








Dataset: Forty soybean cultivars from subsequent harvests

Bruno Rodrigues de **Oliveira**^{1*} , Alan Mario **Zuffo**^{2,*} , Francisco Charles dos Santos **Silva**² , Ricardo **Mezzomo**² , Leandra Matos **Barrozo**² , Tatiane Scilewski da Costa **Zanatta**² , Joel Cabral dos **Santos**² , Carlos Henrique Conceição **Sousa**², Yago Pinto **Coelho**², Aurilucia do Nascimento Silva **Caldas**²

¹ Editora Pantanal, Nova Xavantina, MT, Brazil.

² State University of Maranhão, Balsas, MA, Brazil.

* Correspondences: bruno@editorapantanal.com.br; alan_zuffo@hotmail.com

Received: 2023-10-29

Accepted: 2023-10-30

Published: 2023-11-08

Main Editors

Jorge González Aguilera



Copyright: © 2023. Creative Commons Attribution license: [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/).

For citation: Oliveira, B. R.; Zuffo, A. M.; Silva, F. C.; Mezzomo, R.; Barrozo, L. M.; Zanatta, T. S.; Santos, J. C.; Sousa, C. H.; Coelho, Y. P.; Caldas, A. d. (2023). Dataset: Forty soybean cultivars from subsequent harvests. Trends in Agricultural and Environmental Sciences, (e230005), DOI: 10.46420/TAES.e230005

Abstract: Soybean cultivation is one of the most important because it is used in several segments of the food industry. The evaluation of soybean cultivars subject to different planting and harvesting characteristics is an ongoing field of research. In this paper, we present a dataset obtained from forty soybean cultivars planted in subsequent seasons. The experiment used randomized blocks, arranged in a split-plot scheme, with four replications. The following variables were collected: plant height, insertion of the first pod, number of stems, number of legumes per plant, number of grains per pod, thousand seed weight, and grain yield, resulting in 320 data samples. The dataset presented can be used by researchers from different fields of activity.

Keywords: *Glycine max* L.; grain productivity; sowing times; soybeans.

1. Introduction

Soybeans (*Glycine max* L.) are one of the main agricultural crops in the world, with a current production estimated at 399.50 million tons, according to the USDA. The three largest producers are Brazil, the United States and Argentina, representing approximately 41%, 28% and 12% of global production, respectively. Figure 1 shows this distribution of cultivated areas.

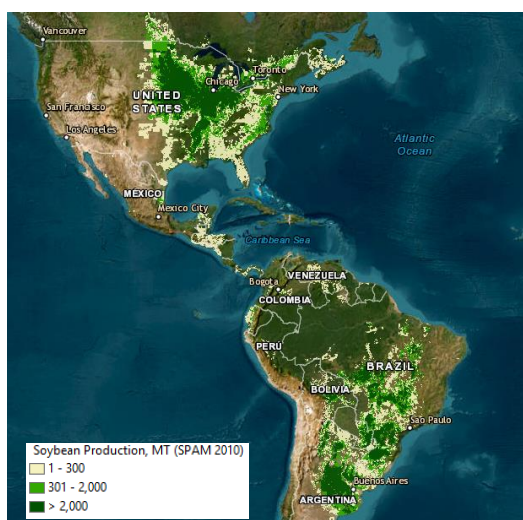


Figure 1. Distribution of soybean cultivation in the three main producing countries. Source: USDA.

Soybean is used for various purposes, such as human and animal food and biofuel production. Most of the soybeans produced in the world are used to make meal and oil. The bran, which is



rich in protein, is mainly used in the poultry, pig and cattle feed industry. The oil is used as a raw material by the industry for the production of refined oil, hydrogenated fats, margarine, and mayonnaise, among other products. Furthermore, soybean oil has become the main raw material for biodiesel production.

2. The experiment

2.1 Location and characterization of the experimental area

The experiment was carried out in the field at farm ‘Pequizeiro’, at the Experimental Station of “Accert Pesquisa e Consultoria Agronomica”, located 10 km from the municipality of Balsas, MA, Brazil (Figure 2), with a latitude of 07°31’57” S, longitude 46°02’08” W and an altitude of approximately 283 m. It was implemented in the 2022/2023 harvest.

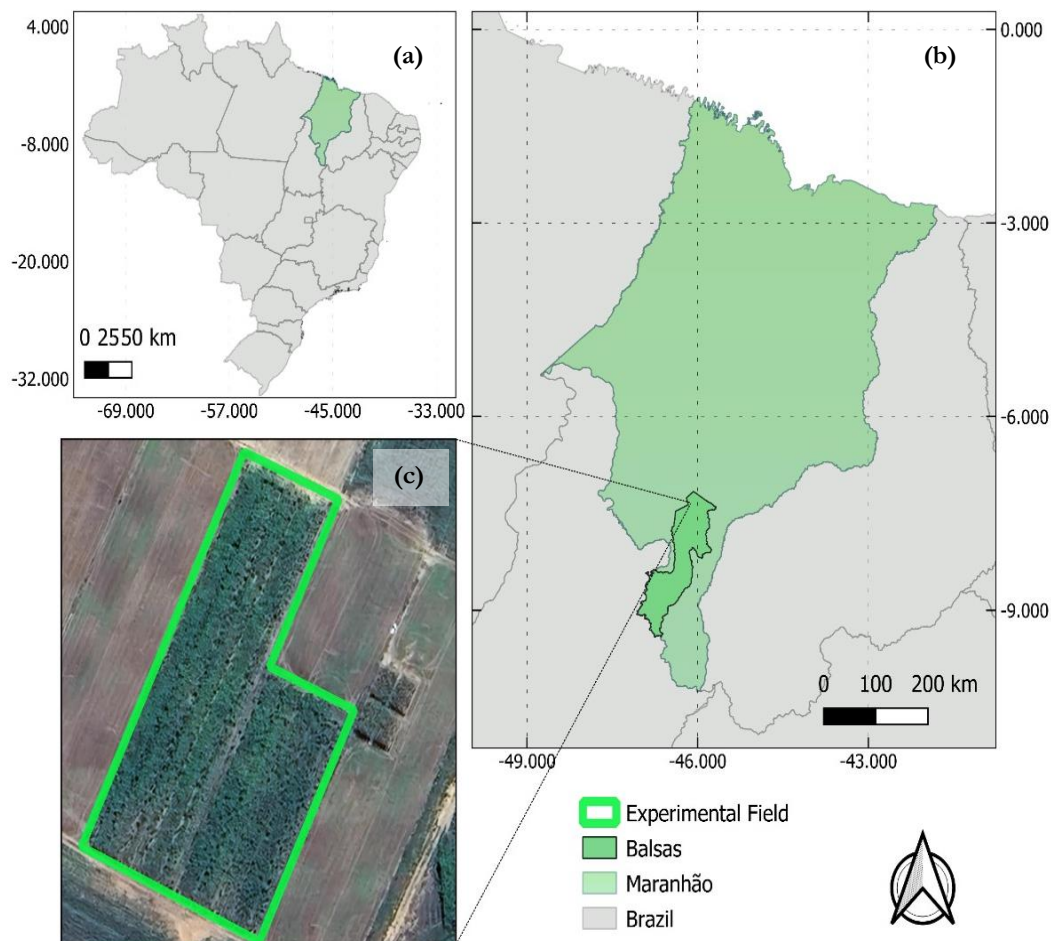


Figure 2. Maps of Brazil (a), Maranhão State (b) and the location of the experimental area (c).

The region's climate, according to Köppen's classification, is hot and humid tropical (Aw), with rainy summers and dry winters (Maranhão, 2002). The total annual rainfall reaches average values of 1175 mm (Passos et al., 2017). Precipitation data during the experiment are shown in Figure 3.

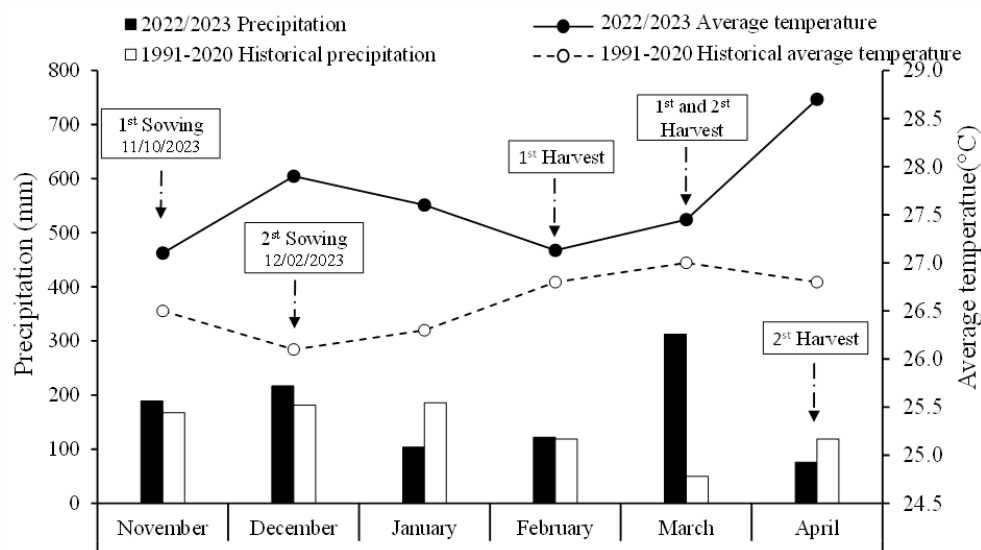


Figure 3. Average monthly precipitation and temperature and history during the execution of the experiment in the year 2022/2023 and the averages for the years 1991 to 2020 for the city of Balsas-MA. Source: Accert (2023) and National Institute of Meteorology (2023).

The soil in the experimental area was classified as a Yellow Oxisol with a sandy texture based on the Brazilian soil classification system (SANTOS et al., 2018). Before starting the experiment, the soils were sampled in the 0-20 cm and 20-40 cm layers. The chemical and physical properties of the analyzed soils are presented in Table 1.

Table 1. Main chemical properties of the soils used in the experiment.

Depth cm	pH	OM	P _{Mehlich-1}	H+Al	Al ³⁺	Ca ²⁺	Mg ²⁺	K ⁺	CEC	B
	H ₂ O	dag/kg	mg dm ⁻³			cmol _c dm ⁻³				%
0-20	6.00	1.29	54.95	1.20	0.01	2.15	0.71	136.00	4.41	72.78
20-40	4.65	0.23	20.72	1.80	0.54	0.95	0.30	70.00	3.23	44.26
	B	Cu	Fe	Mn	Zn	S	TOC	Clay	Silt	Sand
	mg dm ⁻³					dag/kg		%		
0-20	0.22	0.44	113.21	14.28	0.73	6.30	0.75	24.24	9.26	66.49
20-40	0.23	0.40	81.98	4.25	0.37	12.60	0.13			

OM: organic matter. CEC: cation exchange capacity at pH 7.0. B: base saturation. TOC: total organic carbon.

2.2 Experimental design and treatments

The experimental design used was randomized blocks, arranged in a split-plot scheme, with four replications. In the plots, the treatments consisted of two sowing times (season 1: 10/11/2023, season 2: 04/12/2023), and in the subplots, there were 40 soybean cultivars (Table 2). Each experimental unit consisted of eight rows spaced 0.50 m apart and 10 m long, totaling an area of 40 m². As a useful area, the four central lines were considered, with 1 m being ignored at each end, totaling an area of 16 m².

Table 2. Characteristics of the soybean cultivars evaluated in the experiment.

Cultivars	Maturation group	Seeds per meter/linear	Density per meter/linear
FTR 3190 IPRO	9.0	12.5	250.000
FTR 4288 IPRO	8.8	11	220.000
NK 8770 IPRO	8.7	16	320.000
M 8606I2X	8.6	10	200.000
M 8644 IPRO	8.6	11	220.000
ADAPTA LTT 8402 IPRO	8.4	13	260.000
98R30 CE	8.3	20	400.000
FORTALEZA IPRO	8.3	13	260.000
MONSOY 8330I2X	8.3	14	280.000

Cultivars	Maturation group	Seeds per meter/linear	Density per meter/linear
SUZY IPRO	8.3	11	220.000
TMG 22X83I2X	8.3	14	280.000
EXPANDE LTT 8301 IPRO	8.3	14	280.000
FORTALECE L090183 RR	8.3	14	280.000
83IX84RSF I2X	8.3	15	300.000
82HO111 IPRO - HO COXIM IPRO	8.2	15	300.000
82I78RSF IPRO	8.2	18	360.000
SYN2282IPRO	8.2	18	360.000
ATAQUE I2X	8.1	14	280.000
NK 8100 IPRO	8.1	16	320.000
FTR 4280 IPRO	8.0	11	220.000
LYNDA IPRO	8.0	11	220.000
BRASMAX OLÍMPO IPRO	8.0	16	320.000
LAT 1330BT.11	8.0	17	325.000
FTR 3179 IPRO	7.9	15	300.000
97Y97 IPRO	7.9	19	380.000
BRASMAX BÓNUS IPRO	7.9	16	320.000
PAULA IPRO	7.9	14	280.000
NEO 790 IPRO	7.9	16	320.000
LTT 7901 IPRO	7.9	16	305.000
GNS7900IPRO - AMPLA	7.9	19	380.000
79I81RSF IPRO	7.9	14	280.000
ELISA IPRO	7.7	13	260.000
NK 7777 IPRO	7.7	17	340.000
77HO111I2X - GUAPORÉ	7.7	18	360.000
GNS7700IPRO	7.7	18	360.000
FTR 3868 IPRO	7.6	15	300.000
MANU IPRO	7.6	14	280.000
NEO 760 CE	7.6	20	400.000
74K75RSF CE	7.4	20	400.000
96R29 IPRO	6.2	21.9	438.000

2.3 Implementation, fertilization and conduction of experiments

Desiccation was carried out using the products glyphosate + Haloxyfope-P-methyl. After 15 days, soybean was sown mechanically using a fertilizer seeder with a rod-type furrowing mechanism (machete) for nil-tillage systems at a depth of approximately 3 cm, with a spacing of 0.50 m, and the quantity of seeds varied depending on recommendation for each cultivar.

Base fertilization consisted only of the application of 200 kg ha⁻¹ of monoammonium phosphate (MAP). At 30 DAE of soybean plants, 120 kg ha⁻¹ of K₂O was applied, the source of which was potassium chloride.

Soybean seeds were treated with pyraclostrobin + methyl thiophanate + fipronil at a dose of 2 mL p.c. kg⁻¹ of seed inoculated with *Bradyrhizobium japonicum*. The commercial liquid inoculant Simbiose Nod Soja[®] was used at a dose of 150 mL for 50 kg of seeds.

During plant development, the following products were used to manage weeds, pests and diseases: glyphosate, haloxyfop-p-methyl, pyraclostrobin + epoxiconazole, picoxystrobin + benzovindiflupyr, mancozeb, azoxystrobin + cyproconazole, teflubenzuron, chlorpyrifos, cypermethrin and imidacloprid + beta-cyfluthrin.

2.4 Variables evaluated

At harvest (R_8 stage), the following variables were obtained from 10 plants per plot:

- **PH:** plant height (cm) – determined from the soil surface to the insertion of the last leaf using a millimeter ruler;
- **IFP:** insertion of the first pod (cm) – determined from the soil surface to the insertion of the first vegetable;
- **NS:** Number of stems (unit) – through manual counting;
- **NLP:** Number of legumes per plant (unit) – through manual counting;
- **NGP:** Number of grains per plant (unit) – through manual counting;
- **NGL:** Number of grains per pod (unit) – through manual counting;
- **TSW:** Thousand seed weight (g) – according to the methodology described in Brasil (2009);
- **GY:** Grain yield (kg ha^{-1}) – determined by harvesting the useful area of the plot and standardized to a grain moisture level of 13%.

2.5 Experimental field illustration

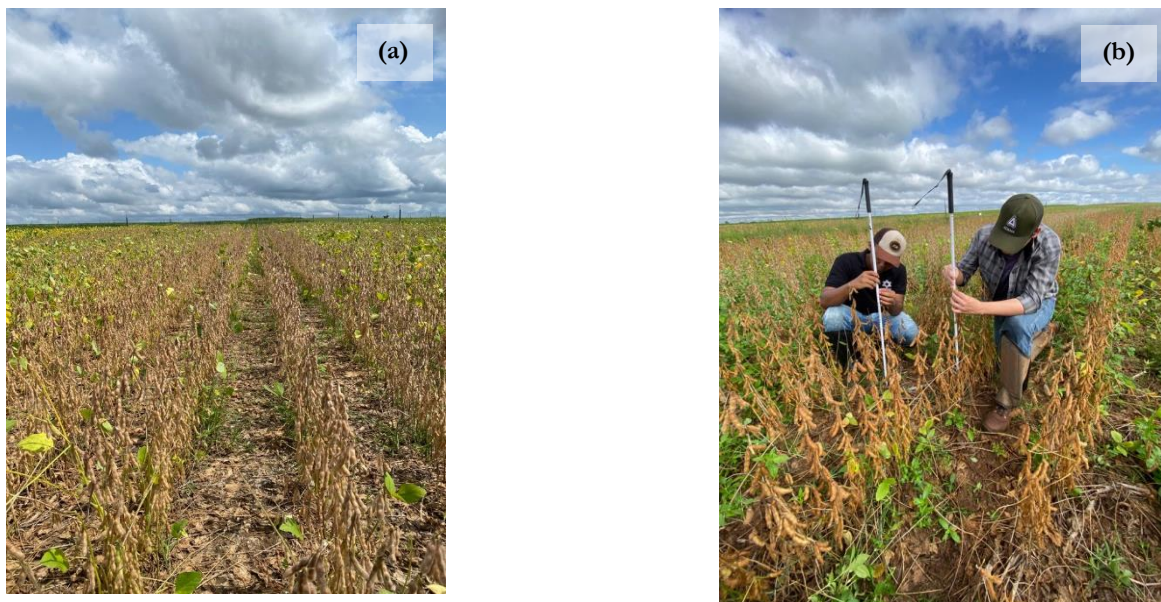


Figure 4. (a) Field experiment showing the harvest at the R_8 stage. (b) Field experiment showing researchers calculating the plant height.

3. Data Description

For each cultivar, there were 4 replications in each harvest season. Therefore, there are $40 \times 4 \times 2 = 320$ data samples. The data are organized in tabular form. The columns are “Season”, “Cultivar”, “Repetition”, “PH”, “IFP”, “NLP”, “NGP”, “NGL”, “NS”, “TSW”, and “GY”. Table 3 shows the data types. Table 4 shows an example of the dataset for two cultivars only. Table 5 provides a statistical summary of the dataset.

Table 3. Data types and values assumed in the dataset.

Column	Data type	Values
Season	Integer	1 or 2
Cultivar	Categorical	Cultivar names according to Table 2.
Repetition	Integer	1, 2, 3 or 4
PH, IFP, NLP, NGP, NGL, NS, TSW, GY	Continuous	Ranging from 0.40 to 4930.00

Table 4. Example of dataset.

Season	Cultivar	Repetition	PH	IFP	NLP	NGP	NGL	NS	TSW	GY
1	NEO 760 CE	1	58.80	15.20	98.20	177.80	1.81	5.20	152.20	3232.82
1	NEO 760 CE	2	58.60	13.40	102.00	195.00	1.85	7.20	141.69	3517.36
1	NEO 760 CE	3	63.40	17.20	100.40	203.00	2.02	6.80	148.81	3391.46
1	NEO 760 CE	4	60.27	15.27	100.20	191.93	1.89	6.40	148.50	3312.58
1	MANU IPRO	1	81.20	18.00	98.80	173.00	1.75	7.40	145.59	3230.99
1	MANU IPRO	2	75.80	20.80	69.20	128.00	1.85	7.20	154.87	3374.80
1	MANU IPRO	3	84.40	15.80	95.40	161.80	1.70	6.80	150.23	3182.76
1	MANU IPRO	4	80.47	18.20	87.80	154.27	1.77	7.13	149.90	3165.72

Table 5. Dataset summary.

	Season	Repetition	PH	IFP	NLP	NGP	NGL	NS	TSW	GY
Mean	1.5	2.5	68.38	15.46	59.08	135.08	2.29	4.07	168.32	3418.55
Std.	0.5	1.1	8.95	3.02	20.06	60.49	0.84	1.47	19.62	503.00
Min	1.0	1.0	47.60	7.20	20.20	47.80	0.94	0.40	127.06	1538.23
25%	1.0	1.7	62.95	13.60	44.35	95.05	2.00	3.00	153.84	3126.61
50%	1.5	2.5	67.20	15.60	54.50	123.00	2.28	3.80	166.15	3397.27
75%	2.0	3.2	74.34	17.33	71.22	161.35	2.48	5.00	183.18	3708.26
Max	2.0	4.0	94.80	26.40	123.00	683.40	14.86	9.00	216.00	4930.00

Std. means standard deviation.

4. Availability

The database is available in.csv format (separated by a comma) and in a Colab Jupyter Notebook, with instructions in the Python (Van Rossum and Drake, 1995) Pandas (McKinney, 2010) library for reading the file (on GitHub only). The dataset is available at

- <https://archive.ics.uci.edu/dataset/913/forty+soybean+cultivars+from+subsequent+harvests>
- <https://github.com/brunobro/dataset-forty-soybean-cultivars-from-subsequent-harvests>

5. References

MARANHÃO - GOVERNO DO ESTADO DO MARANHÃO. Gerência de Planejamento e Desenvolvimento Econômico - GEPLAN. Atlas do Maranhão. São Luís: Universidade Estadual do Maranhão, 2002. p.39.

McKinney, W. (2010, June). Data structures for statistical computing in Python. In Proceedings of the 9th Python in Science Conference (Vol. 445, No. 1, pp. 51-56).

Passos, M. L.; Zambrzycki, G. C.; Pereira, R. S. Balanço hídrico climatológico e classificação climática para o município de Balsas-MA. Revista Scientia Agraria, v.18, n.1, p.83-89, 2017. doi: 10.5380/rsa.v18i1.48584

Santos, H. G. et al. Sistema brasileiro de classificação de solos, Embrapa, ed. 5, 2018.

Van Rossum, G., Drake Jr, F. L. (1995). Python tutorial. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.

USDA. <https://ipad.fas.usda.gov/cropeplorer/cropview/commodityView.aspx?cropid=2222000>

6. Additional Information

6.1 Acknowledgments

The authors would like to thank “Accert Pesquisa e Consultoria Agronomia” for financial support and the availability of the experimental area.

6.2 Funding

There was no funding for this research.

6.3 Conflicts of Interest

We declare that there are no conflicts of interest.