Digital Object Identifier: 10.24054/rcta.v1i47.4076

Use of ChatGPT in teaching the concept of safety factor in geotechnics

Uso de ChatGPT en la enseñanza del concepto de factor de seguridad en geotecnia

MSc. Óscar Hernán Vargas Villamizar [1], PhD. Juan Carlos Ruge Cárdenas [2], PhD. Carlos Felipe Urazán Bonells [2], MSc. Diego Palma Cuero [2]

[1] ***Universidad El Bosque,*** *Facultad de Ciencias Económicas y Administrativas, Bogotá, Colombia.*
[2] ***Universidad Militar Nueva Granada,*** *Facultad de Ingeniería, Bogotá, Colombia.*

Correspondence: *ovargasvi@unbosque.edu.co*

Received: *July 02, 2025.* ***Accepted:*** *December 20, 2025.* ***Published:*** *January 01, 2026.*

How to cite: *O. H. Vargas Villamizar, J. C. Ruge Cárdenas, C. F. Urazán Bonells y D. Palma Cuero, "Use of ChatGPT in teaching the concept of safety factor in geotechnics", RCTA, vol. 1, n.º. 47, pp. 13-25, Jan. 2026.*
Recovered from https://ojs.unipamplona.edu.co/index.php/rcta/article/view/4076

Abstract: This article analyzes the perceptions of undergraduate civil engineering students from a Colombian university regarding the responses generated by ChatGPT when solving an exercise involving the calculation of the safety factor of a geotechnical structure. A mixed analysis of the results was developed using descriptive statistics, supported by Rstudio software. This study contributes to the identification of ChatGPT as a pedagogical tool, whose use implies the reinterpretation of the teacher's role and the adaptation of the types of formative activities, with an emphasis on personalization, autonomous learning, and critical reflection.

Keywords: safety factor, geotechnics, engineering judgment, tertiary education, ChatGPT.

Resumen: En este artículo se analizan las percepciones de estudiantes de pregrado de ingeniería civil de una universidad colombiana acerca de las respuestas generadas por ChatGPT al resolver un ejercicio que involucra el cálculo del factor de seguridad de una estructura geotécnica. Se desarrolló un análisis mixto de los resultados, mediante el uso de estadística descriptiva, apoyado en el software Rstudio. Este estudio contribuye a la identificación de ChatGPT como herramienta pedagógica, cuyo uso implica la reinterpretación del rol del docente y la adaptación del tipo de actividades formativas, con énfasis en la personalización, el aprendizaje autónomo y la reflexión crítica.

Palabras clave: factor de seguridad, geotecnia, criterio ingenieril, educación universitaria, ChatGPT.

1. INTRODUCTION

Large Language Models (LLMs) have the potential to reshape our social structures [1], both because of the opportunities and challenges inherent in the use of these programs, whose responses mimic natural human language, sometimes indistinguishably. Mass access to LLMs is a recent phenomenon, and the November 2022 launch of ChatGPT, the artificial intelligence chat system (chatbot AI) developed by the company OpenAI, is a milestone in popularizing these programs. ChatGPT is estimated to have skills equivalent to those of a 9-year-old child and can pass 92.5% of the tasks included in cognitive tests [1].

Among the advantages of using AI chatbots in higher education is the ability to strengthen autonomous learning through virtual assistants that process high volumes of information and offer personalized recommendations without restriction of time or place [2]. However, AI chatbots also pose a high risk to educational integrity assurance [3] due to the difficulty of detecting chatbot-generated texts. Furthermore, the indiscriminate use of AI chatbots can undermine students' confidence in their abilities and reduce their critical thinking skills, affecting their future academic outcomes [1].

Faced with the challenge posed by AI chatbots in higher education, [4] consider that prohibition is not an option in engineering education, whose body of knowledge demands the incorporation of new technologies. Regarding curricular modifications, the design of activities in which the chatbot supports intermediate tasks, such as performing calculations and generating written proposals, but whose final solution depends on the student's critical thinking, is required [4]. To do this, teachers need to develop a greater understanding of the strengths of AI chatbots (e.g., writing programming code) and where their contributions are still limited (e.g., formulating initial equations) [3].

It has been identified that AI is capable of efficiently solving basic engineering problems [5]. However, AI still presents limitations in integrating ethical and contextual considerations in decision-making processes. This is relevant in the field of civil engineering, where design failures cause substantial economic losses or even the loss of human lives. Apart from the use of technical procedures and mathematical formulas, civil engineering practice relies on the application of engineering judgment, understood as the ability to interpret the conditions of a particular context through experience, practice, and the gradual development of rational thinking in a highly complex creative exercise [6]. Engineering judgment is of special relevance in the development of geotechnical projects, which depend on the characteristics of the soil and the conditions of the physical construction site, both of which are heterogeneous and non-modifiable.

In the process of guaranteeing the stability of a geotechnical structure, it is necessary to determine a factor of safety (FS) by using heuristic methods guided by engineering criteria, which consider aspects such as the uncertainty associated with the context, the risk that is willing to assume, current legislation and accepted technical standards in the discipline. The definition of an FS depends on the criteria of each engineer, based on his experience, technical training and even his personality [7]. This aspect makes teaching FS to an undergraduate student a challenge since it is not a "given" value but corresponds to an elusive concept [7], which is formed through a learning process that is difficult to imitate in the classroom. Because of this, in the university environment, it becomes indispensable to transcend the explanation of technical factors associated with the FS calculation by integrating discussions on the possible biases that a designer adopts in the FS calculation in geotechnics caused by his academic training, the conditions of the geotechnical problem addressed, or the idiosyncrasy of the local context [8]-[12].

In this scenario, AI chatbots' ability to process high volumes of information and generate multiple response scenarios offers an opportunity to teach complex concepts such as FS through critical reflection exercises that allow students to evaluate the methods used and the variables linked to the decision-making process.

2. THEORETICAL BACKGROUND

Despite the high potential of LLMs in educational contexts, [13] highlights their limitations in solving complex mathematical problems, particularly in providing rigorous validations, and [14] indicates that their integration into the curriculum should involve structured activities that require critical evaluation of AI-generated content, comparison with evidence-based resources, and explicit reflection on learning strategies.

Because the emergence of AI chatbots such as ChatGPT is a recent phenomenon, [4] highlights the scarcity of relevant studies on their effects on education. Also, [15] mentions that a limited number of papers have been published about the potential of ChatGPT in the engineering domain. Coinciding with the above, [16] points out a restricted number of publications about the application of ChatGPT in civil engineering. The exploration of ChatGPT's application in geotechnical engineering is still incipient. Table 1 summarizes articles published as of July 2024 in two bibliographic databases (Scopus and Science Direct) on using ChatGPT in geotechnical engineering.

***Table 1:** Studies relating to ChatGPT and geotechnical engineering*

Author(s)	Remarks
[15]	Studies the use of ChatGPT-supported programming in handling complex requirements in geotechnical engineering
[17]	Analyzes the use of ChatGPT to facilitate understanding of earthquake engineering concepts among students, professionals and policymakers.
[18]	Deploys critical reflection on the use of ChatGPT in engineering education and identifies areas of high educational potential, such as mechanical and structural analysis and materials design.
[19]	Evaluates the usefulness of ChatGPT in civil engineering education by posing questions on geomatics and construction surveying. By administering a questionnaire to 44 students at a U.S. university, the authors identified a positive reception of the use of the chatbot by the study participants, which offers new opportunities for improving undergraduate engineering education.
[20]	Highlights the potential of ChatGPT in scientific outreach on disaster prevention
[21]	Analyzes the potential of using LLM for tackling classical geotechnical engineering problems related to soil analysis, slope stability, and foundation design
[22]	Examines the potential ability of ChatGPT to generate finite element code for geotechnical engineering applications from a set of prompts.
[23]	Investigates the potential and limitations of automating the numerical modeling process in geotechnical problems using ChatGPT
[24]	Explores the integration of generative artificial intelligence (AI) into numerical analysis workflows in geotechnical engineering to address the challenges of generating synthetic datasets.

***Source:** own elaboration*

Among the studies identified, only [19] explores the relationship between ChatGPT and the teaching of geotechnical concepts in higher education, constituting one of the first efforts to explore the use of this AI platform for educational purposes among engineering students. Also, in identifying the limitations of their research, [19] point out that they only focused on introductory content on construction surveying and geomatics, thus inviting the exploration of more advanced concepts and analysis beyond these two areas of civil engineering. In connection with the above, [21] identifies that the application of Machine Learning and other AI techniques has shown significant growth in the field of geotechnical engineering, particularly in soil classification, slope stability analysis, and foundation design, prompting a call for educational initiatives to increase AI literacy within the engineering community.

In this scenario, academic integrity emerges as a key pillar in educational institutions when using tools such as ChatGPT. In this regard, [25] recommend that educations allow its use only as a self-learning tool, functioning as a complementary tutor, permitting its use for grammar correction and paraphrasing, but not for content generation. For this purpose, [25] emphasizes that it is essential for teachers and educational leaders to understand the limitations of these models to clearly communicate them to students.

This study is in an underexplored field of research: the integration of AI chatbots in teaching civil engineering in higher education. It focuses on analyzing the perceptions of students about the relevance of the answers generated by ChatGPT when solving an exercise that involves a complex concept of geotechnics, the factor of safety (FS), because its calculation involves the application of civil engineering criteria.

3. METHODOLOGY

3.1. Design and participants

A descriptive study was developed. Among the population, composed of students enrolled in a civil engineering undergraduate program, those enrolled in the Geotechnical Engineering course, taught in the seventh semester of a ten-semester program, were chosen. The sample consisted of 65 students. This course was chosen because of the relevance of the concept of engineering judgment in the activities of its content and because, due to its location in the curriculum, it ensured that all participants had a basic knowledge of the concept of FS, having previously taken the subjects of Soil Mechanics and Foundations.

3.2. Instruments and procedures developed

The instrument consisted of a physical questionnaire, applied in October 2023 and completed in a single day. At the beginning of the session, the purpose of the research was explained to the participants, who were informed that the results would be treated anonymously without affecting their course grades. Informed consent was included at the top of the questionnaire, and information was collected only from those participants who agreed to participate voluntarily.

The questionnaire was completed in a controlled environment to avoid the use of digital tools, taking 90 minutes to complete. It consists of seven questions (four closed and three open) organized in three parts (see Figure 1).

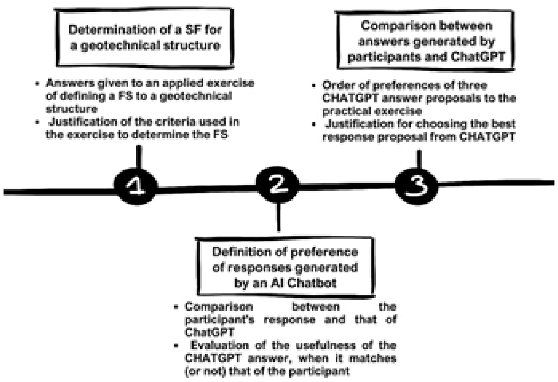

Fig. 1. *Structure of the questionnaire*
Source: *own elaboration.*

In the first part of the questionnaire, the participants answered two questions. In the first one, they were given an FS calculation exercise for a geotechnical foundation with an open answer. In the solution of this exercise, the participants were not allowed to access the internet or know the following questionnaire queries in advance. In the second question, they had to choose, from a list, the method used to find the numerical answer of FS.

The exercise solved by the participants is presented below.

A square foundation in plan has dimensions of 2 m x 2 m. The soil on which the foundation is supported has the following values:

- Effective internal friction angle: 25°
- Effective Cohesion: 20kN/m
- Specific weight: 16,5 kN/m³

If the calculated ultimate bearing capacity of the soil is 1078.29 kN/m², what factor of safety (FS) do you recommend finding the allowable load of the foundation? *Use a single decimal place to write your answer as a number (not a range).*

This same statement, without modifications or additional instructions, was used by the authors as a prompt to request ChatGPT's response to the exercise. The free-access version available in October 2023 was used, whose underlying model was GPT – 3.5 (family gpt-3.5 / gpt-turbo). According to [26], OpenAI did not publish the complete table of hyperparameters nor and official parameter count for the version used in this paper. The free version of ChatGPT was employed because, as reported by [26], between 95% and 97% of users relied on this version in 2023, making it the most accessible option in the university setting.

The wording of the proposed exercise considered the key aspects of generating a prompt (text to communicate with generative AI systems): instruction, context, input data, and output indicator [4].

In this case, the exercise functions as a prompt and follows the same guidelines used in the design of this type of activity in a classroom context. Specifically, the input parameters of the exercise were provided with realistic values to prevent both the chatbot and the student from having to assume any parameter required for solving the problem.

After obtaining the first response to the exercise, the authors asked ChatGPT to generate two additional responses, for a total of three FS options to be used in the foundation analysis. The complete texts of the three ChatGPT solution proposals are presented in Annex 1.

Considering that the focus of this research is the application of the engineering judgment, students did not have access to their electronic devices throughout the activity, nor did they generate prompts themselves, since there was a single prompt corresponding to the statement of the proposed exercise.

As soon as all participants had completed the first part of the questionnaire, they were allowed to consult the second part, in which they compared their answers to the geotechnical exercise with three ChatGPT-generated proposals for the same exercise.

The second part is composed of two questions. In the first, of a closed type, the three proposals generated by ChatGPT in response to the FS calculation exercise were shared, which the participant organized according to the order of preference, assigning number 1 to the one he/she considers the best response, number 2 to the second best proposal, and number 3 to the least relevant proposal as a response to the exercise. In the second open-type question, each participant included a justification for choosing one of the three proposals as the most suitable for answering the problem. In the third part, composed of three open-ended questions, the participant was asked to compare his answer with the one generated by ChatGPT and whether he would be willing to replace his answer with the chatbot's answer or keep his proposal.

3.3. Data Analysis

The information from the physical questionnaires was entered into a spreadsheet and analyzed using RStudio software to perform statistical and graphic analyses of the responses. The tendency of the participants' responses to the FS value for the proposed exercise was analyzed using a frequency analysis (Figure 2) and an interquartile analysis (Figure 3). For this purpose, the following commands were used:

hist (base_de_datos$variable_analizada, xlab=" ", ylab=" ", main=" ").
boxplot (base_de_datos$variable_analizada, xlab=" ", ylab=" ").

Afterwards, the tendency of the criteria used by the participants to choose the value of FS was analyzed. The sequence of the proposals was graphically contrasted with the possible criteria to be used to define the FS: standards and regulations, scientific theories, professional experience, professional discussion and classroom knowledge, the results of which can be seen in Figure 4:

ggplot (base_de_datos, aes(x = variable_analizada, y = variable analizada)) + theme (axis.text.x = element_text(angle = 90, hjust = 1)) + geom_point() + facet_wrap(variable_analizada ~ .) + labs (title = " ", y=" ", x =" ")

Finally, using the "facet wrap" dot plot style, the participants' qualitative justifications related to the alternative of replacing the FS value obtained in the first part of the questionnaire (without external support) with the response offered by their preferred ChatGPT proposal. Figure 6 presents the results for participants who decided to replace their response with the ChatGPT response, and Figure 7 shares the results for those who decided to keep their original response.

To enhance the clarity of the study, a qualitative axial analysis was applied to the criteria for responding the exercise (see figure 4), the reason for the methodological choice (see table 2), and the criterion for redefining the response based on the AI's output (see table 3).

4. RESULTS

The results are presented below, organized according to the three sections of the questionnaire described in the methodology section (see Figure 1).

4.1. Determination of the FS in a geotechnical structure exercise.

Next, Figure 2 presents the frequency in choice of FS by the participants who solved the proposed exercise.

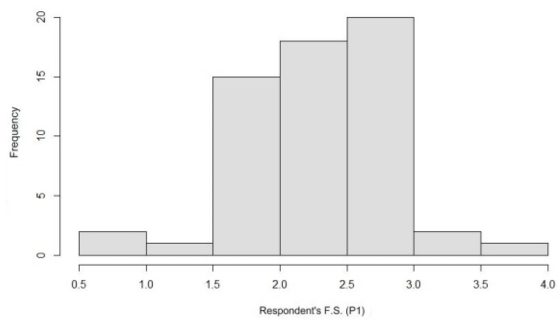

Fig. 2. *Frequency of response to FS exercise.*
Source: *own elaboration.*

According to Figure 2, the FS trend is between 2.5 and 3.0, with a low frequency of responses <1.5 and >3.0. The minimum response value was an FS of 0.5, and the maximum was 3.8. The mean value of the responses was 2.475.
Figure 3 shows that the interquartile range (50% of the participating data or responses) is at an FS between 2.0 and 3.0. The median of the FS data is 2.5. There are extreme outliers close to 0.5 and 3.8. That is, Q1 is between an FS of 0.5 and 2.0, Q2 is between 2.0 and 2.5, Q3 is between 2.5 and 3.0, and Q4 is between 3.0 and 3.8.

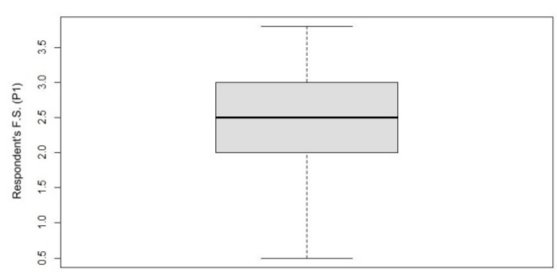

Fig. 3. *Interquartile range of responses to the FS exercise.*
Source: *own elaboration.*

From the above, it is highlighted that there is a very wide range in Q1, which translates into a large number of responses far from the trend values (Q2 and Q3). In addition, Q1 presents a greater range of selection, which gives a higher probability that the user will opt for this range, even if they do not

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

understand the question and randomly answer any alternative. When the participants were asked about the criteria used to answer the exercise, a preference for using the knowledge learned in the classroom was identified, as shown in Figure 4.

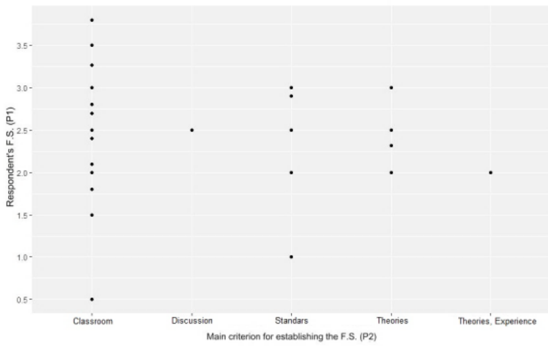

Fig. 4. *Criteria used to respond to the exercise.*
Source: *own elaboration.*

The highest frequency for an FS selection criterion was "knowledge learned in the classroom", distributed throughout the range of values from 0.5 to 3.8. The options on "technical rules and regulations" and "scientific and mathematical theories" are the second and third most employed criteria, with values between 1.0 and 3.0, respectively. Finally, the options associated with knowledge acquired in discussions and professional experiences were the most minor used criteria, with FS responses between 2.0 and 2.5.

From the above, it is established that 71% of the participants supported their choice of FS in what they learned in the classroom, followed by using scientific theories (12%) and existing norms (9%) and, to a lesser extent, they turned to their experience (9%) or by what they learned in discussions with peers (2%).

Based on the above, an axial analysis was conducted to identify the central phenomenon, causal conditions, strategies, and consequences. In this analysis, the source of geotechnical knowledge is considered the cause, and the characteristics of FS selection are interpreted as the consequence. The central phenomenon is the predominance of academic knowledge over empirical results when establishing the FS, and three axial categories were established:

The first category was established as academic and theoretical. It encompasses the body of knowledge acquired in the classroom, supported by scientific theories. It has a relative weight of 83%, making it the dominant category. This category exhibited the greatest dispersion in the FS range (0.5–3.8). The higher the academic judgment, the greater the uncertainty in the decision.

The second category is use of technical standards. It includes the technical regulations governing geotechnical procedures for determining the FS. It has a relative weight of 9% and showed a medium range of uncertainty in determining the Factor of Safety (1.0–3.0).

The third category is confidence in empirical foundation. This refers to knowledge acquired through professional practice and information transmitted conversationally, with a relative weight of 8%. In this case, the decision criterion was strongly constrained to a narrower range (2.0–2.5) as the fear of making an error with the FS increased, ensuring that the response did not deviate from standard values.

4.2. Defining preference of responses generated by an AI Chatbot.

The participants organized the three ChatGPT response proposals according to their preference as a suitable response to the exercise.

The authors of this article conducted a qualitative characterization of the three proposals to better understand the process developed by ChatGPT.

Next, the value of the Factor of safety (FS) for each proposal was presented, along with the approach adopted by ChatGPT in generating each solution, and specific observations regarding each response.

First proposal (FS: 2,5, approach: conventional).

- Recommends a value between 2 and 3, which is in line with the choice of the most experienced geotechnical design engineers.
- Considering the "training" processes of AI chatbots, it can be speculated that ChatGPT has had contact with hundreds of documents that recommend this range as adequate and safe.

Second proposal (FS: 1,5, approach: incorrect).

- The allowable capacity is not calculated using soil load estimation equations but by the value of the stress the soil can physically support together with the foundation.
- The request to ChatGPT did not ask for the calculation of the allowable load, which, to be estimated, requires the value of the structural load coming from a columnar load-bearing system in relation to the foundation area of the

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

footing. Otherwise, the FS must be known in advance.

Third proposal (FS: 2, approach: less conservative than proposal 1).

— It is a popular choice among designers when faced with this geotechnical problem, as it is a safe and cost-effective option.
— It should be noted that the final clarification of ChatGPT informs us that the calculation was simplified, so it did not consider specific design factors, and it urges you to consult a qualified civil engineer.

Figure 5 shows the sequences defined by the participants when organizing the proposals. Each proposal has numerical values of 1, 2, and 3, and 0 indicates that the participant did not register a value when organizing the sequence.

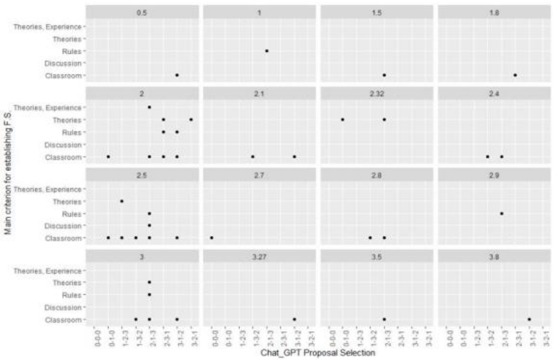

Fig. 5. *Sequences of preference of ChatGPT responses to the exercise on FS.*
Source: *own elaboration.*

The most common organizational sequence was 2-1-3: 45% of the participants chose the second proposal as the best, followed by the first and the third as the least pertinent proposal. In this sequence, the most frequent FS values were between 2.5 and 3.0, followed by 2.0, and the most frequently used criterion was knowledge acquired in the classroom.

The participants wrote a justification for choosing the best ChatGPT proposal among the three available proposals. Table 2 presents a qualitative analysis of the most common organizational sequence (2-1-3).

Table 2: *Qualitative analysis of the most common response sequence*

Reason for choice	Percentage of choice	Justification
Integration of multiple parameters	37,9%	The second proposal was chosen because it uses more data (e.g., soil characteristics) than the other proposals, which are more reliant on assumptions.
Using mathematical formula	31,0%	The second proposal was chosen because it is based on mathematical formulas and Terzaghi's theory, which conveys greater confidence than the other proposals based on elements associated with experience.
Procedure development	17,3%	The second proposal shows a more articulated and robust procedure perceived as more accurate.
Classroom learning	13,8%	The second proposal is solved using a method similar to the one explained in previous courses, such as Foundations.

Source: *own elaboration*

The following most common sequences were 1-3-2 (15% of responses), 3-1-2 (14% of responses), and 2-3-1 (6% of responses), with the most common FS values falling between 2.0 and 3.0. It should be noted that the FS responses with extreme tendency (< 2.0 and > 3.0, out of the interquartile range) are supported by the criterion of what was learned in the classroom. This element indicates that the values with the highest response tendency ($2.0 \geq FS \leq 3.0$) were based on criteria from theories, norms, discussions and experience.

Based on the above, a qualitative axial analysis was conducted, which revealed that the primary category (68.9%) indicates that the decision-making for the FS is based on rigor and objectivity, in contrast to criteria based on intuition or empirical evidence.

From this, the following subcategories emerge: Information density through the integration of multiple parameters (38%), and Theoretical and mathematical support through the application of technical formulations such as Terzaghi's theory (31%).

The second category (17.3%) is the importance of procedural development, where greater weight is given to the procedural method than to the input data. Finally, the third category (13.8%) is the relation of the topic to knowledge acquired in the classroom.

4.3. Comparison between responses generated by the participants and by ChatGPT

In the third part of the questionnaire, participants compared their responses to the exercise with the three solutions generated by ChatGPT. They evaluated the usefulness of what was proposed to understand the concept of FS better. The results are shown in Table 3.

Table 3: *Comparison of participants' responses to ChatGPT*

Question	Yes	No
Does the numerical value of your preferred ChatGPT answer match the one you solved in question 1?	23,7%	76,3%
If your answer is the same as ChatGPT's, did the method used by ChatGPT match the one you carried out?	20,0%	80,0%
Did ChatGPT's response help you improve your judgment when defining a FS?	20,0%	80,0%
If your answer is different from ChatGPT's, would you change the answer to the one given by ChatGPT?	42,2%	57,8%

Source: *own elaboration*

In this regard, a high number of participants identified that their response (76.3%) and their method (80%) did not match that of the chatbot, and a low number of participants (20%) considered that the ChatGPT response helped them improve their criteria for defining FS. However, among the participants whose response did not match ChatGPT's, 57.8% would be willing to replace it with the chatbot. From the above, a qualitative axial analysis emerges, indicating a high technical discrepancy between the student's judgment and the result provided by the AI.

The primary category (76.3%) comprises two subcategories: the first indicates that the results differ significantly between the student's judgment and the AI's output (76.3%), and the second shows that in 80% of cases, the methodological criterion was different.

The second category shows that in 80% of the cases, ChatGPT's response did not help the students redefine their answer. The third category highlights a student confidence issue: although a high percentage indicated that AI did not provide support for changing their FS value, 42% reported being willing to adopt ChatGPT's answer.

Figure 6 groups in graphic form the justifications provided by the participants who decided to replace their response with that provided by ChatGPT.

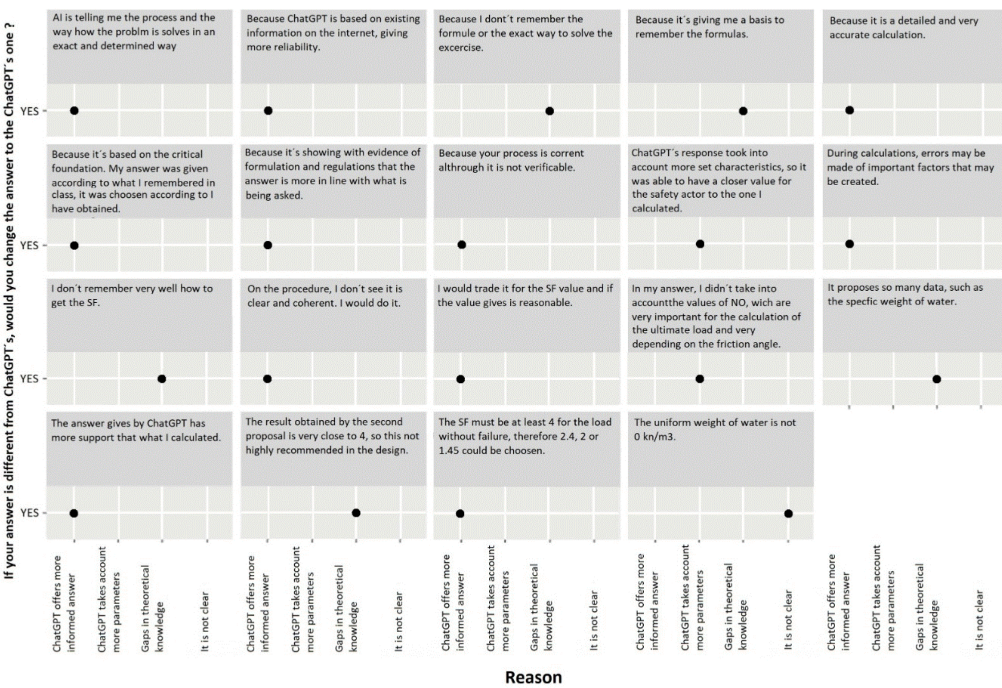

Fig. 6. *Justification for the change of response due to the ChatGPT proposal.*
Source: *own elaboration.*

When analyzing the reasons qualitatively, they are grouped into two categories: the first, linked to the clarity in the wording of the response, and the second, associated with using a more significant number of parameters in the procedure. First, a group of participants believed that ChatGPT offers a more solid justification due to the more articulate explanation of the procedure compared to the one developed by the student independently. From this perspective, it can be inferred that more articulated wording, which includes references to formulas and regulations, conveys reliability in the result. Second, some participants chose ChatGPT's response because it involved more parameters than those used by them. In this regard, including many variables in the calculations (although not all of them are useful, as happens in ChatGPT's second proposal) conveys a sense of security. Likewise, some participants also expressed that the chatbot has advantages associated with using formulas (not always remembered by participants) and access to data libraries.

Figure 7 presents a summary of the justifications expressed by the participants for maintaining their initial response without modification.

Fig. 7. *Justification for maintaining the original response.*
Source: *own elaboration.*

When performing the qualitative analysis of the reasons, these are grouped into two categories: the first category refers to the identification, by a group of participants, of inaccuracies in the argumentation carried out by the chatbot, such as access restrictions to updated standards or the identification of errors in the solution of mathematical equations.

The second category is related to the distrust of another group of participants in the response due to the impossibility of analyzing the process developed by the chatbot. In some cases, it was expressed that the FS offered by ChatGPT are proper but requires adjustments based on engineering criteria; since the calculation of FS is a trial-and-error process, it requires the integration of various factors, which are simplified by the chatbot. Associated with the above, some participants indicated that the chatbot offers help in understanding the procedure but makes mistakes that require human analysis for adjustment.

5. DISCUSSION

This study offers three contributions associated with its general objective. The first contribution is associated with the identification of the potential of

ChatGPT as a pedagogical tool in teaching complex concepts to civil engineering students. The appropriation of complex concepts associated with engineering criteria, such as the calculation of FS in a geotechnical structure, involves multiple analysis variables, and their understanding depends on the experience accumulated by the designer during his academic training and professional practice. For this reason, a tool such as ChatGPT, which allows the generation of multiple response scenarios, increases the possibilities of personalizing teaching content and encouraging critical reflection on various technical concepts and solution methods, which increases student motivation and commitment [18]. This is especially relevant in undergraduate training, where classroom learning is a key factor in students' decision-making process, as the study participants expressed. Thus, AI chatbots such as ChatGPT are established as tools that facilitate student contact with complex concepts through access to multiple examples and consultation possibilities [15].

The second contribution refers to the reinterpretation of the role of the teacher, which, far from losing importance in the face of the emergence of AI chatbots, is reinforced as a guide to the student's autonomous learning. For example, although the interquartile range tended to select an FS between 2.0 and 3.0, corresponding to a conventional approach, the organization sequence preferred by the participants was 2-1-3. In this regard, more than half of the participants chose sequences headed by the second proposal of ChatGPT. However, as previously explained, its approach to addressing the geotechnical problem was incorrect. According to the participants, it is possible that, due to factors such as the inclusion of a more significant number of mathematical formulas and procedural steps or language correction, the second proposal is perceived as more solid than the remaining proposals, which are based on a more conventional interpretation of the FS concept seen from the engineering criterion. It is also important to note that 71% of the participants supported their choice of FS in the knowledge learned in the classroom, which shows the impact of the relationship between teacher and student. Faced with this aspect, the teacher's inducing the student to analyze the chatbot's responses actively is essential to prevent excessive reliance on technology, affecting critical thinking skills and the ability to make autonomous decisions [2].

In this context, incorporating tools such as ChatGPT in teaching civil engineering does not eliminate the teaching task; on the contrary, it becomes essential in its role as a guide. In the face of this technological change, the teacher plays a crucial role in encouraging students to learn independently and providing their knowledge and experience in guiding students in identifying, processing and analyzing information offered by chatbots. The third contribution relates to the importance of adapting the type of training activities, which should focus on incorporating AI chatbots in the development of procedures, with a greater emphasis on critical reflection associated with decision-making. In the exercise proposed in this study, it is identified that the use of ChatGPT without guidance by the teacher is not sufficient to achieve a learning objective: 80% of the participants used a method other than ChatGPT, and the same percentage considered that the chatbot's proposals did not contribute to a greater understanding of the FS concept.

Also, a fundamental objection expressed by participants, which coincides with what was identified by [17], is the concern about the accuracy, updating and authenticity of the data generated by the chatbot. Regarding this, as [2] reminds us, if the data used to train the chatbot is not diverse or is of poor quality, errors will be generated and, in this situation, the teacher must exercise his role as a tutor in guiding the student in the analysis of the quality of the answers. It is worth noting that, among those participants whose answers did not coincide with ChatGPT's, 42.2% would replace their solution with that of the chatbot. This represents a challenge in teaching since the confidence transmitted by these types of tools, which are based on training through millions of entries on the subject, can bias the student and make him think that the chatbot's answer is "correct," even if it does not contribute to a greater understanding of the underlying theoretical concept [2].

Regarding integrating AI chatbots in engineering education, the challenge arises of developing training activities in which the critical analysis of the procedure is prioritized, not the result itself. On the other hand, the opportunity emerges to use the chatbot as an assistant that facilitates immediate access to information for problem-solving (formulas, regulations, among others), but without giving up decision-making, which is the user's responsibility.

In this type of academic exercise, the responses reflect the perspective of a student taking an exam, rather than the application of an engineer's judgment in a professional task.

6. CONCLUSION

This study focuses on a still little-explored field of research by analyzing the contributions of natural language processing tools such as ChatGPT to higher education, particularly the teaching of complex civil engineering concepts.

AI chatbots allow for the evaluation of different solution scenarios, providing a dynamic pedagogical tool, especially in teaching situations where students have limited real-world experience. This article analyzed the possibilities of integrating ChatGPT into the learning of a complex concept such as the FS, which is intrinsically related to engineering judgment, with the teaching process focused on discussing the relevance and applicability of the results proposed by the tool.

However, the challenges associated with the inappropriate use of ChatGPT are also highlighted due to problems not detected by the student in the responses generated or due to the weakening of critical skills and excessive confidence in the tool's effectiveness.

Among the limitations of the study, it focused solely on a specific concept, the factor of safety (FS), within a particular area of civil engineering (geotechnics). Limitations related to its application in a single university course and the use of a single AI chatbot (ChatGPT) as the solution generator should also be noted. For future research, expanding the sample size and including participants from other geographic locations, for example within the Latin American context, could provide additional insights into the possibilities offered by LLMs for teaching engineering judgment in the classroom.

Future research directions include the integration and comparison of results across AI chatbots (e.g., Claude, Gemini, or DeepSeek) and involving participants at different stages of engineering judgment training. This would allow exploration of the advantages and risks of incorporating AI into university curricula for teaching complex engineering concepts.

Finally, the adoption of tools such as ChatGPT, as proposed by [4,17], constitutes an opportunity for the development of research in engineering education, where both the exploration of the modification of curricula and the reflection on the ethical implications of their use in the classroom are involved, in which AI chatbots are understood, with their strengths and limitations, as complements, and not as substitutes for traditional education.

Conflict of interest

The authors declare that they have no conflict of interest related to the preparation of this article. None of the co-authors is the teacher of the course in which the questionnaire was applied.

ACKNOWLEDGEMENT

The authors thank Martha Moreno Igua, professor of the Department of Humanities at Universidad El Bosque, for her valuable guidance in applying discursive strategies in writing this article.

The authors express their gratitude to Universidad Militar Nueva Granada (Colombia) for its willingness and support in the processes that facilitate research outcomes such as this one.

REFERENCES

[1] H. Yu, “Reflection on whether Chat GPT should be banned by academia from the perspective for education and teaching”. *Frontiers in Psychology*, vol. 14:1181712. June, 2023, doi: https://doi.org/10.3389/fpsyg.2023.1181712

[2] K. Fuchs, “Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse?”. *Frontiers in Education*, vol. 8:1166682. May, 2023, doi: https://doi.org/10.3389/feduc.2023.1166682

[3] T. Meng-Lin, W. O. Chong, C. Cheng-Liang, “Exploring the use of large language models (LLM) in chemical engineering education: Building core course problems with Chat-GPT”. *Education for Chemical Engineers*, vol. 44, pp. 71-95. July, 2023, doi: https://doi.org/10.1016/j.ece.2023.05.001

[4] S. Nikolic, S. Daniel, R. Haque, M. Belkina, G. M. Hassan, S. Grundy, S. Lyden, P. Neal y C. Sandison, “ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity”, *European Journal of Engineering Education*, vol. 48, No. 4, pp. 559-614. May, 2023, doi: https://doi.org/10.1080/03043797.2023.2213169

[5] L. M. Sánchez-Ruiz, S. Moll-López, A. Nuñez-Pérez, J. A. Moraño-Fernández, E. Vega-Fleitas, "ChatGPT Challenges Blended Learning Methodologies in Engineering Education: A Case Study in Mathematics". Applied Sciences, vol. 13, No. 10: 6039. May, 2023, doi: https://doi.org/10.3390/app13106039

[6] A. Patel, "Interpretations after a geotechnical failure", in *Woodhead Publishing Series in Civil and Structural Engineering, Geotechnical Interpretations in Field Practice*, A. Patel, ed. Cambridge, MA: Woodhead Publishing, 2024, Chapter 7, pp. 145-152. doi: https://doi.org/10.1016/B978-0-443-14092-1.00017-5

[7] J. C. Musto, "The safety factor: case studies in engineering judgment". International Journal of Mechanical Engineering Education, vol. 38, No. 4. November, 2010, doi: https://doi.org/10.7227/IJMEE.38.4.2

[8] J. Haurylkiewicz, "Critical analysis of the method of safety factors in geotechnics", in *International Conference on Applications of Statistics and Probability in Soil and Structural Engineering*, Sydney, Australia, January - February, 1979.

[9] D. Naylor, "Quantifying safety", *Ground Engineering*, vol. 14, no. 7. 1981.

[10] G. N. Smith, "The use of probability theory to assess the safety of propped embedded cantilever retaining walls". *Geotechnique*, vol. 35, no. 4. 1985.

[11] J. Oliphant, "Controlling the safety of geotechnical structures: a proposed approach". *Geotechnical and Geological Engineering*, Vol. 10. 1992.

[12] J. C. Ruge, O. H. Vargas, J. E. Carmona, "Desafíos en la definición de factores de seguridad en el diseño de estructuras geotécnicas", *Revista Colombiana de Tecnologías de Avanzada*, Vol. 2, No. 40. September, 2022, doi: https://doi.org/10.24054/rcta.v2i40.2354

[13] I. Rizos, N. Gkrekas, "The Impact of LLM on Mathematics education and research at the university", *Social Sciences & Humanities Open*, vol. 12: 101969. September, 2025. DOI: https://doi.org/10.1016/j.ssaho.2025.101969

[14] S. N. Yeh, C. J. R. Siah, "Students' attitudes toward LLMs and its association with metacognitive abilities: a cross-sectional study", *Nurse Education in Practice*, vol. 88: 104567. September, 2025. DOI: https://doi.org/10.1016/j.nepr.2025.104567

[15] D. Kim, T. Kim, Y. Kim, Y. H. Byun, T. S. Yup, "A ChatGPT-MATLAB framework for numerical modeling in geotechnical engineering applications", *Computers and Geotechnics*, vol. 169: 106237. May, 2024. DOI: https://doi.org/10.1016/j.compgeo.2024.106237

[16] M. Aluga, "Application of ChatGPT in civil engineering", *East African Journal of Engineering*, vol. 6, No. 1. pp. 104-112. June, 2023, doi: https://doi.org/10.37284/eaje.6.1.1272

[17] P. P. Ray, "ChatGPT in transforming communications in seismic engineering: Case studies, implications, key challenges and future directions", *Earthquake Science*, vol. 37, No. 4, pp. 352-367. August, 2024, doi: https://doi.org/10.1016/j.eqs.2024.04.003

[18] M. Talha Junaid, M. Barakat, S. Awad, N. Anwar, "Adopting the power of AI Chatbots for Enriching Students Learning in Civil Engineering Education: a Study of Capabilities and Limitations", in *Artificial Intelligence in Education: the Power and Dangers of ChatGPT in the Classroom,* A. Al-Mazouqi, S.A. Salloum, M. Al-Saidat, A. Aburayyam B. Gupta, eds. Switzerland: Springer Nature, 2024, Chapter 3, pp. 25-48. doi: https://doi.org/10.1007/978-3-031-52280-2

[19] S. M. J. Uddin, A. Albert, M. Tamanna, A. Ovid, A. Alsharef, "ChatGPT as an educational resource for civil engineering students", *Computers Applications in Engineering Education*, vol. 32, No. 4. April, 2024, doi: https://doi.org/10.1002/cae.22747

[20] Z. Xue, C. Xu, X. Xu, "Application of ChatGPT in natural disasters prevention and reduction", *Natural Hazards Research*, vol. 3, No. 3, pp. 556-562. September, 2023, doi: https://doi.org/10.1016/j.nhres.2023.07.005

[21] S. Wu, C. Shi, Y. F. Leung, Y. Otake, C. Konishi, M. Zhou, Y. Tao, Z. Cao, T. Nakamura, "Perspectives: LLM agents reshaping the foundation of geotechnical problem-solving", *Geodata and AI*, vol. 4: 100036. September, 2025. DOI: https://doi.org/10.1016/j.geoai.2025.100036

[22] T. Kim, T.S. Yun, H.S., Suh. "Can ChatGPT Implement Finite Element Models for Geotechnical Engineering Applications?", *International Journal for Numerical and Analytical Methods in Geomechanics.* Vol.49, No. 6. April, 2025, *doi:* https://doi.org/10.1002/nag.3956

[23] T. Kim, D, Kim, T.S. Yun. "ChatGPT-driven Numerical Modeling for Geotechnical Applications". *Journal of the Korean Geotechnical Society*, Vol.41, No.4. August, 2025, doi: https://doi.org/10.7843/kgs.2025.41.4.99

[24] A. Parsa-Pajouh. "Application of generative AI to automate numerical analysis and synthetic data generation in geotechnical engineering*". Machine Learning and Data Science in Geotechnics.* Vol. 1. No.1. December, 2025. doi: https://doi.org/10.1108/MLAG-09-2024-0008

[25] A. C. Niloy, M. A. Bari, J. Sultana, R. Chowdury, F. M. Raisa, A. Islam, S. Mahmud, I. Jahan, M. Sarkar, S. Akter, N. Nishat, M. Afroz, A. Sen, T. Islam, M. H. Tareq, M. A. Hossen, "Why do students use ChatGPT? Answering through a triangulation approach", *Computers and Education: Artificial Intelligence*, vol. 6: 100208.

Jan, 2024. DOI: https://doi.org/10.1016/j.caeai.2024.100208

[26] ChatGPT. OpenAI. Accessed: Dec. 1. 2025. [Online]. Available: https://chatgpt.com/