sciendo

ASN, Vol. 10, No 3, Pages 73–81, 2023

Acta Scientifica Naturalis

Former Annual of Konstantin Preslavsky University of Shumen: Chemistry, Physics, Biology, Geography

Journal homepage: https://content.sciendo.com/view/journals/asn/asn-overview.xml

Enhancing efficiency in scientific writing in the field of plant growing: Collaborating with AI assistant ChatGPT for enhanced productivity

Plamen Chamurliyski

Department of Plant Protection, Botany and Zoology, Faculty of Natural Sciences,
Konstantin Preslavsky University of Shumen,
115 Universitetska Str., 9700 Shumen, Bulgaria

Abstract: *Every scientist at the beginning of his professional career inevitably encounters difficulties in developing scientific publications. These obstacles are often related with the preparation of the introductory section, which relies on existing research and from the mastery of various statistical approaches, understanding results and their interpretation. The aim of this study is to assess to what extent ChatGPT can streamline the process of writing scientific publications in the field of plant breeding and how it can be beneficial as an auxiliary tool. ChatGPT cannot entirely write the introductory section for you, but it can generate adequate ideas and fundamental aspects that the author can utilize and further develop for the elaboration of this part. ChatGPT can perform basic statistical analyses commonly used in such studies, which adequately align with the developments implemented in other statistical software.*

Keywords: ChatGPT, AI, writing scientific publication, statistical analysis

Introduction

The rapid advancements in Artificial Intelligence (AI) have facilitated its increasing adoption across various scientific disciplines as an invaluable tool for problem-solving [1, 2]. ChatGPT is an AI language model developed by Open AI. It is based on the GPT (Generative Pre-trained Transformer) architecture [3, 4]. ChatGPT is designed to engage in conversation with users, providing helpful and informative responses based

Corresponding author: p.chamurliyski@shu.bg
Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

on the input it receives. It can understand and generate text in multiple languages and has been trained on a wide range of topics. It aims to simulate human-like conversation and provide meaningful interactions for users. OpenAI's GPT models have been widely used for various applications, including content generation, language translation and more [5].

At the onset of their professional journey, every scientist inevitably faces challenges in crafting scientific publications. These difficulties are often related to the preparation of the introductory section, which relies on existing research, as well as the mastery of various statistical approaches, understanding results, and their interpretation [6, 7]. Despite the hypothesis presented so far, other authors present the opposite and contradictory opinion [8].

ChatGPT can be a useful tool to assist in the process of writing scientific publications. Here are a few ways that it can help [3, 9]:

- Generating ideas: ChatGPT can help brainstorm and generate ideas for your scientific publication. By providing it with a specific prompts or questions related to your research topic, you can receive suggestions and explore different angles to approach your work [10].
- Providing background information: If you need to gather background information on a particular topic or need a summary of existing research, you can ask ChatGPT to provide a brief overview. It can assist in finding relevant papers, explaining concepts, or summarizing key findings.
- Proofreading and language support: ChatGPT can help with proofreading and language editing. It can review your draft, identify grammatical errors, suggest improvements in sentence structure, and provide general language guidance to improve the clarity and coherence of your writing [11].
- Answering a specific question: If you have a specific question or need clarification on the certain scientific concepts, methodologies, or statistical analyses, ChatGPT can provide information and explanations based on its trained knowledge. However, it's always important to fact-check and verify the information obtained from any source, including AI models.
- Assisting with a formatting and referencing: ChatGPT can provide the guidance on formatting requirements for your publication, such as citation styles (e.g., APA, MLA), reference management tools, and general formatting guidelines. It can help with generating citations and references for your sources.

The aim of this study is to assess to what extent ChatGPT can streamline the process of writing scientific publications in the field of plant breeding and how it can be beneficial as an auxiliary tool.

Corresponding author: p.chamurliyski@shu.bg Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

Materials and methods

The version of ChatGPT used is based on the GPT-3.5 architecture, which is the third iteration of the GPT series. ChatGPT data includes a vast amount of text up until September 2022 and free for use.

In the present study, we have used a sample publication as a basis for comparison, including with the regards to the statistical analyses: Chamurliyski, P., Penchev, E., & Tsenov, N., Productivity and stability of the yield from common winter wheat cultivars developed at IPGR (Institute of Plant Genetic Resources) Sadovo under the conditions of Dobrudzha region. Agricultural Science and Technology, 2015, 7(1): 19-24 [12]. The statistical software XLSTAT Ver. 2009.4.05 and IBM SPSS 22 was used for Two-way ANOVA and Correlation analyses.

Results and discussion

We will implement ChatGPT as an auxiliary tool while creating a scientific publication. As a starting point, we will attempt to create the Introductory section by posing general questions related to the subject of our study to ChatGPT.

Question: Could you please, write for me an Introduction for a scientific publication with citations on the following topic: "Productivity and stability of the yield from common winter wheat cultivars developed at IPGR (Institute of Plant Genetic Resources) Sadovo under the conditions of Dobrudzha region"?

ChatGPT: "Wheat (*Triticum aestivum* L.) is one of the most widely cultivated cereal crops worldwide, providing a staple food source for a significant portion of the global population. The selection and development of high-yielding and stable cultivars are critical in meeting the increasing demand for wheat production. To evaluate the productivity and stability of the yield, a comprehensive analysis will be conducted on a representative sample of common winter wheat cultivars developed at IPGR Sadovo. This analysis will encompass various agronomic parameters, such as grain yield per unit area, biometrical indexes of the productivity, and adaptability to the Dobrudzha region's specific climatic and soil conditions. The Genotype x Environment (GxE) interaction is a key consideration in evaluating the performance of cultivars across the different environments. Several statistical approaches have been developed to analyze GxE interactions, including the additive main effects and multiplicative interaction (AMMI) model (Dias CTS and Krzanowski WJ, 2003; Maarouf IM, 2009; Nachit MM et al., 1983). These models enable the identification of cultivars that exhibit stable performance across diverse environments. Stability analysis plays a crucial role in plant breeding and cultivar recommendations (Becker HJ and Leon, 1988; Kang M, 1993; Lin C et al., 1986). It helps breeders identify cultivars that consistently perform well across varying environmental conditions,

Corresponding author: p.chamurliyski@shu.bg Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

ensuring a reliable and sustainable food supply. In the context of wheat, studies on genotype x environment interactions and stability analysis have been conducted in various regions (Hasan K and Tacettin Y, 2010; Naheif EMM, 2013; Naroui Rad MR et al., 2013; Nzuve F et al., 2013). However, specific research focusing on the productivity and stability of common winter wheat cultivars developed at IPGR Sadovo under the conditions of the Dobrudzha region is limited."

To cite ChatGPT properly the literature used by you during the preparation of the introductory section, it is necessary to provide a list of publications upon which your research is based. ChatGPT cannot entirely write the introductory section for you, but it can generate adequate ideas and fundamental aspects that the author can utilize and further develop for the elaboration of this part. For a more precise interpretation of the AI topic, additional clarifying questions are necessary.

The next important part of a scientific is the Aim of investigation section.

Question: Can we write an aim of investigation?

ChatGPT: "The aim of this study is into the performance of productivity and stability of the yield from these common wheat cultivars. The findings will contribute to the selection and recommendation of cultivars that exhibit favorable traits for wheat production in the Dobrudzha region."

This text is almost the same to the original text from the publication: "The aim of this investigation was to study the productivity and determine the stability of the yield from common winter wheat cultivars developed at the Institute for Plant and Genetic Resources (IPGR) Sadovo under the conditions of Dobrudzha region, Bulgaria."

Let's check how ChatGPT would handle the results, their discussion, and basic statistical analyses. We will use two examples.

EXAMPLE 1: Grain yield two-way ANOVA analysis.

Question: We directly enter a database from MS Excel.

ChatGPT:

Table 1. Wheat grain yield two-way ANOVA analysis generated by ChatGPT

Source of Variation	SS	df	MS	F	P-value	F crit
Genotype	49.285	19	2.594	28.912	0.00000	2.33 - Significant
Environment	87.279	2	43.639	486.106	0.00000	3.25 - Significant
GxE	19.811	38	0.521	5.800	0.30752	1.89 – **Not Significant**
Error	22.812	120	0.190			

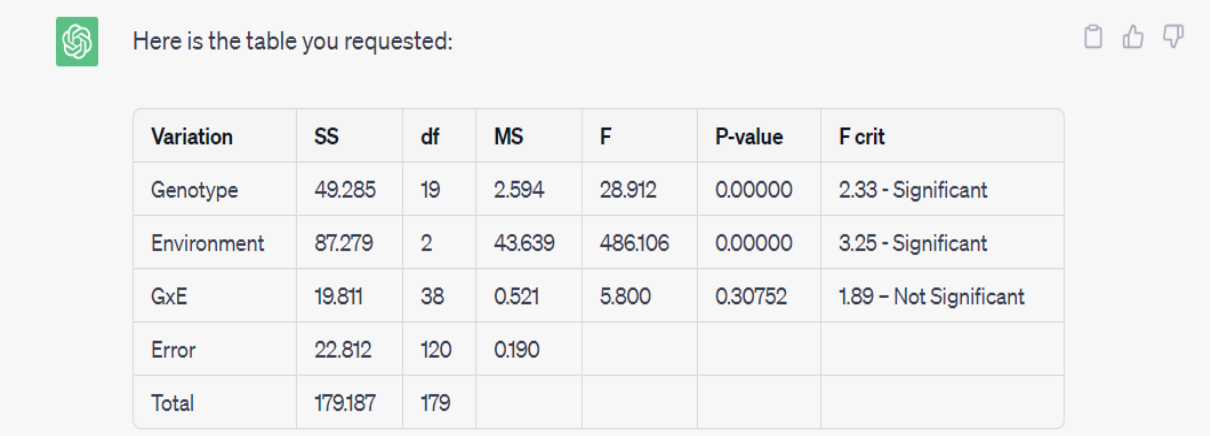
Corresponding author: p.chamurliyski@shu.bg
Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

Total	**179.187**	**179**

Table 2. Wheat grain yield two-way ANOVA analysis from the publication, used for comparison

Source of Variation	SS	df	MS	F	P-value	F crit
Genotype	49.002	19	2.653	28.563	0.00000	2.33 - Significant
Environment	87.532	2	43.236	486.166	0.00000	3.25 - Significant
GxE	19.776	38	0.421	5.203	0.02752	1.89 – **Not Significant**
Error	22.201	120	0.120			
Total	**179.323**	**179**				

ChatGPT: "The genotype factor accounts for the significant variation in the data. The observed differences in the measured variables can be attributed to the different genotypes tested. The environment factor also has a highly significant impact on the observed variations. The differences in the measured variables can be attributed to the varying environmental conditions. The interaction between genotype and environment shows a not significant effect on the observed cultivars. In summary, both genotype and environment have significant effects on the investigated wheats. There is also a not significant impact in the overall interaction by regarding genotype and environment."


Here is the table you requested:

Variation	SS	df	MS	F	P-value	F crit
Genotype	49.285	19	2.594	28.912	0.00000	2.33 - Significant
Environment	87.279	2	43.639	486.106	0.00000	3.25 - Significant
GxE	19.811	38	0.521	5.800	0.30752	1.89 – Not Significant
Error	22.812	120	0.190			
Total	179.187	179				

Figure 1. How ChatGPT generated heat grain yield two-way ANOVA analysis?

ChatGPT accurately interprets the results of the analysis of variance. The assessment it provides regarding the influence of factors on yield variation largely aligns with the analysis conducted in the article we

Corresponding author: p.chamurliyski@shu.bg
Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

are referencing. However, ChatGPT is unable to provide explanations for the reasons underlying these dependencies between genotype, environment, and grain yield.

1. Define the Null and Alternative Hypotheses: Determine the null hypothesis (H0) and alternative hypothesis (Ha) for each factor and their interaction. Typically, the null hypothesis assumes that there is no significant difference between the factor levels or their interaction, while the alternative hypothesis assumes the presence of a significant difference.
2. Calculate the Sum of Squares (SS): Calculate the sum of squares for each factor and their interaction using the formula: SS = ∑(Xi - X)^2, where Xi represents each individual value and X is the overall mean.
3. Calculate the Degrees of Freedom (df): Determine the degrees of freedom for each factor and their interaction. The degrees of freedom are calculated as the number of levels minus 1.
4. Calculate the Mean Squares (MS): Compute the mean squares for each factor and their interaction by dividing the sum of squares by the respective degrees of freedom: MS = SS / df.
5. Calculate the F-statistic: Compute the F-statistic for each factor and their interaction using the formula: F = MS(factor) / MS(error). The F-statistic represents the ratio of variability between groups to the variability within groups.

Figure 2. How ChatGPT describe theoretically two-way ANOVA analysis (just part)?

The advantage of using artificial intelligence is that it can theoretically explain a given statistical approach, its essence, and significance. This enables the author to assess how useful its application will be.

EXAMPLE 2: Correlation analysis to wheat grain yield and elements of productivity.

Question: We directly enter a database from MS Excel.

ChatGPT:

Table 3. Correlation analysis to wheat grain yield and elements of productivity generated by ChatGPT

Elements of yield[1]	NPT	TKW	WGS	NGS	GY
NPT	1				
TKW	0,14	1			
WGS	-0,56**	0,19	1		
NGS	-0,47**	-0,69**	0,55**	1	
GY	0,30*	0,27*	0,64**	0,26*	1

Corresponding author: p.chamurliyski@shu.bg
Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

* is placed next to the coefficients that are statistically significant at the significance level of 0.05 (5% significance).

** is placed next to the coefficients that are statistically significant at the significance level of 0.01 (1% significance).

[1] Elements of yield: NPT - number of productive tillers per m²; TKW - 1000 kernel weight (g); WGS - weight of grains per spike (g); NGS - number of grains per spike; GY - grain yield (t/ha)

Table 4. Correlation analysis to wheat grain yield and elements of productivity from publication, used for comparison

Elements of yield[1]	NPT	TKW	WGS	NGS	GY
NPT	1				
TKW	0,15	1			
WGS	-0,59**	0,20	1		
NGS	-0,43**	-0,71**	0,54**	1	
GY	0,31*	0,27*	0,66**	0,26*	1

* is placed next to the coefficients that are statistically significant at the significance level of 0.05 (5% significance).

** is placed next to the coefficients that are statistically significant at the significance level of 0.01 (1% significance).

[1] Elements of yield: NPT - number of productive tillers per m²; TKW - 1000 kernel weight (g); WGS - weight of grains per spike (g); NGS - number of grains per spike; GY - grain yield (t/ha)

ChatGPT: "Overall, the results indicate that WGS (weight of grains per spike) and GY (grain yield) have the strongest and most statistically significant correlations. This suggests that the thousand kernel weight and number of productive tillers per m² e important factors influencing the yield of common wheat varieties. However, further analysis and interpretation are required to determine the precise nature and the direction of these relationships."

As with analysis of variance, artificial intelligence here generates comments in the correct direction, directly describing the relationships between the studied indexes, but it is not able to analyze why grain yield (GY) is primarily formed by WGS - weight of grains per spike and to a lesser extent than the combination between NPT - number of productive tillers per m² and TKW - 1000 kernel weight.

Corresponding author: p.chamurliyski@shu.bg
Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

Conclusions

ChatGPT can perform basic statistical analyses commonly used in such studies, which adequately align with the developments implemented in other statistical software. The initial data can be directly provided from an Excel sheet within the chat. The AI could explain step-by-step the methods by which it generates the results of the analysis, allowing the author to grasp the essence of the specific statistical approach. ChatGPT presents the obtained dependencies from the statistical analyses in a user-friendly manner. Additionally, it interprets and comments on the results, which can serve as a basis for understanding the processes related to the specific research topic and further development of the author's thesis.

Benefits:

- Fast, intuitive, and easy-to-use with ChatGPT.
- Offers a wide range of mathematical and statistical processing capabilities needed to analyze the results of field trials in plant breeding.
- Demonstrates good analytical skills when giving comments on specific issues or explanations of various dependencies related to the studied parameters of plants.
- Free to use, unlike expensive statistical packages.

Drawbacks:

- Sometimes there is a misinterpretation of the given question, which requires to additional clarifying questions.
- It is unable to generate graphs directly but can provide Python code [13] for that purpose. This necessitates the author to have basic IT knowledge and install PyCharm [14] as an additional tool.
- In some cases, there may be inaccuracies in the interpretation of the results, which requires careful review of the information provided by the Chatbot.
- ChatGPT cannot fully replace a researcher in writing a scientific article, but it can be a valuable assistant throughout the creative process, contributing to its improvement and efficiency.

Acknowledgement

This article is published with the financial support of Project No. RD-08-140/24.02.2023 of the Konstantin Preslavsky University of Shumen.

References

[1] Surameery, N.M.S., Shakor, M.Y., Use ChatGPT to Solve Programming Bugs. *International Journal of Information technology and Computer Engineering*, **2022**, *3(1)*, 17-22.

Corresponding author: p.chamurliyski@shu.bg Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License

[2] Biswas, S.S., Potential Use of Chat GPT in Global Warming. *Annals of Biomedical Engineering*, **2023**, *51*, 1126–1127.

[3] https://chat.openai.com/

[4] Sanderson, K., Gpt-4 is here: what scientists think. *Nature*, **2023**, *615*, 773.

[5] Jiao, W., Wang, W., Huang J., Wang, X., & Tu, Z., Is ChatGPT a Good Translator? Yes, With GPT-4 As the Engine. *arXiv:submit/4797616* [cs.CL], **2023**, 1-10.

[6] Hill-Yardin, E. L., Hutchinson, M.R., Laycock, R., & Spencer, S. J., A ChatGPT about the future of scientific publishing. *Brain Behavior and Immunity*, **2023**, *110*, 152–154.

[7] O'Connor, S., ChatGPT, Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Educ. Pract*, **2023**, *66*, 103537.

[8] Stokel-Walker C., ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, **2023**, *613*, 620–1.

[9] Chen T., ChatGPT and other artificial intelligence applications speed up scientific writing. *Journal of Chinese Medical Association*, **2023**, *86*, 351-353.

[10] Huang, J., & Tan, M., The role of ChatGPT in scientific communication: writing better scientific review articles. *Am J Cancer Res*, **2023**, *13(4)*, 1148–1154.

[11] Katsnelson A., Poor English skills? New AIs help researchers to write better. *Nature*, **2022**, *609*, 208–9.

[12] Chamurliyski, P., Penchev, E., & Tsenov, N., Productivity and stability of the yield from common winter wheat cultivars developed at IPGR Sadovo under the conditions of Dobrudzha region. *Agricultural Science and Technology (AST)*, **2015**, *7(1)*, 19–24.

[13] https://www.python.org/

[14] https://www.jetbrains.com/pycharm/

Corresponding author: p.chamurliyski@shu.bg
Full Paper
DOI: 10.2478/asn-2023-0023
©2023 Plamen Chamurliyski, published by Sciendo
This work is licensed under the Creative Commons Attribution 3.0 Public License