Review of:

NorSand4AI: A Comprehensive Triaxial Test Simulation Database for NorSand Constitutive Model Materials

…submitted for publication in Geoscientific Model Development (an open-access journal of the European Geosciences Union).

General Comments

The ms offers an approach to using AI for assessing soil behaviour, with a focus on static liquefaction. This is novel. But, the nature of this novel contribution is unclear in both the Abstract and the Introduction – the key ideas do not occur until L62-68, with further relevant comments in L128-133. Further, the Abstract is too long with more detail than appropriate while not highlighting the main ideas and contribution.

More generally, the ms focuses on details of the process rather than results. For example, there is no figure comparing the AI output with the NS model used for example soil properties or, better, several examples). This leaves an interested reader uncertain on how well the AI offered works, with Figures 2 and 3 not helping if you are not already familiar with details of AI.

A related concern is the range of soil properties over which the AI has been trained. Table 1 provides the ranges used, and most seem reasonable (or at least in accord with what might be expected for 'sands'). There are issues with these ranges as it appears each soil property has been treated as independent whereas H_0 is commonly inversely correlated with λ and also directly correlated with G – if the soil is stiff elastically, then it will be stiff plastically. Thus, there appear to be property combinations that are not physically representative. Further, in the case of OCR the lower limit is unity, not 0.5: all critical state models do not allow under-consolidation.

These concerns lead me to the conclusion that the paper needs re-writing to properly present the ideas, likely moving some detail to an appendix, with an expanded presentation to illustrate the performance of AI as it might be utilized by a practicing engineer. Additional comments follow to assist this reworking on the ms.

Detailed Comments

L63. All critical state models have particular idealizations, the most obvious of which is that all of them (at least to date) are based on 'remoulded' (or disturbed) soil properties. You might argue that 'structure' might appear as apparent OCR (a common view with clays, see Bjerrum's Rankine Lecture) but then where does that leave ψ ? Should there be a cohesion component to the strength ? Could M vary with strain ? These points are not a criticism of your approach as such, but rather aspects that need identifying to the reader. Indeed, and this is something you might discuss in a revised ms, it seems to me your AI could be quickly used to assess a new set of data to evaluate if indeed ‘structure’ might be something that needed considering. If I understand correctly, you allude to this in L128-133.

L83. An open-source Python implementation is an excellent idea/contribution, but it really is a different subject than AI. Present the Python work as a second paper, not necessarily to this journal.

L161. NS is closer to original cam clay (OCC; Schofield & Wroth) than MCC, with NS and OCC yield surfaces having the same shape and the same flowrule.

L177/Table 1. The sampling ranges used really should be presented as a distinct new section, as it is a new topic. Indeed, you might even move it forward to the discussion of training NN as these properties are not intrinsic to NS (with the exception of H, which you could slave to λ, say using $H_0 = 2/\lambda$). I also wonder if there should not be a figure, say plotting χ vs λ with the points using a different symbol for M – not comprehensive, but it would illustrate the space occupied by your realized ‘training’ cases; such a figure might usefully follow L195-208.

L231. I do not understand what comprises your dataset developed using NorSandTXL. Is it a set of strains and stresses, and if so at what strain increments ? Are you capturing what amounts to a numerical triaxial tests with 100 steps ? 300 steps ? I am guessing that a ‘dataset’ amounts to an array [n,4]; this needs clarifying.

L288. This needs a figure illustrating the q-e1 with 40 vs 4000 points. This is a quite extreme compression to my eyes, as 40 points always seems too few when recording a lab test. And please indicate whether you run out to 10%, 15% or what strain, keeping in mind the critical state is a large-strain condition. I am also surprised that you regress on e – I would have use the dilatancy; and, it may be appropriate to treat drained and undrained tests differently.

Section 6, Validation. I found this section unconvincing. The aim is to recover best-estimate soil properties from a [40,3] reduced dataset. There are 10 properties and 4 state measures (Table 3). Actually, thinking as I type you could remove ν as it is rarely measured and commonly just assumed as a “not unreasonable” value. What is needed are plots showing ‘truth’ (= known property of training dataset) on x-axis vs ‘prediction’ (= recovered using AI) on y axis; I would focus on just a few properties/state to keep the ms to a reasonable size, say: λ, χ, G, ψ. Such plots will allow a reader to quickly assimilate how well the procedures presented in the paper work.