CC1

Dear authors

I hope this message finds you well. I have had the opportunity to review your manuscript and would like to provide my feedback. The study presents an interesting and innovative approach with significant potential within the scope of the journal. However, several issues need to be addressed, where revisions could greatly improve the clarity, rigor, and overall quality of the work, and I recommend further revisions:

1. **Language and Editing**: To enhance the readability and accuracy of the manuscript, I recommend having it reviewed by a professional English editing service with expertise in geosciences or soil science. This will help ensure the language meets the standards expected for a publication of this nature.

2. **Clarity of the Research Gap**: While the manuscript addresses an important topic, the research gap could be more clearly articulated. A more concise explanation of the gap, along with a discussion of how this study contributes to soil science and its innovative aspects, would help strengthen the introduction.

3. **Objectives and Hypotheses**: The objectives need to be better defined and aligned with the research gap. In addition, it would be more appropriate to position the hypotheses after the objectives for better coherence.

4. **Methodology**: The methodology section would benefit from a more detailed and logical presentation, ensuring that it follows a clear, chronological sequence. Some aspects of the methodology are unclear, and it would be helpful to clarify the reasoning behind the chosen procedures to ensure the study is reproducible.

5. **Results and Discussion**: The results and discussion section could be revised to better adhere to scientific writing standards. The results should be presented clearly, followed by a more in-depth and up-to-date discussion. The current structure of the section could be improved to ensure a smoother flow and better integration of the findings.

6. **Limitations**: The limitations of the study could be explored in more depth. There are additional factors that were not addressed in the manuscript, such as mineralogical analysis, soil types and classes, and landscape dynamics, which could impact the findings, particularly with respect to the magnetic susceptibility of the soil.

7. **Conclusions**: The conclusions section could be more concise and focused on the main findings of the study. At present, the conclusions do not fully align with the research objectives or the results, and further clarification is needed to reflect the study's actual contributions.

I believe that with these revisions, the manuscript will be much stronger and more aligned with the expectations of the journal. I appreciate the opportunity to review this work and hope my feedback proves helpful. Please feel free to contact me if you need further clarification.

Response: thank you very much for reviewing this work. The comments and suggestions are much appreciated. Because specific comments are the most beneficial for pointing out particular aspects, the work is now benefited by them, as listed below.

CC1-1: Please, verify recent articles which used the Nested Leave One Out Cross Validation method for soil analyses containing small dataset and geophysical data:

https://doi.org/10.5194/gmd-15-1219-2022

https://doi.org/10.1016/j.soilad.2024.100024

Response: I checked out the sources and the described method.

CC1-2: This is a key aspect of your work. You should clearly elaborate on the knowledge gap that this study aims to address, highlighting its innovative elements and practical contributions. Merely emphasizing the lack of prior research is not a sufficient justification for conducting this study.

Response: the paragraph just next to this sentence already highlights the elements that are indeed required to motivate the study:

"The main hypothesis is that soil κ can support characterizing soil mineralogy, which also influences the permanent component of CEC. Therefore soil κ may significantly enhance the accuracy of CEC PTFs, which can help evaluating field CEC rapidly, and at low cost. To improve predictions of field CEC by integrating soil κ we develop and test uni- and multivariate polynomial PTFs based on data of diverse soil types sampled in Europe. In addition, we explore soil κ measured in-situ and in laboratory at different frequencies to give insights into the κ-CEC relationship and investigate how clay content affects the relationship between κ and CEC.**"**

CC1-3: only clay content? what about clay type (mineralogy) which strongly influence soil k due to presence or absence of ferrimagnetic minerals?

Response: we do not have such data for this study, so it was not possible to evaluate.

CC1-4: Please, exemplify what minerals

Response: mostly ultrafine magnetite and maghemite

CC1-5: This part of the methodology is unclear and requires additional details to ensure reproducibility. Please reorganize the section and provide further clarification, including specific steps, tools, and datasets used.

Response: All the features are mentioned explicitly and the number of data points, so the dataset used is well described, including previous information about how the features were obtained. I think that the steps are clear, and the criteria are explicit, so the reproducibility is ensured. Otherwise, could you particularly mention some aspect?

Additionally, all steps can be checkout and reproduced using the open-source code.

CC1-6: More detailed explanations of the statistical analyses are essential to ensure the reproducibility of this study. Certain parts of the methodology, particularly in the first paragraph, appear irrelevant or better suited for discussion. Clarify the specific analyses conducted, their purpose, the insights they provided, and the software used for the analysis. Additionally, explicitly state the database utilized, as relevant information regarding these aspects is currently missing in this section.

Response: the reproducibility is fully ensured because all the code used is open source. The purpose is discussed, this is, distinguishing between independent and masked effects. The insights cannot be presented in the methodology section.

Because the result of this calculation is not dependent on the used software, it is not mentioned. However, I specific mention to Python code is added at the end of the introduction: 'To ensure transparency and reproducibility, all the collected data and developed code for this work is publicly available in an open source Python software: Mendoza Veirana, 2024.'

CC1-7: please use recent reference

Response: references stand by the trustability of the information they provide, and this is not dependent on time.

RC1

General Summary

The manuscript investigates the relationship between soil magnetic susceptibility (κ) and cation exchange capacity (CEC) beyond the site level across various European soils to improve pedotransfer functions (PTFs) for CEC using near-surface electromagnetic geophysics. The authors considered several properties to develop univariable and multivariable regressions.

I appreciate the authors' valuable contribution to developing PTFs for CEC, considering the wide range of samples and properties. I believe this manuscript is suitable for inclusion in the special issue - Agrogeophysics: illuminating soil's hidden dimensions, as it provides important insights into the performance of geophysical methods in soil. However, I believe that some points can be improved. Therefore, I recommend a minor revision.

All the best!

General comments –

RC1-1: The introduction is detailed and covers key concepts effectively. However, the flow could be streamlined to ensure a clear connection between the hypothesis and the motivation for the study.
Response: thanks, this is an observation is right. Here is an improved version of the connection between hypothesis and motivation:

'To the best of our knowledge, the κ- CEC relationship has not been studied beyond the site level (Siqueira et al., 2010). This limited scope represents a significant research gap, as the broader applicability of κ for CEC prediction remain largely unexplored across diverse soil types and conditions.

The main hypothesis is that soil κ can support characterizing soil mineralogy, which also influences the permanent component of CEC. Therefore soil κ may significantly enhance the accuracy of CEC PTFs. This study directly addresses the identified gap by systematically examining the κ – CEC relationship using a new comprehensive dataset. The potential to develop more robust, widely applicable CEC PTFs underscores the significance of this work, with implications for sustainable land management, precision agriculture, and environmental monitoring.

To improve predictions of field CEC by integrating soil κ, this study focuses on develop and test uni- and multivariate polynomial PTFs based on data of diverse soil types sampled in Europe. In addition, we explore soil κ measured in-situ and in laboratory at different frequencies to give insights into the κ-CEC relationship and investigate how clay content affects the relationship between κ and CEC. While the methodology of this study focusses on soil and geophysical data collection, data analysis and model development, delving into the underlying physicochemical mechanisms of soil mineralogy that would link κ and CEC are out of our scope but is highlighted as an important direction for future research.**'**

RC1-2: The Results and Discussion section should be improved by incorporating in-depth discussions

Response: thanks for the suggestion. This section has been expanded along discussions.

RC1-3: Future improvements should be discussed with identified limitations

Response: this section was expanded by discussing how to bridge the current limitations.

Specific comments –

RC1-4: Lines 37-38 – "Defined as the ability of a soil to hold and exchange cations .." Please rewrite the sentence for clarity.

Response: such lines now read:

'CEC, which refers to a soil's capacity to retain and exchange positively charged ions (Khaledian et al., 2017), is highly correlated with the soil clay content due to a larger colloid surface for particle exchanges'

RC1-5: Line 70 – The statement highlighted the novelty of the study. Could you please explain the research gap further and include why the study is significant in addressing this research gap? It would be nice to highlight the importance of this study.

Response:

> 'To the best of our knowledge, the κ- CEC relationship has not been studied beyond the site level (Siqueira et al., 2010). This limited scope represents a significant research gap, as the broader applicability of κ for CEC prediction remain largely unexplored across diverse soil types and conditions.
>
> The main hypothesis is that soil κ can support characterizing soil mineralogy, which also influences the permanent component of CEC. Therefore soil κ may significantly enhance the accuracy of CEC PTFs. This study directly addresses the identified gap by systematically examining the κ – CEC relationship using a new comprehensive dataset. The potential to develop more robust, widely applicable CEC PTFs underscores the significance of this work, with implications for sustainable land management, precision agriculture, and environmental monitoring. '

RC1-6: Lines 74-77 – Please rephrase the objectives for clarity – the way objectives are presented in the manuscript is a bit complicated.

Response: a re-phrased and added scope is now in the text:

> 'To improve predictions of field CEC by integrating soil κ, this study focuses on develop and test uni- and multivariate polynomial PTFs based on data of diverse soil types sampled in

Europe. In addition, we explore soil κ measured in-situ and in laboratory at different frequencies to give insights into the κ-CEC relationship and investigate how clay content affects the relationship between κ and CEC. While the methodology of this study focusses on soil and geophysical data collection, data analysis and model development, delving into the underlying physicochemical mechanisms of soil mineralogy that would link κ and CEC are out of our scope but is highlighted as an important direction for future research.'

RC1-7: Line 79 – "(Mendoza Veirana, 2024)" – please remove the brackets

Response: 'To ensure transparency and reproducibility, all the collected data and developed code for this work is publicly available in an open source Python software: Mendoza Veirana, 2024.'

RC1-8: Lines 83 – Please provide how many sites and samples from each country.

Response: In line 85:

'Specifically, 6 sites in Belgium contributed 38 samples, one site in the Netherlands contributed 6 samples, and one site in Serbia contributed 5 samples. This distribution ensures representation of diverse soil types and textures across the three countries.'

RC1-9: Lines 92-93 – "Undisturbed soil samples (100 cm^3) were collected manually, by pushing standard steel rings horizontally into the soil profile wall at the same locations where $\kappa*$ was measured" – slight suggestion for rephrasing.

Response: thanks, changed as suggested.

RC1-10: Lines 93-94 – Volumetric or gravimetric water content? "After drying them for 24 hours at 105 °C " – this should be gravimetric water content.

Response: Thank you for pointing this out. In our study we worked with volumetric water content. We first determined the mass of water lost during oven-drying the 100 cm^3 cores, then divided that mass loss by the core volume to obtain θ. We will therefore revise the sentence to read:

"Undisturbed soil samples (100 cm^3) were collected manually, by pushing standard steel rings horizontally into the soil profile wall at the same locations where κ_* was measured. After the cores were weighed fresh and oven-dried for 24 h at 105 °C, volumetric water content (θ) was calculated from the water-mass loss divided by the core volume, and bulk density (b_d) from the oven-dry mass divided by the same volume (Grossman and Reinsch, 2002). "

RC1-11: Grossman and Reinsch, 2002 and Ciesielski et al., 1997a, 1997b – the fonts are different from the rest

Response: thanks, changed as suggested.

RC1-12: Lines 126-128 – Please move the first sentence to the introduction or discussion. This content is more appropriate under the introduction and discussion sections than the methodology.

Response: thanks, changed as suggested.

RC1-13: Equation 2 – "R^2=0.94" should be corrected as "R^2 = 0.94"

Response: thanks, changed as suggested.

RC1-14: Figure 5 – 0,0 and 0,40 overlapped in the CEC axis and s axis – please correct them. That is a very nice figure.

Response: thanks, changed as suggested.

RC1-15: Lines 204 – "(Glover, 2015; Wunderlich et al., 2013)" Please change the font.

Response: thanks, changed as suggested.

RC1-16: Results and discussion – This section should be improved with in-depth discussions of the results, especially in the 3.3 and 3.4 sections. Please expand the potential reasons for your results a little bit further by considering relevant literature.

Response: thanks. Relevant literature is scare since there are no studies analysing the link at a cross-site scale. The reason for the results were expanded:

'The strong performance of σ and κ_* as predictors of CEC in sandy soils (median test R^2 = 0.85) is particularly noteworthy. σ is known to be influenced by several factors including soil water content, salinity, and the concentration of dissolved ions, which collectively can reflect the variable component of CEC (Glover, 2015). In sandy soils, which typically have lower water and nutrient retention capacities, σ can provide a dynamic measure of the available exchangeable cations at a given time. Concurrently, the strong predictive capacity of κ_* suggests it captures a different, yet complementary, aspect of CEC. In soils with low clay content, and therefore limited colloid surface area, the permanent component of CEC is more significantly influenced by minerals. The fact that κ_*, measured in-situ, performed better than laboratory κ suggests that the undisturbed soil structure and field conditions are crucial for this relationship, possibly reflecting the spatial arrangement and contact of these minerals within the soil matrix. '

RC1-17: Line 214 – It would be nice if you could change the "4. Limitations" to "4. Limitations and future directions" and discuss future improvements of the proposed methodology.

Response: change as suggested. This section was expended and further improvements are suggested:

'The current study, while providing novel insights, has several limitations that also point towards important future research directions.

Firstly, the main limitation of the analyzed results are related to the dataset size, although diverse in terms of European soil types, is relatively small. A larger sample size could improve the statistical relevance of the findings and improve the robustness and generalizability of the developed PTFs. Future work should aim to expand the database with more samples covering an even wider range of soil properties and parent materials.

Secondly, all collected samples come from non-tropical regions, where organic matter content and bacterial activity do not significantly influence soil κ. In contrast, these factors may contribute substantially to higher soil CEC in other environments (Seybold et al., 2005). Therefore, the results are valid for the sampled sites that belong to European soils, and applications to scenarios beyond this range of soils should be approached with caution.

Thirdly, a significant limitation is the lack of direct mineralogical analysis, especially for clay and iron oxide fractions. While κ offers an indirect proxy for ferrimagnetic mineralogy, detailed characterization (e.g., via X-ray diffraction) is needed for a mechanistic understanding of the κ - CEC link. Identifying specific clay minerals (like kaolinite vs. smectite) and their abundance would clarify their CEC contributions and interactions with magnetic minerals. This is a crucial step to move beyond empirical correlations towards a process-based understanding

Fourthly, while field-measured κ proved useful, the reasons for its superiority over laboratory-measured κ_{lf} or κ_{fd} in the PTFs warrant further exploration. This could involve investigating the effects of soil structure, moisture content (which are preserved in in-situ κ measurements). A deeper understanding of how these factors influence different κ measurements could lead to optimized measurement strategies.

Finally, the model shown in **Error! Reference source not found.** is valid for samples with clay content between 2.9% to 16.1%, σ between 0.55 mS/m to 39 mS/m, κ_* between 8 to 320 μ, and CEC between 1.6 meq/100g to 8.7 meq/100g.As larger and more comprehensive datasets become available, exploring advanced modelling techniques, such as machine learning algorithms, may capture more complex, non-linear relationships. Assessing the scalability of the κ-CEC relationship from point measurements to field-scale predictions using proximal sensing platforms, for example, vehicle-mounted EMI sensors providing dense κ data, would be beneficial.'

RC2

This manuscript attempts to use pedotransfer functions (PTFs) to explore the relationship between soil magnetic susceptibility (κ) and cation exchange capacity (CEC).

The idea of investigating a potential relationship between these properties based on mineralogical connections is interesting. However, the manuscript requires further refinement in multiple aspects. Below are some suggestions for improvement:

General Comments

RC2-1: The manuscript should undergo a thorough formatting check to ensure consistency before submission, including font uniformity (e.g., L94, L204).

Response: thanks, this was corrected.

RC2-2: The discussion section is quite weak and needs substantial improvement. A more in-depth discussion should be provided, particularly on the mechanisms linking magnetic susceptibility to CEC.

Response: thanks, this section has been expanded.

RC2-3: Limitations: The authors should engage more with relevant literature, particularly on mineralogy and other geophysical approaches for predicting CEC. This will help contextualize the study's limitations and highlight its contributions more effectively.

Response: Thanks. A detailed mineralogical investigation and its causal link to κ and CEC, while important, is beyond the scope of this initial data analysis and modelling study, as stated in our objectives. Our findings provide a strong basis for future mineralogical research, a point now further emphasized in the revised "Limitations and Future Directions." Other geophysical methods are discussed in the introduction, and we have now added more relevant literature to the limitations section to better contextualize our work and its contributions.

Specific Comments

RC2-4: L55: What do you mean by "even though they generally correlate well with CEC"? How strong is this correlation? Please provide supporting data or references.

Response: such data is now next to their references

"Additionally, results have shown that σ and soil κ are independent (Maier et al., 2006), even though they generally correlate well with CEC.

Soil magnetic susceptibility has been correlated positively with CEC in studies focusing on soil type identification (Mello et al., 2020) (Pearson's correlation 0.4), soil characterization (Siqueira et al., 2010) (Pearson's correlation 0.68), paleoclimatic reconstruction (Maher, 1998) (Pearson's correlation

0.95 for Podsol and 0.73 for Cambisol samples),, and electromagnetic induction applications (McLachlan et al., 2022) (variable correlation).”

RC2-5: L82: What was your sampling strategy? Serbia is geographically distant from the other sites, and the total number of soil samples appears to be quite small.

Response: “This distribution ensures representation of diverse soil types and textures across the three countries.”

RC2-6: L83: Rather than only citing your previous paper, please include key details about the soils, such as soil types and the time of sample collection.

Response: soil types are mentioned in Table 1. No need to go to my previous paper, it is just there because part of the data was already published.

RC2-7: L97: ISO 11164 (which year?) specifies pretreatment procedures for soil samples. However, within my understanding, it does not apply to particle size analysis. What specific method was used for particle size determination?

Response: you are right. This was rephrased:

> “Clay, silt and sand content (denoted as Clay, Silt, Sand, respectively, expressed in %) was measured following the pipette method (NF X31-107, 2003) after sieving at 2 mm, content of humus, CEC was determined by CoHex method (Ciesielski et al., 1997a, 1997b). “

RC2-8: L98: How was CEC measured? What does "CoHex" refer to—is it a commercial product or a specific method? Please clarify.

Response: method, this is now specified.

RC2-9: Table 1: The table is not well-structured. Consider referencing other studies and including key statistical indicators such as minimum, mean, median, and maximum values.

Response: minimum and maximum are already there, these are the first and second number of each interval

RC2-10: L135: Why did you choose the median value to split calibration and test datasets?

Response: because it is the most simple way to discriminate regarding Clay content. Also, this is data-dependant.

RC2-11: L140: Is there any reference supporting the use of the median R^2 test in similar studies?

Response: No, but this is a common practice in data science modelling. Specifically in our case, I highlighted such thing at the very beginning of the section:

"The absence of previous attempts at developing a CEC PTF using soil κ data that can be generalized beyond site-specific highlights the importance of thorough data exploration"

I added a reference about this statistical learning procedure:

"The best polynomial degree (linear or quadratic) was determined by the highest median of the R^2 test scores over the 100 repetitions (Tibshirani et al., 2001). Finally, model implementation was performed after tuning and feature selection using all the samples of each subset."

RC2-12: Figure 2: Why are some correlations missing, such as between CEC and depth?

Response:

"Figure 1 Spearman rank correlation heatmap showing significant P-values≤0.005 for the 49 soil samples, missing correlations have P-values>0.005."

RC2-13: Clay and humus pairing: The choice of this combination is unclear. Did you simply sum the humus and clay content? If so, what is the rationale behind this approach?

Response: I did not sum both, but both are used as predicting variables, this is, calculating CEC based on humus and clay:

F(Clay, humus)= CEC

This is also mentioned in Methods (line 145):
'The top four combinations in terms of test performance were compared to the standard combination of *Clay* and *Humus* content, also, single features were considered ($Clay$, σ, and κ_*). '

Also in the introduction (line 50):

'Commonly, CEC PTFs are expressed in function of clay content and humus, and less frequently pH and soil depth (Khaledian et al., 2017; Seybold et al., 2005).'

RC2-14: Figure 5: This figure is unclear, making it difficult to interpret the data points. Consider presenting it in a clearer format. Additionally, you should include a predicted vs. measured CEC comparison for the test sites.

Response: Reviewer 1 comment 12 states: 'That is a very nice figure.'

I think that the clarity of the figure is relative. From my side, it is clear because it shows three sides of the cube, providing perspective in the relationship between σ, κ , and CEC.

On the other hand, there are no test sites as such because based on the strategy to train the model, the test samples are chosen regardless of the sites to avoid biases. Also, because there are between 2 and 7 samples per site, which makes difficult to compare sites performances. This is the reason why are all the 25 samples of the sandy group compared together in Figure 5.