SCIENCEDOMAIN *international*
www.sciencedomain.org

SDI FINAL EVALUATION FORM 1.1

PART 1:

Journal Name:	International Journal of Plant & Soil Science
Manuscript Number:	**2013_IJPSS_8827**
Title of the Manuscript:	**Seedling Emergence and Seed Germination of Shepherd's needle (Scandix pecten-veneris) as Affected by Seed Weight and Burial Depth – The title was changed!**

PART 2:

FINAL EVALUATOR'S comments on revised paper (if any)	**Authors' response to final evaluator's comments**
The revised paper shows a high number of important improvements over the first submission but it still raises a number of questions some of them resulting from the new analyses done. Before going into them I would like to come back to the use of Turkey's HSD, the use of MET, the use of soil temperature data, and Bond et al paper. The authors abandoned Turkey's HSD but say that it is 'the common method for many authors'. Unfortunately they are correct. Statistically wrong analyses are really widespread and few journals enforce strict approaches like Weed Research or Annals of Applied Biology do. The first is sceptical about using STP's the second forbids them without appeal. It might be true that Seed Science & Technology and many other scientific journals accept MET. Again, unfortunately the authors are correct but this does not make MET (or a number of other indices) less misleading or more meaningful as the example I gave before shows. However you insist in using it. So please tell me what exactly does it represent, germination or emergency-wise, a MET of 31.8±2.1 days (or any other value of table 5)? What do you infer from MET data 2008-2009, 15 cm (table 5)? MET for 25 Nov. is two times higher than 15 Mar.? Exactly what is the meaning of this difference? Anyway, I fear that authors will stick by MET. After all a lot of people have used it before (this reminds me the story of water photolysis in photosynthesis)… In that case, please provide the reader with the range of possible MET values for each experiment so that the reader can assess your data against some reference interval! The authors measured soil temperatures at 5 cm depth and state that they found differences between the years. Ok! So what might be the relationship between differences at 5 cm depth and seedling emergence from 15 cm? As for Bond et al., how did they control upward and downward movements? Well, there you have a point. I don't know! But I'm not refereeing their paper. I'm refereeing yours, so my question remains. As for the sand part. Bond et al. used sand all right. Their equation points that maximum depth for successful emergence of 25 mg seeds would be 8.0 cm. You will agree that in soil this depth would probably be lower (at least not higher) but you report emergences from 15 cm depth. I still think that it would be nice to acknowledge this discrepancy and discuss it. Before going into new details of the revised version I would like to look at two aspects that I missed before. Now in line 57 (before in line 56) you present texture data for 0-30 cm depth. Given that you only experimentally went down to 15 cm what is the relevance of texture for soil depth twice as high? Now in lines 295-298 (before in lines 238-142) the authors say that during field experiment installation (what year? what date?) soil was very close to saturation and soil used to cover seeds was not compressed properly (pity that things were not done properly, especially in soil compression when soil depths are involved). Then the authors say: 'This may have altered soil's layer characteristics and, perhaps, the results that	

were observed'.
Which results exactly?

Now for new stuff.

1) The link in reference 10 was unusable and it still is!
By the way, reference 10 now appears in the text before reference 9.

2) You say that your previous work showed that seeds had to be left on the mother plant about 2 months (…); this is interesting information. Just write it down explicitly!

3) The authors present their work in view of the control of *Scandix pecten-veneris* which is a weed and in their answers they justify sowing in 15 March because Scandix is also sown in Crete from October to April. However interesting this may be it is meaningless unless the authors present this information in their paper and somehow put it in the context of their research, which they did not do. Otherwise the question I made previously remains, because as I can see it, the omission is somehow fishy!

4) Formerly I argued for the importance of year of harvest.
The authors are against the argument based on what they call grand means.
Well, in their view grand means result from pooling together data from different samples which is one of the oldest tricks in the book to obscure data.
Below I will come to pooling again. By now I would say that in highly structured designs (like those in the paper) looking at the grand means in Table 3 and saying that they are similar is bizarre!
First because 15 Mar. 2010-2011 (25% of data!) is clearly different. Second because its obscures what might be interesting to discuss: in 2008-2009 emergence in 25 Nov. was higher than in 15 Mar. at shallow depths (2.5 and 5 cm), the reverse at deeper depths, while in 2010-2011 the pattern was different. Why?

5) In this revised paper the authors regress seedling emergence and MET to soil depth by pooling together data from 4 different samples.
Pooling data is a tricky business and its adequacy is a matter of debate and discussion, but pooling data from samples that obviously come from different statistical populations is not debatable, is completely wrong.
Data from 15 Mar. 2010-2011 cannot be pooled with the other three periods of time (see Table 3) and the regression done is not acceptable. A number of alternatives can be envisaged. Multiple regressions with dummy variables with repetitions are one of them (in that case please don't forget to test for lack of fit!).
Also, what about the P-values of the coefficients, which are much more important that the P-value from the overall ANOVA, especially when total df are small (df=5).
Look at them and you'll find as I did that your choice of quadratic in MET (Fig. 5) is obviously wrong.
In addition, if comparisons among models are to be made, please always provide adjusted coefficients of determination. Also be consistent with decimals. Look at R^2 in table 4 and 5 (or F-values…). Also what were the criteria to opt for quadratic instead of linear? Also, rewrite lines 180-182 in order to be understood (what is the meaning, if any, of minimizing parameters for the linear trend?)

6) The experimental design of experiment 2.2 apparently was RCB, separately for each year, one treatment per year with two levels (small-light and big-heavy seeds) and four replications (the plots).
Apparently each plot 0.5 m × 0.5 m provided one value of emergence and MET. So, for each year there were 4 values of emergence or MET for small seeds and 4 values of emergence or MET for big seeds.
Total n=8.
How can such design be analysed as a two-way ANOVA?
Not nested ANOVA I guess, because you tested treatments and replications MS against Error MS.
Are you implying that your design is a two-way ANOVA without repetitions? Why?
It hardly makes sense! After all, your so-called replications in table 1 would be a random variable and then your two-way ANOVA would be a mixed model without repetitions.
I simply cannot see the difference from the experimental design analysed in table 2.
Please explain.

SCIENCEDOMAIN *international*
www.sciencedomain.org

SDI FINAL EVALUATION FORM 1.1

I feel from the beginning that this work is worth being published. This is the reason why, as a referee, I'm seriously investing in it. But please, seek for help especially in statistical matters! (It doesn't make much sense to choose RCB over split-blocks or split-plots because of the kind of experiment and studied variables; any basic statistical book on experimental design especially those written from an agronomic point of view would make clear). All the best.	

Note: Anonymous Reviewer