Geoderma Regional 22 (2020) e00289

Contents lists available at ScienceDirect

Geoderma Regional

journal homepage: www.elsevier.com/locate/geodrs

Predicting spatial variability of selected soil properties using digital soil mapping in a rainfed vineyard of central Chile

Lwando Mashalaba [a], Mauricio Galleguillos [b,c,*], Oscar Seguel [a], Javiera Poblete-Olivares [b]

[a] Universidad de Chile, Facultad de Ciencias Agronómicas, Departamento de Ingeniería y Suelos, Casilla 1004, Santiago de Chile, Chile
[b] Universidad de Chile, Facultad de Ciencias Agronómicas, Departamento de Ciencias Ambientales y Recursos Naturales Renovables, Chile
[c] Center for Climate and Resilience Research (CR)^2, Chile

ARTICLE INFO

Article history:
Received 14 January 2020
Received in revised form 25 April 2020
Accepted 28 April 2020

Keywords:
Digital soil mapping
Soil properties
Vineyard
Random Forest model
Environmental covariates
Remote sensing
Alfisols

ABSTRACT

Soil physical properties influence vineyard behavior, therefore the knowledge of their spatial variability is essential for making vineyard management decisions. This study aimed to model and map selected soil properties by means of knowledge-based digital soil mapping approach. We used a Random Forest (RF) algorithm to link environmental covariates derived from a LiDAR flight and satellite spectral information, describing soil forming factors and ten selected soil properties (particle size distribution, bulk density, dispersion ratio, Ksat, field capacity, permanent wilting point, fast drainage pores and slow drainage pores) at three depth intervals, namely 0–20, 20–40, and 40–60 cm at a systematic grid (60 × 60 m^2). The descriptive statistics showed low to very high variability within the field. RF model of particle size distribution, and bulk density performed well, although the models could not reliably predict saturated hydraulic conductivity. There was a better prediction performance (based on 34% model validation) in the upper depth intervals than the lower depth intervals (e.g., R^2 of 0.66; nRMSE of 27.5% for clay content at 0–20 cm and R^2 of 0.51; nRMSE of 16% at 40–60 cm). There was a better prediction performance in the lower depth intervals than the upper depth intervals (e.g., R^2 of 0.49; nRMSE of 23% for dispersion ratio at 0–20 cm and R^2 of 0.81; nRMSE of 30% at 40–60 cm). RF model overestimated areas with low values and underestimated areas with high values. Further analysis suggested that Topographic position Index, Topographic Wetness Index, aspect, slope length factor, modified catchment area, catchment slope, and longitudinal curvature were the dominant environmental covariates influencing prediction of soil properties.

1. Introduction

Soil mapping is regarded as key for guiding decision-makers in natural resource assessment, environmental modeling, and land use studies. However, it requires the knowledge of experienced pedologists for the stages of soil mapping (Kempen et al., 2012; Resende et al., 2014). Many environmental and agro-economic activities require accurate information about the spatial variability of soil types and soil properties. Therefore, this information is being generated increasingly through digital soil mapping (DSM) techniques (Minasny and McBratney, 2016). Digital soil mapping (DSM) aims at the creation of an population of a spatial soil information system by numerical models inferring the spatial and temporal variation of soil types and soil properties from soil observation and knowledge and from related environmental variables (Lagacherie and McBratney, 2007). DSM has been applied at local, regional (Heung et al., 2014; Lacoste et al., 2011), national or a global scale in mapping soil classes and properties (Arrouays et al., 2014). The conducted review on the use of DSM for soil mapping (Caten et al., 2012), indicated that approaches used up to 2011 show three main classification models applied for DSM (artificial neural networks, logistic regression, and decision tree). McBratney et al. (2003) stated that advances in remote sensing and information system have created the way for DSM, which couples soil point data with statistically correlated auxiliary data. This approach overcomes the imitations of the traditional mapping methods by reducing both the workload and costs involved (Giasson et al., 2015). However, it only allows the prediction of one variable at a time. The auxiliary data include the soil forming factors that were proposed by Jenny (1941), which was later modified by McBratney et al. (2003) into a now widely used Scorpan model in which is a set of soil properties at the same location (s), climate (c), organism (o), relief (r), parent material, age (a) and spatial location (n).

In DSM, this auxiliary data is mostly derived from digital elevation model (DEM) and available conventional maps. In addition, further advances are foreseen with the availability of satellite imagery data with high spatial, spectral and temporal resolution to improve mapping accuracy (Forkuor, 2014) at a given location in the landscape. The combination of the environmental covariates derived from DEM with radar imagery data has great potential for improving prediction accuracy for

* Corresponding author at: Universidad de Chile, Facultad de Ciencias Agronómicas, Departamento de Ciencias Ambientales y Recursos Naturales Renovables, Chile.
E-mail address: mgalleguillos@renare.uchile.cl (M. Galleguillos).

https://doi.org/10.1016/j.geodrs.2020.e00289
2352-0094/

a targeted soil properties. While DSM can be said to now be operational, there are still unresolved methodological issues regarding to the number of predictor variables that can be used to predict the property of interest. Soil properties vary spatially and temporally within a landscape and within the soil profile (Mulla and McBratney, 2002). Thus, soil properties determined for soil samples taken in close proximity are often highly correlated relative to samples taken further apart. Several studies reported that soil properties vary at different spatial scales (Outeiro et al., 2008; Goovaerts, 1998) primarily due to heterogeneity of internal factors and anthropogenic impacts generating complex spatial soil patterns (Kilic et al., 2012; Liu et al., 2009) and land use (Saglam and Dengiz, 2012). Spatial variability of soil properties results in the change in the values of certain soil properties over space and time (Ettema and Wardle, 2002). According to recent studies, understanding of spatial variability of soil physicochemical characteristics in both its static (e.g. texture and mineralogy) and dynamic (e.g. water content, compaction, organic matter, etc.) forms is necessary for site-specific management of agricultural practices, as it is directly contributing to variability in crop yield and quality (Jabro et al., 2010; Silva Cruz et al., 2011). Soil physical properties influence vineyard behavior, therefore the knowledge of their spatial variability is essential for making vineyard management decisions. Chile is the world's eighth-largest producer of wine and fifth-largest exporter (Felzensztein et al., 2011) and the Chilean wines are positioned as the countrys' most emblematic and best-known world ambassador. The Maule region of Southern Chile is one of the most wine producers in the country and most vineyards produce rainfed vine, which indicates that the distribution of water in the profile solely depends on soil properties and environmental variables such as precipitation and evapotranspiration.

Ubalde et al. (2007) stated that the aim of modern oenology is to produce wines of recognized quality and typicality, which can then be differentiated in a market with growing demand. Therefore to achieve all this, it is essential to consider that the potential quality of the wine is established in vineyards. However, one of the challenges facing the vineyards managers is how to manage the yield and quality variability of the vineyard to identify uniform batches of good-quality fruits (Bramley, 2005). The soil in vineyards is subject to frequent traffic associated with soil tillage, weed control or plant protection and harvesting of vines. Unamunzaga et al. (2014) noted that there is little work that has been conducted at high resolution on soil properties at depths lower than 0.30 m which are of special relevance to perennial crops. It is noted that wine quality is often strongly influenced by spatial variability of soil properties such as soil texture and soil depth due to their relationship with soil water holding capacity (WHC) because vine behavior is closely related to water uptake (Van Leeuwen et al., 2004). Therefore, the main objective of this study was to model and map the selected soil properties of the 29ha vineyard from knowledge-based digital soil mapping and the application of remote sensing data.

2. Material and methods

2.1. Site description

The study was carried out in a 29 ha vineyard situated 20 km away to the west from the Cauquenes city, in the Maule Region, central Chile (36° 02′ 27.18′′ SL, 72° 28′ 09.47′′ WL) at 202 m asl (Fig. 1).

The site is characterized by a sub-humid Mediterranean climate with winter rainfall. According to Uribe et al. (2012), the mean annual precipitation of the area is 690 mm, mainly concentrated in winter months (June–July). The temperature regimes are moderate with cold winters, with maximum average temperatures ranging between 14 and 29 °C and the minimum between 3 and 12°C. The total annual evapotranspiration of the area is 1128 mm, with minimum and maximum (40–162 mm) occurring in July and January, respectively. According to CIREN (Centro de Información de Recursos Naturales Chile) (1997), the soils are belonging to Cauquenes soil Associations and classified as Ultic Palexeralf, which corresponds to deep soils (≥100 cm depth) with increasing clay content in the deeper horizons and slopes gradient ranging from 1 to >30% in a hilly landscape. The soils are developed in situ from weathered granite, rich in quartz and feldspars. The vineyard correspond to "Pais" variety and is more than 50 years old with a plant density of approximately 10,000 plants per ha with a spacing of 1×1 m^2. The vines are produced under dryland condition, without irrigation.

2.2. Soil sampling and storage

To study the spatial and vertical variability of selected soil physical properties, the following methodology was performed. A systematic sampling grid (60×60 m) was used, and 62 soil pits were opened as close as possible to vine rows (Fig. 1). At each sampling point, three undisturbed core samples were vertically taken with soil core sampler from each sampling depth (0–20 cm, 20–40 cm and 40–60 cm) and 186 disturbed soil samples ($\pm$ 4 kg) were collected from each sampling point and depth for further analysis. The sampling density was two soil pits per hectare and all the sampling points were geo-referenced using global positioning system (GPS) receiver (accuracy of $\pm$4 m). Undisturbed samples were used for the determination of soil bulk density, the water retention curve at low suctions (>100 kPa) and Ksat. Disturbed soil samples were thoroughly air-dried, mixed and 300 g of representative soil was ground to pass 2 mm sieve and the amount that could not pass through was also recorded then was used for the determination of particle size distribution and microaggregate stability. None sieved soil samples were used for water retention relationship at higher suctions (>100 kPa) using a pressure plate apparatus.

2.3. Laboratory analysis

2.3.1. General soil physical properties

Under laboratory conditions, disturbed soil samples (2 mm sieved, 186 samples in total) were used to particle size distribution, according to methodologies detailed in Sandoval et al. (2012). To quantify the porous system, bulk density (BD) was measured using the cylinder method (Grossman and Reinsch, 2002).

2.3.2. Soil water retention curves and pore size distribution

To prepare the soil samples for water retention measurements, the first step was to saturate the soil cores using a capillary rise saturation method. Thereafter, the soil water retention curve was done using a sandbox (Eijkelkamp) and pressure plate apparatus (Soil Moisture equipment) according to the method described Sandoval et al. (2012) applying increasing pressures to saturated samples (0.2, 6, 33, 500 and 1500 kPa). The volumetric water content at field capacity (FC) and the permanent wilting point (PWP) were considered when the equilibrium was reached at −33 and −1500 kPa, respectively. The pore size distribution was derived from the water retention curves data for each sampling point according to Hartge and Horn (2009), calculating the fast drainage pores (FDP > 50 μm) as the difference between water content at metric equilibrium of −0.2 and the equilibrium at −6 kPa; the slow drainage pores (SDP 10–50 μm) were calculated based on the difference between water content at −6 and −33 kPa.

2.3.3. Micro aggregate stability

The stability of the microaggregates (1–2 mm) was determined by the dispersion ratio (DR) method according to Seguel et al. (2003). Two soil samples of 50 g with aggregates between 1 and 2 mm in diameter were obtained by sieving air-dried soil; one of the subsamples was subjected to a slight dispersion in 150 cm^3 distilled water, while the other sample was subjected to a drastic dispersion with the same amount of distilled water and 25 cm^3 of sodium pyrophosphate. Both samples were left overnight. The drastically dispersed sample was then mechanically shaken for 5 min in a 75 cycle Hamilton Beach blender. Finally, samples were poured into 1000 mL measuring

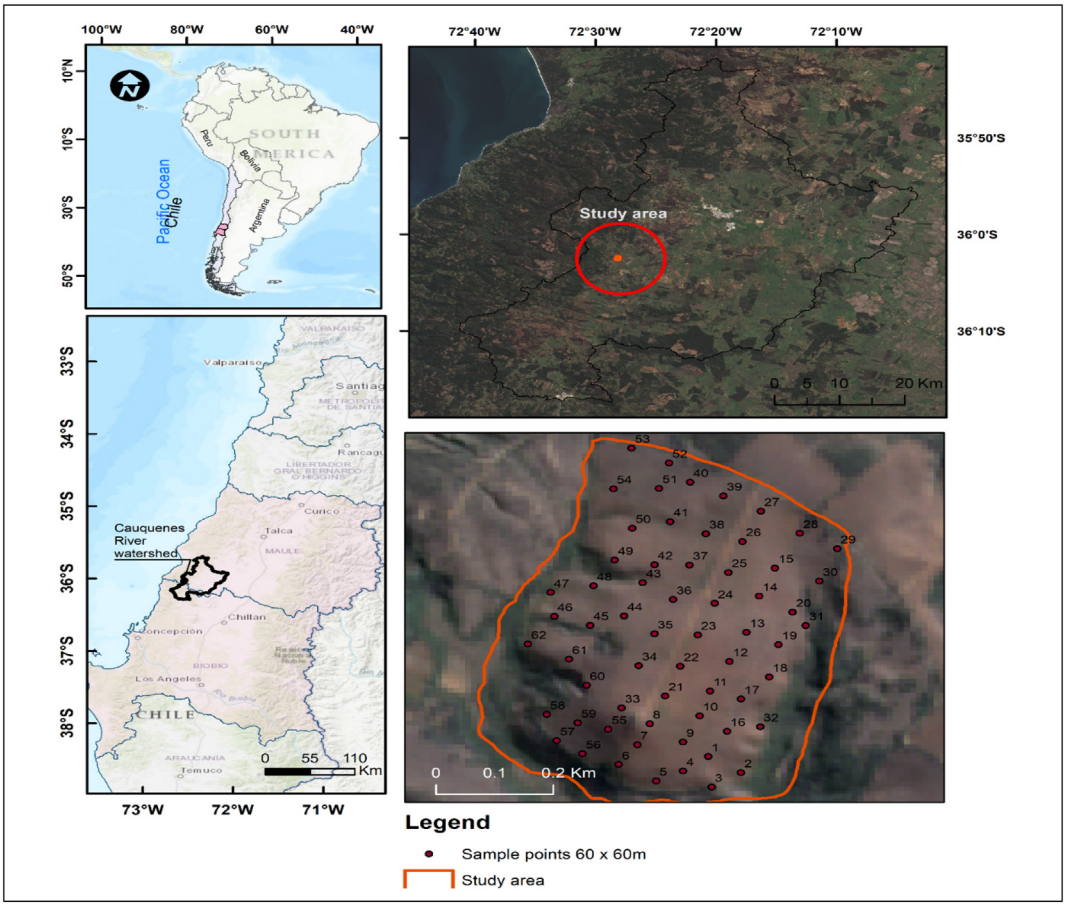

Fig. 1. Location of the study area located in the Maule region, Cauquenes, Chile central with 62 sampling points and a distance of 60 m × 60 m.

cylinders filled up to 1000 cm^3 with distilled water. The density of the suspension was measured with a hydrometer along with the temperature 40 s after the start of the decanting process. The clay and silt content of both samples was calculated by the Bouyoucos hydrometer method based on Stoke's law (Dane and Topp, 2002) and the DR was calculated using the equation:

$$DR = \frac{(S+C)Sd}{(S+C)dd} \cdot 100 \quad (1)$$

where *DR* is the dispersion ratio, *(S + C) Sd* is the percentage of clay + silt content of slightly dispersed samples (without sodium pyrophosphate and no mechanical agitation) and *(S + C)dd* is the percentage of total clay and silt (with drastic dispersion). The lower values of *DR* indicates the highest stability.

2.3.4. Saturated hydraulic conductivity

Saturated hydraulic conductivity, Ksat (LT^{-1}), values of the soils were measured for all depths (0–20, 20–40 and 40–60 cm) in the laboratory using the constant head soil core method (Eijkelkamp) (SSSA, 2008). The samples were fully saturated using capillarity method then water was allowed to flow through the soil at a steady rate under a constant hydraulic head gradient. The measurements of volume were taken after 1 h and 5 h, respectively for all the samples. The Ksat processed using Darcy's formula of water flow:

$$Ksat = \frac{4\,VL}{\pi \cdot d_c^2 \cdot \Delta t \cdot \Delta H} \quad (2)$$

where V (L^3) is the volume of water collected during time interval Δt, *L* is the length of soil sample in the core, ΔH *(L)* is the difference in elevation between the water level in the reference tube and water level in the side arm of the outflow dripper and d_c is the inside diameter of the core sample.

2.4. Digital elevation model and remotely sensed imagery attributes

The LiDAR point cloud was acquired in 2009 with an average point cloud density of 4.64 points m^{-2} and a spatial resolution of 5 m. The digital crown height model (DCM) was calculated by the difference between the DSM and the DTM. Then the 5 m × 5 m digital elevation model (DEM) and Sentinel-2A optical images from November 2018 (available from the European Space Agency, ESA), were used to calculate topographic attributes and vegetation indices in SAGA GIS 1.3 (Conrad, 2014) and with R, respectively.

Terrain attributes were calculated, including slope, aspect, plan curvature, profile curvature, flow line curvature, elevation (DEM), catchment area, wetness index, stream power index, length slope factor (LS Factor), flow accumulation, convergence index, curvature classification, Multiresolution index of valley bottom flatness (MRVBF), multiresolution index of ridge top flatness (MRRTF), topographic position index (TPI), analytical Hillshading, and topographic wetness index (TWI). These terrain attributes give information about the depositions, water flow and accumulations. For example curvature, flow accumulation controls the movement of water and fine materials within the field. Topographic position index is the difference between the elevation at a cell and the average elevation in a neighbourhood surrounding that cell. Positive values means that the cell is higher than its neighbours (indicate ridges and hills) while negative values means the cell is lower (indicates valleys). Atmospherically corrected imagery, were extracted from Copernicus data hub and used to calculate the spectral indices

such as, Green-Red Vegetation Index (GRVI), and three-band spectral index (TBSI-V, TBSI-T and TBSI-W) computed according to formulations of Verrelst et al. (2015), Tian et al. (2014) and Wang et al. (2012), respectively. This data was obtained for 62 georeferenced pixels where the measurements of soil physical properties were carried out and samples were collected. Preliminary processing was carried out on these records to remove outliers due to the presence of clouds. Only images without the presence of clouds were used in the analysis. The technical aspect can be obtained from Van derWerff and Van der Meer (2016).

2.5. Selection of environmental covariates

Due to the large amount of data available for use as covariates in modeling, it was necessary to use a data-mining technique to select the most appropriate dataset as an optimal set of predictors to run the model, affording the lowest error. When selecting some environmental covariates, it is advisable to select the most appropriate variables that are involved in the processes that determine soil properties. The process of selecting appropriate predictors was done using a methods of variable selection, which consist of selecting a subset of characteristics from a set of complete data, maintaining high precision according to Ladha and Deepa (2011). The method was the Recursive Feature Elimination (RFE) algorithm which is an example of "backward" feature deletion, in an iterative process according to Guyon et al. (2002). The variables that were eliminated are the ones that had a negative influence on the model and are not included in final models. RFE calculates a measure of model performance, which in this case was the root of the mean square error (RMSE) according to Cabezas and Galleguillos (2016). Then, collinearity was checked among predictors following Lopatin et al. (2016) methodology.

2.6. Statistical analysis

2.6.1. Descriptive statistics

Statistical parameters which are generally regarded as indicators of the central tendency and spread of the data were analysed. They included the determination of mean, minimum, maximum values, standard deviation, range, coefficient of variation (CV), kurtosis and skewness. The normal frequency distribution was decided through the evaluation of skewness according to Paz-Gonzalez et al. (2000). The data were analysed using the SPSS version 25.0 software. The CV was used to assess the variability of the different data set. Normality test was carried out using the Quantile-quantile plots, Shapiro-Wilk, Kolmogorov-Smirmov, histogram plot and therefore the data that did not follow the normal distribution was log transformed to stabilize the variance. The normality tests were recalculated using the log-transformed data, as asymmetry in the distribution of data has an important effect on the variability analyses. The correlation between soil physical properties and environmental variables was tested using the Pearson correlation coefficient accepting a confidence level of 95%.

2.6.2. Statistical modeling

After the environmental covariates were selected for each of the selected soil properties, the predictive models were generated. The non-parametric Random Forest (RF) model proposed by Breiman (2001) was chosen. RF uses a collection of decision tree classifiers, where each forest tree has been trained using a bootstrap sample of individuals from the data, and each division attribute in the tree is chosen from the random subset of attributes (Reif et al., 2006). The RF model is insensitive to noise and allows the incorporation of many environmental covariates without a problem, and that make RF more useful compared to other prediction models. For this study, we applied the parameterization describes in Castillo-Riffart et al. (2017). All the procedures were carried out using the packages "Random Forest" (Liaw and Wiener, 2002), "caret" (Kuhn, M (Contributions from Jed Wing), et al., 2017) of the R-project software.

2.6.3. Random forest model performance

For validation of purposes, the best RF models were embedded in bootstrap with 500 iterations. To evaluate the prediction performance by the soil properties models, the data set was divided into two separate subsets by a random selection procedure using the "sample" function in R package before modeling. In using the sample function, approximately 66% of the total population ($n = 40.3$) were earmarked for model calibration while the remaining 34% of the population ($n = 20.7$) was used for model testing. The following two parameters were computed on the validation subset, using the R statistical software package (R Development Core Team, 2013).

$$\text{nRMSE} = [\text{RMSE}/[\max\ (\text{number of attributes}) - \min\ (\text{number of attributes})]] \times 100$$

where RMSE was calculated as;

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}\right)^2} \quad (3)$$

where y denotes reference parameter values, $\hat{y}$ predicted value, and n sample size.

$$R^2 = \frac{\sum_{j=1}^{n}\left(P_i - \overline{\theta_i}\right)^2}{\sum_{j=1}^{n}\left(O_i - \overline{\theta_i}\right)^2} \quad (4)$$

where n denotes data point, O_i and $P_i$are observed and predicted soil selected properties values at the i^{th} point and θ_i are the respective means. High values of R^2 and low values of nRMSE indicate high model quality.

2.6.4. Predictive maps of selected soil properties

The best models generated for each of the soil spatial variables. For this, the maps were obtained as proposed by Castillo-Riffart et al. (2017), where maps were calculated based on 500 iterations of Bootstrap, using the best performed models selected in each iteration. The environmental covariates used for the generation of the models are spatially explicit, all correspond to "rasters", which allows extrapolation to areas that were not covered by the pixels used for the construction of the models, which correspond to the sampling plots (Castillo-Riffart et al., 2017;). Finally, coefficient of variation values maps for selected soil properties predictions, as obtained from the 500 bootstrap iterations were produced for the entire study field.

3. Results

3.1. Descriptive statistics and Pearson correlation

The descriptive statistics calculated for each of the soil properties studied and for all the sampling depths (0–20, 20–40 and 40–60 cm) are presented in Table 1. The soil texture of the study area varied from silt loam to silty clay loam. The values of bulk density for surface soil (0–20 cm) ranged between 1.17 Mg m^{-3} and 1.67 Mg m^{-3}, with an average BD of 1.48 Mg m^{-3}, whereas BD for subsoil (40–60 cm) varied between 1.46 Mg m^{-3} and 1.77 Mg m^{-3}, with a mean of 1.68 Mg m^{-3}. Surface soil microaggregate stability (DR) values varied between 21% and 76%, with an average of 43.1%, and DR values for the subsoil ranged from 13% and 85%, with an average of 28%. The values of Ksat for surface soil ranged between 0.0 cm h^{-1} and 290 cm h^{-1}, with a mean of 50.6 cm h^{-1}, whereas Ksat for lower depth varied between 0.0 cm h^{-1} and 330 cm h^{-1}, with a mean of 18.4 cm h^{-1}. In addition, the values of surface volumetric water content at FC and PWP values varied between 14% and 29%, 6.0% and 20%, respectively, with the means of 20% and 12%, respectively, whereas FC and PWP for the lower soil depth varied between 17% and 47%, 11% and 29%, with the means of 28% and 20%, respectively. The values of surface soil FDP and SDP varied between 2%

Table 1
Descriptive statistics for selected soil properties at 0–20, 20–40 and 40–60 cm soil depths.

Variables	Min	Max	Mean	SD	CV	Median	Range	Skewness	Kurtosis
0–20 cm soil depth									
Clay (%)	6.8	32.7	14.9	5.3	35.1	13.1	26.0	1.17	1.07
Silt (%)	9.4	29.6	20.4	3.5	17.4	20.4	20.2	- 0.38	1.30
Sand (%)	48.4	78.9	64.7	6.4	9.9	65.9	30.5	−0.38	0.03
BD (Mg m^{-3})	1.17	1.67	1.48	0.12	7.9	1.48	0.5	−0.41	−0.16
DR (%)	21.2	75.5	43.1	11.3	26.3	42.3	54.3	0.43	0.12
Ksat (cm h^{-1})	0.0	290.1	50.6	52.8	104.4	32.2	290.1	2.27	6.31
FC (%)	14.0	29.0	20.0	4.0	17.8	20.0	15.0	0.58	0.08
PWP (%)	6.0	20.0	12.0	3.0	23.3	12.0	14.0	0.53	0.02
FDP (%)	2.0	25.0	15.0	4.0	27.9	15.0	23.0	−0.03	0.60
SDP (%)	2.0	8.0	5.0	1.0	26.6	5.0	6.0	0.30	0.63
20–40 cm soil depth									
Clay (%)	9.7	45.1	29.4	8.5	28.9	30.7	35.4	−0.14	−0.64
Silt (%)	13.5	26.3	18.7	2.9	15.5	18.8	12.8	0.33	0.01
BD (Mg m^{-3})	1.43	1.73	1.59	0.07	4.40	1.60	0.30	−0.28	−0.54
DR (%)	11.7	76.4	30.0	13.1	47.5	23.4	64.7	1.52	2.14
Ksat (cm h^{-1})	0.0	186.3	22.6	36.3	197.3	2.5	186.3	2.77	7.45
FC (%)	15.0	36.0	25.0	4.0	17.2	26.0	21.0	−0.34	0.38
PWP (%)	9.0	30.0	19.0	4.0	20.8	19.0	21.0	−0.14	0.34
FDP (%)	4.0	20.0	9.0	3.0	34.1	9.0	16.0	0.97	0.92
SDP (%)	1.0	7.0	3.0	1.0	35.8	3.0	6.0	0.85	1.12
40–60 cm soil depth									
Clay (%)	15.6	46.3	34.3	7.3	21.2	36.2	30.7	−0.62	−0.34
Silt (%)	3.3	26.2	16.5	3.7	22.1	16.0	22.9	−0.17	2.62
Sand (%)	35.2	68.4	49.2	7.2	14.7	48.1	33.3	0.46	−0.41
BD (Mg m^{-3})	1.46	1.77	1.68	0.07	4.47	1.63	0.31	−0.07	−0.54
DR (%)	13.0	85.4	27.7	14.8	49.3	25.6	72.4	1.60	2.42
Ksat (cm h^{-1})	0.0	330.7	18.4	52.0	229.9	2.5	330.7	4.04	18.62
FC (%)	17.0	47.0	28.0	5.0	16.7	28.0	30.0	0.93	3.12
PWP (%)	11.0	29.0	20.0	4.0	17.8	21.0	18.0	−0.24	0.40
FDP (%)	4.0	16.0	8.0	3.0	33.0	8.0	12.0	0.47	−0.39
SDP (%)	1.0	7.0	3.0	1.0	36.7	3.0	6.0	1.29	2.05

N = 62, SD = standard deviation, CV = coefficient of variation (%), BD = bulk density, DR = dispersion ratio, Ksat = saturated hydraulic conductivity, FC = field capacity, PWP = permanent wilting point, FDP = fast drainage pores, SDP = slow drainage pores.

and 25%, 2% and 8%, respectively, with the means of 15.0% and 5%, respectively. Whereas the lower soil depth FDP and SDP varied between 4% and 16.0%, 1.0% and 7.0%, respectively, with the means of 8.0% and 3.0%, respectively. Generally, the mean values of clay, BD, FC and PWP increased with the increasing soil depth (Table 1), whereas the mean values of sand content decreased with soil depth.

The coefficient of variation values (CV) was used to interpret the variability in soil properties. The criteria proposed by Gomes and Garcia (2002) was used to classify the soil properties into low (<10%), medium (10%–20%), high (20%–30%) and very high (>30%) variability. The lowest and the highest CVs obtained for surface soil were 7.9% for BD and 104.4% for Ksat, respectively, and lowest and highest CVs for the lower soil depth (40–60 cm) were 4.47% for BD, and 230% for Ksat. Based on CVs, except for BD in all the depths and sand content in surface soil, both surface and subsoil had high to very high spatial variability. Zhou et al. (2010) noted that the CV is the most important factor in defining the variability of soil properties than other parameter such as SD, mean and median (Xing-Yi et al., 2007). In addition, variability of clay, silt, BD, FC and PWP decreased with soil depth (Table 1), whereas variability of sand, DR, Ksat, FDP and SDP increased with the soil depth. Logarithmic transformation was used for the data with a skewness above 1.0. In spite of skewness and kurtosis of the distribution of the soil properties, the mean and median values were similar with means equal to or almost equal to the median.

Pearson correlations between all analysed parameters for soil properties of the entire data set of 62 sampling points for all the sampling depths (0–20 cm, 20–40 cm and 40–60 cm) are shown in Table 2.

In the surface soil (0–20 cm), the clay content was significantly ($p < .01$) positively correlated with properties such as FC ($r = 0.49$), PWP ($r = 0.74$), but negatively with sand ($r = -0.83$), FDP ($r = 0.27$) and SDP ($r = -0.27$) (Table 2). The volumetric water content at FC was significantly ($p < .01$) positively correlated with properties such as clay (r = 0.49), and silt ($r = 0.34$), but negatively with sand ($r = -0.59$) and Ksat ($r = -0.44$) (Table 2). However, there was no significant correlation between clay content and aggregate stability test (DR), and Ksat. Correlation results between the properties of sub-soils (20–40 cm) depth shows that BD was significantly ($p < .05$) positively correlated with properties such as DR ($r = 0.26$), but significantly ($p < .01$) negatively correlated with Ksat ($r = -0.40$) and FDP ($r = -0.32$) (Table 2). Nevertheless, DR was significantly positively correlated with silt ($r = 0.36$), BD (r = 0.26) and significantly negatively correlated with clay ($r = -0.25$). On the other hand, Ksat at the depth of 20–40 cm was significantly positively correlated with sand (r = 0.27) and significantly negatively correlated with BD (r = −0.40), FC ($r = -0.50$), PWP ($r = -0.39$) and FDP ($r = -0.30$). Lastly, the Pearson correlation results of sub-soils (40–60 cm) depth shows that Ksat was significantly ($p < .01$) negatively correlated with clay content ($r = -0.49$) and significantly positive correlated with sand and clay content ($r = 0.31$ and r $p = 0.34$, respectively), while DR was significant positively correlated with silt (r = 0.31), sand (r = 0.34) and significant negatively correlated with clay (r = −0.49) see Table 2.

3.2. Statistical models

3.2.1. Particle size distribution

Based on the averaged importance measured across 500 runs of the Random Forest model, the environmental variables that had the greatest influence on the model error rate for soil surface clay, sand and silt content, respectively were: elevation, topographic wetness index; catchment slope, slope; convexity and green-red vegetation index, respectively (Fig. 2.1 to 2.5). The environmental variables that

Table 2
The Pearson correlation relationships between variables at 0–20 to 40–60 cm soil depth.

Variables	Clay	Silt	Sand	BD	DR	Ksat	FC	PWP	FDP	SDP
0–20 cm soil depths										
Clay	1									
Silt	0.03	1								
Sand	−0.83**	−0.58**	1							
BD	0.08	0.05	−0.09	1						
DR	−0.16	0.31*	−0.04	0.25*	1					
Ksat	−0.10	−0.47**	0.34**	−0.48**	−0.24	1				
FC	0.49**	0.34**	−0.59**	0.14	0.05	−0.44**	1			
PWP	0.74**	0.28*	−0.76**	0.18	−0.04	−0.43**	0.77**	1		
FDP	−0.27*	−0.01	0.23	−0.70**	−0.14	0.45**	−0.44**	−0.40**	1	
SDP	−0.27*	0.32*	0.04	−0.17	0.14	−0.12	−0.04	−0.06	0.43**	1
20–40 cm soil depth										
Clay	1									
Silt	−0.27*	1								
Sand	−0.94**	−0.07	1							
BD	−0.01	−0.02	0.02	1						
DR	−0.29*	0.36**	0.17	0.26*	1					
Ksat	−0.24	−0.04	0.27*	−0.40**	0.12	1				
FC	0.66**	0.01	−0.69**	0.18	−0.06	−0.50**	1			
PWP	0.73**	−0.09	−0.73**	0.15	−0.16	−0.39**	0.80**	1		
FDP	−0.40**	0.10	0.38**	−0.36**	0.08	0.37**	−0.75**	−0.55**	1	
SDP	−0.51**	0.17	0.47**	−0.09	0.08	0.01	−0.30*	−0.36**	0.14	1
40–60 cm soil depth										
Clay	1									
Silt	−0.26*	1								
Sand	−0.87**	−0.24	1							
BD	−0.26*	−0.04	0.28*	1						
DR	−0.49**	0.31*	0.34**	0.15	1					
Ksat	−0.02	−0.002	0.02	−0.16	−0.04	1				
FC	0.38**	−0.05	−0.36**	−0.02	0.17	−0.09	1			
PWP	0.61**	−0.05	−0.59**	−0.09	−0.11	−0.16	0.69**	1		
FDP	−0.24	0.00	0.24	−0.15	0.09	−0.06	−0.44**	−0.49**	1	
SDP	−0.42**	0.20	0.32*	−0.30*	0.07	−0.04	−0.39**	−0.42**	0.46**	1

*N = 62, *, ** Significant at p < .05 and p < .01, respectively.*

had the greatest influence on subsurface particle size distribution model error rate for final predictive models, averaged across 500 runs, were modified catchment area, longitudinal curvature; modified catchment area, topographic position index; and SAGA TWI and green-red vegetation index, respectively (Fig. 2.1 to 2.5). The average error rate for the final models used to generate the predictive map of clay, sand and silt content for all the sampled depth was ranging between 16 and 28%, 8.6–13% and 14–19%, respectively (Fig. 3b and Supplementary Material I). The particle size distribution results also showed that the combination of the various environmental variables can account for R^2 of 0.51 to 0.66, 0.48 to 0.52, and 0.42 to 0.50 of the variation in clay, sand, and silt, respectively (Fig. 3a and Supplementary Material I).

The variation in soil texture shows a progressive transition from the coarse texture (sand) along the fringes of the northern part of the vineyard to finer texture towards the Southern part. When looking at the predictive maps for surface clay content, there were few areas where more than 25% of clay content was predicted (Fig. 4.1). The areas that were predicted to have a greater amount of clay content were predominantly in basins and the low lying areas in the Southern and Southern Western part of the vineyard, and along the footslope of hills. In addition, when looking at the predictive map for subsoil (40–60 cm) clay content, there was a large area were more than or equal to 31% clay was predicted (Fig. 4.3) and the map showed moderate overall variation with the majority of the area being classified as having 31% to having between 37% and 42% clay content (Fig. 4.3). On the other hand, the obtained sand predictive maps show that the RF model predicted the following minimum and maximum values for surface (55–78%), and subsoil (37–60%) (Supplementary Material II). According to the surface sand predictive map, the sand content is predicted to be high in upper elevation compared to lower lying elevations (Supplementary Material II). The predictive maps showed that there were larger areas where more than 16, 15 and 9% silt content was predicted for all the three depths, respectively (Supplementary Material II). The higher values of the silt content are predicted to be along the fringes of the Western part of the vineyard. The maps showed moderate to high variability at all depths. In addition, the coefficient of variation maps obtained from 500 bootstrap prediction of the RF model (Fig. 5 and Supplementary Material III) showed that sand content values were predicted with low variation (0–12%), while the prediction for clay and silt content showed moderate variation. There was a tendency towards higher variation in low elevation areas compared to areas of high elevation (Fig. 5 and Supplementary Material III).

3.2.2. Bulk density

Based on the averaged importance measured across 500 runs of the Random Forest model, the environmental variables that had the greatest influence on the model error rate for soil bulk density for 0–20, 20–40 and 40–60 cm, respectively were: aspect, SAGA TWI; longitudinal curvature, Ls factor; analytical Hillshading and TBSI-W, respectively (Fig. 2.1 to 2.5). The average error rate (nRMSE) for the final models used to generate the predictive map of bulk density for all the sampled depth was ranging between 4 and 7% (Fig. 3b and Supplementary Material I). The bulk density results also showed that the combination of the various environmental variables can account for R^2 of 0.26 to 0.51 of the variation in bulk density (Fig. 3a and Supplementary Material I). The model predicted the minimum and maximum bulk densities for all the three sampling depths; 1.27–1.65 Mg m^{-3}, 1.40–1.70 Mg m^{-3}, and 1.53–1.74 Mg m^{-3}, respectively (Fig. 4.4 to 4.6). The predictive map for surface bulk density showed little overall spatial variation, with the majority of the area being classified as having less than or equal to 1.46 Mg m^{-3} bulk density. The areas that were predicted to have greater values of bulk density were predominantly in the

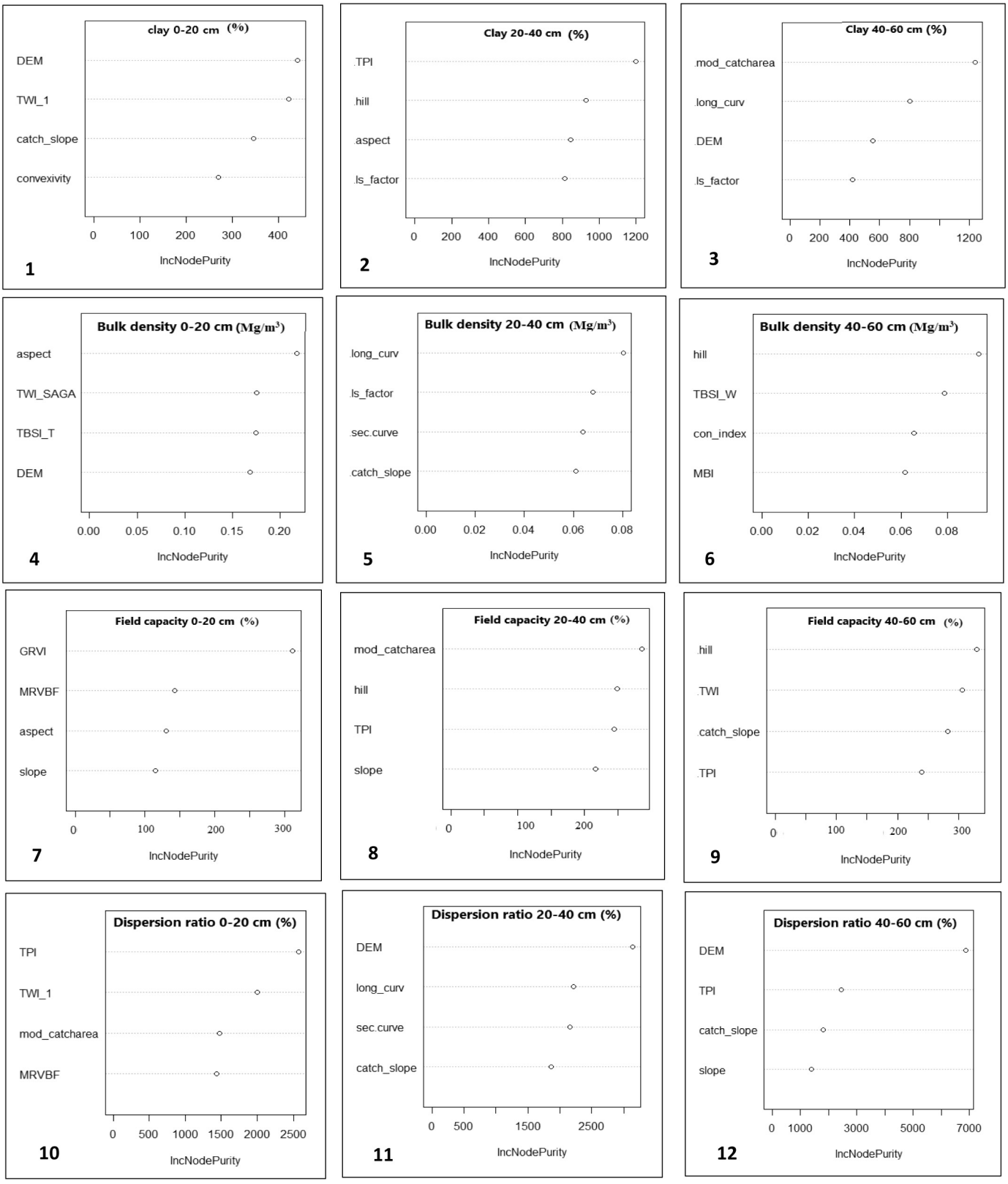

Fig. 2. Illustration of variable importance derived from Random Forest model for prediction of soil properties within the vineyard.

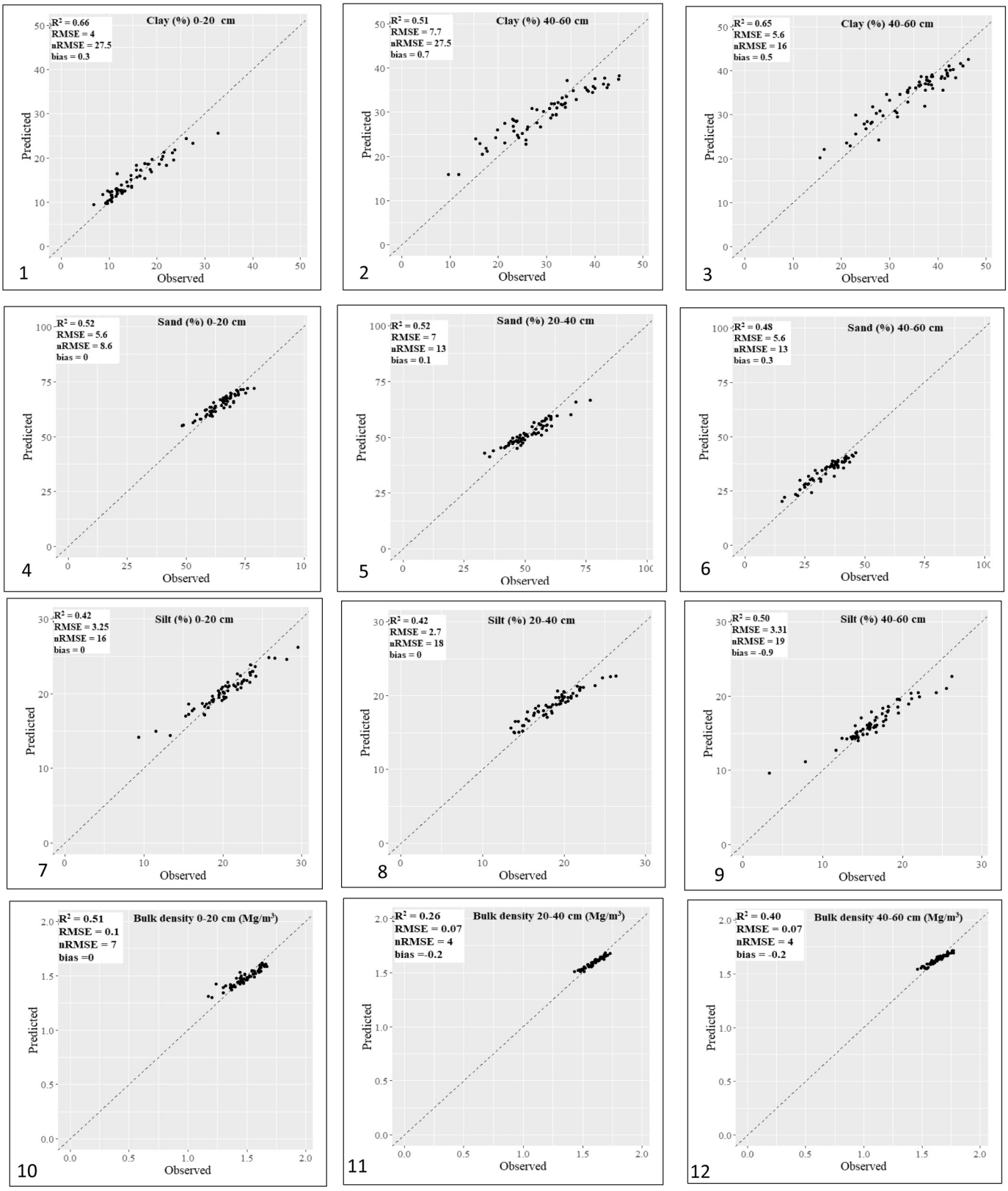

Fig. 3. Performance of random forest model (RFM) in modeling selected soil properties.

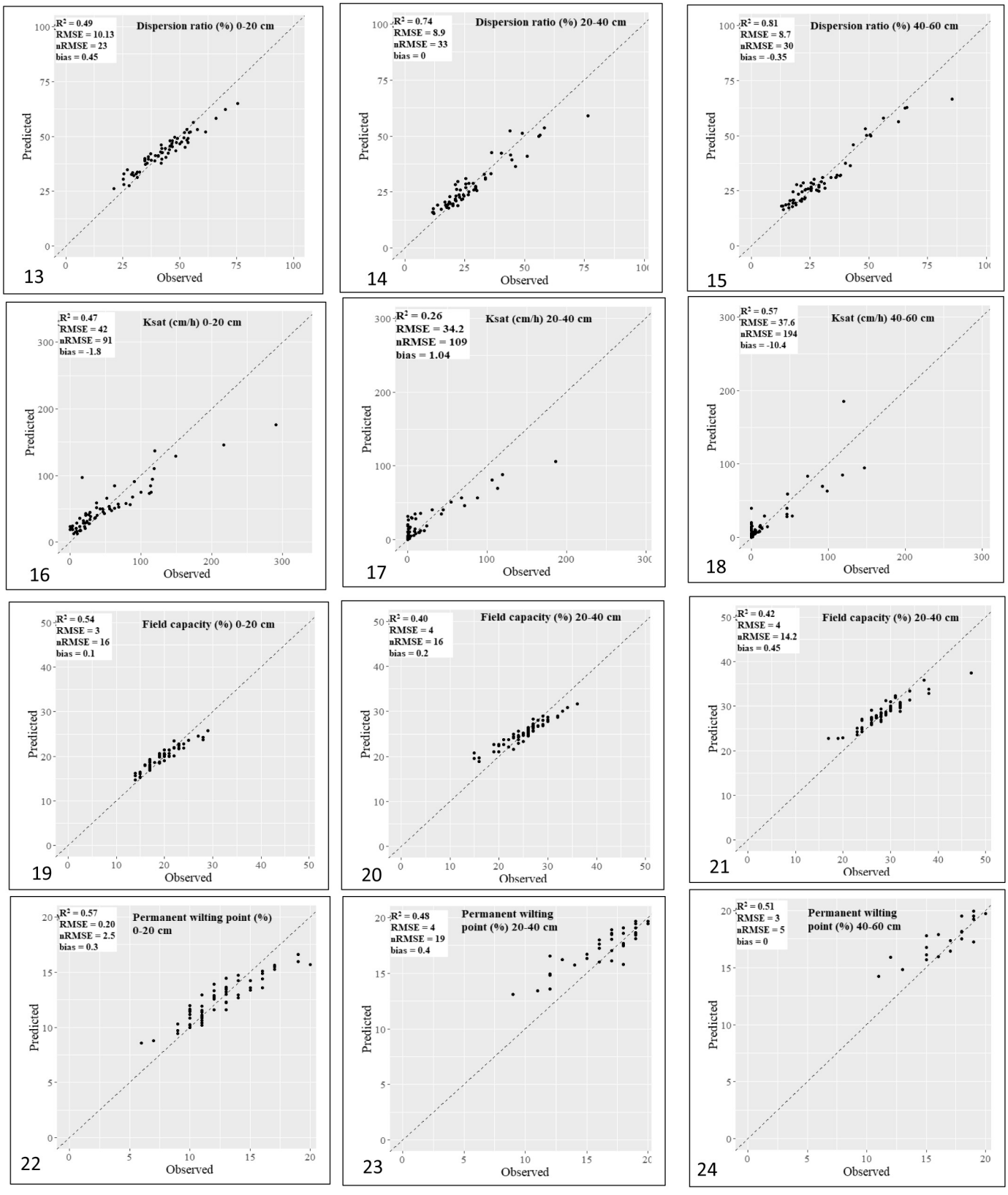

Fig. 3 (*continued*).

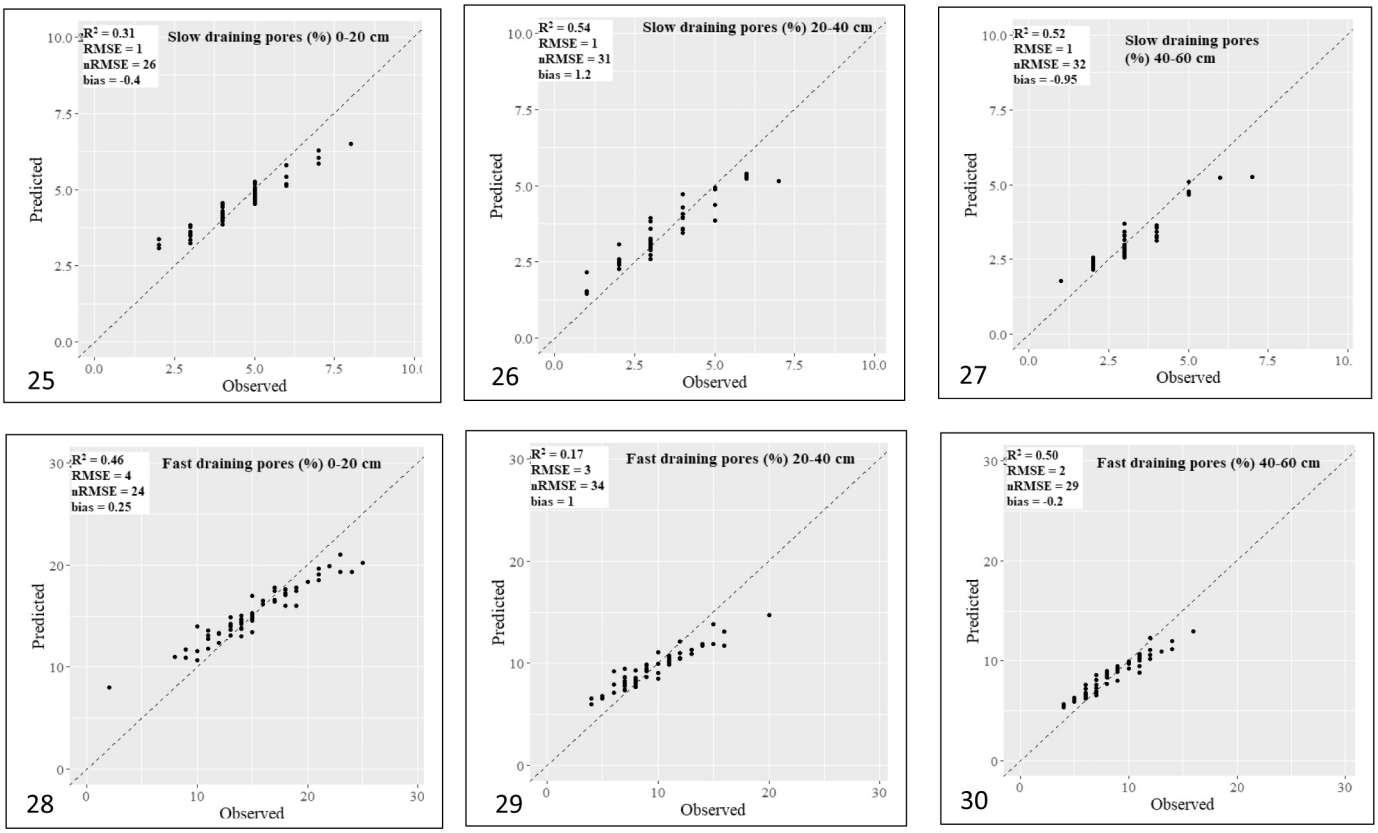

Fig. 3 (*continued*).

Western part of the field. The predictive map of the bulk density at the lower soil depths (40–60 cm) showed little variation, the majority of the area is being classified as having less than or equal to 1.63 Mg m^{-3} (Fig. 4.6). The higher predicted values of bulk density are at the flat slope and the lowest values are at the east towards the Southern part of the vineyard. The CV map showed that bulk density was predicted with low variation (0–7%).

3.2.3. Dispersion ratio

The averaged variable importance from 500 runs of the Random Forest model for soil surface, subsurface and lower depth dispersion ratio showed that topographic position index, topographic wetness index; elevation, longitudinal curvature and elevation, topographic position index, respectively were the important variables (Fig. 2.1 to 2.5). The error rate for the Random Forest model predicting for all the sampling depths (0–20, 20–40 and 40–60 cm) was ranging between 23 and 33% and bias −0.5-0 (Fig. 3 and Supplementary Material I). In addition, the dispersion ratio results also showed that the combination of the various environmental variables can account for R^2 of 0.49 to 0.81 of the variation in dispersion ratio (Fig. 3). The minimum and maximum dispersion ratio values predicted by the RF model at all the depths; 26.2–62.3%, 13.6–64%, and 16.4–61.9%, respectively (Fig. 3 and Supplementary Material I). The predictive map of soil surface dispersion ratio showed high overall variation, with the majority of the area being classified as having greater 44.2% (Fig. 4.10). The areas that were predicted to have a greater amount of dispersion ratio were predominantly in edges of the Western and Eastern part of the field (Fig. 4.11). However, the lower predicted dispersion ratio values are distributed within the field. The predictive map for subsurface dispersion ratio again showed high overall variation, with the majority of the area being classified as having less than 26.2% (Fig. 4.12). The highest and lowest dispersion ratio predicted values are found in the Western and Eastern part of the field, respectively (Fig. 4.12).

3.2.4. Saturated hydraulic conductivity

The averaged variable importance from 500 runs of the Random Forest model for soil surface, subsurface and lower depth Saturated hydraulic conductivity showed that longitudinal curvature, TBSI-T; topographic position index, TWI-SAGA and flow accumulation, catchment area, respectively played an important role in predicting the Ksat (Fig. 2.1 to 2.5). The error rate for the Random Forest model predicting for all the sampling depths (0–20, 20–40 and 40–60 cm) was ranging between 91 and 194% and −10.5–1.04 (Fig. 3b and Supplementary Material I). In addition, the Ksat results also showed that the combination of the various environmental variables can account for R^2 of 0.26 to 0.57 of the variation in Ksat (Fig. 3a and Supplementary Material I). The RF model predicted the minimum and maximum saturated hydraulic conductivity values for all the three sampling depths; 9.2–144 cm hr^{-1}, 0.79–149 cm $hr^{-1}$and 0.27–88.6 cm hr^{-1}, respectively (Supplementary Material II). All of the predictive maps showed strong variation, with the majority of the areas being classified as having Ksat of 76.7, 37.8 and 22.4 cm hr^{-1}, respectively (Supplementary Material II). The predictive map of the surface Ksat showed that the high predicted values were scattered within the field. While on the other hand, the predictive map of the subsurface Ksat showed that greater values were predicted on the edges of the West facing slope (Supplementary Material II). The CV maps obtained from RF model showed that Ksat values were predicted with high variation more especially in the lower soil depth (Supplementary Material III). These results correspond with the error rate (nRMSE) which were the highest for Ksat.

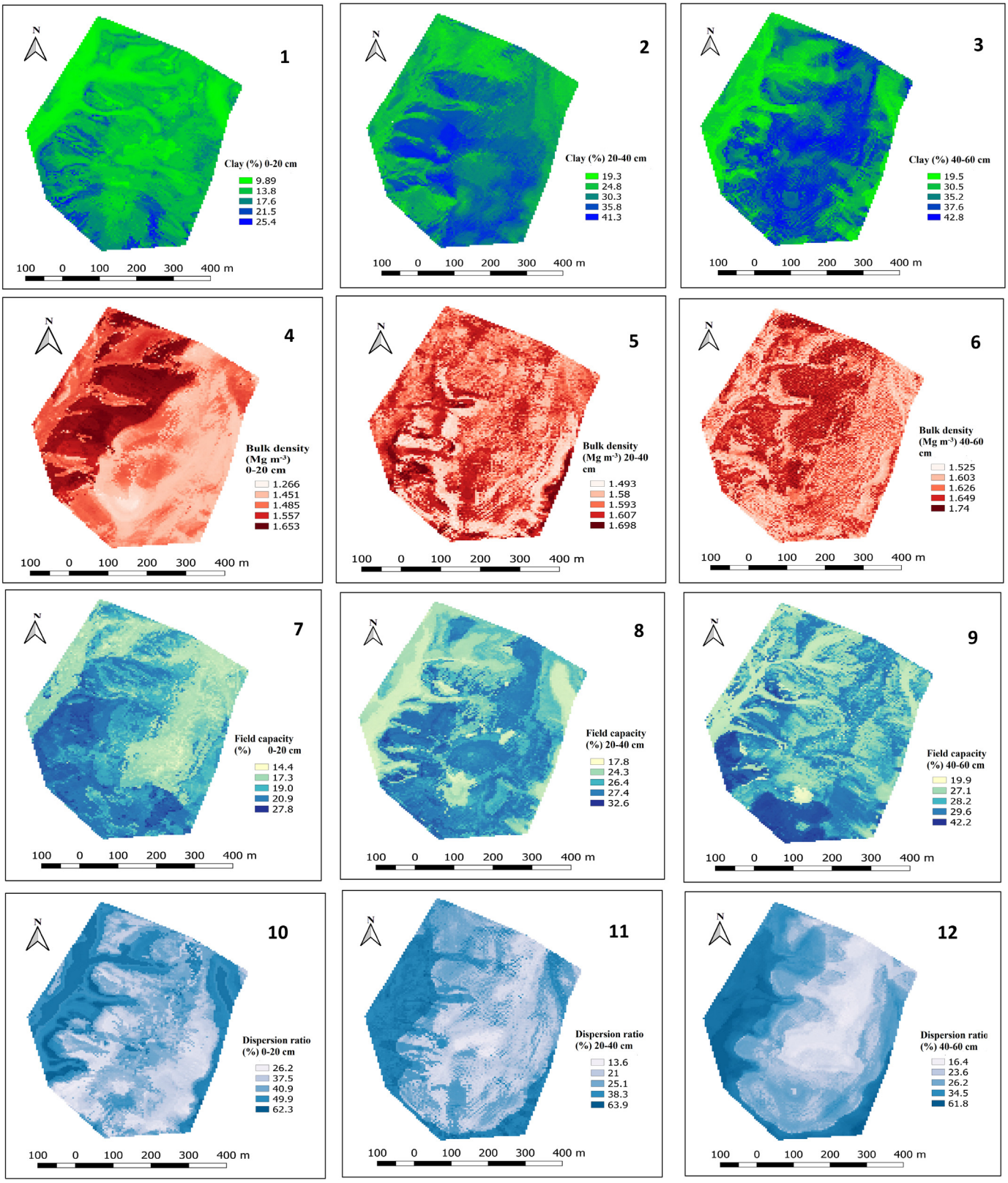

Fig. 4. Spatial distribution of predicted soil properties using Random Forest model in vineyard for three layers.

3.2.5. Soil water retention characteristics

Based on the averaged importance measured across 500 runs of the Random Forest model, the environmental variables that had the greatest influence on the model error rate for soil surface FC, PWP, SDP and FDP, respectively were: green-red vegetation index, MrVBF; elevation, MrVBF; modified catchment area, elevation and analytical Hillshading, longitudinal curvature, respectively (Fig. 2.1 to 2.5). The environmental variables that had the greatest influence on 40–60 cm soil depth soil water retention characteristics (FC, PWP, SDP and FDP) model error rate for final predictive models, averaged across 500 runs, were analytical Hillshading, topographic wetness index; topographic position index, topographic wetness index; modified catchment area, SAGA TWI and analytical Hillshading, modified catchment area, respectively (Fig. 2.1 to 2.5). The average error rate for the final models used to generate the predictive map of FC, PWP, SDP and FDP for all the sampled depth was ranging between 16 and 32%, 15–19%, 26–32% and 24–34%, respectively (Fig. 3b). The particle size distribution results also showed that the combination of the various environmental variables can account for R^2 of 0.40 to 0.54, 0.48 to 0.57, 0.31 to 0.54 and 0.20 to 0.50 of the variation in soil water retention characteristics, respectively (Fig. 3a).

The predictive map for surface water content at permanent wilting point (PWP) and field capacity showed high variation and moderate to high variation, with the majority of the area being classified as having less than 14% and 24%, respectively (4.7 and Supplementary Material II). The areas that were predicted to have a greater amount of water content at PWP and FC were predominantly in the areas with low elevations (4.7 and Supplementary Material II). On the other hand, the predictive map of the water content at PWP for subsurface (20–40 cm) showed high variation within the field, with the majority of the area being classified as 24% water content. Also, the predictive map for the water content at PWP (40–60 cm) showed moderate or medium variation within the field, with the majority of the area being classified as being 22% water content (Supplementary Material II). The RF model predicted high values of FC water content at the Southern part of the field as shown in Fig. 4.7. On the other hand, the predictive map of subsurface water content at FC showed high variability, with the majority of the areas being classified as having 29% (Fig. 4.8). The areas that were predicted to have a greater amount of FC water content were predominantly distributed across the field (Fig. 4.17). When looking at the predictive map for lower depth water content at FC, with the majority of the areas being classified as having less than 31% of water content at field capacity (Fig. 4.18). The predictive map of 40–60 cm soil depth FC showed that greater values were predicted at the lower areas of the field and the South to South Western parts of the field.

In addition, the predictive map of surface slow and fast drainage pores showed high variation within the field, with the majority of the areas classified as having 4% and less than 11%, respectively. The high volume of slow draining pores was predicted in the low lying slope on the Western side of the field (Supplementary Material II). The predictive map of the subsurface SDP showed high variation, with the majority of the area being classified as having between 3 and 2% (Supplementary Material II). The areas that were predicted to have a greater amount of SDP were predominantly in Northern and lower values are predicted in the Southern part of the vineyard. The predictive map of the lower depth (40–60 cm) SDP showed high variation, with the majority of the areas being classified as having less than 3%. The higher values are distributed in the Western side of the vineyard. The highest pores are distributed in the Eastern direction of the vineyard and the lowest were are at the Western part of the field. The predictive map for subsurface FDP again showed high variation, with the majority of the area being classified as being 6%. The low FDP values are scattered within the field. On the other hand, the predictive map of the lower depth (40–60 cm) FDP showed high spatial distribution within the field, with the majority of the area being classified as 11%. The lower FDP (<6%) are predicted to be at the flat area of the vineyard, which is situated in the middle of the field (Fig. 4.24). The CV maps obtained from 500 bootstrap prediction of the RF model showed that PWP and FC values were predicted with relatively moderate variation, while the prediction for SDP and FDP showed higher variation (Fig. 5 and Supplementary Material III).

4. Discussion

4.1. Descriptive statics and statistical models

4.1.1. Particle size distribution

The studied site produces vines under dryland conditions, therefore the distribution and movement of soil moisture content and nutrients within the soil profile mainly depend on soil properties. Particle size distribution is one of the properties that drive other soil properties, for example, determining the potential soil moisture that drives crop yield potential (Akpa et al., 2014). The results of the particle size distribution (Table 1) showed high variability in fine particles (3–29% silt and 6–46% clay). The soil particle size distribution is vital in most hydrological, ecological, and environmental risk assessment models (Liess et al., 2012). The clay content would be influenced by depositions (Ayuba et al., 2007; Sharu et al., 2013) and erosion processes at the surface (Amusan et al., 2005; Salako et al., 2006). The mean sand content is higher than clay and silt contents for each soil depth, which is commonly found in soil rich in quartz and granitic parent material as it is the case for this study (Sauer, 2010; Graham and O'Geen, 2010). The high sandiness and its decrease with soil depth could be due to the larger particle size of sand and its decreased transportability while silt and clay sizes are smaller and lighter hence easily moved in suspension both vertically and horizontally.

The coefficient of variations (CV) showed higher variation or heterogeneity in clay fraction compared to silt and sand fractions and the results are in agreement with the findings by Odeh et al. (2003) and Oku et al. (2010), but in contrast to Buchanan et al. (2012) and Adhikari et al. (2013) who reported a higher variability in sand content compared to clay and silt contents. The high variation could be due to depositions and differences in landscape. The low kurtosis values for clay, silt, and sand contents at all soil depths indicates lack of outliers in their data sets. The Clay content had a positive significant correlation with other attributes such as water contents at field capacity and permanent wilting point (Table 2). The reason for the positive correlation between clay and water content at both field capacity and the wilting point could be due to the pore size and water holding capacity of the clay fraction at all soil depths (Rowls et al., 2003). However, the negative correlation between clay content, fast drainage pores, slow drainage pores is explained by the pore sizes. While the silt content was negatively correlated with the saturated hydraulic conductivity because of its particle size which does not allow a constant flow of water within the soil profile due to its small particle size compared to sandy soils. However, on the other hand, a sand fraction is completely different from the clay and silt because it is characterized by a larger volume of macropores which drains water quickly when exposed to suction (Jury and Horton, 2004). Sand had a significant negative correlation with water retention at field capacity and at the permanent wilting point at all soil depths. This is associated with high volume of macropores (FDP) that releases water quickly when exposed to suction leading to low water available for crops.

In predicting the spatial variability of the particle size distribution for three soil depths within the vineyard, RF model performed significantly better at the top 40 cm compared with the lower sampling depths for all the particles. The opposite was observed for clay content, with the performance improvements with depth. Opposite results were reported by several authors (Henderson et al., 2005; Minasny et al., 2006; Malone et al., 2013). This could be accounted for by the nature of the environmental covariates used (Adhikari et al., 2013) and the effect of lower data

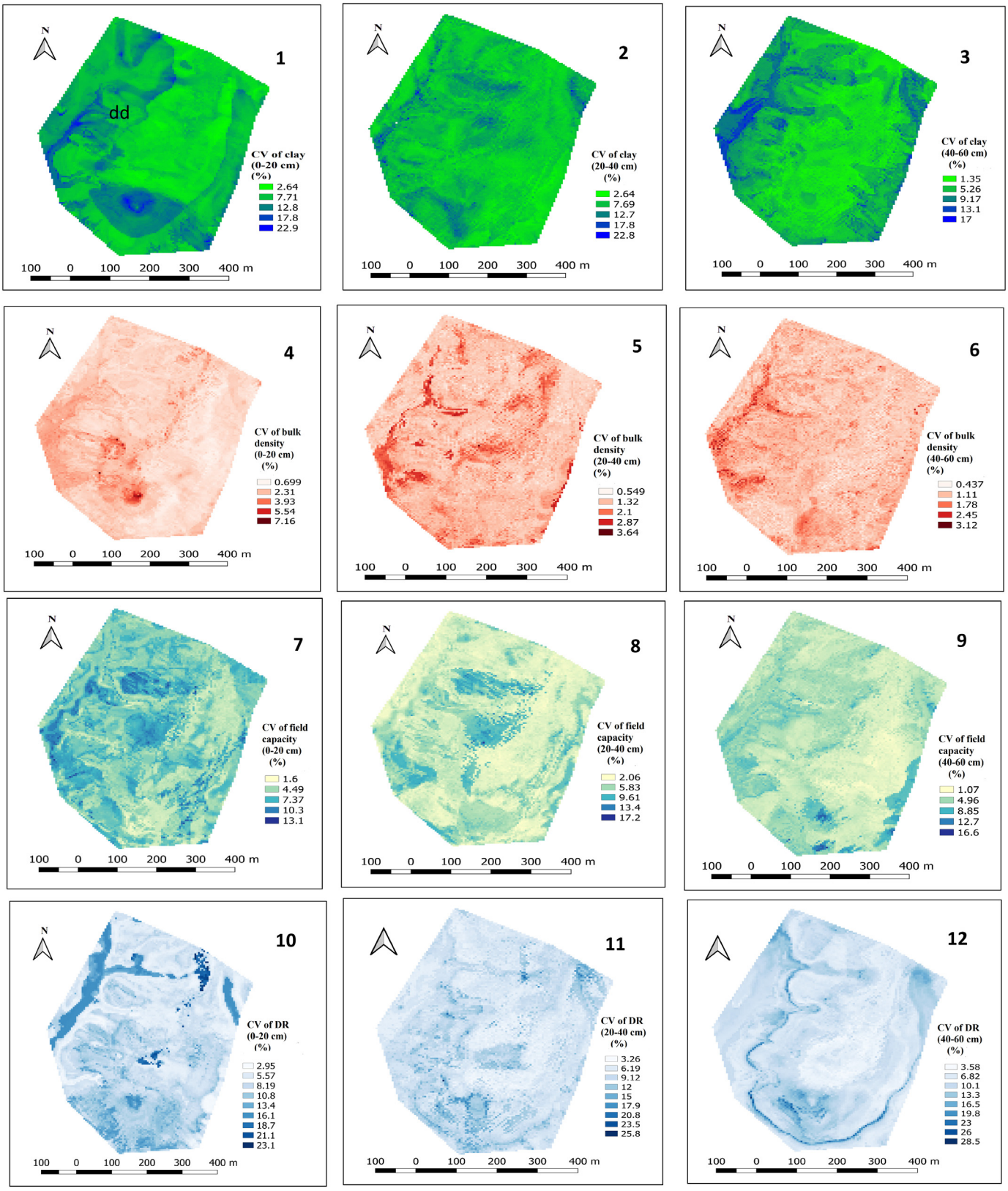

Fig. 5. Coefficient of variation (CV) maps in % obtained from the 500 bootstrapped RF model runs.

density with depth. In addition, most of the predictor variables used are more based on land surface features which are likely to have a strong influence on the topsoil than subsoil properties. The prediction accuracy for lower depth soil properties could be improved by the inclusion of predictor variables such as Gamma-radiometric and or electromagnetic induction (EM38) that allows accurate prediction of soil properties up to approximately 2 m soil depth (Cook et al., 1996; Rawlins et al., 2009; Priori et al., 2014).

In terms of prediction accuracy, clay content had the highest error rate (nRMSE) values across all depths whereas the lowest nRMSE was associated with the modeling of sand content at all soil depths. This variation is in disagreement with the work reported by studies using similar modeling approaches (Akpa et al., 2014; Niang et al., 2014; Buchannan et al., 2012). However, the results are in agreement with the work reported by Odeh et al. (2003) which studied the spatial composition of soil particle size fraction as compositional data, which came to a different conclusion as this current study. The variables that proved important for predicting particle size distribution included elevation, topographic position index, modified catchment area, green-red vegetation index, slope, MrVBF, topographic wetness index, and longitudinal curvature. Although not highly ranked as an important variables in the models, the influence of flow accumulation and aspect can be seen in the predictive maps for subsurface soil clay content. However, the relative importance of the variables differs with depth and from one particle to another. Other authors have also reported the relationship between terrains attributes and soil properties, especially particle size distribution (Moore et al., 1993; Odeh et al., 1995; Greve et al., 2012a, 2012b), with terrain attributes explaining 20 and 88% of the variation (Thompson et al., 2006).

This could be explained to their impact on the vertical and horizontal movement of soil particles through erosion and deposition in low areas of the field. Sand content is also influenced by geology and rates of weathering. However, the influence of geology and soil type on spatial distribution soil texture in Chile is not well documented and few studies are available (Bernhard et al., 2018). The spatial distributions of soil texture in general effect and control runoff generation, slope stability, depth of accumulation, and soluble salt content (Yoo et al., 2006; Gochis et al., 2010; Crouvi et al., 2013). For soil properties such as texture, it becomes important to also pay attention to other observations made at the field site to determine what the dominant processes were influencing soil properties.

4.1.2. Bulk density

The sandy texture also influenced the bulk density values (Chaudhari et al., 2013), which were considered slightly high ($1.48 < 1.59 < 1.63$ Mg m^{-3}) for all soil depths (see Table 1). The average bulk density within the field increases with the soil depth. This behavior could be explained by the influence of soil organic matter, which decrease the soil bulk density by its own low value of bulk density (Seguel and Horn, 2006). Based on CVs values of bulk density, all soil showed low spatial variability within the vineyard. The low variability of bulk density could also be because of uniform parent material in the area. The obtained results are in agreement with the work of Osama et al. (2008) that aimed at investigating the spatial variability of penetration resistance. According to Ellerbrook et al. (2005) and Eynard et al. (2006), soil organic matter interacts with other soil properties to influence water behavior in soils. The skewness and kurtosis showed that the data set had low trails. The Pearson correlation results show that bulk density had a significant ($p < .01, 0.05$) negative correlations with Ksat and fast drainage pores and significant ($p < .01, 0.05$) positive correlation with sand, and dispersion ratio. The significant negative correlation with Ksat and Fast drainage pores is associated with total porosity, where the increase in bulk density will affect the total porosity and reduce water penetration.

The Random Forest model was able to take a set of noisy data and still identify correlations between soil bulk density and environmental covariates that could be interpreted as landscape processes acting on the soil. The variables that proved important for predicting bulk density included aspect, topographic wetness index SAGA, longitudinal curvature, slope length factor, analytical Hillshading, and TBSI-W for predicting soil bulk density, which is associated with soil mineralogy. From the list of important covariates used to predict the bulk density, aspect, longitudinal curvature and analytical Hillshading proved to be the most important variables of them all. Therefore the removal of these variables from the list of predictors would increase the error rate by reducing accuracy. Some of these covariates influence the organic matter, which affects the bulk density of the soil surface (Liu et al., 2015; Hu et al., 2019). These variables include curvatures and slope length which are associated with erosion and distribution of surface materials such as leaf litter and the directional flow. The RF model had the lowest error rate for predicting spatial variability in soil bulk density for all soil depths. However, surface soil layer had higher R^2 and nRMSE compared to lower soil depths. This variation could be explained by different environmental covariates used for prediction of bulk density at each sampling depths. Some covariates such as remote sensing data as in this case are well known to directly predict topsoil variables compared lower depths. Similar results were reported by Adhikari et al. (2014). As it was expected, the predictive maps showed an increase in bulk density with soil depths. Looking at its distribution, soils in the northern part of the field appeared to have high density than soils of the central and Eastern vineyard, especially in the 0–20 cm depth. The spatial variability of bulk density in the vineyard could also be associated with variation in texture and organic matter (Adams, 1973) within the field.

4.1.3. Dispersion ratio

The soil micro-aggregate stability was evaluated by the dispersion ratio method (DR), for the three sampling depths. The mean values of the DR index associated with micro-aggregate stability showed an increase with soil depth; however, the lowest values of DR index indicate high micro-aggregate stability within the soil profile. This was probably due to the increase in clay content and presence of Fe oxides in the lower soil depths, as was suggested by Brunel et al. (2010) in the same soil series with different tillage system. The micro-aggregate stability results are in agreement with the work reported by Seguel et al. (2003) which stated that clay content plays an important role in the formation of micro-aggregates. It is well known that organic material binds the soil particles together to resist the degradation and improves the porosity of the soil, thus reduces the slaking and dispersion (Chenu et al., 2000). The CV values of micro-aggregate indicate high to very high variability within the field in all the three sample depths according to the criteria proposed by Gomes and Garcia (2002). However, in the study area, the high variability of stability could be ascribed to pedogenic processes.

The predictive maps of DR showed the increase of the aggregate stability (lower DR values) with soil depth and the spatial variability, with the 40–60 cm soil depths showing more variability. The reason for the variation could be explained by organic matter content across the field and the variation in clay content (Seguel et al., 2003). The RF models of microaggregate stability showed that TPI, TWI, modified catchment area, MRVBF, digital elevation model, longitudinal curvature, secondary curvature, catchment slope, and slope were the dominant factors affecting the stability. Several studies also showed that spatial variability of aggregate stability is closely related to the terrain attributes such as slope, curvatures, and aspect through their impacts on various soil properties (Rhoton and Duiker, 2008) which are associated with erosion and distribution of surface materials such as leaf litter, clay particles and the detainment of water. However, Rhoton et al. (2006) noted that studies on the spatial variability of aggregate stability and its relationship with topography are rare and focused on the assessment of soil aggregate stability in different parts of the slope than on their direct relationship with topographic derivatives (Canton et al., 2009). Therefore based on

this study, topographic wetness index and slope played more impact in the prediction of spatial variability of aggregate stability, through their impact on various properties (Rhoton and Duiker, 2008). In terms of prediction accuracy, the surface depth had the lowest error rate (nRMSE) whereas the highest error rate was associated with the prediction of dispersion ratio at all 20–40 and 40–60 cm soil depth. This could be attributed to the nature of the environmental variables used (Adhikari et al., 2013) and the effect of lower data density with depth. However, all the depths has demonstrated promising results and high values of the determinant coefficient in the case of dispersion ratio.

4.1.4. Saturated hydraulic conductivity

The results obtained in this study showed that saturated hydraulic conductivity (Ksat) decreases with soil depth (50.6 > 22.6 > 18.4 cm hr^{-1}) see Table 1. The reason for this variation within the soil profile is associated with soil physical properties such organic matter, soil texture, bulk density and total porosity (Hillel, 1980). The Ksat of the surface was on average about two times greater than for the lower soil depths, therefore increased Ksat values in the surface depth could be due to lower bulk density owing to the presence of organic material and high total porosity. A similar conclusion was reported by Rasse et al. (2000) and Iqbal et al. (2005). The CV values of saturated hydraulic conductivity (Ksat) increases with soil depth (104.4 < 197.3 < 229.9%) and the Ksat is considered as highly variable within the field according to Mulla and McBratney (2002). The high CV values of Ksat are high because it is affected by many other soil properties. There are several arguments regarding the magnitude of the spatial variability of saturated hydraulic conductivity across the agricultural fields from which, Biggar and Nielsen (1976) reported Ksat as the one with the highest variability. For example, Jury and Horton (2004) indicated the values of coefficient of variation for saturated hydraulic conductivity in the range of 50–300%. Therefore the present study is in agreement with the work that was reported by the above literature. The Ksat data was log-transformed because it was skewed, as a result, mean values and median values were not similar. The Pearson correlation results show that Ksat had a significant ($p < .01, 0.05$) negative correlations with silt, bulk density, water content at field capacity and permanent wilting point for 0–20 cm and 20–40 cm depths, except for the lower depth (40–60 cm) which showed that statistically there was no significant correlation between Ksat and other soil physical properties. Ksat at 0–20 cm and 20–40 cm depths also showed significant ($p < .01, 0.05$) positive correlation with sand, and fast drainage pores (Table 2). The obtained significant correlation between Ksat and other soil properties is in agreement with the work of Candemir and Gülser (2012) which found that Ksat significantly increased with increasing sand and decreasing clay content. The results reported by Iqbal et al. (2005) stated that increased Ksat values in the surface horizon could be due to lower bulk density owing to the presence of macro-porosities.

An appropriate estimation of saturated hydraulic conductivity (Ksat) was necessary, as it is an essential part of management practices including, irrigation, drainage, flood protection and erosion control in addition to water flow and transport modeling in soil. As expected, the Ksat decreased with soil depths. The trend is associated with other soil properties such as macropores, texture, bulk density, organic matter and water retention (Kotlar et al., 2019). The predictive maps of Ksat showed higher spatial variability within the vineyard, compared to other predicted soil properties. There were areas within the field that showed very low saturated hydraulic conductivity more especially the areas at the centre of the field, which makes them more at risk of being eroded. The variables that proved important for predicting saturated hydraulic conductivity for all soil depths included longitudinal curvature, TBSI-T, topographic position index, topographic wetness index, flow accumulation and catchment area. Therefore the removal of any of these variable from the final predictive model would result in the increase of error rate, thus reducing model accuracy. These variables played an important role in the prediction of Ksat because they control erosion and distribution of fine materials such as clay particles and the accumulation of water. For an example, TWI described soil moisture pattern in the watershed based on topography. However, the relative importance of these variables varies with soil depths. During the fieldwork, it was evident that some areas that had low Ksat were washed off or eroded. The RF model had the highest error rate for predicting spatial variability in Ksat. However, the error rate and precision varied with soil depths. The variation of in prediction accuracy of the models can be associated with the nature of environmental covariates used for the prediction of the Ksat at each soil depth. The prediction accuracy can be improved by using the soil properties (clay, sand, silt, bulk density, water content, and organic matter) that affect Ksat as input for predictive models (Nemes et al., 2005). The values are in accordance with the previous studies using similar methods. Agyare et al. (2007) while estimating Ksat obtained R^2 and RMSE about 0.6 and 0.42, respectively. On the other hand, Merdun et al. (2006) obtained R^2 range and RMSE varied from 0.44 to 0.95 and 0.020 to 3.51, respectively. The study area produces vines under dryland condition, which means that the distribution of water within the profile mainly depends on other soil properties and precipitation. Therefore it was necessary to predict saturated hydraulic conductivity using machine learning (RF model) because measuring Ksat is time-consuming and very costly, it varies much within time and space.

4.1.5. Soil water retention characteristics

The results of water retention characteristics showed that the water content at field capacity and permanent wilting point increase with soil depth. The reason for this trend is associated with relatively higher clay content, low total porosity and high bulk density in lower soil depths which allow clay particle to hold moisture to its micropores and less water is released (Warrick, 2002). The pore size distribution mean values decreases with soil depth, this is associated with particle size distribution. For example, the decrease in fast-draining pores with soil depth is influenced by the increase in clay content in lower depths. The contribution of slow drainage pores is necessary to improve the amount of usable water for the plant growth according to Hartge and Horn (2009). All the CV values of soil water retention characteristics showed high and very high variability (10% to >30%) within the field. Similar results for water retention data was obtained by Malla et al. (1996), observing the variance tends to decrease with increasing depth. Water release through more uniform pores may explain it, particularly for high water tension values. In this study, a lower variance for high water retention values may be explained by water retention caused by absorption rather than capillarity since it's controlled by porosity (Warrick, 2002). The availability of water content within the soil profile is influenced by many other soil physical properties such as texture, organic matter, and pore size distribution depending on soil structure (Warrick, 2002). A high significant positive correlation was found between water content at field capacity (33 kPa) and permanent wilting point (1500 kPa) and soil properties such as clay, silt contents, and dispersion ratio. This was expected because the increase in organic matter helps improve the water holding capacity of the soil and clay content helps retain more water. However, the water that is absorbed by micropores of clay fraction is not available for the plants. On the other hand, significant ($p < .05, < 0.01$) negative correlation was observed between PWP, FC and sand content, saturated hydraulic conductivity (Table 2). This was expected because sandy soil contains a high volume of macropores that drains water quickly when exposed to suction, which then leaves the soil with less water at field capacity and at permanent wilting point (Jury and Horton, 2004). The reduction of water content at field capacity and the wilting point when sand content increase reduces the water retention within the profile, which would then leave the plants water-stressed and thus others dying (Hillel, 1980).

This study would not be complete if we did not characterize the pore size distribution, because pores play an important role in vertical and

horizontal movement of water and nutrients within the soil profile. The soil pore sizes in the vineyard were classified as fast draining pores (non-capillary) pores, slow draining pores (coarse capillary) and water-holding pores according to Behaver and Gardner (1972). The predictive maps of pores size distribution showed the decrease of fast-draining pores (FDP, >50 μm) with soil depth (see Supplementary Material II to 4.24), the reason for the variation could be explained by associated with soil texture (clay, sand, and silt) content, soil structure and organic matter across the field. Similar results were reported by Seguel et al. (2015). These pores are responsible for the vertical flow of water that is available for root uptake and helps in solute movement in the soil. The study by Chen and Wagenet (1992) showed that FDP comprise only a small portion of the total soil voids, but under some conditions, vertical flow through macropores dominate during infiltration. The topsoil contains a high volume of FDP because of the domination of coarse-textured soil which drain first after rainfall and play a significant effect of soil moisture content at field capacity. The main reason for the characterization of the pores was the fact that the ability of pores to conduct water is mainly controlled by pore size, continuity, and distribution of pores in the soil. However, on the other hand, predictive maps of slow draining pores (SDP, 10–50 μm) showed a decrease of this kind of pores with soil depth. This could be associated with soil texture and soil aggregation. These pores help the horizontal and upward movement of water due to the presence of cohesion and adhesion forces. The spatial variability of slow draining pores is associated with other soil properties such as soil texture, organic matter, soil depth and soil aggregation or arrangement of soil particles.

In terms of prediction accuracy, slow drainage and fast drainage pores had the highest error rate (nRMSE) values across all depths whereas the lowest nRMSE was associated with the modeling of water content at field capacity and permanent wilting point at all soil depths. This could be associated with environmental variables used in prediction of soil water retention characteristics. The predictive maps of water retention characteristics showed the increase of the moisture content (FC and PWP) with soil depth. According to Yoon et al. (2007), at low matric tension, the water retention depends on soil structure, while at higher matric tension (PWP) water retention depends on the soil texture. For example, clay content contains a large volume of micropores which holds more water. However, some of that adsorbed water is available for plants and some are not available to be taken up by plants. The predictive maps showed the high prediction of soil moisture content in the areas with low elevation, which is associated with the depositions of fine materials more especial in the south to southwest areas. Moisture content combined with other soil properties plays an important role in rainfed vine production, as it influences the movement of nutrients within the soil. The spatial variability demonstrated by the predictive maps of water retention characteristics is associated with the spatial of other soil properties such as silt, clay, sand, bulk density, and organic matter which affects water holding capacity and water uptake (Van Leeuwen et al., 2004). The areas with the lesser available water content at lower depth are more likely to be water stress and affect the vine size because vine behavior is closely related to water uptake. The horizontal spatial variability of available water content is in agreement with the work reported by Tsegaye and Hill (1998).

The variables that proved important for predicting soil water retention characteristics included elevation, topographic position index, modified catchment area, green-red vegetation index, analytical Hillshading, MrVBF, topographic wetness index, and longitudinal curvature among others. However, the relative importance varied with soil depths. The predictors such as elevation, topographic position index, topographic wetness index and curvature were expected to affect the water content because they affect the local climate (microclimate), controls depositions and accumulation of water. Predictor variables such as MrVBF represent areas of accumulation of sediments according to the topographic surrounding context and has demonstrated its relevance to explain soil physical properties among different landcovers for the same study region (Soto et al., 2019). In this sense, Taylor et al. (2013) reported MRVBF as a relevant predictor of water table depths under Mediterranean landscape of southern France. Therefore lower areas had the lowest MRVBF and accumulation of sediments was possible. It was also interesting to see remote sensing predictors such as green-red vegetative index playing an important role in predicting the soil moisture content of the topsoil because vegetation vigour can be indirectly used to predict the soil properties.

The RF model showed promising results in predicting the soil water content, however, it showed little accuracy during the prediction of available water content. The prediction results obtained in this study are in agreement with the work reported by Szabó et al. (2019) which mapped soil hydraulic properties based on random forest based on Pedotransfer functions and geostatistics. Even though the final RF model performed fairly well, it can still be improved by reducing the number of predictors and using predictor variables related to water distribution (channel network, valley depth…etc.). The accurate final models would be used to design sustainable agricultural system management strategies responsive to fluctuating soil moisture regime within the vineyard because it is essential for modeling agricultural system productivity. The final models showed a slight tendency to overestimate areas with low pore sizes and to produce a slight underestimate of areas high pore sizes. This could be associated with model calibration. Less to none work published about the prediction of draining pores, reason why the predicted pore sizes results can be used as input for further hydraulic property analysis within the vineyard.

5. Conclusion

A field study was conducted to investigate and characterize the spatial variability of selected soil properties using digital soil mapping within the vineyard. The following conclusions are inferred from the study:

Low to very high spatial variability in selected soil properties was observed across the experimental field but the magnitude varied with soil depth. The soil properties that showed a considerable degree of variation were clay content, saturated hydraulic conductivity, available water content and dispersion ratio. The applied DSM approach including Random Forest model as a relatively tool in the field of soil science for prediction of selected soil properties yielded promising results as the accuracy of the model and generated prediction maps were acceptable. Soil particle distribution is directly correlated with soil draining pores, water content and hydraulic conductivity. All these results indicate that a better understanding of spatial variability of soil properties at a depth deeper than 40 cm can help improve the quality and yield from the vineyard.

The main environmental predictors for soil properties variability in the vineyard were analytical Hillshading, topographic wetness index, topographic position index, longitudinal curvature, secondary curvature, modified catchment area, GRVI, TBSI-W, TBSI-V and TBSI-T, MRVBF, DEM, aspect, convergence index, and length slope factor. The remote sensing data also played a role in the prediction and can be used with success as input for digital soil mapping. Random Forest model provided a promising framework for the spatial prediction of soil properties as the accuracy of the model performance was acceptable. RF model predicted particle size distribution, bulk density, and water content at permanent wilting point significantly well. The RF model can be improved by applying predictor variables that are directly related to the variable of interest. As DSM mapping showed satisfactory concordance with the conventional soil map, in combination with the other observations made in this study, the application of pedometric methods such as DSM algorithms should be seriously considered as a complementary approach to conventional methods for mapping soil properties in the Mediterranean vineyards. The saturated hydraulic conductivity was predicted with high variation and bulk density was predicted with low variation according to the CV maps. Finally, the

produced predictive maps can be used for modeling agricultural system productivity within the vineyard and can be used for site-specific management.

Declaration of Competing Interest

The authors declare that they have no financial or personal relationship (s) that may have inappropriately influenced them in writing this article.

Acknowledgement

We would like to thank Fondecyt 1171560: "Assessing Spatio-temporal impacts of global change on water and biomass production processes at catchment scale: a synergistic approach based on remote sensing and coupled hydrological models to improve sustainable management of forest ecosystems", and the Center for Climate and Resilience Research (ANID/FONDAP/15110009) for financing our research. Finally, we would like to thank Arauco forest for supplying us with the LiDAR data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geodrs.2020.e00289.

References

Adams, W.A., 1973. The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils. J. Soil Sci. 24 (1), 10–17.
Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., McBratney, A. B., Greve, M.H., 2013. High-resolution 3-D mapping of soil texture in Denmark. Soil Sci. Soc. Am. J. 77, 860–876.
Adhikari, K., Minasny, B., Greve, M.B., Greve, M.H., 2014. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. Geoderma 214–215, 101–113.
Agyare, W.A., Park, S.J., Vlek, P.L.G., 2007. Artificial neural network estimation of saturated hydraulic conductivity. Vadose Zone J. 6 (2), 423–431.
Akpa, S., Bishop, T.F., Odeh, I.A., Hartemink, A.E., 2014. Digital mapping of soil particle-size fractions for Nigeria. Soil Sci. Soc. Am. J. 78 (6), 1953–1956.
Amusan, A.A., Olayinka, A., Oyedele, D.J., 2005. Genesis, classification, and management requirements of soils formed in windblown material in the Guinea Savanna area of Nigeria. Commun. Soil Sci. Plant Anal. 36, 2015–2031.
Arrouays, D., Grundy, M.G., Hartemink, A.E., Hempel, J.W., Heuvelink, G.B.M., Hong, S.Y., et al., 2014. GlobalSoilMap. Adv. Agron. 125, 93–134.
Ayuba, S.A., Akamigbo, F.O.R., Itsegha, S.A., 2007. Properties of soils in river Katsina-Ala catchments area, Benue State, Nigeria. Niger. J. Soil Sci. 17, 24–29.
Bernhard, N., Moskwa, L.M., Schmidt, K., et al., 2018. Pedogenic and microbial interrelations to regional climate and local topography: new insight from a climate gradient (arid to humid) along the coastal cordillera of Chile. Catena 170, 335–355.
Biggar, J.W., Nielsen, D.R., 1976. Spatial variability of leaching characteristics of a field soil. Water Resour. Res. 12, 78–84.
Bramley, R.G.V., 2005. Understanding variability in winegrape production systems. 2. Within vineyard variation in quality over several vintages. Aust. J. Grape Wine Res. 11, 33–42.
Buchanan, S., Triantafilis, J., Odeh, I., Subansinghe, R., 2012. Digital soil mapping of compositional particle-size fractions using proximal and remotely sensed ancillary data. Geophysics 77 (4), 201–211.
Cabezas, J., Galleguillos, M., Pérez-Quezada, J.F., 2016. Predicting vascular plant richness in a heterogeneous wetland using spectral and textural features and a random forest algorithm. IEEE Geosci. Remote Sens. Lett. 13, 646–650.
Candemir, F., Gülser, C., 2012. Influencing factors and prediction of hydraulic conductivity in fine textured-alkaline soils. Arid Land Res. Manag. 26 (1), 15–31.
Castillo-Riffart, I., Galleguillos, M., Lopatin, J., Pérez-Quezada, J.F., 2017. Predicting vascular plant diversity in anthropogenic peatlands: comparison of modeling methods with free satellite data. Remote Sens. 9 (7), 681.
Chaudhari, P.R., Dodha Ahire, V., Vidya Ahire, D., Chkravarty, M., Maity, S., 2013. Soil bulk density as related to soil texture, organic matter content and available total nutrients of Coimbatore soil. Int. J. Sci. Res. Publ. 3 (2), 2250–3153.
Chen, C., Wagenet, R.J., 1992. Simulation of water and chemicals in macropore soils. Part 1. Representation of the equivalent macropore influence and its effect on soil water flow. J. Hydrol. 130, 105–126.
Chenu, C., Le Bissonnais, Y., Arrouays, D., 2000. Organic matter influence on clay wettability and soil aggregate stability. Soil Sci. Soc. Am. J. 64, 1479–1486.
CIREN (Centro de Información de Recursos Naturales Chile), 1997. Estudio Agrológico Región VII. Descripciones de suelos, materiales y símbolos Santiago, Chile. (Publicación 117).
Conrad, O., 2014. System for Automated Geoscientific Analyses (SAGA). Version: 2.1.2 Available online: http://www.saga–gis.org (accessed on 29 July 2014).
Cook, S.E., Corner, R.J., Groves, P.R., Grealish, G.J., 1996. Use of airborne gamma radiometric data for soil mapping. Aust. J. Soil Res. 34, 183–194.
Crouvi, O., Pelletier, J.D., Rasmussen, C., 2013. Predicting the thickness and aeolian fraction of soils in upland watersheds of the Mojave Desert. Geoderma 195, 94–110.
Dane, J., Topp, G., 2002. Methods of soil analysis. Part 4. Physical Methods. SSSA Book Ser. 5. SSSA, Madison, WI 1692 p.
Ettema, C.H., Wardle, D.A., 2002. Spatial soil ecology. Trends Ecol. Evol. 17, 177–183.
Eynard, A., Schumacher, T.E., Kohl, R.A., Malo, D.D., 2006. Soil wettability relationships with soil organic carbon and aggregate stability. Proceedings of the 18th World Congress of Soil Science, Philadelphia, July 9–15 Pennyslavia, USA.
Felzensztein, C., Echecopar, G., Deans, K., 2011. Marketing Strategy, Innovation and Externalities: The Case of the Chilean Wine Cluster, BALAS Conference. Universidad Adolfo Ibanez Santiago, April, 2011.
Forkuor, G., 2014. Agricultural Land Use Mapping in West Africa Using Multi-sensor Satellite Imagery. University of Wuerzburg, Wuerzburg, Germany, p. 191.
Giasson, E., Caten, A.T., Bagatini, T., Bonfatti, B., 2015. Instance selection in digital soil mapping: a study case in Rio Grande do Sul, Brazil. Ciência Rural 45 (9), 1592–1598.
Gochis, D.J., Vivoni, E.R., Watts, C.J., 2010. The impact of soil depth on land surface energy and water fluxes in the North American monsoon region. J. Arid Environ. 74, 564–571.
Gomes, F.P., Garcia, C.H., 2002. Estatrstica aplicada a experimentos agronómicos e florestais. [statistics Applied to Agronomic Experiments and Forestry.]. FAELQ, Piracicaba, p. 309.
Goovaerts, P., 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. Biol. Fertil. Soils 27, 315–334.
Graham, R.C., O'Geen, A.T., 2010. Soil mineralogy trends in California landscapes. Geoderma 154 (3–4), 418–437.
Greve, M.H., Kheir, R.B., Greve, M.B., Bøcher, P.K., 2012a. Quantifying the ability of environmental parameters to predict soil texture fractions using regression-tree model with GIS and LiDAR data: the case study of Denmark. Ecol. Indic. 18, 1–10.
Greve, M.H., Kheir, R.B., Greve, M.B., Bøcher, P.K., 2012b. Using digital elevation models as an environmental predictor for soil clay contents. Soil Sci. Soc. Am. J. 76, 2116–2127.
Grossman, R.B., Reinsch, T.G., 2002. The solid phase. Bulk density and linear extensibility. In: Dane, J.H., Topp, G.C. (Eds.), Methods of Soil Analysis. Part 4. Physical Methods. Soil Science Society of America, Madison, Wisconsin, USA, pp. 201–228 Book Series N° 5.
Guyon, I., Weston, J., Barnhill, S., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46, 389–422.
Hartge, K., Horn, R., 2009. Die physikalische Untersuchung von Böden. Praxis Messmethoden Auswertung. 4. Vollst. Überarbeitete Auflage . Schweizerbart Vorlage, Stuttgart.
Henderson, B.L., Elisabeth, N.B., Christopher, J.M., Simon, D.A.P., 2005. Australia-wide predictions of soil properties using decision trees. Geoderma 124, 383–398.
Heung, B., Bulmer, C.E., Schmidt, M.G., 2014. Predictive soil parent material mapping at a regional-scale: A random Forest approach. Geoderma 214-215, 141–154.
Hillel, D., 1980. Fundamental of Soil Physics. Academic Press, New York.
Hu, C., Wright, A., Lian, G., 2019. Estimating the spatial distribution of soil properties using environmental variables at a catchment scale in the loess hilly area, China. Int. J. Environ. Res. Public Health 16 (3), 491.
Iqbal, J., Thomasson, J.A., Jenkins, J.N., Owens, P.R., Whisler, F.D., 2005. Spatial variability analysis of soil physical properties of alluvial soils. Soil Sci. Soc. Am. J. 69 (4), 872–882.
Jabro, J.D., Stevens, W.B., Evans, R.G., Iversen, W.M., 2010. Spatial variability and correlation of selected soil properties in the Ap horizon of a CRP grassland. Appl. Agric. Eng. 26 (3), 419–428.
Jenny, H., 1941. Factors of Soil Formation - a System of Quantitative Pedology. McGraw-Hill, New York.
Jury, W.A., Horton, R., 2004. Soil Physics. 6th ed. John Wiley & Sons, New York 370p.
Kempen, B., Brus, D.J., Stoorvogel, J.J., Heuvelink, G.B.M., Vries, F., 2012. Efficiency comparison of conventional and digital soil mapping for updating soil maps. Soil Sci. Soc. Am. J. 76, 2097–2115.
Kilic, K., Kilic, S., Kocyigit, R., 2012. Assessment of spatial variability of soil properties in areas under different land use. Bulgar. J. Agr. Sci. 18, 722–732.
Kotlar, A.M., Iversen, B.V., de Jong van Lier, Q., 2019. Evaluation of parametric and non-parametric machine-learning techniques for prediction of saturated and near-saturated hydraulic conductivity. Vadose Zone J. 18, 141–180.
Kuhn, M (Contributions from Jed Wing), Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., et al., 2017. Caret: Classification and Regression Training R package version. 6.0–76.
Lacoste, M., Lemercier, B., Walter, C., 2011. Regional mapping of soil parent material by machine learning based on point data. Geomorphology 133, 90–99.
Ladha, L., Deepa, T., 2011. Feature selection methods and algorithms. Int. J. Comput. Sci. Eng. 3 (5), 1787–1797.
Lagacherie, P., McBratney, A.B., 2007. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: Lagacherie, P., et al. (Eds.), Digital Soil Mapping: An Introductory Perspective. Elsevier, New York, pp. 3–22.
Liaw, A., Wiener, M., 2002. Classification and regression by RandomForest. R News 2 (3), 18–22.
Liess, M., Glaser, B., Huwe, B., 2012. Uncertainty in the spatial prediction of soil texture comparison of regression tree and random forest models. Geoderma 170, 70–79.
Liu, X., Zhang, W., Zhang, M., Ficklin, D.L., Wang, F., 2009. Spatio-temporal variations of soil nutrients influenced by an altered land tenure system in China. Geoderma 152, 23–34.

Liu, S., An, N., Yang, J., Dong, S., Wang, C., Yin, Y., 2015. Prediction of soil organic matter variability associated with different land use types in mountainous landscape in southwestern Yunnan province, China. Catena 133, 137–144.
Lopatin, J., Dolos, K., Hernández, H.J., Galleguillos, M., Fassnacht, F.E., 2016. Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in Central Chile. Remote Sens. Environ. 173, 200–210.
Malone, B.P., McBratney, A.B., Minasny, B., 2013. Spatial scaling for digital soil mapping. Soil Sci. Soc. Am. J. 77 (3), 890–902.
McBratney, A.B., Mendoça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.
Merdun, H., Cinar, O., Meral, R., Apan, M., 2006. Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. Soil Tillage Res. 90 (1–2), 108–116.
Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. Geoderma 264, 301–311.
Minasny, B., McBratney, A.B., Mendonça-Santos, M.L., Odeh, I.O.A., Guyon, B., 2006. Prediction and digital mapping of soil carbon storage in the lower Namoi Valley. Aust. J. Soil Res. 44, 233–244.
Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57, 443–452.
Mulla, D., McBratney, A.B., 2002. Soil spatial variability. In: Warrick, A. (Ed.), Soil Physics Companion. CRC Press, Boca Raton, pp. 343–373.
Nemes, A., Rawls, W.J., Pachepsky, A., 2005. Influence of organic matter on the estimation of saturated hydraulic conductivity. Soil Sci. Soc. Am. J. 69 (4), 1330–1337.
Niang, M.A., Nolin, M., Jego, G., Perron, I., 2014. Digital mapping of soil texture using RADARSAT-2 polarimetric SAR data. Soil Sci. Soc. Am. J. 78, 673–684.
Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67, 215–226.
Odeh, I.O.A., Todd, A.J., Triantafilis, J., 2003. Spatial prediction of soil particle-size fractions as compositional data. Soil Sci. 168, 501–515.
Oku, E., Essoka, A., Thomas, E., 2010. Variability in soil properties along an Udalf toposequence in the humid forest zone of Nigeria. Kasetsart J. (Nat. Sci.) 44, 564–573.
Osama, M., Tomoyasu, I., Kazunari, F., Kunihiko, Y., 2008. Assessment of spatial variability of penetration resistance and hardpan characteristics in a cassava field. Aust. J. Soil Res. 46 (3), 210–218.
Outeiro, L., Aspero, F., Ubeda, X., 2008. Geostatistical methods to study spatial variability of soil cations after a prescribed fire and rainfall. Catena 74, 310–320.
Paz-Gonzalez, A., Vieira, S.R., Castro, M.T.T., 2000. The effect of cultivation on the spatial variability of selected properties of an umbric horizon. Geoderma 97 (3–4), 273–292.
Priori, S., Bianconi, N., Costantini, E.A., 2014. Can g-radiometrics predict soil textural data and stoniness in different parent materials? A comparison of two machine-learning methods. Geoderma 226–227, 354–364.
Rasse, D.P., Smucker, A.J.M., Santos, D., 2000. Alfalfa root and shoot mulching effects on soil hydraulic properties and aggregation. Soil Sci. Soc. Am. J. 64 (2), 725.
Rawlins, B.G., Marchant, B.P., Smyth, D., Scheib, C., Lark, R.M., Jordan, C., 2009. Airborne radiometric survey data and a DTM as covariates for regional scale mapping of soil organic carbon across Northern Ireland. Eur. J. Soil Sci. 60, 44–54.
Reif, D.M., Motsinger, A.A., McKinney, B.A., Crowe Jr., J.E., Moore, J.H., 2006. Feature Selection Using a Random Forests Classifier for the Integrated Analysis of Multiple Data Types. IEEE, Computational Intelligence and Bioinformatics and Computational Biology, pp. 1–8.
Resende, M., Curi, N., Rezende, S.B., Corrêa, G.F., 2014. Gênese: aspectos gerais. In: Pedologia: base para distinção de ambientes. 6. ed. Editora UFLA, Lavras, MG.
Rhoton, F.E., Duiker, S.W., 2008. Erodibility of a soil drainage sequence in the loess uplands of. Mississippi Catena 75 (2), 164–171.
Rhoton, F.E., Emmerich, W.E., Goodrich, D.C., Miller, S.N., McChesney, D.S., 2006. Soil geomorphological characteristics of a semiarid watershed. Soil Sci. Soc. Am. J. 70 (5), 1532–1540.
Saglam, M., Dengiz, O., 2012. Influence of selected land use types and soil texture interactions on some soil physical characteristics in an alluvial land. Int. J. Agro. Plant Prod. 3, 508–513.
Salako, F.K., Tian, G., Kirchhof, G., Akinbola, G.E., 2006. Soil particles in agricultural landscapes of a derived savanna in southwestern Nigeria and implications for selected soil properties. Geoderma 137, 90–99.
Sandoval, M., Dörner, J., Seguel, O., Cuevas, J., Rivera, D., 2012. Métodos de Análisis Físicos de Suelos. Departamento de Suelos y Recursos Naturales Universidad de Concepción Publicación N° 5. ISBN 956-227-293-1. Chillán, Chile.
Sauer, D., 2010. Approaches to quantify progressive soil development with time in Mediterranean climate—I. Use of field criteria. J. Plant Nutr. Soil Sci. 173, 822–842.
Seguel, O., Horn, R., 2006. Structure properties and pore dynamics in aggregate beds due to wetting-drying cycles. J. Plant Nutr. Soil Sci. 169, 221–232.
Seguel, O., García de Cortázar, V., Casanova, M., 2003. Variación en el tiempo de las propiedades físicas de un suelo con adición de enmiendas orgánicas. Agricultura Técnica 63 (3), 287–299.
Seguel, O., Farias, E., Luzio, W., Casanova, M., Pino, I., Parada, A.M., Videla, X., Nario, A., 2015. Physical properties of soil after change of use from native forest to vineyard. Agro Sur J. 43 (2), 23–39.
Sharu, M.B., Yakubu, M., Noma, S.S., Tsafe, A.I., 2013. Land evaluation of agricultural landscape in Dingyadi district, Sokoto state, Nigeria. Niger. J. Basic Appl. Sci. 21, 148–156.
Silva Cruz, J., de Assis Júnior, R.N., Rocha Matias, S.S., Camacho-Tamayo, J.H., 2011. Spatial variability of an Alfisol cultivated with sugarcane. Ciencia e investigación Agraria 38 (1), 155–164.
Soto, L., Galleguillos, M., Seguel, O., Sotomayor, B., Lara, A., 2019. Assessment of soil physical properties' statuses under different land covers within a landscape dominated by exotic industrial tree plantations in south-Central Chile. J. Soil Water Conserv. 74 (1), 12–23.
SSSA, 2008. Methods of Soil Analysis Part 5— Mineralogical Methods. SSSA Book Ser. 5.5. SSSA, Madison.
Szabó, B., Szatmári, G., Takács, K., Laborczi, A., Makó, A., Rajkai, K., Pászto, L., 2019. Mapping soil hydraulic properties using random-forest-based pedotransfer functions and geostatistics. Hydrol. Earth Syst. Sci. 23 (6), 2615–2635.
Taylor, J., Jacob, F., Galleguillos, M., Prévot, L., Guix, N., Lagacherie, P., 2013. The utility of remotely-sensed vegetative and terrain covariates at different spatial resolutions in modelling soil and watertable depth (for digital soil mapping). Geoderma 193-194, 83–93.
Thompson, J.A., Moore, A.C., Austin, R.F., Yewtukhiw, E.M., 2006. Multiscale terrain analysis to improve landscape characterization and soil mapping. 18th World congress of Soil Science Philadelphia Pennsylvania USA.
Tian, Y.C., Gu, K.J., Chu, X., Yao, X., Cao, W.X., Zhu, Y., 2014. Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. Plant Soil 376 193–109.
Tsegaye, T., Hill, R.L., 1998. Intensive tillage effects on spatial variability of soil physical properties. Soil Sci. 163 (2), 143–154.
Ubalde, J.M., Sort, X., Poch, R.M., Porta, M., 2007. Influence of edapho-climatic factors on grape quality in Conca de Barbera vineyards (Catalonia, Spain). J. Int. Sci. Vigne et du Vin 41, 33–41.
Unamunzaga, O., Besga, G., Castellón, A., Usón, M.A., Chéry, P., Gallejones, P., Aizpurua, A., 2014. Spatial and vertical analysis of soil properties in a Mediterranean vineyard soil. Soil Use Manag. 30, 285–296.
Uribe, J.M., Cabrera, R., La Fuente, A., Paneque, M., 2012. Atlas Bioclimático de Chile. Universidad de Chile, Santiago de Chile.
Van derWerff, H.M.A., Van der Meer, F.D., 2016. Sentinel-2A MSI and Landsat 8 OLI provide data continuity for geological remote sensing. Remote Sens. 8, 883.
Van Leeuwen, C., Friant, P., Chone, X., Tregoat, O., Koundouras, S., Dubourdieu, D., 2004. Influence of climate, soil and cultivar on terroir. Am. J. Enol. Vitic. 55, 207–217.
Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Marí, J., Clevers, J.G., Camps-Valls, G., et al., 2015. Experimental sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – a comparison. ISPRS J. Photogramm. Remote Sens. 108, 260–272.
Wang, W., Yao, X., Yao, X., Tian, Y., Liu, X., Ni, J., Cao, W., Zhu, Y., 2012. Estimating leaf nitrogen concentration with three–band vegetation indices in rice and wheat. Field Crop Res. 129, 90–98.
Warrick, A.W., 2002. Soil Physics Companion. CRC Press, Boca Raton, USA.
Xing-Yi, Sui, Y.Y., Zhang, X.D., Meng, K., Herbert, S.J., 2007. Spatial variability of nutrient properties in black soil of Northeastz China. Pedosphere 17 (1), 19–29.
Yoo, K., Amundson, R., Heimsath, A.M., Dietrich, W.E., 2006. Spatial patterns of soil organic carbon on hillslopes: integrating geomorphic processes and the biological c cycle. Geoderma 130, 47–65.
Yoon, Y., Kim, J., Hyun, S., 2007. Estimating soil water retention in a selected range of soil pores using tension disc infiltrometer data. Soil Tillage Res. 97, 107–116.
Zhou, H., Chen, Y., Li, W., 2010. Soil properties and their spatial pattern in an oasis on the lower reaches of the Tarim River, Northwest China. Agric. Water Manag. 97, 1915–1922.