Unmanned aircraft for monitoring elephant grass genotypes in energy biomass production

Ricardo Guimarães Andrade, Marcos Cicarini Hott, Walter Coelho Pereira de Magalhães Junior, Juarez Campolina Machado, Cristiano Amancio Vieira Borges

Brazilian Agricultural Research Corporation (Embrapa), Embrapa Dairy Cattle, Brazil

Abstract— *Elephant grass is a promising plant for economic and sustainable energy production. However, adapted cultivars and efficient strategies for selecting genotypes aimed at energy biomass production is essential. Remote sensing techniques provide spatiotemporal information from plants in an agile, non-destructive and non-invasive way. The present study aimed to use remote sensors onboard an unmanned aerial vehicle (UAV) to monitor elephant grass genotypes and assist in plant phenotyping for energy biomass production. The experimental plots were imaged in the visible and near infrared bands. Imaging was carried out in 66 experimental plots in the José Henrique Bruschi Experimental Field (CEJHB), located in Coronel Pacheco, MG, Brazil. The experiment was arranged in a randomized block design with three replications, and 22 elephant grass genotypes were evaluated. The aggregated index iMAP$_{NDRE}$ was strongly correlated with the dry matter production observed in the field, therefore a method with potential application for estimating the biomass of elephant grass genotypes. Thus, sensors aboard UAV platforms can assist breeders to select the best elephant grass genotypes for energy production.*

I. INTRODUCTION

The demand for sustainable and renewable energy sources that can be alternatives to fossil fuels has been growing in recent years [1]. The production of energy from plant biomass is one of the economically viable alternatives. In this case, elephant grass (Pennisetum purpureum Schum.) has very promising potential for energy production when compared to other energy biomass sources such as sugarcane and eucalyptus.

Some of the advantages of using elephant grass as bioenergy source are the great productive potential, rapid growth and short production cycle. However, there are challenges to be overcome such as the lack specific cultivars for energy production and efficient strategies for selection of genotypes for quality of biomass in energy use. The development of automated phenotyping tools, aiming to specific needs, can optimize the resources used in the selection and development of cultivars.

Automated data collection techniques and remote monitoring already contribute to boost agriculture and the rational use of natural resources [2, 3, 4]. The concept of smart farms has already been started, and devices such as Unmanned Aerial Vehicles (UAVs) have been helping farmers in activities including cargo transportation (fertilizers or pesticides), and monitoring of livestock and crops.

UAVs are platforms that embed sensors that provide a close panoramic view of the fields and are effective in generating data to extract knowledge or more accurate information about cropped areas to assist farmers in planning and decision-making [5].

Sensors that collect data in the visible range (RGB sensors) are the most common, but, they are cost-effective, as, in addition to the various spectral indices that can be generated, they have other applications such as a digital terrain model (MDT), 3D model, image orthomosaic, volume estimation and contour lines. On the other hand, sensors that capture data in the near infrared (700 to 1,100 nm) and thermal (5,000 to 12,000) ranges are very useful, for example, to identify vegetation stresses or plant that are more vigorous and excel in production and productivity.

The need for non-destructive, inexpensive, and large-scale experimentation makes remote sensing and data processing technologies fundamental to improving the performance and efficiency of plant phenotyping [6, 7]. Thus, remote sensing phenotyping methods have the advantage of gathering information from plants in a non-destructive and non-invasive way, both in space and time.

Recent technological advances have contributed to precision and high-throughput surveys to the benefit of large-scale field phenotyping [5]. From the foregoing, therefore, this study aimed to use remote sensors onboard an unmanned aerial vehicle (UAV) to monitor elephant grass genotypes and assist in plant phenotyping for energy biomass production.

II.   MATERIAL AND METHODS

Study area characterization

The study was conducted at the José Henrique Bruschi Experimental Field (CEJHB) of Embrapa Dairy Cattle (Figure 1), in Coronel Pacheco (MG), Brazil.

Based on the Köppen-Geiger climate classification, the study area is located in a transition zone of Aw climate (tropical climate with dry winter season) and Cwa (temperate humid climate with dry winter and hot summer). However, there is a predominance of Cwa in the region of the municipality where the meteorological station of the National Institute of Meteorology (INMET) is located. According to the INMET climatological normals from 1981 to 2010, the annual average air temperature is 21.4°C and the average annual rainfall

volume is 1620.6 mm. July (12.6 mm) and January (355.1 mm) have the lowest and highest rainfall, respectively.

In the municipality of Coronel Pacheco-MG, 10% of the area has flat relief, 10% mountainous relief and 80% wavy relief. The maximum and minimum altitudes are 1,070 m and 409 m, respectively. The municipal seat has an altitude of 484 m. The altitude of the area of the experiment is between 414 and 418 m (Figure 1) and a clayey-textured dystrophic yellow latosol predominates.

Installing the experiment and sample data collection

The experiment was arranged in a randomized block design with three replications, 22 elephant grass genotypes, 66 experimental plots standardized with three 4-m rows and plots spaced 1.2 m apart, net plot established at 3.6 m² (3 x 1.2 m) to minimize the plot border effect (Figure 1). The total area of the experiment is 0.6 hectares. The growth period and plant height varied according to cutting, with approximately 20 cm of residue.

Biomass (Dry Matter Yield – DMY) samples were collected in the experimental plots to estimate the productivity of each elephant grass genotype. Standardized cuttings were carried out in the net plots, at 20 cm height above ground level. Each sample was identified, ground and weighed individually to estimate the total fresh weight of the sample (FWS). Dry matter was estimated using a random sub-sample of fresh matter, which was labeled and weighed (fresh sub-sample weight – FSSW). The sub-samples were dried in an oven at 55°C for a minimum period of 72 hours, until all the moisture was removed. Then, the subsamples were weighed to obtain the dry matter weight (DMW). Afterwards, the percentage of dry matter was calculated for each sub-sample (%DM):

$$\%DM = DMW * 100 / FSSW$$

Next, the productivity per hectare was estimated:

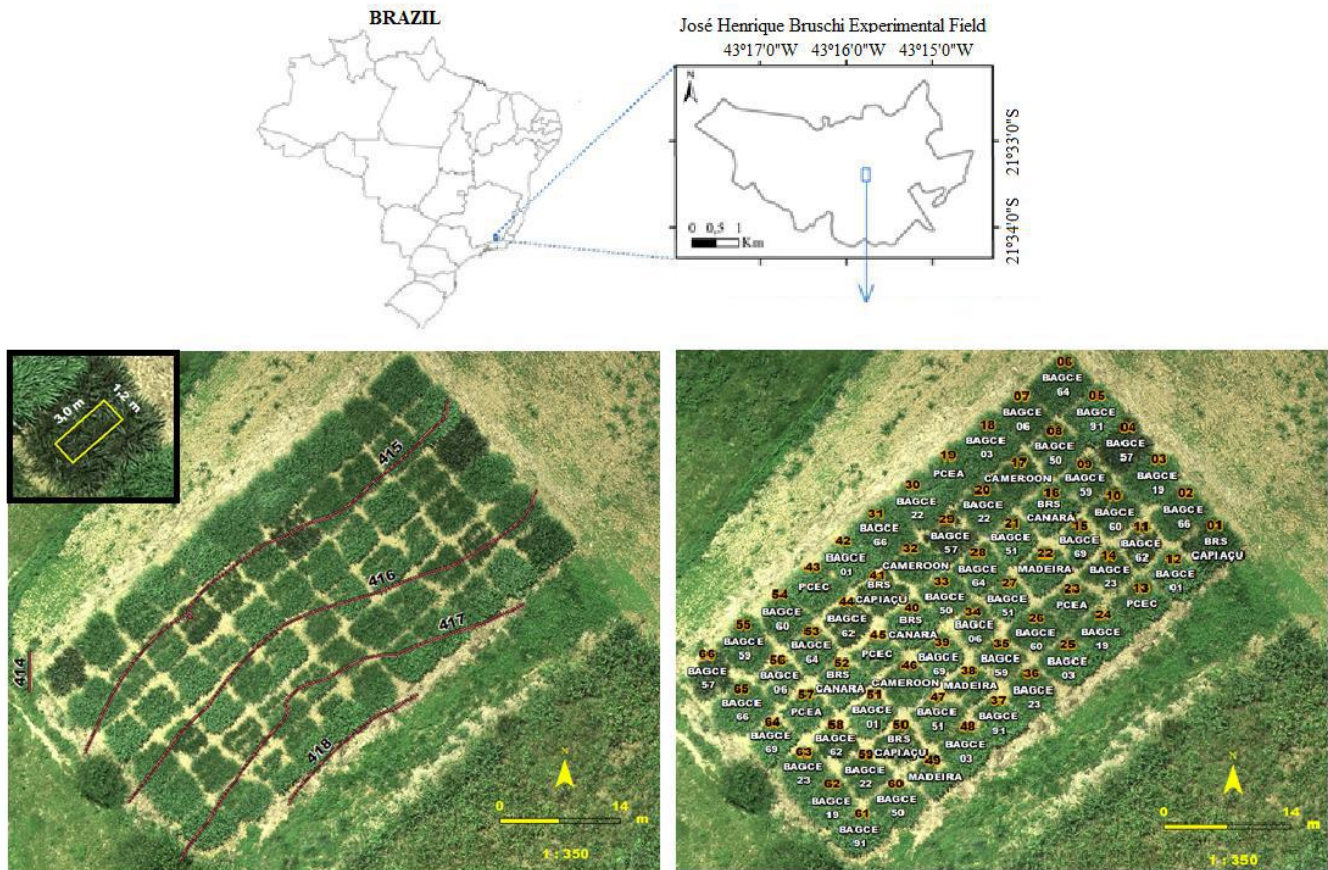$$DM \ (kg/ha) = FWS * \%DM * 10,000 / Net \ plot$$

Fig. 1: Location of the José Henrique Bruschi experimental field - Coronel Pacheco, Minas Gerais, Brazil. Contour curves and identification of the sixty-six elephant grass plots and net plot for the genotype evaluations.

Aerial surveys and vegetation indices

The aerial surveys were carried out on 02/26/19, 04/18/19 and 11/04/19 using the Inspire 1 Pro rotary-wing UAV, quadcopter, featuring a Sentera Multispectral Double 4K camera with gimbal for imaging in the visible and near infrared ranges, as described in Table 1.

Table 1: Specifications of each band of the spectral range were obtained with the use of Sentera Multispectral Double 4K Gimbaled imaging

Bands	Wavelength center (nm)	Band width (nm)
Blue	446	60
Green	548	45
Red	650	70
NIR	720	40
Red Edge	840	20

The flight plans were carried out according to technical compliance requirements, so that the results or products of the aerial survey could be compared on similar bases, equalizing variables such as flight height, pixel size of ground images (Ground Sample Distance - GSD), sensor calibration, percentage of image overlap, wind speed, brightness, shadow positioning, time of day, angle of view, sun position, etc.

The flight plan was parameterized as follows: (i) 70 m flight height; 2 cm GSD; 8 m/s maximum speed, 5 min flight time using battery; lateral and frontal overlap of the images 75% and 85%, respectively. Based on this flight plan configuration, it took 5 flight lines and 74 images to cover the entire area and generate the orthomosaic in the Pix4D Mapper Pro 4.125 software.

In the present study, the NDRE (Normalized Difference Red Edge [8]) and NDVI (Normalized Difference Vegetation Index [9]) were used according to the following equations:

$$NDRE = \frac{\rho_{nir} - \rho_{rededge}}{\rho_{nir} + \rho_{rededge}} \qquad \text{Eq. 1}$$

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$$

Eq. 2

Where ρ_{Green}, ρ_{Red}, ρ_{Blue}, $\rho_{RedEdge}$ and ρ_{NIR} are the spectral bands corresponding to the Green, Red, Blue, Red Edge, and near infrared (NIR) channels, respectively.

i-MulticriteriaAnalysisPlants (iMAP)

Throughout the experiment, the Embrapa Dairy Cattle Remote Sensing and Geoprocessing team used an aggregate index that allows analyzing the agronomic characteristics of any grouping of plants of the same genotype, which can be estimated from the images captured by the sensors embedded in the platform VANT. This index is based on the multicriteria analysis that calculates standardized anomalies between the agronomic characteristics of the canopy such as perimeter, area, height, volume, vigor, uniformity, fresh weight and dry weight. The index equation is given by:

$$iMAP_{IV} = \sum \left(\frac{CA - \bar{CA}}{\sigma_{CA}} \right)$$

Eq. 3

Where: iMAP is the aggregated index of multi-criteria analysis; IV is the designation of the vegetation index selected to assess the canopy vigor of the genotypes; CA is the agronomic characteristic selected for the composition of the aggregate index and calculated for each plot of the trial; $\bar{CA}$ is the mean of the agronomic characteristic distribution for the 66 plots of the trial; σ_{CA} is the standard deviation of the agronomic characteristic distribution for the 66 plots of the trial.

The aggregated indices iMAP$_{NDVI}$ and iMAP$_{NDRE}$ were generated for the NDVI and NDRE vegetation indices, respectively. They express the association of vegetative vigor with the average volume reached by the genotype. Then, correlations between biomass and the aggregated indices iMAP$_{NDVI}$ and iMAP$_{NDRE}$ were examined.

III.   RESULTS AND DISCUSSION

Figure 2 shows the spatiotemporal analysis of the elephant grass experimental plots. It shows heterogeneity in terms of growth and vigor of the 22 genotypes distributed in the 66 plots.

It is visible the variability in the shades of green and the exposure of background soil or faults in some plots in the RGB mosaics (Figures 2A and 2D). At the same time, in the images with estimates of NDVI (Figures 2B and 2E) and NDRE (Figures 2C and 2F), the variations in the

indices both within each plot and between plots were evident. These variations may be related to differences in vigor and soil faults or exposure within each net plot that may be imperceptible to an observer in the field. In this case, the influence of decaying organic matter may have occurred along the edges of the planting area (border effect).
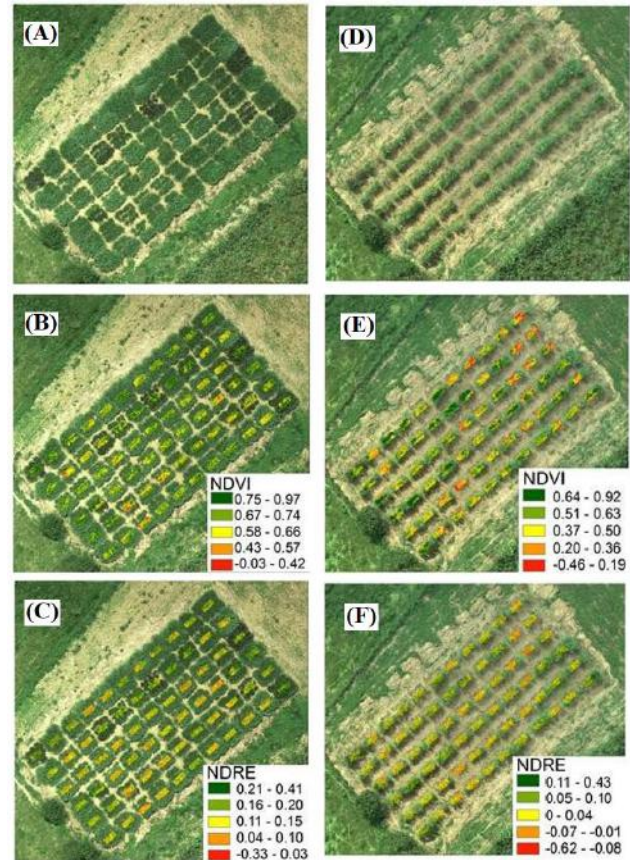

Fig. 2: Monitoring of elephant grass experimental plots using RGB and near infrared (NIR) sensors onboard unmanned aerial vehicle (UAV) to obtain RGB mosaics (A and D) and the vegetation indices NDVI (B and E) and NDRE (C and F) of the net plots. Image dates: 02/26/2019 (Figures A, B, C) and 04/18/2019 (Figures D, E, F).

Analysis of data distribution

Pearson's linear correlation coefficient assumes the existence of a linear relationship between variables, in a way that it can be used to estimate the degree of intensity or strength of this relationship. The existence of data clusters, outliers, sharp left or right asymmetries and other abnormalities must be investigated in advance and, if possible, mitigated. Generally, asymmetry, whether milder or more pronounced, is the problem that is most observed for variables of different natures, and often the use of some transformation, e.g., logarithmic, is enough to make the

distribution approximately symmetrical and favor its use in any statistical analyses.

The variable dry matter yield (kg), which was estimated by the traditional field method and indicates the biomass produced by the genotype, and the variable volume (m³) of the plants showed positive asymmetry with an extension of the right tail, with this deviation being detected by the Shapiro-Wilk test (Figures 3A and 3C, $p < 0.0001$). In both cases, the problem was mitigated by the logarithmic transformation, which made the distributions visibly more symmetrical, as evidenced by the histograms in Figures 3B and 3D and by the descriptive levels (p-values) that reached values greater than 0.19.
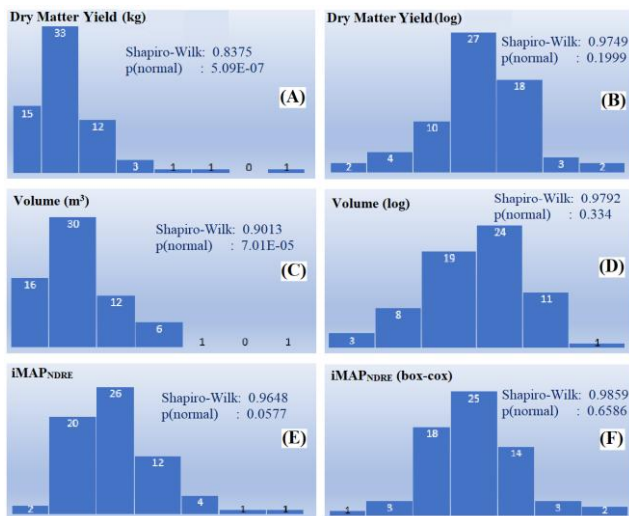

Fig. 3: Histograms of frequency distribution of dry matter yield (A), volume (C) and aggregated index iMAP_NDRE (E), as well as their forms transformed by decimal logarithm (B, D) or by Box-Cox power transformation (F), followed by the Shapiro-Wilk Normality Test.

The aggregated index iMAP_NDRE showed a slight positive asymmetry, which would not be sufficient to reject the distribution normality hypothesis (Figure 3E; p=0.0577). Nevertheless, to maximize as much as possible the linearity between this variable and any other in the study, its values were also transformed by the Box-Cox Transformation, making the distribution even more symmetrical (Figure 3F; p=0.6586). On the other hand, the distributions of the NDRE and NDVI vegetation indices, as well as the aggregate iMAP_NDVI index, showed only slight asymmetries (Figure 4) and no strong deviations from normality (p > 0.0809), ruling out the need for a transformation. Although, in general, all correlations can be refined, specifically from a greater number of aerial surveys, as expected, the iMAP_NDRE index showed the

highest correlation with dry matter yield or biomass produced by elephant grass genotypes (Figure 5). This index is a combination of the volume estimates and the NDRE. In addition, this combination enabled better classification of elephant grass genotypes.
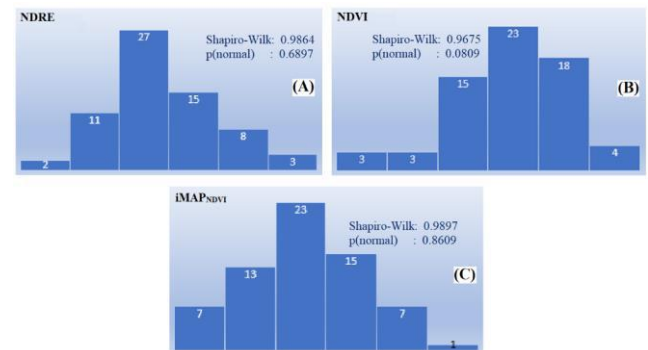

Fig. 4: Histograms of frequency of NDRE (A) and NDVI (B) vegetation indices and the aggregated iMAP_NDVI (C) index, followed by the Shapiro-Wilk Normality Test.
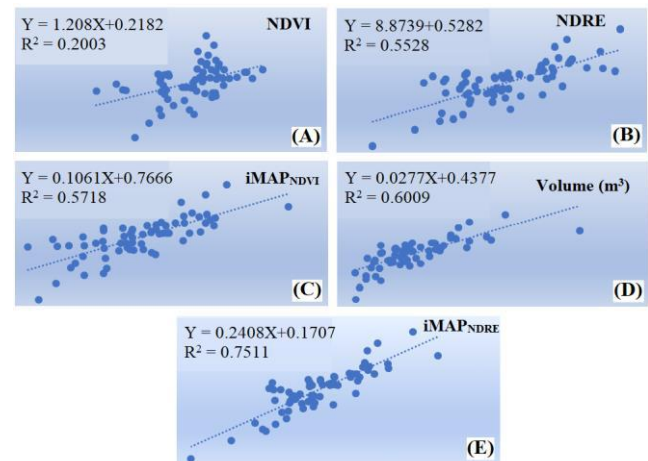

Fig. 5: Scatter plots of the relationship between biomass observed in the field with NDVI (A), NDRE (B), iMAP_NDVI (C), volume (D), iMAP_NDRE (E) data.

The application of the Scott and Knott's (1974) grouping method to the iMAP_NDRE aggregate index data, with a significance level of 5%, allows the discrimination of eight genetic materials (Table 2); and, if biomass (DMY) is used for selection, also at 5% significance level, virtually the same materials are indicated. This agreement of results indicates that, for the samples of materials involved in this study, the aggregated index iMAP_NDRE proved to be very useful and effective for selection aiming at the production of biomass.

Table 2: Clustering of means by the Scott and Knott method, α = 0.05, for the variables iMAP$_{NDRE}$ and dry matter yield (DMY), emphasizing the selection of the best materials. Standard error of the mean: 0.9 (iMAP$_{NDRE}$) and 1.9 (DMY)

Selection rank	Genotype	iMAP$_{NDRE}$			DMY		
		Mean	Cluster	Rank	Mean	Cluster	Rank
1	BAGCE 19	3.9	a	*1*	10.9	a	*3*
2	Cameroon-BAG 38	2.5	a	*2*	9.3	a	*4*
3	BAGCE 66	2.1	a	*3*	11.3	a	*2*
5	BRS Capiaçu	1.8	a	*4*	11.6	a	*1*
4	BAGCE 57	1.7	a	*5*	9.2	a	*5*
5	BAGCE 23	1	a	*6*	8.6	a	*6*
6	BAGCE 3	0.7	a	*7*	7.5	b	*7*
7	BAGCE 62	0.5	a	*8*	6.4	b	*10*
	BAGCE 60	0.2	b	*9*	7	b	*8*
	BRS Canará	-0.1	b	*10*	6	b	*13*
	BAGCE 91	-0.2	b	*11*	6	b	*12*
	BAGCE 69	-0.5	b	*12*	6.3	b	*11*
	BAGCE 1	-0.7	b	*13*	6.5	b	*9*
	PCEC	-0.8	b	*14*	4.3	b	*19*
	BAGCE 6	-0.8	b	*15*	5.2	b	*16*
	BAGCE 22	-0.9	b	*16*	5.9	b	*14*
	Madeira	-1	b	*17*	4.5	b	*18*
	BAGCE 64	-1.4	b	*18*	5.8	b	*15*
	BAGCE 51	-1.7	b	*19*	4.7	b	*17*
	PCEA	-1.7	b	*20*	3.7	b	*20*
	BAGCE 59	-2.2	b	*21*	3.3	b	*21*
	BAGCE 50	-2.3	b	*22*	2.7	b	*22*

Figure 6 describes, in a synthetic way, the data collected for the variables iMAP$_{NDRE}$ and biomass (DMY), allowing an easy comparison between the performances of each genotype. For each genetic material, the boxplot (box and whisker) allows the exact identification of each of the three replications: the minimum of the three values is limited by the lower whiskers; the maximum of them is limited by the upper whiskers; and the remaining value, intermediate between these two, and corresponding to the median of the three, is represented by the thickest horizontal line. Additionally, the set of means, designated by the central points (solid, filled circle), shows the differences between the genetic materials and the decreasing trend between the best-ranked genotype and the one that achieved the worst performance in the experiment. It is noteworthy that, to allow the juxtaposition of the distributions of the two variables of

different natures - iMAP$_{NDRE}$ and DMY (original scales on the right side) - the variables were standardized by subtracting the mean and dividing by standard deviation, to be on a similar scale, having mean 0 and standard deviation 1 (represented by the ordinate axis on the left side).

It is also found that, generally, the eight materials grouped by Scott-Knott as the best based on iMAP$_{NDRE}$ – in addition to BAGCE 60 ranked ninth – were precisely those with a pair of means equal to or higher than the overall means of the two characteristics, iMAP$_{NDRE}$ and DMY, estimated at 0 (s/unit) and 6.66 kg/ha, respectively. In addition, the best of the entire experiment – BAGCE 19 – showed mean around two standard deviations above the overall mean found for iMAP$_{NDRE}$ and achieved mean

above one standard deviation above the estimated overall   mean for biomass.

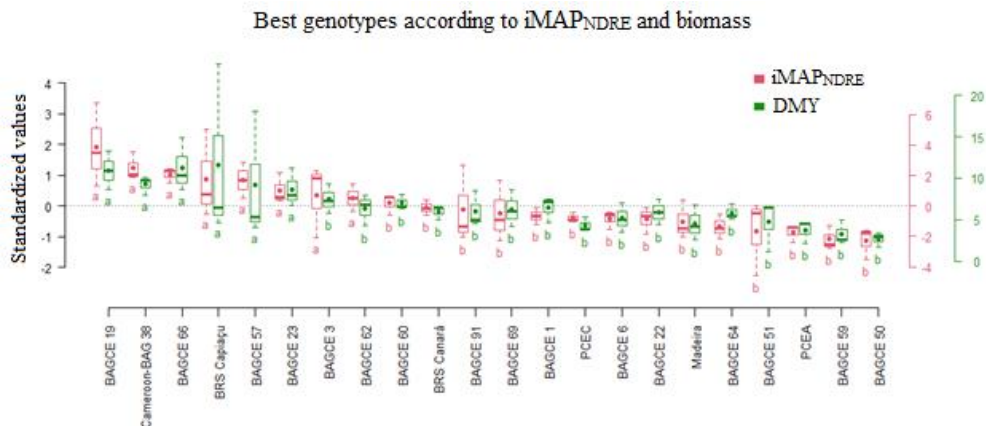Best genotypes according to iMAP$_{NDRE}$ and biomass

Fig. 6: Boxplot of frequency distributions per genetic material observed for the variables iMAP$_{NDRE}$ and energy biomass (DMY) and arranged in descending order of mean for iMAP$_{NDRE}$. The mean and median of each distribution are represented by the central solid point and the thickest horizontal line, respectively, while the whiskers limit the minimum and maximum values observed for the genotype. Means of the same variable followed by different letters were grouped into distinct clusters by the Scott-Knott method ($\alpha = 0.05$). Variables were standardized (subtracting the mean and dividing by standard deviation) before the juxtaposition, in the figure, with the overall means of the characteristics delimited by the dotted horizontal line.

IV.   CONCLUSION

The aggregated index iMAP$_{NDRE}$ showed a strong correlation with the dry matter production observed in the field and has potential application for estimating the biomass of elephant grass genotypes. Thus, the use of sensors onboard UAV platforms can help breeders to select the best elephant grass genotypes for energy use.

REFERENCES

[1] Morais, R. F., Souza, B. J., Leite, J. M., Soares, L. H. B., Alves, B. J. R., Boddey, R. M., Urquiaga, S. (2009). Elephant grass genotypes for bioenergy production by direct biomass combustion. Pesquisa Agropecuária Brasileira, 44(2), pp. 133-140.

[2] Tripicchio, P., Satler, M., Dabisias, G., Ruffaldi, E., Avizzano, C. A. (2015). Towards Smart Farming and Sustainable Agriculture with Drones. INTERNATIONAL CONFERENCE ON INTELLIGENT ENVIRONMENTS. IEEE Computer Society, pp. 140-143.

[3] Ribeiro-Gomes, K., Hernandez-Lopez, D., Ballesteros, R.; Moreno, M. A. (2016). Approximate georeferencing and automatic blurred image detection to reduce the costs of UAV use in environmental and agricultural applications. Biosystems Engineering, 151, pp. 308-327.

[4] Zhang, D., Zhou, X., Zhang, J., Lan, Y., Xu, C., Liang, D. (2018). Detection of rice sheath blight using an unmanned aerial system with high-resolution color and multispectral imaging. PLOS ONE, 13 (5), e018747, pp. 1-14.

[5] Andrade, R. G., Hott, M. C., Magalhães Junior, W. C. P., Oliveira, P. S. d'., Oliveira, J. S. (2019). Monitoring of Corn Growth Stages by UAV Platform Sensors. International Journal of Advanced Engineering Research and Science, 6, pp. 54-58.

[6] Menzel, M. I., Tittmann, S., Bühler, J., Preis, S., Wolters, N., Jahnke, S., Walter, A., Chlubek, A., Leon, A., Hermes, N., Offenhäuser, A., Gilmer, F., Blümler, P., Schurr, U., Krause, H. (2009). Non-invasive determination of plant biomass with microwave resonators. Plant, Cell & Environment, 32(4), pp. 368-379.

[7] Sousa, C. A., Cunha, B. A. D. B., Martins, P. K., Molinari, H. B. C., Kobayashi, A. K., Souza Junior, M. T. (2015). Nova abordagem para a fenotipagem de plantas: conceitos, ferramentas e perspectivas. Revista Brasileira de Geografia Física, 8 (número especial), pp. 660-672.

[8] Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C.; Riley, E., Thompson, T., Lascano, R. J., Li, H., Moran, M. S. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In: International Conference on Precision Agriculture, 5., 2000, Madison. Proceedings. Madison: WI 53711, USA. 15p.

[9] Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. In: Earth resources technology satellite-1 Symposium, 3., 1973, Greenbelt. Proceedings… Greenbelt: NASA SP-351 I, pp. 309-317.