General comments

The authors aim to improve an existing nutrient availability metric for forest ecosystems. They present their results well-structured and clearly. A nutrient availability metric would definitely be of great importance, however, the paper has two major drawbacks:

1. In the introduction the authors write that they aim at developing a globally valid metric while using data from Sweden only. It should be clarified that the upgraded metric is only valid for Sweden. Possibly, a global metric is not achievable at all, since nutrient availability is not limited by the same factors in different ecosystems worldwide. Hence, a metric for Sweden might be usable for other boreal, but not for tropical ecosystems. These considerations should be discussed.
2. The performance of the metric is bad. Instead of discussing the – non-existing – relationships, the authors should rather discuss the possible reasons for the failure of the metric. One possible reason is data quality. The authors should describe the soil sampling design and the methods used for chemical analyses. Inventory data might not be suitable to find relationships between parameters even though they exist, because (soil) variability would require a large number of replications, which is often not affordable in inventories.

Specific comments

l. 17 The coefficients of determination are that poor that you should not write "Normalized productivity increased with decreasing soil C:N ratio, while SOC exhibited an empirical optimum."

l. 21 The coefficient of determination of the upgraded metric is still poor and should not be called "a significant fraction".

l. 34 and l. 36 "among terrestrial ecosystems" and "global scale" is misleading. You should clarify that in the present paper only Sweden is considered.

l. 47 "are more indicative" to what and for what?

l. 48 What is meant with "the size of the soil solution"?

l. 57 to 60 Yes! Perhaps the fact that a global metric might not be achievable at all, since nutrient availability is not limited by the same factors in different ecosystems worldwide, should be discussed here, too.

l. 78 to 84 If the goal is a "global metric", data from the Swedish forest inventory service do not represent "a substantial variation in nutrient availability". Which were the additional variables? P?

l. 89 Here you restrict your results on Swedish forests – you should already earlier mention that your goal is a metric for Sweden – not a global metric.

l. 94 to 134 List all the parameters (soil, tree/productivity, climate/meteorology) and explain how they were measured.

l. 121 to 130 Explain in more detail how the scores were derived and what you can find in the look-up tables.

l. 137 to 138 You should state that you call the two alternative ways to calculate normalized productivity "method 1" and "method 2" in the following.

l. 138 Name the advantages and drawbacks.

l. 179 Is "method 1" referring to the method used to calculated the normalized productivity?

l. 193 to 200 You used half of the dataset from southern Sweden to upgrade the metric and tested the upgraded metric with the rest of the dataset including middle and northern Sweden, right? Did you also upgrade the metric using method 2 (half of the dataset for calibration, half for validation)?

l. 216 to 217 Why not using both SOC and TN in separate models and choosing the model that better fits the data?

l. 225 "… related to normalized productivity (Table 2), however, the coefficient of determination was small (0,002 to 0,146)."

l. 227 I cannot see from Fig. 5 that "the effect became more pronounced towards the south".

l. 232 to 234 According to Table 3 also other variables than SOC, pH and C:N were included in the multiple regression models. Why?

l. 239 to 240 On the one hand you write that SOC and C:N "consistently describe a distinct, clear effect on normalized productivity" on the other hand you write that R^2 is "at least a few percent" – this is contradictive. Both the figures and the R^2 show that there is no distinct, clear effect on normalized productivity.

l. 225 to 241 This part should be rewritten since the variance in productivity is hardly explained by the soil variables (see also discussion l. 272 to 313).

l. 251 to 252 The relationship between residuals and variables is very weak.

l. 254 To my opinion, the results do not show "that SOC, C:N and pH are important factors influencing nutrient availability in Sweden."

l. 264 to 265 From Figure 8 no enhancement of the upgraded metric can be deduced. The performance is still bad.

l. 272 to 313 The variance in productivity is not explained well by the soil variables. Instead of discussing the (significant but very weak) relationships, you should rather discuss why you do not find relationships were you would expect them. What about the soil sampling design and the methods used for chemical analyses? You did not describe them in the material and methods section. Could the sampling design (number of replications) explain the bad model performance? How is the variability of soil variables in Sweden?

l. 323 … but again with a very small coefficient of determination.

l. 330 "… the nutrient availability metric was intended to be improved by …"

l. 335 I cannot agree. The relationships were significant but only a very small part of the variance was explained by the variables used.

l. 336 to 339 Why are especially stable N isotope signatures and ion exchange resin bags of interest? What about other methods?

Table 2 Are you sure that R^2 is the same for all three regions for the parameters SOC, N stock, sand, clay and pH_{KCl}? This is quite unrealistic and the data shown in Figure 5 (SOC) lead to the assumption that R^2 of the North is largest and that of the South smallest. In line 226 to 227 you write that the effect becomes more pronounced towards the south (C:N), however, in Figure 5 the relationship is worse for the south than for the other regions and a larger R^2 – as written in Table 2 – seems to be quite unrealistic.

Table 4 and 5 The coefficients of determination are similar and do not point on a better implementation of the parameters in the upgraded metric.

Figure 1 I have problems understanding the legend of Figure 1.

Technical corrections

l. 12 …to test which combination <u>of</u> soil factors…

l. 63 to 64 rephrase this sentence to avoid the twofold use of "recent(ly)"

l. 96 explain the abbreviations

l. 118 ~~mass~~ <u>stock</u> [kg <u>m^{-2}</u>]; if really the mass is meant, the formula is wrong

l. 157 "we therefore we split" delete one "we"

l. 164 "SOC" – stock or content? "Total N" – total N content? "N stock" – total N stock?

l. 167 Name the software used.