig zu wissen, dass die Serienanalyse verhältnismäßig bis unverhältnismäßig lange dauern kann, weil der in der SAS-Prozedur MIXED verwendete Algorithmus zur Schätzung der Varianzkomponenten mit Hilfe der Restricted Maximum Likelihood (REML) Methode sehr zeitraubend ist. Für eine nichtorthogonale Versuchsserie muss man unter Umständen eine Stunde und mehr Zeit einplanen.

2. Zur Realisierung multipler Mittelwertvergleiche für die Analyse von Versuchsserien mit SAS 6.12

Die im Programm 1 (s. u.) betrachteten Programmzeilen entstammen dem SAS-Programmsystem RESI (MOLL 2000) zur Auswertung von Versuchsserien. Hauptaugenmerk liegt auf der Realisierung der multiplen Mittelwertvergleiche, wobei davon ausgegangen wird, dass die beiden Faktoren Orte und Jahre jeweils mehr als eine Stufe haben. Die Macrovariable NOBO steht für die programmgesteuerte Verwendung der Option NOBOUND, die im balancierten Fall genutzt wird. SALPHA beinhaltet das Signifikanzniveau für die Versuchsserie. Die Macrovariablen SMODEL und SRAND definieren das Modell der Versuchsserie.

Die ersten vier LSMEANS-Anweisungen realisieren alle Vergleiche für die Tukey-Prozedur. Die nächsten drei LSMEANS-Anweisungen beziehen sich auf die Bonferroni-Prozedur, die in bestimmten Fällen (gemischtes Modell, Sorten nicht orthogonal und Wahl eines multiplen Mittelwertvergleichs mit versuchsbezogenem Risiko 1. Art) „zusätzlich“ herangezogen wird. Die Vergleiche `sorten*orte*jahre` fehlen hier, da bei fixen Wechselwirkungen `Sorten*Orte*Jahre` kein gemischtes Modell vorliegt und folglich die Bonferroni-Prozedur nicht zur Anwendung kommt.

Die Dunnett-Prozedur zum Vergleich der Effekte des Prüffaktors ist gegen einen bestimmten Standard, beispielsweise den resistenten Standard `=RR=`, aufgeführt. Die Vergleiche `sorten*orte`, `sorten*jahre` und `sorten*orte*jahre` für die Standards verbergen sich hinter den mit `%include` beginnenden Anweisungen. Es bedarf einiger Hilfsmittel und Kenntnisse zu `%include` und `put`-Anweisungen (Wolf F. LESENER persönl. Mitteilung; SAS Institute 1992), um die entsprechenden LSMEANS-Anweisungen schreiben zu können.

Das Programm 1 zeigt, dass alle möglichen Tests durchgeführt werden, von denen aber nur die Ergebnisse der vom Nutzer ausgewählten, sachlogisch zutreffenden und überhaupt durchführbaren Tests präsentiert werden. Mit Hilfe der MAKE-Anweisungen im Programm 1 werden die Ergebnisse nicht sofort in das Output-Fenster (NOPRINT) sondern auf SAS-Dateien ausgegeben. Die Datei `test_m` nimmt die Ergebnisse der Varianztabelle der fixen Effekte, `lsmean_m` die Mittelwertschätzungen und `diff_m` die Ergebnisse der multiplen Vergleiche auf.

Programm 1: Varianzanalyse und multiple Mittelwertprozeduren mit der Prozedur MIXED

```
proc mixed data=gesamt &NOBO;
  class sorten orte jahre blocks;
  model merkmal = &SMODEL / ddfm=satterth;
  random &SRAND;
  lsmeans sorten / adjust=tukey alpha=&SALPHA;
  lsmeans sorten*orte / adjust=tukey alpha=&SALPHA;
  lsmeans sorten*jahre / adjust=tukey alpha=&SALPHA;
  lsmeans sorten*orte*jahre / adjust=tukey alpha=&SALPHA;
  lsmeans sorten / adjust=bon alpha=&SALPHA;
  lsmeans sorten*orte / adjust=bon alpha=&SALPHA;
  lsmeans sorten*jahre / adjust=bon alpha=&SALPHA;
  lsmeans sorten / adjust=dunnett diff=control('=RR=') alpha=&SALPHA;
  ...
  %include lsmean1/source2;
  %include lsmean2/source2;
  %include lsmean3/source2;

  make 'ClassLevels' out = _null_ noprint;
  make 'CovParms' out = _null_ noprint;
  make 'Parms' out = _null_ noprint;
  make 'Fitting' out = _null_ noprint;
  make 'REML' out = _null_ noprint;
  make 'tests' out = test_m noprint;
  make 'LSMeans' out = lsmean_m noprint;
  make 'Diffs' out = diff_m noprint;
run;
```

Im Programm 2 wird für einen bestimmten Standard - beispielsweise =RR= - und die Vergleiche der Mittelwerte `sorten*orte` gezeigt, wie die LSMEANS-Anweisungen zusammengestellt werden können. Die SAS-Datei `orte_all` enthält alle Orte. In die Datei `lsmean1` werden Zeilen geschrieben wie sie im Programm 3 aufgeführt sind. Die `%include`-Anweisung (Programm 1) fügt diese Programmzeilen in PROC MIXED ein und führt sie aus.

Darauf hingewiesen werden soll, dass PROC MIXED unter SAS/AF nur grundsätzlich so reagiert wie unter SAS/STAT – zumindest mit SAS 6.12. Beim Übergang auf SAS 8.1 führt die Option NOPRINT in der MAKE-Anweisung zu einem Fehler und zum Abbruch der Abarbeitung, da sie nicht mehr unterstützt wird. Es müssen dann die ODS-Anweisungen herangezogen werden.

In SAS 6.12 wurden LSMEANS-Anweisungen für zufällige Effekte völlig richtig mit einer Fehlerausschrift im Log-Fenster bedacht und ihre Ausführung übergangen. In SAS 8.1 erfolgt auch eine Fehlermitteilung, aber nun wird die Abarbeitung der Prozedur abgebrochen. Jetzt müssen die LSMEANS-Anweisungen hinsichtlich zufälliger Effekte vorher dahingehend getestet werden, dass sie nicht in die Prozedur übernommen werden. Das könnten entsprechend belegte `%include`-Anweisungen realisieren. Die ODS-Befehle übernehmen die Zielstellung der MAKE-Anweisungen. Das Prozedur hätte dann unter SAS 8.1 die im Programm 4 vorgestellte Form.

Programm 2: Zusammenstellen der LSMEANS-Anweisungen für die Dunnett-Prozedur

```
filename lsmean1 'c:\temp1.sas';
filename lsmean2 'c:\temp2.sas';
filename lsmean3 'c:\temp3.sas';

data null ;
  vgl = 'xxxxxxxx';
  file lsmean1;
  vgl = '=RR=';
  set orte_all;
  put "lsmeans sorten*orte/adjust=dunnett diff=control("
      vgl
      " " "
      orte $ 12.
      " ");";
  ...
run;
```

Programm 3: Einzufügende LSMEANS-Anweisungen

```
lsmeans sorten*orte/adjust=dunnett diff=control('=RR=' 'ORT1');
lsmeans sorten*orte/adjust=dunnett diff=control('=RR=' 'ORT2');
lsmeans sorten*orte/adjust=dunnett diff=control('=RR=' 'ORT3');
...
```

Programm 4: PROC MIXED unter SAS 8.1

```
ods listing exclude
  covparms fitstatistics tests3 iterhistory modelinfo lrt
  convergencestatus classlevels dimensions lsmeans diffs;
ods output tests3=test_m
  lsmeans=lsmean_m
  diffs=diff_m;
proc mixed data=gesamt &NOBO;
  class sorten orte jahre blocks;
  model befall = &SMODEL / ddfm=satterth;
  random &SRAND;
  lsmeans sorten / adjust=tukey alpha=&SALPHA;
  %include tukeyso/source2;
  %include tukeysj/source2;
  %include tukeysoj/source2;
  lsmeans sorten / adjust=bon alpha=&SALPHA;
  %include bon_so/source2;
  %include bon_sj/source2;
  lsmeans sorten / adjust=dunnett diff=control('=RR=') alpha=&SALPHA;
  lsmeans sorten / adjust=dunnett diff=control('=LL=') alpha=&SALPHA;
  lsmeans sorten / adjust=dunnett diff=control('=SS=') alpha=&SALPHA;
  lsmeans sorten / adjust=dunnett diff=control("&svgl") alpha=&SALPHA;
  %include lsmean1/source2;
  %include lsmean2/source2;
  %include lsmean3/source2;
run;
```

3. Zur Berechnung der Grenzdifferenzen und der Überschreitungswahrscheinlichkeiten der Tukey- und der Dunnett-Prozedur

Die Grenzdifferenz der Tukey-Prozedur zum Vergleich der Mittelwerte des Prüffaktors A (Sorten) ist bekanntlich $HSD = q_{1-\alpha; FG_{Rest}, a} * s_{\bar{q}} / \sqrt{2}$, wenn keine Wechselwirkungen signifikant sind. Dabei sind q das $(1-\alpha)$ -Quantil der studentisierten Spannweiten-Verteilung und $s_{\bar{q}}$ der Standardfehler der Mittelwertdifferenzen. Mit der Berücksichtigung signifikanter Wechselwirkungen ändert sich der Test und damit auch die Berechnung der zutreffenden Grenzdifferenz.

SAS berücksichtigt beim Vergleich der beispielsweise Sorten*Orte-Mittelwerte, ob der Vergleich auf derselben Stufe des Prüffaktors Sorten bzw. des Faktors Orte vorgenommen wird und verwendet die dafür die dem Modell entsprechenden Freiheitsgrade ($_{DF}$) und Standardfehler der Mittelwertdifferenzen ($_{SE}$), die auch in die SAS-Datei `diff_m` (Programm 1) geschrieben werden. Für die Berechnung der Überschreitungswahrscheinlichkeit ($_{ADJP}$) verwendet SAS die bei der Tukey-Prozedur zugrunde gelegte Verteilung der Mittelwertdifferenzen q unter Berücksichtigung der Anzahl Faktorkombinationen, also beispielsweise der Zahl der Sorten*Orte-Mittelwerte. In der Tabelle 1 ist die Anzahl der Faktorkombinationen durch die Produkte a_j , a_o , a_{oj} und a_u ausgewiesen, wobei a für die Anzahl der Stufen des Prüffaktors A, j für die Anzahl Jahre, o für die Anzahl Orte und u für die Anzahl Umwelten stehen.

Tabelle 1: Berechnung der Grenzdifferenzen bei signifikanten Wechselwirkungen für die Tukey-Prozedur

Vergleich der A-Mittelwerte bei signifikanter Wechselwirkung	Grenzdifferenz der Tukey-Prozedur	q-Quantil (SAS)	q-Quantil (RESI)
AxJ	$HSD^{AJ/J} = q_{1-\alpha; FG_{Rest}, a} \cdot s_{\bar{d}_{AJ/J}} / \sqrt{2}$	$1 - p_{q_{1-\alpha; FG_{Rest}, aj}}$	$1 - p_{q_{1-\alpha; FG_{Rest}, a}}$
AxO	$HSD^{AO/O} = q_{1-\alpha; FG_{Rest}, a} \cdot s_{\bar{d}_{AO/O}} / \sqrt{2}$	$1 - p_{q_{1-\alpha; FG_{Rest}, ao}}$	$1 - p_{q_{1-\alpha; FG_{Rest}, a}}$
AxOxJ	$HSD^{AOJ/OJ} = q_{1-\alpha; FG_{Rest}, a} \cdot s_{\bar{d}_{AOJ/OJ}} / \sqrt{2}$	$1 - p_{q_{1-\alpha; FG_{Rest}, aoj}}$	$1 - p_{q_{1-\alpha; FG_{Rest}, a}}$
AxU	$HSD^{AU/U} = q_{1-\alpha; FG_{Rest}, a} \cdot s_{\bar{d}_{AU/U}} / \sqrt{2}$	$1 - p_{q_{1-\alpha; FG_{Rest}, au}}$	$1 - p_{q_{1-\alpha; FG_{Rest}, a}}$

Wenn für den Vergleich der Sortenmittelwerte die Wechselwirkung von Sorten*Orte signifikant ist; die Dreifachwechselwirkung sei es nicht, dann werden die Sorten*Orte-Mittelwerte auf jeweils derselben Stufe des Ortes verglichen. Das bedeutet, dass auf jeder Stufe der Orte a Sorten-Mittelwerte in Verbindung mit der Sorten*Orte-Wirkung miteinander verglichen werden. Dieses berücksichtigt das zur Berechnung der Grenzdifferenz verwendete q-Quantil (Tab. 1). SAS legt beim Vergleich der beispielsweise Sorten*Orte-Mittelwerte auf jeweils derselben Stufe des Ortes die Anzahl von ao zu vergleichenden Mittelwerten zugrunde. Damit zeigt SAS für diese Fälle ein konservatives Testverhalten, d. h. die Nullhypothese wird länger beibehalten. Das Programm 5 beinhaltet die Berechnung der Grenzdifferenz der Tukey-Prozedur (hsd) und der Überschreitungswahrscheinlichkeit (prob) bei signifikanter fixer Wechselwirkung Sorten*Orte. Die Macrovariable a gibt die Anzahl der auf jeder Stufe der Orte zu vergleichenden Mittelwerte des Prüffaktors A (Sorten) an.

Programm 5: Berechnung der Grenzdifferenz und der Überschreitungswahrscheinlichkeit

```

data tuk;
  set diff_m
    (keep=_effect_ sorten_orte _sorten_orte
      _df_ _se_ _diff_ _adjp_ _adjust_
      where= (_diff_ ^= .));
  if _adjust_ =: "Tukey";
  if _effect_ = "SORTEN*ORTE";
  if orte = _orte;
  . . .
  q = abs(_diff_)*sqrt(2)/_se_;
  hsd = probmc("RANGE", ., 1-&salph, _df_, &a)*_se_/sqrt(2);
  prob = 1-probmc("RANGE", q, ., _df_, &a);
  . . .
run;

```

Somit wird in RESI die bisher übliche Form der Berechnung der Grenzdifferenz bei signifikanter Wechselwirkung realisiert. Die Programme können auch relativ unproblematisch an eine neue Entwicklung angepasst werden, sei es die „automatische“ Übernahme der SAS-Vorgehensweise und damit ein konservatives Testen oder beispielsweise eine weitere Korrektur, etwa in Richtung einer geeigneten Aufteilung der Irrtumswahrscheinlichkeit α auf jede Stufe.

Literatur

- MOLL, E. (2000): RESI - Bewertung der partiellen Resistenz von Getreidesortimenten - Eine SAS Anwendung - Version 3.1, Mitteilungen aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft, Heft 374, S. 35-61
- MOLL, E. und FLATH, K. (2001): Die Bewertung von partieller Resistenz mit Hilfe der SAS-Applikation RESI Zeitschrift für Agrarinformatik, im Druck
- MOLL, E., PIEPHO, H.-P. und KRÜGER, F. (2000): Grundlagen der statistischen Auswertung einer Versuchsserie zur Bewertung der partiellen Resistenz von Getreidesortimenten Mitteilungen aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft, Heft 374, S. 26-34
- SAS Institute (1992): SAS® Language: Reference, Version 6, First Edition, SAS Institute Inc., Cary
- WALTHER, U., FLATH, K., MOLL, E., Prochnow, J. und SACHS, E. (2000): Methodische Anleitung zur Bewertung der partiellen Resistenz von Sorten bzw. Linien unter Berücksichtigung epidemiologischer Aspekte Mitteilungen aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft, Heft 374, S. 8-25