SOIL Discuss.,
https://doi.org/10.5194/soil-2018-28-RC1, 2018

Interactive comment on "Using deep learning for Digital Soil Mapping" *by* José Padarian et al.

Anonymous Referee #1

Received and published: 30 September 2018

Interesting paper using deep learning (in particular convolutional neural networks) for digital soil mapping. The paper is well written, concise and to the point. The methods are easy to understand, but (disclaimer) I am familiar with CNNs.

I agree with the authors that CNNs are a possibility to incorporate landscape information in the DSM models.

Section 4.1. Data: how many data points? How many data points for each depth interval? I know it is all described in the Padarian et al 2017, but it is quite important information to evaluate data augmentation effects. It is also mentioned in section 4.5, but without depth information.

Section 4.2. Data augmentation: I think the comparison should show the results of cubist with data augmentation. Data augmentation is something often associate with neural networks modelling, but it could be used for any models. I do not think it is a

level comparison if one model has X sampling points and the other has X*4 sampling points. Figure 4 shows the reduction of RMSE with data augmentation. However in figure 5 the results without data augmentation are not compared with cubist. I also think that the authors should at least discuss (if not test) the potential effects of data augmentation on the spatial auto-correlation of the data. Unfortunately, I do not think that the results presented by the authors are enough to describe the effect of data augmentation on the final outcomes.

Section 4.5 Training and validation: I am a bit confused about how the various datasets are named (see comments for figure 5 as well). My understanding is (further comments are based on this): Test : 10% of data never used in the fitting of the model. Remaining data (90%): bootstrap splitting in training and validation (1/3)

Results Figure 4: The largest reduction in RMSE happened for the deepest depth interval. I assume that this interval was also the one with fewest data points (as it is usually in soil datasets). Is this an effect of data augmentation? I.e. stronger effect when there is the lowest number of points?

Figure 5: How is it possible that the test dataset (i.e. the one that was never used in the model fitting) has a lower error than train itself or even validation? (see above comment for my interpretation of these labels) If this is not a label problem, I think the authors should really explain and discuss this, because, to me, it seems a problem of overfitting. I would also like to see a comparison with cubist model with data augmentation.

Section 5.3 Wavelet decomposition method. Maybe it should be briefly mentioned in the methods sections? Maybe a short summary of cubist implementation (i.e. with covariates and main parameters) should also be mentioned in the methods. However as cubist is a more common approach, I think this is less necessary than for the wavelet approach.

Section 5.5. Maybe comparison of some spatial structure metrics?

Section 5.6 Uncertainty. The maps show only a small area and the text is a bit confusing. For example: “greater reductions in higher (larger?) areas of the landscape” is not very clear or precise. Maybe it would be better to show a measure of the proportion of the landscape that has a great reduction of intervals width?

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2018-28, 2018.