

Tipps zur SAS-Programmierung in der PIAFStat-Umgebung

Andrea Zenk
Landesforschungsanstalt für
Landwirtschaft und Fischerei MV
Dorfplatz 1
18276 Gülzow
a.zenk@lfa.mvnet.de

Volker Michel
Landesforschungsanstalt für
Landwirtschaft und Fischerei MV
Dorfplatz 1
18276 Gülzow
v.michel@lfa.mvnet.de

Jörg Schmidtke
BioMath GmbH
Schnickmannstraße 4
18055 Rostock
joerg.schmidtke@biomath.de

Zusammenfassung

PIAFStat als ein Modul des PIAF-Systems verwaltet in einer Verfahrensbibliothek mit den PIAFStat-Verfahren SAS-Code für die Datenanalyse und biostatistische Auswertung von landwirtschaftlichen Feldversuchen und Versuchsserien [1]. Zur Erstellung von PIAFStat-Verfahren sind umfassende Kenntnisse sowohl in der PIAFStat-Syntax als auch in der SAS-Welt notwendig. In diesem Beitrag wird über Erfahrungen zur SAS-Programmierung für PIAFStat berichtet. Es wird erläutert, wie ein sinnvoller Grad von Automatismus und Flexibilität im Verfahren erreicht werden kann. Daneben werden Tipps zu SAS direkt gegeben. Zur Veranschaulichung dieser Tipps wird ein aktuelles PIAFStat-Verfahren genutzt [8].

Schlüsselwörter: PIAF, PIAFStat, SAS-Programmierung

1 Einleitung

In der Landesforschungsanstalt für Landwirtschaft und Fischerei Mecklenburg-Vorpommern wird seit 2001 erfolgreich das PIAF-System zur Erfassung, Berechnung und Archivierung aller Feldversuchsdaten eingesetzt [2]. Ein Teil des PIAF-Systems ist die Software PIAFStat, die in einer Verfahrensbibliothek SAS-Code für die Datenanalyse und biostatistische Auswertung von landwirtschaftlichen Feldversuchen und Versuchsserien verwaltet und den SAS-RUN steuert [6].

In PIAFStat kann man von zwei verschiedenen Nutzerprofilen sprechen. Als Hauptnutzer werden die Anwender verstanden, die ohne SAS-Kenntnisse Standardverfahren zur Versuchsauswertung einsetzen. Daneben gibt es einen zweiten Nutzerkreis, der die Standardverfahren von PIAFStat für die Hauptnutzer entwickelt. Im vorliegenden Beitrag werden für diesen zweiten kleineren Nutzerkreis Erfahrungen zur SAS- Programmierung in PIAFStat weitergegeben.

2 Informationen zu PIAFStat

Das Programm PIAFStat ist ein Softwareprodukt der Firma BioMath GmbH, deren erste Version 2001 entstand. Finanziert von der Bund-Länder-Gemeinschaft, wurde und wird PIAFStat von einer Gruppe von Fachleuten des Sortenwesens der Länder konzeptionell mitentwickelt und begleitet. Warum war diese Software PIAFStat notwendig? Nach der Einführung des PIAF-Projektes (**P**lanung **I**nformation und **A**uswertung von **F**eldversuchen) in die alltäglichen Praxis des Sortenversuchswesens der Länder [2] wurden folgende Forderungen gestellt:

Für die statistischen Analysen im Feldversuchswesen wollte man das SAS-System nutzen. Es wurde geplant, gemeinsame Verfahren für alle PIAF-Nutzer zu erstellen, die optionale individuelle Möglichkeiten enthalten. Insbesondere sollten auch SAS-Unkundige die Verfahren anwenden können [7].

Mit der Entwicklung der Software PIAFStat ist es gelungen, diese Forderungen zu erfüllen. *Durch Demonstration des Verfahrens DEMO (einfaktorielle Blockanlage – Beispiel KSFE 2012) wird PIAFStat vorgestellt und gezeigt, dass die Forderungen erfüllt wurden.*

Die aktuelle Version PIAFStat 6.34 präsentiert sich als einfaches, anwenderfreundliches Programm. Es wird das unkomplizierte Starten von SAS-Programmen im Batch-Modus mit automatischer Bindung an Daten einer PIAF-Schnittstelle ermöglicht. Voraussetzung dafür ist jedoch, dass in der Verfahrensbibliothek von PIAFStat die notwendigen SAS-Programme enthalten sind (Abb. 1). Diese SAS-Programme sind nicht Bestandteil der Software, sondern werden von PIAFStat-Nutzern, die SAS-Kenntnisse besitzen, in PIAFStat erstellt und allen Anwendern zur Verfügung gestellt.

Sort	Name	Label	Info-Datei	Version	Autor	D
0001	1	----- EINZELVERSUCH -----				1
0001	A-BL	einfaktorielle Blockanlage (vorrangig aml. Mittelprüfung)	A_BI.pdf	221	Michel / Zenk	2
0001	A-BL1	einfaktorielle Blockanlage - Beispiel KSFE 2012	A_BL007.pdf	553	Michel / Zenk	2
0010	A/B BZW. B/A	zweifaktorielle Spaltanlage Standard M-V	A_B_SPA.pdf	75	Michel / Zenk	1
2000	2	----- S E R I E N -----				1
2005	PHI	Hohenheim-Gülzower Serienauswertung - optim. Datentransformation	Info_phi.pdf		Michel / Zenk / Möhring / Pieph	1
2010	VK	Hohenheim-Gülzower Serienauswertung - Varianzkomponenten	Info_vk.pdf	69	Michel / Zenk / Möhring / Pieph	1
2020	MW	Hohenheim-Gülzower Serienauswertung - Bestimmung der Mittelwerte	Info_mw.pdf	221	Michel / Zenk / Möhring / Pieph	1

Abbildung 1: Auszug aus einer Verfahrensbibliothek in PIAFStat

So wurden von Mitarbeitern der Landesforschungsanstalt für Landwirtschaft und Fischerei Mecklenburg-Vorpommern Standardverfahren für die Auswertung von Einzelversuchen und Serien entwickelt, die von vielen Kollegen im Versuchswesen der Bundesländer routinemäßig angewendet werden.

3 Aufbau eines PIAFStat-Verfahrens

Ein PIAFStat-Verfahren ist in zwei Bereiche gegliedert.

Im Bereich DEKLARATION werden PIAFStat-typische syntaktische Elemente als Platzhalter definiert:

- K – Klassifikations-Merkmale (K-Merkmale)
- A – Analyse-Merkmale (A-Merkmale)
- O – Optionen
- P – Funktion
- B – Block
- L – Layout.

Herzstück der PIAFStat-Elemente sind die Optionen. Mit den Optionen kann ein Verfahren so gestaltet werden, dass der Anwender in den Ablauf des SAS-Programms interaktiv eingreifen kann. Mit den Platzhaltern K- und A-Merkmale können strukturell festgelegte ASCII-Daten (ADS-Schnittstelle von PIAF) eingebunden werden.

Die weiteren Platzhalter Funktion und Block beinhalten Teile von SAS-Code, die an verschiedenen Stellen unterschiedlich eingebunden werden [6].

Im Bereich PROGRAMM folgt der eigentliche SAS-Code. Im SAS-Code werden die definierten Platzhalter so positioniert, dass bei Start des Verfahrens ein korrektes SAS-Programm entsteht.

4 Ablauf in PIAFStat

Durch den Befehl ‚SAS ausführen‘ wird in PIAFStat ein Verfahren gestartet. Es öffnet sich ein Fenster mit drei Reitern (Abb.2). Hier können die im Deklarationsteil definierten Elemente K-Merkmale, A-Merkmale und Optionen interaktiv in ihren Einstellungen verändert werden. Durch Setzen der Haken wird z.B. im Bereich der Optionen der Wechsel von OFF auf ON vollzogen, es wird die entsprechende Option aktiviert. Dadurch ermöglicht PIAFStat es, ein flexibles Verfahren zu erstellen.

Nach dem OK erfolgt durch PIAFStat die Generierung des SAS-Codes. Dieser SAS-Code wird in einer Datei PIAFSTAT.SAS abgelegt und besteht aus mehreren Abschnitten. Im ersten Abschnitt wird eine Reihe von globalen Makrovariablen (Tab. 1) definiert.

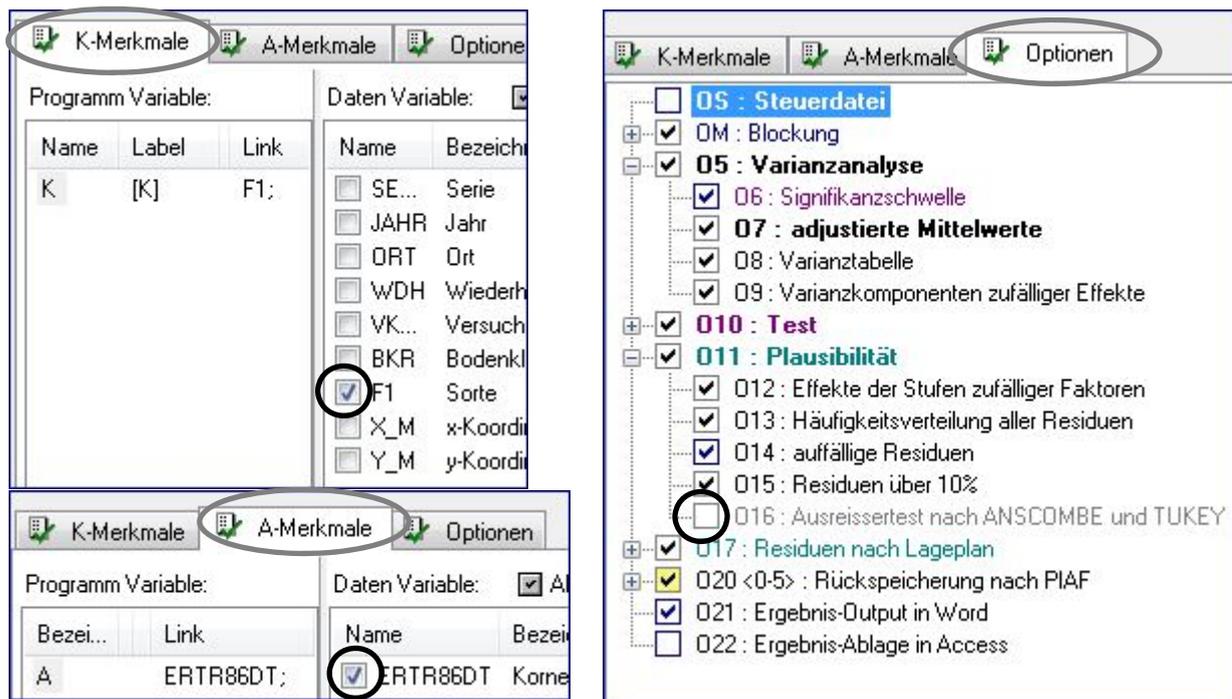


Abbildung 2: SAS ausführen in PIAFStat

Tabelle 1: Auswahl der von PIAFStat gebildeten globalen Makrovariablen

Variable	Bemerkung
VERZRES	Verzeichnis zur Ergebnis-Abspeicherung
VERZTEMP	temporäres Verzeichnis
VERFAHRENNAME	Name des PIAFStat-Verfahrens
VERFAHRENAUTOR	Autor des PIAFStat-Verfahrens
VJAHR	Versuchsjahr
VORTBEZ	Versuchsort
VWDH	Anzahl der Wiederholungen

Wird eine PIAF-Schnittstelle mit ASCII-Daten als Datenquelle genutzt, dann generiert PIAFStat SAS-Datasets (Tab. 2) mit vordefinierten Bezeichnungen.

Im Anschluss an die Bildung der globalen Variablen und – gegebenenfalls – der Datasets folgt das eigentliche SAS-Programm mit den entsprechend den Vorgaben übersetzten Platzhaltern. SAS wird im Batch-Modus gestartet und arbeitet den Code ab.

Tabelle 2: Aus ASCII-Daten gebildete SAS-Datasets

Dataset-Name	Inhalt
Vers	Beschreibung der Versuche
Fak	Faktoren und Faktorstufen der Versuche
Merk	Merkmale der Versuche
Daten	Daten der Versuche
Datense	SE-Daten der Versuche
Grenzd	Grenzdifferenzen der Versuche

5 Tipps für die Verfahrenserstellung

Es ist vorteilhaft, zum Programmieren, Korrigieren und Weiterentwickeln von PIAFStat-Verfahren den Enhanced Editor in SAS zu nutzen. Durch Laden der Datei PIAFSTAT.SAS kann das Verfahren in SAS direkt bearbeitet werden. In PIAFStat selbst steht dafür nur ein einfacher Editor zur Verfügung. Die Datei PIAFSTAT.SAS wird in einem temporären Verzeichnis abgelegt, sobald das Verfahren einmal in PIAFStat gestartet worden ist. Im SAS-Code sind die von PIAFStat automatisch durchgeführten Routinen zur Bildung der Makrovariablen und Datasets ersichtlich. Das Programm selbst folgt nach dem Kommentar `/*Verfahrensbibliothek*/`. Die Ersetzungen der PIAFStat-Platzhalter durch PIAFStat wurden im SAS-Code hier bereits vorgenommen. Das ist nach dem Verändern des Codes für PIAFStat zu beachten, die Platzhalter sind wieder einzurichten.

Das Korrigieren und Weiterentwickeln der Verfahren in SAS direkt erfordert, dass man sich im SAS-Code schnell und sicher zurechtfindet und den Bezug zur PIAFStat-Bibliothek nicht verliert. Die Nutzung mehrerer Funktionen, Optionen und Blöcke in einem Verfahren kann dies erschweren. Kommentarzeilen erleichtern das Zuordnen dieser Programm-Abschnitte in PIAFStat. Diese Kommentarzeilen sollten den Namen von Funktion, Option und/oder Block und die Information Beginn oder Ende enthalten.

Beispiel:

```
/*Funktion Macros - Block BBL - Beginn */;
/*BBL - Zeile gleich WDH*/;
%macro bl;
%if &Z= %then %do;
%let Z=&B;
%end;
%mend;
/*Funktion Macros - Block BBL - Ende */;
```

Die Nutzung der von PIAFStat automatisch gebildeten globalen Makrovariablen und Datasets bietet sich bei der Verfahrenserstellung an.

Beispiel:

```
libname e "&VERZRES";
data daten;
set daten;
XKOORD=X_M;
call symput('XKOORD',XKOORD);
run;
```

Mitunter findet sich für eine zu lösende Problemstellung ein von einem Kollegen erstelltes SAS-Makro, das dieser zur Weiterverwendung zur Verfügung gestellt hat. Zu empfehlen ist, dieses Makro als eigene Funktion in ein PIAFStat-Verfahren zu integrieren. Bei einem Update des Makro kann auch die PIAFStat-Funktion unkompliziert erneuert werden. So wurden z.B. die an der Universität Hohenheim entwickelten und von Jens Möhring geschriebenen Makros `%boxcox`, `%varianzkomponenten`, und `%lsm` in die PIAFStat-Verfahren PHI, VK und MW der Hohenheim-Gülzower Serienauswertung integriert [9]. Ein weiteres Beispiel ist das Makro `%Mult`, das von Hans-Peter Piepho entwickelt und Andreas Büchse angepasst wurde. Es dient der Visualisierung signifi-

kanter Mittelwertvergleiche in gewohnter Buchstabendarstellung nach der Methode von Newman & Keuls in MIXED und findet in einem PIAFStat-Verfahren zur varianzanalytischen Auswertung einfaktorieller Blockanlagen Anwendung.

Für den Anwender ist es wichtig, dass Ergebnisse in den bekannten Standard-Formaten ausgegeben werden und sofort für weitere Berechnungen oder zur Einarbeitung in Berichte zur Verfügung stehen. Gewünscht werden deshalb insbesondere Word- und Excel-Dateien. Bei sehr umfangreichen Verfahren kann es zu einer Vielzahl von Ergebnis-Outputs kommen. Um die Übersicht zu behalten, empfiehlt es sich, die Ergebnisse zu bündeln und je Verfahrensdurchlauf in einer rtf-Datei und einer xls-Datei abzuspeichern. Dazu wird über ODS z.B. die rtf-Datei mehrfach geöffnet und geschlossen.

Beispiel:

```
/* ods rtf öffnen*/
ods rtf file= RtfName ;
/* Ausgabe 1*/;
title "Plausibilitätsprüfung";
proc print data= plausi;
run;
/* ods rtf temporär schließen*/
ods rtf exclude all;
. . .
/* weiterer SAS-Code*/
. . .
/* ods rtf temporär öffnen*/
ods rtf select all;
/* Ausgabe 2*/;
/* Grafik Häufigkeitsverteilung der Residuen*/
goptions;
ods graphics on;
proc sgrender data=gesamt template="graph.myplot";
run;
ods graphics off;
/* ods rtf schließen*/
ods rtf close;
```

Neben Abspeichern externer Dateien (*xls, *rtf oder andere) ist auch das Löschen dieser aus SAS heraus oftmals ein notwendiger Programmschritt. Die Nutzung der X-Befehle ermöglicht das sehr einfach. Mit der Option ‚noxwait‘ können einfache DOS-Befehle in SAS abgesetzt werden, nach dessen Ausführung die Kontrolle wieder an SAS zurückgegeben wird [5].

Beispiel:

```
option noxwait;
* Löschen der Datei keff.xls im Pfad C:\temp;
x "cd C:\temp";
x "del keff.xls";
```

Auch Access ist als Archiv für Ergebnistabellen hervorragend nutzbar. Zudem können Informationen, die SAS aus einer externen Quelle beziehen soll, sehr einfach über Access nach SAS eingespielt werden.

Beispiel:

```
/* Ergebnistabellenablage in Access*/
proc export data=zusa outtable="Mittelwerte"
dbms=access replace;
database="&PFAD\&Kultur..mdb";
run;
/*Info-Datei einlesen*/
proc import dbms=ACCESS table="Infodatei PIAFStat" out=infodat;
database="&PFAD\&Kultur..mdb";
run;
```

Nachteilig ist, dass die Kenntnis von Access nicht vorausgesetzt werden kann. Eine häufige Forderung von PIAFStat-Nutzern war, dass die Ergebnisdateien logische Namen erhalten und nicht durch einen Neustart des Verfahrens aus Versehen überschrieben werden. Um dies zu erreichen, werden in Standardverfahren in MV zur Bildung der Dateinamen die globalen Variablen von PIAFStat sowie die Makrovariable ‚SYSJOBID‘ verwendet [3]. Dadurch bekommen die Dateinamen neben Versuchsjahr und Ortsbezeichnung eine von SAS generierte vierstellige Nummer, die sich bei jedem Verfahrensablauf verändert.

Beispiel für die Bildung des Dateinamens:

```
&VJAHR.&VORTBEZ.&SYSJOBID..rtf
```

In vielen Fällen hat sich die Nutzung automatischer Makrovariablen und automatischer SQL-Makrovariablen [4] als sehr hilfreich erwiesen. Mit ihrer Hilfe können unter anderem wichtige Systeminformationen genutzt und im Output als Information mit ausgegeben werden. Folgende Tabelle zeigt eine Auswahl interessanter Makrovariablen.

Tabelle 3: Automatische Makrovariablen und SQL-Makrovariablen

Variable	Inhalt
SYSDATE	Aktuelles Datum des Programmstartes
SYSENV	Ausführungsmodus, Fore=interaktiv, Back=Batchjob
SYSERR	Fehlerinformation (Return Code) von Prozeduren oder Datensritten
SYSMACRONAME	Name des aktuellen Makros
SQL OBS	Anzahl zurückgelieferter Beobachtungen
SQLRC	Ausführungsstatus, 0= ohne Fehler ausgeführt

6 Fazit

PIAFStat ist ein kleines, anwenderfreundliches Programm, mit dem auch dem SAS-unkundigen Kollegen die Arbeit mit SAS-Programmen ermöglicht wird. Diese SAS-Programme werden als PIAFStat-Verfahren in einer Verfahrensbibliothek erstellt und verwaltet.

Die Programmierung in PIAFStat erfordert in erster Linie die Kenntnis der SAS-Sprache. Jedoch können die Programme durch PIAFStat-typische syntaktische Elemente so gestaltet werden, dass der Anwender optimal unterstützt wird und interaktiv in den Programmablauf eingreifen kann.

Ein wesentlicher Vorteil von PIAFStat: leistungsstarke, moderne und oftmals aufwendig in SAS programmierte PIAFStat-Verfahren können von vielen Kollegen genutzt werden. Das wird z.B. in den Länderdienststellen zum Sortenwesen und zum Pflanzenschutz praktiziert, dort werden PIAFStat-Verfahren gemeinschaftlich genutzt. Diese Verfahren stehen im Internet unter www.piafstat.de/dload.htm jedem PIAFStat-Anwender zur Verfügung.

Literatur

- [1] Michel, V.; Schmidtke, J.; Zenk, A.: Die Bibliothek biometrischer Verfahren in PIAFStat. Online Source (<http://www.landwirtschaft-mv.de>) (2008)
- [2] Michel, V.; Zenk, A.: Konzipierung, Entwicklung und Einführung von PIAF. Mitteilungen der Landesforschungsanstalt für Landwirtschaft und Fischerei Mecklenburg-Vorpommern 37(2007)59-60.
- [3] SAS Institute: Macro Language Elements: Automatic Macro Variables. SAS(R) 9.2 Macro Language: Reference, Online Source (20.11.2008)
- [4] SAS Institute. Using PROC SQL with the SAS Macro Facility. SAS(R) 9.2 SQL Procedure User's Guide, Online Source (26.2.2009)
- [5] SAS Institute. Running Windows or MS-DOS Commands from within SAS - 9.3. SAS(R) 9.3 Companion for Windows, Online Source (9.7.2011)
- [6] Schmidtke, J.: Handbuch PIAFStat, Version 6.34. (19.10.2011)1-58.
- [7] J. Schmidtke, K. Voit. PIAFStat - Datenanalyse für landwirtschaftliche Feldversuche. 5. KSFE, Proceedings der 5. KSFE, (2001) 317-321.
- [8] Zenk, A.: DEMO / einfaktorielle Blockanlage - Beispiel KSFE 2012. PIAFStat-Verfahren(2012)
- [9] Zenk, A.; Michel, V.; Möhring, J.: Einbindung neuer Methoden zur Routineauswertung von landwirtschaftlichen Versuchen mit Hilfe von SAS-Macros. 9. KSFE - SAS: Verbindung von Theorie und Praxis, Proceedings der 9.KSFE, (2005)407-417.