Optimization of maize yield by fuzzy regression

Alexander N. Sadovski
Agricultural academy, Institute of Soil Science, Agrotechnologies and Plant Protection "Nikola Poushkarov", Bulgaria
E-mail: *bsc.ieas@yahoo.com*

Citation
Sadovski, A. N. (2023). Optimization of maize yield by fuzzy regression. *Bulgarian Journal of Crop Science, 60*(1) 84-89.

Abstract
The purpose of the present study is to determine the optimal values of the main nutrients nitrogen, phosphorus, potassium, and silicon with the help of the theory of fuzzy sets in the conditions of a field experiment with maize on leached smolnitsa and alluvial-meadow soil. The design of multifactorial experiments allows the assessment of actions and interactions of four factors, varying on three levels. The fuzzy set theory provides a strict mathematical framework in which vague conceptual phenomena can be precisely and rigorously studied. The use of fuzzy regression techniques is appropriate in field experiments when the results obtained are influenced by multiple random factors during the growing season. Statistical analysis of data from two years (2020 and 2022) establishes trends in maize nutrition.

Keywords: field experiment; maize; yield; fuzzy regression

INTRODUCTION

Field experiments are studies using an experimental design that occur in a natural setting. When researchers conduct experiments, they study how the manipulation of independent variables, or variables that remain constant, cause a change in a dependent variable, or a factor that changes. A field experiment is the main and most objective method for studying the theoretical and practical problems of agriculture (Shanin, 1977).

The optimal values of the main nutrients nitrogen (N), phosphorus (P), potassium (K), and silicon (Si) can be established after performing a regression analysis of the experimental data. The results of the field trials are influenced by multiple random factors during the growing season. This justifies the use of fuzzy regression methods. Different methods have been applied to find a solution for fuzzy linear systems. The least squares method is used (Diamond, 1988). A general solution of m×n fuzzy linear systems is given in (Mikaeilvand & Noeiaghdam, 2012), where the original system is replaced by two m×n crisp linear systems.

The theory of fuzzy sets allows us to structure in the best way everything that is not separated by very precise boundaries, for example, the storage of different soils with nutrients. This makes these methods suitable for establishing a tolerance for fertilizing crops with different substances.

The purpose of the present study is to determine the optimal values of the main nutrients nitrogen, phosphorus, potassium, and silicon with the help of the theory of fuzzy sets in the conditions of a field experiment with maize on leached smolnitsa and alluvial-meadow soil.

MATERIALS AND METHODS

Data from field experiments with maize on leached smolnitsa (Pellic Vertisol) and alluvial-meadow soil (Eutric Fluvisol) (FAO, 2015) derived in 2020 and 2022 by a team under a project funded by the Scientific Research Fund (Lozanova et al., 2022; Petkova & Sadovski, 2022; Sadovski et al., 2022) are used. Mineral fertilizers - N (ammonium nitrate), P (superphosphate), K (potassium sulfate), and Si (diatomic earth, which represents 89-95% silicon in amorphous form) are applied. The experiment includes 9 variants of fertilization in three replications with the size of the experimental parcels - 25 m^2.

The design of multifactorial experiments allows the assessment of actions and interactions of four factors, varying on three levels. It is a 1/2 replication of a 2^4 factor scheme with added control variant. The design of treatments is presented in Table 1.

The fuzzy set theory provides a strict mathematical framework in which vague conceptual phenomena can be precisely and rigorously studied. It can also be considered as a modeling language well suited for situations in which fuzzy relations, criteria, and phenomena exist.

Following Zimmermann (1992), a fuzzy number may be defined as *F = (b, g, h);* where *b* denotes the center (or mode), *g* and *h* are the left spread (L) and right spread (R), respectively, L and R denote the left and right shape functions. A popular fuzzy number is the triangular fuzzy number (see Figure 1).

Table 1. Field experiment – quantities of fertilizers (kg/da)

Variant	N	P	K	Si
1	10	8	6	1.4
2	20	8	6	2.8
3	10	16	6	2.8
4	20	16	6	1.4
5	10	8	12	1.4
6	20	8	12	2.8
7	10	16	12	2.8
8	20	16	12	1.4
9	0	0	0	0

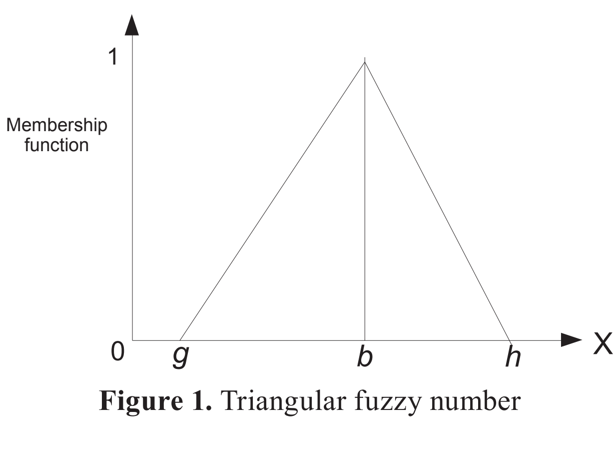

Figure 1. Triangular fuzzy number

In this paper, the main aim is using of a method where right - hand - side is a fuzzy vector and the coefficients matrix is crisp. Crisp means - something clearly defined, and deterministic in character. When we have crisp explanatory variables X_i, (i = 1, ..., n) and a fuzzy dependent variable $Y_i \equiv$ (b, g, h), (i = 1, ..., m), a model capable to incorporate the possible influence of the magnitude of the centers on the spreads, can be taken into account (D'Urso, 2003).

From all replications of experimental data X_i, (i = 1, …, 9) are derived the quantities

g_i = Min (X_i); b_i = Average (X_i); h_i = Max (X_i).

There is a simple solution approach to solving a general fuzzy system of linear equations (Mosleh et al., 2011). In the case of a fully fuzzy linear system $\tilde{A} \otimes \tilde{x} = \tilde{b}$ with a new notation $\tilde{A}(A, M, N)$, where *A*, *M*, and *N* are three crisp matrices, of the same size $\tilde{A}$, the matrices *A*, *M*, and *N* are called the center matrix, the left and right spread matrices, respectively.

In our paper, the coefficient matrix is considered as real crisp, whereas the unknown variable vectors are considered fuzzy. In this case, the matrices *M* and *N* are zero matrices. Using matrix notation we have

$$A \otimes x = \tilde{b},$$

or in expanded form

$$\begin{cases} (a_{11} \otimes x_1) \oplus (a_{12} \otimes x_2) \oplus \ldots \oplus (a_{1n} \otimes x_n) = \tilde{b}_1 \\ (a_{21} \otimes x_1) \oplus (a_{22} \otimes x_2) \oplus \ldots \oplus (a_{2n} \otimes x_n) = \tilde{b}_2 \\ . \\ (a_{m1} \otimes x_1) \oplus (a_{m2} \otimes x_2) \oplus \ldots \oplus (a_{mn} \otimes x_n) = \tilde{b}_m \end{cases} \quad (1)$$

where the crisp coefficient matrix is

$$A = (a_{ij}), \quad (i = 1, \ldots, m;\ j = 1, \ldots, n)$$

and $\tilde{b}_i = (b_i, g_i, h_i)$ are nonnegative fuzzy numbers.

For calculate (1) the following simple sequence is used:

1. Singular value decomposition is made

$$A = U\Sigma V^t, \quad (2)$$

where *U* and *V* are orthogonal matrices; and Σ is a diagonal matrix.

2. Pseudo-inverse matrices Σ^+ and $A^+ = V\Sigma^+U^t$ are found.

The following dependencies exist

$$Ax = b,$$

$$Ay + Mx = g, \quad (3)$$
$$Az + Nx = h.$$

3. From them consecutively the unknown values are calculated

$$x = A^{+}b,$$
$$y = A^{+}(g - MA^{+}b), \quad (4)$$
$$z = A^{+}(h - NA^{+}b).$$

From calculated values of x, y, and z we can find the fuzzy solution

$$\tilde{b} = (b, g, h),$$

where $b = Ax$, $g = Ay$, $h = Az$.

Calculations by the method described are done with the free software package GNU Octave, version 6.4.0.

Regression analysis of data for the yield of Maize from a field experiment in Bozhurishte and Tsalapitsa is performed using the equation

$Y = b_0 + b_1N + b_2P + b_3K + b_4Si + b_5NP + b_6PK + b_7KSi$ (5)

In this regression equation, the obtained fuzzy solution for the average yield g is considered as the dependent variable Y. Analysis was performed with Prism 9 software (GraphPad, 2020).

RESULTS AND DISCUSSION

The unknown variable vectors we are looking for are:

Yield = ($Y\text{-}_g$, $Y\text{-}_b$, $Y\text{-}_h$),

Explanatory variables for the analysis (see Table 2) are:

Table 2. Values of crisp coefficients matrix A

Variant	Explanatory variables						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	10	8	6	1.4	80	48	8.4
2	20	8	6	2.8	160	48	16.8
3	10	16	6	2.8	160	96	16.8
4	20	16	6	1.4	320	96	8.4
5	10	8	12	1.4	80	96	16.8
6	20	8	12	2.8	160	96	33.6
7	10	16	12	2.8	160	192	33.6
8	20	16	12	1.4	320	192	16.8
9	0	0	0	0	0	0	0

X_1 = N (kg/da),
X_2 = P (kg/da),
X_3 = K (kg/da),
X_4 = Si (kg/da),
X_5 = NP interaction,
X_6 = PK interaction,
X_7 = KSi interaction.

Fuzzy variables vectors for Bozhurishte – 2020 and 2022 years, as well as for Tsalapitsa – 2020 and 2022 years are presented in Tables 3 – 6.

Multiple regression analysis of fuzzy solutions for the average yield g is given in Tables 7 – 10.

Corresponding optimums of the fuzzy solutions of the main nutrients' average yields are shown in Table 11, which is easy to interpret.

The quantity of Nitrogen in Bozhurishte applied in 2020 has a consequence that leads to a smaller required dose for 2022. The same goes for Potassium. Phosphorus is needed in almost the same amount for both years. The difference in Silicon norms is insignificant.

Alluvial-meadow soil in Tsalapitsa needs higher doses of Nitrogen for both years. The same applies to Potassium. The need for Phosphorus is greater in the second year. The soil in Tsalapitsa needs a significant amount of Silicon in the first year, which is sufficient for the second year as well.

CONCLUSIONS

The optimal values of the main nutrients nitrogen (N), phosphorus (P), potassium (K), and silicon (Si) for Maize are established after performing a regression analysis of the field experimental data. The interpretation of the obtained results shows the need for higher doses of Nitrogen and Potassium for the alluvial-meadow soil. The leached chernozem needs a sufficient amount of Phosphorus in both years. Maize in Tsalapitsa requires a larger amount of Silicon as stock fertilization.

In this article, we show the efficiency of the proposed method for solving non-least-square linear fuzzy regression. This scheme for finding the positive solution of the fuzzy systems, when parameters are positive, it turns out quite satisfactory. It can be concluded that the fuzzy set techniques are promising for future research in other agricultural crops as well.

Table 3. Fuzzy variables vectors Bozhurishte 2020

Variant	Input data			Output results		
	g	b	h	$Y\text{-}_g$	$Y\text{-}_b$	$Y\text{-}_h$
1	723.25	770.13	815.67	797.87	819.13	839.21
2	1074.40	1105.54	1136.68	999.78	1056.50	1113.10
3	584.23	810.52	1036.80	658.85	859.52	1060.30
4	949.56	973.30	997.03	874.94	924.30	973.49
5	789.40	826.80	864.19	714.78	777.80	840.65
6	1139.27	1163.34	1187.41	1213.90	1212.30	1211.00
7	1289.88	1289.88	1289.88	1215.30	1240.90	1266.30
8	1059.55	1059.55	1059.55	1134.20	1108.50	1083.10
9	538.03	576.97	615.91	0.00	0.00	0.00

Table 4. Fuzzy variables vectors Bozhurishte 2022

Variant	Input data			Output results		
	g	b	h	$Y\text{-}_g$	$Y\text{-}_b$	$Y\text{-}_h$
1	431.69	583.23	734.77	440.20	592.56	744.92
2	417.13	566.28	715.44	408.62	556.95	705.29
3	357.58	420.54	483.51	366.09	429.87	493.66
4	282.41	468.95	655.49	273.90	459.62	645.34
5	448.82	475.24	501.66	440.31	465.91	491.51
6	302.08	345.04	388.00	310.59	354.37	398.15
7	359.28	360.05	360.82	350.77	350.72	350.67
8	348.20	447.06	545.91	356.71	456.39	556.06
9	179.69	209.12	238.56	0.00	0.00	0.00

Table 5. Fuzzy variables vectors Tsalapitsa 2020

Variant	Input data			Output results		
	g	b	h	$Y\text{-}_g$	$Y\text{-}_b$	$Y\text{-}_h$
1	630.18	831.48	953.67	706.06	869.19	958.72
2	1344.15	1359.80	1379.90	1268.30	1322.10	1374.90
3	1163.03	1241.45	1311.22	1238.90	1279.20	1316.30
4	1181.85	1372.43	1499.76	1106.00	1334.70	1494.70
5	1045.59	1083.77	1118.54	969.71	1046.10	1113.50
6	1046.99	1257.29	1469.92	1122.90	1295.00	1475.00
7	1045.25	1073.55	1094.85	969.37	1035.80	1089.80
8	1169.62	1257.68	1317.85	1245.50	1295.40	1322.90
9	641.39	666.87	680.85	0.00	0.00	0.00

Table 6. Fuzzy variables vectors Tsalapitsa 2022

Variant	Input data			Output results		
	g	b	h	Y_{-g}	Y_{-b}	Y_{-h}
1	973.30	1019.60	1065.90	1071.30	1094.60	1117.90
2	1181.88	1188.28	1194.69	1083.80	1113.30	1142.70
3	892.84	1052.57	1212.30	990.89	1127.60	1264.30
4	1081.90	1138.28	1194.65	983.85	1063.30	1142.70
5	1009.80	1208.02	1406.25	911.75	1133.00	1354.30
6	948.53	1187.53	1426.53	1046.60	1262.50	1478.50
7	1368.86	1553.60	1738.34	1270.80	1478.60	1686.40
8	1043.40	1228.44	1413.47	1141.40	1303.40	1465.40
9	332.13	405.62	479.11	0.00	0.00	0.00

Table 7. Regression analysis of Bozhurishte 2020 fuzzy solution

Variable	b	Std.err.	|t|	p
N	124.300	0.0061	20348	<0,0001
P	102.500	0.0060	17015	<0,0001
K	-77.340	0.0064	12045	<0,0001
Si	-339.100	0.0319	10631	<0,0001
NP	-9.098	0.0005	17512	<0,0001
PK	4.699	0.0004	12759	<0,0001
KSi	23.470	0.0021	11153	<0,0001

Table 8. Regression analysis of Bozhurishte 2022 fuzzy solution

Variable	b	Std.err.	|t|	p
N	92.940	0.0012	76084	<0,0001
P	84.860	0.0012	70411	<0,0001
K	-29.030	0.0013	22601	<0,0001
Si	-191.100	0.0064	29953	<0,0001
NP	-7.769	0.0001	74774	<0,0001
PK	2.571	0.0001	34910	<0,0001
KSi	-9.039	0.0004	21475	<0,0001

Table 9. Regression analysis of Tsalapitsa 2020 fuzzy solution

Variable	b	Std.err.	|t|	p
N	-13.810	0.0159	870	0.0007
P	4.407	0.0159	278	0.0023
K	99.540	0.0164	6090	0.0001
Si	381.100	0.0816	4673	0.0001
NP	3.269	0.0013	2424	0.0003
PK	-4.505	0.0009	4806	0.0001
KSi	-24.290	0.0054	4535	0.0001

Table 10. Regression analysis of Tsalapitsa 2022 fuzzy solution

Variable	b	Std.err.	|t|	p
N	133.600	0.0176	7576	<0,0001
P	143.600	0.0176	8152	<0,0001
K	-45.700	0.0182	2516	0.0003
Si	-373.000	0.0906	4117	0.0002
NP	-11.320	0.0015	7555	<0,0001
PK	4.203	0.0010	4035	0.0002
KSi	13.200	0.0060	2217	0.0003

Table 11. Fuzzy optimums of the nutrients (kg/da)

Location	Year	N	P	K	Si
Bozhurishte	2020	18.73	13.66	14.45	0.56
	2022	3.93	11.96	0.00	0.19
Tsalapitsa	2020	20.27	4.22	15.69	3.31
	2022	23.18	11.80	28.26	0.00

Disclaimer

The Author does not imply the expression of any opinion about the software mentioned in the article.

ACKNOWLEDGMENTS

The publication is a result of the work on Project KP-06-H 36/15 of 17. 12. 2019 *"Innovative meth-*

od for optimal fertilization of agricultural crops", funded by the Research Fund, Bulgarian Ministry of Education and Science.

REFERENCES

Diamond, P. (1988). Fuzzy least squares. *Information Sciences*, 46, pp. 141-157.

D'Urso, P. (2003). Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data. *Computational Statistics & Data Analysis*, *42*(1-2), 47-72.

FAO. (2015). World reference base for soil resources 2014. FAO, Rome, p. 203. http://www.fao.org/3/i3794en/I3794en.pdf

GraphPad. (2020). Prism 9. GraphPad Software LLC. San Diego. https://www.graphpad.com/

Lozanova, V., Simeonova, Ts., Benkova, M. Nenova, L. Gerasimova, I. Petkova, Z. Katsarova, A. Nenov, M. Atanassova, I. Nikolova, M. Harizanova, M., & Sadovski, A. (2022). Results of fertilization of wheat with nitrogen, phosphorus, potassium and silicon fertilizer in field experiment on alluvial-meadow soil. Sci. Sess. "Ecology and Agrotechnology - Basic Science and Practical Implementation", October 21-22. 2021. N. Pushkarov Institute, Sofia, 102-108. (Bg.).

Mikaeilvand, N., & Noeiaghdam, Z. (2012). The general solution of m× n Fuzzy Linear Systems. *Middle-East Journal of Scientific Research*, *11*(1), 128-133.

Mosleh, M., Abbasbandy, S., & Otadi, M. (2011). A method for solving fully fuzzy linear system. *Mathematics Scientific Journal*, *7*(2), 59-70.

Octave Project. (2021). GNU Octave. https://www.octave.org/

Petkova, Z., & Sadovski, A. (2022). Studies to optimization of wheat nutrition. *Rastenievadni nauki,* 2022, 59(1), 7-12.

Sadovski, A. (2020). An Introduction to Mathematical Agronomy: Theoretical and Experimental bases. LAP Lambert Academic Publishing. ISBN 978-620-2-51889-5, 284 p.

Sadovski, A., Atanassova, I., & Petkova, **Z.** (2022). Simulation of crops yield grown on leached smolnitsa and alluvial-meadow soil. *Ecological Engineering and Environment Protection*, No 1, pp. 62-69.

Shanin, Y. (1977). Methodology of the field experiment. BAS, Sofia (in Bg).

Zimmermann, H. J. (1992). Fuzzy Set Theory and its Application. Second, Revised edition. Kluwer Academic Press, Dordrecht, p. 407.