Answers to reviewer comments and list of changes to Geoscientific Model Development Discussion paper 6, 7077–7116, 2013

M. Nussbaum, A. Papritz, A. Baltensweiler, L. Walthert

April 15, 2014

1 Answers to comments by anonymous referee #1 (RC C2487) of 31 January 2014

Thank you for your positive feedback. We comment on your suggestions for changing our article in the subsequent text. Your main criticism relates to the fact that we used only one statistical approach in our analysis and did not include any methodological comparison of results obtained by our and other statistical approaches (regression kriging [RK] and machine learning method [ML] methods, in particular).

We agree that such a comparison would be of some interest as there are only few published studies comparing ML with geostatistical methods to date. However, our main intention when writing the paper was to document our new robust external-drift kriging (EDK) approach in detail to make it fully transparent to the reader how we calculated mean soil organic carbon (SOC) stocks. Hence, we see no easy way to shorten the text to compensate for the inclusion of new material. Adding additional material would therefore likely extend the paper beyond acceptable length.

When selecting an approach for the analysis of our data we abstained from using RK and ML methods for the following reasons:

RK is an ad hoc approach and a (theoretical) sub-optimal variant of EDK. Regression coefficients are estimated by ordinary least squares (OLS) that does not — as opposed to generalized least squares (GLS), which is used in EDK — take into account the correlation structure in the residuals when estimating the regression coefficients.

The differences of spatial predictions by RK and EDK are expected to be small (e.g. Minasny and McBratney, 2007). However, the quantification of prediction accuracy — a main focus of our paper — remains likely too optimistic with RK. Two sources of uncertainty contribute to the standard errors of plug-in EDK predictions (i.e. when the uncertainty of the variogram parameters is ignored): (1) The uncertainty about the estimated regression coefficients and (2) the uncertainty arising from spatial interpolation of the regression residuals (simple kriging variance). RK accounts only for the second component (often

in an inconsistent way as mostly ordinary [OK] instead of simple kriging is used). The RK prediction standard errors are therefore very likely too small.

One could combine the OK variances with the covariance matrix of the regression coefficients as returned by the OLS fit in the same way as EDK does. However uncertainties of regression coefficients are notoriously underestimated by OLS in case of spatially correlated residuals. Furthermore, why should one want to compensate in an ad hoc manner for a flaw of RK when EDK provides a consistent solution to the problem?

ML methods do not — unlike kriging — provide consistent approaches to handle change-of-support. Therefore, the quantification of the uncertainty of regional and national SOC stock estimates with ML remains difficult. To report standard errors of such estimates was a main objective of our work. Their calculation requires the covariance matrix of the point prediction errors (see P7092 L1-20, section 2.3.4). So far, it remains unclear how this covariance matrix could be computed for Random Forest or boosted classification and regression trees (BRT). Martin *et al.* (2011) who used BRT for mapping SOC stocks across France did not provide a correct answer to this problem.

The specific comments are very welcome as the suggested changes clearly improve the clarity of the text. We accepted most of the suggestions, and we comment below only on those points where we (partly) disagree and did not change the text as proposed.

P7085 L8 Rename the title as Soil bulk density To our understanding "bulk density" refers to the density of whole soil including rock fragments (particles with diameter > 2 mm). As mentioned on P7085 L9 we considered the density of soil fraction with particle size ≤ 2 mm (corresponding to the term "Feinerde" in the Swiss and German soil taxonomy). The use of the term "bulk density" would therefore be misleading.

P7086 L9–16 Change of formula for calculating SOC stocks We use the formula to explicitly state how the correction of the soil density for rock fragment content was done. In your suggestion this is mentioned by using "ρ_0 the soil bulk density corrected for rock fragments" and does not include details. However, we changed the wording of the paragraph and avoid now the term "volume of horizon per unit area" and use "thickness of horizon" (denoted by symbol D_i) instead.

P7095 L11–24 Shift sub-section to Material and methods Here we disagree. This subsection describes the structure of the fitted models, and this is clearly a result of our work.

Table 1 Covariates were selected with the aim to optimize the predictive power of the models. The reason for different terrain covariates being chosen for top- (0–30 cm) or bulk soil (0–100 cm) is simply because they were better predictors in either model. In general, we strove in model building for parsimonious models containing as few covariates as possible, and we took care to not include any

covariate where the sign of its coefficient would contradict the dependency expected by a soil scientist. Thus, we do not believe that we can causally interpret the results of the regression analysis — this is notoriously difficult for models fitted to observational data — and see therefore no need to extend the respective part of the discussion.

Table 2 R^2 (including its robust variant) is a relative measure to describe the strength of the linear dependence of observed and predicted values. It does not depend on the variance of the data, unlike the root mean squared error (RMSE), which is an absolute measure for the precision of the predictions and thus depends on the variance of the observations. In our analysis, we used the *relative* RMSE, obtained by standardizing the prediction errors by the observations. Hence, we expect the relative RMSE to depend on the coefficient of variation (CV) of the data.

Considering the CVs of SOC stocks in the validation set (Table S1 in Supplement) we can see that a larger CV for the compartment 0–100 cm (0.63) than for 0–30 cm (0.51) is in accordance with larger relative RMSE for 0–100 cm (0.56) than for 0–30 cm depth (0.49, Table 2). The ratio of relative RMSEs for 0–100 cm and 0–30 cm (1.14) is somewhat smaller than the respective ratio of CVs (1.23) which is what we qualitatively expect if the R^2 is larger for 0–100 cm than for 0–30 cm soil depth.

2 Answers to comments by Philippe Lagacherie (RC C2581) of 17 February 2014

Many thanks for your positive comment. Your only concern is the missing evaluation of the *"added value of using the proposed robust external-drift kriging approach because a comparison with a more classical model is not provided in the paper"*. Unfortunately, we are not quite sure what is meant by "a more classical model".

We explained in our answer to the general comment by referee #1 (AC C2823) why we did not use regression kriging. Here we comment on the advantage of robust external-drift kriging (EDK) over

a) non-robust conventional EDK and

b) use of multiple linear regression models fitted by ordinary least squares (OLS) or robustly (MM estimator) to the data under the (wrong) assumption that the residuals are spatially uncorrelated.

To compare the precision of the respective predictions with the validation data we used the same set of covariates as for the final robust EDK model (robEDK) and did not go through the model building process again for these methods. In Table 1 below we report for the mentioned approaches the statistics of the relative prediction errors as in Table 2 of the article.

In general, the differences in the precision of the predictions by the four methods were small: For topsoil stocks (0–30 cm depth) ignoring autocorrelation impaired the precision somewhat. For bulk stocks (0–100 cm) this was the case only when the

regression model was estimated by OLS. In fact MM-regression performed best for these data. When comparing robust and customary EDK we see again small differences only, which is in accordance with the cross-validation results (see section 3.2 of article). robEDK, although not performing consistently best for all criteria, performed on average well. Hence, taking autocorrelation into account and using robust procedures offered some slight advantage over a more customary analysis.

3 Answers to comments by anonymous referee #3 (RC C2582) of 17 February 2014

Many thanks for treating our article with so much care and for giving us detailed feedback. In the sequel we would like to comment on your suggestions.

Validation with independent data

Your main criticism relates to our claim that we validate the precision of the predictions with "independent" data. Your view is that "validation with independent data" should be reserved for a comparison of predictions with extra data (collected by probability sampling, cf. Brus *et al.*, 2011) and that our validation approach does not differ from cross-validation exercises done earlier by other authors. We carefully evaluated these arguments, but in the end only partially agree:

We certainly accept that model assessment is most meaningful when predictions are compared to validation data that are 1) newly and independently collected from the calibration data by a randomized sampling design and 2) are not used in any phase of the model building process. Unfortunately, such a procedure can be rarely used in practice as it requires considerable funding and time and cannot be used for obvious reasons when only legacy data are available.

One is then bound to data splitting strategies and cross-validation for assessing the predictive power of a statistical model. You argue that there is no difference

Table 1: Statistics of relative prediction errors of soil organic carbon (SOC) stocks in two depth compartments (0–30 cm, 0–100 cm) for the validation set ($n = 175$). The statistics are reported for the method used in the article (robEDK), for non-robust external-drift kriging (EDK), and predictions by non-robustly (OLS) or robustly fitted (MM estimator) linear regression models (ignoring residual autocorrelation).

	model	BIAS	RMSE	R^2	robBIAS	robRMSE	robR^2	CRPS
0–30 cm	robEDK	0.135	0.488	0.346	0.070	0.388	0.337	0.221
	EDK	0.128	0.483	0.349	0.063	0.394	0.342	0.220
	MM est.	0.142	0.519	0.286	0.072	0.407	0.279	0.229
	OLS	0.143	0.500	0.335	0.077	0.389	0.321	0.222
0–100 cm	robEDK	0.152	0.556	0.477	0.066	0.420	0.403	0.247
	EDK	0.147	0.553	0.473	0.067	0.425	0.401	0.248
	MM est.	0.149	0.566	0.482	0.074	0.402	0.408	0.245
	OLS	0.162	0.569	0.468	0.082	0.428	0.391	0.249

between our data splitting approach and cross-validation as done earlier by Martin *et al.* (2009, 2011) and Meersmans *et al.* (2012). We are sorry to say that we do not share this view: Following Hastie *et al.* (2009, chap. 7) our validation set ($n = 175$) corresponds to a *test set* (used for estimating the generalization error, i.e. the prediction error for new data), and our calibration set ($n = 858$) has the combined role of a *training* (used for parameter estimation) and a *validation* set (used for estimating prediction errors during model selection). When no separate training and validation sets are available then cross-validation is often used for parameter estimation and model selection and for choosing tuning parameters of employed algorithms. This is exactly what we did: We chose the relevant covariates of the regression model and the robustness tuning constant by cross-validation using only the calibration data. Then we estimated the generalization error by applying the final model to the test set (i.e. to our validation data).

The "external validation" reported by Martin *et al.* (2009, 2011) and Meersmans *et al.* (2012) does not have the same significance as our validation results: Martin *et al.* (2009, 2011) used cross-validation (with the full data set) for selecting the tuning parameters of their algorithm and kept these parameters then fixed for estimating again by cross-validation the generalization error. Clearly, such a procedure gives a too optimistic estimate of the generalization error. Some information in the cross-validation subset, currently being predicted and used for estimating the generalization error, had been used before for model building. Also the reported "external validation" in Meersmans *et al.* (2012) provided only a distorted estimate of the generalization error because the structure of the regression model and its coefficients were not re-estimated in cross-validation (Meersmans, personal communication). Unfortunately, this is common practice in the geostatistical community, but is clearly not correct (Hastie *et al.*, 2009, sec. 7.10.2). Such a "validation" scheme tends to be overly optimistic when reporting the precision of predictions for new data. According to our knowledge only Mishra *et al.* (2009, 2010, 2012) and Wiesmeier *et al.* (2012) did not use measurements of the validation set for their model calibration. This justifies to explicitly mention these studies and to oppose them to the model validation strategy used by Martin *et al.* (2011) and Meersmans *et al.* (2012) who did the "external validation" (partly) with fixed parameters estimated from the full data set.

Moreover, we would like to comment on the concern that our validation data might not be spatially representative for SOC stocks of Swiss forests because it had been obtained by splitting a legacy data set (and not by independent probability sampling): 134 out of 175 soil profiles of our validation set stem from a survey where the sites were arranged on a 8×8 km grid. Data from these sites should be spatially representative because the grid was placed without consideration of SOC stocks. The additional 38 sites of this survey had to be assigned to the calibration set because there were no other data available for certain parts of the country. Another 38 soil profiles of the validation set stem from regional studies where the sites were arranged on 1×1 km grids. Again, these should be representative for local conditions and should not be influenced by sampling bias. You recommend that we use for validation the data available from the national and two cantonal soil monitoring networks as described in section 3.4 of Nussbaum *et al.* (2012). However,

the sites of these networks have been purposively selected similar to the majority of the WSL data and are likely not more representative than the data that we assembled for validation. Furthermore, soil sampling and analysis differed in these surveys from the procedure used by WSL and this likely adds some extra variation when we evaluate the precision of our predictions with this data. We have therefore good reason to believe that the generalization error, estimated with our validation set, provides a fair picture of the prediction error of our model for new data. Finally, although we could have repeated splitting the data randomly into calibration and validation sets (and then building and assessing the model for each such split) we abstained from such a procedure because we wanted to have a single final model that can be used in future applications to predict SOC stocks. Furthermore random splitting would have precluded the creation of a spatially representative validation data set as described above.

Further general comments

We further identified the following issues in the general part of your review, on which we would like to comment:

1. On page C2584 (line 25) of the review you suggested to discuss the added value of robust measures for evaluating the precision of SOC predictions. We see no real need for this as the inclusion would extend the length of the article: We mentioned in the article why we computed BIAS and RMSE of the *relative* prediction errors. We used robust variants of these statistics because non-robust quality measures are possibly influenced by only few observations that are poorly predicted. Robust quality measures provide a better picture how well a method predicts the majority of the observations. A detailed discussion of continuous ranked probability score (CRPS) can be found in Gneiting *et al.* (2007), to which we refer repeatedly in the paper, and we see no need to expand on this.

2. On page C2585 (line 1–10) you inquire about the increase in prediction errors when residual autocorrelation is neglected and predictions are computed from the regression models only. We have dealt with this issue in our answer to the review by Philippe Lagacherie (AC C2824).

3. You suggested to include graphs of the variograms in the article. We see again no need for this as such plots can be easily generated from the estimated variogram parameters listed in Table S5 of the Supplementary Material.

4. We refer now in the introduction of our article to the important review paper by Minasny *et al.* (2013). We were quite happy to see that our study addresses two issues that the authors explicitly mentioned as weaknesses of past studies: Validating SOC predictions with independent data and qualifying the precision of predictions by modelled standard errors.

5. Concerning the advantages of LASSO (least absolute shrinkage and selection operator) we refrain from adding a detailed discussion. First, non-specialist

information about this procedure is available (e.g. Hastie *et al.*, 2009, section 3). Second, we used LASSO only as a screening tool to find a preliminary set of covariates. The final set of covariates was selected by cross-validating the robustly fitted geostatistical models.

Specific comments

We accepted most of the suggestions as they improve the clarity of the article. We discuss only the comments where we (partly) disagree or some clarification is needed:

P7081 L23 We have changed the text and replaced "common model" by "the same set of fitted parameters". For a linear model the covariance of the prediction errors can be easily computed from the covariance matrix of the fitted regression coefficients (see any textbook on linear regression), for ML methods, it is not clear how to compute these covariances irrespective whether there is residual autocorrelation or not (see our answer to comments by referee #1, AC C2823).

P7082 L25 According to Table 2 in Martin *et al.* (2011) SOC stocks are more variable in French forests than on cultivated land. There is not yet harmonized soil data available to check this for Switzerland. Nevertheless, we mention now that forest soil stocks might be more variable and use this as a further justification for a separate analysis of the respective data.

P7083 L11 These percentages are correct. 29 % of the *total* area of Switzerland as opposed to 45.5 % of the *vegetated* area of Switzerland is covered by forests.

P7087 L5–15 This is a fair comment, which points to a weakness of our analysis (in particular of the validation scheme): We re-computed therefore the "median mass of soil particles < 2 mm" assigned to geotechnical map units using only the calibration data ($n = 858$) and re-fitted the model for topsoil SOC stocks (0–30 cm) to this data, however, without repeating the full model building process. Table 2 below lists for the validation set the statistics of the relative prediction errors of the re-calibrated model. Comparison with Table 2 of the article reveals that the statistics hardly changed. We abstain therefore from re-computing the predictions of regional and national SOC stocks as the figures published by Nussbaum *et al.* (2012) and listed in our article that are currently used for Switzerland's GHG inventory.

P7094 L6–9 Gneiting *et al.* (2007) give an exhaustive account on quality measures to validate probabilistic forecasts and we see no need to expand on this.

P7095 L20–25 We too refer to our answer to the comments by Philippe Lagacherie (AC C2824).

P7099 L2–9 We computed BIAS and RMSE of the *relative* prediction errors because this seems a natural choice for a lognormal model, where the coefficient of variation (and not the variance) is constant. Other studies have reported

absolute BIASes and RMSEs. The (robust) R^2 is therefore the only criterion available for a cross-study comparison.

Table 2: Statistics of relative prediction errors of soil organic carbon (SOC) stocks in topsoil (0–30 cm depth) for the validation set ($n = 175$). The model was fitted to the data where the covariate "mass of soil particles < 2 mm assigned to geotechnical map units" was computed only from the calibration data set ($n = 858$).

BIAS	RMSE	R^2	robBIAS	robRMSE	robR^2	CRPS
0.135	0.488	0.355	0.063	0.390	0.345	0.221

4 List of changes to the manuscript

According to our answer to the reviewers we changed the manuscript as follows:

P7079 L5 Added "SOC" in "... stocks. SOC stock estimates".

P7079 L6-7 Changed "... accounts in ..." to "... uses ...".

P7079 L7 Added "SOC" in "... for SOC stock changes".

P7079 L8 Changed "... grass- and other land ..." to "... grassland and other land cover types ...".

P7079 L10 Moved "(Brassel and Lischke, 2001)" to the end of the sentence.

P7079 L20 Deleted "therefore".

P7079 L25 Changed "... are estimated by averaging LM or ML point predictions for the nodes ..." to "... are estimated by averaging point predictions done by LM or ML for the nodes ...".

P7079 L28 Changed "This restriction generally..." to "The restriction to contiguous spatial covariates as only predictors generally...".

P7080 L16-17 Changed "Besides precision, also honest modelling of estimation errors of national SOC stocks matters for GHG inventories ..." to "Besides precise SOC estimates, standard errors of national SOC stocks are needed for GHG inventories ...".

P7080 L23-24 Changed wording from "... from a common model ..." to "... from the same set of fitted parameters ...".

P7080 L26 Removed "and in Martin *et al.* (2011, personal communication, 2013)".

P7081 L3 Changed "... estimates, and only few ..." to "... estimates. As pointed out by Minasny *et al.* (2013) only few ...".

P7081 L12 Added "Minasny *et al.* (2013)".

P7081 L28 Inserted "..., who used non-parametric tree-based methods, .. ".

P7082 L13, P7083 L12 Changed "... a.s.l. ..." to "... above sea level ...".

P7082 L19 Inserted "being" before "digitized".

P7082 L22-23 Changed "... carbon (OC) cropland and still 1.2 times ..." to "... carbon (OC) than cropland and still 1.2 times ...".

P7082 L23 Added a sentence: "Lastly, Martin *et al.* (2011) showed for France that forest SOC stocks are more variable than stocks on cultivated land.".

P7084 L7 Added "(Minasny *et al.*, 2013)".

P7084 L15 Changed "... formed the ..." to "... were used as ...".

P7085 L5 Inserted "(e.g. Krogh *et al.*, 2003; Meersmans *et al.*, 2008; Xu *et al.*, 2011)".

P7085 L10 Changed "... ca. ..." to "... about ...".

P7085 L13, L14 Changed "... levels ..." to "... categories ...".

P7085 L15 Changed "... soil samples ..." to "... soil horizons ...".

P7085 L25 Added "... than the PTF by Adams, but it was worse than the one by Jalabert *et al.* (2010) who recalibrated their PTF by ML methods.".

P7086 L10 Changed "... from the total volume V_i of the horizon per unit area [m^3 m^{-2}] ..." to "... from the thickness D_i of the horizon [m] ..." and replaced V_i by D_i in equation 1.

P7088 L16-17 Changed "... indices, computed by single and multi-flow algorithms (Tarboton, 1997), were available, too." to "... indices were computed by single and multi-flow algorithms (Tarboton, 1997).".

P7091 L2-5 Changed "After having tested the final model with the validation set (see below), we fitted it once more to the merged calibration and validation data to get better parameter estimates and to use all data for mapping." to "To get better parameter estimates for final mapping SOC stocks by EDK we fitted the model to the merged calibration and validation data. This was done after computing validation statistics (see above).".

P7092 L21 to P7094 L14 Moved section "2.3.5 Evaluating predictive performance of statistical models" to the end of page P7090 L26. This section has now the new number "2.3.3".

P7093 L7 Changed "... where MAD stands for median absolute deviation." to "... where $S(s_i)$ stands for calculated, $\tilde{S}(s_i)$ for predicted SOC stocks and MAD for median absolute deviation.".

P7096 L4 Changed "... loess smoothers ..." to "... loess scatterplott smoothers (Cleveland, 1979) ...".

P7096 L9 Changed "... the (non-)robust R^2, the models explained about 35 % of the variation of calculated SOC stock in 0–30 cm and 40–48 % of stocks ..." to "... the robust R^2, the models explained about 34 % of the variation of calculated SOC stock in 0–30 cm and 40 % of calculated SOC stocks ...".

P7098 L6 Added after "... profile sites." the sentences: "From the validation set only 14 of 175 sites were within these zones. Neglecting spatial autocorrelation but using the same set of covariates would slightly lower the precision of SOC stock estimates for the 0–30 cm and increase for the 0–100 cm depth compartment (Table 6 in Supplement)."

Changed "... within these zones from predictions ..." to "... within the estimated range of spatial dependence from predictions ...".

P7098 L12 Added after "... REML and kriging." the sentence: "This is reflected in the slight increase of robRMSE (0.6 % for SOC stock predictions in 0–30 cm and 0.5 % for stocks in 0–100 cm) compared to a non-robust fit of the model with the same covariates (Table 6 in Supplement).".

Table 2 Changed table caption "Statistics of relative soil organic carbon (SOC) stock prediction errors computed for the validation set in two depth compartments 0–30 cm and 0–100 cm" to "Statistics of relative prediction errors of soil organic carbon (SOC) stocks in two depth compartments (0–30 cm, 0–100 cm) for the validation set.".

Figure 3 In figure caption replaced "Boxplots of soil organic carbon ..." by "Boxplots of calculated soil organic carbon ...".

Supplement page 6, equation 3 and 4 Inserted $\hat{\theta}$ index for $\hat{Z}$ and $\hat{\beta}$.

Supplement page 9, Table 6 Added Table 1 in the answer to P. Lagacherie 2 (above) as new table to the Supplement.

References

Brassel, P. and Lischke, H., editors (2001). *Swiss National Forest Inventory: Methods and Models of the Second Assessment*. Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf.

Brus, D. J., Kempen, B., and Heuvelink, G. B. M. (2011). Sampling for validation of digital soil maps. *European Journal of Soil Science*, **62**(3), 394–407.

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, **74**, 829–836.

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society Series B*, **69**(2), 243–268.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning; Data Mining, Inference and Prediction*. Springer, New York, second edition.

Jalabert, S. S. M., Martin, M. P., Renaud, J.-P., Boulonne, L., Jolivet, C., Montanarella, L., and Arrouays, D. (2010). Estimating forest soil bulk density using boosted regression modelling. *Soil Use and Management*, **26**(4), 516–528.

Krogh, L., Noergaard, A., Hermansen, M., Greve, M. H., Balstroem, T., and Breuning-Madsen, H. (2003). Preliminary estimates of contemporary soil organic carbon stocks in Denmark using multiple datasets and four scaling-up methods. *Agriculture, Ecosystems & Environment*, **96**(1–3), 19–28.

Martin, M. P., Lo Seen, D., Boulonne, L., Jolivet, C., Nair, K. M., Bourgeon, G., and Arrouays, D. (2009). Optimizing pedotransfer functions for estimating soil bulk density using boosted regression trees. *Soil Science Society of America Journal*, **73**(2), 485 – 493.

Martin, M. P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boulonne, L., and Arrouays, D. (2011). Spatial distribution of soil organic carbon stocks in France. *Biogeosciences*, **8**(5), 1053–1065.

Meersmans, J., De Ridder, F., Canters, F., De Baets, S., and Van Molle, M. (2008). A multiple regression approach to assess the spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). *Geoderma*, **143**(1–2), 1–13.

Meersmans, J., Martin, M. P., Lacarce, E., De Baets, S., Jolivet, C., Boulonne, L., Lehmann, S., Saby, N. P. A., Bispo, A., and Arrouays, D. (2012). A high resolution map of French soil organic carbon. *Agronomy for Sustainable Development*, **32**(4), 841–851.

Minasny, B. and McBratney, A. B. (2007). Spatial prediction of soil properties using EBLUP with the Matérn covariance function. *Geoderma*, **140**(4), 324–336.

Minasny, B., McBratney, A., Malone, B., and Wheeler, I. (2013). Digital mapping of soil carbon. *Advances in Agronomy*, **118**, 1–47.

Mishra, U., Lal, R., Slater, B., Calhoun, F., Liu, D., and Van Meirvenne, M. (2009). Predicting soil organic carbon stock using profile depth distribution functions and ordinary kriging. *Soil Science Society of America Journal*, **73**(2), 614–621.

Mishra, U., Lai, R., Liu, D., and Van Meirvenne, M. (2010). Predicting the spatial variation of the soil organic carbon pool at a regional scale. *Soil Science Society of America Journal*, **74**(3), 906–914.

Mishra, U., Torn, M. S., Masanet, E., and Ogle, S. M. (2012). Improving regional soil carbon inventories: Combining the IPCC carbon inventory method with regression kriging. *Geoderma*, **189 – 190**(0), 288 – 295.

Nussbaum, M., Papritz, A., Baltensweiler, A., and Walthert, L. (2012). Organic carbon stocks of swiss forest soils. Final report, Institute of Terrestrial Ecosystems, ETH Zürich and Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Zürich and Birmensdorf.

Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, **33**(2), PP. 309–319.

Wiesmeier, M., Spörlein, P., Geuss, U., Hangen, E., Haug, S., Reischl, A., Schilling, B., Lützow, M., and Kögel-Knabner, I. (2012). Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. *Global Change Biology*, **18**(7), 2233–2245.

Xu, X., Liu, W., Zhang, C., and Kiely, G. (2011). Estimation of soil organic carbon stock and its spatial distribution in the Republic of Ireland. *Soil Use and Management*, **27**(2), 156–162.