

Interactive comment on “Qualitative soil moisture assessment in semi-arid Africa: the role of experience and training on inter-rater reliability” by M. Rinderer et al.

Anonymous Referee #2

Received and published: 21 April 2015

General Comments:

This manuscript describes the testing of a soil “wetness” classification scheme with three groups of individuals (experts, students, and farmers) in Tanzania. The methods and results presented here have importance for both “experts” and farmers that seek to determine optimal conditions for seeding and for maintaining crop vigor. The paper was well-written and straight-forward. One point in particular that I feel warrants further discussion is how transferable the wetness scheme and training is between sites/soils. For example, the classification scheme refers to conditions in which bricks could be made, which appears to be a highly localized bit of knowledge. Overall, it seems that

C1189

this process involved a fair degree of “calibration” to local practices and soil conditions, so how dependent is the process on the experts coming into a given locale? Similarly, could these methods be developed into fact sheets or training guides that could be used by non-expert individuals in multiple locations (in theory at least)?

I wonder why there was no testing (or at least discussion) of whether these practices resulted in any kind of improvement in farming practices and/or yields. That seems to be the ultimate goal of the effort, yet as far as I can tell was not mentioned.

I am a bit confused by what constitutes the “control” in these experiments. Is there a single person who decides the “true” wetness classification of the sampling points, or is that determination reliant on the median response from the group? Based on the limited differences between volumetric water contents in the dry (1 & 2) and wet (5-7) classes, this seems like important information when judging the accuracy of responses.

Finally, the authors argue against reducing the number of wetness categories. However, Figure 3 in particular leads me to question the value of having so many different classes, when there is little or no quantitative moisture differences between many of them. At the same time, even after training many of the farmers appear to have misclassified samples by multiple categories (Figure 8), which seems to suggest that there may be little benefit in having seven different classes.

Specific Comments:

P3035, L3: How realistic is the assumption that vertical soil moisture is close to equilibrium? This assumption seems suspect to me, but could possibly be verified by repeated tests through time or depth. A citation here would help.

P3038, L5: It is not clear why maximum attainable CK value (CKmax) would be less than 1, since the previous statement states that perfect agreement would be $CK = 1$. What is the value of CK/CK_{max} if CK_{max} is not a constant?

P3041, L3-13: How were the samples divided into “halves”? The meaning of this entire

C1190

paragraph was not clear to me.

P3042, L8: Practically speaking, is there any difference in outcome if a farmer rates a soil as too wet versus too dry?

P3043, L5: I wonder why this information about the misclassification of 6 classes due to ticking error was not included in the results section, since on P3041, L1 there is mention of confusing assessment form for the April test.

P3044, L6-8: Since the median volumetric water contents were practically identical between the two driest classes and three wettest classes, how does one resolve small scale changes in soil moisture using them? Or are the classification schemes capturing differences in the soil matric potential, even if they do not appear to be significant changes in soil moisture?

P3045, L7-8: In my opinion, the line "The study also shows that the qualitative wetness classes are reflecting the quantitative differences in volumetric water content" is debatable, since the results and discussion both reflect that many of the wetness classes did not have statistically significant differences between them. For example, Figure 3 indicates that Class 6 may have had slightly higher VWC values compared to Class 7.

Table 1: The statement "water liquefies" does not make sense to me. Water is generally liquefied.

Figure 4: It seems like this could be put into a single chart, since there appears to be mostly repeated data between a) and b). At the same time, it seems like this information is wholly contained within Figure 6 (though certainly in a different form). Is the figure necessary?

Figures 6, 7, and 8: It took me a while to interpret these figures. Maybe the caption could be improved to better convey how to interpret the information (possibly by using an example sampling point and an example test person).

C1191

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 3029, 2015.

C1192