Selection of alfalfa cultivars adapted for tropical environments with repeated measures using PROC MIXED of SAS® System

G. M. L. de Assis[1]*, A. C. Ruggieri[2], M. E. Z. Mercadante[3], G. M. F. de Camargo[2] and J. M. Carneiro Júnior[1]

[1]*Brazilian Agricultural Research Corporation (Embrapa), Agroforestry Research Centre of Acre State (Embrapa Acre), Rodovia BR 364, km 14, C. P. 321, CEP 69908-970, Rio Branco, AC, Brazil,* [2]*São Paulo State University (Unesp), Jaboticabal, SP, Brazil and* [3]*Animal Science Institute, Sertãozinho, SP, Brazil*

Abstract

Although alfalfa (*Medicago sativa* L.) is a leguminous herbage widely used in temperate regions as animal feed, there is not much research in tropical regions to develop cultivars adapted to these environmental conditions. The utilization of adapted cultivars with adequate management practices is important to improve productivity, quality and persistence of cultivated pastures. The objectives of this study were to verify the genetic variability among alfalfa cultivars and to rank them using mixed model methodology. A total of 35 alfalfa cultivars were evaluated in the rainy and dry seasons, from 1996 to 2000, in plots of $2.8\,m^2$ in Sertãozinho, São Paulo, Brazil. The experimental design was a randomized complete block with three replications. Longitudinal data of dry matter yield were analyzed using PROC MIXED of SAS® System. Several covariance structures were tested and the spherical spatial structure was selected. The results show that the genetic variability was statistically significant only for the dry season. Moreover, the interaction among cultivars and harvests variance was highly significant for both seasons. The empirical best linear unbiased predictions of cultivar effects were obtained, allowing for the selection of the superior cultivars MH 15, 5715, SW 8210, Rio, High, 5888, Monarca, Victoria, Florida 77 and Falcon. Crioula, the most common cultivar in Brazil, showed low forage potential in Sertãozinho. Results indicate potential for use of more productive cultivars of alfalfa to produce animal feed in tropical environments.

Keywords: covariance structure; genetic evaluation; herbage production; longitudinal data; *Medicago sativa*; PROC MIXED

Introduction

Alfalfa (*Medicago sativa* L.) is one of the most disseminated forage legumes in the temperate regions of the world. It can be used for hay, silage and pasture. With the intensification of milk and meat production systems, there is an increased need to investigate the potential use of alfalfa as a source of feed for cattle in the tropics. Alfalfa can contribute to the productivity and profitability of milk and meat production systems as a result of its symbiotic nitrogen fixation, its ability to mobilize nutrients from the deepest layers of the soil, its high

*Corresponding author. E-mail: giselle@cpafac.embrapa.br

nutritional quality and forage productivity. However, there has been little focus on the potential use of alfalfa cultivars in tropical regions. The utilization of adapted cultivars in combination with adequate herbage management practices is important for the improvement of pasture productivity, quality and persistence.

Genetic evaluation of perennial forage species usually takes several years with repeated samples having to be taken from the same plant. This, in addition to missing data, makes statistical analyses quite complex. Usually, in a repeated-measure design, it is assumed that errors are independent, so that all the observations within an experimental unit are equally correlated. However, in a typical repeated-measure experiment, two measurements taken at adjacent times are more correlated than two measurements taken at different times (Littell et al., 1996). In order to accommodate the correlation structure of the repeated measures, specific methods of analysis are required.

Plant breeders have traditionally estimated variance and covariance components using the method of moments on the basis of analysis of variance (ANOVA). The drawbacks of ANOVA include ignorance of the distributional properties of estimators when data are unbalanced and the possibility of obtaining estimates outside of the established parameter range (Liu et al., 1997). An alternative is the restricted maximum likelihood (REML) method (Patterson and Thompson, 1971; Holland, 2006), which has been used largely by animal breeders and recently by tree and forage breeders (Assis et al., 2008; Resende et al., 2008). Besides the estimation of genetic parameters, breeders are also interested in identifying superior genotypes. The usual methodologies, based on least-square estimation, are not the best suited for analysis of data from perennial plant-breeding programmes (Resende, 2002). Several authors (Duarte and Vencovsky, 2001; Smith et al., 2001; Resende, 2004; Welham et al., 2004; Resende and Duarte, 2007) have emphasized the importance of considering genotypes as random effects. Resende and Duarte (2007) do not recommend the use of multiple comparison tests in cultivar trials when the number of treatments is higher than four. Thus, in order to select superior genotypes, the empirical best linear unbiased prediction (EBLUP) of the breeding values is used (Smith et al., 2001; Resende, 2002; Furlani et al., 2005).

Therefore, if appropriate statistical models are defined in agreement with the experimental designs and also in agreement with the objectives of breeding programme, mixed model methodology (Henderson, 1973) can be used for the genetic evaluation of plants. Analyses based on mixed model methodology can be accomplished through the MIXED procedure of SAS® System (Littell et al., 1998), which allows access to the covariance structure among measures over time. It is also possible to accomplish efficient statistical tests for fixed and random effects. In addition, one can ignore all measures of a certain genotype if some observations are lost.

The objectives of this study were to verify the genetic variability among alfalfa cultivars grown in a tropical environment using variance components estimation and then to rank them using PROC MIXED of SAS® System.

Materials and methods

Data and experimental design

This experiment is part of the National Evaluation Program of Alfalfa Cultivars (RENACAL), coordinated by the Brazilian Agricultural Research Corporation (Embrapa). It was developed by the Advanced Center for Technological Beef Cattle Agribusiness Research, located in Sertãozinho, São Paulo, Brazil. This Center is located at an altitude of 548 m, 21°08′S latitude and 47°59′W longitude. The climate is humid tropical, classified as AW according to Koppen, with a rainy season (RS) in the summer and a dry season (DS) in the winter. The soil is classified as oxissoil with high fertility.

A total of 35 alfalfa cultivars were evaluated from 1996 to 2000 in plots of 2.8 m². The experiment was designed in complete randomized blocks with three replications. The plants were harvested when they reached 10% of flowering and the data were collected at unequal intervals of time. There were 23 harvests in the RS and eight in the DS. The first harvest, which occurred on 12 September 1996, was not included in the analyses. Alfalfa cultivars were harvested on 42, 77, 116, 160, 194, 397, 427, 467, 495, 525, 567, 771, 804, 826, 854, 889, 939, 1134, 1183, 1217, 1245 and 1288 days after the first harvest in the RS and 244, 293, 327, 609, 658, 714, 987 and 1033 days after the first harvest in the DS. Four incomplete agricultural years were considered. Year 1, year 2, year 3 and year 4 encompass harvests from October 1996 to September 1997, from October 1997 to September 1998, from October 1998 to September 1999 and from October 1999 to March 2000, respectively. The trait evaluated at each harvest was dry matter yield (DMY; kg/ha). This experiment became unbalanced for RS because the block 1 replications of harvest number 13 were accidentally lost.

Model and statistical analyses

The following mixed linear model was considered:

$$y_{ijk} = \mu + c_i + b_j + \delta_{ij} + h_k + ch_{ik} + \varepsilon_{ijk},$$

where y_{ijk} is the observation of cultivar i on block j from harvest k; μ is a constant inherent to all observations; c_i is the random effect of cultivar i; b_j is the random effect of block j; δ_{ij} is the residual term referred to interaction of cultivar i with block j; h_k is the fixed effect of harvest k; ch_{ik} is the random effect of interaction among cultivar i and harvest k; and ε_{ijk} is the random error.

The analyses were done for RS and DS separately. Initially, data of each season were submitted to sphericity test to verify the possibility to perform split plot ANOVA. The sphericity condition is satisfied when the variances of differences between the harvests are all the same (Huynh and Feldt, 1970). These tests were performed using the REPEATED command through PROC GLM of SAS® (Littell et al., 1998), and the obtained covariance matrices for both seasons did not satisfy sphericity condition. So it is not appropriate to analyze these repeatedly measured data considering a split plot design.

Hence, the data were then analyzed by PROC MIXED, available in SAS® System (Littell et al., 1998). A RANDOM statement was used to specify the random effects and to model variation between experimental units, whereas a REPEATED statement was added to model the covariance structure within experimental units (**R** matrix).

PROC MIXED provides an assortment of covariance types from which to select (Littell et al., 1996); however, several of them are not adequate to unequally spaced measures. The compound symmetry (CS) and unstructured (UN) structures are still appropriate, but CS assumes that the correlations remain constant and UN is often too general. Spatial structures are also available in PROC MIXED and are useful for unequally spaced longitudinal measurements. Thus, five different covariance structures were evaluated to model covariances within experimental units: two traditional structures – CS and UN – and three spatial structures – SP(POW), SP(GAU) and SP(SPH). SP means Spatial and POW, GAU and SPH mean Power, Gaussian and Spherical, respectively. More details about these structures are found in Verbeke and Molenberghs (2000).

Two different methods to fit criteria were considered to compare and select the most appropriate covariance structure: Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC). The AIC and SBC are adjusted versions of REML log likelihood. It was used to impose a penalty according to the number of parameters estimated (Littell et al., 1998). The penalty imposed by SBC is more severe than the one imposed by AIC.

Variance components were estimated by REML (Patterson and Thompson, 1971), and approximate Wald and likelihood ratio tests were performed to verify the significance of components. Based on these variance components, EBLUP of cultivar effects was obtained for each season. Thus, cultivars were classified

and it was possible to select those that are genetically superior for dry matter production.

An example of SAS statements used to perform the above analysis:

```
PROC MIXED DATA = DRY COVTEST;
CLASS CULT HARVEST REPETITION SUBJECT;
MODEL DMY = HARVEST;
RANDOM CULT REPETITION CULT × REPETITION CULT × HARVEST/SOLUTION;
REPEATED HARVEST/TYPE = SP(POW) (DAYS) SUB = SUBJECT R RCORR;
RUN;
```

Some options were added to data analysis, like COVTEST, SOLUTION, R, SUB and RCORR:

(1) through option COVTEST, one can print the resulting asymptotic standard errors and associated Wald tests for variance components;
(2) using the SOLUTION option, the predictions for the random effects parameters (EBLUP) are printed;
(3) with the R option, blocks of the estimated R matrix can be displayed (the first block determined by the 'SUB = effect' is the default displayed block);
(4) the SUB option defines SUBJECT (each CULTIVAR × REPETITION, which is the experimental unit) to be a blocking factor, where data from different blocks are assumed to be independent; and
(5) the RCORR option produces the correlation matrix corresponding to blocks of the estimated **R** matrix.

This example uses SP(POW) structure, which was previously selected. The TYPE option specifies the structure using the continuous variable DAYS to indicate time levels, which is a copy of HARVEST in the DRY dataset.

Genotypic values ($\mu + c$) of each cultivar obtained from EBLUP were presented only for DS, since the hypothesis that variance component of cultivar effect for RS is equal to zero was not rejected. Additionally, upper and lower confidence limits were calculated for each cultivar.

Since the interaction cultivar × harvest was significant, regression analysis for all harvests (RS and DS) were performed for the ten cultivars that had highest dry matter production in DS.

Results

Descriptive analysis

The results of DMY (kg/ha) per harvest for each cultivar during the RS and DS for each agricultural year are

Table 1. Dry matter yield production means (kg/ha) per harvest for each cultivar, considering rainy season (RS) and dry season (DS) of each agricultural year

| Cultivar | Year 1 | | Year 2 | | Year 3 | | Year 4 |
	RS	DS	RS	DS	RS	DS	RS
Valley Plus	2847.9	1251.9	2116.8	1385.5	1556.1	647.3	1350.9
WL 516	2592.5	1180.4	2237.9	1151.6	1679.0	824.3	1502.2
Alfa 200	2801.8	1229.9	2236.5	1282.0	1768.2	854.3	1565.8
Falcon	2680.4	1323.8	1928.8	1242.6	1681.6	860.7	1423.1
SW 8210	2977.0	1376.2	2384.8	1381.7	1858.7	853.8	1648.7
SW 8112A	2704.1	1142.9	2117.0	1202.1	1633.0	816.5	1529.8
Alto	2951.8	1320.5	2357.3	1414.9	1750.2	860.0	1582.6
Rio	3042.0	1473.4	2167.9	1275.2	1670.9	884.2	1458.1
ICI 990	2452.2	1098.3	1904.6	1193.4	1703.1	781.3	1452.6
Monarca	3115.1	1326.1	2357.8	1275.3	1894.6	963.1	1555.1
Victoria	3008.1	1422.0	2262.9	1341.2	1923.9	725.1	1518.7
Esmeralda	2853.5	1119.5	2134.5	1173.1	1736.4	783.0	1464.6
Costera	3059.1	1230.9	2096.2	1043.8	1478.4	393.5	1241.6
Semit 711	3030.4	1263.6	2222.4	1207.3	1660.2	577.6	1358.4
Semit 921	2572.9	1082.5	2058.1	1119.5	1774.8	773.5	1492.0
Araucana	3030.9	1272.5	2302.4	1239.1	1898.4	757.7	1488.9
Maricopa	2939.1	1079.6	2306.3	1175.2	1871.0	685.6	1511.5
Sutter	2755.0	1103.9	2233.9	1094.8	1740.1	584.8	1462.9
P 30	2855.3	1135.6	1998.7	955.8	1342.7	575.6	1279.6
P 205	2961.2	1096.9	1923.2	846.4	1378.3	473.2	1147.5
F 708	3029.5	1120.7	2204.6	1194.1	1619.4	668.8	1505.0
F 686	3022.3	1305.0	2340.8	1201.3	1695.5	627.8	1372.5
El Grande	2786.9	933.6	2141.4	1131.6	1625.2	765.4	1491.8
5929	2708.7	1046.1	2047.6	1268.6	1834.2	822.7	1682.1
Florida 77	2925.1	1315.1	2162.6	1294.2	1728.0	899.5	1617.0
5888	2885.5	1185.5	2255.5	1493.7	1865.6	867.4	1666.9
5715	2712.2	1272.7	2262.9	1434.3	1942.2	1016.6	1818.9
MH 4	3091.0	1333.9	2184.3	1202.7	1546.8	734.1	1416.9
MH 15	3179.6	1433.6	2250.8	1361.7	1811.4	912.9	1581.2
BR 1	2948.9	1157.9	2142.3	1209.4	1696.3	809.1	1616.8
BR 2	3035.5	1319.1	2463.0	1249.3	1824.3	823.2	1601.6
BR 3	2911.6	1196.9	2221.7	1121.6	1697.2	786.9	1592.0
BR 4	2669.9	1234.7	2073.9	1140.2	1564.1	685.5	1491.9
SW9210 A	2913.3	1105.3	2192.5	1246.8	1716.1	944.1	1685.1
Crioula	2842.7	1208.4	2262.9	1055.5	1573.3	652.4	1726.9

shown in Table 1. The average of DMY considering all cultivars and all harvests was 1979 kg/ha per harvest for RS and 1105 kg/ha per harvest for DS. The highest average DMY was observed for RS, as was expected, and was responsible for about 81.3–85.0% of annual DMY. The total annual DMY varied from 12,900 kg/ha (P205 cultivar) to 16,200 kg/ha (Monarca); from 10,900 kg/ha (P205) to 13,400 kg/ha (Monarca and BR 2) for RS and from 1900 kg/ha (P205) to 2900 kg/ha (SW 8210, Rio, 5715 and MH 15) for DS. Unequal seasonal DMY distribution could be observed for all cultivars. However, special attention is necessary to select superior productive genotypes that have a better production distribution during the year. Thus, it is also important to study the trait seasonal DMY distribution to verify whether genetic variability exists among cultivars. A possible cause of the pronounced unequal seasonal distribution is the fact that there was no effective

irrigation of the experimental area. Thus, the lack of water in the DS was probably a determinant factor contributing to the low DMY observed.

High standard deviations were observed, mainly for RS. These high values are associated with high DMY at the beginning of the experiment and low DMY at the end of the experiment. For instance, DMY of cultivar MH 15 in a single harvest ranged from 5509 to 419 kg/ha in RS and from 2219 to 552 kg/ha in DS. These accentuated decreases in yields over the years indicate that a meticulous study of the persistence of the cultivars must be performed.

Covariance structure and statistical tests

Selection of covariance structure was based on goodness of fit for criteria AIC and SBC. The larger the value of
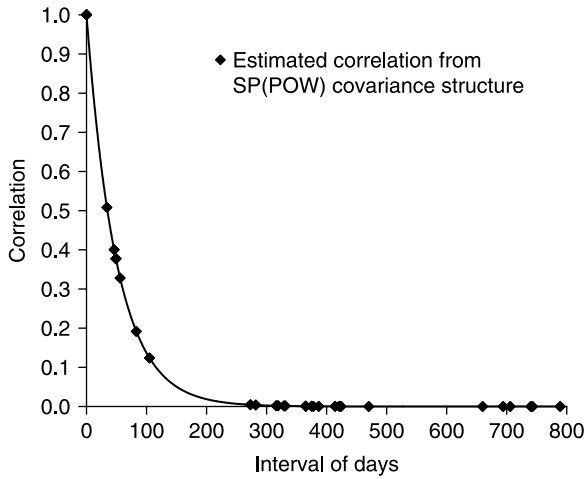
Fig. 1. Estimated correlations from SP(POW) covariance structure between harvests (expressed in interval of days) of alfalfa cultivars evaluated in dry season, showing the decay in correlation between harvests.

AIC or SBC, the better the structure. For both criteria, the best structure was SP(POW) for RS and DS. REML failed to converge when analyses were performed with UN structure, probably because it requires estimation of a large number of variance and covariance parameters.

Covariance components and correlations between harvests were estimated at RS and DS using SP(POW) structure, which models the covariance between two measurements at times t_1 and t_2 (Littell *et al.*, 1996) as follows:

$$\text{cov}(y_{t_1}, y_{t_2}) = \sigma^2 \rho^{|t_1 - t_2|},$$

where ρ is an autoregressive parameter, estimates were 0.9803 and 0.9554 for RS and DS, respectively, and σ^2 is the residual variance; estimates are shown below. The covariance matrix is printed for experimental unit number 1. It indicates that it is the covariance submatrix for the repeated measures from cultivar 1 drawn from block 1. All the other subjects (cultivars *i versus* block *j*) are assumed to have the same covariance matrix.

All correlations were of medium (for DS) and low (for RS) values or equal to zero (for RS and DS). It was also verified that there is no correlation between harvests

when different agricultural years are compared. The most significant correlations were between adjacent harvests, which were illustrated for DS in Fig. 1. For RS, correlations between adjacent harvests at year 1 varied from 0.13 to 0.21; at year 2 varied from 0.00 to 0.25; at year 3 varied from 0.10 to 0.37; and at year 4 varied from 0.11 to 0.28. For DS, they were highest, ranging from 0.38 to 0.51 at year 1; from 0.33 to 0.38 at year 2 and equal to 0.40 at year 3. Thus, a linear association only appears between adjacent harvests in the same year, mainly at DS. It emphasizes the necessity to evaluate perennial alfalfa cultivars during different years and harvests.

Variance components estimated by REML as well as the standard errors associated with each estimate and approximate Wald tests are in Table 2. The hypothesis $H_0 : \sigma^2_{cult} = 0$ was only rejected at the 16% level of significance for RS and at the 2% level of significance for DS. This result indicates that genetic variability among cultivars is larger at DS. Thus, selection of superior cultivars will be more efficient if it is based on predicted genotypic values of cultivars in this season. The rank of 35 alfalfa cultivars, based on EBLUP of cultivar effect evaluated in DS considering the SP(POW) structure, is presented in Table 3. Upper and lower confidence limits for each cultivar are also shown in Table 3.

The hypothesis $H_0 : \sigma^2_{cult \times harvest} = 0$ was rejected at 1 and 0.1% level of significance for RS and DS, respectively (Table 2). It indicates that there is a significant interaction among cultivars and harvests. Simple linear regression of dry matter production (y) on days after the first harvest (x) has presented the best fit of the data, when comparing with the quadratic linear regression. The regression equations obtained are presented in Table 4. The productivity over the years of ten alfalfa cultivars selected in DS (Table 3) was studied from regression analyses considering all harvests and can be observed in Fig. 2.

Discussion

Some experiments in Brazil have also shown the absence of genetic variability of alfalfa cultivars. Monteiro *et al.* (1998) evaluated 17 alfalfa cultivars in a subtropical

Table 2. Cultivar ($\hat{\sigma}^2_{cult}$), interaction cultivar with harvest ($\hat{\sigma}^2_{cult \times harvest}$) and residual ($\hat{\sigma}^2$) variance component estimates (VC) obtained by REML, their standard errors (SE) and approximate Wald test (Pr Z) for rainy season (RS) and dry season (DS)

	RS			DS		
	VC (kg^2/ha^2)	SE (kg/ha)	Pr Z	VC (kg^2/ha^2)	SE (kg/ha)	Pr Z
$\hat{\sigma}^2_{cult}$	2645	2644	0.16	6292	2902	0.02
$\hat{\sigma}^2_{cult \times harvest}$	6400	2598	0.01	3965	1265	0.001
$\hat{\sigma}^2$	114,882			31,148		

Table 3. Rank of 35 alfalfa cultivars based on EBLUP of cultivar effect ($\mu + c$) in kilograms (kg) per hectare (ha), lower confidence limit (LCL) and upper confidence limit (UCL) of confidence interval of the genotypic effect evaluated in dry season

Dry season			
Cultivar	$\mu + c$ (kg/ha)	LCL	UCL
MH 15	1198.2	1092.0	1304.5
5715	1198.0	1091.8	1304.3
SW 8210	1184.6	1078.4	1290.9
Rio	1182.2	1076.0	1288.5
Alto	1180.7	1074.5	1287.0
5888	1173.6	1067.4	1279.9
Monarca	1166.5	1060.3	1272.8
Victoria	1162.9	1056.7	1269.2
Florida 77	1158.6	1052.4	1264.9
Falcon	1143.6	1037.4	1249.9
BR 2	1138.3	1032.1	1244.6
Alfa 200	1133.7	1027.5	1240.0
Valley Plus	1130.5	1024.3	1236.8
Araucana	1119.3	1013.1	1225.6
MH 4	1118.8	1012.6	1225.1
SW9210 A	1115.3	1009.1	1221.6
F 686	1098.7	992.5	1205.0
BR 1	1096.3	990.1	1202.6
SW 8112A	1094.9	988.7	1201.2
WL 516	1092.9	986.7	1199.2
5929	1091.1	984.9	1197.4
Semit 711	1083.5	977.3	1189.8
BR 3	1082.7	976.5	1189.0
BR 4	1081.0	974.8	1187.3
Esmeralda	1079.6	973.4	1185.9
ICI 990	1079.4	973.2	1185.7
F 708	1067.9	961.7	1174.2
Semit 921	1059.7	953.5	1166.0
Maricopa	1057.9	951.7	1164.2
Crioula	1050.9	944.7	1157.2
El Grande	1033.7	927.5	1140.0
Sutter	1030.3	924.1	1136.6
Costera	1012.4	906.2	1118.7
P 30	1004.9	898.7	1111.2
P 205	958.5	852.3	1064.8

Approximate Wald test and standard errors are based on asymptotic properties and are not very reliable if the degrees of freedom to estimate the variance component are small (Littell et al., 1996; Steel et al., 1997). Another approximate test is possible when a set of model parameters can be set to zero. This is called the likelihood ratio test, drawing from likelihood theory. It is computed taking the difference between the -2 REML log likelihood of the model containing the random effect, whose component will be tested, and the model without it. This difference is compared to a chi-square (χ^2) distribution. The hypothesis $H_0 : \sigma^2_{cult} = 0$ was only rejected at the 14% level of significance ($\chi^2 = 1.2$) for RS and at the 0.2% level of significance ($\chi^2 = 8.2$) for DS. The hypothesis $H_0 : \sigma^2_{cult \times harvest} = 0$ was rejected at the 0.5% ($\chi^2 = 6.8$) and the 0.01% ($\chi^2 = 13.6$) levels of significance for RS and DS, respectively. These p-values were obtained by taking half of the probability of a greater χ^2 from a χ^2 distribution with one degree of freedom. It can be noted that likelihood ratio tests generated smaller p-values than Wald tests, but the discussion made above is still appropriate.

Since the hypothesis $H_0 : \sigma^2_{cult \times harvest} = 0$ was rejected, the classification of cultivars for RS and DS changes over different harvests. Consequently, the identification of the most productive cultivars depends on the evaluation during several harvests. It can be appropriately analyzed by a mixed model methodology using cultivars ranking based on EBLUP of genotypic values that consider all harvests to predict these effects.

Regression analyses were also performed to study the cultivars productivity over the years, since the interactions between cultivars and harvests were significant. According to a simple linear regression, it is important to verify: (a) the intercept value of the equation, which is related to the cultivar productivity; and (b) the angular coefficient, which is related to the persistence of the cultivar. Since all angular coefficients were negative (Table 4), the closer they are to zero the greater the persistence of cultivar. It is interesting to note that the cultivar 5715 showed a less pronounced reduction in dry matter production over the years, being a persistent and productive genotype in tropical regions (Fig. 2). Estimated DMY per harvest was 1385 kg/ha after 3.5 years of evaluation. On the other hand, cultivar MH 15 showed a sharper decrease in DMY, although its estimated production was 2883 kg/ha 42 days after the first harvest. Thus, observing the rank of alfalfa cultivars (Table 3) and the linear regression equations estimated (Table 4), it is possible to verify that the most productive cultivars are not always the more persistent.

The genetic variability was more expressive in DS; therefore, selection of superior genotypes based on DMY should consider EBLUP for this season. The first

climate in Brazil and did not find significant differences among cultivars for annual dry matter production. The results obtained for dry matter production per harvest during summer and winter indicated significant differences only for the 1992 winter. The authors used the ANOVA and classified cultivars as fixed effects. On the other hand, several experiments carried out in Minas Gerais, Brazil, showed significant differences among alfalfa cultivars when DMY per harvest was the evaluated trait (Souza-Sobrinho et al., 2004). These data were also analyzed using variances in a split plot design. Thus, genetic variability detection is influenced by the evaluated cultivars, the experimental precision, the genetic–environmental interaction and the statistical methodology used.

Table 4. Simple linear regression equations of dry matter production (y) on days after the first harvest (x), considering all harvests for 35 alfalfa cultivar evaluated at tropical environment

Cultirvar	Regression equation	Cultivar	Regression equation
MH 15	$\hat{y} = -1.3643x + 2740.4150$	SW 8112A	$\hat{y} = -1.0333x + 2391.6924$
5715	$\hat{y} = -0.7522x + 2353.6331$	WL 516	$\hat{y} = -0.9503x + 2319.0806$
SW 8210	$\hat{y} = -1.1681x + 2638.4204$	5929	$\hat{y} = -0.7759x + 2235.0113$
Rio	$\hat{y} = -1.3947x + 2693.0319$	Semit 711	$\hat{y} = -1.4648x + 2642.2494$
Alto	$\hat{y} = -1.2317x + 2622.3599$	BR 3	$\hat{y} = -1.1549x + 2542.6172$
5888	$\hat{y} = -0.9786x + 2486.5520$	BR 4	$\hat{y} = -1.1068x + 2395.8410$
Monarca	$\hat{y} = -1.2387x + 2689.1517$	Esmeralda	$\hat{y} = -1.0936x + 2429.0193$
Victoria	$\hat{y} = -1.2084x + 2627.4736$	ICI 990	$\hat{y} = -0.7140x + 2072.9148$
Florida 77	$\hat{y} = -1.1610x + 2556.7319$	F 708	$\hat{y} = -1.3376x + 2614.9212$
Falcon	$\hat{y} = -0.9901x + 2340.6077$	Semit 921	$\hat{y} = -0.7870x + 2174.7498$
BR 2	$\hat{y} = -1.2649x + 2698.9151$	Maricopa	$\hat{y} = -1.0984x + 2485.2811$
Alfa 200	$\hat{y} = -1.0435x + 2472.5285$	Crioula	$\hat{y} = -1.1873x + 2516.7548$
Valley Plus	$\hat{y} = -1.3230x + 2535.9575$	El Grande	$\hat{y} = -1.0827x + 2357.6762$
Araucana	$\hat{y} = -1.1960x + 2584.9869$	Sutter	$\hat{y} = -1.0832x + 2387.2934$
MH 4	$\hat{y} = -1.5345x + 2708.4661$	Costera	$\hat{y} = -1.6519x + 2656.4668$
SW9210 A	$\hat{y} = -1.0463x + 2496.4134$	P 30	$\hat{y} = -1.4726x + 2522.4590$
F 686	$\hat{y} = -1.4443x + 2690.1968$	P 205	$\hat{y} = -1.5802x + 2529.4666$
BR 1	$\hat{y} = -1.1601x + 2544.5187$		

ten cultivars with higher genotypic values were: MH 15, 5715, SW 8210, Rio, Alto, 5888, Monarca, Victoria, Florida 77 and Falcon. Crioula, the most common cultivar in Brazil, showed low forage potential in Sertãozinho. Results demonstrate that more productive alfalfa cultivars can be used in tropical pastures.

The use of multiple comparison tests is not appropriate when the number of treatments is greater than four, since they have low capacity to detect significant differences (Steel *et al.*, 1997; Resende and Duarte, 2007). Moreover, treatments should be treated as random effects. Thus, lower and upper confidence limits of the genotypic effects were calculated for each cultivar. Confidence limits are the lower and upper boundaries of a confidence interval, i.e. the values that define the range of a confidence interval (Snedecor and Cochran, 1989). The interval estimate gives an indication of how much uncertainty there is in the estimate of the true parameter. The narrower the interval is, the more precise the estimate.

This statistical approach, based on a mixed model methodology, has proven to be an interesting way to analyze data from perennial herbage-breeding programmes. It allows testing and selecting of appropriate covariance structure within cultivars and among harvests and also indicates superior cultivars that can be selected by cultivar effects of EBLUP. Moreover, it allows for analysis of the possible interactions between cultivars and harvests. Correlations between harvests were estimated, providing a better understanding of the relationship between production and the harvests during different seasons of the year.
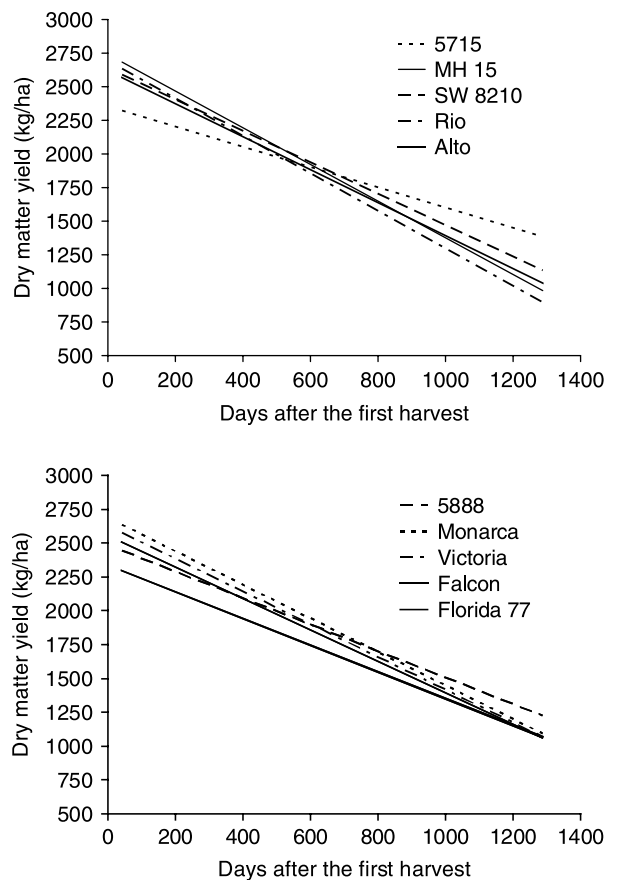
Fig. 2. Productivity (dry matter yield – kg/ha) over the years (42–1288 days after the first harvest) of ten alfalfa cultivars selected in dry season from EBLUP, considering all harvests.

Acknowledgements

References

Assis GML, Valentim JF, Carneiro Júnior JM, Azevedo JMA and Ferreira AS (2008) Selection of forage peanut genotypes for ground cover and aerial biomass production during establishment period using mixed model methodology. *Brazilian Journal of Animal Science* 37(11): 1905–1911.

Duarte JB and Vencovsky R (2001) Estimação e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. *Scientia Agricola* 58: 109–117.

Furlani RCM, Moraes MLT, Resende MDV, Furlani Júnior E, Gonçalves PS, Valério Filho WV and Paiva JR (2005) Estimation of variance components and prediction of breeding values in rubber tree breeding using the REML/BLUP procedure. *Genetics and Molecular Biology* 28: 271–276.

Henderson CR (1973) Sire evaluation and genetic trends. *Proceedings of Animal Breeding Genetic Symposium in Honor of Dr. J.L. Lush*. Champaign, IL: ASAS/ADSA, pp. 10–28.

Holland JB (2006) Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc Mixed. *Crop Science* 46: 642–654.

Huynh H and Feldt LS (1970) Conditions under which mean square rations in repeated measure designs have exact F-distributions. *Journal of American Statistical Association* 65: 1582–1589.

Littell RC, Milliken GA, Stroup WW and Wolfinger RD (1996) SAS® system for mixed models. Cary, NC: SAS Institute Inc., p. 633.

Littell RC, Henry PR and Ammerman CB (1998) Statistical analysis of repeated measures data using SAS procedures. *Journal of Animal Science* 76: 1216–1231.

Liu BH, Knapp SJ and Birkes D (1997) Sampling distributions, biases, variances and confidence intervals for genetic correlations. *Theorical and Applied Genetics* 94: 8–19.

Monteiro ALG, Costa C and Silveira AC (1998) Dry matter production and seasonal distribution and chemical composition of alfalfa cultivates (*Medicago sativa* L.). *Brazilian Journal of Animal Science* 27: 868–874.

Patterson HD and Thompson R (1971) Recovery of inter-block information when block size are unequal. *Biometrics* 58: 545–554.

Resende MDV (2002) *Genética biométrica e estatística no melhoramento de plantas perenes*. Brasília: Embrapa Informação Tecnológica, p. 975.

Resende MDV (2004) *Métodos estatísticos ótimos na análise de experimentos de campo*. Colombo: Embrapa Florestas, p. 65.

Resende MDV and Duarte JB (2007) Precisão e controle de qualidade em experimentos de avaliação de cultivares. *Pesquisa Agropecuária Tropical* 37(3): 182–194.

Resende MDV, Resende RMS, Jank L and Valle CB (2008) Experimentação e análise estatística no melhoramento de forrageiras. In: Resende MS, Valle CB and Jank L (eds) *Melhoramento de Forrageiras Tropicais*. Campo Grande, MS: Embrapa Gado de Corte, pp. 195–293.

Smith A, Cullis B and Gilmour A (2001) The analysis of crop variety evaluation data in Australia. *Australian and New Zealand Journal of Statistics* 43: 129–145.

Snedecor GW and Cochran W (1989) *Statistical Methods*. 8th ed. Ames, IA: Iowa State University Press, p. 491.

Souza-Sobrinho F, Ledo FJS, Pereira AV, Botrel MA, Evangelista AR and Viana MCM (2004) Estimates of repeatability for alfalfa dry matter production. *Ciência Rural* 34: 531–537.

Steel RGD, Torrie JH and Dickey DA (1997) Principles and procedures of statistics: a biometrical approach. New York: McGraw Hill Companies, p. 666.

Verbeke G and Molenberghs G (2000) *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics. New York: Springer, p. 568.

Welham S, Cullis B, Gogel B, Gilmour A and Thompson R (2004) Prediction in linear mixed models. *Australian and New Zealand Journal of Statistics* 46: 325–347.