Stability Analysis with **agricolae**

Felipe de Mendiburu[1], Muhammad Yaseen[2]

2020-05-01

Contents

1. Professor of the Academic Department of Statistics and Informatics of the Faculty of Economics and Planning.National University Agraria La Molina-PERU.

2. Department of Mathematics and Statistics, University of Agriculture Faisalabad, Pakistan.

1 Stability Analysis

In **agricolae** there are two methods for the study of stability and the AMMI model. These are: a parametric model for a simultaneous selection in yield and stability "SHUKLA'S STABILITY VARIANCE AND KANG'S", (Kang, 1993) and a non-parametric method of Haynes, based on the data range.

1.1 Parametric Stability

Use the parametric model, function `stability.par`.

Prepare a data table where the rows and the columns are the genotypes and the environments, respectively. The data should correspond to yield averages or to another measured variable. Determine the variance of the common error for all the environments and the number of repetitions that was evaluated for every genotype. If the repetitions are different, find a harmonious average that will represent the set. Finally, assign a name to each row that will represent the genotype (Kang, 1993). We will consider five environments in the following example:

```
options(digit=2)
f <- system.file("external/dataStb.csv", package="agricolae")
dataStb<-read.csv(f)
stability.par(dataStb, rep=4, MSerror=1.8, alpha=0.1, main="Genotype",console=TRUE)
```

```
INTERACTIVE PROGRAM FOR CALCULATING SHUKLA'S STABILITY VARIANCE AND KANG'S
                    YIELD - STABILITY (YSi) STATISTICS
```

```
Genotype
Environmental index  - covariate

Analysis of Variance
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Total	203	2964.1716			
Genotypes	16	186.9082	11.6818	4.17	<0.001
Environments	11	2284.0116	207.6374	115.35	<0.001
Interaction	176	493.2518	2.8026	1.56	<0.001
Heterogeneity	16	44.8576	2.8036	1	0.459
Residual	160	448.3942	2.8025	1.56	<0.001
Pooled Error	576		1.8		

```
Genotype. Stability statistics
```

	Mean	Sigma-square	.	s-square	.	Ecovalence
A	7.4	2.47	ns	2.45	ns	25.8
B	6.8	1.60	ns	1.43	ns	17.4
C	7.2	0.57	ns	0.63	ns	7.3
D	6.8	2.61	ns	2.13	ns	27.2
E	7.1	1.86	ns	2.05	ns	19.9
F	6.9	3.58	*	3.95	*	36.5
G	7.8	3.58	*	3.96	*	36.6
H	7.9	2.72	ns	2.12	ns	28.2
I	7.3	4.25	**	3.94	*	43.0
J	7.1	2.27	ns	2.51	ns	23.9
K	6.4	2.56	ns	2.55	ns	26.7
L	6.9	1.56	ns	1.73	ns	16.9
M	6.8	3.48	*	3.28	ns	35.6
N	7.5	5.16	**	4.88	**	51.9
O	7.7	2.38	ns	2.64	ns	24.9
P	6.4	3.45	*	3.71	*	35.3
Q	6.2	3.53	*	3.69	*	36.1

```
Signif. codes:  0 '**' 0.01 '*' 0.05 'ns' 1

Simultaneous selection for yield and stability  (++)
```

	Yield	Rank	Adj.rank	Adjusted	Stab.var	Stab.rating	YSi	...
A	7.4	13	1	14	2.47	0	14	+
B	6.8	4	-1	3	1.60	0	3	
C	7.2	11	1	12	0.57	0	12	+
D	6.8	4	-1	3	2.61	0	3	
E	7.1	9	1	10	1.86	0	10	+
F	6.9	8	-1	7	3.58	-4	3	
G	7.8	16	2	18	3.58	-4	14	+
H	7.9	17	2	19	2.72	0	19	+
I	7.3	12	1	13	4.25	-8	5	
J	7.1	10	1	11	2.27	0	11	+
K	6.4	3	-2	1	2.56	0	1	

```
L   6.9    7        -1         6    1.56           0    6
M   6.8    6        -1         5    3.48          -4    1
N   7.5   14         1        15    5.16          -8    7    +
O   7.7   15         2        17    2.38           0   17    +
P   6.4    2        -2         0    3.45          -4   -4
Q   6.2    1        -3        -2    3.53          -4   -6


 Yield Mean: 7.1
 YS    Mean: 6.8
 LSD (0.05): 0.45
 - - - - - - - - - - -
 +   selected genotype
 ++  Reference: Kang, M. S. 1993. Simultaneous selection for yield
 and stability: Consequences for growers. Agron. J. 85:754-757.
```

For 17 genotypes, the identification is made by letters. An error variance of 2 and 4 repetitions is assumed.

1.1.1 Analysis

```
output <- stability.par(dataStb, rep=4, MSerror=2)
names(output)
```

```
[1] "analysis"   "statistics" "stability"
```

```
print(output$stability)
```

```
   Yield Rank Adj.rank Adjusted Stab.var Stab.rating YSi ...
A   7.4   13         1        14    2.47           0   14    +
B   6.8    4        -1         3    1.60           0    3
C   7.2   11         1        12    0.57           0   12    +
D   6.8    4        -1         3    2.61           0    3
E   7.1    9         1        10    1.86           0   10    +
F   6.9    8        -1         7    3.58          -2    5
G   7.8   16         2        18    3.58          -2   16    +
H   7.9   17         2        19    2.72           0   19    +
I   7.3   12         1        13    4.25          -4    9    +
J   7.1   10         1        11    2.27           0   11    +
K   6.4    3        -2         1    2.56           0    1
L   6.9    7        -1         6    1.56           0    6
M   6.8    6        -1         5    3.48          -2    3
N   7.5   14         1        15    5.16          -8    7
O   7.7   15         2        17    2.38           0   17    +
P   6.4    2        -2         0    3.45          -2   -2
Q   6.2    1        -2        -1    3.53          -2   -3
```

The selected genotypes are: A, C, E, G, H, I, J and O. These genotypes have a higher yield and a lower variation. to see output$analysis, the interaction is significant.

If for example there is an environmental index, it can be added as a covariate In the first five locations. For this case, the altitude of the localities is included.

```
data5<-dataStb[,1:5]
altitude<-c(1200, 1300, 800, 1600, 2400)
```

```
stability <- stability.par(data5,rep=4,MSerror=2, cova=TRUE, name.cov= "altitude",
file.cov=altitude)
```

1.2   Non-parametric Stability

For non-parametric stability, the function in **agricolae** is `stability.nonpar()`. The names of the genotypes should be included in the first column, and in the other columns, the response by environments (Haynes et al., 1998).

1.2.1   Analysis

```
data <- data.frame(name=row.names(dataStb), dataStb)
output<-stability.nonpar(data, "YIELD", ranking=TRUE)
names(output)
```

```
[1] "ranking"    "statistics"
```

```
output$statistics
```

```
  MEAN es1 es2  vs1 vs2 chi.ind chi.sum
1  7.1 5.6   24 0.72  47     8.8      28
```

1.3   AMMI

The model AMMI uses the biplot constructed through the principal components generated by the interaction environment-genotype. If there is such interaction, the percentage of the two principal components would explain more than the 50% of the total variation; in such case, the biplot would be a good alternative to study the interaction environment-genotype (Crossa, 1990).

The data for AMMI should come from similar experiments conducted in different environments. Homogeneity of variance of the experimental error, produced in the different environments, is required. The analysis is done by combining the experiments.

The data can be organized in columns, thus: environment, genotype, repetition, and variable.

The data can also be the averages of the genotypes in each environment, but it is necessary to consider a harmonious average for the repetitions and a common variance of the error. The data should be organized in columns: environment, genotype, and variable.

When performing AMMI, this generates the Biplot, Triplot and Influence graphics, see Figure 1.

For the application, we consider the data used in the example of parametric stability (study):

1.3.1   AMMI structure

```
  str(AMMI)
```

```
function (ENV, GEN, REP, Y, MSE = 0, console = FALSE, PC = FALSE)
```

1.3.2   plot.AMMI structure, plot()

```
  str(plot.AMMI)
```

```
data(plrv)
model<-with(plrv,AMMI(Locality, Genotype, Rep, Yield, console=FALSE))
names(model)
```

```
[1] "ANOVA"     "genXenv"  "analysis" "means"     "biplot"    "PC"
```

```
model$ANOVA
```

```
Analysis of Variance Table

Response: Y
            Df Sum Sq Mean Sq F value  Pr(>F)
ENV          5 122284   24457  257.04 9.1e-12 ***
REP(ENV)    12   1142      95    2.57  0.0029 **
GEN         27  17533     649   17.54 < 2e-16 ***
ENV:GEN    135  23762     176    4.75 < 2e-16 ***
Residuals  324  11998      37
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model$analysis
```

```
    percent acum Df Sum.Sq Mean.Sq F.value   Pr.F
PC1    56.3   56 31  13369     431   11.65 0.0000
PC2    27.1   83 29   6428     222    5.99 0.0000
PC3     9.4   93 27   2242      83    2.24 0.0005
PC4     4.3   97 25   1028      41    1.11 0.3286
PC5     2.9  100 23    696      30    0.82 0.7059
```

```
pc <- model$analysis[, 1]
pc12<-sum(pc[1:2])
pc123<-sum(pc[1:3])
```

```
oldpar<-par(cex=0.4,mar=c(4,4,1,2))
plot(model,type=1,las=1,xlim=c(-5,6))
```

```
par(oldpar)
```

In this case, the interaction is significant. The first two components explain 83.4 %; then the biplot can provide information about the interaction genotype-environment. With the triplot, 92.8% would be explained.

To triplot require klaR package. in R execute:

```
plot(model,type=2,las=1)
```

1.4   AMMI index and yield stability

Calculate AMMI stability value (ASV) and Yield stability index (YSI) (Purchase, 1997; N. et al., 2008).

```
data(plrv)
model<- with(plrv,AMMI(Locality, Genotype, Rep, Yield, console=FALSE))
index<-index.AMMI(model)
# Crops with improved stability according AMMI.
print(index[order(index[,3]),])
```
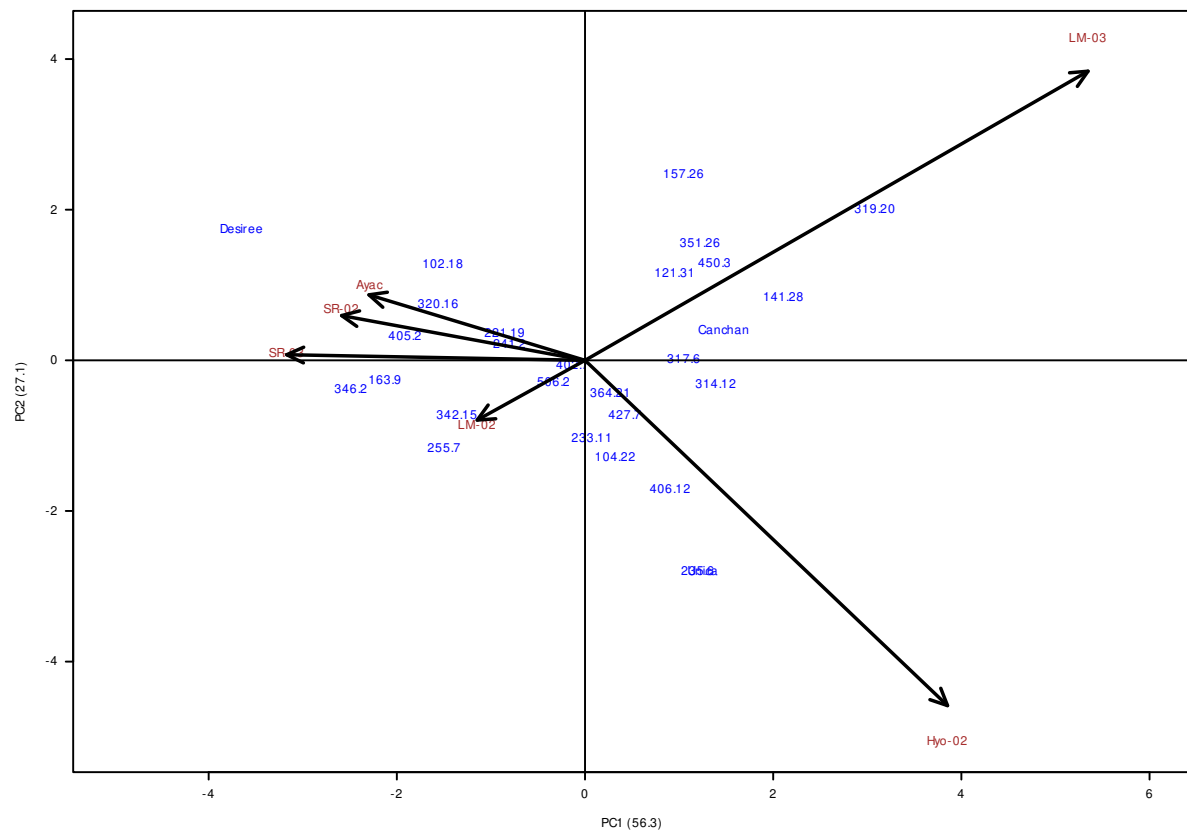
Figure 1: Biplot

```
         ASV  YSI rASV rYSI means
402.7    0.28  20    1   19    27
364.21   0.72  12    2   10    34
506.2    0.75  14    3   11    33
233.11   1.06  21    4   17    29
427.7    1.15  12    5    7    36
104.22   1.46  19    6   13    31
241.2    1.68  29    7   22    26
221.19   1.80  34    8   26    23
317.6    2.19  18    9    9    35
121.31   2.29  25   10   15    30
406.12   2.56  23   11   12    33
314.12   2.92  30   12   18    28
342.15   2.92  37   13   24    26
351.26   2.98  22   14    8    36
Canchan  3.10  35   15   20    27
450.3    3.14  22   16    6    36
157.26   3.29  22   17    5    37
320.16   3.32  39   18   21    26
255.7    3.33  33   19   14    31
102.18   3.38  43   20   23    26
235.6    3.76  25   21    4    39
Unica    3.84  24   22    2    39
405.2    3.98  39   23   16    29
163.9    4.43  51   24   27    21
141.28   4.47  26   25    1    40
346.2    5.18  51   26   25    24
319.20   6.72  30   27    3    39
Desiree  7.78  56   28   28    16
```

```
# Crops with better response and improved stability according AMMI.
print(index[order(index[,4]),])
```

```
         ASV  YSI rASV rYSI means
141.28   4.47  26   25    1    40
Unica    3.84  24   22    2    39
319.20   6.72  30   27    3    39
235.6    3.76  25   21    4    39
157.26   3.29  22   17    5    37
450.3    3.14  22   16    6    36
427.7    1.15  12    5    7    36
351.26   2.98  22   14    8    36
317.6    2.19  18    9    9    35
364.21   0.72  12    2   10    34
506.2    0.75  14    3   11    33
406.12   2.56  23   11   12    33
104.22   1.46  19    6   13    31
255.7    3.33  33   19   14    31
121.31   2.29  25   10   15    30
405.2    3.98  39   23   16    29
233.11   1.06  21    4   17    29
314.12   2.92  30   12   18    28
402.7    0.28  20    1   19    27
```

```
Canchan 3.10   35   15   20    27
320.16  3.32   39   18   21    26
241.2   1.68   29    7   22    26
102.18  3.38   43   20   23    26
342.15  2.92   37   13   24    26
346.2   5.18   51   26   25    24
221.19  1.80   34    8   26    23
163.9   4.43   51   24   27    21
Desiree 7.78   56   28   28    16
```

References

Crossa, J. (1990). "Statistical Analyses of Multilocation Trials," in *Advances in agronomy* (Elsevier), 55–85.

Haynes, K. G., Lambert, D. H., Christ, B. J., Weingartner, D. P., Douches, D. S., Backlund, J. E., et al. (1998). Phenotypic stability of resistance to late blight in potato clones evaluated at eight sites in the United Stated. 75, 211–217.

Kang, M. S. (1993). Phenotypic stability of resistance to late blight in potato clones evaluated at eight sites in the United Stated. *Agronomy Journal* 85, 754–757.

N., S., H.Sabaghpour, S., and Dehghani, H. (2008). The use of an AMMI model and its parameters to analyse yield stability in multienvironment trials. *Journal of Agricultural Science* 146, 571–581.

Purchase, J. L. (1997). Parametric analysis to describe genotype environment interaction and yield stability in winter wheat.