ISSN: 1692-7257 - Volume 1 – Issue 47 - 2026

ORIGINAL

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

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

Óscar Hernán Vargas Villamizar[1]
Juan Carlos Ruge Cárdenas[2]
Carlos Felipe Urazán Bonells[2]
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.**

Received: 02 July 2025. Revised: 02 September 2025. Accepted: 20 December 2025. Published: 01 January 2026.
Corresponding Author: Óscar Hernán Vargas Villamizar

How to cite: Ó. H. Vargas Villamizar, J. C. Ruge Cárdenas, C. F. Urazán Bonells, and D. Palma Cuero, "Uso de ChatGPT en la enseñanza del concepto de factor de seguridad en geotecnia", *Revista Colombiana de Tecnologías de Avanzada (RCTA)*, vol. 1, no. 47, ene. 2026, art. no. 4076, doi: 10.24054/rcta.v1i47.4076

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

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.

INTRODUCTION

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

Among the advantages of using AI chatbots in higher education is the capacity 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 represent a high risk for ensuring educational integrity[3], because of the difficulty in detecting texts generated by chatbots. Likewise, the indiscriminate use of AI chatbots can undermine students' confidence in their personal abilities and reduce their critical thinking skills, which will affect their academic results in the future[1].

Faced with the challenge posed by AI chatbots in higher education,[4]consider that prohibition is not an option in engineering teaching, whose body of knowledge demands the incorporation of new technologies. As for curricular modifications, activities must be designed 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[4].

In view of this, teachers must develop a greater understanding of the strengths of AI chatbots (for example, writing programming code) and of the aspects in which their contributions are still limited (for example, the formulation of initial equations)[3].

AI has been identified as capable of solving basic engineering problems efficiently[5]. However, AI still has limitations in the integration of ethical and contextual considerations into decision-making processes. This is relevant in the field of civil engineering, in which design failures cause substantial economic losses, or even the loss of human lives.

Apart from the use of technical procedures and mathematical formulas, the practice of civil engineering 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 creative exercise of high complexity[6]. Engineering judgment is especially relevant in the development of geotechnics projects, in which one depends on the characteristics of the soil and on the conditions of the physical construction site, both conditions being heterogeneous and non-modifiable.

In the process of ensuring the stability of a geotechnical structure, it is necessary to determine a safety factor (SF) through the use of heuristic methods, guided by engineering judgment, which take into account aspects such as the uncertainty associated with the context, the risk that one is willing to assume, current legislation and the technical standards accepted in the discipline[7]. The definition of an SF depends on each engineer's judgment, based on their experience, their technical training and even their personality[7]. This makes teaching the SF to an undergraduate student a challenge, since it is not a "given" value, but rather corresponds to an elusive concept[7], which is formed through a learning process that is difficult to replicate in the classroom. Given this, in the university setting it is essential to go beyond the explanation of technical factors associated with the calculation of the SF, through the integration of discussions on the possible biases that a designer adopts in the calculation of the SF in geotechnics, caused by their academic training, the conditions of the geotechnical problem addressed, or the idiosyncrasy of the local context[8]-[12].

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

THEORETICAL FRAMEWORK

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

Because the emergence of AI chatbots such as ChatGPT is a recently appearing phenomenon,[4]highlight the scarcity of relevant studies on their effects in the field of education. Likewise,[15]mention that a limited number of articles have been published on the potential of ChatGPT in the engineering field. In agreement with the above,[16]points out the existence of a restricted number of publications on the application of ChatGPT in civil engineering. Exploration of the application of ChatGPT in the field of geotechnics is still incipient.

Table 1 presents a summary of articles published, as of September 2025, in two bibliographic databases (Scopus and Science Direct) on the use of ChatGPT in the field of geotechnics.

Table 1: Studies relating ChatGPT to geotechnics

Author(s)	Comments
[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 legislators.
[18]	Develops a critical reflection on the use of ChatGPT in engineering education. 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 university in the United States, the authors identified a positive reception of the chatbot's use among the study participants, which offers new opportunities to improve tertiary education in engineering.
[20]	Highlights the potential of ChatGPT in scientific outreach on disaster prevention.
[21]	Analyzes the potential of using LLMs to address classic geotechnics problems associated with soil analysis, slope stability and foundation design.
[22]	The study examines the capacity ChatGPT may have to generate finite element code in geotechnical engineering applications from a set of prompts
[23]	This study investigates the potential and limitations of automating the numerical modeling process in geotechnical problems using ChatGPT.
[24]	This study 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, making it one of the first efforts to explore the use of this AI platform for educational purposes among engineering students. In identifying the limitations of their research,[19]also note that they focused only on introductory content on construction surveying and geomatics, and therefore invite the exploration of more advanced concepts and analysis beyond these two areas of civil engineering. Related to this,[21]note that the application of Machine Learning and other AI techniques has shown significant growth in the field of geotechnics, especially in soil classification, slope stability analysis and foundation design, and therefore calls for educational initiatives to increase knowledge about AI within the engineering community.

Given this scenario, academic integrity stands as a key pillar in educational institutions regarding the use of tools such as ChatGPT. On this point,[25]recommends that educators allow it only as a self-learning tool, in the manner of a complementary tutor, permitted as a grammar and paraphrasing checker but not as a content generator. To this end,[25]states that it is essential for teachers and educational administrators to understand the limitations of these models in order to communicate them to students.

This study is situated in a little-explored field of research: the integration of AI chatbots in the teaching of civil engineering in higher education. In particular, it focuses on analyzing students' perceptions of the pertinence of the responses generated by ChatGPT when solving an exercise involving a complex geotechnics concept, the safety factor (SF), because its calculation entails the application of engineering judgment.

METHODOLOGY

Design and participants

A descriptive study was conducted. From the population, made up of students enrolled in an undergraduate program in civil engineering, those enrolled in the Geotechnics course were selected; this course is taught in the seventh semester of a ten-semester program. 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, given its location in the curriculum, it ensured that all participants had a basic knowledge of the concept of SF, having previously taken the subjects of Soil Mechanics and Foundations.

The instrument consisted of a physical questionnaire, administered in October 2023 and completed in a single session. At the beginning of the session, the purpose of the research was explained to the participants, and they were informed that the results would be treated anonymously, without affecting their course grade. Informed consent was included at the top of the questionnaire, and only the information of participants who agreed to participate voluntarily was collected. The questionnaire was completed in a controlled environment, to prevent the use of digital tools, with a time of 90 minutes to complete it. The questionnaire consists of seven questions (four closed and three open) organized into three parts (see Figure 1).

Instruments and procedures developed

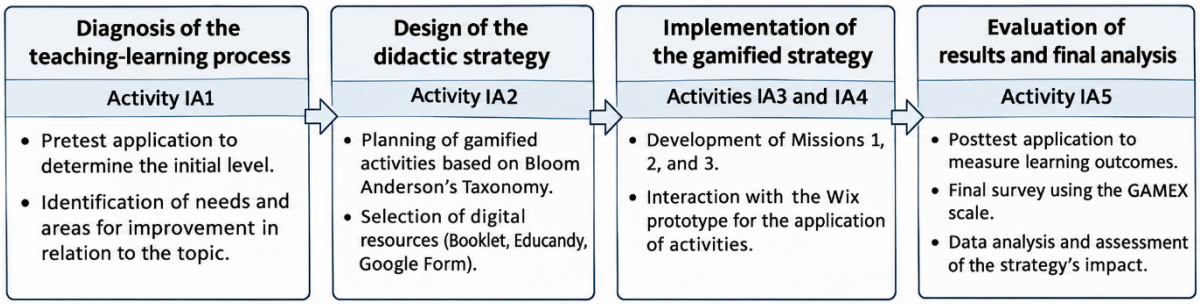

Fig. 1. *Structure of the questionnaire*
Source: Own elaboration.

In the first part of the questionnaire, the participants answered two questions: in the first, they were given an SF calculation exercise for a geotechnical foundation, with an open response. In the solution of this exercise, access to the internet was not permitted, nor was it permitted for the participants to know the following questions of the questionnaire in advance. In the second question, they had to choose, from a list, the method used to find the numerical response for SF.

The exercise solved by the participants is presented below.

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

- Effective internal friction angle: 25°
- Effective cohesion: 20kN/m
- Unit weight: 16.5 kN/m³

If the ultimate bearing capacity calculated for the soil is 1078.29 kN/m², what safety factor (SF) do you recommend to find the allowable bearing capacity of the foundation? Write your response as a number (not as a range).

This same statement, without modifications or additional instructions, was used by the authors as a prompt to request the response to the exercise from ChatGPT. The free-access version available in October 2023 was used, whose underlying model was GPT -3.5 (gpt-3.5 / gpt-turbo family). According to[26], OpenAI did not publish the complete hyperparameter table or an official number of parameters for the version used in this article. The free version of ChatGPT was used because, according to[26], between 95% and 97% of users used this version in 2023, making it the most accessible option in the university setting.

In the statement of the proposed exercise, the key aspects in the generation of a prompt (text to communicate with generative AI systems) were taken into account: 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 formulation of this type of activity in a classroom. That is, the input parameters of the exercise were provided with realistic values in order to prevent, for both the chatbot and the student, the assumption of any parameter required for the solution of the problem.

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

Considering that the focus of this research is the application of engineering judgment, the students did not have access to their electronic devices at any point during the activity, nor did they write prompts, since there is a single one, corresponding to the statement of the proposed exercise.

Once all participants had completed the first part, they were allowed to consult the second part of the printed questionnaire, in which they compared their responses to the geotechnical exercise with the three proposals generated by ChatGPT for the same exercise.

The second part consists of two questions. In the first, a closed-ended question, the three proposals generated by ChatGPT as a response to the FS calculation exercise were shared, and the participant ranked them in order of preference, assigning the number 1 to the one considered the best response, the number 2 to the second-best proposal, and the number 3 to the least appropriate proposal as a response to the exercise. In the second question, an open-ended one, each participant included a justification of the reason for choosing one of the three proposals as the most suitable for answering the exercise. In the third part, consisting of three open-ended questions, the participant was asked to compare their response with the one generated by ChatGPT, and was asked whether they would be willing to replace their response with the chatbot's, or whether they would

keep their proposal.

Data analysis

The information from the printed questionnaires was entered into a spreadsheet and analyzed using Rstudio software for the statistical and graphical analysis of the responses. The response trend of the participants regarding the FS value for the proposed exercise was analyzed through 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=" ").

Next, the trend of the criterion used by the participants to choose the FS value 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, whose results can be consulted in Figure 4. The graphical analysis was performed using the following command:

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 a dot-plot style with "facet wrap", the participants' qualitative justifications were 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 the participants who decided to replace their response with that of ChatGPT, and Figure 7 presents the results of those who decided to keep their original response.

To make the study clearer, an axial qualitative analysis was applied to the criteria for answering 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 result (see Table 3).

RESULTS

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

Determination of the FS in a geotechnical structure exercise.

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

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 data or participant 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 in a range between 2.0 and 2.5, Q3 is in a range between 2.5 and 3.0, and Q4 is between 3.0 and 3.8.

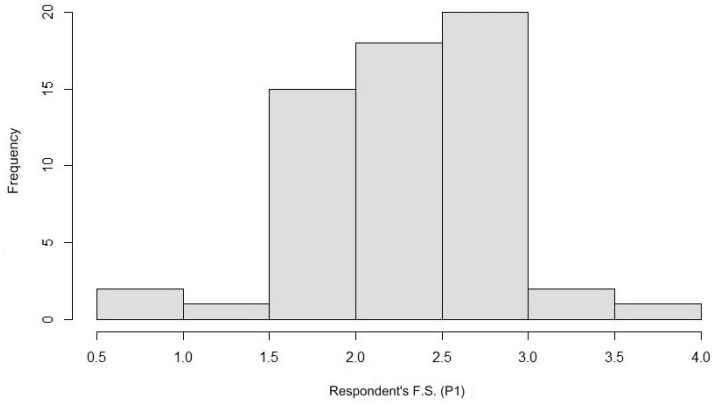

Fig. 2. *Frequency of the response to the FS exercise.*
Source: own elaboration.

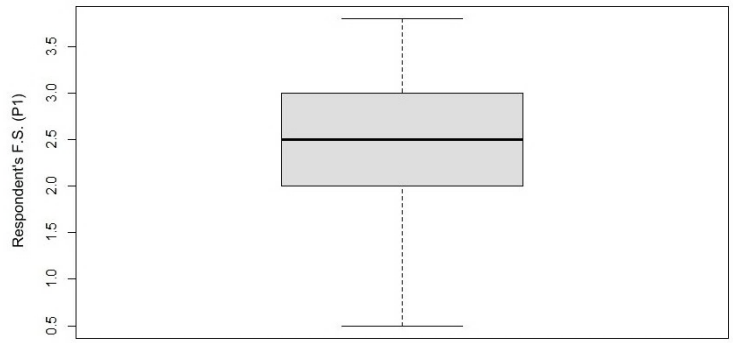

Fig. 3. *Interquartile range of the response to the FS exercise.*
Source: own elaboration.

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 larger selection range, which gives a higher probability that the user will choose this range, even in the event of not understanding the question and answering any alternative at random. When asking the participants 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.

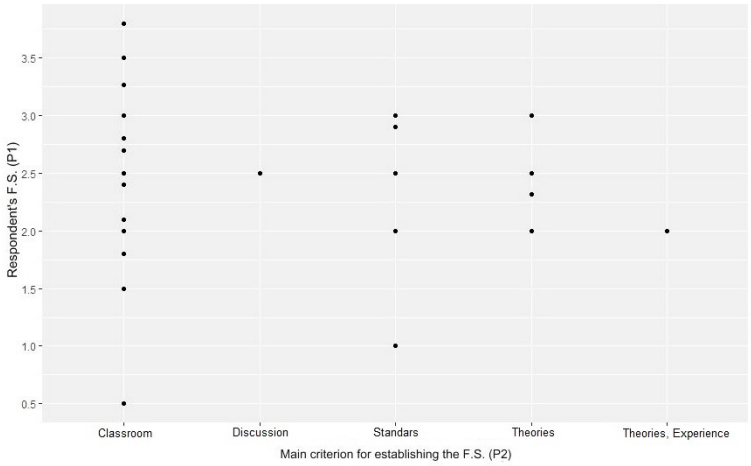

Fig. 4. *Criteria used to answer the exercise.*
Source: own elaboration.

The highest frequency for a safety factor (SF) selection criterion was “knowledge learned in the classroom,” distributed across the entire range of values, from 0.5 to 3.8. The options on “technical standards and regulations” and “scientific and mathematical theories” are the second and third most used criteria, with values between 1.0 and 3.0, respectively. Finally, the options associated with knowledge acquired in professional discussions and experiences were the least used criteria, with SF responses between 2.0 and 2.5.

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

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

The main category was established as academic and theoretical. It accumulates the compendium of knowledge acquired in the classroom supported by scientific theories. It has a relative weight of 83%, being the dominant category. This category presented the greatest dispersion in the SF range (0.5 – 3.8). The greater the academic criterion, the greater the uncertainty in the decision.

The second category is the use of technical regulations. It comprises the technical regulations that govern geotechnical procedures for determining the SF. It has a relative weight of 9%. This category presented a medium range of uncertainty

when determining the safety factor (1.0 – 3.0).

The third category is confidence in the empirical foundation. It refers to knowledge acquired in professional practice and in information transmitted conversationally, and had a relative weight equivalent to the remaining 8%. In this case, the decision criterion was strongly reduced to a narrower range (2.0 – 2.5) as the fear of erring with the SF increased, causing the response not to move away from standard values.

Definition of preference for responses generated by an AI chatbot.

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

The authors of this article carried out a qualitative characterization of the three responses, in order to gain a greater understanding of the process developed by ChatGPT. Below, the safety factor (SF) value for each of the proposals, the type of approach adopted by ChatGPT when generating each solution, and specific observations about each response text are presented.

First proposal (SF: 2.5, approach: conventional).

- It recommends a value between 2 and 3, which aligns with the choice of most designers with experience in geotechnical design.
- 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 (SF: 1.5, approach: incorrect).

- The calculation of the allowable capacity is not given by equations for estimating the soil load, but by the value of the stress that the soil, together with the foundation, can physically withstand.
- The request submitted to ChatGPT did not ask for the calculation of the allowable load, which, to be estimated, requires the value of the structural load from a moment-resisting frame system in columns, in relation to the foundation area of the footing. Otherwise, the FS must be known beforehand.

Third proposal (SF: 2, approach: less conservative than the first proposal).

- It is a popular option among designers when facing this type of geotechnical problem, by resorting to a safe and cost-effective option.
- ChatGPT's final clarification should be highlighted, as it states that the calculation is simplified, and therefore did not consider specific design factors, and it recommends consulting a certified civil engineer.

Figure 5 presents the sequences defined by the participants when organizing the proposals. The numerical values of 1, 2 and 3 are available for each proposal; and 0 indicates that the participant did not record a value when organizing the sequence.

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

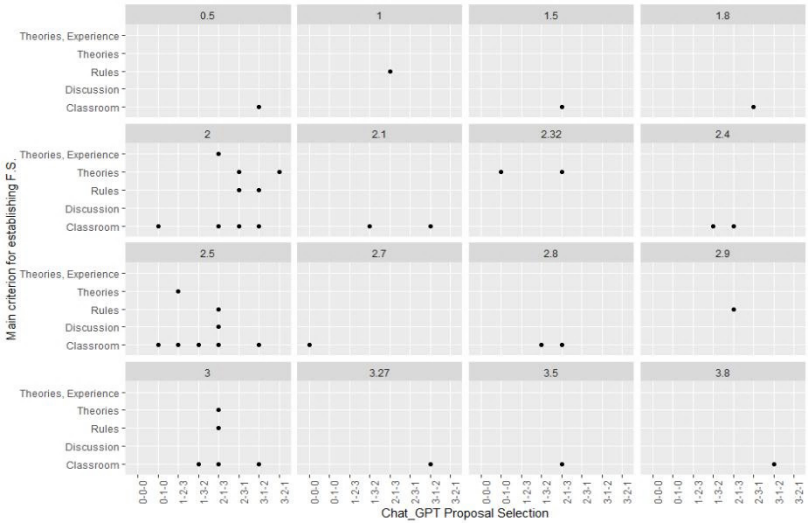

Fig. 5. *Preference sequences of ChatGPT's responses to the exercise on FS.*
Source: own elaboration.

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 (for example, soil characteristics), compared with the other proposals, which depend more heavily on assumptions.
Use of a mathematical formula	31,0%	The second proposal was chosen because it is based on mathematical formulas and Terzaghi's theory, which conveys greater confidence compared with the other proposals, which are based on elements associated with experience.
Development of a procedure	17,3%	The second proposal shows a more articulated and robust procedure, which is perceived as more truthful.
Classroom learning	13,8%	The second proposal is solved with a method comparable to the one explained in previous courses, such as Foundations.

Source: own elaboration

The next most common sequences were as follows: 1-3-2 (15% of the responses), 3-1-2 (14% of the responses), and 2-3-1 (6% of the responses), and the most common FS values lie in the range between 2.0 and 3.0. It should be noted that the FS responses with an extreme tendency (< 2.0 and > 3.0, outside the interquartile range) rest on the criterion of what was learned in the classroom. This indicates that the values with the strongest response tendency ($2.0 \geq FS \leq 3.0$).were based on criteria derived from theories, standards, discussions, and experience.

From this, a qualitative axial analysis was carried out, which showed that the main category (68.9%) indicates that decision-making for the FS is based on a criterion of rigor and objectivity, in contrast with a criterion based on intuition or empirical evidence.

From this, the following subcategories emerge:

The density of the information through the integration of multiple parameters (38%) and the theoretical and mathematical support from the application of technical formulations such as Terzaghi's theory (31%).

The second category (17.3%) is the importance of procedural development. In this case, greater importance is given to the procedural method than to the input data. Finally, the third category (13.8%) is the relationship of the topic to be addressed with the knowledge acquired in classrooms.

In the third part of the questionnaire, participants compared their response to the exercise with the three solutions generated by ChatGPT and assessed how useful the proposed solutions were for gaining a better understanding of the concept of FS. The results are presented in table 3.

Comparison between responses generated by the participants and by ChatGPT

Table 3: comparison of the participants' responses versus ChatGPT

Question	Yes	No
Does the numerical value of your preferred ChatGPT response match the one you solved in question 1?	23,7%	76,3%
If your response 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 an FS?	20,0%	80,0%
If your response differs from ChatGPT's, would you change your response 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 the chatbot's, and a low number of participants (20%) considered that ChatGPT's response helped them improve their judgment when defining an FS. Nevertheless, among the participants whose response did not agree with ChatGPT's, 57.8% would be willing to replace it with the chatbot's. From this, an axial-type qualitative analysis emerges, which establishes that there is a high technical discrepancy between the student's judgment and the result provided by the AI.

The main category (76.3%) establishes two subcategories: the first indicates that the results differ strongly between the student's judgment and the AI's result (76.3%), and the second states that in 80% of the cases the methodological judgment was different.

As a second category, 80% of the cases state that ChatGPT's response did not help the student redefine their response. As a third category, a student trust problem is established: although a high percentage stated that the AI did not support changing their opinion regarding the FS value, 42% say they are willing to change their response to ChatGPT's.

Figure 6 graphically groups the justifications provided by the participants who decided to replace their response with the one provided by ChatGPT.

A qualitative analysis of the reasons groups them into two categories. First, a group of participants considered that ChatGPT offers a more solid substantiation, because its explanation of the procedure is more articulate than the one the student develops autonomously. From this perspective, one may infer that a more articulate wording, which includes mentions of formulas and standards, conveys reliability in the result.

Second, some participants chose ChatGPT's response because it involved more parameters than they used. In this regard,

including a greater number of variables in the calculations (even if not all of them are useful, as occurs in ChatGPT's second proposal) conveys a sense of security. Some participants also state that the chatbot has advantages associated with the use of formulas (which the participants do not always remember) and with access to data libraries.

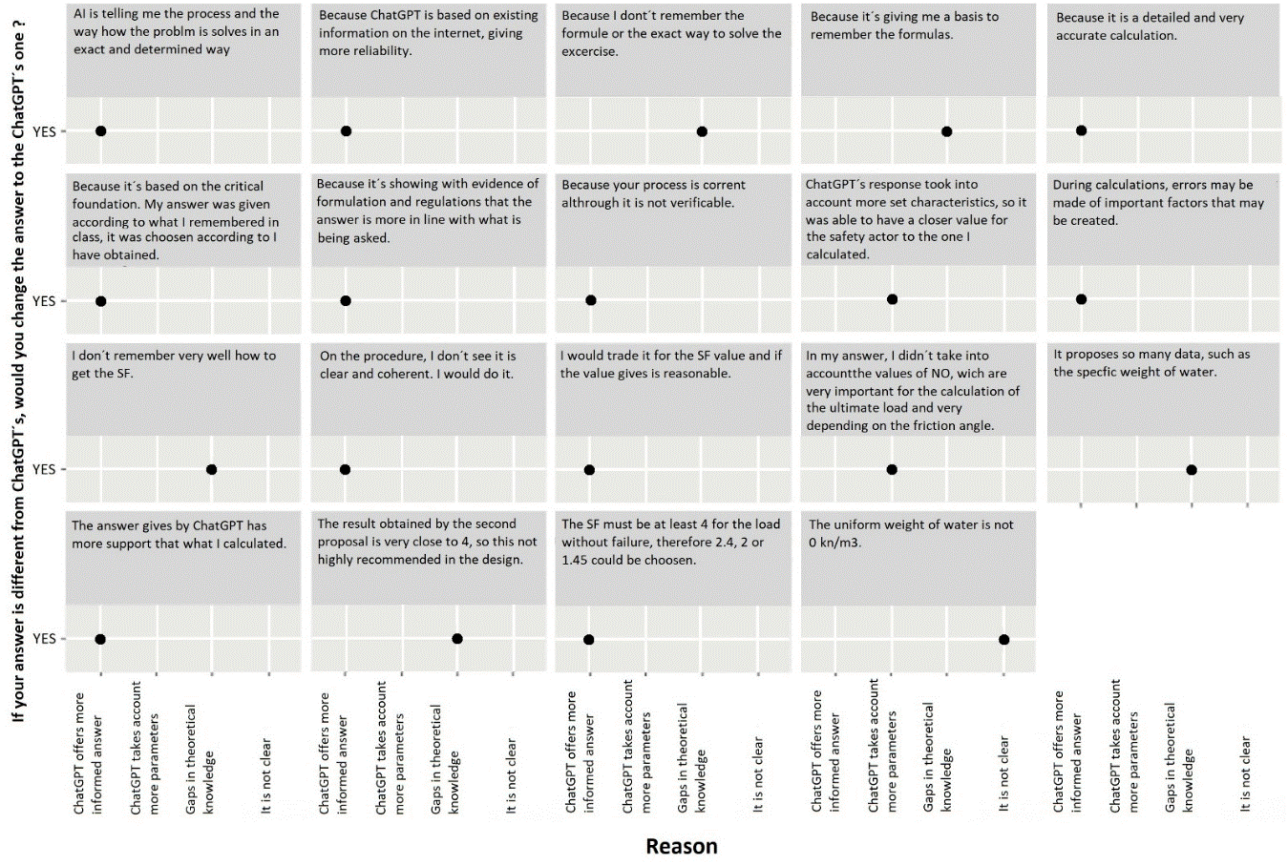

Fig. 6. *Justification for changing the response in light of ChatGPT's proposal.*
Source: own elaboration.

Figure 7 presents a compendium of the justifications expressed by the participants for keeping their initial response unchanged.

The qualitative analysis of the reasons groups them 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 restrictions on access to updated standards or the identification of errors in solving mathematical equations.

The second category is related to the distrust of another group of participants regarding the response, due to the impossibility of analyzing the process developed by the chatbot. In some cases, participants expressed that the FS offered by ChatGPT are useful but require adjustments based on engineering judgment, since the calculation of FS is a trial-and-error process and therefore requires the integration of diverse factors, which the chatbot simplifies. Related to the above, some participants indicated that the chatbot offers help in understanding the procedure but makes errors that require human analysis for their adjustment.

DISCUSSION

This study offers three contributions associated with its general objective. The first contribution is associated with identifying ChatGPT's potential as a pedagogical tool in teaching complex concepts to civil engineering students. The appropriation of complex concepts associated with engineering judgment, 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 their 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 fostering critical reflection on various technical concepts and solution methods, which increases students' motivation and engagement[20]. This is especially relevant in undergraduate training, in which classroom learning is a key factor in students' decision-making process, as expressed by the study participants. Thus, AI chatbots such as ChatGPT stand as tools that facilitate the student's contact with complex

concepts, through access to multiple exercises and possibilities for consultation[17].



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

The second contribution refers to the reinterpretation of the teacher's role, which, far from losing importance with the emergence of generative AI chatbots, is reinforced as a guide for the student's autonomous learning. For example, although the interquartile range showed a tendency to select an FS between 2.0 and 3.0, which corresponds 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 ChatGPT's second proposal, even though, as explained previously, its approach to addressing the geotechnical problem was incorrect.

According to what the participants expressed, it is possible that, due to factors such as the inclusion of a greater number of mathematical formulas and procedural steps, or language correction, the second proposal is perceived as more robust than the remaining proposals, which rely on a more conventional interpretation of the concept of FS seen from the perspective of engineering judgment. It is also important to note that 71% of the students supported their choice of FS with the knowledge learned in the classroom, which shows the impact of the teacher-student relationship. In view of this, the teacher's action of inducing the student to actively analyze the chatbot's responses is fundamental to prevent excessive trust in the technology, which can affect critical thinking skills and the capacity for autonomous decision-making[2].

In this context, incorporating tools such as ChatGPT into civil engineering teaching does not eliminate the teaching task; on the contrary, that task becomes more important in its role as a guide. Faced with this technological change, the teacher plays a key role in encouraging the student's autonomous learning and in contributing their knowledge and experience to guiding the student in identifying, processing and analyzing the information offered by chatbots. The third contribution concerns the importance of adapting the type of formative activities, which should focus on incorporating AI chatbots into the development of procedures, with greater emphasis on the critical reflection associated with decision-making. In the exercise proposed in this study, the use of ChatGPT without guidance from the teacher is found not to be sufficient to achieve a learning objective:

80% of the participants used a method different from that of ChatGPT, and the same percentage considered that the chatbot's proposals did not contribute to a greater understanding of the concept of FS.

Also, a key objection expressed by the participants, which coincides with what was identified by[19], is the concern about the accuracy, updating and authenticity of the data generated by the chatbot. In this regard, as[2]recalls, if the data used to train the chatbot are not diverse, or are of poor quality, errors will be generated, and in this situation the teacher must exercise their role as tutor in guiding the student in the analysis of the quality of the responses. It is worth highlighting that, among those participants whose response did not coincide with that of ChatGPT, 42.2% would replace their solution with that of the chatbot. This represents a challenge in teaching, since the confidence conveyed by this type of tool, which relies on training through millions of inputs on the topic, can bias the student and make them think that the chatbot's response is "correct", even though it does not contribute to a greater understanding of the underlying theoretical concept[2].

Regarding the integration of the use of AI chatbots in engineering teaching, the challenge arises of developing formative activities in which critical analysis of the procedure is prioritized, and 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 renouncing decision-making, which is the user's responsibility.

Given this type of academic exercise, it is clear that the stance of the responses is that of a student answering an exam, rather than the application of an engineer's judgment in a professional exercise.

CONCLUSIONS

This study focuses on a research field that is still little explored, by analyzing the contributions of natural language processing tools such as ChatGPT in higher education and, in particular, in the teaching of complex civil engineering concepts.

AI chatbots allow different solution scenarios to be evaluated, which offers a dynamic pedagogical medium, especially in teaching situations in which the student has little experience in real settings. This article analyzed the possibilities of integrating ChatGPT into the learning of a complex concept such as the SF, intrinsically related to engineering judgment, in which the teaching process focuses 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 generated responses, or to the weakening of critical skills and excessive confidence in the effectiveness of the tool.

Among the limitations of the study, it focused only on a specific concept, the safety factor (SF) in a particular area of civil engineering (geotechnics). Mention should also be made of the limitations associated with the application in a single university course, and the use of a single AI chatbot (ChatGPT) as the generator of solution responses. As a projection for future research, expanding the sample size and including participants in other geographic locations, for example, within the Latin American context, could offer more inputs for the analysis of the possibilities that LLMs offer for teaching engineering judgment in the classroom.

Regarding future lines of work, the integration and comparison of results among AI chatbots (for example, Claude, Gemini or DeepSeek) and the involvement of participants at different stages of training in engineering judgment would make it possible to explore the advantages and risks of involving AI in university curricula in the teaching of complex concepts in the field of engineering.

Finally, the adoption of tools such as ChatGPT, as proposed by[4],[19], constitutes an opportunity for the development of research in engineering education, involving both the exploration of curriculum modification and reflection on the ethical implications of its use in the classroom, in which AI chatbots, with their strengths and limitations, are understood as complements, and not as substitutes for traditional education.

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/

ACKNOWLEDGMENTS

The authors thank Martha Moreno Igua, professor in the Department of Humanities at Universidad El Bosque, for her valuable guidance in the application of discursive strategies in the writing process of this article. The authors thank Universidad Militar Nueva Granada (Colombia) for its willingness and support in the processes that facilitate research results such as the present one.

FUNDING

No financing

CONFLICT OF INTEREST

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

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: Óscar Hernán Vargas Villamizar, Juan Carlos Ruge Cárdenas, Carlos Felipe Urazán Bonells, Diego Palma Cuero.

Writing–review and editing: Óscar Hernán Vargas Villamizar, Juan Carlos Ruge Cárdenas, Carlos Felipe Urazán Bonells, Diego Palma Cuero.