

Interactive comment on “Multi-source data integration for soil mapping using deep learning” by Alexandre M. J.-C. Wadoux et al.

Behrens

thorsten.behrens@uni-tuebingen.de

Received and published: 23 January 2019

Dear authors,

I like the idea of the paper because the integration of multi-source data and modeling with uncertain measurements is as important as the extension of the toolbox for contextual modeling. In addition, CNNs are a promising new technique, at least for some machine learning applications.

General remarks:

The paper is not easy to read (at least for a non-native speaker) and requires some restructuring and revision.

C1

The title starts with "Multi-source data integration". CNNs are explained in detail, but the integration of multi-source data is not adequately explained and discussed.

The explanation of the CNNs lacks some detail.

Why did you only use three covariables?

What is the difference between a window and a filter?

Specific comments

Most of the specific comments are hints for more detailed explanations or rephrasing.

Abstract

The abstract is very condensed.

At the end you mention "different window size of input covariates matrix". This is not clear and needs an explanation.

1 Introduction

"Digital Soil Mapping (DSM) techniques are now commonly used to predict a soil property".

What was the purpose of DSM techniques before they were used to predict soil properties? ;-)

"...pre-processing, subjective decisions based on the resolution to which covariates must be treated as input to the model or modeller's choice regarding neighbouring size."

Can you explain this a little? The only point I really see is the size of the neighborhood. Some approaches require less preprocessing than traditional terrain analysis. And you can simply set neighbourhood size as large as possible. For most approaches the resolution is the same for all input data. So I don't see any real "drawbacks" here. It's probably the same as setting up a CNN.

C2

In your paper you use three covariates. I assume that processing large spatial data sets with a CNN is a challenge. However, some of the other methods quoted here have no problem handling hundreds of covariables/scales. So not sure if they really have "drawbacks".

"The CNN proposed here has the advantage that it relies on the local representation of covariates so as to leverage the spatial information contained in the vicinity of a sampled point."

This is what the other contextual approaches do.

"(i) develop the framework of Convolutional Neural Networks for contextual spatial modelling"

What is the difference to approach presented by Padarian et al. (2018) especially when you write that "Padarian et al. (2018) have shown that it is possible to use CNN for soil mapping while accounting for contextual covariate information"?

2.1 Model definition

"Note that unlike classical geostatistics, measurements of the soil property are assumed independent and identically distributed."

Is this different with CNN?

Please explain the ReLU activation function and why you chose it.

2.2 CNN

"In this paper we use the vicinity information of the measured soil property."

At this point, it's not clear how you're doing it. I propose to revise this paragraph

"In convolutional neural network, at least one layer is a convolution (Goodfellow et al., 2016)."

What is a convolution. Again, I suggest to revise this paragraph. Is it "network" or C3

"networks"? Can a layer be a convolution?

"We apply a 2D convolution using the filter F "

What kind of filter is it?

"...modification to include the case where we have $c = 3$ environmental covariates"

There is no equation showing how this is done.

"Filters detect features"

How do they detect them? And what are these features.

"Then, the last convolution returns an image of size 1×1 and with a number of channels. this is a vector that we can pass to a fully connected layer."

Change "this" to "This".

So the final layer just has one neuron? Or three? This is not clear.

I suggest to revise this section to better explain it to the soil science community.

Flatten and dropout layers are not explained but mentioned in Table 1.

Why was that specific network structure chosen? Subjective decisions?

Figure 1: I suggest to split it in two separate figures. Does the convolution reduce the image size or just the pooling?

2.3 Parameter estimation

Is this "parameter estimation" or "learning (with backpropagation)"

2.4 Multi-source data integration

The term "multi-source data integration" is in the title and is the interesting and important part of this article. However, this section does not show how it really works. How is the function updated? An example would be good.

2.5 Quality of predictions

I'm not sure if the formulas for RMSE are R^2 are required.

3.2.1

Why no cross-validation? Single subsets are generally not recommended for validation.

Why only 1000 trees? I suggest setting it to at least 2000 trees and then optimizing mtry and nodesize. However, this is not crucial as it will not significantly increase the prediction accuracy of RF in this study. The difference is the lack of spatial context.

3.2.2

Why 90:10 in this case?

What "window size"? Filter size? In Table 1 I find 3x3.

What are the "hyperparameters of the model architecture"? Ok, it follows. I suggest to restructure.

Why is the "window size" optimized separately? What were the initial settings of the hyperparameters when optimizing the window size?

Figure 6

The maps look noisy and probably show artificial horizontal and vertical stripes that have nothing to do with the input data and look like artifacts from the convolution. But without examining the data, this is only a guess.

5 Discussion

"The proposed modelling approach explicitly accounts for the TC measurement error in the model calibration"

Can you produce uncertainty maps reflecting influence of the measurement error?

What is the effect on the final maps as well as prediction accuracy when setting all

C5

weights to 1?

"To the best of our knowledge, our study is the first to account for measurement error for mapping using machine learning."

I suggest to put more emphasis on this.

"The window size of the input images had a significant impact on model's accuracy measure, as tested on an independent test set. This is because the size of the input image is closely related to the amount of contextual information we supply to our model."

Can you explain the relationship between the window size, the size of the input image, and the amount of contextual information? A figure would be good.

"...but larger window size worsened the prediction accuracy. In a similar context, this confirms the results found by Behrens et al. (2010)."

This is not what is shown in the ConMap paper quoted by you, but the multi-scale one in which we have tested comparable window sizes. Multi-scale digital terrain analysis and feature election for digital soil mapping Geoderma 155 (3-4), 175-185).

Interactive comment on SOIL Discuss., <https://doi.org/10.5194/soil-2018-39>, 2018.