Soil Use and Management, Volume 34, Issue 4, December 2018, pp. 533-541
DOI:10.1111/sum.12446

Digital soil assessment for quantifying soil constraints to crop production: a case study for rice in Punjab, India

EZEKIEL I. OKONKWO, RON CORSTANJE & GUY J.D. KIRK

Cranfield Soil and Agrifood Institute, Cranfield University, Cranfield, MK43 0AL, UK

Correspondence: E. I. Okonkwo. E-mail: e.okonkwo@cranfield.ac.uk

Running Title: Digital Soil Assessment

Abstract

Assessments of land capability for particular functions such as food production need to allow for uncertainties both in the criteria used to specify the function and in information on relevant soil properties. In this paper we evaluate the use of digital soil assessment (DSA) for dynamic assessment of soil capability allowing for both uncertainties and spatial variability in soil properties, and flexibility in the values of assessment criteria. We do this for soil constraints to rice production in the state of Punjab, India, where soil salinity and alkalinity are potentially important constraints to cropping. In DSA, spatial predictions of soil properties and associated uncertainties made with digital soil mapping (DSM) are used to assess soil functions. We use a combination of DSM and Monte Carlo simulation methods to estimate the spatial variation in soil electrical conductivity (ECe) and pH to 20 cm depth in soils across Punjab. We then use the estimates and associated uncertainties to assess the likelihood that soil salinity or alkalinity or both could constrain rice production. Results show that allowing for prediction uncertainties of soil attributes results in far smaller areas affected by salinity (1.2 versus 2.0 Mha) and alkalinity (3.0 vs 3.2 Mha). Results also show the importance of correctly setting threshold values for constraint criteria and the flexibility of the DSA approach for setting thresholds.

Keywords: Digital soil mapping, soil spatial variability, rice, problem soils

Introduction

Land capability assessment for crop production is often limited by the availability of high-resolution spatial soil data, and assessment criteria are therefore correspondingly coarse. Digital soil mapping (DSM) has emerged as a credible alternative to traditional soil survey methods (Lagacherie, 2008; Minasny & McBratney, 2016). As well as being cheaper and faster, DSM has the advantage over traditional soil mapping that its statistical models are repeatable and uncertainty of outputs can be estimated. This facilitates the translation of quantitative information on soil properties generated through DSM into assessments of constraints to soil functions, such as supporting crop production. This is the purpose of digital soil assessment (DSA; Carré *et al.*, 2007; van Zijl *et al.*, 2014; Harns et al., 2015).

In traditional soil survey, soil suitability assessments are made according to categorical classifications tailored to the spatial resolution of soil maps. As such they may not adequately reflect the true spatial variability of the soil and they cannot convey uncertainty. By contrast, DSA incorporates the high-resolution spatial information on soil properties and associated uncertainties provided by DSM into a flexible, quantitative framework for assessing soil functions. This process has been applied in identifying environmental risks and in land evaluation, such as for crop production (Kidd *et al.*, 2015; Malone *et al.*, 2015). But full DSA methods are yet to be widely applied, in spite of the widespread use of DSM methods. This is particularly so for areas where existing soil information is sparse. But such areas are where DSA may have the greatest potential.

In this paper, we evaluate DSA for assessing soil constraints to cropping in areas with limited existing soil data. The methodology we develop involves mapping soil properties by DSM and using a stochastic simulation technique and specified threshold values for soil constraints to determine the spatial extent of the constraints. We do this for the lowland rice growing areas of Punjab in the Indo-Gangetic plain of India, as a case study. A coarse-resolution assessment of soil constraints to rice production globally was made by Haefele et al. (2014) producing one mapping unit per 17,400 ha across 112 countries. We aim for a much finer resolution. We chose Punjab because of its importance for food production, and because of the extent of salinity and alkalinity (high pH) in the soils (Sidhu *et al.*, 1995), and because recent studies indicate a general decline in soil fertility in the Indo-Gangetic plain due to abiotic, biotic and socio-economic factors (Bhandari *et al.*, 2002). We have the additional complications with lowland rice that the changes in soil chemistry with flooding for rice mean the soil conditions during the rice season may be quite different to those in the air-dry aerobic soil on which most measurements are made (Ponnamperuma, 1972; Kirk, 2004). Also long-term growth of lowland rice produces permanent changes in the soil (Greenland, 1997). The aim of the paper is to assess the use of DSA for gauging soil constraints to cropping in areas with limited soil data, as for rice soils in Punjab. We develop methods for allowing quantitatively for uncertainties in the soil information. We also demonstrate the flexibility of the DSA approach to correctly set threshold values for soil constraints in poorly quantified soil systems, such as for paddy rice soils.

Materials and Methods

In brief, we used DSM methods to model the relationships between soil attributes and environmental covariates to produce 250-m resolution digital soil maps with corresponding uncertainties. We then used DSA methods to map the extent of soil salinity and alkalinity constraints to rice production, as defined by set criteria, using Monte Carlo methods to allow for the effects of uncertainties.

Study area

Punjab is in northwest India (Fig. 1). It covers an area of 50,362 km^2 and is divided into 22 Districts with three Divisions and 46 Sub-divisions. The predominant crops are rice and wheat. The climate is predominantly sub-tropical semi-arid and monsoonic. The mean annual rainfall is 705 mm and varies from 1200 mm at Pathankot to < 300 mm at Abohar. The monsoon season is from July to September. Dry conditions prevail from October to the end of June, except for light showers from December to February. The general slope is from northeast to southwest and the elevation is 180 to 300 m AMSL in the plains; and 300 to 700 m AMSL in the Siwaliks. The major geological units are Siwalik tertiary deposits and recent alluvium, and the main landforms are alluvial plain, piedmont plain, Siwalik Hills and aeo-fluvial plain.

The soils of the Siwaliks and piedmont plain are deep to very deep, well drained and with loamy sand to sandy loam texture. The soils of the alluvial plain (90% of the area) are very deep, well to moderately-well drained with textures varying from sandy loam to silty clay loam. The soils have been classified by the USDA system (Soil Survey Staff, 2014) into seven sub-orders (Aquepts, Fluvents, Ochrepts, Orthents, Argids and Psamments) and fifteen sub-groups by the National Bureau of Soil Survey and Land Use Planning (NBSS&LUP), India (Sidhu et al., 1995). The dominant soils are red Ochrepts covering 50% of the area.

Digital soil mapping

Input datasets. Legacy soil data were obtained from the National Bureau of Soil Survey and Land Use Planning (NBSS, Sidhu *et al.*, 1995). This included soil maps covering the Punjab at 1:500,000 scale and depth-resolved data on soil properties for representative profiles at 34 locations (Fig. 1). The soils are classified into four orders and 15 sub-groups of USDA system (Soil Survey Staff, 1986) with 124 mapping units. The soil map sheets appended with the report were digitized. A digital elevation model (DEM) at 30 m resolution was generated from CartoDEM of the National Remote Sensing Centre, India (NRSC ISRO, 2015). From the DEM, common soil covariates such as elevation, slope and flow direction, representing landscape morphology, erosion and deposition processes, were determined. Vegetation data at 250 m resolution were derived from monthly averages of sixteen-day

MODIS products (MOD13Q1) of the United States Land Processes Distributed Active Archive Centre (NASA, 2016). Land use and land cover at 30 m resolution was extracted from GlobeLand30 dataset (Chen *et al.*, 2014). Climatic variables for the 1950-2000 periods at 1 km spatial resolution were generated from the WorldClim database (Hijmans et al., 2005). Landforms were derived by digitizing 1:240,000 map sheets obtained from World Soil Survey Archive and Catalogue (WOSSAC, 2017). All data sources were brought to a common 250 m resolution.

Depth-averaged data. Analytical data for 34 representative soil profiles were obtained from the NBSS (Sidhu *et al.*, 1995). An equal-area quadratic spline function (Malone *et al.*, 2009) was fitted to the values of soil ECe (representing salinity) and pH at the depths reported to a maximum depth of 2 m, in order to derive depth-averaged soil properties over a nominal rice rooting depth of 0–20 cm The spline function *f*(*z*) was obtained by minimising the following equation applied to the measured property values y_i at n depths z_i:

$$\frac{1}{n}\sum_{i=0}^{n}(y_i - \bar{f}_i)^2 + \lambda\int_{z_0}^{z_n}\left[f'(z)\right]^2 dz \qquad (1)$$

where $\bar{f}_i$, is the mean value of $f(z)$ over the ith depth interval, $f'(z)$ is the first derivative of $f(z)$ and λ is a smoothing parameter (here 0.1) (Bishop *et al.*, 1999). The first term in Equation (1) represents the fit of the spline to the data; the second term measures the roughness of the function $f(z)$. The parameter λ controls the trade-off between the fit and the roughness of the spline. The observed ECe and pH values had a log-normal distribution and were log-transformed before splines were fitted. Mean values over 0–20 cm depth were then obtained by integrating the spline function over 0–20 cm. These values were used for the DSA maps.

Modelling. Artificial neural networks (ANN) based on multilayer perceptron (MLP) were used to model the relationship between the desired soil properties (ECe and pH) and environmental covariates. A total of 20 networks were trained and the five best-performing networks were retained. The data were randomly divided into 80 % for model calibration and 20 % for validation. We evaluated regression model performance using 20% of the soil datasets. We used three criteria to evaluate the accuracy of the spatial prediction models of ECe and pH: root mean square error (RMSE), coefficient of determination (R^2) and adjusted coefficient of determination (adjusted R^2). We used neural networks on the basis that predictions can be made simultaneously at each depth interval, and that the resulting model can capture a greater variety of non-linear relationships than a traditional logistic regression model. The ANN model was then used to generate spatial predictions of ECe and pH across the study area. ArcGIS software was used to generate raster grids of ECe or pH spatial distributions for each modelled depth at each pixel within the study area.

Digital soil assessment

We focused on potential soil constraints that are widespread in Punjab: high salinity and high pH. The numerical thresholds set to define constraints are critical to any mapping exercise. A strength of DSA is that it is easy to assess the consequences of varying threshold values, and to set appropriate values. We first used threshold values from the Fertility Capability Classification (FCC) system of (Sanchez et al., 2003), which includes modifications for lowland rice following Sanchez and Buol (1985). In the FCC system, high salinity is defined as ECe (measured in a saturated paste) > 4 dS m^{-1} and high pH is defined as pH (1:1 in H_2O) > 7.3. However, these values are only crude approximations for lowland rice given that, for practical reasons, they must be determined on the soil when air-dry and therefore aerobic, whereas the relevant values for lowland rice may be very different given the changes in soil chemistry that occur following soil flooding. Further, ECe values only indicate the total salt content, not its composition, whereas salinity stress is sensitive to the soil solution composition, particularly the ratio of Na to other cations. Based on actual lowland rice responses, Fairhurst et al. (2007) give the following values for yield reductions in rice due to salinity: ECe > 4 dS m^{-1} slight yield reduction (10–15%); ECe > 6 dS m^{-1} moderate yield reduction (20–50%); ECe > 10 dS m^{-1} > 50% yield reduction in susceptible cultivars. Accordingly we set a critical value of ECe > 8 dS m^{-1} for high salinity. For high pH, the principal effects on rice are deficiencies of micronutrients, particularly of Zn. However, pH is only one of the determinants. Particularly, Zn deficiency in rice is also associated with high soil organic C content and perennially wet conditions (Kirk, 2004). Further, flooding alkaline soils for rice generally produces a decrease in pH towards neutral, even in soils with aerobic pH > 9 (Kirk, 2004). According to IRRI (2018), pH 8.8–9.2 is considered as non-stress, pH 9.3–9.7 moderate stress and pH > 9.7 higher stress. Accordingly we set a critical value of pH > 9.3 for high pH.

Allowing for uncertainty in soil constraint assessment

We used Monte Carlo simulations to generate likelihood scenarios for constraint levels and assess the real extent of likely constrained soils (Fisher, 1991; Heuvelink, 1998; Nol *et al.*, 2010; Nelson *et al.*, 2011). The basis of the Monte Carlo method is the simulation of possible realisations (pixel-by-pixel) of the input variables to obtain a distributional description of the likelihood of these to exceed the constraint criteria. One hundred realisations of the model were obtained from the probability distributions based on the predicted mean and standard error of ECe and pH values at top soil level (0–20 cm). The probabilities were estimated as the number of occurrences of each constraint class divided by the total number of realisations. Pixels were then classified based on values above the defined critical levels for ECe and pH (ECe > 4 or 8 dS m^{-1}; pH > 7.3 or 9.3). If the probability of the critical value exceeding the set level was less than 50 %, the soil was classified as not saline or not alkaline; if the probability was greater than 50% but less than 80 %, the soil was classified as possibly saline or possibly alkaline; otherwise, the soil was classified as likely saline or likely alkaline. The resulting likelihood maps (salinity and alkalinity) were displayed as a soil constraint map of Punjab.

The remaining steps were undertaken only for grid cells with rice cultivation, as identified by Gumma *et al.* (2011). We intersected the raster cell values representing the area of each rice cell with the constraint levels of ECe and pH to compute the areal coverage of soil constraints to rice production. We aggregated and tabulated these data to the district level and report the results in hectares and percentages.

Results

Descriptive statistics of ECe and pH

Figure 2 shows the dependence of predicted soil ECe and pH (0–20 cm depth) on auxiliary variables. It shows the most important predictors of both ECe and pH are soil class and landform. The DSM outputs of the spatial distribution of ECe and pH at the standard soil depths are displayed in Figures S1 and S2 (Supplementary Information). The figures indicate that salinity decreases with increasing depth. The mapping further shows that the Southwest region of the state drained by the Satluj and Beas rivers are largely saline. The soils on the plain, underlain by alluvium, are often alkaline in the subsoil or throughout the profile.

The results of the model evaluation and the assessment of the uncertainty of DSM outputs used for DSA are shown in Table 1. The RMSE shows a low prediction accuracy for soil ECe (RMSE = 5.44 dS m^{-1}) compared with the prediction of soil pH (RMSE = 0.44). Similarly, the adjusted coefficient of determination (adjusted R^2) indicates that prediction accuracy of the model was 40 % for soil ECe and 54 % for soil pH (Table 1). Allowing for prediction uncertainties improved the accuracy of prediction for soil pH and ECe as indicated by the coefficient of variation (CV) (Table 1). Figure 3 shows the distribution of the predicted soil pH and ECe values across the study area.

Digital soil assessment

The outputs from the soil constraint assessments using alternative DSA techniques (with and without allowing for prediction uncertainties) and with alternative threshold values for the soil constraints (FCC values versus our own values) are shown in Figure 4 and Table 2. The results in Table 2 are that, with the FCC criteria, and not allowing for prediction uncertainties, over 2 Mha (43 % of the total land area) of Punjab is predicted to have salinity problems and 3.2 Mha (64 %) alkalinity problems. A further 24 % is classified as possibly saline and 10 % as possibly alkaline. These numbers far larger than realistic for Punjab, given that it is one of the main rice production areas in India. Allowance for model uncertainties reduced the estimates somewhat: 1.16 Mha (23 %) likely saline and 2.98 Mha (59 %) likely alkaline (Fig. 4a and Table 2). However these numbers are still larger than realistic. The reason for these larger numbers compared with values in Figure 3 is because the classification was based on estimated probabilities of ECe and pH which allowed for those values that are near or slightly above the critical levels to define the constraints.

Figure 4b and Table 2 show results using more stringent critical values of soil salinity and alkalinity, better matching paddy soil conditions, and allowing for model uncertainties. The areas categorised as likely or possibly affected by salinity or alkalinity or both are greatly diminished compared with the FCC criteria, demonstrating the sensitivity of predictions to the criteria. These results are far more consistent with expectations for Punjab based on the productivity of rice there. From results for individual districts using the more stringent criteria (Table S1), about 9% (0.25 Mha) of all districts are likely saline, and 5% (0.14 Mha) likely alkaline. None of the areas is classified as possibly saline or alkaline. The saline soils are mostly located in southwest of the Punjab, which is a semi-arid alluvial plain. This is most likely due to accumulation of soluble salts brought in from higher elevation areas. Salinity is also found in patches in other areas, possibly due to poor quality irrigation water and management.

Discussion

Advantages of the DSA approach compared with more-traditional soil suitability assessments include: (i) the incorporation of uncertainty in the assessment results in the assessment as likelihood of a soil constraint, reflecting both the underlying spatial variability and the relative quality of the available data on the soil; and (ii) it is relatively straightforward to run reassessments with new suitability criteria as new data are obtained, over-turning past assumptions, or when a new cropping or soil-crop combination is being considered for which the original suitability criteria are inappropriate. The importance of allowing for uncertainties and having appropirate constraint criteria is illustrated by the senistivity of our constraint estimates to the various criteria we tested. Malone *et al.* (2015) showed similar sensitivities in the land suitability assessment approach they used. Harms *et al.* (2015) discuss the importance of quantifying uncertainties in land suitability assessments.

The sample support in this study is sparse, but equivalent to those used in other studies (e.g. Hengl et al., 2017) for areas where it is difficult to obtain hard data on soil variability. We have not sought to produce a definitive map of soil constraints for the Indian Punjab. Rather we aimed to illustrate how, for areas of the world where data is sparse but the assessment of soil capability critical for food supply, this approach can allow for the consequences of data scarcity.

The results show the sensitivity of predictions to the criteria used to define constraints. The areas predicted to be affected by salinity or alkalinity were far smaller with the revised, more-stringent criteria than with the FCC criteria. These criteria are more appropriate for lowland rice than the FCC criteria because, in most rice soils, the biogeochemical changes that occur following flooding for rice cause decreases in both salinity and alkalinity. Hence, higher threshold values are appropriate for the un-flooded aerobic soil in which the underlying measurements were made. The results demonstrate the importance of setting appropriate threshold criteria for judging soil constraints for particular applications. The

DSA approach provides a quantitative basis for setting such criteria. In analysing gaps between potential crop yields – set by agro-climatic constraints in a given region – and actual yields in the region (e.g. Laborte et al., 2012; Silva et al., 2017), DSA can be used to quantify the potential contributions of soil constraints versus socio-economic, farm- management and other factors. Threshold values for soil constraints can be set using DSA in regions for which non-soil constraints to production are well defined.

This research highlights how digital soil constraint assessments can flexibly assess the functional capacity of soils for particular applications, taking account of uncertainties arising from incomplete understanding of soil variability in the region being assessed. The DSA approach provides a quantitative basis for assessing what level of measured soil information is required to support land suitability assessments, and hence the need for ground-based soil survey. As requirements for soil information increasingly focus on what the soil does (soil functions) rather than its properties, this study illustrates the benefits of using a DSA approach, incorporating information from DSM to describe and model soil functioning in a manner which allows for different evaluations of those functions.

Conclusions

This study provides an example where soil attributes generated from limited existing legacy soil information were incorporated into a DSA framework to assess soil constraints in Punjab rice production. Using this approach, it was possible to quantitatively assess the importance of incorporating prediction uncertainties in soil constraint assessments. The research demonstrates the flexibility of DSA for matching assessment critical values to actual constraints. Because the assessment of the soil constraints was done digitally, it is possible to continually update mapping to improve accuracy.

Acknowledgements

This research was funded by the Nigeria Tertiary Education Trust Fund (TETFund) through a scholarship to EO and by a grant from the UK's Biotechnology and Biological Sciences Research Council (BBSRC, Grant Ref. BB/J011584/1) under the Sustainable Crop Production Research for International Development (SCPRID) programme, a joint multi-national initiative of BBSRC, the UK Government's Department for International Development (DFID) and (through a grant awarded to BBSRC) the Bill & Melinda Gates Foundation (BMGF). We thank Parvesh Chandna and JK Ladha of the International Rice Research Institute (IRRI) for help in obtaining the Punjab soil data. We also thank Joanna Zawadzka, Thomas Mayr and Stephen Hallett of Cranfield for helpful discussions.

References

Bhandari, A.L., Ladha, J.K., Pathak, B.H., Padre, A.T., Dawe, D., Gupta, R.K., 2002. Yield and Soil Nutrient Changes in a Long-Term Rice-Wheat Rotation in India. Soil Sci. Soc. Am. 66, 162–170.

Bishop, T.F.., McBratney, A.B., Laslett, G.M., 1999. Modeling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma 91, 27–45. doi:10.1016/S0016-7061(99)00003-8

Carré, F., McBratney, A.B., Mayr, T., Montanarella, L., 2007. Digital soil assessments: Beyond DSM. Geoderma 142, 69–79. doi:10.1016/j.geoderma.2007.08.015

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 103, 7–27. doi:10.1016/j.isprsjprs.2014.09.002

Fairhurst, T.H., Witt, C., Buresh, R.J., Dobermann, A., 2007. A practical Guide to Nutrient Management, 2nd ed. International Rice Research Institute, Manila, Philippines.

Fisher, P., 1991. Modelling soil map-unit inclusions by Monte-Carlo simulation. Int. Jounal Geogr. Inf. Syst. 5, 193–208.

Greenland, D.J., 1997. The sustainability of rice farming. International Rice Research Institute, CAB International Publishing.

Gumma, M.K., Nelson, A., Thenkabail, S.P., Singh, N.A., 2011. Mapping rice areas of South Asia using MODIS multitemporal data. J. Appl. Remote Sens. 5, 1–26. doi:10.1117/1.3619838

Haefele, S.M., Nelson, A., Hijmans, R.J., 2014. Soil quality and constraints in global rice production. Geoderma 235–236, 250–259. doi:10.1016/j.geoderma.2014.07.019

Harns, B., Brough, D., Philip, S., Bartley, R., Clifford, D., Thomas, M., R., W., Gregory, L.J., 2015. Digital soil assessment for regional agricultural land evaluation. Glob. Food Sec. 5, 25–35.

Hengl, T., Leenaars, J.G.B., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B.M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., Wheeler, I., Kwabena, N.A., 2017. Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutr. Cycl. Agroecosystems 109, 77–102. doi:10.1007/s10705-017-9870-x

Heuvelink, G.B.M., 1998. Error propagation in environmental modelling with GIS. Taylor & Francis.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. doi:10.1002/joc.1276

IRRI, 2018. Rice Knowledge Bank. [WWW Document]. URL http://www.knowledgebank.irri.org (accessed 3.16.18).

Kidd, D., Webb, M., Malone, B., Minasny, B., McBratney, A., 2015. Digital soil assessment of agricultural suitability, versatility and capital in Tasmania, Australia. Geoderma Reg. 6, 7–21. doi:10.1016/j.geodrs.2015.08.005

Kirk, G.J.D., 2004. The Biogeochemistry of Submerged Soils. John Wiley & Sons, Chiester, UK. doi:10.1093/aob/mci162

Laborte, A.G., de Bie, K. (C. a. J.M.., Smaling, E.M. a., Moya, P.F., Boling, A. a., Van Ittersum, M.K., 2012. Rice yields and yield gaps in Southeast Asia: Past trends and future outlook. Eur. J. Agron. 36, 9–20. doi:10.1016/j.eja.2011.08.005

Lagacherie, P., 2008. Chapter 1. Digital soil mapping: a state of the art., Digital soil mapping with limited data. Springer Science+Business Media B.V. doi:10.1017/CBO9781107415324.004

Malone, B.P., Kidd, D.B., Minasny, B., McBratney, A.B., 2015. Taking account of uncertainties in digital land suitability assessment. PeerJ 3, 1–21. doi:10.7717/peerj.1366

Malone, B.P., McBratney, A.B., Minasny, B., Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. Geoderma 154, 138–152. doi:10.1016/j.geoderma.2009.10.007

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. doi:10.1016/S0016-7061(03)00223-4

NASA, 2016. MODIS13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250 Grid SIN V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5057/MODIS/MOD13Q1.V006.

Nelson, M.A., Bishop, T.F.A., Triantafilis, J., Odeh, I.O.A., 2011. An error budget for different sources of error in digital soil mapping. Eur. J. Soil Sci. 62, 417–430. doi:10.1111/j.1365-2389.2011.01365.x

Nol, L., Heuvelink, G.B.M., Veldkamp, A., de Vries, W., Kros, J., 2010. Uncertainty propagation analysis of an N2O emission model at the plot and landscape scale. Geoderma 159, 9–23. doi:10.1016/j.geoderma.2010.06.009

NRSC ISRO, 2015. Cartosat-1 30 m Digital Elevation Model (CartoDem) Version-3 R1. National Rempote Sensing Centre, Indian Space Research Organisation.

Ponnamperuma, F.N., 1972. The chemistry of submerged soils. Adv. Agron. 24, 29–26.

Sanchez, P.A., Buol, S.W., 1985. Constraints to Rice Production and Wetland Soil Characteristics, in: Wetland Soils: Characterization, Classification and Utilization. International Rice Reasearch Institute, Los Banos Philippines.

Sanchez, P.A., Palm, C.A., Buol, S.W., 2003. Fertility capability soil classification: A tool to help assess soil quality in the tropics. Geoderma 114, 157–185. doi:10.1016/S0016-7061(03)00040-5

Sidhu, G.., Walia, C.S., Lal, T., Rana, K.P.C., Seghal, J., 1995. Soils of Punjab for

Optimising Land Use. NBSS Publ.45 (Soils of India Series 4). National Bureau of Soil Survey and Land Use Planning. Nagpur, India, 75p + 2 sheets soil map (1:500, 000 scale).

Silva, J.V., Reidsma, P., Laborte, A.G., Van Ittersum, M.K., 2017. Explaining rice yields and yield gaps in Central Luzon, Philippines: An application of stochastic frontier analysis and crop modelling. Eur. J. Agron. 82, 223–241.

Soil Survey Staff, 2014. Keys to Soil taxonomy. 12th Edition, USDA-Natural Resources Conversation Service, Washington DC.

Soil Survey Staff, 1986. Keys to Soil Taxonomy, 3rd ed. United States Department of Agriculture, Washington.

van Zijl, G.M., Bouwer, D., van Tol, J.J., le Roux, P.A.L., 2014. Functional digital soil mapping: A case study from Namarroi, Mozambique. Geoderma 219–220, 155–161. doi:10.1016/j.geoderma.2013.12.014

WOSSAC, 2015. World Soil Survey Archive and Catalogue. Cranfield University, United kingdom.

Table 1 Model evaluation and descriptive statistics of predictions of soil ECe and pH and associated uncertainties at 0–20 cm depth. RMSE = Root mean squared error of predicted soil ECe and pH, R^2 = Coefficient of determination of the prediction accuracy, Adjusted R^2 = Adjusted Coefficient of determination of the prediction accuracy, MC = Monte Carlo simulation analysis of prediction uncertainties

				Observed			Predicted			MC		
	RMSE	R^2	Adjusted R^2	Mean	Stdev	CV	Mean	Stdev	CV	Mean	Stdev	CV
ECe (dS m^{-1})	5.44	0.97	0.40	3.41	4.39	128.6	2.98	4.56	153	2.95	3.57	121
pH	0.44	0.98	0.54	8.03	0.63	7.88	8.02	0.64	7.98	8.51	0.63	7.46

Table 2 Extent of soil salinity (ECe) and alkalinity (high pH) constraints to rice production in Punjab assessed using FCC and alternative constraint criteria and DSA techniques with and without allowing for prediction uncertainties using Monte Carlo (MC) methods. Data are areas in Mha with percentages of total area in parenthesis

	Saline			Alkaline		
	Likely	Possibly	Not	Likely	Possibly	Not
Whole area (5,010 Mha)						
FCC criteria only	2,140 (43)	1,207 (24)	1,663 (33)	3,219 (64)	806 (16)	985 (20)
FCC criteria with MC	1,161 (23)	501 (10)	3,348 (67)	2,980 (59)	982 (20)	1,048 (21)
Rice area (2,800 Mha)						
FCC criteria only	1,061 (38)	931 (33)	809 (29)	1,976 (71)	370 (13)	455 (16)
FCC criteria with MC	540 (19)	201 (7)	2,059 (74)	1,894 (68)	395 (14)	511 (18)
Alternative criteria with MC	247 (9)	0 -	2,554 (91)	148 (5)	0 -	2,652 (95)

Figures Captions

Figure 1 Study area indicating the locations of the representative soil profiles with analytical data used for soil property mapping

Figure 2 The dependence of (a) soil ECe and (b) soil pH at 0–20 cm depth on auxiliary variables: LDF, landform; Soil_Tax, USDA soil class; Flowdr, Flow Direction; Lulc30, land cover; RedR, Red Reflectance Band 1; Prec, precipitation; MIR: mid-infrared reflectance band 7; Tmean, mean temperature; NIR: near-infrared reflectance band 2; Dem, digital elevation model; NDVI, normalised difference vegetation index.

Figure 3 Distribution of predicted soil pH and ECe at 0–20 cm depth. Bars are predicted values; red line is a fitted normal distribution.

Figure 4 Soil constraints maps allowing for prediction uncertainties with two sets of constraint criteria: (a) the FCC criteria, and (b) more stringent criteria that are more appropriate for lowland rice.

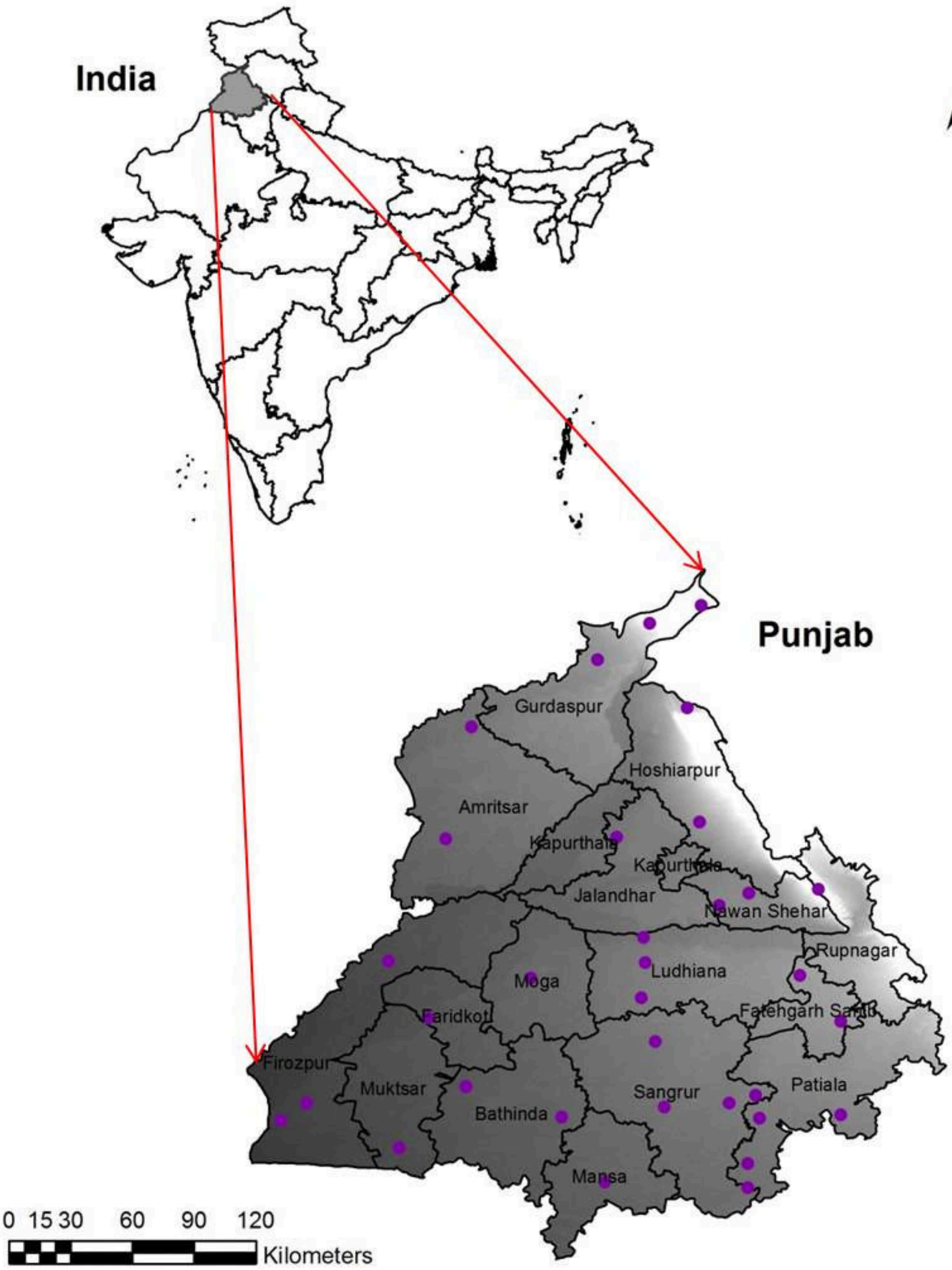

Figure 2 Study area indicating the locations of the representative soil profiles with analytical data used for soil property mapping.

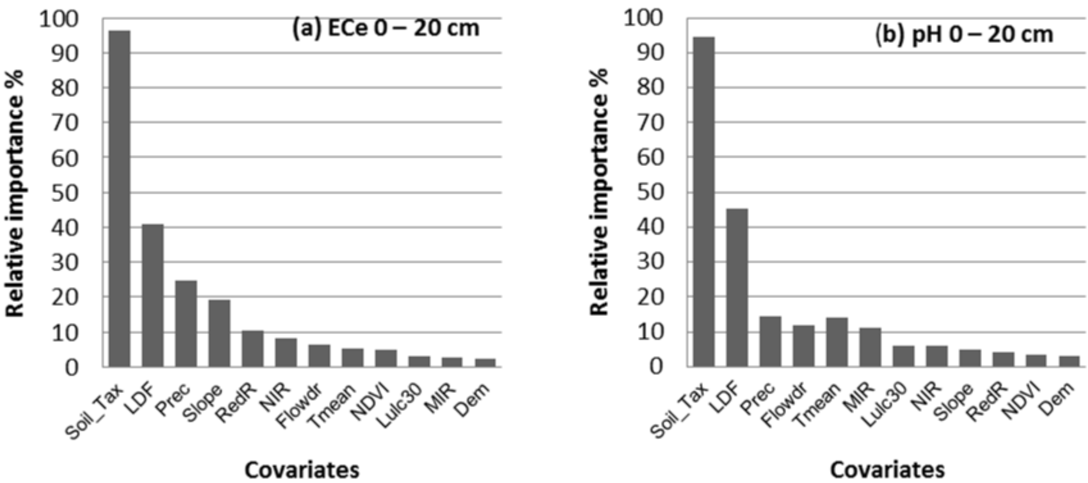

Figure 2 The dependence of (a) soil ECe and (b) soil pH at 0–20 cm depth on auxiliary variables: LDF, landform; Soil_Tax, USDA soil class; Flowdr, Flow Direction; Lulc30, land cover; RedR, Red Reflectance Band 1; Prec, precipitation; MIR: mid-infrared reflectance band 7; Tmean, mean temperature; NIR: near-infrared reflectance band 2; Dem, digital elevation model; NDVI, normalised difference vegetation index.

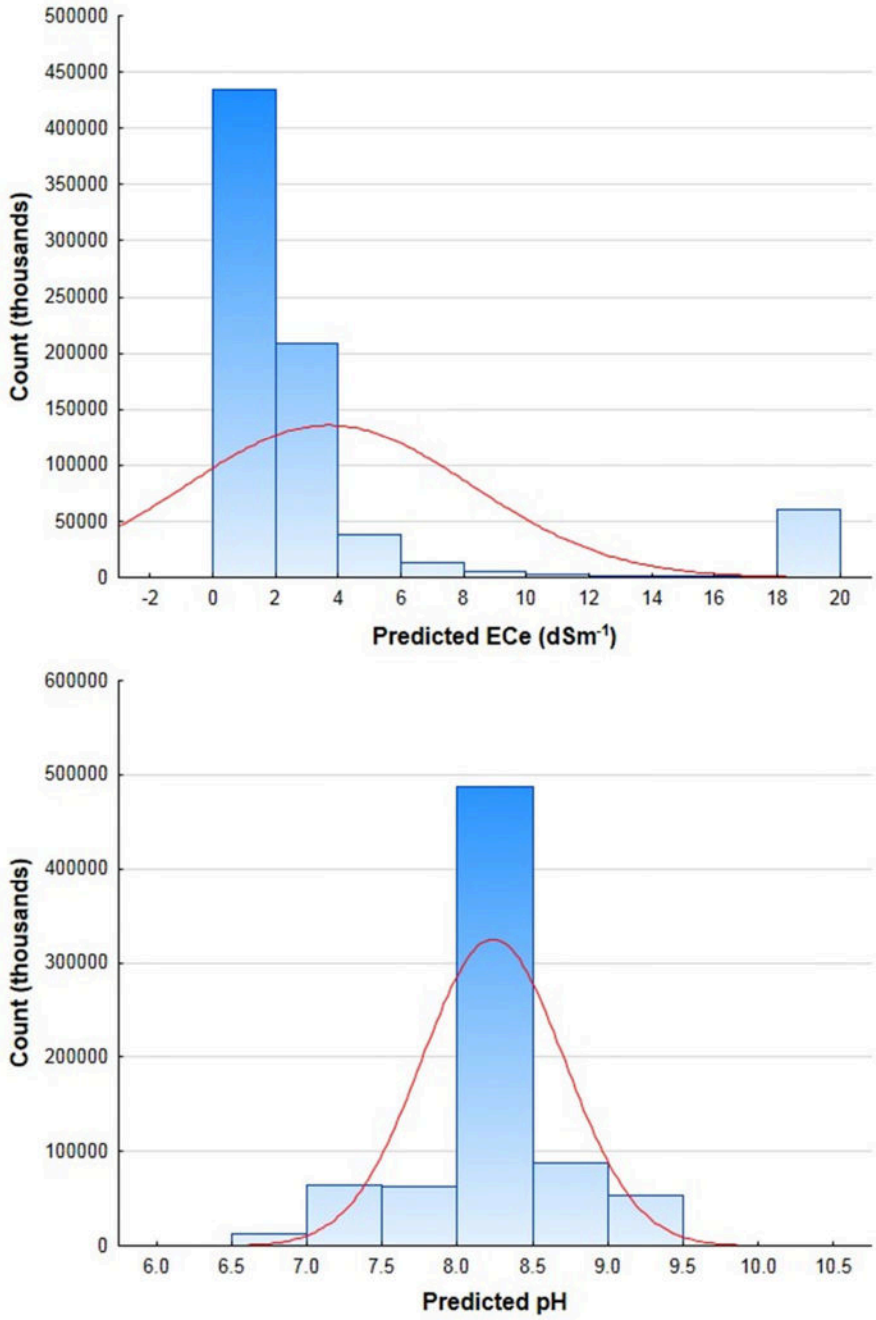

Figure 3 Distribution of predicted soil pH and ECe at 0–20 cm depth. Bars are predicted values; red line is a fitted normal distribution.

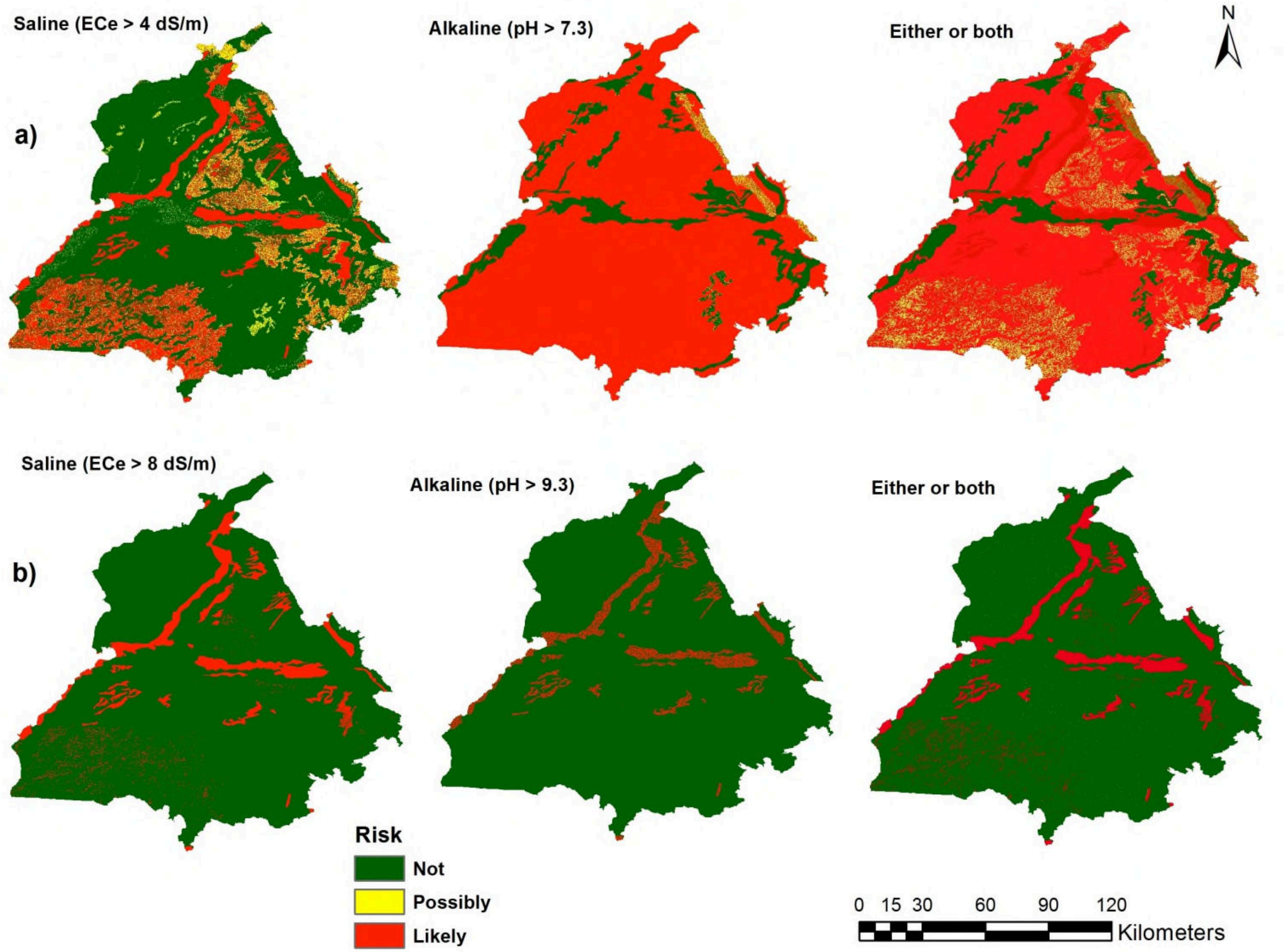

Figure 4 Soil constraints maps allowing for prediction uncertainties with two sets of constraint criteria: a) the FCC criteria and b) more stringent criteria that are more appropriate for lowland rice.