UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS AGRONÓMICAS

ESCUELA DE POSTGRADO

PREDICIENDO DE LA VARIABILIDAD ESPACIAL DE LAS PROPIEDADES DEL SUELO SELECCIONADAS MEDIANTE EL MAPEO DIGITAL DE SUELO EN UN VIÑEDO DEL SECANO DE CHILE CENTRAL

PREDICTING SPATIAL VARIABILITY OF SELECTED SOIL PROPERTIES USING DIGITAL SOIL MAPPING IN A RAINFED VINEYARD OF CENTRAL CHILE

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN MANEJO DE SUELOS Y AGUAS

LWANDO MASHALABA

Directores de Tesis
DR. MAURICIO GALLEGUILLOS TORRES
DR. OSCAR SEGUEL SEGUEL

Profesores Consejeros
DR. MARCO PFEIFFER JAKOB
DR. CRISTIAN MATTAR BADER

SANTIAGO DE CHILE

2019

UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS AGRONÓMICAS
ESCUELA DE POSTGRADO

PREDICIENDO DE LA VARIABILIDAD ESPACIAL DE LAS PROPIEDADES DEL SUELO SELECCIONADAS MEDIANTE EL MAPEO DIGITAL DE SUELO EN UN VIÑEDO DEL SECANO DE CHILE CENTRAL

Tesis presentada como parte de los requisitos para optar al Grado de Magíster en Manejo de Suelos y Aguas

LWANDO MASHALABA

	Calificaciones
DIRECTORES DE TESIS	
Mauricio Galleguillos Torres Ingeniero Agrónomo, M.Sc, PhD.	6,8
Oscar Seguel Seguel Ingeniero Agrónomo, Dr. Cs. Agr.	7,0
PROFESORES CONSEJEROS	
Marco Pfeiffer Jakob Ingeniero Agrónomo, M.Sc, PhD.	6,6
Cristian Mattar Bader Ingeniero en Recursos Naturales Renovables, M.Sc, PhD.	6,5

Santiago, Chile
2019

Acknowledgements

Firstly, I 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" for financing my research.

Taking a look back at the end of this long journey, I cannot help but wonder if I would be standing at the finishing line, which once seemed so far away, without the people who have been there with me along the way to give me direction, guidance, and encouragement. They have carried me through the times where it just seemed too hard to go on.

Chilean International development Cooperation Agency (AGCID), for the financial support that enabled me pursue and accomplish my Master of Science degree studies.

Department of Rural Development and Agrarian Reform (DRDAR) from South Africa for allowing me to resign and pursue my MSc. degree in Chile.

Prof. Mauricio Galleguillos my supervisor for his enduring patience, full commitment to my success, and unconditional efforts of providing the best possible academic training have been the source of my conviction, confidence, and motivation to reach higher standards. As a good teacher and mentor, he saw my strengths and weaknesses, and pushed me to reach my full potential.

Prof. Oscar Seguel my co-supervisor stood firmly by me during the times of confusion and frustration. He devoted a tremendous amount of time and effort to help me redefine a set of achievable goals and then made sure every step I took was a step closer to those goals. His painstaking attention to detail and accuracy in research, specifically during the development of this dissertation are two qualities I will spend the rest of my career trying to achieve. The staff members in the Faculty of Science Agronomy, especially Mr H. Perez, Mr F. Montoya, Cristobal Puelma and Mr R. Berthin Ticona for their assistance during field work and laboratory experiments and Javiera Poblete Olivares for her assistance during the statistical modelling.

Lastly, my thanks and appreciation should go to my Uncle, Dr. Yandisa Mashalaba, for his much needed love and support through good and difficult times and relatives and friends for whatever they have done and patience to help me accomplish this work.

Dedicated this thesis to my family:

Grandmother, dad, and Evelyn Villanueva Tancara.

Who have filled my cup with joy and made all my trials worthwhile.

Table of content

List of tables

List of figures

1. LITERATURE REVIEW

1.1 State of the art of digital soil mapping (DSM)

1.1.1 Soil data and auxiliary information

One cannot begin to talk about digital soil mapping without the acknowledgement of researchers who introduced the term in the early 1980s (Burgess and Webster, 1980; McBratney *et al*., 1982), which lead to Digital soil mapping (DSM) that was defined by Schull *et al*. (2003) as the computer-assisted production of digital maps of soil types and soil properties by the use of field and laboratory observation methods coupled with spatial and non-spatial soil inference systems, to become a main subject of soil sciences. Several authors stated that digital soil mapping applies pedometric, which is defined as the application of mathematical and statistical methods for the study of the distribution and genesis of soil, to map or predict the spatial and temporal variability of soils (McBratney *et al*., 2000, 2003; Grunwald, 2006). The use of DSM is aimed at extending the functionality of soil information systems from the storage of digitised conventional soil maps to the production of soil maps to meet current and future demand for accurate, up-to-date soil information. Even though DSM is the preferred soil mapping method, studies showed that traditional soil mapping and digital soil mapping are based on a similar principle (Roecker *et al*., 2010). To support this idea, Seeno (2015) concluded that both types of mapping can be interpreted as predictive because both look to the known relationship between soil and the more visible environmental variables to make inference about the properties and behaviour of soil.

Soil mapping, in general, requires (i) a predefined model of soil formation, and (ii) data on soil properties and on other environmental variables that have a significant impact on soil formation and thus on the spatial distribution of the soil properties (Dobos *et al*., 2006). Therefore in this sense, traditional soil mapping and digital soil mapping do not differ much (Roecker *et al*., 2011). The main reason that makes them similar is the fact that both approaches need input data on soil and covariates characterizing the environment where the soil formation takes place. However, the major difference between the approaches is the way in which the model derives the soil information from the input data. The traditional models are based on empirical studies and qualitatively defined correlation that formulates a mental model in the surveyor's mind used to understand and characterize the soil resources (Roecker *et al*., 2011). The main problem with this approach is that it requires intensive fieldwork, time-consuming and too costly. The decisions are made mainly on the field, where all environmental covariates can be directly observed and information on the soil can be deduced. On the other side, digital soil mapping is based on hard soil data as well. Just like in traditional soil mapping, profile information is required to train the models and to understand the soil resources of the area. One of the advantages of digital soil mapping is that it is cheap, manages time and requires digital data sources as input variables for quantitative models (Dobos *et al*., 2006). The advancement of technology made it easy for digital soil mapping of soil properties. In 1941, Jenny introduced a well-recognized equation that identified 5 major

soil formation factors namely, climate, organisms, relief, parent material and time which was extracted from the hypothesis that was first proposed by Dokuchaev (1883) Eq. 1.

$$S = f\ (cl, o, r, p, t) \tag{1}$$

To successfully predict the soil properties, a surveyor needs good quality and adequate resolution input data. However, Jenny's approach is focused on the prediction of soil physiochemical and biological properties on a given location, but Roecker *et al.* (2011) noted that Jenny did not consider the soil as a continuum, where the soil properties at a given location depends on the geographic position and also on the soil properties at neighbouring location. Some properties are difficult and expensive to measure but can be successfully and accurately predicted from other available properties of the same area or field. It was then that McBratney *et al*. (2003) modified the equation introduced by Jenny (1941) and added two more variables in Equation 1 to make them 7. The new equation is well-known as SCORPAN.

$$Sa\ or\ Scl = f\ (SCORPAN) + \mathcal{E} \tag{2}$$

Where *Sa* is soil attributes, *Scl* is referring to soil class, *S* is available soil properties at the same area, *C* is climate, *O* is organisms, *R* is a relief, *P* is parent material, *A* is age or sometimes referred to as time and *N* is known as geographic position. McBratney *et al*. (2003) also noted that along the way there would be some errors that might be committed when predicting the soil properties and introduced the error in the equation that will account for committed errors during the survey.

1.2 Predicting and modelling of soil variability

The prediction of soil properties at lower depth is not an easy task to do, because some of the predictor variables such as remote sensing spectral indices used in DSM are more correlated with soil properties of the top soil compared to lower (Taylor *et al.*, 2013). Incorporating spatial dependence in statistical prediction models for categorical soil attributes such as soil types and/or texture class is not as straightforward as for continuous attributes. Therefore, many digital soil mapping application focus on methods that ignore spatial dependence in the categorical soil variable, such as multinomial logistic regression (Debella-Gilo and Etzelmüller, 2009; Hengl *et al.*, 2007; Kempen *et al*., 2009), and classification and regression trees (Minasny and McBratney, 2007; Nelson and Odeh, 2009; Stum *et al.*, 2010). However, variability of soil properties can also be quantified using classical statistics, which involves the determination of mean, range and coefficient of variation (CV); on another hand, geostatistical tools are also used, which for example determine: semivariogram, autocorrelation, coefficient of determination, cross semivariogram, kriged and co-kriged maps (Jabro *et al.*, 2010). McBratney *et al.* (2003) stated that different interpolation techniques have been used with varying degrees of success in order to create more accurate property maps.

On the other side, an experimental variogram is usually used to measure the variability between pairs of points at various distances (Deutsch and Journel, 1998). This is the reason that the correlations at various distances can be established to come up with values non-sampled field location (Tola *et al*., 2017). However, the work conducted by Taylor

et al. (2013) proved that predicting regression models based on remote sensing information that describes topography and vegetation vigour can also be used to characterize the spatial variability of soil physical properties with success. Akpa *et al.* (2014) motivated the use of DSM and regards it as a promising approach that combines field observations and laboratory data, with remote and proximal soil observations to the spatial prediction of soil attributes. Many scientists are now considering the use of DSM, which lead to McBratney *et al*. (2003) formalizing DSM in the now widely used Scorpan model.

Bacis Ceddia *et al*. (2009) emphasized that soil sampling allows the characterization of several soil attributes which may be estimated at unsampled sites through existing models. However, Siquiera *et al.* (2014) stated that sampling density also plays an important role in obtaining representative information to make a proper decision about the variability of properties across the landscape and or field. Isaaks and Srivastava (1989) concluded that deterministic models are considered to be more appropriate when there is enough information on physical and chemical properties. However, Webster (2000) could not agree nor deny that, instead, he stated that use of deterministic models for the understanding of both soil formation and their attributes does not result in accurate estimation, because of the great complexity among soil properties.

Vieira (2000) further added that geostatistic focus is based on a probabilistic model, and has been successfully used in soil science for quantitative description of spatial variability, which may support predictions about the phenomena or property investigated. However, Brown *et al.* (2006) also motivated the use of remote sensing images of bare soil and spectral reflectance of soil samples to accurately and rapidly estimate soil properties across the field. Bacis Ceddia *et al*. (2009) indicated that it is very important for researchers to identify landscape features when using soil mapping procedures to predict soil properties. Much work has been done trying to correlate landscape features (altitude, slope, and shape) with physical soil properties (Kreznor *et al.*, 1989; Pachepsky *et al*., 2001; Sobieraj *et al*., 2002; Rezaei and Gilkes, 2005). For example, in a study done by Novaes Filho *et al.* (2007) spatial variability of soil colour and texture were considered feasible to be used in models of DSM in Southern of Amazon.

McBratney *et al*. (2000) noted that prediction can be made at unobserved locations using the environmental variables at those locations. With this technique, soil properties can be predicted using their interrelationships with the environmental covariates, such as digital elevation models, remotely sensed data and physical attributes obtained through laboratory analysis of soil samples (Mayr *et al*., 2008; Odeh and McBratney, 2000). Akpa *et al*. (2014) also indicated that models used in spatial variability studies are often based on compositional ordinary kriging (COK), regression kriging (RK), multiple linear regression (MLR), generalized linear model (GLM), regression tree models (RTM) and recently random forest model (RFM) with varying scale. However, kriging is known to have some theoretical and practical shortcomings such as probabilities that are outside the interval (D'Or and Bogaert, 2004; Papritz, 2009).

From all the methods used to characterize and quantify the spatial variability of soil properties within the agricultural fields, many authors suggested that Geostatistics has proven to be useful, and has been successfully used by soil scientists and engineers (Webster and Oliver, 2001; Iqbal *et al*., 2005). Complementary, new methods of determining spatial variabilities such as DSM and remote sensing have been approved

and have attracted the attention of researchers (Taylor *et al*., 2013). The literature further indicates that semivariogram and cross-semivariogram have been continuously used to characterize and model spatial variance of the data to assess how data points are related with separation distance in the field (Nielsen and Wendroth, 2003). To prove the success of geostatistics in the field of soil science, Adhikary *et al*. (2009) applied with success the Geostatistics to study and analyse the spatial behaviour of soil texture and provided the maps which assisted the farmers to design the land management and the obtained results were convincing.

Due to DSM's increasing popularity, the number of models and the procedures for optimizing modelling performance has increased (Minasny and McBratney, 2016). Some models applied to soil mapping include the following: geostatistical (Kempen *et al*., 2012); fuzzy membership (Nolasco-Carvalho *et al*., 2009; Taghizadeh-Mehrjardi *et al*., 2015; Rizzo *et al*., 2016); and *clorpt* techniques based on environmental correlation, usually applying, CART, and regression models (Brungard *et al*., 2015; Heung *et al*., 2016; Chagas *et al*., 2017). Breiman (2001) used CART to develop the RF method that features enhanced accuracy without substantially increasing the number of calculations. RFM grows trees to the maximum depth possible, but the degree to which splits at nodes are randomly selected can be controlled (Svetnik *et al*., 2003) and it inherits the advantages of CARTs for low calculation load and high explanatory power. Amongst its advantages, Iverson *et al.* (2008) stated that RFM can also be used to determine the relative importance of input variables and that makes it a preferred algorithm to predict the soil variability. However, its accuracy of the models depends on the quality of information available to perform the interpolation of unsampled areas, in this sense, it is advisable to use variables that showed positive influence with the predicted property.

1.3 Selection of input variables

Selection of input variables for models is one of the critical tasks to do, as to perform a representative mapping of soil properties, it is important to select the most appropriate and relevant environmental covariates. Also, environmental variables connected to soil processes are also connected to each other. Swanson *et al*. (1988) stated that just as geomorphology determines soil properties, these environmental variables determine the pattern of vegetation across the landscape. The authors made an example of vegetation that is influenced by both soil properties and topographic position but also by climate, and by geology. The traditional survey begins by looking at the coarsest scale relevant to the area of interest. Dent and Young (1981) noted that it is necessary to identify the types of landscapes that make up the area, and then the landforms within those landscapes.

This strategy was later applied by other researchers in DSM through the use of terrain attributes determined from digital terrain model (DTM) and imagery across a range of scale sizes to predict variability of soil physical properties across the field (Miller *et al*., 2015). Based on the study conducted by Miller *et al*. (2015) it is very clear that the interconnectedness of the environmental variables used to predict soil properties allows for substitution with data mining models when there is limited information available for different environmental variables. Several studies noted data source for generating the environmental variables includes remote sensing data (Poggio *et al*., 2013; Toghizadih-

Mehrjardi *et al*., 2014) such as normalized difference vegetation index (NDVI) which is directly related to organic material.

The generated environmental variables are then directly used as inputs in machine learning tools such as regression models. However, till to date, it is not yet clear how many covariates can be used as predictors. It is also noted or suggested that a number of covariates should depend on the number of sampling points. The contribution of each variable is measured and the best-performed variables will then be included in the final model. Poggio *et al.* (2013) noted that being able to use remote sensing data for digital soil mapping makes it an incredibly powerful tool to apply in areas that have little field data. However, remote sensing alone is unable to give the information for subsoil properties.

1.4 Data sources of environmental covariates for DSM

1.4.1 Remote sensing data

Satellite remote sensing data has emerged as a vital tool in soil resources survey and generation of information which helps to evolve the optimum land use plan for sustainable development at a scale ranging from regional to micro level (Singh, 2016) and also helps to obtain data from an area that cannot be sampled and it being cost effective. The surface features reflected on satellite image provide enough information to accurately delineate the boundaries which are accomplished effectively through the systematic interpretation of satellite imageries (Singh, 2016). The literature review by McBratney *et al.* (2003) shows that soil properties that govern the spectral reflectance are colour, texture, mineralogy, organic matter, water content and some physicochemical properties.

Some properties are directly related to the surface colour and thus relatively easy to map when the soil is bare and visible spectra is used to detect colour. Dobos *et al*. (2006) gave an example with iron-oxide and organic matter content, the soil water contents and texture as good examples of direct properties. The prediction of subsoil attributes is generally more difficult than topsoil because many environmental covariates used DSM only generate a topsoil response (e.g. visible and near infrared, imagery and gamma radiometry) (Taylor *et al*., 2013). However, other soil features like many chemical properties of deeper horizons can be detected only indirectly, through the type and the condition of the surface vegetation. As alternatives, many types of proximal sensors (electromagnetic magnetic induction and ground penetrating radar) have been used for soil studies (Lombardi and Lualdi, 2019). However, these sensors are said to produce a signal that is difficult to interpret or to deconstruct into individual subsoil attributes or different subsoil layer responses.

Another approach to deriving subsoil attributes is to incorporate a vegetation – soil inference system into the mapping. McKenzie and Ryan (1999) stated that compound remote sensing indices such as NDVI which generally reflects biomass status, have been shown to correlate well with the distribution of the organic matter.

Based on the study by Taylor *et al*. (2013) vegetative covariates have generally only been derived from the visible and near infrared regions of the electromagnetic spectrum (EMS) and usually only used empirically with models. A logical further development was to combine DEM-derived and remote sensing data to improve prediction models (Dobos *et al.,* 2000). The use of combined terrain data and remote sensing imagery has been especially interesting for medium scale-surveys with grid resolutions from 20 – 200 m, although there has been an increasing number of field-site studies (Dobos et al., 2006). If they are appropriately selected, the remote sensing images can reflect the overall environmental conditions, types and condition of vegetation influenced by the soil properties, surface roughness and moisture content.

1.4.2 Digital elevation models (DEM) and terrain parameters

Previous studies stated that terrain attributes derived from digital elevation models are frequently used in digital soil mapping as auxiliary covariates in the construction of soil prediction models (Kempen *et al*., 2011; Caten *et al*., 2013; Teske *et al*., 2014). In recent years, it has been suggested that DEMs and information extracted from it may be limited with regards to the spatial resolution and error magnitude, and can differ in the behaviour of terrain features (Moura-Bueno *et al*., 2016). Hutchinson and Gallant (2000) stated that a DEM can be generated from topographic data of field survey, interpolation of vector bases (e.g. elevation points) extracted from topographic maps, pairs of stereoscopic images on aerial photographs and by satellite images obtained by optical sensors and or orbital radar. Several authors who worked closely with digital terrain model gave an idea that extracted parameters can be used for modelling of soils, vegetation, land use and geological features (Minella and Merten, 2012; Caten *et al*., 2011 and Samuel-Rosa *et al*., 2013).

The basic terrain features (elevation, slope gradient, aspect, flow direction and curvature) are relatively simple and easy to derive and they influence the prediction of soil properties (McBratney *et al*., 2003; Smith *et al*., 2006). Several geomorphometric parameters, such as slope gradient, aspect, profile and elevation can be computed from DEMs and the study that was conducted by Moore *et al*. (1993) found significant correlations between quantified terrain attributes and measured soil properties

1.5 Problem statement

In the Mediterranean South of Chile, where the vine has a high potential for quality wine production without the need of irrigation, there are many areas with no to very limited publicly available soil spatial variability data at detailed scale. The spatial variability of soil physical properties is regarded as one of the problems affecting water distribution and the crop development, especially in rainfed conditions where environmental conditions (evapotranspiration), terrain attributes (slope gradient, aspect and curvature) and soil properties (texture, depth, soil water retention, organic matter, etc.) controls the availability of water in soil profile within the field. Considering the spatial variability is one of the major soil characteristics, variation in soil properties appears to be a key driver of vineyard yield and quality variability (Bramley and Lamb, 2003). Wine is one of the

most development axes in Chile, reaching a total of 120,000 ha planted with vines (INE, 2008). The strong spatial variability of physical properties that controls the distribution of water will lead to the growth variability of the vine in some areas within the field. Some areas of the field can suffer and excessive water stress which then affects the vine development because of the variation in texture and depth which influence the water holding capacity (WHC) and water uptake (van Leeuwen *et al*., 2004). Poor vine development affects the quality and price of wines produced in such rainfed areas.

Chile is approaching its maximum capacity regarding the allocation of natural resources, water and soils for irrigation. Chile is a country with diverse geography and landscape formation and therefore it has limited space for agriculture because of mountains and hills, which then force removal of trees and natural vegetation in the hills in order to gain space for production of avocados and vines, among others, as it is the case in the Maule Region and other regions (Ferreyra *et al*., 2001; Seguel *et al*., 2015). The topography and particularly hillside aspect influence the microclimate of the area, hence the variation in soil properties across the landscape influences the distribution of water within the area during the rainfall distribution, which then leads to uneven plant growth.

Vrsic *et al*. (2011) noted that relief plays an important role in the vineyard because, for centuries, these have mainly occupied hillside as it is the case in most vineyards in Chile and other countries with a geographical formation similar to Chile. The soils from the upper slopes are frequently removed by erosion and deposited in the low-lying slopes, which leads to the variation in soil depth, texture and water content. Little work has been done at high resolution on spatial variability of soil properties in the vineyard at a lower than 0.3 m and that makes it difficult for those responsible for the particular site to model the site-specific management and making sure of the uniform quality of vines.

Adhikary *et al*. (2009) postulated that soil water and texture are the most important soil physical properties that are known for governing nearly all of the other physical properties. Therefore high-resolution maps of texture are essential for hydrological and ecological modelling as well as agricultural management (Zhao *et al*., 2009). This idea gives evidence that combined variation of texture and soil water across the field would definitely lead to spatial variation in organic matter (Kong *et al*., 2009), penetration resistance and pedogenesis (Western *et al*., 2003). Based on documented studies, the temporal and spatial variability of soil texture and soil water may lead to structural differences in soil quality (Kettler *et al*., 2001) and hydrologic cycle (Western *et al*., 2003) across the field which all then lead to variability in crop quality or development. Therefore quantifying the variability of these properties using appropriate methods would then provide some assistance in developing, modelling and designing site-specific management, which includes water management, prevention of soil degradation and avoids environmental problems.

1.6 Justification

High-quality soil data can be used by multiple agencies, landowners and managers to aid in soil conservation and restoration efforts. The main reason for the development of this study is to provide the necessary information about the methods that can be used to map and predict the soil physical properties by means of knowledge-based DSM approach.

Characterizing the spatial variability of soil physical properties across the hillsides or catena is one of the useful focus for modelling, planning and managing the soils through the site-specific management. The use of spatial variability characterization methods will help to construct the maps of the areas across the hill indicating those areas that require the attention. It would be potentially used to describe spatial patterns by remote sensing information and predict the values of soil properties at unsampled locations using predicting regression models such as random forest. Thus, the potential has existed for vines and wine producers to acquire detailed geo-referenced information about vineyard performance and to use this information to tailor production of both vines and resultant wines according to expectations of vineyard performance, and desired goals in terms of both yield and quality (Bramley and Proffitt, 1999). The produced predictive soil maps provides important information about the properties and condition of the land and can be used as a guide for land use planning. In addition they can be used as part of the land management guide, such as monitoring the land drainage capabilities of the area.

CHAPTER 2. PREDICTING SPATIAL VARIABILITY OF SELECTED SOIL PROPERTIES USING DIGITAL SOIL MAPPING IN A RAINFED VINEYARD OF CENTRAL CHILE

RESUMEN

Las propiedades físico-hidráulicas del suelo influyen en el comportamiento del viñedo, por lo tanto, el conocimiento de su variabilidad espacial es esencial para tomar decisiones de manejo. El estudio tuvo como objetivo modelar y mapear las propiedades físicas del suelo mediante un enfoque basado en conocimiento de mapeo digital del suelo. Este estudio tuvo como objetivo predecir la variabilidad espacial de las propiedades físicas del suelo para un viñedo de secano utilizando datos de detección remota y el modelo de Random Forest (RFM). El sitio de estudio es un viñedo de 29 ha, ubicado cerca de la ciudad de Cauquenes en la Región del Maule de Chile. Se dividió en cuadrículas regulares (60 x 60 m^2), recogiendo y georreferenciando muestras de suelo perturbadas y no perturbadas a tres profundidades (0-20 cm, 20-40 cm y 40-60 cm). En 62 ubicaciones se analizaron los perfiles del suelo a 60 cm de profundidad para: distribución del tamaño de partícula, materia orgánica, densidad aparente, porosidad total, densidad de partículas, índice de estabilidad estructural, índice de repelencia, estabilidad de microagregados, resistencia a la penetración, resistencia al corte, conductividad hidráulica saturada, características de retención de agua del suelo, distribución de poros por tamaño y estabilidad de macroagregados. La estadística descriptiva mostró baja a muy alta variabilidad en las propiedades. Utilizamos un algoritmo de RF para vincular las covariables ambientales que describen los factores de formación del suelo y once propiedades del suelo seleccionadas en los intervalos de profundidad. La precisión del modelo se midió por R^2, nRMSE, RMSE y sesgo. El modelo RF para arcilla, arena, relación de dispersión, y punto de marchitez permanente (PWP) funcionó adecuadamente, aunque los modelos no pudieron predecir de manera confiable el AWC, la porosidad total y el contenido de materia orgánica. Los mapas de predicción mostraron que los valores más altos de arcilla, capacidad de campo (FC), y PWP se encontraban en el lado occidental del campo donde hay una elevación más baja en el paisaje. Hubo un mejor desempeño de predicción en el contenido de arcilla de la superficie que en el subsuperficial (por ejemplo, R^2 de 0,66; RMSE de 4,3% para 0-20 cm y R^2 de 0,51; RMSE 5,6% a 40-60 cm). En general, la distribución del tamaño de partícula muestra variaciones marcadas en el viñedo con un mayor contenido de arena en comparación con el contenido de limo y arcilla. Hubo un menor rendimiento de predicción en de la relación de dispersión de superficie que la subsuperficial (por ejemplo, R^2 de 0,49; RMSE de 10,1% para los 0-20 cm y R^2 de 0,81; RMSE de 8,7% a 40-60 cm). El modelo de RF sobreestimó las áreas con valores bajos y Sobreestimó las áreas con valores altos. El análisis de importancia variable mostró que el Índice de Posición Topográfica, el Índice de Humedad Topográfica, el aspecto, el Factor de longitud de pendiente, el área de captación modificada, el gradiente de pendiente del área de captación y la curvatura longitudinal se encontraban entre las covariables que más influían en la predicción.

Palabras clave: mapeo digital del suelo, propiedades del suelo, viñedo, modelo de bosque aleatorio, covariables ambientales.

ABSTRACT

Soil physico-hydraulic properties influence vineyard behavior, therefore the knowledge of their spatial variability is essential for making vineyard management decisions. The study aimed to model and map soil physical properties by means of knowledge-based digital soil mapping approach. The study site, a vineyard of 29 ha located near Cauquenes city in Maule Region, Chile, was divided in a regular square grid (60 x 60 m^2), collecting and georeferencing disturbed and undisturbed samples from three different soil depths (0-20 cm, 20-40 cm and 40-60 cm). At 62 locations soil profiles to 60 cm were analysed for particle size distribution, soil organic matter, bulk density, total porosity, particle density, structural stability index, repellency index, microaggregate stability, penetration resistance, shear strength, saturated hydraulic conductivity, soil water retention characteristics, pore size distribution and macroaggregate stability. The descriptive statistics showed low to very high variability within the field. We used a RF algorithm to link environmental covariates describing soil forming factors and eleven selected soil properties at three depth intervals. The model accuracy was measure by R^2, nRMSE, RMSE and bias. RF model of clay, sand, dispersion ratio, and permanent wilting point (PWP), performed well, although the models could not reliably predict the available water content (AWC), total porosity and organic matter. Prediction maps showed that the highest amounts of clay, FC, and PWP were on the western side of the field where there is lower elevation in landscape. There was a better performance in the upper depth intervals than the lower depth intervals (e.g., R^2 of 0.66; RMSE of 4.3 % for clay content at 0–20 cm and R^2 of 0.51; RMSE of 5.6 % at 40-60 cm). Overall, the particle size distribution show marked variations across the vineyard with a higher sand content compared with silt and clay contents. There was a better performance in the lower depth intervals than the upper depth intervals (e.g., R^2 of 0.49; RMSE of 10.1 % for dispersion ratio at 0–20 cm and R^2 of 0.81; RMSE of 8.7 % at 40-60 cm). RF model overestimated areas with low values and underestimated areas with high values. Analysis of variable importance showed that Topographic Position Index, Topographic Wetness Index, aspect, Length Slope Factor, modified catchment area, catchment slope, and longitudinal curvature were the covariates with the highest influence on the prediction.

Keywords: Digital soil mapping, soil properties, Vineyard, Random Forest model, Environmental covariates.

2. INTRODUCTION

Soil mapping is regarded as key for guiding decision makers in natural resource assessment, environmental modelling, and land use studies. However, it requires the knowledge and experience of a senior pedologist 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 properties (Ma *et al*., 2017). This information is being generated through the application of digital soil mapping (DSM) methods. DSM is regarded as the creation and 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, national or a global scale. The conducted review on the use of DSM for soil mapping in Brazil (Caten *et al*., 2012; Heung *et al*., 2014; Lacoste *et al*., 2011), indicated that approaches used up to 2011 show three main classification models applied for DSM (artificial neural networks, logistic regression, and decision tree). 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 physical properties results in the change in the values of certain soil properties over space (Ettema and Wardle, 2002). According to recent studies, understanding of spatial variability of soil physiochemical 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).

Detailed digital soil maps showing the spatial heterogeneity of soil properties consistent with the landscape are required for site-specific management of plant nutrients, land use planning and process-based environmental modelling. Brady and Weil (2008) stated that spatial variability of soil properties can be categorized into three broad classes namely; large-, medium-, and small-scale. The variability of soil properties across landscapes and regions is attributed to the medium and large-scale, respectively. However, several authors noted that there is still few studies describe spatial variability of multiple soil properties and their inter-relationship at a landscape scale (Bruland *et al*., 2006; Liu *et al*., 2010; Paz-Gonzalez *et al*., 2000; Rivero *et al*., 2007; Stutter *et al*., 2009).

Through the researches that have been conducted on this subject, Pereira and Ubeda (2010) regard mapping of spatial variability of soil properties as one of the important tools to understand how processes change in space and time. From the study that was conducted in the soils of North-Eastern of Sao Paulo state, Souza *et al.* (2004) noticed that small variations in the landscape form defined different spatial variability in the soil physical properties. These results were in agreement with the work reported by Souza *et al*. (2003), which evaluated the effect of landforms on the anisotropy of soil physical properties, observed higher spatial variability of soil physical properties in the concave landscape landform when compared to a linear one. Tsegaye and Hill (1998) conducted

the study on the effect of intensive tillage on spatial variability of soil, the results showed that all the measured soil properties except saturated hydraulic conductivity (Ksat) showed weakly spatial dependency for 6 – 9 cm and moderately spatial dependency for 27 – 30 cm soil depths. Serrano *et al*. (2014) noted that heterogeneity in soil properties with depth and across the landscapes can be accounted for by several interacting factors that operate with different intensities and at different scales and acting simultaneously. This all proves the fact that spatial variability can be observed both vertically and horizontally which completes the variability across the profile.

It is well documented that particle size, soil water content, plant available water (PAW) and penetration resistant showed a clear horizontal spatially variability structure capture by soil map units (Tsegaye and Hill, 1998). There have been several talks about the magnitude of the spatial variability of Ksat across the agricultural fields from which, Biggar and Nielsen (1976) reported Ksat as the one with the highest variability in the field. For example, Jury and Horton (2004) indicated the values of coefficient of variation for Ksat in the range of 50 - 300%.

Chile is regarded as the world's eighth largest producer of wine and fifth largest exporter (Felzensztein *et al*., 2011) and the Chilean wines are positioned as the country's 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 soil 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. Based on the statistics given by Ferrero *et al*. (2005) in highly mechanized viticulture, the number of tractor passes per year is estimated to be up to 20 – 30 in traditional cultivation.

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 behaviour is closely related to water uptake (van Leeuwe *et al*., 2004).

The vineyards in Chile are largely distributed in hilly terrains and in such areas, relief controls the distribution of soil types in a landscape and is regarded as a source of variability (Jenny, 1980; De Gryze *et al*., 2008; Ruth and Lennartz, 2008), it has been reported by many authors as a source of spatial variability. The characterization of spatial variability of soil properties is essential in making site-specific management and other decision within the agricultural field. There are certain tools that were developed to help in determining the magnitude of spatial variability of soil and crop yield; classical statistics, which includes the determination of mean, range, standard deviation,

coefficient of variation and skewness were reported to be an important tool that can be used (Ramzan *et al*., 2017). However other researchers reported the importance of geostatistical methods in the implementation of site-specific management systems (Najafian *et al*., 2012).

On the other hand, Taylor *et al*. (2013) proved that predicting regression models based on remote sensing information can also be used to characterize the spatial variability of soil physical properties with success. Akpa *et al*. (2014) motivated the use of DSM and regards it as a promising approach to the spatial prediction of soil attributes. Moreover, remotely sensed imagery can be used as a data source supporting DSM (Ben-Dor *et al*., 2008; Slaymaker, 2001). One of the reasons that make remote sensing data to be more important in DSM technique is the fact that it facilitates mapping inaccessible areas by reducing the need for extensive time-consuming and costly fieldwork. Several authors noted terrain attributes such as elevation, plan and profile curvatures, relative slope position influence soil properties, and classification (Mehnatkesh *et al*. 2013; Umali *et al*. 2012) and landscape hydrology.

In addition, various papers in the literature have emphasized the importance of integrating terrain attributes with soil and crop variables when modeling yield and soil parameters (Beaudette *et al*. 2013; Brown *et al*. 2004; Norouzi *et al*. 2010; Zhang *et al*. 2012). Therefore DSM offers the use of soil, digital terrain model, and remote sensing data to map various soil properties at low costs. In addition, remote sensing helps to overcome errors in locating and plotting soil boundaries as also in generating soil map of inaccessible areas. Kudrat *et al*. (1992) explained that the dynamic inter-relationship between physiography and soil is utilized while deriving information on soil from satellite data. McBratney *et al*. (2003) added that different interpolation technique have been used with varying degree of success in order to create more accurate soil property maps.

2.1 Research question

Is it possible to generate successfully predictive maps of spatial variability of soil physical properties within a vineyard using digital soil mapping methods?

2.2 General and specific objectives

2.2.1 General objective

To predict and characterize spatial variability of soil physical properties in a rainfed vineyard situated in a hill of granitic catena at central Chile using DSM.

2.2.2 Specific objectives

- To study the impact of auxiliary data, such as terrain attributes and satellite images for the prediction of spatial variability of soil properties.

- Determine importance of each environmental covariates on spatial variability of soil physical properties.
- To model and map soil physical properties by means of knowledge-based digital soil mapping approach.
- Establish correlations amongst the soil physical properties.

3. MATERIAL AND METHODS

3.1 Site description

The study was carried out in a 29 ha vineyard situated 20 km away to the west from Cauquenes city, in the Maule Region, central Chile (36º 02' 27.18" SL, 72º 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 ℃ and the minimum between 3 and 12 ℃. 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 (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 x 1 m^2.

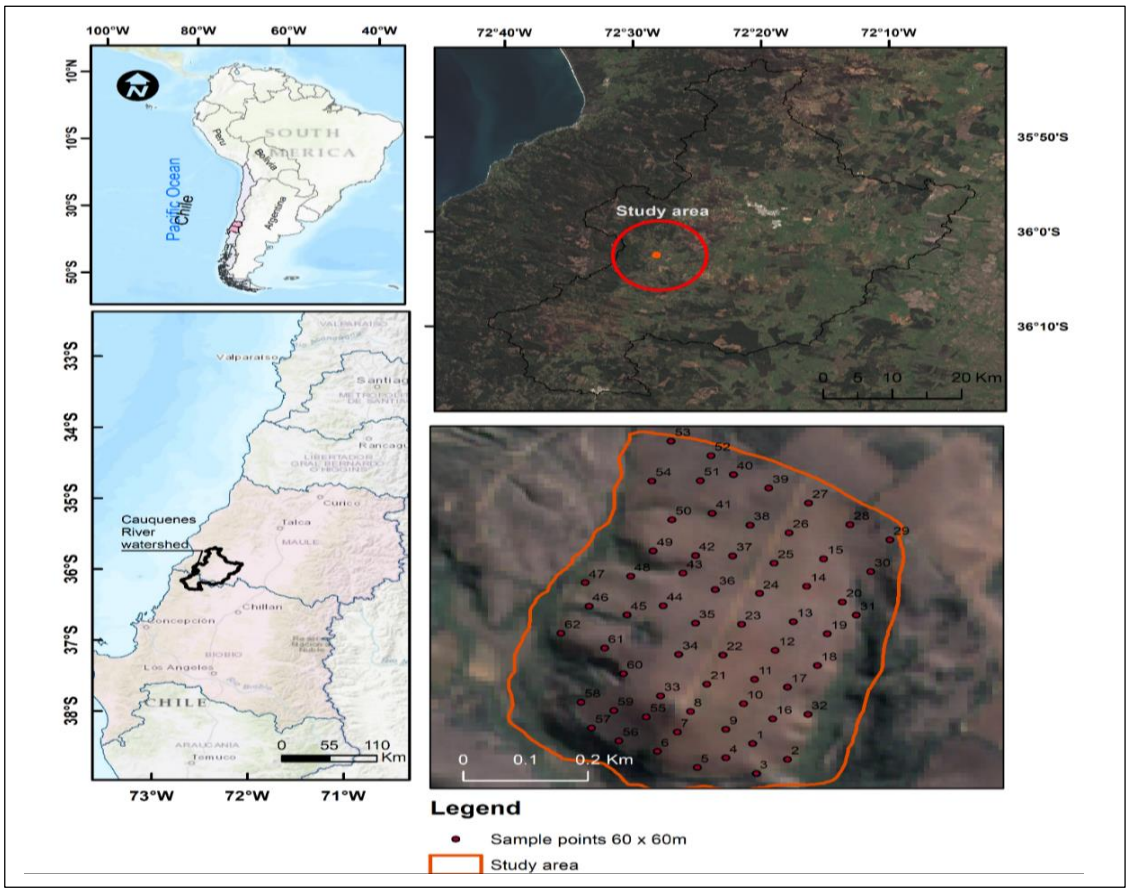


Figure 1. Location of the study area in the Maule Region, Cauquenes watershed at central Chile, with the detail of 62 sampling points (pits) and a distance of 60 x 60 m^2.

3.2 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^2) 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 (± 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 ± 4 m). Undisturbed samples were used for the determination of soil bulk density, the water retention curve at low suctions (< 100 kPa) and Ksat, the water repellency, and aggregate stability. 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, organic matter, particle density, and microaggregate stability. None sieved soil samples were used for water retention relationship at higher suctions (> 100 kPa) using a pressure plate apparatus.

3.3 Laboratory analysis

3.3.1 General soil physical properties

Under laboratory conditions, disturbed soil samples (2 mm sieved, 186 samples in total) were used to determine particle density (PD) and particle size distribution, both 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) and total porosity (*f*) was calculated with the relation:

$$f = [1 - \frac{BD}{PD}] \cdot 100\ [\%] \quad (3)$$

3.3.2 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, water retention curve was done using 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 matric 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, and the water available pores

(WAP, 0.2 – 10 μm) was determined as the difference between water content at – 33 and – 1500 kPa.

3.3.3 Aggregate stability

3.3.3.1 Macro aggregate stability

Macro-aggregate stability was determined by the dry and wet sieving method (Hartge and Horn, 2009). The disturbed samples were air-dried, 200 g subsample of aggregates (dry soil exposed to air and up to 3 cm diameter) were randomly taken from each sample and placed on a set of sieves with diameters of 19, 9.5, 6.35, 4.75, 3.33, and 2 mm. The subsample was sieved for 2 min at a frequency of 60 Hz and the sample water content was determined. Once the dry sieving was finished, the aggregate mass left in each sieve was determined and the partial fraction of each size (ratio between sieve soil mass and total soil mass) was calculated and corrected for its water content to determine the initial distribution of aggregate sizes. A second sieving was then done under water where the set of sieves was shaken with 5 cm upward and downward movements for 5 min at 60 rpm. The soil remaining in each sieve was dried at 105 °C for 16 h or more to determine the aggregate fraction for each size. The variation of weighted mean diameter (Δ WMD) was determined with the results of the dry and wet sieving (Hartge and Horn, 2009) by the equation:

$$\Delta WMD = \sum_{i=1}^{n} \frac{(n_{i_1} \times d_i) - (n_{i_2} \times d_i)}{n_{i_1}} \quad (4)$$

Where *d* is the aggregate mean diameter in range *i* that corresponds to each range of sieves being used, n_{i_1} is the dry mass of the fraction in range *i* for dry sieving, and n_{i_2} is dry mass in range *i* of wet sieving. Therefore, with the results of both sievings (dry and wet), a stability analysis is performed for each natural aggregate size range from a known distribution of aggregates obtained in dry sieving. The lowest ΔWMD value indicates the highest soil stability (Hartge and Horn, 2009).

3.3.3.2 Micro aggregate stability

Micro-aggregate stability was determined by the dispersion ratio (DR) method described by Seguel *et al.* (2003). Two soil samples of 50 g with aggregates between 1 and 2 mm 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 20 cm^3 of sodium pyrophosphate. Both samples were left to rest during the night. The drastically dispersed sample was then mechanically shaken for 10 min in a 75 cycle Hamilton Beach blender. Finally, the samples were poured into 1L measuring 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) using the equation:

$$DR = \frac{(S+C)Sd}{(S+C)dd} \cdot 100 \qquad (5)$$

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.

3.3.4 Organic matter (OM)

Organic matter was determined using the calcination method, according to Schulte *et al*. (1991). Approximately 10 g of air-dried < 2 mm soil fractions were placed into 30 mL crucibles. Then samples were placed in a muffle oven at 360℃ for 16 h. Before weight determination, the samples were placed in a desiccator to reach room temperature. This method calculates OM (%) by comparing the mass of a sample before and after the soil has been ignited, using the formula:

$$OM = \frac{Preignition\ mass\ (g) - post\ ignition\ mass\ (g)}{Preignition\ mass\ (g)} \cdot 100 \qquad [\%] \quad (6)$$

3.3.5 Soil water repellency

Soil aggregates with diameters of approximately 3 cm were collected from the three soil profile depths (0-20, 20-40, and 40-60 cm) at all the 62 sampling points (soil pits). The sorptivity of each aggregate was measured according to Leeds-Harrison *et al*. (1994). In this method, water infiltrates into each aggregate from a small area (± 4 mm) as shown in (Fig.2) which produces an expanding wetting bulb that does not reach the boundary of the aggregate during measurement. The balance used was accurate to 1 mg, which is less than 2 % of the smallest total mass of water infiltrated during the test. A water-repellency index (*R*) was determined from the sorptivity measurements of two wetting liquids with different soil-liquid contact angles according to Tillman *et al*. (1989). It was evaluated from sorptivity measurements that were conducted at -1 cm pressure head for both water and 95 % methanol. Sorptivity (S) at -1 cm pressure is given by:

$$Q_{(-1)} = \frac{4\,b\,S_{(-1)}^2 r}{f}, \qquad (7)$$

Where the subscript -1 signifies the pressure head at which the measurements were being made, Q is a steady rate of water flow, b (0.55) is a parameter that depends on the soil-water diffusivity function, r is the radius of the infiltrometer tip, *S* being sorptivity at -1 head and *f* is the fillable air-porosity. For non-repellent soils, the sorptivity of a 95 % ethanol to water solution, S_E is related to the sorptivity of pure water, S_W, by:

$$Sw = \left[\frac{(\mu_m/\gamma_m)^{1/2}}{(\mu_W/\gamma_W)^{1\prime 2}}\right] \cdot S_E \ , \qquad (8)$$

Where μ_m is the viscosity of 95 % methanol at 20 ℃ (0.001 Ns m^{-2}), γ_m is the surface tension of 95 % methanol at 20 ℃ (0.023 N m^{-1}), μ_W is the viscosity of water at 20 ℃ (0.0010 N m^{-2}), and γ_W the surface tension of water at 20°C (0.073 N m^{-1}). Using these values, Equation (8) is reduced to:

$$S_W = 1.78\, S_E. \quad (9)$$

Then the index *R* therefore becomes:

$$R = 1.78 \left(\frac{S_E}{S_W}\right), \quad (10)$$

With *R* = 1.0 signifying a totally non-repellent soil. According to Tillman *et al*. (1989), when $S_E < S_W$ (*R* < 1.78) the soil is non-repellent.

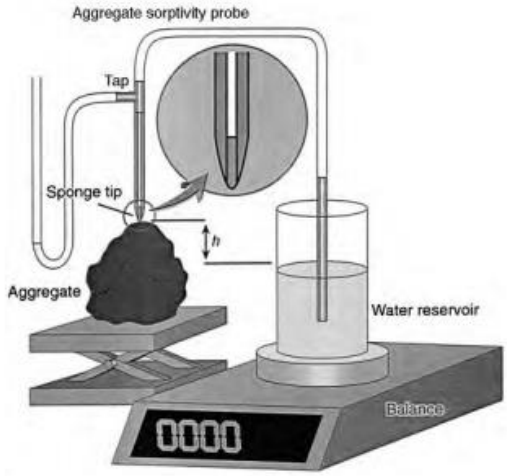


Figure 2. The infiltration devise used to measure the sorptivity of individual soil aggregates (Hallett and Young, 1999).

3.3.6 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:

$$Ksat = \frac{4\,VL}{\pi \cdot d_c^2 \cdot \Delta t \cdot \Delta H} \quad (11)$$

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.

3.3.7 Penetration resistance and shear strength

The vertical penetration resistance (PNT) was measured *in situ* with a 30º conical tip hand-push penetrometer and 1 cm^2 (McKyes, 1989). The PNT measurements in each soil sampling point were made by pushing vertically the penetrometer to the soil at an approximated speed of 2 cm s^{-1} with three repetitions (Bradford, 1986). The measurements were done after a high-intensity rainfall, ensuring the soil was close to field capacity. In each soil sampling point, measurements were taken with three repetitions in each depth; surface, 10 cm, 30 cm and 50 cm. In addition, the soil shear strength (horizontal resistance) was then measured using a handheld shear vane at the same depths as penetration resistance, with three repetitions in each sampling depth.

3.3.8 Structural stability index (SSI)

The SSI indicates the risk of structural degradation, therefore it was inferred from the index proposed for mineral tropical soil, especially those found in the West African Savanna (Pieri, 1992). The higher the value of the SSI, the more stable is the soil. The values below 9 % represent an unstable structure. Therefore the SSI was calculates using the following formula:

$$SSI = \left[\frac{OM}{(silt + clay)}\right] \cdot 100 \quad (\%) \qquad (12)$$

Where silt, clay and the organic matter content (OM) are in percent.

3.4 Digital Elevation Model and Remotely Sensed Imagery

3.4.1 Digital Elevation Model (DEM)

The LiDAR point cloud was acquired in 2009 using a Harrier 54/G4Dual System sensor mounted on a Piper PA-24 Comanche airplane, achieving an average point cloud density of 4.64 points m^{-2}. The pulse and scanning frequencies were 100 kHz and 100 Hz, respectively, with a field vision angle of approximately 22.5° and a laser pulse wavelength of 1550 nm. Using the "lasground.exe" tool of LAStools software (Isenburg, 2014), the point cloud was classified as ground and no ground returns, subsequently obtaining from these a digital terrain model (DTM) and digital surface model (DSM) respectively with the "las2dem.exe" tool, both with a spatial resolution of 5 m (see appendix). A digital crown height model (DCM) was calculated by the difference between the DSM and the DTM. The SAGA GIS software Conrad, (2014) was used to calculate, based on the DTM, the first-order topographical variables, landform variables associated with morphometric, hydrology and topography. These 26 variables acquired from LiDAR are summarized in Table 1.

Table 1. Description of the environmental covariates derived from 5 m resolution DEM.

Type of variables		Variable names	References
Terrain attributes	*Topographical derivatives*	Slope aspect	
		Secondary curvature	Zevenbergen and Thorne (1987)
		Longitudinal curvature	
		Elevation	
Landform variables	*Hydrological variables*	Catchment area	Böhner and Selige (2006)
		SAGA Wetness Index	Moore *et al*. (1991)
		LS Factor	Melton (1965)
		Flow Accumulation	Beuer *et al*. (1985)
		Modified catchment area	Freeman (1991)
			Fairfield and Leymarie (1991)
		Catchment slope gradient	Quinn *et al*. (1991)
		Topographic wetness index (x2)	
	Morphometry	Convergence Index	Koethe and Lehmeier (1996)
		Curvature Classification	Dikau (1988)
		Multiresolution Index of Valley Bottom Flatness (MRVBF)	Gallant and Dowling (2003)
		Multiresolution Index of the Ridge Top Flatness (MRRTF)	Gallant and Dowling (2003)
		Topographic Position Index(TPI)	Guisan *et al*. (1999)
		TPI Based Landform Classification	Guisan *et al*. (1999)
		Relative Height and Slope Positions	Boehner and Selige (2006)
		Terrain Surface Texture	Iwahashi and Pike (2007)
		Terrain Surface Convexity	Iwahashi and Pike (2007)
		Upslope and Downslope Curvature	Freeman (1991)
			Zevenbergen and Thorne (1987)
		Terrain Surface Classification	Iwahashi and Pike (2007)
		Diurnal Anisotropic Heat	Boehner and Antonic (2009)
	Lighting, Visibility	Analytical Hillshading	Tarini *et al*. (2006)

3.4.2 Spectral remote sensing data

Sentinel-2A optical images from November 2018 (freely available from the European Space Agency, ESA), were used. For this study, Sentinel-2 band 4 (B4; 10 m spatial resolution; 665 nm), band 2 (B2; 10 m spatial resolution; 490 nm), band 3 (B3; 10 m spatial resolution; 560 nm) atmospherically corrected imagery, were extracted from Copernicus data hub and used to calculate the spectral indices as summarized in Table 2. This data was obtained for the 62 geo-referenced pixels where the measurements of soil physical properties were carried out and soil samples were collected. Preliminary processing was carried out on these records to remove outliers due to the presence of clouds. Only the images without the presence of clouds were used in the analysis. The technical aspect can be obtained from Van der Werf and Van de Meer (2016).

Table 2. Description of the spectral indices derived from sentinel 2A remote sensing.

Spectral indices	Abbr.	Reference
Green-Red vegetation index	GRVI	Motohka *et al.* (2010)
Three band spectral index	TBSI-T	Tian *et al*. (2014)
Three band spectral index	TBSI-W	Wang *et al.* (2012)
Three band spectral index	TBSI-V	Verrelst *et al.* (2015)

3.5 Selection of predictors

Due to the large amount of data available for use as covariates in modelling, 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. The process of selecting appropriate predictors was done using two methods of attribute 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 first method applied was the Boruta method (Kursa and Rudnicki, 2010) which is a wrapping type selection algorithm, that is, the classifier is used as a black box to interpret the subsets of characteristics based on its predictive power, returning a ranking of characteristics (Ladha and Deepa, 2011; Kursa and Rudnicki, 2010). The selection is made "backwards", starting with all the variables and eliminating them from one, wherein each step the variable that allows the error to be reduced is eliminated (Ladha and Deepa, 2011). Kursa and Rudnicki (2010) noted that to use the Boruta algorithm, the following steps should be applied:

(a) Create copies of all variables, these copies would be "shadow variables".
(b) Mix all the attributes, training a Random Forest (RF) classifier in the system.
(c) Calculate a measure of importance for shadow variables and real predictors.
(d) Find the maximum value of importance within the shadow variables.
(e) Compare the real predictors with the maximum importance values of the shadow variables.
(f) It gives a score (z-score) to those real predictors that are more important than the shadow variables.
(g) Eliminate predictors that have less importance than shadow variables.

(h) Remove all shadow variables.
(i) Repeat the procedure until a score is assigned to all the predictors or until the previously established limit of executions has been reached.

On the other hand, the Recursive Feature Elimination (RFE) algorithm is an example of "backward" feature deletion, in an iterative procedure (Guyon *et al*., 2002). According to Guyon *et al*. (2002), RF trains a classifier, which in this case was also RF, then calculates the classification criteria for all the characteristics and finally eliminates the function with the smallest classification criterion. Both methods eliminate variables as the classifier is iterated. Although both methods calculate a measure of importance, unlike Boruta, RFE calculates a measure of accuracy, which in this case was the root of the mean square error (RMSE) according to Cabezas *et al*. (2016). Then, considering that neither of the two methods considers collinearity among the predictors, it is that the selected predictors were compared with each other, choosing those predictors with better ranking and Pearson correlation coefficient (r) with the response variable and smaller among them. The best-performed predictors were then used in the final model.

3.6 Statistical analysis

3.6.1 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%.

3.6.2 Generation of statistical models

The main objective behind the use of RF is to achieve an improved predictive accuracy by growing a large number of de-correlated trees. Breiman (2001) stated that this is done to obtain a prediction accuracy by averaging the prediction values from all the trees in the ensemble for each observation RF is thus especially beneficial for data sets with a large number of predictors that may be correlated. The RF method deals with an ensemble of trees and has evolved into an important non-parametric method that has capabilities to fit interactions that may be highly non-linear. Cutler *et al*. (2007) suggested that RF may also be used to deal with irrelevant predictor variables and robust outliers in the predictor variables list.

The RF method is a bagging method and uses recursive partitioning to form regression trees. Therefore each regression tree that is created is then independently grown until its maximum size is reached based on the training data set, known as the bootstrap sample consisting of 66% of the total population (Liaw and Weiner, 2002). On the other side RF seeks to add randomness into the data by selecting random subsets of input variables to establish the most efficient split at each tree node, thus reducing bias and maintaining diversity with the data since no pruning is performed (Breiman, 2001). The remaining 34% of the data also known as out-of-bag data is used for the model prediction. The ensemble predicts the data using the difference in the mean square error of the out-of-bag data and the data that is used to grow the homogeneous regression trees.

Once the predictors were chosen for each of the response variables (soil properties), the predictive models were generated. For this, two approaches were used, a regression testing the nonparametric Random Forest (RF) model proposed by Breiman (2001). 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 algorithm, for classification and regression, consists of (Liaw and Wiener, 2002):

(a) Draw ntree bootstrap samples of the original data, where ntree corresponds to the number of trees to grow, this should not be a very small number. In this case, the ntree used was 500 (Fassnacht *et al*., 2014a, Lopatin *et al*., 2016, Castillo-Riffart *et al*., 2017).
(b) For each of the bootstrap samples, a classification or regression tree grows without pruning, with the following modifications: in each node, instead of choosing the best division among all the predictors, randomly sample a random sample mtry (specifies how many random features will be selected to grow a single tree) of the predictors and choose the best division between these variables. For regressions, the default value of mtry corresponds to 1/3 of the total descriptors (Svetnik *et al*., 2003). The mtry was adjusted for each model, depending on the number of variables used.
(c) Predict new data by adding the predictors of ntree trees. All the procedures were carried out using the packages "Random Forest" (Liaw and Wiener, 2002), "caret" (Kuhn *et al*., 2017) and "stats" (R Core Team, 2017) of the R-project software.

3.6.3 Model accuracy

For validation of purposes, the best RF models were embedded in bootstrap with 500 iterations. In each bootstrap iterations, 80 times were drawn with replacement from 80 available samples. In this procedure, on average 34% of the total number of samples were not drawn. These samples were subsequently used as holdout samples for an independent validation according to Fassnacht *et al*. (2014b). The model performances of RF was compared based on differences in coefficient of determination R^2 and normalized RMSE between predicted and observed soil physical properties values of the hold-out samples in the bootstrap. To enable sound comparisons between four response variables, the normalized RMSE (nRMSE) will be used and calculated as:

nRMSE= [RMSE/ [max (number of attributes) – min (number of attributes)]] x 100

Where RMSE will be calculated as;

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}} \sum_{j=1}^{n} \left(y_j - \hat{y}\right)^2 \quad (13)$$

With y denotes reference parameter values, $\hat{y}$ estimated value, and *n* number of samples.

$$R^2 = \frac{\sum_{j=1}^{n} (P_i - \overline{\theta_1})^2}{\sum_{j=1}^{n} (O_i - \overline{\theta_1})^2} \quad (14)$$

Where *n* denotes data point, O_i and P_i are observed and predicted soil selected physical 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. The bias of prediction was measured as one minus the slope of regression without intercept of the predicted versus observed values.

3.6.4 Predictive variability maps

Predictive maps of selected soil properties were calculated for each depth based on the obtained best-performed models.

4. RESULTS

4.1 Soil properties

4.1.1 Descriptive statistics

The analysis of the collected data of selected soil parameters was first achieved through the conventional statistics (minimum, maximum, arithmetic mean, standard deviation, range, coefficient of variation (CV), kurtosis and skewness) as given in Tables 3 to 5. The descriptive statistics of the selected soil physical properties from the surface (0-20 cm) are given in Table 3. The CV values 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%) variabilities. However, the high or low rank of the CV of a given variable is not necessarily the same for another variable (Costa et al., 2008). Mulla and McBratney (2000) proposed the CV classification for different variables be: SOM with a CV value of 21-41 % is moderate to high, while bulk density with the CV values of between 3-26 % is regarded as low to moderate, clay with a CV ranging between 16-53% is classified as moderate to high and sand content with a CV ranging between 3-37 % is classified as low to moderate variability, porosity with a CV that ranges between 7-11 % regarded as low variable, saturated hydraulic conductivity with the CV ranging between 48-352 % is regarded as high variable, water content at field capacity with an CV ranging between 4-20 % is categorised as low to moderate variable and water content at permanent wilting point with an CV value of 14-45% is categorized as moderate to high.

The results obtained in this study indicate low to very high variability of soil physical properties within the top horizon (0-20 cm). The greatest and the least CVs for soil physical properties were obtained for Ksat (104%) and PD (3.1%), indicating very high variability for Ksat relative to other soil properties and the low value for PD indicates low variability relative to other properties. However, the majority of soil properties falls under high variability category. Cerri and Magalhães (2012) stated that high CV is the first indicator of data heterogeneity. Based on the skewness and kurtosis, most of the variables were satisfactorily described by a normal distribution (Table 3) and did not require transformation. Two possible exceptions were Ksat and available water content.

The skewness for the normality should be less than 3 and it was found to be in a range of (-0.03 – 2.27). In spite of skewness and kurtosis of the distribution of some soil properties, the mean and median values were similar with the mean being equal to or almost equal to the median; except for clay (14.9 – 13.1), SSI (15.3 – 14.0) and Ksat (50.6 – 32.2). Minimum, maximum and range for other variables are also shown in (Table 3). Nevertheless, CV is the most discriminating factor for describing the variability of soil properties than the other parameters such as SD, mean, median according to Xing-Yi *et al*. (2007).

Table 3. Descriptive statistics for selected soil physical properties in the top soil (0-20 cm).

Variables	Min	Max	Mean	SD	CV	Median	Range	Skewness	Kurtosis
Clay (%)	6.8	32.7	14.9	5.3	35.0	13.1	26.0	1.17	1.07
Silt (%)	9.4	29.6	20.4	3.5	17.0	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
OM (%)	3.13	9.79	5.26	1.16	22.0	5.07	6.7	1.20	2.45
BD (Mg m^{-3})	1.17	1.67	1.48	0.12	7.9	1.48	0.50	-0.41	-0.16
f (%)	33.0	54.8	42.0	4.4	10.0	41.7	21.8	0.6	0.3
PD (Mg m^{-3})	2.36	2.68	2.55	0.08	3.1	2.57	0.30	-0.36	-0.88
SSI (%)	9.7	28.0	15.3	4.3	28.0	14,0	18.2	1.57	1.88
R Index	0.81	4.36	2.27	0.81	35.0	2.16	3.55	0.42	-0.28
DR (%)	21.2	75.5	43.1	11.3	26.0	42.3	54.3	0.43	0.12
PNT (kPa)	90.0	365.0	215.0	55.6	26.0	205	275	0.17	-0.34
Shear strength	72.0	336.0	164.9	53.5	32.0	160	264	0.63	0.28
Ksat (cm h^{-1})	0.0	290.1	50.6	52.8	104.0	32.2	290.1	2.27	6.31
FC (%)	14.0	29.0	20.0	4.0	18.0	20.0	15.0	0.58	0.08
PWP (%)	6.0	20.0	12.0	3.0	23.0	12.0	14.0	0.53	0.02
AWC (%)	4.0	18.0	7.0	2.0	31.0	7.0	14.0	1.92	6.16
FDP (%)	2.0	25.0	15.0	4.0	28.0	15.0	23.0	-0.03	0.60
SDP (%)	2.0	8.0	5.0	1.0	27.0	5.0	6.0	0.30	0.63
ΔWMD (mm)	0.1	8.47	4.56	2.14	47.0	4.22	8.42	-0.05	-0.77

N = 62, SD = standard deviation, CV = coefficient of variation (%), OM = organic matter, BD = bulk density, PD = particle density, f = total porosity, SSI = structure stability index, R = repellence index, DR = dispersion ratio, PNT = penetration resistance, Ksat = saturated hydraulic conductivity, FC = field capacity, PWP = permanent wilting point, AWC = available water content, FDP = fast drainage pores, SDP = slow drainage pores, ΔWMD = weighted mean diameter.

The descriptive statistics of the soil physical properties belonging to 20-40 cm soil depth are given in Table 4. Probability distributions of the soil properties were evaluated using skewness and kurtosis, and those not normally distributed were subjected to logarithmic transformation (PD, SSI, R index, DR, and PNT). PD and Ksat have minimum and maximum values of skewness, respectively (Table 4) and some properties were negatively skewed. Moreover, minimum and maximum values of kurtosis are shear strength and Ksat, respectively (Table 4). 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; exceptions were DR, PNT, shear strength and Ksat (Table 4). Ranking of CVs for soil physical properties into different classes, for most soil properties exceeded 20%, indicating significantly variability (Table 4). PD, BD and *f* had the lowest CV (3.7%-9.6%); silt, sand and FC had the medium CVs (15.5%-17.2%); clay, OM, SSI, PNT, shear strength, and PWP had high CVs (20.4%-28.9%); and finally, R-index, DR, FDP, SDP, ΔWMD and AWC had the highest CVs (35.6%-197.3%), indicating that soil properties were ordinarily heterogeneous. In overall, minimum and maximum values of CV is particle density (PD) and Ksat, respectively (Table 4). Among the soil properties analysed, the SD for SDP, OM, PD, BD and R index were the lowest (0.07-1.0), while that for PNT was the greatest (57.0). The Ksat values varied from 0.00 to 186.3 and have a mean of 22.6 cm h^{-1} (Table 4).

Table 4. Descriptive statistics for selected soil physical properties in the subsurface (20-40 cm).

Variables	Min	Max	Mean	SD	CV	Median	Range	Skewness	Kurtosis
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
Sand (%)	33.5	76.8	51.9	8.2	15.8	50.3	43.4	0.39	0.57
OM (%)	2.10	6.55	4.33	0.88	20.39	4.35	4.45	-0.21	0.20
BD (Mg m^{-3})	1.43	1.73	1.59	0.07	4.40	1.60	0.30	-0.28	-0.54
f (%)	24.6	46.5	39.1	3.8	9.6	39.2	21.9	-0.95	2.41
PD (Mg m^{-3})	2.16	2.73	2.62	0.10	3.66	2.65	0.57	-2.28	7.20
SSI (%)	4.2	15.0	9.1	1.9	21.0	8.9	10.8	0.61	1.01
R Index	0.68	3.02	1.68	0.60	35.63	1.69	2.34	0.38	-0.66
DR (%)	11.7	76.4	30.0	13.1	47.5	23.4	64.7	1.52	2.14
PNT (kPa)	130.0	385.0	276.4	57.0	20.6	287.0	255.0	-0.78	0.36
Shear strength	116.0	332.0	240.7	55.8	23.2	251.0	216.0	-0.41	-0.68
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
AWC (%)	3.0	17.0	7.0	2.0	33.9	7.0	14.0	1.98	6.79
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
ΔWMD (mm)	0.33	12.01	6.81	2.62	38.46	6.68	11.68	-0.01	-0.33

N = 62, SD = standard deviation, CV = coefficient of variation (%), OM = organic matter, BD = bulk density, PD = particle density, f = total porosity, SSI = structure stability index, R = repellence index, DR = dispersion ratio, PNT = penetration resistance, Ksat = saturated hydraulic conductivity, FC = field capacity, PWP = permanent wilting point, AWC = available water content, FDP = fast drainage pores, SDP = slow drainage pores, ΔWMD = weighted mean diameter.

The descriptive statistics of the soil physical properties within the deeper horizon (40-60 cm) is presented in Table 5. The statistics result obtained indicates moderate to high skewness, and the values of skewness vary from -0.62 to 4.04 depicting moderate to high skewness (Table 5); Ksat had the highest value of skewness, while PNT had the lowest skewness. Moreover, minimum and maximum values of kurtosis are PNT (-0.74) and AWC (21.09), respectively (Table 5). The mean and median values were similar with means being equal to or almost equal to the median; exceptions were Ksat, ΔWMD, PNT and DR (Table 5). CVs for most of the soil physical properties exceeded 20% indicating significant spatial variability. BD, *f* and PD had the lowest CVs (3.2%-8.6%); FC, PWP and sand had the medium CVs (14.7%-17.8%); clay, silt, OM, SSI, PNT, shear strength and ΔWMD had high CVs (21.2%-29.1%); finally, R index, DR, Ksat, AWC, FDP and SDP had very high CVs (33.0%-229.9%) indicating that soil properties were heterogeneous (Table 5). Among the soil properties analysed, the SD for SDP, BD and PD were the lowest (0.07-1.0), while PNT was the greatest (99.0). Ksat had the highest CV (229.9%) relative to all the other soil properties, while PD showed the lowest variability (3.2%). PNT, shear strength and Ksat showed the highest range relative to other soil properties (Table 5).

Table 5. Descriptive statistics for selected soil physical properties in the subsoil (40-60 cm).

Variables	Min	Max	Mean	SD	CV	Median	Range	Skewness	Kurtosis
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
OM (%)	1.59	5.56	3.67	0.88	24.0	3.51	3.97	0.20	-0.62
BD (Mg m^{-3})	1.46	1.77	1.63	0.07	4.5	1.63	0.31	-0.07	-0.54
f (%)	32.0	46.0	38.0	3.3	8.6	37.7	14.0	0.26	-0.38
PD (Mg m^{-3})	2.47	2.93	2.63	0.08	3.2	2.65	0.46	0.30	1.03
SSI (%)	2.8	10.9	7.3	1.7	23.8	7.1	8.1	0.02	-0.49
R Index	0.8	1.47	1.67	0.61	36.7	1.47	0.67	1.16	1.20
DR (%)	13.0	85.4	27.7	14.8	49.3	25.6	72.4	1.60	2.42
PNT (kPa)	160	550.0	364.2	99.0	27.2	372.5	390	-0.04	-0.74
Shear strength	132	500.0	304.2	88.5	29.1	304.0	368	0.42	0.03
Ksat (cm h^{-1})	0.0	330.7	18.4	52	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
AWC (%)	2.0	29.0	8.0	3.0	44.2	7.0	27.0	3.95	21.09
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
ΔWMD (mm)	3.5	12.1	7.5	2.0	26.4	7.8	8.6	0.16	-0.55

N = 62, SD = standard deviation, CV = coefficient of variation (%), OM = organic matter, BD = bulk density, PD = particle density, f = total porosity, SSI = structure stability index, R = repellence index, DR = dispersion ratio, PNT = penetration resistance, Ksat = saturated hydraulic conductivity, FC = field capacity, PWP = permanent wilting point, AWC = available water content, FDP = fast drainage pores, SDP = slow drainage pores, ΔWMD = weighted mean diameter

As expected the soil properties varied with depth; OM, sand and porosity decreased with depth while BD, clay, PD, shear strength, PNT and MWD increased with depth (Table 3 to 5).

4.1.2 Pearson correlations

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 6, 7 and 8, respectively. In the top-soils (0-20 cm) depth, the clay content was significantly ($p < 0.01$) positively correlated with properties such as shear strength (r = 34 %), FC (r = 49 %), PWP (r = 74 %) and significantly ($p < 0.05$) positively correlated with OM (r = 31 %), but negatively with sand (r = -83 %), SSI (r = -32 %), FDP (r = 27 %) and SDP (r = -27 %) (Table 6). The FC was significantly ($p < 0.01$) positively correlated with properties such as clay (r = 49 %), silt (r = 34 %), OM (r = 37 %), but negatively with sand (r = -59 %) and Ksat (r = -44 %) (Table 6). As expected SOM was negatively correlated with sand content and bulk density (Table 6). However, there was no significant correlation between clay content and functional soil properties like aggregate stability test (DR and ΔWMD), Ksat and PNT.

Correlation results between the properties of sub-soils (20-40 cm) depth shows that BD was significantly ($p < 0.05$) positively correlated with properties such as DR (r = 26 %), but significantly ($p < 0.01$) negatively correlated with *f* (r = -76 %), Ksat (r = -40 %) and

FDP (r = -32 %) (Table 7). SSI was significantly ($p < 0.01$) positively correlated with sand (r = 42 %), OM (r = 60 %) and significant ($p \leq 0.05$) positively correlated with *f*, but significant ($p < 0.05$) negatively correlated with clay (r = -32 %) (Table 7). There was no significant correlation between ΔWMD and any of the other soil properties at the sampling depths, except for the topsoil (Table 6, 7 and 8). Nevertheless, DR was significantly positively correlated with silt (r = 36 %), BD (r = 26 %), PD (r = 30 %) and significantly negatively correlated (20-40 cm) with clay (r = -29 %), OM (r = -47%), SSI (r = -32 %) and PNT (r = -38 %). On the other hand, Ksat at the depth of 20-40 cm was significantly positively correlated with sand (r = 27 %), *f* (r = 32 %) and FDP (r = 37 %) significantly negatively correlated with BD (r = -40 %), FC (r = -50 %) and PWP (r = -39 %). Lastly, the Pearson correlation results of sub-soils (40-60 cm) depth shows that PNT was significant ($p < 0.01$) positively correlated with clay content (r = 50 %), shear strength (r = 39 %) and PWP (r = 38 %), but significant negatively correlated with DR (r = -41 %) and SDP (r = -27 %) (Table 8). The *f* was significant ($p < 0.01$) positively correlated with PD (r = 54 %) and SDP (r = 43 %). Ksat showed no dependence with other properties, while DR was significant positively correlated with silt (r = 31 %), sand (r = 34 %) and AWC (r = 37 %) and significant negatively correlated with clay (r = -49 %) and PNT (r = -41 %) (Table 8). Generally, there is a positive correlation between sand and Ksat and negative correlation between sand and shear strength and water retention (FC and PWP) among all the depths (except for sand at 40-60 cm). The same for the relationship between OM and water retention at FC and PWP and OM and SSI.

Table 6. Pearson correlation of selected soil physical properties in the surface (0-20 cm).

Variables	Clay	Silt	Sand	OM	BD	*f*	PD	SSI	R	DR	PNT	Shear	Ksat	FC	PWP	AWC	FDP	SDP	ΔWMD
Clay	100																		
Silt	3	100																	
Sand	-83**	-58**	100																
OM	31*	-2	-24	100															
BD	8	5	-9	-13	100														
f	-6	-13	12	5	-92**	100													
PD	4	-18	7	-19	30*	9	100												
SSI	-32*	-49**	53**	67**	-18	14	-10	100											
R index	-1	3	-1	-1	-29*	29*	-3	1	100										
DR	-16	31*	-4	-17	25*	-21	15	-23	4	100									
PNT	4	-9	2	-26*	22	-14	23	-16	-17	-6	100								
Shear	34**	1	-29*	-10	19	-16	10	-29*	-4	-7	53**	100							
Ksat	-10	-47**	34**	-6	-48**	49**	-4	22	1	-24	8	-1	100						
FC	49**	34**	-59**	37**	14	-16	-4	-13	-14	5	-21	13	-44**	100					
PWP	74**	28*	-76**	39**	18	-19	-2	-20	-5	-4	-3	31*	-43**	77**	100				
AWC	-18	13	8	8	4	-5	-3	8	-18	12	-26*	-20	-14	57**	-8	100			
FDP	-27*	-1	23	-1	-70**	61**	-32*	17	15	-14	-16	-32*	45**	-44**	-40**	-20	100		
SDP	-27*	32*	4	-17	-17	9	-25	-12	3	14	-24	-30*	-12	-4	-6	0.2	43**	100	
ΔWMD	0.10	16	-17	-18	26*	-13	34**	-28*	-6	2	16	17	-0.4	3	6	-0.1	-5	-5	100

*N = 62, *, ** Significant at p = 0.05 and 0.01, respectively. OM = organic matter (%), BD = bulk density (Mg m^{-3}), PD = particle density (Mg m^{-3}), f = total porosity (%), SSI = structure stability index (%), R = repellence index, DR = dispersion ratio(%), PNT = penetration resistance (kPa), Ksat = saturated hydraulic conductivity (cm h^{-1}), FC = field capacity (%), PWP = permanent wilting point (%), AWC = available water content (%), FDP = fast drainage pores (%), SDP = slow drainage pores (%), ΔWMD = weighted mean diameter (mm), Shear = shear strength (kPa).*

Table 7. Pearson correlation of selected soil physical properties in the subsurface (20-40 cm).

Variables	Clay	Silt	Sand	OM	BD	*f*	PD	SSI	R	DR	PNT	Shear	Ksat.	FC	PWP	AWC	FDP	SDP	ΔWMD
Clay	100																		
Silt	-27*	100																	
Sand	-94**	-7	100																
OM	52**	-17	-46**	100															
BD	-1	-2	2	-11	100														
f	-9	7	7	-10	-76**	100													
PD	-14	9	11	-27*	-6	70**	100												
SSI	-32*	-24	42**	60**	-12	32*	-18	100											
R index	-10	-1	11	2	-19	21	9	16	100										
DR	-29*	36**	17	-47**	26*	1	30*	-32*	-17	100									
PNT	24	-24	-17	21	3	-17	-22	5	3	-38**	100								
Shear	45**	-12	-43**	24	10	-8	-3	-10	9	-15	35**	100							
Ksat	-24	-4	27*	-16	-40**	32*	6	9	-2	12	0.4	-17	100						
FC	66**	1	-69**	47**	18	-23	-16	-12	-11	-6	14	36**	-50**	100					
PWP	73**	-9	-73**	53**	15	-21	-16	-9	-4	-16	14	37**	-39**	80**	100				
AWC	-1	13	-4	2	-4	1	-3	-3	-11	9	1	4	-7	36**	-23	100			
FDP	-40**	10	38**	-32*	-36**	33**	11	1	12	8	-14	-38**	37**	-75**	-55**	-34**	100		
SDP	-51**	17	47**	-20	-9	17	15	18	-10	8	-25	-39**	1	-30*	-36**	2	14	100	
ΔWMD	11	-5	-10	8	-9	11	9	1	-3	4	-7	-1	-0.2	10	7	4	-11	-10	100

*N = 62, *, ** Significant at p = 0.05 and 0.01, respectively. OM = organic matter (%), BD = bulk density (Mg m^{-3}), PD = particle density (Mg m^{-3}), f = total porosity (%), SSI = structure stability index (%), R = repellence index, DR = dispersion ratio (%), PNT = penetration resistance (kPa), Ksat = saturated hydraulic conductivity (cm h^{-1}), FC = field capacity (%), PWP = permanent wilting point (%), AWC = available water content (%), FDP = fast drainage pores (%), SDP = slow drainage pores (%), ΔWMD = weighted mean diameter (mm), Shear = shear strength (kPa).*

Table 8. Pearson correlation of selected soil properties in the subsoil (40-60 cm)

Variables	Clay	Silt	Sand	OM	BD	*f*	PD	SSI	R	DR	PNT	Shear	Ksat	FC	PWP	AWC	FDP	SDP	ΔWMD
Clay	100																		
Silt	-26*	100																	
Sand	-87**	-24	100																
OM	37**	-15	-30*	100															
BD	-26*	-4	28*	-9	100														
f	11	5	-13	-4	-80**	100													
PD	-19	1	19	-21	7	54**	100												
SSI	-19	-31*	34**	79**	10	-12	-8	100											
R index	-9	1	9	-9	-3	8	11	-1	100										
DR	-49**	31*	34**	-16	15	16	-5	4	5	100									
PNT	50**	-18	-42**	31*	1	-4	-7	6	-13	-41**	100								
Shear	38**	-12	-33**	31*	-4	-5	-15	10	-6	-16	39**	100							
Ksat	-2	-0.2	2	-5	-16	10	-7	-4	-2	-4	6	-18	100						
FC	38**	-5	-36**	46**	-2	-4	-12	20	9	17	20	20	-9	100					
PWP	61**	-5	-59**	48**	-9	3	-9	8	-13	-11	38**	44**	-16	69**	100				
AWC	-14	2	13	10	5	-7	-4	17	25	37**	-15	-3	6	62**	-14	100			
FDP	-24	0	24	-22	-15	4	-15	-5	-17	9	-9	-11	-6	-44**	-49**	-5	100		
SDP	-42**	20	32*	-22	-30*	43**	30*	-3	-14	7	-27*	-36**	-4	-39**	-42**	6	46**	100	
ΔWMD	2	-0.4	-2	-22	-18	23	14	-23	5	-21	4	4	10	-10	10	-23	-10	15	100

*N = 62, *, ** Significant at p = 0.05 and 0.01, respectively. OM = organic matter (%), BD = bulk density (Mg m^{-3}), PD = particle density (Mg m^{-3}), f = total porosity (%), SSI = structure stability index (%), R = repellence index, DR = Dispersion ratio(%), PNT = penetration resistance (kPa), Ksat = saturated hydraulic conductivity (cm h^{-1}), FC = field capacity (%), PWP = permanent wilting point (%), AWC = available water content (%), FDP = fast drainage pores (%), SDP = slow drainage pores (%), ΔWMD = weighted mean diameter (mm), Shear = shear strength (kPa).*

4.2 Statistical models

4.2.1 Model performance

The performance results of RFM as a predictor in modelling soil physical properties were summarized in terms of R^2, RMSE, nRMSE and bias for all 500 bootstraps values for all the sampling depths (Table 9). The particle size distribution results showed that the combination of the various predictor 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. The prediction of clay content showed systematically higher R^2 (0.66, 0.65, 0.51), across all depths. However, the silt content showed the lowest R^2 (0.42, 0.42 and 0.50) across all depths, respectively. The values of R^2 showed the trend within the soil depths for all the particles, for example, in clay content is higher for the first 20 cm, then decreased with depth 0.66 > 0.51 > 0.65 (see Table 9). Sand showed the decrease with the soil depth. However, the opposite was observed for silt content. In terms of prediction accuracy, clay content showed the highest nRMSE (27.5, 26 and 16 %), bias (0.3, 0.7 and 0.5) values across all depths and the lowest nRMSE (8.6, 13, and 13) was associated with the prediction sand at all depths interval (Table 9). The lowest bias values were observed in silt content.

The performance results of the random forest model as a predictor in modelling bulk density (BD) is summarised in Table 9. The best model fit of BD was found at surface depths (median bootstrap R^2 of 0.51; nRMSE of 7 %). While the worst fit of BD was observed for 20-40 cm soil depths (median bootstrap R^2 of 0.26; nRMSE of 4 %). The summary of root means square error and bias can also be obtained from Table 9. On the other hand, the performance results of the random forest model as a predictor in modelling soil organic matter (OM) is summarised in Table 9. The best model fit of BD was found at subsurface depths (median bootstrap R^2 of 0.53; nRMSE of 18 %). While the worst fit of OM was observed for both surface (20-40 cm) and 40-60 cm soil depths (median bootstrap R^2 of 0.31; nRMSE of 21 and 23 %, respectively). The summary of root means square error and bias can also be obtained from (Table 9)

The results of the dispersion ratio (DR) showed R^2 ranges from 0.49 to 0.81 across all the soil depths (Table 9). The best model fit was found when predicting DR at the lower soil depth (40-60 cm) (median bootstrap R^2 of 0.81, nRMSE of 30 % and bias of -0.4). While the worst fit was observed for the upper depth (0-20 cm) (median bootstrap R^2 of 0.49, nRMSE of 23 % and bias of -0.5) (Table 9). The values of R^2 and nRMSE increased with soil depth, while RMSE decreased within the soil depths.

The saturated hydraulic conductivity (Ksat) results demonstrate that R^2 ranges from 0.26 to 0.57 across all the soil depths. The best model fit was found when predicting Ksat at the lower depth (median bootstrap R^2 of 0.57, nRMSE of 194 % and bias -10.4). While the worst fit was observed for 20-40 cm (median bootstrap R^2 of 0.26, nRMSE of 109 and bias of 1.0) (see Table 9). The values of RMSE are summarized in Table 9. In addition, the water retention characteristics results showed that the combination of the various predictor variables can account for R^2 of 0.40 to 0.57, 0.42 to 0.54 and 0.21 to 0.31 of the variation in permanent wilting point (PWP), field capacity (FC) and available water content (AWC) respectively. The best fit was found when predicting PWP (median bootstrap R^2 of 0.57,

0.48 and 0.54; nRMSE of 19 %, 19% and 15% and bias of 0.3, 0.4 and 0 for all the depths, respectively). However, the worst fit was observed for FC for two soil depths (median bootstrap R^2 of 0.54, 0.40 and 0.42; nRMSE of 16 %, 16 % and 32 %; bias of 0.1, 0.2 and 0.45) (see Table 9). The values of average error rates increase with the increase in soil depths.

The pore size distribution results showed that the combination of various predictor variables can account for R^2 of 0.31 to 0.54 and 0.20 to 0.50 of the variation in slow draining pores (SDP) and fast draining pores (FDP), respectively. The best fit was found when predicting SDP (median bootstrap R^2 of 0.31, 0.54 and 0.50; nRMSE of 26 %, 31 % and 32%; bias -0.4, 1.2 and 0.9) (see Table 9). However, the FDP showed the lowest R^2 (0.20, 0.46 and 0.50) across all the sampling depths. The average error rate (nRMSE) and bias for the final model used to generate the predictive map of FDP were 26 %, 31% and 29% and 0.2, 1.0 and -0.2, respectively (see Table 9). The model performance results of other soil properties can also be obtained in Table 9.

In general, the dispersion ratio showed the best fit than other soil physical properties and available water content showed the worst fit. In addition, Ksat showed the highest average error rates across all the depths (91, 109 and 194 %) See Table 9. Based on model validation, all the models showed a systematic tendency to overestimate small values and underestimate high values (Fig.4). This effect was strong in Ksat models (Fig. 4.16, 4.17 and 4.18) and least in BD models (Fig. 4.10, 4.11 and 4.12).

4.2.2 Variable importance

The "importance" function in the random forest package was used to access the importance of predictor variables used to predict soil properties. Variable importance was calculated as the increase in node impurity weighted by the probability of reaching that node. The higher values showed the more importance of the variable (Fig. 3). Based on the average importance measured across 500 runs of the random forest model, the environmental variables that had the greatest influence on the model error rate for surface (0-20 cm) clay content were; DEM, catchment slope, topographic wetness index (TWI) and convexivity (Fig. 3.1). The predictor variables that had the greatest influence on subsurface (20-40 cm) clay content model error rate, averaged across 500 runs, were topographic position index (TPI), analytical Hillshading, aspect and LS factor (Fig. 3.2). The predictor variables that had the greatest influence on subsoil (40-60 cm) were DEM, longitudinal curvature, LS factor and modified catchment area (Fig. 3.3). The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm sand content were DEM, catchment slope, slope and secondary curvature; catchment slope, topographic position index, LS factor, analytical Hillshading; modified catchment area, topographic position index, aspect and catchment area, respectively (Fig. 3.4, 3.5 and 3.6). The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm silt content were; Analytical Hillshading, convexity, green-red vegetative index (GRVI), Multiresolution Index of the Ridge Top Flatness (MrRTF); Multiresolution Index of Valley Bottom Flatness (MrVBF), GRVI, TWI_1, TPI; Modified catchment area; TWI-SAGA, convergence index and GRVI, respectively (Fig. 3.7, 3.8, and 3.9).

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 bulk density were; aspect, DEM, TWI-SAGA, and TBSI-T (Fig. 3.10). The environmental variables that had the greatest influence on subsurface bulk density model error rate, averaged across 500 runs, were; secondary curvature, catchment slope, LS factor, and longitudinal curvature (Fig. 3.11). The environmental variables that had the greatest influence on 40-60 cm bulk density model error rate, averaged across 500 runs, were; analytical Hillshading, TBSI-W, convergence index and MBI (Fig. 3.12).

The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm dispersion ratio model were; TPI, modified catchment area, TWI_1, MrVBF (Fig. 3.13); Secondary curvature, longitudinal curvature, catchment slope, DEM (Fig. 3.14); DEM, catchment slope, TPI and longitudinal curvature, respectively (Fig. 3.15). The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm saturated hydraulic conductivity (Ksat) model were; longitudinal curvature, TBSI-T, TPI and MrRTF (Fig. 3.16); TPI, SAGA-TWI, modified catchment area and LS factor (Fig. 3.17); flow accumulation, catchment area, VRM and TPI, respectively (Fig. 3.18).

The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm permanent wilting point (PWP) model were; MrVBF, TPI, GRVI, DEM (Fig.3.19); Aspect, TPI, longitudinal curvature, analytical Hillshading (Fig. 3.20); Analytical Hillshading, TPI, TWI and catchment slope (Fig. 3.21), respectively. The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm field capacity water content model were; GRVI, MrVBF, aspect, slope (Fig. 3.22); Analytical Hillshading, modified catchment area, slope, TPI (Fig. 3.23); TWI, analytical Hillshading, catchment slope and TPI (Fig. 3.24), respectively.

The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm slow draining pores model were; Modified catchment area, DEM, catchment slope, GRVI (Fig. 3.25); catchment area, slope, modified catchment area, MrVBF (Fig. 3.26); Modified catchment area, TWI-SAGA, TBSI-W and MrVBF (Fig. 3.27), respectively. The environmental variables that had the greatest influence on the model error rate for 0-20 cm, 20-40 cm and 40-60 cm fast draining pores were; Analytical Hillshading, longitudinal curvature, convergence index, secondary curvature (Fig. 3.28); Modified catchment area, TPI, analytical Hillshading, slope (Fig. 3.29); Analytical Hillshading, modified catchment area, TBSI-W and TWI (Fig. 3.30), respectively.

Table 9. Performance of Random Forest model as a predictor in modelling soil properties

Properties	Depth (cm)	R^2	RMSE	nRMSE	Bias
Clay (%)	0-20	0.66	4.3	27.5	0.3
	20-40	0.51	7.7	26	0.7
	40-60	0.65	5.6	16	0.5
Sand (%)	0-20	0.52	5.6	8.6	0
	20-40	0.52	7.0	13	0.1
	40-60	0.48	5.6	13	0.3
Silt (%)	0-20	0.42	3.3	16	0
	20-40	0.42	2.7	14	0
	40-60	0.50	3.3	19	-0.9
OM (%)	0-20	0.31	1.1	21	-0.1
	20-40	0.53	0.8	18	0.7
	40-60	0.31	0.9	23	0.6
BD (Mg m^{-3})	0-20	0.51	0.1	7.0	0
	20-40	0.26	0.1	4.0	-0.2
	40-60	0.40	0.1	4.0	-0.2
PD (Mg m^{-3})	0-20	0.43	0.1	2.8	0
	20-40	0.15	0.1	4.0	-0.1
	40-60	0.45	0.1	3.0	-0.1
f (%)	0-20	0.49	3.9	9.1	-0.4
	20-40	0.30	3.6	9.0	0.1
	40-60	0.28	3.2	8.0	0.1
DR (%)	0-20	0.49	10.1	23.0	-0.5
	20-40	0.74	8.9	33	0
	40-60	0.81	8.7	30	-0.4
Ksat (cm h^{-1})	0-20	0.47	42	91	-1.8
	20-40	0.26	34.2	109	1.0
	40-60	0.57	37.6	194	-10.4
PWP (%)	0-20	0.57	3.0	19.0	0.3
	20-40	0.48	4.0	19.0	0.4
	40-60	0.54	3.0	15.0	0
FC (%)	0-20	0.54	3.0	16.0	0.1
	20-40	0.40	4.0	16.0	0.2
	40-60	0.42	1.0	32.0	0.4
SDP (%)	0-20	0.31	1.0	26.0	-0.4
	20-40	0.54	1.0	31.0	1.2
	40-60	0.50	1.0	32.0	0.9
FDP (%)	0-20	0.46	4.0	24.0	0.2
	20-40	0.20	3.0	34.0	1.0
	40-60	0.50	2.0	29.0	-0.2
AWC (%)	0-20	0.31	2.0	28.0	0.6
	20-40	0.21	2.0	34.0	2.1
	40-50	0.22	3.0	39.0	1.0

OM = organic matter, BD = bulk density, PD = particle density, f = total porosity, DR = dispersion ratio, PNT = penetration resistance, Ksat = saturated hydraulic conductivity, FC = field capacity, PWP = permanent wilting point, AWC = available water content, FDP = fast drainage pores, SDP = slow drainage pores, RMSE = root mean square error, nRMSE = normalized root mean square error, RFM = Random Forest model.

4.2.3 Prediction maps

The results present the patterns of soil texture for the three layers, as predicted by RFM (see Fig. 5.1-5.9). 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. 5.1). The areas that were predicted to have a greater amount of clay content were predominantly in the low lying slope of the toposequence in the southern and southern western part of the vineyard. Random Forest predicted surface clay content to range between 10-25 %. The largest area of the surface in the vineyard is predicted to be at most 14% of clay content. The predictive map of the subsurface (20-40 cm) clay content show that the RFM predicted clay content to be between 19 and 41 % (Fig. 5.2). High clay content was predicted most frequently in the lower-lying areas southern part of the study field.

The predictive map of the subsoil (40-60 cm) clay content shows that RFM predicted the lower and upper values of the clay content to be 20-43%. When looking at the predictive map for subsoil (40-60 cm) clay content, there was a large area where more than or equal to 31% clay was predicted (Fig. 5.3). The clay content is high from the northern to the lower slope of the southern part of the vineyard 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. 5.3).

The study present here the patterns of sand content within vineyard for the three soil depths (0-20, 20-40 and 40-60 cm, respectively), as predicted by RFM and the predictive maps are shown in Fig. 5.4 to 5.6. The obtained predictive maps show that RFM predicted the following minimum and maximum values for surface (55-78 %), subsurface (41-70%) and subsoil (37-60%). According to the surface sand predictive map, the sand content is predicted to be high in the upper elevation compared to the low lying or valley bottom of the toposequence (see Fig. 5.4). Also, the map shows moderate spatial variability within the vineyard. On the other hand, the predictive map of the subsurface sand shows that there were few areas where greater than 60% of sand was predicted across all the studied field (see Fig. 5.5). However, there was a greater margin between the minimum and maximum predicted values. The predictive map for the subsoil sand content showed that there were few areas where more than 54% of sand content was predicted (see Fig. 5.6). The areas that were predicted to have a greater amount of sand content were predominantly at the edges of the western part of the vineyard. However, the lower values of the sand content were predicted to be distributed in the middle of the field until the southern part of the field (Fig. 5.6).

The predictive maps of silt content at three different soil sampling depths are demonstrated in Fig. 5.7 to 5.9. The RFM predicted the minimum and maximum values of the silt content at the three layers to be 16-25%, 15-24%, and 9-33%, respectively. 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 (Fig. 5.7, 5.8 and 5.9). 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 in all the depths.

The study presented here the patterns of soil bulk density (BD) for the three layers (0-20, 20-40 and 40-60 cm), as predicted by RF model (Fig. 5.10, 5.11 and 5.12). The model predicted the minimum and maximum BD 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. The predictive map for surface BD 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} BD (Fig. 5.10). Their areas that were predicted to have greater values of BD were predominantly in the western part of the field. On the other hand, the predictive map of subsurface BD showed little overall variation and the majority of the area being classified as having 1.60 Mg m^{-3} (Fig. 5.11). The low values (1.40-1.54 Mg m^{-3}) of the BD in the subsurface is predicted in the low lying slope from the eastern side of the field towards the southern part of the vineyard.

The predictive map of the BD 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. 5.12). The higher predicted values of BD are at the flat slope and the lowest values are at the east towards the southern part of the vineyard.

The dispersion ratio (DR) which is the indicator of the micro-aggregate stability was predicted by RFM for three sampling depths (Fig. 5.13, 5.14 and 5.15). The minimum and maximum DR values predicted by RFM at all the depths; 26.2-62.3 %, 13.6-64%, and 16.4-61.9%, respectively. The predictive map of soil surface DR showed high overall variation, with the majority of the area being classified as having greater 44.2% (Fig. 5.13). The areas that were predicted to have a greater amount of DR were predominantly in edges of the western and eastern part of the field (Fig. 5.13). However, the lower predicted DR values are distributed within the field. The predictive map for subsurface DR again showed high overall variation, with the majority of the area being classified as having less than 26.2 %. (Fig. 5.14). The highest and lowest DR predicted values are found in the western and eastern part of the field, respectively (Fig. 5.14).

The results presented here show the patterns of saturated hydraulic conductivity (Ksat) for three sampling layers, as predicted by RFM (Fig. 5.16, 5.17 and 5.18). The model predicted the minimum and maximum Ksat values for all the three sampling depths; 9.2-144 cm h^{-1}, 0.79-149 cm h^{-1} and 0.27-88.60 cm h^{-1}, respectively. The all 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 h^{-1}, respectively (Fig. 5.16 to 5.18). The predictive map of the surface Ksat showed that the high predicted values were scattered within the field (Fig. 5.16). 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 (Fig. 5.17). The predictive map of the lower depth Ksat showed that greater values were on the south western side of the field (Fig. 5.18). However, there were few areas with high Ksat.

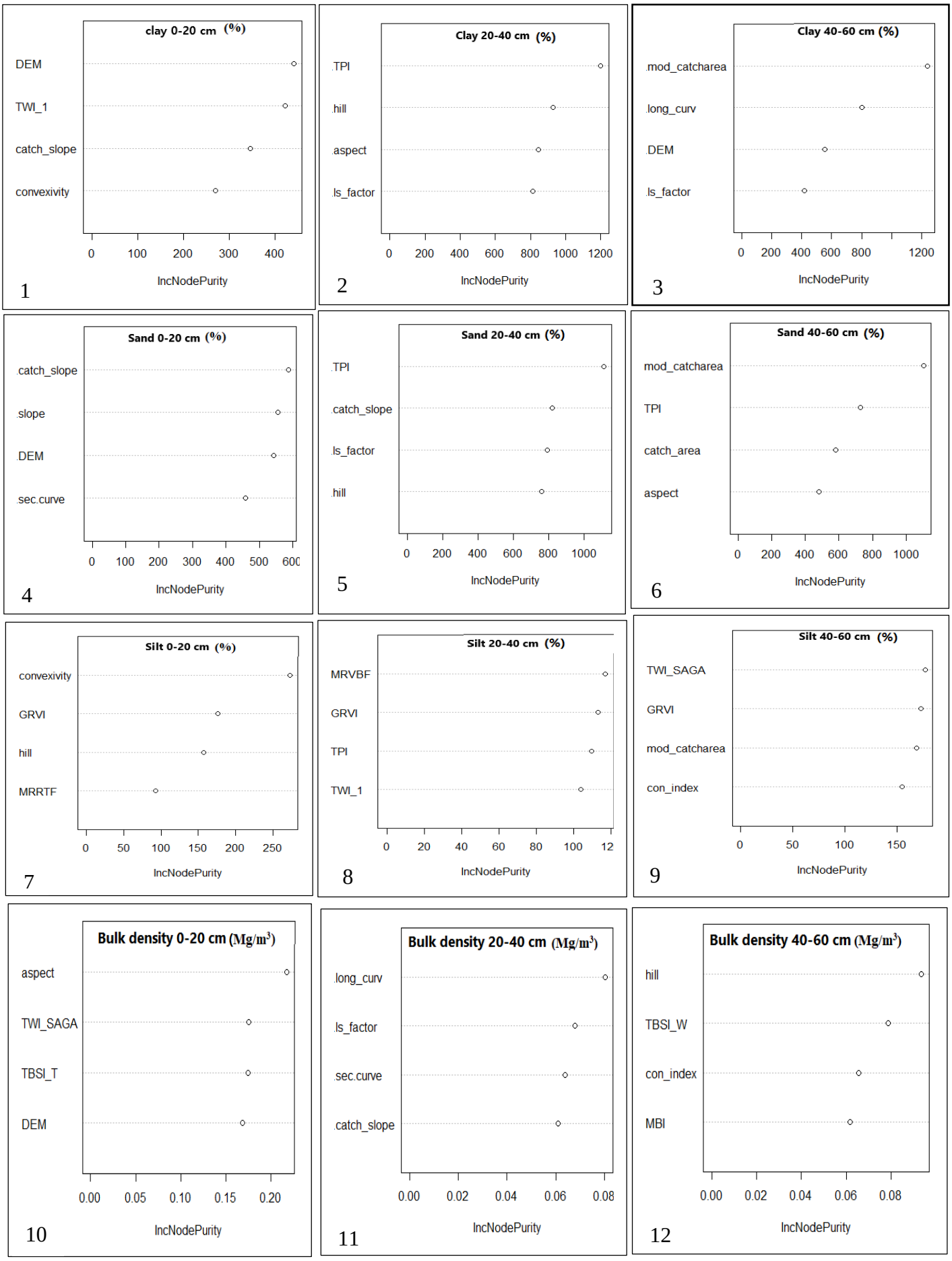
The model predicted the minimum and maximum water content at PWP values for all the three sampling depths; 7-17 %, 12-27 % and 15-28 %, respectively. The predictive map for surface permanent wilting point (PWP) water content showed high variation, with the majority of the area being classified as having less than 14 % PWP water content (Fig. 5.19). The areas that were predicted to have a greater amount of water content at PWP (14-16 %) were predominantly in the low lying slopes (Fig. 5.19). On the other hand, the predictive map of the water content at PWP (20-40 cm) showed high variation within the field, with the majority of the area being classified as 24 % water content (Fig. 5.20). Also,

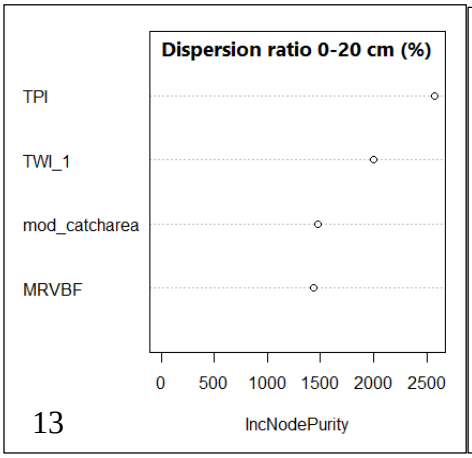
the predictive map for the water content at PWP (40-60 cm) showed moderate or medium variation within the field, with the majority of the area being classified as being 22 % water content (Fig. 5.21).

The model predicted the minimum and maximum water content at field capacity (FC) values for all the three sampling depths; 14-28 %, 18-33 % and 19-42 %, respectively. The predictive map for surface water content at FC showed moderate to high variation, with the majority of the area being classified as having 24 % (Fig. 5.22). The RFM predicted high values of FC water content at the southern part of the field as shown in (Fig. 5.22). 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. 5.23). The areas that were predicted to have a greater amount of FC water content were predominantly distributed across the field (Fig. 5.23). 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 % (Fig. 5.24). The predictive map of FC showed that greater values were predicted at the lower areas of the field and the south to south western parts of the field (Fig. 5.24).

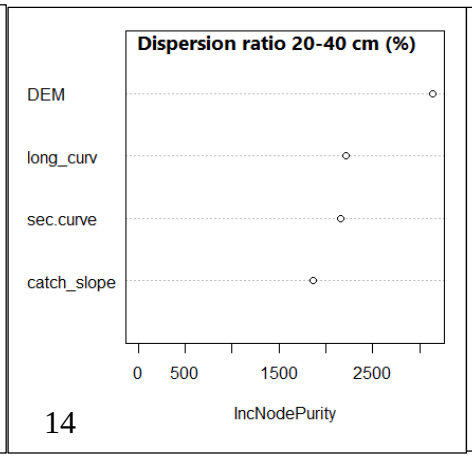
RFM predicted the spatial distribution of slow draining pores (SDP) within the field for three different sampling depths (Fig. 5.25 to 5.27). The model predicted the minimum and maximum SDP values for all the three sampling depths; 2-6 %, 1-6 % and 1-6 %, respectively. The predictive map of surface slow draining pores showed high variation within the field, with the majority of the areas classified as having 4 % (Fig. 5.25). The high volume of slow draining pores was predicted in the low lying slope on the western side of the field (Fig. 5.25). The predictive map of the subsurface SDP showed high variation, with the majority of the area being classified as having between 1.3-5.6 % (Fig. 5.26). 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 (Fig. 5.26). 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 % (Fig. 5.27). The higher values are distributed in the western side of the vineyard (Fig. 5.27).

The RFM predicted the minimum and maximum fast-draining pores (FDP) values for all the three sampling depths; 7-23 %, 7-14 % and 4-11 %, respectively (Fig. 5.28 to 5.30).The predictive map of soil surface FDP showed high variation, with the majority of the area being classified as having less than 11 % (Fig. 5.28). The highest pores are distributed in the eastern direction of the vineyard and the lowest were are at the western part of the field (Fig. 5.28). The predictive map for subsurface FDP again showed high variation, with the majority of the area being classified as being 6 % (Fig. 5.29). The low FDP values are scattered within the field. On the other hand, the predictive map of the lower depth FDP showed high spatial distribution within the field, with the majority of the area being classified as 11 % (Fig. 5.30). 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. 5.30).

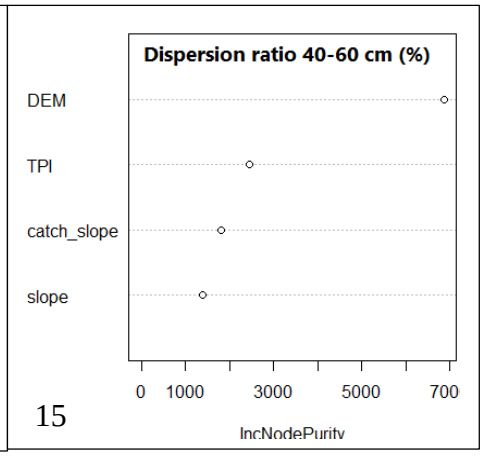

clay 0-20 cm (%)
DEM
TWI_1
catch_slope
convexivity
0 100 200 300 400
IncNodePurity
1
Clay 20-40 cm (%)
TPI
hill
aspect
ls_factor
0 200 400 600 800 1000 1200
IncNodePurity
2
Clay 40-60 cm (%)
mod_catcharea
long_curv
DEM
ls_factor
0 200 400 600 800 1200
IncNodePurity
3
Sand 0-20 cm (%)
catch_slope
slope
DEM
sec.curve
0 100 200 300 400 500 600
IncNodePurity
4
Sand 20-40 cm (%)
TPI
catch_slope
ls_factor
hill
0 200 400 600 800 1000
IncNodePurity
5
Sand 40-60 cm (%)
mod_catcharea
TPI
catch_area
aspect
0 200 400 600 800 1000
IncNodePurity
6
Silt 0-20 cm (%)
convexivity
GRVI
hill
MRRTF
0 50 100 150 200 250
IncNodePurity
7
Silt 20-40 cm (%)
MRVBF
GRVI
TPI
TWI_1
0 20 40 60 80 100 12
IncNodePurity
8
Silt 40-60 cm (%)
TWI_SAGA
GRVI
mod_catcharea
con_index
0 50 100 150
IncNodePurity
9
Bulk density 0-20 cm (Mg/m3)
aspect
TWI_SAGA
TBSI_T
DEM
0.00 0.05 0.10 0.15 0.20
IncNodePurity
10
Bulk density 20-40 cm (Mg/m3)
long_curv
ls_factor
sec.curve
catch_slope
0.00 0.02 0.04 0.06 0.08
IncNodePurity
11
Bulk density 40-60 cm (Mg/m3)
hill
TBSI_W
con_index
MBI
0.00 0.02 0.04 0.06 0.08
IncNodePurity
12



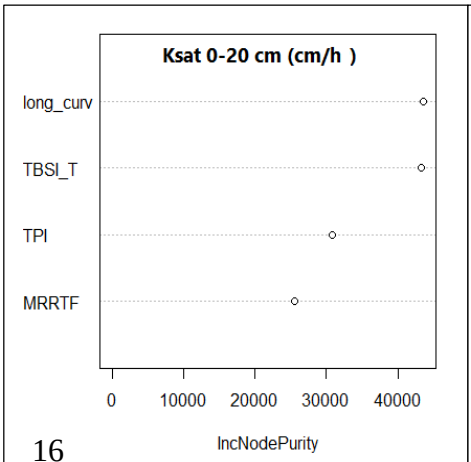
13



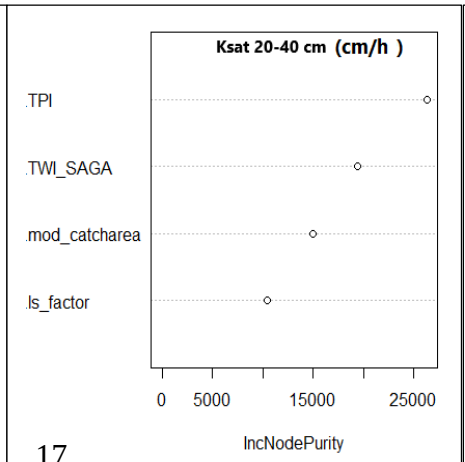
14



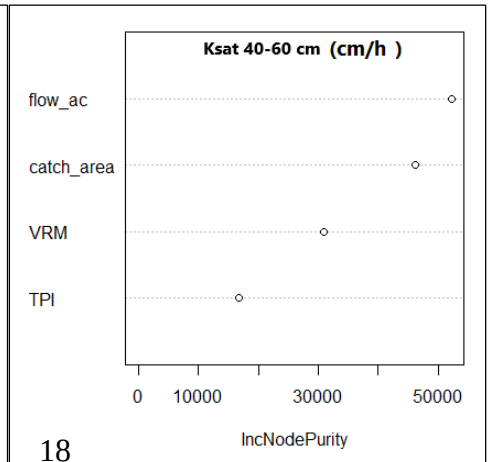
15



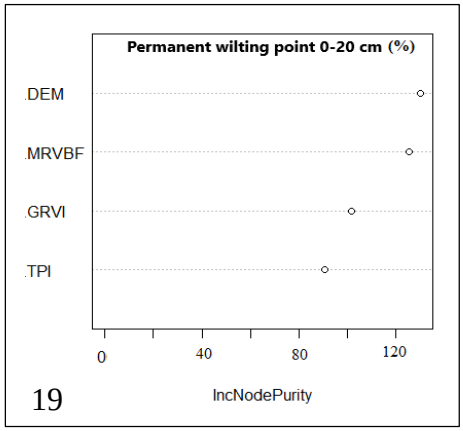
16



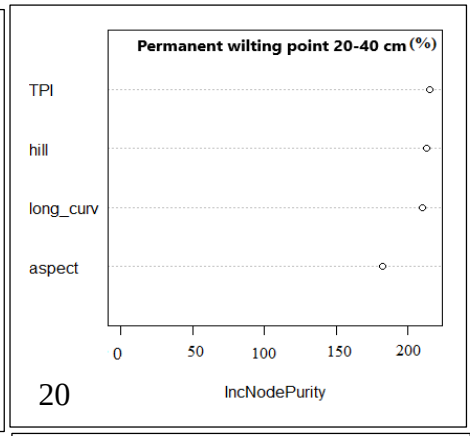
17



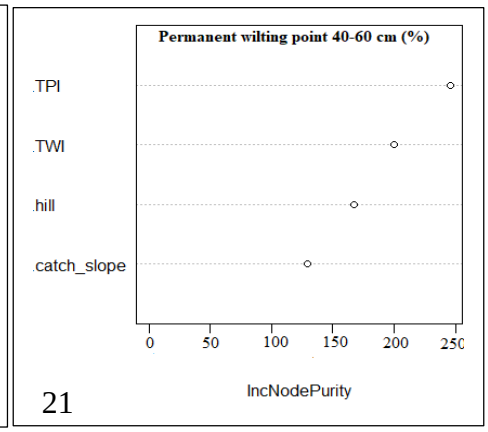
18



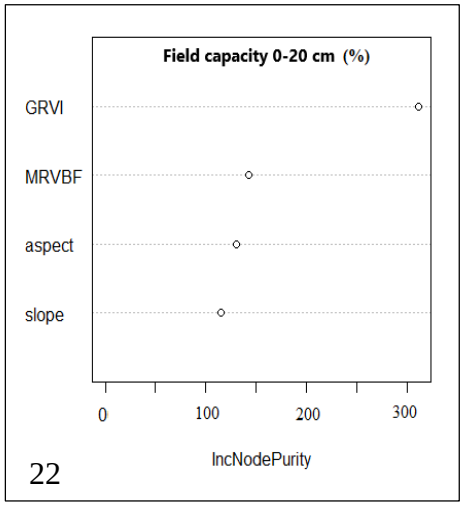
19



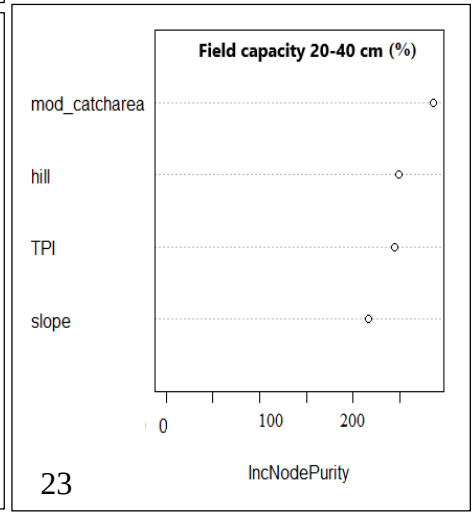
20



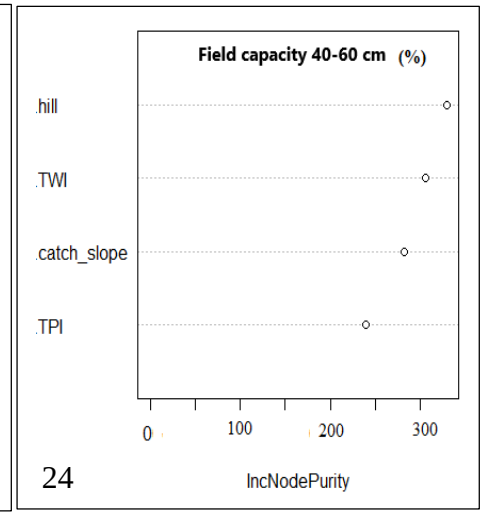
21



22



23



24

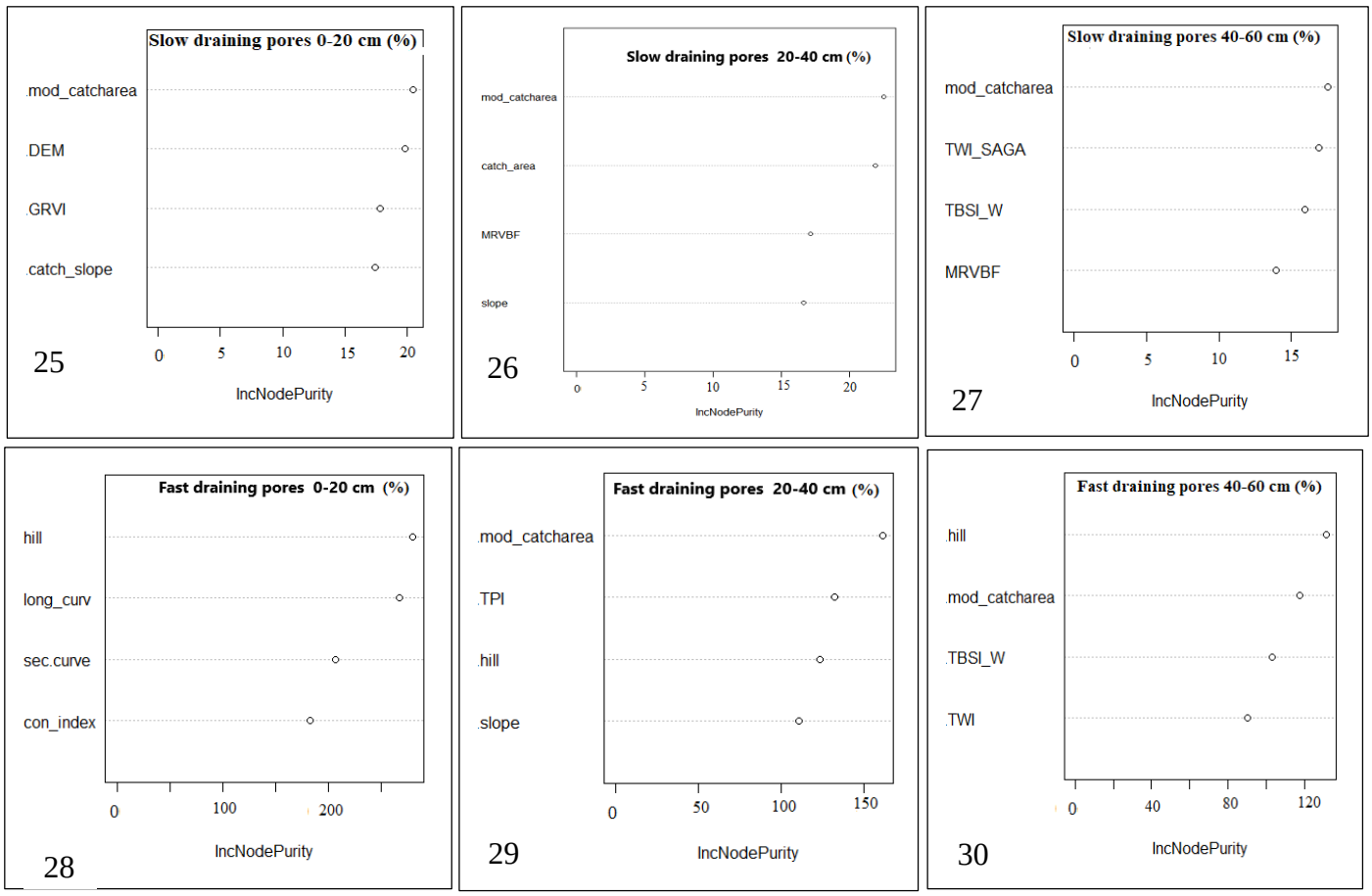
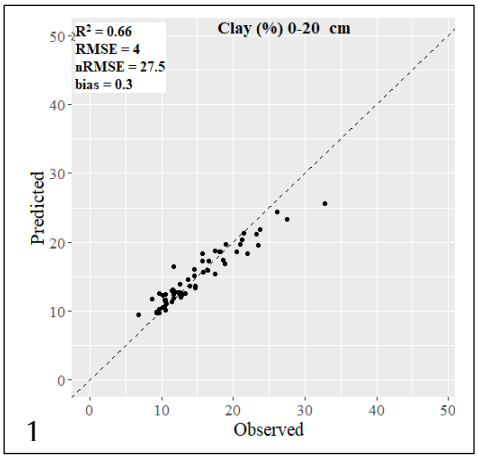
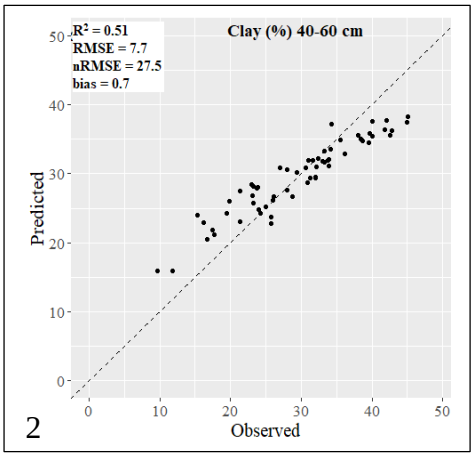
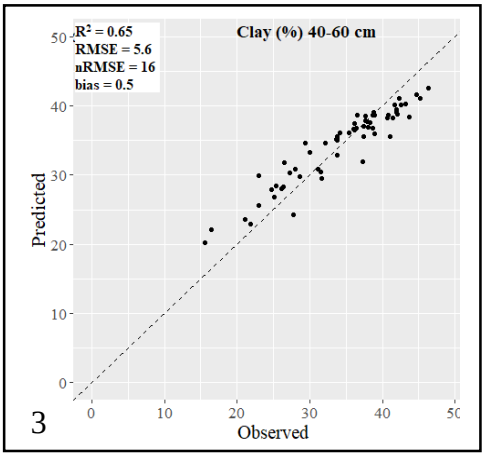
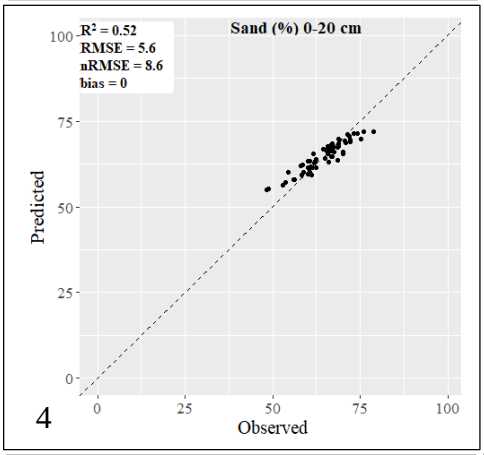


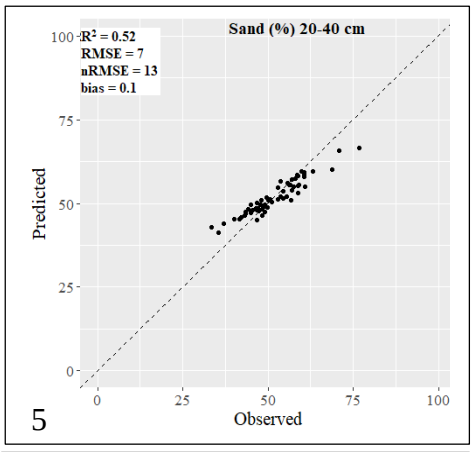
Figure 3. Illustration of variable importance derived from Random Forest model for prediction of soil properties within the vineyard.

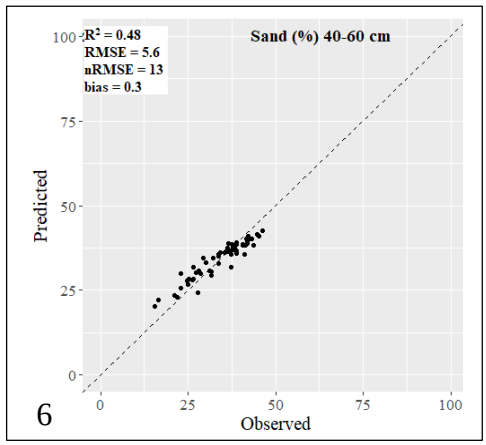

R² = 0.66
RMSE = 4
nRMSE = 27.5
bias = 0.3
Clay (%) 0-20 cm
Predicted
Observed
1

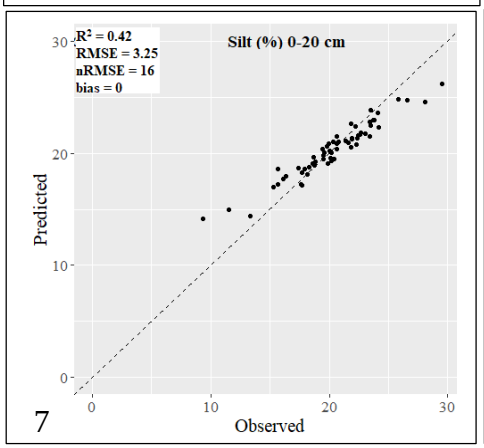

R² = 0.51
RMSE = 7.7
nRMSE = 27.5
bias = 0.7
Clay (%) 40-60 cm
Predicted
Observed
2

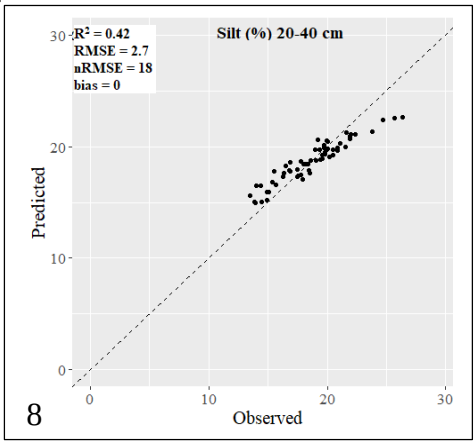

R² = 0.65
RMSE = 5.6
nRMSE = 16
bias = 0.5
Clay (%) 40-60 cm
Predicted
Observed
3

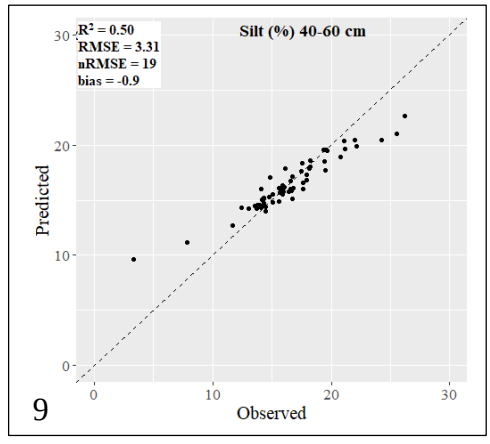

R² = 0.52
RMSE = 5.6
nRMSE = 8.6
bias = 0
Sand (%) 0-20 cm
Predicted
Observed
4

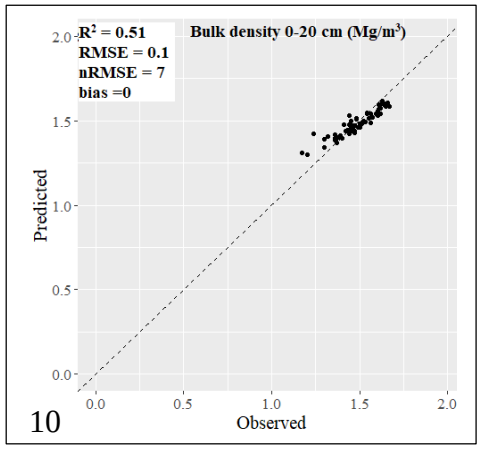

R² = 0.52
RMSE = 7
nRMSE = 13
bias = 0.1
Sand (%) 20-40 cm
Predicted
Observed
5

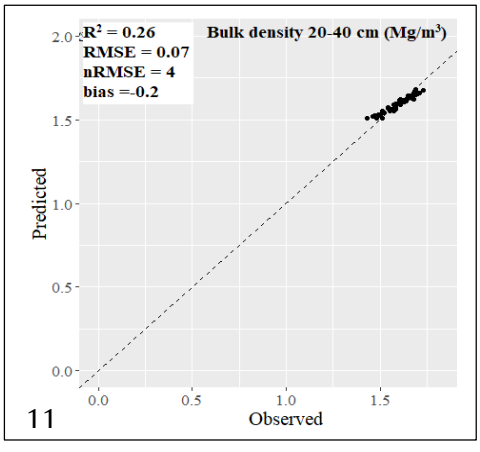

R² = 0.48
RMSE = 5.6
nRMSE = 13
bias = 0.3
Sand (%) 40-60 cm
Predicted
Observed
6

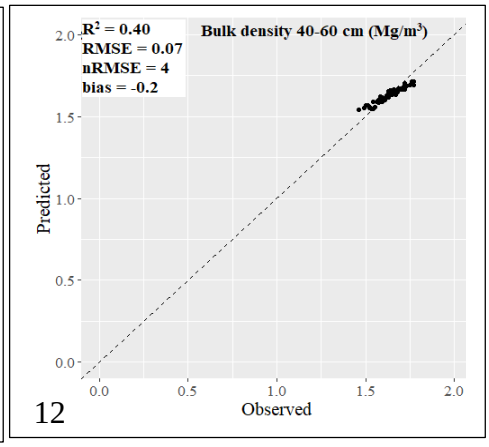

R² = 0.42
RMSE = 3.25
nRMSE = 16
bias = 0
Silt (%) 0-20 cm
Predicted
Observed
7

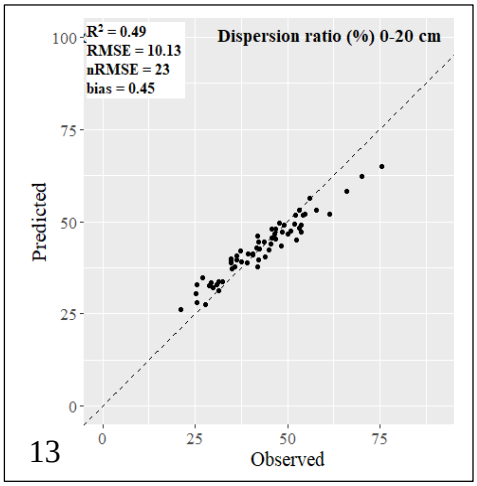

R² = 0.42
RMSE = 2.7
nRMSE = 18
bias = 0
Silt (%) 20-40 cm
Predicted
Observed
8

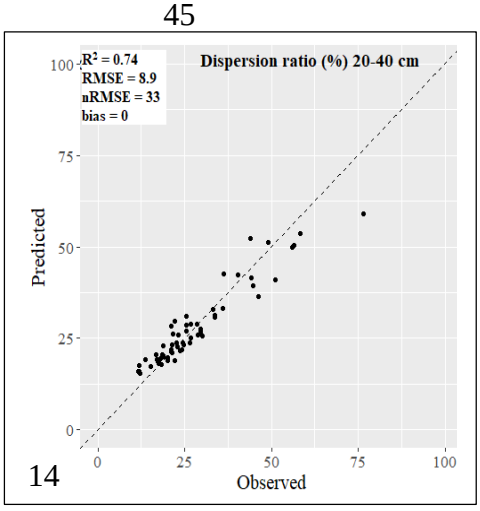

R² = 0.50
RMSE = 3.31
nRMSE = 19
bias = -0.9
Silt (%) 40-60 cm
Predicted
Observed
9

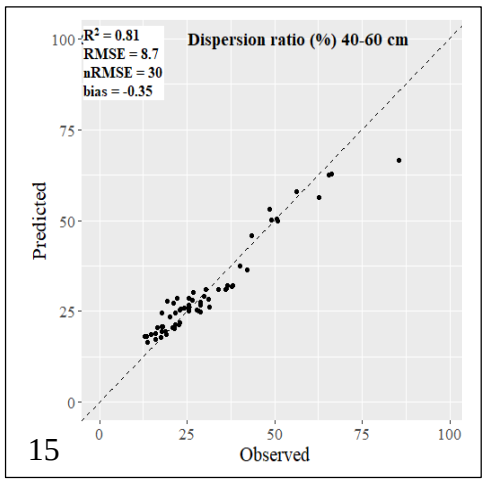

R² = 0.51
RMSE = 0.1
nRMSE = 7
bias =0
Bulk density 0-20 cm (Mg/m³)
Predicted
Observed
10

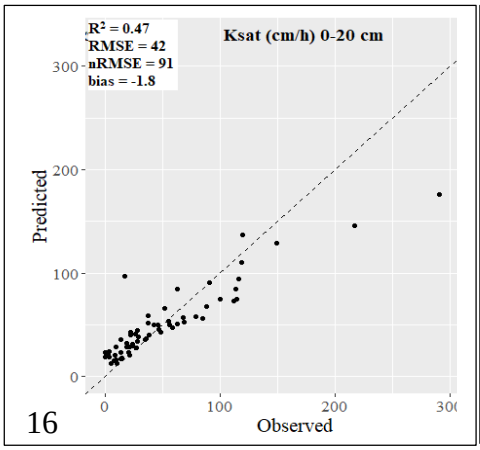

R² = 0.26
RMSE = 0.07
nRMSE = 4
bias =-0.2
Bulk density 20-40 cm (Mg/m³)
Predicted
Observed
11

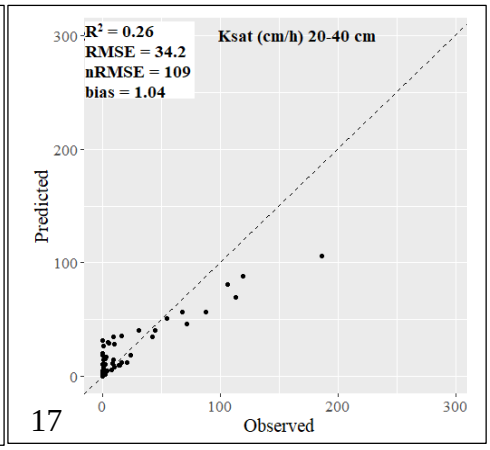

R² = 0.40
RMSE = 0.07
nRMSE = 4
bias = -0.2
Bulk density 40-60 cm (Mg/m³)
Predicted
Observed
12

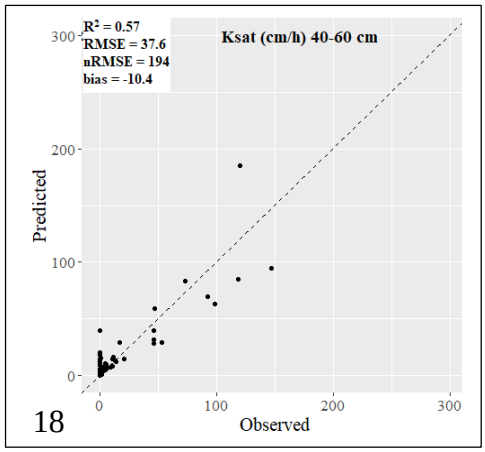

R² = 0.49
RMSE = 10.13
nRMSE = 23
bias = 0.45
Dispersion ratio (%) 0-20 cm
Predicted
Observed
13

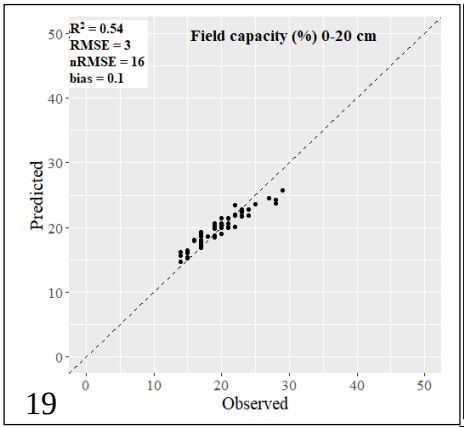

R² = 0.74
RMSE = 8.9
nRMSE = 33
bias = 0
Dispersion ratio (%) 20-40 cm
Predicted
Observed
14

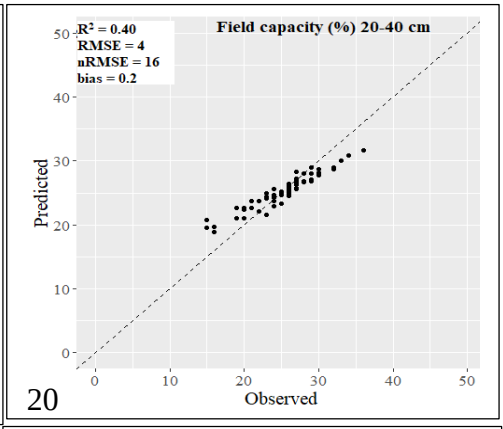

R² = 0.81
RMSE = 8.7
nRMSE = 30
bias = -0.35
Dispersion ratio (%) 40-60 cm
Predicted
Observed
15

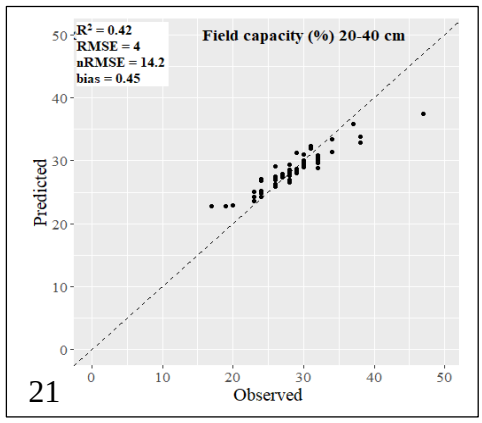

R² = 0.47
RMSE = 42
nRMSE = 91
bias = -1.8
Ksat (cm/h) 0-20 cm
Predicted
Observed
16

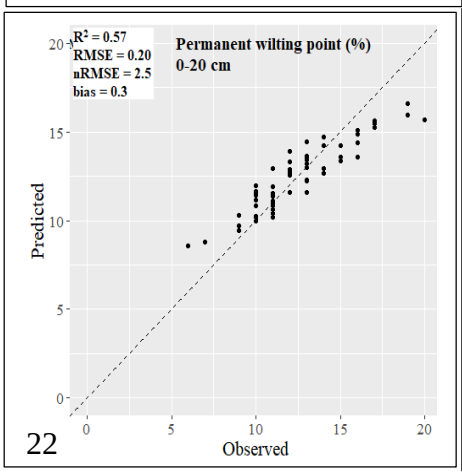

R² = 0.26
RMSE = 34.2
nRMSE = 109
bias = 1.04
Ksat (cm/h) 20-40 cm
Predicted
Observed
17

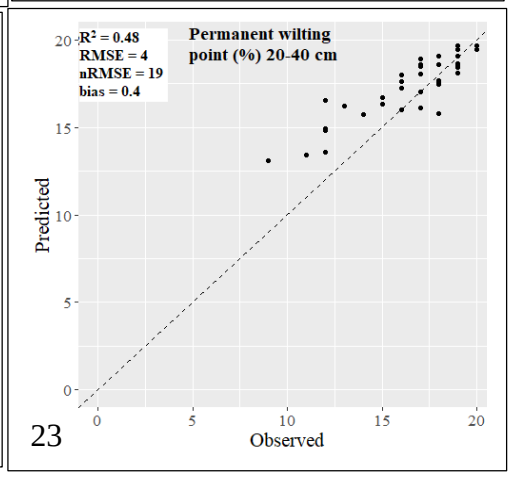

R² = 0.57
RMSE = 37.6
nRMSE = 194
bias = -10.4
Ksat (cm/h) 40-60 cm
Predicted
Observed
18

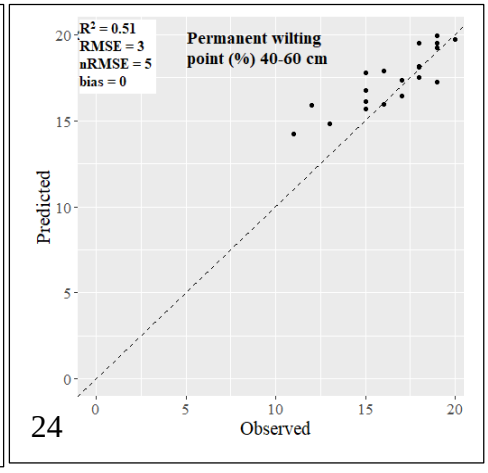

R² = 0.54
RMSE = 3
nRMSE = 16
bias = 0.1
Field capacity (%) 0-20 cm
Predicted
Observed
19


R² = 0.40
RMSE = 4
nRMSE = 16
bias = 0.2
Field capacity (%) 20-40 cm
Predicted
Observed
20


R² = 0.42
RMSE = 4
nRMSE = 14.2
bias = 0.45
Field capacity (%) 20-40 cm
Predicted
Observed
21


R² = 0.57
RMSE = 0.20
nRMSE = 2.5
bias = 0.3
Permanent wilting point (%) 0-20 cm
Predicted
Observed
22


R² = 0.48
RMSE = 4
nRMSE = 19
bias = 0.4
Permanent wilting point (%) 20-40 cm
Predicted
Observed
23


R² = 0.51
RMSE = 3
nRMSE = 5
bias = 0
Permanent wilting point (%) 40-60 cm
Predicted
Observed
24

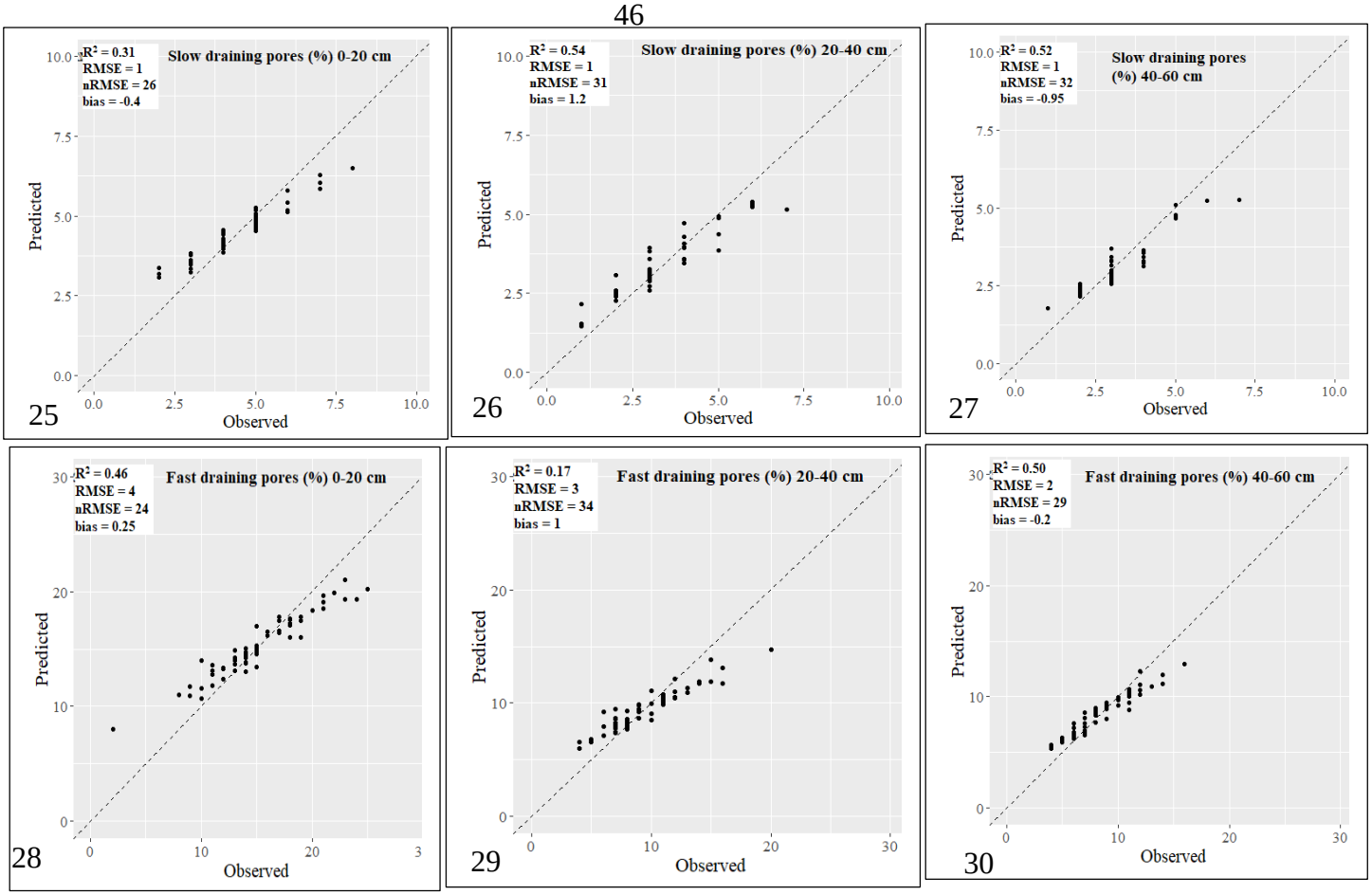
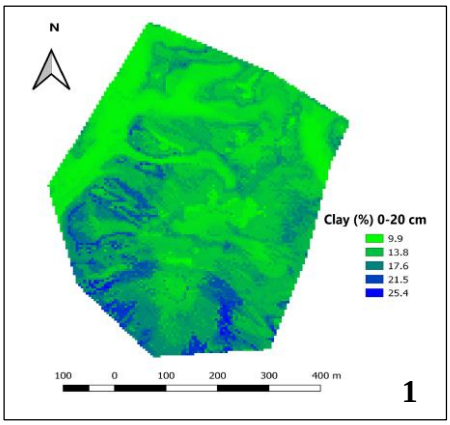
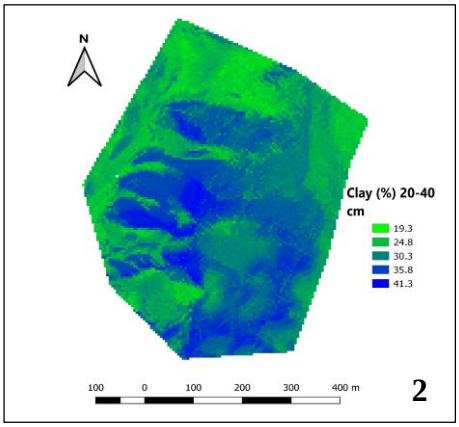
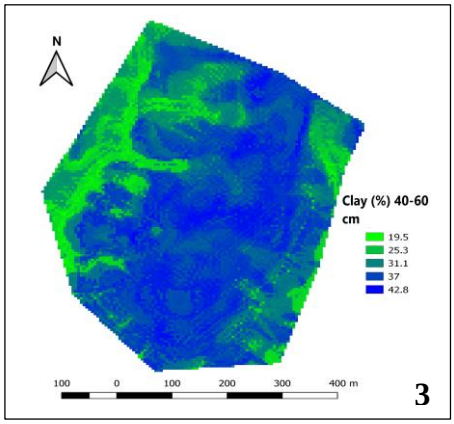
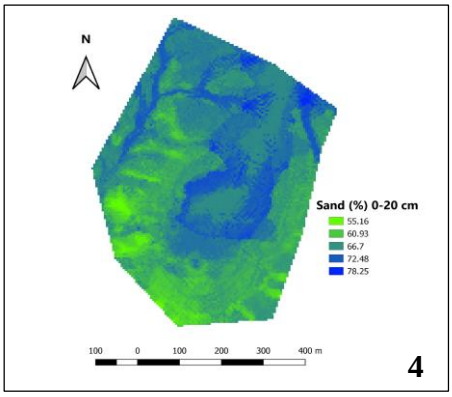


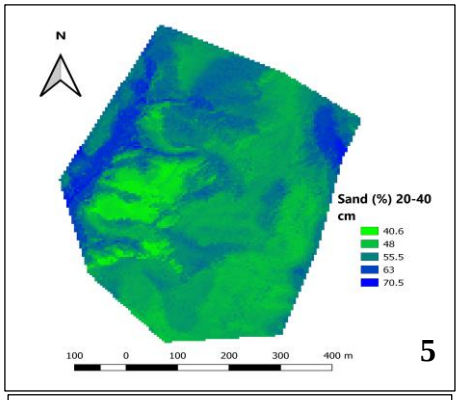
Figure 4. Validation of soil physical properties for the three soil sampling depths.

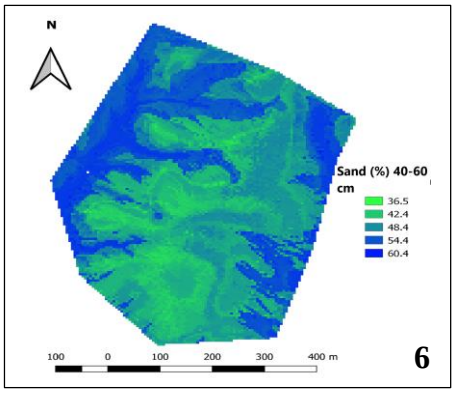

N
Clay (%) 0-20 cm
9.9
13.8
17.6
21.5
25.4
100 0 100 200 300 400 m
1

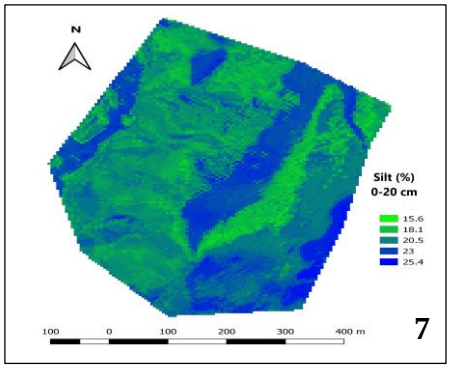

N
Clay (%) 20-40 cm
19.3
24.8
30.3
35.8
41.3
100 0 100 200 300 400 m
2

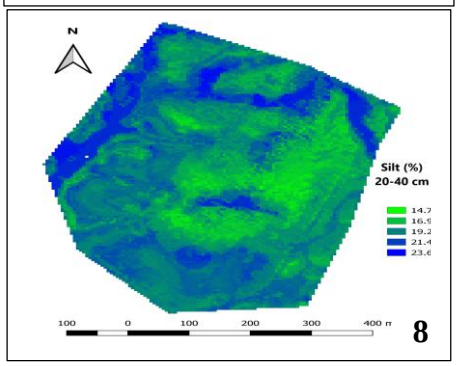

N
Clay (%) 40-60 cm
19.5
25.3
31.1
37
42.8
100 0 100 200 300 400 m
3

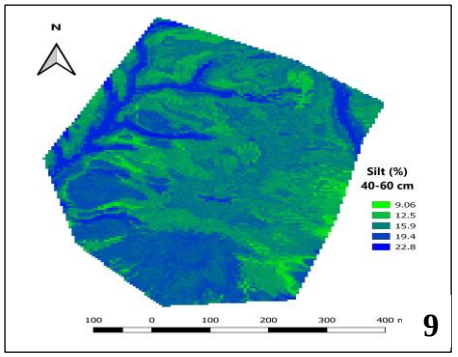

N
Sand (%) 0-20 cm
55.16
60.93
66.7
72.48
78.25
100 0 100 200 300 400 m
4

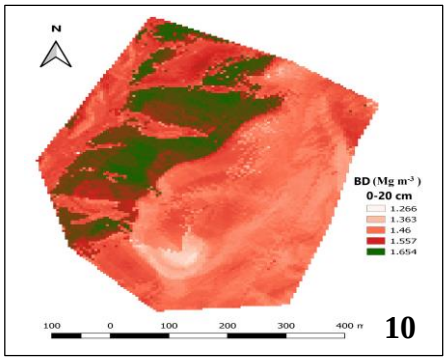

N
Sand (%) 20-40 cm
40.6
48
55.5
63
70.5
100 0 100 200 300 400 m
5

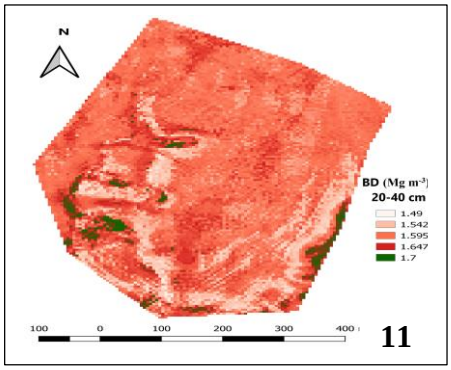

N
Sand (%) 40-60 cm
36.5
42.4
48.4
54.4
60.4
100 0 100 200 300 400 m
6

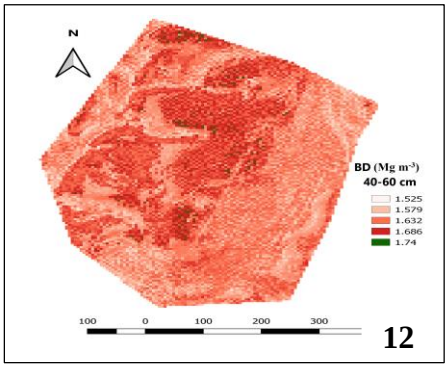

N
Silt (%) 0-20 cm
15.6
18.1
20.5
23
25.4
100 0 100 200 300 400 m
7

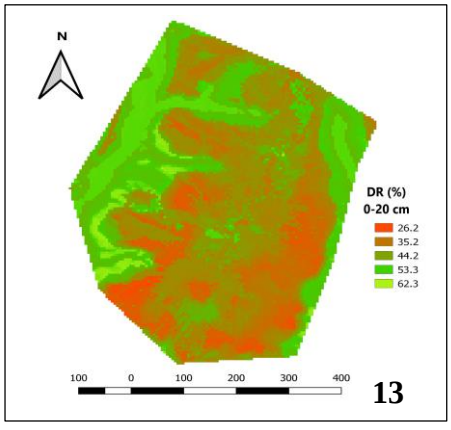

N
Silt (%) 20-40 cm
100 0 100 200 300 400 m
8

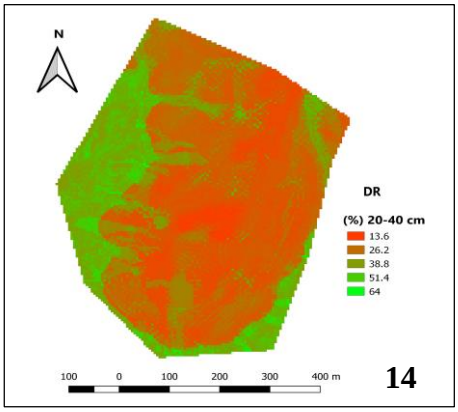

N
Silt (%) 40-60 cm
9.06
12.5
15.9
19.4
22.8
100 0 100 200 300 400 m
9

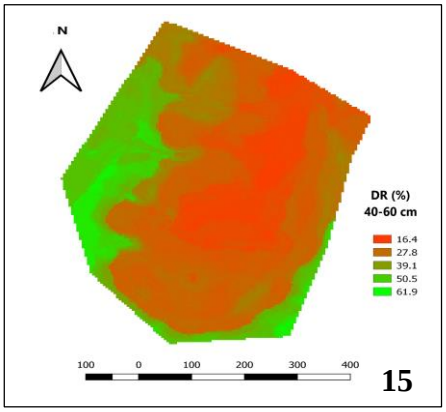

N
BD (Mg m-3) 0-20 cm
1.266
1.363
1.46
1.557
1.654
100 0 100 200 300 400 m
10


N
BD (Mg m-3) 20-40 cm
1.49
1.542
1.595
1.647
1.7
100 0 100 200 300 400 m
11


N
BD (Mg m-3) 40-60 cm
1.525
1.579
1.632
1.686
1.74
100 0 100 200 300
12


N
DR (%) 0-20 cm
26.2
35.2
44.2
53.3
62.3
100 0 100 200 300 400
13


N
DR (%) 20-40 cm
13.6
26.2
38.8
51.4
64
100 0 100 200 300 400 m
14


N
DR (%) 40-60 cm
16.4
27.8
39.1
50.5
61.9
100 0 100 200 300 400
15

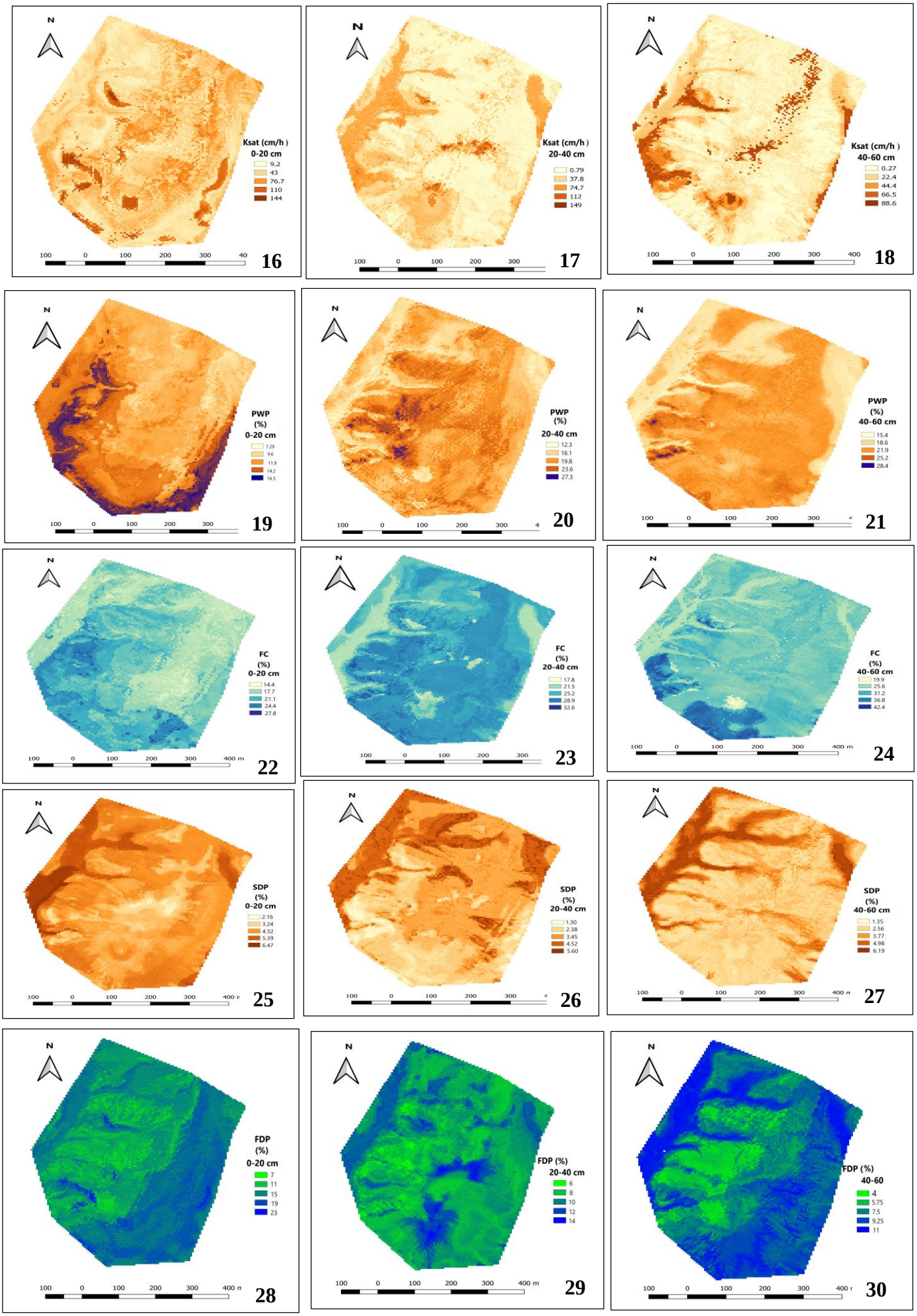


Fig. 5. Spatial distribution of predicted soil properties using RF model in vineyard for three layers.

5. DISCUSSION

5.1. Descriptive statistics

The studied site produces vines under dryland conditions, therefore the distribution and movement of soil moisture content and nutrients within the profile mainly depend on soil physical properties. Particle size distribution is one of the properties that drive other soil properties, for example, determining the potential soil water content that drives crop yield potential (Akpa *et al*., 2014). The results of particle size distribution (Tables 3 to 5) showed a high variability in fine particles (3-29 % silt and 6-46 % clay). The soil particle size distribution, i.e., sand, silt and clay, are vital in most hydrological, ecological, and environmental risk assessment models (Liess *et al*., 2012). The mean and variance of particle size varied with depth. The clay content increases from the top 20 cm depth with a peak at the 40 to 60 cm depth likely caused by clay illuviation and eluviation (Ayuba *et al*., 2007; Sharu *et al*., 2013) and movement of clay particles due to soil erosion (Amusan *et al*., 2005; Salako *et al*., 2006). The mean sand content is higher than that of clay and silt contents for each soil depth, which is commonly found is soil rich in quartz and granitic parent material as it is the case for this study (Sauer *et al*., 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. Changes in the sand content were to a greater extent influenced along the topo-sequence while that of silt and clay were greater along with profile depth.

The standard deviations (SD) and coefficient of variations (CV) showed moderate to high variation or heterogeneity in clay fraction compared with 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. The high variation could be due to depositions and differences in landscape. The skewness and kurtosis coefficient were used to describe the shape of the data distribution; the values for particle size distribution are slightly within the normal distribution range, which is 0 for skewness and 3 for kurtosis. The low kurtosis values for clay, silt, and sand in all the soil depth indicates that their data set have light trails and/or lack of outliers. The soil particle size fractions are one of the most important properties to influence soil physiochemical properties (e.g hydraulic properties, nutrient status, and water holding capacity) due to their correlation with other variables as indicated in Tables 6 to 8. Clay content had a positive significant correlation with other attributes such as shear strength, organic matter, water contents at field capacity and permanent wilting point. 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 in all the soil depths (Rawls *et al*., 2003). However, the negative correlation between clay content, fast drainage pores, slow drainage pores is explained by the pore sizes because clay soils hold more water through their micropores.

The obtained results showed that silt content negatively influenced the saturated hydraulic conductivity because of its particle size which doesn't 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 than micropores which drains first when suction starts during drying (Jury and Horton, 2004). Sand has low water holding capacity as it is shown by the results obtained in this study, where it proved to have a significant negative correlation with water retention at field capacity and at the permanent wilting point for all the soil depths. This is due to a high volume of macropores (fast drainage pores) that releases water quickly when exposed to suction leading to low water available for crops. The mean values of bulk density and particle density showed an increase in soil depth. This behaviour could be explained by the influence of organic matter, which decrease the soil bulk density by its own low value of bulk density (Seguel and Horn, 2006) and same applies with particle density (Sandoval *et al*., 2012; Adams, 1973). Nevertheless, no significant correlations were found among these properties and only particle density was negatively correlated with organic matter at 20-40 cm (Table 7).

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}), as well as particle density ($2.55 < 2.62 < 2.63$ Mg m^{-3}) for the three depths (Table 3, 4 and 5, respectively). Those soil particle density values can be explained by the dominance of the quartz mineral in the sand fraction of the soils across the vineyard, as was reported by Silva *et al*. (2001). However, it is the opposite with total porosity because the porosity decreased with soil depth and again this is associated with the decrease in organic matter and within the profile and high proportion of sand compared to clay or silt.

Based on CV values of bulk density, total porosity and particle density, both surface and subsurface depths soil had low spatial variability according to the CV values of bulk density, porosity set by Mulla and McBratney (2000) and their variability magnitude. The low variability of bulk density within the field is in agreement with the work of Osama *et al*. (2008) that aimed at investigating the spatial variability of penetration resistance. The low variability of bulk density and particle density could also be because of uniform parent material and land management in the area. In addition, bulk density is regarded as one of the properties that are less variable. The skewness and kurtosis showed that the data set had low trails, except for particle density at the lower depth, the reason could be the presence of an outlier. However, organic matter showed high variability in all the sampling depth which could be due to differences in terrain elevation variables (aspect, slope gradient, etc.) model of the area (Dahlgren *et al*., 1997).

As expected, organic matter had a significant positive correlation with water retention characteristics (FC, PWP) at all the sampling depth, this is due to the fact that presence of organic matter improves water holding capacity (Rawls *et al*., 2003), soil structure and structural stability index (Soto *et al*., 2018). However, the significant negative correlation is observed between organic matter and penetration resistance at the first soil depth (0-20 cm) but there is a positive correlation in the lower soil depth. On the other hand, total porosity showed a significant positive relationship with other soil properties like saturated hydraulic conductivity, fast draining pores and repellence index especially at the first two sampling depths. The reason could be the fact that total porosity is the summation of both macro and microporosity which then allows the hydraulic processes to occur. The results water retention characteristics showed that the water content at field capacity, permanent

wilting point and available water content increase with the increase in soil depth. The reason for this trend is associated with relatively higher clay content, low total porosity and high bulk density in the lower depth which allow clay particle to hold moisture to its micropores and less water is released (Warrick, 2002).

The water that is adsorbed by micropores is not available for plant uptake. The pore size distribution means 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 moderate and 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). High significant positive correlations were found when water content at field capacity (33 kPa) and permanent wilting point (1,500 kPa) are correlated with clay and silt contents, organic matter, penetration resistance and shear strength. 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 ($p < 0.05, < 0.01$) negative correlation was observed between PWP, FC and sand content, saturated hydraulic conductivity (Tables 6 to 8). This was expected because sandy soil contains a high volume of macropores that all fast drainage of water 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 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). The pore size distribution also plays an important role in water movement within the soil. In this sense, it was observed that fast drainage pores (FDP) had a significant positive correlation with total porosity, sand, and saturated hydraulic conductivity, and significant negative relationship with bulk density, water content at field capacity and permanent wilting point in almost all the horizon. The skewness and kurtosis of FC, PWC, SDP, and FDP showed that the data set lacked the outliers, except for available water content in all the soil depth.

The results obtained in this study showed that saturated hydraulic conductivity (Ksat) decreases with the increase in soil depth ($50.6 > 22.6 > 18.4$ cm h^{-1}) (Tables 3 to 5). The reason for this variation within the soil profile is due to the fact that Ksat it is affected by many 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. Similar conclusion was reported by Rasse *et al*. (2000) and Iqbal *et al*. (2005).

The CV values of Ksat increases with soil depth (104.4 < 197.3 < 229.9 %) and because the values are greater than 48 %, the Ksat is considered as highly variable within the field, according to Mulla and McBratney (2000). There are several talks about the magnitude of the spatial variability of Ksat 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 CV for Ksat in the range of 50-300%. Therefore the present study is in agreement with the work that was reported by the above literature.

The skewness and kurtosis of Ksat showed that the data set were not normally distributed and therefore the data was log transformed to make it normal because it was positively skewed and contained slight trails. As result, mean values and median values were not similar. The Pearson correlation results show that Ksat had a significant ($p < 0.01$ and < 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 ($p < 0.01$, 0.05) positive correlation with sand, total porosity and fast drainage pores (Table 6, 7 and 8). 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.

Soil aggregate stability is one of the indicators of soil quality and plays a very important role in the movement and storage of water and air throughout the soil profile. For this study the soil macroaggregates and micro-aggregate stability were evaluated by dispersion methods (ΔWMD and DR, respectively), and structural stability index (SSI) was determined for the three sampling depths. The mean values of the DR index, associated with micro-aggregate stability, showed an increase with the increase in 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 lowest soil depths as was suggested by Brunel *et al*. (2016) in the same soil series with different tillage systems. 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. On the other hand, the ΔWMD associated with macro-aggregate stability showed an increase with the increase in soil depth, which indicates a decrease in macro-aggregate stability with the increase in soil depth. The reason for this variation is associated with the decrease of soil organic matter in depth. 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 soil aggregate stability is related to soil structure, therefore structural stability index (SSI) results showed that there is a decrease with the increase in soil depth. The high values of SSI in the upper soil depths indicates the resistance of the soil to structural degradation. However, in the lower soil depth, there is a risk of structural degradation according to the lowest values of OM. The reason for this change is related to the variation in soil organic matter within the soil profile, as it is well known that OM that helps bind the soil particles together to form a stable soil structure (Reynolds *et al*., 2009; Soto *et al*., 2018). Serme *et al*. (2016) use the SSI to assess the soil structural stability with

different amounts of organic amendments and concluded that the SSI increases with the rise in soil organic matter content. The larger content of organic matter in the upper horizon promotes the stability as the effect of slow humectability of aggregates, generating organic-mineral bonds with soil particles and preventing the dispersive effect of water (Ellies *et al*., 1995; Chenu *et al*., 2000; Dexter, 2002). However, high amounts of clay and silt contents are observed in the lower soil depths. The CV values of micro-aggregate, macro-aggregate and structural stability index indicates 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 influenced by the micro-topographical variations operating over a different period of time which influences the presence of soil organic carbon (Minasny *et al*., 2008). The dataset indicates moderate to high skewness. The values of skewness vary from -0.01 to 1.60 for all the soil depths (Tables 3 to 5). Highly skewed variables indicate that these properties have a local distribution; that is high values were found for some sampling points, but most values were low. However, these high soil test values may not always be an outlier but a management induced variation (Vasu *et al*., 2017).

Soil compaction is one of the most complex phenomena that are difficult to measure directly due to its multifaceted relationship with other physical properties of the soil (Horn and Fleige, 2009). This phenomenon involves disruption and reduction of larger pores which then lead to an increase in bulk density; however, only the shear strength was negatively correlated with FDP (coarse pores) in the upper two soil layers and no relationships were observed with BD values. The mean values of soil PNT and shear strength increased with the increasing soil depth (Tables 3 to 5). The average PNT and shear strength values of the study area were related to the soil moisture regime which depends soil texture. The shear strength increased with soil depth due to the nature of soils in the region defined as residual soils from granitic material with less intense pedogenic process at depth. Similar results were reported by Soto *et al*. (2018). There is an inversely proportional relationship between penetration resistance, shear strength, and root growth because high penetration resistance restricts root growth and or penetration (Horn *et al*., 2007). However, for this study, the resistance values were not critical for both surface and subsurface. Different authors presented different critical values of PNT, e.g between 1.0 and 3.5 MPa (Merotto and Mudstock, 1999) and 2.0 and 8.0 MPa (Schoeneberger *et al*., 2012) as values of moderate and high restrictions for rooting. Nevertheless, Taylor and Bran (1991) determined that PNT values higher than ≈ 0.2 MPa (200 kPa) generate a lineal decrease in root development.

The CV values of PNT and shear strength indicates high to very high variability in penetration resistance within the field. The high variability is associated to differences in textural properties and moisture content. These differences have resulted from the processes of deposition that reflect the composition and properties of material deposited spatially and temporarily, which contribute to the heterogeneity within the field. The results showed a significant positive correlation between PNT and shear strength, clay content and PWP and a significant negative correlation to organic matter, available water content, dispersion ration, slow draining pores, and sand content. The obtained results were in agreement with the work that was reported by Gülser *et al*. (2016) and Acuna (2013). When water content decreases, an increase in mechanical resistance occurs because of the diminution of cohesion within the solid fraction of the silty soils. On the other hand, there is a significant positive correlation between soil shear strength and clay

content, PNT, FC, PWP, and significant negative correlation to sand content, structural stability index, fast draining pores, and slow draining pores.

When studying water repellency, it is important to examine soil aggregates as individual structural entities. The measurement of soil water repellency for each soil aggregate was done using a device that was proposed by Leeds-Harrison *et al*. (1994) which measures the hydraulic properties of individual soil aggregate. The mean values of soil water repellency index for all the aggregates decreased with the increasing soil depth (Tables 3, 4 and 5). The mean *R*-index values showed that surface if water repellent because the *R* index is greater than 1.78 while the subsurface is regarded as hydrophilic, since the *R* index values are less than 1.78. The main reason could be that the stability of the aggregates is mainly influenced by the quantity and quality of organic matter in the soil (Piccolo and Mbagwu, 1999), where some compounds favor stability to water, either by a cementing action in the points of contact or by its water repellency, which reduces the speed of humidification of aggregates (Chenu *et al*., 2000). The excessive hydrophobicity of the soil in the surface, is caused by the presence of coatings hydrophobic on solid particles, results in a reduction in the rate of wetting and a water repellency in the soil (Hallet, 2007; Urbanek *et al*., 2007). This water repellency is due to the accumulation of certain types of hydrophobic organic compounds, which do not present a degree of alteration very advanced, and that they have their origin in the decomposition of organisms, mainly plants and microbial (Doerr *et al*., 2000). In general, the repellency increases with increases in organic matter and decreases with increases in soil clay contents (Harper *et al*., 2000).

Based on the CV values of *R* index, both surface and subsurface soil had very high spatial variability because the CV values are greater than 30 % according to Gomez and Garcia (2002). In general, CV values increased with increasing soil depth. Similar results were reported by Liu *et al.* (2009) which demonstrated high variability of soil water repellency within the field. In spite of skewness and kurtosis of the distribution of soil water repellency, the mean and median values were similar with means being equal or almost equal to the median. There is a significant negative correlation between the *R* index and bulk density and significant positive with total porosity within the surface. However, the correlation was not significant between *R* index and other soil properties in the subsurface (Tables 6 to 8).

5.2 Statistical models

In predicting the spatial variability of the particle size distribution (clay, sand, and silt) for three different soil depths within the vineyard, RFM performed significantly better at the top 20 cm compared with the lower sampling depths for all the particles. The opposite was observed for silt content, with the performance improvements with depth. Similar 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 density with depth. Most of the environmental covariates used in this research more based on land surface features are likely to have a stronger relationship or correlation with topsoil than subsoil properties. Several studies suggested that prediction performance for lower depths could be improved by the inclusion of covariates such as Gamma-radiometric or

electromagnetic induction (EM38) which allows accurate prediction of soil properties up until approximately 2 meters (Cook *et al*., 1996; Rawlins *et al*., 2009; Priori *et al*., 2014). However, considering the extent of Chile, the cost of acquiring such data may be extremely high to offset the extra benefit. In terms of prediction accuracy, sand content had the highest RMSE values across all the depths whereas the lowest RMSE was associated with the modeling of silt content at all the depths. This variation is in agreement 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 disagreement 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. There was a large influence of analytical Hillshading, topographic position index, DEM, LS Factor, modified catchment area, catchment slope and topographic wetness index on the particle size distribution. 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 the low lying 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 and soil particle size distribution 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). The predictive maps showed that soil texture varied considerably across the field (Fig. 5.1-5.9). Soil high in clay content were observed at lower elevations in a landscape (Schimel *et al*., 1985; Bonifacio *et al*., 1997) due to the influence of erosion relative to other soil forming processes.

RFM 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 model identified aspect, topographic wetness index, three-band spectral index and elevation as important covariates for predicting surface bulk density, which is associated with soil mineralogy. From the list of important covariates used to predict the bulk density, aspect proved to be the most important variable of them all. Therefore the removal of the aspect 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). The model for subsurface bulk density listed longitudinal curvature, slope length factor, secondary curvature and catchment slope as major players to influence spatial variability pattern of soil bulk density. The best effect of curvature was its combinations, particularly in concave and convex and as a result of this effect, bulk density was predicted to be high at the south to southwest which is characterized by domination on concave curvature.

Validation of bulk density prediction RFM at all depths using R^2 and RMSE indices showed that surface soil layer had higher prediction accuracy (i.e., higher R^2 and lower RMSE) compared to the layers at lower soil depths. These values were 0.51 and 0.10, and 0.26 and 0.10. 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 bulk density distribution, soils in the north of the field appeared to have high density than soils at the centre of the vineyard and the eastern area, 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.

Soil aggregate stability is considered as an important indicator of soil quality in the landscapes witnessing land degradation due to erosion by water. The stability of the microaggregates (1-2 mm) was determined by the dispersion ratio (DR) method. 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 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 played an important role in predicting the microaggregate stability. Several studies also showed that spatial variability of aggregate stability is closely related to the terrain attributes such as slope, curvature, and aspect through their impacts on various soil properties (Rhoton and Duiker, 2008).

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). The RF model performed significantly better at the lower soil depth compared with the 20 cm. The Validation of microaggregate stability prediction RFM at the surface using R^2, nRMSE and bias showed that lower depth (40-60 cm) had higher prediction accuracy (i.e., higher R^2, lower bias, and nRMSE) compared to the topsoil. The values of R^2, nRMSE and bias for topsoil and subsoil were: 0.49, 23 and - 0.5; 0.81, 30 and -0.4, respectively. The differences in model accuracy are associated with environmental covariates. However, all the depths has demonstrated promising results and high values of the determinant coefficient in the case of dispersion ratio.

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. The predictive Ksat results showed the minimum and maximum predicted values of saturated hydraulic conductivity within the study area for all the three sampling depths (0-20 cm, 20-40 cm, and 40-60 cm). 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 soil properties. There were areas within the field that showed very low Ksat more especially the areas at the centre of the field, which makes them more at risk of being eroded. There was a large influence of longitudinal curvature, TBSI-T, TPI, MrRTF, TWI-SAGA, modified catchment area, LS factor, flow accumulation and VRM (Fig. 3.16, 3.17 and 3.18) on the saturated hydraulic conductivity. However, the relative importance of these variables

varies with depths. Topographic wetness played a crucial role in predicting the Ksat due to the geographic formation of the vineyard that is composed of different slopes. Therefore TWI described soil moisture pattern in the watershed based on topography. The valley bottom and the centre of the vineyard are characterized by low slope, which permitted the water accumulation. During the fieldwork, it was evident that some areas that had low Ksat were washed off or eroded.

The validation of Ksat prediction RF models at the surface using R^2, nRMSE and bias showed that lower depth had higher prediction accuracy (i.e., higher $R^2 = 0.57$, higher bias = -10.4 and nRMSE = 194) compared to the topsoil with values of R^2, nRMSE and bias for topsoil: $R^2 = 0.47$, bias = -1.8 and nRMSE = 91, respectively. The variation of in prediction accuracy of the models can be associated with 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, 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.

The water storage capacity of the vineyard was quantified by water holding capacity (difference between water content at field capacity and permanent wilting point). The predictive maps of water retention characteristics showed the increase of water content (FC, and PWP) with soil depth, the higher amount of available water content can be associated with good quality soil structure. 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 the low slope gradient, 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).

For describing spatial variation of the soil hydraulic properties; water content at field capacity (FC), and permanent wilting point (PWP) at three different depths the most important environmental covariates were DEM (elevation), MrVBF, TPI, longitudinal curvature, topographic wetness index, catchment area, Green-Red vegetation index

(GRVI), modified catchment area, analytical Hillshading, TBSI-V, length-slope factor across the field. However, the relative importance varied with soil depths. The predictors such as elevation and topographic curvature were expected to affect the water content because they affect the local climate (microclimate). It was also interesting to see remote sensing predictors such as green-red vegetative index and TBSI-V 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 final RFM were used for estimating the deterministic component for each soil hydraulic property (PWP, and FC). The performance of the final RF models is summarized in Table 9. R^2 varies between 0.54–0.57, and 0.42– 0.54 for PWP, and FC, respectively. RMSE was 0.03–0.04, and 0.01–0.03 cm^3 m^{-3} for PWP, FC and AWC, respectively. The variation in accuracy can be associated with environmental covariates and their direct influence.

The RFM showed promising results in predicting the soil 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 RFM 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 vines are very sensitive to moisture content and they become water-stressed when there is very limited water content.

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 within the soil properties. 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 Baver and Gardner (1972). The predictive maps of pores size distribution showed the decrease of fast-draining pores (FDP, >50 μm) with soil depth (see Fig. 5.28 to 5.30), 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 other reason could be associated to the present of organic matter which influences the soil structure. The main reason to characterize these 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 (Fig. 5.25-5.27). 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. These pores hold water against gravity, however, the water can still be able to be absorbed by the vines. The SDP predictive maps for all the depths demonstrate that the capillary pores are more in the areas with flat slope, more especial in the West and Northwest direction of the vineyard (Fig. 5.25 to 5.27). Because flatter areas respond to sediment deposition from the surrounding hillsides with the accumulation of fine particles such as clay and silt, which contains a high amount of micropores.

For describing spatial variation of the slow draining and fast-draining pores for all the sampling depths, modified catchment area, digital terrain model (DEM), catchment area, TWI-SAGA, analytical Hillshading, longitudinal curvature, and TPI played a significant role in predicting the pore size distribution due to their high node purity (Fig. 3.25-3.30). The other predictor variables to influence the prediction included Green-Red vegetation index, MrVBF, slope, and TBSI-W (Fig. 3.25-3.30). However, their relative importance varies with depth. Predictor variables such as MrVBF represent areas of accumulation of sediments according to the topographic surrounding context. Taylor *et al*. (2013) reported MrVBF as a relevant predictor of water table depths under the Mediterranean landscape. The slope also played a role in the final model because it influences the movement of water and particles within the field. Higher values of Multi-resolution Index of Valley Bottom Flatness (MrVBF) represent areas where the accumulation of sediments are not possible, such as highly elevated areas of the vineyard. Therefore the lower areas had the lowest MrVBF and accumulation of sediments was possible. The performance of the final RF models is summarized in Table 9. R^2 varies between 0.31–0.54, and 0.20–0.50 for SDP and FDP, respectively. nRMSE was 26–32 % and 24–34% for SDP and FDP, respectively. Bias was -0.4-1.2 and -0.2-1.0 for SDP and FDP, respectively. The variation in accuracy can be associated with environmental covariates and their direct influence in modeling the soil properties. 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.

CONCLUSIONS

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

Low to very high spatial variability in soil properties were observed across the experimental field but the magnitude varied with soil depth. The six soil properties that showed a considerable degree of variation were clay, repellency index, shear strength, saturated hydraulic conductivity, available water content (AWC), and dispersion ratio. The applied DSM approach including Random Forest model (RFM) as a relatively tool in the field of soil science for prediction of soil physical properties yielded promising results as accuracy of the model and generated prediction maps were acceptable.

There was a negative and positive correlation amongst the soil properties, OM had a significant positive correlation with soil water retention characteristics (AWC, PWP, and FC) and 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 higher depths 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, Green-Red vegetation index (GRVI), Three band spectral index (TBSI-W), Three band spectral index (TBSI-V) and Three band spectral index (TBSI-T), Multi resolution Index of Valley Bottom Flatness (MrVBF), Digital Elevation Model (DEM), slope aspect, convergence index, and length slope factor.

The remote sensing data played a role in the prediction and can be used with success as input for DSM for the surface depths. RFM provided a promising framework for the spatial prediction of soil properties as the accuracy of the model performance was acceptable. However RFM overestimated areas with lower values and underestimated areas with higher values. RFM predicted dispersion ratio, clay content, sand content, FC, PWP, and Ksat significantly well. The RFM can be improved by applying predictor variables that are directly related to the variable of interest. The prediction error may be improved by broadening the sampling design to include more replicates.

The performance in soil depth could be improved by inclusion of covariates such as Gamma radiometric and electromagnetic induction. As DSM 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. Finally, the produced predictive maps can be used for modeling agricultural system productivity in the near future.

BIBLIOGRAPHY

Acuna, J.C. 2013. The effects of winter cover crops and soil compaction treatments on soil properties and soybean production in Illinois [Master's Thesis]. Champaign (IL): University of Illinois Urbana-Champaign.

Adams, W.A. 1973. The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils. *Journal of Soil Science* 24(1): 10–17.

Adhikari, K., B. Minasny, M.B. Greve, and M.H. Greve. 2014. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. *Geoderma* 214–215: 101–113.

Adhikari, K., R.B. Kheir, M.B. Greve, P.K. Bøcher, B.P. Malone, B. Minasny, A.B. McBratney and M.H. Greve. 2013. High-resolution 3-D mapping of soil texture in Denmark. *Soil Science Society of American Journal* 77: 860–876.

Adhikari, K., A. Guadagnini, G. Toth, and T. Hermann. 2009. Geostatistical Analysis of Surface Soil Texture from Zala County in Western Hungary. pp: 219–224. *In*: Proceedings of the International Symposium on Environment, Energy and Water in Nepal: Recent Researches and Direction for Future.

Akpa, S., T.F. Bishop, I.A. Odeh and A.E Hartemink. 2014. Digital mapping of soil particle-size fractions for Nigeria. *Soil Science Society of America Journal* 78(6): 1953-1956.

Amusan, A.A., A. Olayinka and D.J. Oyedele. 2005. Genesis, classification, and management requirements of soils formed in windblown material in the Guinea Savanna area of Nigeria. *Communications in Soil Science and Plant Analysis* 36: 2015–2031.

Agyare, W.A., S.J. Park and P.L.G. Vlek. 2007. Artificial neural network estimation of saturated hydraulic conductivity. *Vadose Zone Journal* 6(2): 423–431.

Ayuba, S.A., F.O.R. Akamigbo and S.A. Itsegha. 2007. Properties of soils in River Katsina-Ala catchments area, Benue State, Nigeria. *Nigerian Journal of Soil Science* 17: 24 – 29.

Bacis Ceddia, M., S.R. Vieira, A.L. Oliveira Villela, L. dos Santos Mota, L.H. Cunha dos Anjos, and D. Fonseca de Carvalho. 2009. Topography and spatial variability of soil physical properties. *Agricultural Science* 66: 338-352.

Bauer, J., H. Rohdenburg and H.R. Bork. 1985. Ein Digitales Reliefmodell als Vorraussetzung fuer ein deterministisches Modell der Wasser- und Stoff-Fluesse. *Landschaftsgenese und Landschaftsoekologie* 10: 1-15.

Baver, L.D., and W.R. Gardner. 1972. Soil physics. 4th ed. John Wiley & Sons, New Jersey, USA.

Beaudette, D.E., P. Roudier and A.T. O'Geen. 2013. Algorithms for quantitative pedology: A toolkit for soil scientists. *Computers and Geosciences* 52: 258–268.

Ben-Dor, E., D. Heller and A. Chudnovsky. 2008. A novel method of classifying soil profiles in the field using optical means. *Soil Science Society of America Journal* 72: 1–13.

Bernhard, N., L.M. Moskwa, K. Schmidt, 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., and D.R. Nielsen. 1976. Spatial variability of leaching characteristics of a field soil. *Water Resource Research* 12:78–84.

Boehner, J., and O. Antonic. 2009. Land-surface parameters specific to topo-climatology. *Developments in Soil Science* 33: 195-226.

Böhner, J., and T. Selige. 2006. Spatial prediction of soil attributes using terrain analysis and climate regionalisation. pp: 13-27. *In*: J. Böhner, K. McCloy and J. Strobl (Eds.) SAGA - Analysis and modelling applications. Alemania: Goettinger Geographische Abhandlungen.

Bonifacio, E., E. Zanini, V. Boero and M. Franchini-Angela. 1997. Pedogenesis in a soil catena on serpentinite in north-western Italy. *Geoderma* 75: 33–51.

Bradford, J.M. 1986. Penetrability. pp: 463-478. *In*: Klute, A. (Ed.). Methods of soil analysis. 2nd ed. Madison: American Society of Agronomy, *Soil Science Society of America*.

Bramley, R.G.V. 2005. Understanding variability in winegrape production systems. 2. Within vineyard variation in quality over several vintages. *Australian Journal of Grape and Wine Research* 11: 33–42.

Bramley, R.G.V. and A.P.B. Proffitt. 1999. Managing variability in viticultural production. *Grapegrower and Winemaker* 427: 11-16.

Bramley, R.G.V. and D.W. Lamb. 2003. Making sense of vineyard variability in Australia. pp. 33–54. *In*: Precision viticulture (eds. R. Ortega and A. Esser) Proceedings of an international symposium held as part of the IX Congreso Latinoamericano de Viticultura y Enología, Chile. Centro de Agricultura de Precisión, Pontificia Universidad Católica de Chile, Facultad de Agronomía e Ingeniería Forestal, Chile.

Brady, N.C. and R.R. Weil. 2008. The Nature and Properties of Soils. 14th ed. Pearson Prentice Hall, Upper Saddle River, New York, USA.

Breiman, L. 2001. Random forests. *Machine Learning* 45: 5–32.

Breiman, L. 1996. Bagging Predictors. *Machine Learning* 24(2): 123–140.

Breiman, L., J.H. Friedman, R.A. Olshen and C.I. Stone. 1984. Classification and regression trees. Belmont, California, Wadsworth.

Brown, C., M. Fingas and R. Marois. 2006. Oil spill Remote Sensing flights around Vancouver Island. pp: 921–930. *In*: Proceedings 29th Arctic and Marine Oil Spill Program (AMOP) Tech. Seminar, Vancouver, Canada, June. 6–8.

Brown, D.J., M.K. Clayton and K. McSweeney. 2004. Potential terrain controls on soil color, texture contrast and grain-size deposition for the original catena landscape in Uganda. *Geoderma* 122: 51–72.

Bruland, G.L., S. Grunwald, T.Z. Osborne, K.R. Reddy and S. Newman. 2006. Spatial distribution of soil properties in Water Conservation Area 3 of the Everglades. *Soil Science Society of America Journal* 70(5): 1662–1676.

Brunel-Saldias, N., I. Martínez, O. Seguel, C. Ovalle, and E. Acevedo. 2016. Structural characterization of a compacted alfisol under different tillage systems. *Journal of Soil Science and Plant Nutrition* 16(3) 689-701.

Brungard, C.W., J.L. Boettinger, M.C. Duniway, S.A. Wills and T.C. Edwards Jr. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma* 239-240: 68-83.

Buchanan, S., J. Triantafilis, I. Odeh and R. Subansinghe. 2012. Digital soil mapping of compositional particle-size fractions using proximal and remotely sensed ancillary data. *Geophysics* 77(4): 201-211.

Burgess, T.M. and R. Webster. 1980. Optimal Interpolation and Isarithmic Mapping of Soil Properties: I. The Variogram and Punctual Kriging. *Journal of Soil Science* 31: 315-331.

Cabezas, J., M. Galleguillos and J.F. Pérez–Quezada. 2016. Predicting vascular plant richness in a heterogeneous wetland using spectral and textural features and a random forest algorithm. *IEEE Geoscience and Remote Sensing Letters* 13: 646–650.

Candemir, F., and C. Gülser. 2012. Influencing factors and prediction of hydraulic conductivity in fine textured-alkaline soils. *Arid Land Research Management* 26(1): 15-31.

Cantón, Y., A. Solé-Benet, C. Asensio, S. Chamizo and J. Puigdefábregas. 2009. Aggregate stability in range sandy loam soils relationships with runoff and erosion. *Catena* 77(3): 192-199.

Castillo-Riffart, I., M. Galleguillos, J. Lopatin and J.F. Pérez-Quezada. 2017. Predicting vascular plant diversity in anthropogenic peatlands: Comparison of Modeling Methods with Free Satellite Data. *Remote Sensing* 9(7): 681.

Chagas, C.S., H.S.K. Pinheiro, W. Carvalho Junior, L.H.C. Anjos, N.R. Pereira and S.B Bhering. 2017. Data mining methods applied to map soil units on tropical hillslopes in Rio de Janeiro, Brazil. *Geoderma Regional* 9: 47-55.

Chaudhari, P.R., Dodha V. Ahire, Vidya D. Ahire, M. Chkravarty, and S. Maity. Soil Bulk Density as related to Soil Texture, Organic Matter Content and available total Nutrients of Coimbatore Soil. *International Journal of Scientific and Research Publications* 3(2): 2250-3153.

Cerri, D.G.P., and P.S.G. Magalhães .2012. Correlation of Physical and Chemical Attributes of Soil with Sugarcane Yield. *Pesquisa Agropecuaria Brasileira* 47: 613-620.

Chenu, C., Y. Le Bissonnais and D. Arrouays. 2000. Organic matter influence on clay wettability and soil aggregate stability. *Soil Science Society of America Journal* 64: 1479-1486.

Chen, C., and R.J. Wagenet. 1992. Simulation of water and chemicals in macropore soils. Part 1. Representation of the equivalent macropore influence and its effect on soil water flow. *Journal of Hydrology* 130: 105-126.

CIREN (Centro de Información de Recursos Naturales Chile). 1997. Estudio Agrológico Región VII. Descripciones de suelos, materiales y símbolos. Santiago, Chile. (Publicación 117).

Cook, S.E., R.J. Corner, P.R. Groves and G.J. Grealish. 1996. Use of airborne gamma radiometric data for soil mapping. *Australian Journal Soil Research* 34: 183–194.

Conrad, O. System for Automated Geoscientific Analyses (SAGA). Version: 2.1.2. Available online: http://www.saga–gis.org (accessed on 29 July 2014).

Costa, N.H. de A.D. et al. 2002. Novo método de classificação de coeficiente de variação para a cultura do arroz de terras altas. Pesquisa Agropecuária Brasileira, Brasília 37(3): 243-249.

Crouvi, O., J.D. Pelletier and C. Rasmussen. 2013. Predicting the thickness and aeolian fraction of soils in upland watersheds of the Mojave desert. *Geoderma* 195: 94-110.

Cutler, D.R, T.C Jr. Edwards., K.H. Beard, A. Cutler, K.T. Hess, J. Gibson and J. Lawler. 2007. Random forests for classification in ecology. *Ecology* 88: 2783–2792.

Dahlgren, A.R., L.T. Bottinger, L.G. Huntington and A.R. Amundson. 1997. Soil development along an elevation transect in the western Sierra Nevada, California. *Geoderma* 78: 207–236.

Dane, J., and G. Topp. 2002. Methods of soil analysis. Part 4. Physical methods. SSSA Book Series 5. Soil Science Society of America, Madison, Wisconsin.

De Gryze, S., J. Six, H. Bossuyt, K. Van Oost and R. Merckx. 2008. The relationship between landform and distribution of soil C, N and P under conventional and minimum tillage. *Geoderma* 144: 180–188.

Debella-Gilo, M., and B. Etzelmüller. 2009. Spatial prediction of soil classes using digital terrain analysis and multinomial logistic regression modeling integrated in GIS: Examples from Vestfold County, Norway. *Catena* 77(1): 8–18.

Dent, D. and A.Young. 1981. Soil Survey and Land Evaluation. Georg Allen and Unwin Publishers, London.

Deutsch, C.V., A.G. Journel. 1998. GSLIB: Geostatistical Software and User's Guide, 2nd Ed. Oxford University Press, New York, USA.

Dexter, A., 2002. Soil structure: the key to soil function, *Advances in Geoecology* 35: 57-69.

Dikau, R. 1988. Entwurf einer geomorphographisch-analytischen systematik von reliefeinheiten. *Heidelberger Geographische Bausteine* 5: 1-45.

Dobos, E., F. Carré, T. Hengl, H.I. Reuter and G. Tóth. 2006. Digital Soil Mapping as a support to production of functional maps. EUR 22123 EN, Office for Official Publications of the European Communities, Luxemburg.

Dobos, E., E. Micheli, M.F. Baumgardner, L. Biehl and T. Helt. 2000. Use of combined digital elevation model and satellite radiometric data for regional soil mapping. *Geoderma* 97(3-4): 367-391.

Doerr, S. H., R.A. Shakesby and R.P.D. Walsh. 2000. Soil water repellency: its causes, characteristics and hydro-geomorphological significance. *Earth-Science Reviews* 51(1-4): 33–65.

Dokuchaev, V. V. 1883. Russkij Cernozem. Sankt Petersburg

D'Or, D. and P. Bogaert. 2004. Spatial prediction of categorical variables with the Bayesian Maximum Entropy approach: The Ooypolder case study. *European Journal of Soil Science* 55(4): 763–775.

Ellerbrock, R.H., H.H. Gerke, J. Bachmann and M.O. Goebel. 2005. Composition of organic matter fractions for explaining wettability of three forest soils. *Soil Science Society of America Journal* 69: 57-66.

Ellies, A., Grez, and R., Ramírez, C. 1995. Potencial de humectación y estabilidad estructural de los agregados de suelos sometidos a diferentes manejos. *Agricultura Técnica* 55(3-4): 220-225.

Ettema, C.H. and D.A. Wardle. 2002. Spatial soil ecology. *Trends Ecology and Evolution* 17: 177–183.

Eynard, A., T.E. Schumacher, R.A. Kohl, and D.D. Malo, 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.

Fairfield, J., and P. Leymarie. 1991. Drainage networks from grid digital elevation models. *Water Resources Research* 27: 709-717.

Fassnacht, F., C. Neumann, M. Förster, H. Buddenbaum, A. Ghosh, A. Clasen, *et al*. 2014b. Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 7: 2547–2561.

Fassnacht, F., F. Hartig, H. Latifi, C. Berger, J. Hernández, P. Corvalán. *et al*. 2014a. Importance of simple size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment* 154: 102-114.

Felzensztein, C., G. Echecopar and K. Deans. 2011. Marketing strategy, innovation and externalities: the case of the Chilean wine cluster, BALAS Conference, Universidad Adolfo Ibanez, Santiago, April, 2011.

Ferrero A., B. Usowicz and J. Lipiec. 2005. Effects of tractor traffic on spatial variability of soil strength and water content in grass covered and cultivated sloping vineyard. *Soil and Tillage Research* 84: 127–138.

Ferreyra, R., G. Sellés, P. Gil, P. Maldonado, G. Cabezas y V. Rodríguez. 2001. Diagnóstico de la situación de las plantaciones de frutales en cerro, Provincia de Quillota. Instituto de Investigaciones Agropecuarias. Informe nº 4, PROVALTT Quillota, Julio.

Freeman, G.T. 1991. Calculating catchment area with divergent flow based on a regular grid. *Computers and Geosciences* 17: 413-22.

Gallant, J.C. and T.I Dowling. 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research* 39(12): 1347-1359.

Gomes, F.P. and C.H. Garcia. 2002. Estadística aplicada a experimentos agronómicos e florestais. [Statistics Applied to Agronomic Experiments and Forestry.] FAELQ, Piracicaba.

Gochis, D. J., E.R. Vivoni, and C.J. Watts. 2010. The impact of soil depth on land surface energy and water fluxes in the North American monsoon region, *Journal of Arid Environments* 74: 564-571.

Goovaerts, P., 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. *Biology and Fertility of Soils* 27: 315–334.

Graham, R. C., and A.T. O'Geen. 2010. Soil mineralogy trends in California landscapes. *Geoderma* 154(3-4): 418–437.

Greve M. H., R. B. Kheir, M. B. Greve, and P. K. Bøcher. 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. *Ecology Indicators* 18: 1–10.

Greve M. H., R. B. Kheir, M. B. Greve and P. K. Bøcher. 2012b. Using digital elevation models as an environmental predictor for soil clay contents. *Soil Science Society of America Journal* 76: 2116–2127.

Grossman, R.B., and T.G. Reinsch. 2002. The solid phase. Bulk density and linear extensibility. pp: 201-228. *In:* Dane, J.H. and G.C. Topp (Eds.). Methods of soil analysis. Part 4. Physical Methods. Soil Science Society of America. Book Series Nº 5. Madison, Wisconsin, USA.

Grunwald, S. 2006a. What do we really know about the space–time continuum of soil-landscapes? pp. 3–36. In S. Grunwald (ed.) Environmental soil-landscape modeling: Geographic information technologies and pedometrics. CRC Press, Boca Raton, FL.

Guisan, A., S.B. Weiss and A.D. Weiss. 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* 143: 107-122.

Gülser C., I. Ekberli, F. Candemir, and Z. Demir. 2016. Spatial variability of soil physical properties in a cultivated field. *Eurasian Journal of Soil Science* 5(3): 192-200.

Guyon, I, J. Weston and S. Barnhill. 2002. Gene selection for cancer classification using support vector machines. *Machine Learning* 46: 389–422.

Hallet P. 2007. An introduction to soil water repellency. International Society for Agrochemical Adjuvants (ISAA). 8th International Symposium on Adjuvants for Agrochemicals (ISAA2007). Scottish Crop Research Institute.

Hallett, P.D. and I.M. Young. 1999. Change to water repellence of soil aggregates caused by substrate-induced microbial activity. *European Journal of Soil Science* 50: 35-40.

Harper R., I. Mckissock, R. Gilkes, D. Carter and P. Blackwell. 2000. A multivariate framework for interpreting the effects of soil properties, soil management and land use on water repellency. *Journal of Hydrology* 231-232: 371-383.

Hartge, K., und R. Horn. 2009. Die physikalische Untersuchung von Böden. Praxis Messmethoden Auswertung. 4. Vollst. Überarbeitete Auflage. Schweizerbart Vorlage, Stuttgart.

Henderson, B.L., N.B. Elisabeth, J.M. Christopher, and D.A.P. Simon. 2005. Australia-wide predictions of soil properties using decision trees. *Geoderma* 124: 383–398.

Hengl, T., 2007. A practical guide to geostatistical mapping of environmental variables. EUR 22904, Luxembourg. Office for Official Publications of the European Communities.

Hengl, T., G. B. M. Heuvelink and D. G. Rossiter. 2007. About regression-kriging: From equations to case studies. *Computers and Geosciences* 33(10): 1301–1315.

Heung B, H.C. Ho, J. Zhang, A. Knudby, C.E. Bulmer, M.G. Schimdt. 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping *Geoderma* 265: 62-77.

Heung B, C.E. Bulmer and M.G. Schmidt .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.

Horn, R., and H. Fleige. 2009. Risk assessment of subsoil compaction for arable soils in Northwest Germany at farm scale. *Soil and Tillage Research* 102: 201-208.

Horn, R., J. Vossbrink, S. Peth, and S. Becker. 2007. Impact of modern forest vehicles on soil physical properties. *Forest Ecology and Management* 248(1-2): 56–63.

Hu, C., A. Wright and G. Lian. 2019. Estimating the Spatial Distribution of Soil Properties Using Environmental Variables at a Catchment Scale in the Loess Hilly Area, China. *International Journal of Environmental Research and Public Health*, 16(3): 491.

Hutchinson, M.F., J.C. Gallant. 2000. Digital elevation models and representation of terrain shape. pp. 29–50. *In*: Wilson, J.P., and J.C. Gallant. (Eds.), Terrain Analysis: Principles and Applications. John Wiley and Sons, New York.

Iqbal, J., J.A. Thomasson, J.N. Jenkins, P.R. Owens and F.D. Whisler. 2005. Spatial Variability Analysis of Soil Physical Properties of Alluvial Soils. *Soil Science Society of America Journal* 69(4): 1338.

Iqbal, J., J.J. Read, A. Thomasson and J.N. Jenkins. 2005. Relationships between soil-landscape and dryland cotton lint yield. *Soil Science Society of America Journal* 69: 872–882.

Isaaks, H. E. and R.M. Srivastava. 1989. An Introduction to Applied Geostatistics, Oxford University Press, New York.

Isenburg, M. LAStools-Efficient Tools for LiDAR Processing, Software for Rapid Converting, Filtering, Viewing, Gridding, and Compressing of Lidar. Version: 140221. Available online: http://lastools.org (accessed on 30 July 2014).

INE, 2008. Catastro vitivinícola, informe anual 2008. Instituto Nacional de Estadísticas, Chile. Departamento de Estadísticas Agropecuarias y Medioambientales.

Iverson, L. R., A. M. Prasad, S. N. Matthews and M. Peters. 2008. Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *Forest Ecology and Management* 254: 390-406.

Iwahashi, J. and R.J. Pike. 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology* 86 (3–4): 409–440.

Jabro J.D., W.B. Stevens, R.G. Evans, and W.M. Iversen. 2010. Spatial variability and correlation of selected soil properties in the Ap horizon of a CRP grassland. *Application of Agricultural Engineering* 26(3): 419–428.

Jenny, H. 1980. The soil resource: origin and behaviour. Springer, New York.

Jenny, H., 1941. Factors of soil formation - a system of quantitative pedology. McGraw-Hill, New York.

Jury, W.A., and R. Horton. 2004. Soil physics. 6th ed. John Wiley & Sons, New York.

Kempen, B., D.J. Brus and J.J. Stoorvogel. 2011. Three-dimensional mapping of soil organic matter content using soil type–specific depth functions. *Geoderma* 162(1-2): 107-123.

Kempen, B., D.J. Brus, G.B.M. Heuvelink and J.J. Stoorvogel. 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: a multinomial logistic regression approach. *Geoderma* 151: 311-326.

Kempen, B., D.J. Brus, J.J. Stoorvogel, G.B.M. Heuvelink and F. Vries. 2012. Efficiency comparison of conventional and digital soil mapping for updating soil maps. *Soil Science Society America Journal* 76: 2097-2115.

Kettler, T.A., J.W. Doran, and T.L. Gilbert. 2001. Simplified Method for Soil Particle-Size Determination to Accompany Soil-Quality Analyses. *Soil Science Society of American Journal* 65: 849-852.

Kilic, K., S. Kilic and R. Kocyigit. 2012. Assessment of spatial variability of soil properties in areas under different land use. *Bulgarian Journal of Agricultural Science* 18: 722–732.

Koethe, R. and F. Lehmeier. 1996. SARA - System zur Automatischen Relief-Analyse. User Manual, 2. Edition [Dept. of Geography, University of Goettingen, unpublished]

Kotlar, A.M., B.V. Iversen, and Q. de Jong van Lier. 2019. Evaluation of parametric and nonparametric machine-learning techniques for prediction of saturated and near-saturated hydraulic conductivity. *Vadose Zone Journal* 18: 180-141.

Kong, X., T.H. Dao, J. Qin, H. Qin, C. Li and F. Zhang. 2009. Effects of soil texture and land use interactions on organic carbon in soils in North China cities' urban fringe. *Geoderma* 154: 86-92.

Kreznor, W.R., K.R. Olson, W.L. Banwart and O.L. Johnson. 1989. Soil, landscape, and erosion relationships in a northwest Illinois watershed. *Soil Science Society of America Journal* 53: 1763-1771.

Kudrat, M., A. K. Tiwari, S. K. Saha and S. K. Bhan. 1992. Soil resource mapping using IRS 1A LISS-II digital data – A case study of Kandi area adjacent to Chandigarh, India. *International Journal on Remote Sensing* 13(17): 3287-3302.

Kuhn, M (Contributions from Jed Wing), S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper *et al*. 2017. Caret: Classification and Regression Training. R package version 6.0-76.

Kursa, M. B. and W. R. Rudnicki. 2010. Feature Selection with the Boruta Package. *Journal of Statistical Software* 36(11): 1-13.

Lacoste, M., B. Lemercier and C. Walter. 2011. Regional mapping of soil parent material by machine learning based on point data. *Geomorphology* 133: 90–99.

Ladha, L. and T. Deepa. 2011. Feature Selection Methods and Algorithms. *International Journal on Computer Science and Engineering* 3(5): 1787-1797.

Lagacherie, P., and A.B. McBratney. 2007. Spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. pp. 3–22. *In*: P. Lagacherie *et al*. (ed.) Digital soil mapping: An introductory perspective. *Elsevier*, New York.

Leeds-Harrison, P.B., E.G. Youngs and B. Uddin. 1994. A devise for determining the sorptivity of soil aggregates. *European Journal of Soil Science* 45: 269-272.

Liaw, A. and M. Wiener. 2002. Classification and regression by RandomForest. *R News* 2(3): 18–22.

Liess, M., B. Glaser and 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, S., N. An, J. Yang, S. Dong, C. Wang and Y. Yin. 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.

Liu, X.B., X.Y. Zhang, Y.X. Wang, Y.Y. Sui, S.L. Zhang, S.J. Herbert and G. Ding . 2010. Soil degradation: a problem threatening the sustainable development of agriculture in Northeast China. *Plant Soil and Environment* 56: 87-97.

Liu, X., W. Zhang, M. Zhang, D.L. Ficklin and F. Wang. 2009. Spatio-temporal variations of soil nutrients influenced by an altered land tenure system in China. *Geoderma* 152: 23–34.

Lombardi, F., and M. Lualdi. 2019. Step-Frequency Ground Penetrating Radar for Agricultural Soil Morphology Characterisation. *Remote Sensing* 11(9): 1075-1109.

Lopatin, J., K. Dolos, H.J. Hernández, M. Galleguillos and F.E. Fassnacht. 2016. Comparison generalized linear model and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. *Remote Sensing of Environment* 173: 200–210.

Lopatin, J., M. Galleguillos, F.E. Fassnacht, A. Ceballos, and J. Hernández. 2015. Using a multistructural object-based LiDAR approach to estimate vascular plant richness in Mediterranean forests with complex structure. *IEEE Geoscience and Remote Sensing Letters* 12: 1008–1012.

Malone, B. P., A.B. McBratney and B. Minasny. 2013. Spatial Scaling for Digital Soil Mapping. *Soil Science Society of America Journal* 77(3): 890-902.

Ma, Y., B. Minasny and C. Wu. 2017. Mapping key soil properties to support agricultural production in Eastern China. *Geoderma Regional* 10: 144–153.

Mayr, T.R., R.C. Palmer and H.J. Cooke. 2008. Digital soil mapping using legacy data in the Eden valley, UK. pp: 291-301. *In*: A.E. Hartemink, A.B. McBratney and M.D.L. Mendonça Santos, (Eds.). Digital soil mapping with limited data. Springer, London.

McBratney, A.B., M.L. Mendoça Santos, and B. Minasny. 2003. On digital soil mapping. *Geoderma* 117(1-2): 3-52.

McBratney, A.B, I.O.A Odeh, T.F.A. Bishop, M.S. Dunbar and T.M. Chatar. 2000. An overview of the pedometric techniques for use in soil survey. *Geoderma* 97: 293-327.

McBratney, A.B., R. Webster, R.G. McLaren and R.B. Spiers. 1982. Regional variation of extractable copper and cobalt in the topsoil of South-east Scotland. *Agronomie* 2(10): 969-982.

McKenzie, N.J., and P.J. Ryan. 1999. Spatial prediction of topsoil properties using environmental correlation, *Geoderma* 89(1-2): 67-94.

McKyes, E. 1989. Agricultural engineering soil mechanics. Elsevier. New York.

Mehnatkesh, A., S. Ayoubi, A. Jalalian, and K.L Sahrawat. 2013. Relationships between soil depth and terrain attributes in a semi-arid hilly region in western Iran. *Journal of Mountain Science* 10(1): 163–172.

Menezes, M.D. de, S.H.G. Silva, C.R. de Mello, P.R. Owens and N. Curi. 2018. Knowledge-based digital soil mapping for predicting soil properties in two representative watersheds. *Scientia Agrícola* 75(2): 144–153.

Melton, M.A. 1965. The geomorphic and paleoclimatic significance of alluvial deposits in Southern Arizona. *The Journal of Geology* 73(1): 1-38.

Merotto A, C.M. Mundstock. 1999. Wheat root growth as affected by soil strength. *Revista Brasileira Ciencia do Solo* 23: 197–202.

Merdun, H., O. Cinar, R. Meral, and M. Apan. 2006. Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. *Soil and Tillage Research* 90(1-2): 108–116.

Miller, B. A., S. Koszinski, M. Wehrhan and M. Sommer. 2015. Impact of multi-scale predictor selection for modeling soil properties *Geoderma* 239: 97–106.

Minasny B and A.B. McBratney. 2016. Digital soil mapping: a brief history and some lessons. *Geoderma* 264: 301-311.

Minasny, B. and A.B. McBratney. 2007. Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes. *Geoderma* 142: 285– 293.

Minasny, B., A.B. McBratney and S. Salvador-Blanes. 2008. Quantitative models for pedogenesis- A review. *Geoderma* 144:140-157.

Minasny, B., A.B. McBratney, M.L. Mendonça-Santos, I.O.A. Odeh and B. Guyon. 2006. Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. *Australian Journal Soil Research* 44: 233 – 244.

Minella, J. P. G. e G.H. Merten. 2012. Índices topográficos aplicados à modelagem agrícola e ambiental. *Ciência Rural*, Santa Maria 9(42): 1575-1582.

Moore, I.D., P.E. Gessler, G.A. Nielsen, G.A. Peterson. 1993. Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal* 57: 443-452.

Moore, I.D., R.B. Grayson and A.R. Ladson. 1991. Digital Terrain Modelling: A review of hydrological, geomorphological, and biological aplications. *Hydrological Processes* 5: 3-30.

Motohka, T., K.N. Nasahara, H. Oguma and S. Tsuchida. 2010. Applicability of Green-Red Vegetation Index for Remote Sensing of vegetation Phenology. *Remote Sensing* 2: 2369-2387.

Moura-Bueno, J.M., R.S.D. Dalmolin, A. ten Caten, L.F.C. Ruiz, P.V. Ramos and A.C. Dotto. 2016. Assessment of Digital Elevation Model for Digital Soil Mapping in a Watershed with Gently Undulating Topography. *Revista Brasileira Ciência* Solo 40: 1-15.

Mulla, D.J., and A.B. McBratney. 2000. Soil Spatial Variability. In: Malcolm E, Sumner J (eds) Handbook of soil science. CRC Press, Boca Raton.

Najafian, A., M. Dayani, H.R. Motaghian and H. Nadian. 2012. Geostatistical Assessment of the Spatial Distribution of Some Chemical Properties in Calcareous Soils. *Journal of Integrative Agriculture* 11(10): 1729–1737.

Nemes, A., W.J. Rawls and A. Pachepsky. 2005. Influence of Organic Matter on the Estimation of Saturated Hydraulic Conductivity. *Soil Science Society of America Journal* 69(4): 1330-1337.

Nelson, M.A., and O. A. Odeh. 2009. Digital soil class mapping using legacy soil profile data: a Comparison of genetic algorithm and Classification tree approach. *Australian Journal of Soil Research* 47(6): 632-649.

Niang, M.A., M. Nolin, G. Jego and I. Perron. 2014. Digital mapping of soil texture using RADARSAT-2 polarimetric SAR data. *Soil Science Society America Journal* 78: 673–684.

Nielsen, D., and O. Wendroth. 2003. Spatial and Temporal Statistics – Sampling Field Soils and their Vegetation, John Wiley and Sons - Verlag, Germany.

Novaes Filho, J.P., E.G. Couto, V.A. Oliveira, M.S. Johnson, J. Lehmann e S.S. Riha. 2007. Variabilidade especial de atributos físicos de solo usada na identificação de classes pedológicas de microbacias na Amazônia meridional. *Revista Brasileira de Ciência do Solo* 31: 91-100.

Nolasco-Carvalho CC, W.Franca-Rocha and J.M. Ucha. 2009. Mapa digital de solos: uma proposta metodológica usando inferência fuzzy. *Revista Brasileira de Engenharia Agrícola e Ambiental* 13: 46-55.

Norouzi, M., S. Ayoubi1, A. Jalalian, H. Khademi, and A.A. Dehghani. 2010. Predicting rainfed wheat quality and quantity by artificial neural network using terrain and soil

characteristics. *Acta Agriculturae Scandinavica Section B-Soil and Plant Science* 60: 341–352.

Odeh, I.O.A., A.J. Todd and J. Triantafilis. 2003. Spatial prediction of soil particle-size fractions as compositional data. *Soil Science* 168: 501–515.

Odeh, I.O.A., and A.B. McBratney. 2000. Using AVHRR images for spatial prediction of clay content in the lower Namoi valley of eastern Australia. *Geoderma* 97: 237–254.

Odeh, I.O.A., A.B. McBratney and D.J. Chittleborough. 1995. Further results on prediction of soil properties from terrain attributes: Heterotopic cokriging and regression-kriging. *Geoderma* 67: 215–226.

Oku, E., A. Essoka, and E. Thomas. 2010. Variability in soil properties along an Udalf toposequence in the humid forest zone of Nigeria. *Kasetsart Journal of Natural Sciences* 44: 564–573.

Osama, M., I.Tomoyasu, F. Kazunari and Y. Kunihiko. 2008. Assessment of spatial variability of penetration resistance and hardpan characteristics in a cassava field. *Australian Journal of Soil Research* 46(3): 210-218.

Outeiro, L., F. Aspero and X. Ubeda. 2008. Geostatistical methods to study spatial variability of soil cations after a prescribed fire and rainfall. *Catena* 74: 310–320.

Pachepsky, Y.A., D.J. Timlin and W.J. Rawls. 2001. Soil water retention as related to topographic variables. *Soil Science Society of America Journal* 65: 1787- 1795.

Papritz, A. 2009. Limitations of indicator krigging for predicting Data with Trend. pp. 1-6. *In*: Conford, D. *et al.* (Ed.), StatGIS Conference Proceedings. Milos, Greece.S

Paz-Gonzalez, A., S.R. Vieira and M.T.T. Castro. 2000. The effect of cultivation on the spatial variability of selected properties of an umbric horizon. *Geoderma* 97(3–4): 273–292.

Pereira, P., and X. Úbeda. 2010. Spatial distribution of heavy metals released from ashes after a wildfire. *Journal of Environmental Engineering and Landscape Management* 18: 13-22.

Piccolo A., and J. Mbagwu. 1999. Role of hydrophobic components of soil organic matter in soil aggregate stability. *Soil Science Society of America Journal* 63: 1801–1810.

Pieri C.J.M.G. 1992. Fertility of Soils: A Future for Farming in the West African Savannah. Springer-Verlag, Berlin, Germany.

Poggio, L., A. Gimona and M.J. Brewer. 2013. Regional scale mapping of soil properties and their uncertainty with a large number of satellite-derived covariates. *Geoderma* 209-210: 1–14.

Priori, S., N. Bianconi, and E.A. Costantini. 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.

Quinn, P.F., K.J. Beven, P. Chevallier and O. Planchon. 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes* 5: 59-79.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [en línea]. Recuperado en: < https://www.R-project.org/ > Consultado el: 21 de agosto de 2018.

Ramzan S., M.A Wani and M.A. Bhat. 2017. Assessment of spatial variability of soil fertility parameters using geostatistical techniques in temperate Himalayas. *International Journal of Geosciences* 8: 1251-1263.

Rasse, D. P., A.J.M. Smucker and D. Santos. 2000. Alfalfa Root and Shoot Mulching Effects on Soil Hydraulic Properties and Aggregation. *Soil Science Society of America Journal* 64(2): 725-731.

Rawlins, B.G., B.P. Marchant, D. Smyth, C. Scheib, R.M. Lark and C. Jordan. 2009. Airborne radiometric survey data and a DTM as covariates for regional scale mapping of soil organic carbon across Northern Ireland. *European Journal Soil Science* 60: 44–54.

Rawls, W.J., Y.A. Pachepsky, J.C. Ritchie, T.M. Sobecki, and H. Bloodworth. 2003. Effect of soil organic carbon on soil water retention. *Geoderma* 116(1-2): 61–76.

Reif, D.M., A.A. Motsinger, B.A. McKinney, J.E. Crowe Jr. and J.H. Moore. 2006. Feature Selection using a Random Forests Classifier for the Integrated Analysis of Multiple Data Types. pp: 1-8. *In: IEEE, Computational Intelligence and Bioinformatics and Computational Biology*. 5-7 October. Chiang Mai. Thailand.

Resende, M., N. Curi, S.B. Rezende e G.F. Corrêa. 2014. Gênese: aspectos gerais. pp. 109-147. *In*: Pedologia: base para distinção de ambientes. 6th Ed. Lavras, MG: Editora UFLA.

Reynolds, W.D., C.F. Drury, C.S. Tan, C.A. Fox and X.M. Yang. 2009. Use of indicators and pore volume- function characteristics to quantify soil physical quality. *Geoderma* 152: 252-263.

Rezaei, S.A., and R.J. Gilkes. 2005. The effects of land bulk density cape attributes and plant community on soil physical properties in rangeland bulk density. *Geoderma* 125: 145-154.

Rhoton, F.E., S.W. Duiker. 2008. Erodibility of a soil drainage sequence in the loess uplands of Mississippi. *Catena* 75(2): 164-171.

Rhoton, F.E., W.E. Emmerich, D.C. Goodrich, S.N. Miller, and D.S McChesney. 2006. Soil geomorphological characteristics of a semiarid watershed. *Soil Science Society of America Journal* 70(5): 1532-1540.

Rizzo R., J.A.M. Demattê, I.F. Lepsch, B.C. Gallo and C.T. Fongaro. 2016. Digital soil mapping at local scale using a multi-depth Vis-NIR spectral library and terrain attributes. *Geoderma* 274: 18-27.

Rivero, R.G., S. Grunwald, T.Z. Osborne, K.R. Reddy and S. Newman. 2007. Characterization of the spatial distribution of soil properties inwater conservation area 2A, Everglades, Florida. *Soil Science* 172(2): 149–166.

Roecker, S., D.Howell, C. Haydu-Houdeshell and C. Blinn. 2010. A qualitative comparison of conventional soil survey and Digital Soil Mapping approaches. pp. 369– 384. *In*: *Digital Soil Mapping: Bridging Research, Environmental Application, and Operation* (eds. J. L. Boettinger, D. W. Howell, A. C. Moore,A. E. Hartemink & S. Kienast-Brown), Springer, Dordrecht, Netherlands.

Ruth, B., and B. Lennartz. 2008. Spatial variability of soil properties and rice yield along two catenas in Southeast China. *Pedosphere* 18: 409–420.

Sauer, D. 2010. Approaches to quantify progressive soil development with time in Mediterranean climate—I. Use of field criteria. *Journal of Plant Nutrition and Soil Science* 173: 822-842.

Saglam, M., and O. Dengiz. 2012. Influence of selected land use types and soil texture interactions on some soil physical characteristics in an alluvial land. *International Journal of Agronomy and Plant Production* 3: 508–513.

Samuel-Rosa, A., R. Simão Diniz Dalmolin, P. Miguel, J. Zalamena and D. Pinheiro Dick. 2013. The effect of intrinsic soil properties on soil quality assessments. *Revista Brasileira de Ciência do Solo* 37: 1236-1244.

Sandoval M., J. Dörner, O. Seguel, J. Cuevas, and D. Rivera. 2012. Métodos de Análisis Físicos de Suelos. Departamento de Suelos y Recursos Naturales Universidad de Concepción. Publicación Nº 5. Chillán, Chile.

Salako, F.K., G. Tian, G. Kirchhof, and G.E. Akinbola. 2006. Soil particles in agricultural landscapes of a derived savanna in southwestern Nigeria and implications for selected soil properties. *Geoderma* 137: 90–99.

Schulte, E. E., C. Kaufmann and B.J. Peter. 1991. The influence of sample size and heating time on soil weight loss-on-ignition. *Communication in Soil Science and Plant Analysis* 22: 159–168.

Schull, P., J. Franklin, O.A. Chadwick and D. McArthur. 2003. Predictive soil mapping-a review. *Progress in Physical Geography* 27(2): 171-197.

Schimel, D., M.A. Stillwell and R.G. Woodmansee. 1985. Biogeochemistry of C, N, and P in a soil catena of the shortgrass steppe. *Ecology* 66: 276–282.

Schoeneberger, P.J., D.A. Wysocki, E.C. Benham, and Soil Survey Staff. 2012. Field book for describing and sampling soils, Version 3.0. Natural Resources Conservation Service, National Soil Survey Center, Lincoln, Nebraska.

Seguel, O., E. Farías, W. Luzio, M. Casanova, I. Pino, A.M. Parada, X. Videla, and A. Nario. 2015. Physical properties of soil after change of use from native forest to vineyard. *Agro Sur Journal* 43(2): 23-39.

Seguel, O., V. García de Cortázar y M. Casanova. 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., and R. Horn. 2006. Structure properties and pore dynamics in aggregate beds due to wetting-drying cycles. *Journal Plant Nutrition of Soil Science* 169: 221-232.

Seeno, E.A. 2015. Predicting spatial variability in soil properties using DSM across Malheur National forest. Master´s thesis. Oregon State University. Department of Crop and Soil Science. Corvallis. Unites State of America.

Serrano, J. M., S. Shahidian and J.R. Marques da Silva. 2014. Soil phosphorus retention in a Mediterranean pasture subjected to differential management. *European Journal of Soil Science* 65(4): 562–572.

Serme, I., K. Ouattara, B. Ouatara and S. Taonda. 2016. Short term impact of tillage and fertility management on Lixisol structural degradation. *International Journal of Agricultural Policy and Research* 4(1): 1-6.

Sharu, M.B., M. Yakubu, S.S. Noma and A.I. Tsafe. 2013. Land evaluation of agricultural landscape in Dingyadi district, Sokoto state, Nigeria. *Nigerian Journal of Basic and Applied Science* 21:148–156.

Silva, S. H. G., A. F. dos S Teixeira, M. D. de Menezes, L. R. G. Guilherme, F. M. de S Moreira and N. Curi. 2017. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF*). Ciência e Agrotecnologia* 41(6): 648– 664.

Silva Cruz, J., R.N. de Assis Júnior, S.S. Rocha Matias and J.H. Camacho-Tamayo. 2011. Spatial variability of an Alfisol cultivated with sugarcane. *Ciencia e Investigação Agraria* 38(1): 155-164.

Silva, M.B., L.H.C. Anjos, M.G. Pereira e R.A.M. Nascimento. 2001. Estudo de toposseqüência da baixada litorânea fluminense: efeitos do material de origem e posição topográfica. *Revista Brasileira de Ciência do Solo* 25: 965-976.

Singh, A. 2016. Managing the water resources problems of irrigated agriculture through geospatial techniques: an overview. *Agriculture Water Management* 174: 2– 10.

Siqueira, D.S., J. Marques jr, G.T. Pereira, R.S. Barbosa, D.B. Teixeira, and R.G. Peluco. 2014. Sampling density and proportion for the characterization of the variability of Oxisol attributes on different materials. *Geoderma* 232-234: 172-182.

Smith, M.P., A.X. Zhu, J.E. Burt and C. Stiles. 2006. The effects of DEM resolution and neighborhood size on digital soil survey. *Geoderma* 137: 58-69.

Slaymaker, O. 2001. The role of remote sensing in geomorphology and terrain analysis in the Canadian Cordillera. *International Journal of Applied Earth Observation and Geoinformation* 3(1): 11–17.

Sobieraj, J.A., H. Elsenbeer, R.M. Coelho and B. Newton. 2002. Spatial variability of hydraulic conductivity along a tropical rainforest catena. *Geoderma* 108: 79-90.

Soto, L., M. Galleguillos, O. Seguel, B. Sotomayor, and A. Lara. 2018. Assessment of soil physical properties' statuses under different land covers within a landscape dominated by exotic industrial tree plantations in south-central Chile. *Journal of Soil and Water Conservation* 74(1): 12–23.

Souza, C.K., J. Marques Júnior, M.V. Martins Filho, e G.T. Pereira. 2003. Influência do relevo na variação anisotrópica dos atributos químicos e granulométricos de um Latossolo em Jaboticabal - SP. *Engenharia Agrícola* 23: 486-495.

Souza, Z.M., J. Marques Júnior, e G.T. Pereira. 2004. Variabilidad espacial de atributos físicos do solo em diferentes formas do relevo sob cultivo de cana-de-açúcar. *Revista Brasileira de Ciência do Solo* 28: 937-944.

SSSA. 2008. Methods of Soil Analysis Part 5— Mineralogical Methods. Book Series 5, Soil Science Society of America, Madison.

Stum, A.K., J.L. Boettinger, M.A. White, and R.D. Ramsey. 2010. Random forests applied as a soil spatial predictive model in arid Utah. *Progress in Soil Science* 2: 179-190.

Stutter, M.I., D.G. Lumsdom, M.F. Billett, D. Low and L.K. Deeks. 2009. Spatial Variability in Properties Affecting Organic Horizon Carbon Storage in Upland Soil. *Soil Science Society of America Journal* 73: 1724-1732.

Svetnik, V., A. Liaw, C. Tong, J. Culberson, R.P. Sheridan and B.P. Feuston. 2003. Random forest: A classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computation Science* 43(6): 1947–1958.

Swanson, F.J., T.K. Kratz, N. Caine and R.G. Woodmansee. 1988. Landform effects on eco- system patterns and processes. *BioScience* 38: 92- 98.

Szabó, B., G. Szatmári, K. Takács, A. Laborczi, A. Makó, K. Rajkai and L. Pásztor. 2019. Mapping soil hydraulic properties using random-forest-based pedotransfer functions and geostatistics. *Hydrology and Earth System Sciences* 23(6): 2615–2635.

Tarini, M., P. Cignoni and C. Montani. 2006. Ambient Occlusion and Edge Cueing to Enhance Real Time Molecular Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12(5): 1237-1244.

Taghizadeh-Mehrjardi R, k. Nabiollahi, B. Minasny and J. Triantafilis. 2015. Comparing data mining classifiers to predict spatial distribution of USDA-family soil groups in Baneh region, Iran. *Geoderma* 253-254: 67-77.

Taghizadeh-Mehrjardi, R., B.F. Minasny, F. Sarmadian, and B.P. Malone. 2014. Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma* 213: 15–28.

Taylor, J., A.F. Jacob, M. Galleguillos, L. Prévot, N. Guix and P. Lagacherie. 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.

Taylor, H.M., and G.S. Brar. 1991. Effect of soil compaction on root development. *Soil and Tillage Research* 19(2-3): 111-119.

ten Caten, A., R.S.D. Dalmolin, F.A. Pedron, L.F.C. Ruiz and C.A. Silva. 2013. An appropriate data set size for digital soil mapping in Erechim, Rio Grande do Sul, Brazil. *Revista Brasieira Ciencia Solo* 37: 359-366.

ten Caten, A., R.S.D. Dalmolin, M.L. Mendonça-Santos e E. Giasson. 2012. Mapeamento digital de classes de solos: características da abordagem brasileira. *Ciencia Rural* 42: 1989-1997.

ten Caten, A., R. Dalmolin, F. Pedron, and M. Mendonça Santos. 2011. Extrapolação das relações solo-paisagem a partir de uma área de referência. *Ciência Rural* 41(5): 812-816.

Teske R., E. Giasson and T. Bagatini. 2014. Comparação do uso de modelos digitais de elevação em mapeamento digital de solos em Dois Irmãos, RS, Brasil. *Revista Brasileira Ciência Solo* 38: 1367-1376.

Thompson, J.A, A.C. Moore, R.F. Austin and E.M. Yewtukhiw. 2006. Multiscale Terrain Analysis to Improve Landscape Characterization and Soil Mapping. 18th World Congress of Soil Science Philadelphia, Pennsylvania, USA.

Tillman, R.W., D.R. Scotter, M.G. Wallis, and B.E. Clothier. 1989. Water-repellency and its measurement by using intrinsic sorptivity. *Australian Journal of Soil Research* 27: 637- 644.

Tian, Y.C., K.J. Gu, X. Chu, X. Yao, W.X. Cao and Y. Zhu. 2014. Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. *Journal of Plant and Soil* 376: 193–109

Toghizadih-Mehrjardi, R., F. Sarmadian, B. Minasny, J. Triantafilis and M. Omid. 2014. Digital Mapping of Soil Classes Using Decision Tree and Auxiliary Data in the Ardakan Region, Iran. *Arid Land Research and Management* 28(2): 147–168.

Tola, E., K.A. Al-Gaadi, R. Madugundu, A.M. Zeyada, A.G. Kayad and C.M. Biradar. 2017. Characterization of spatial variability of soil physicochemical properties and its impact on Rhodes grass productivity. *Saudi Journal of Biological Sciences* 24(2): 421–429.

Tsegaye, T., and R. L. Hill. 1998. Intensive tillage effects on spatial variability of soil physical properties. *Soil Science* 163(2): 143–154.

Ubalde, J.M., X. Sort, R.M. Poch and M. Porta. 2007. Influence of edapho-climatic factors on grape quality in Conca de Barbera vineyards (Catalonia, Spain). *Journal International des Sciences de la Vigne et du Vin* 41: 33–41.

Umali, B.P., D.P. Oliver, S. Forrester, D.J. Chittleborough, J.L. Hutson, R.S. Kookana, and B. Ostendorf. 2012. The effect of terrain and management on the spatial variability of soil properties in an apple orchard. *Catena* 93: 38-48

Unamunzaga, O., G. Besga, A. Castellón, M.A. Usón, P. Chéry, P. Gallejones and A. Aizpurua. 2014. Spatial and vertical analysis of soil properties in a Mediterranean vineyard soil. *Soil Use and Management* 30: 285-296.

Urbanek, E., P. Hallett, D. Feeney and R. Horn. 2007. Water repellency and distribution of hydrophilic and hydrophobic compounds in soil aggregates from different tillage systems. *Geoderma* 140:147-155.

Uribe, J.M., R. Cabrera, A. de la Fuente and M. Paneque. 2012. Atlas Bioclimático de Chile. Universidad de Chile, Santiago de Chile.

Van derWerff, H.M.A and F.D. Van der Meer. 2016. Sentinel-2A MSI and Landsat 8 OLI Provide Data Continuity for Geological Remote Sensing. *Remote Sensing* 8(11): 1-16.

Van Leeuwen, C., P. Friant, X. Chone, O. Tregoat, S. Koundouras and D. Dubourdieu. 2004. Influence of climate, soil and cultivar on terroir. *America Journal of Enology and Viticulture* 55: 207– 217.

Vasu, D., S.K. Singha, N. Sahua, P. Tiwarya, P. Chandrana, V.P. Duraisamib, V. Ramamurthyc, M. Lalithac and B. Kalaiselvic. 2017. Assessment of Spatial

Variability of Soil Properties Using Geospatial Techniques for Farm Level Nutrient Management. *Soil and Tillage and Research* 169: 25-34.

Verrelst, J., J.P. Rivera, F. Veroustraete, J. Muñoz-Marí, J.G. Clevers, G. Camps-Valls, *et al*., 2015. Experimental sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – a comparison. *ISPRS Journal of Photogramm and Remote Sensing* 108: 260– 272.

Vieira, S. R. 2000. Geoestadística em estudos de variabilidade espacial do solo. *Tópicos em Ciência do Solo* 1: 1-54.

Vrsic, S., A. Ivancic, B. Pulko and J. Vadhuver. 2011. Effect of soil management systems on erosion and nutrition loss in vineyard on steep slopes. *Journal of Environmental Biology* 32: 289–294.

Warrick, A.W. 2002. Soil physics companion. CRC Press, Boca Raton, USA.

Wang, W., X. Yao, X. Yao, *et al*., 2012. Estimating leaf nitrogen concentration with three–band vegetation indices in rice and wheat. Field Crops Research, 129: 90–98.

Webster, R. 2000. Is soil variation random? *Geoderma* 97: 149- 163.

Webster, R., and M. A. Oliver. 2001. Geostatistics for Environmental Scientists. Hoboken, New Jersey. John Wiley and Sons Inc.

Western, A.W., R.B. Grayson, G. Blöschl and D.J. Wilson. 2003. Spatial variability of soil moisture and its implication for scaling. pp: 119-142. *In*: Pachepsky, Y., Radcliffe, D., E., Selim, H. M. (Eds). Scaling methods in soil physics. Boca Raton, FL: CRC Press.

Xing-Yi., Y. Y. Sui, X.D. Zhang, K. Meng and S.J. Herbert. 2007. Spatial Variability of Nutrient Properties in Black Soil of Northeast China. *Pedosphere* 17(1): 19–29.

Yoo, K., R. Amundson, A.M. Heimsath and W.E. Dietrich. 2006. Spatial patterns of soil organic carbon on hillslopes: Integrating geomorphic processes and the biological cycle. *Geoderma* 130: 47-65.

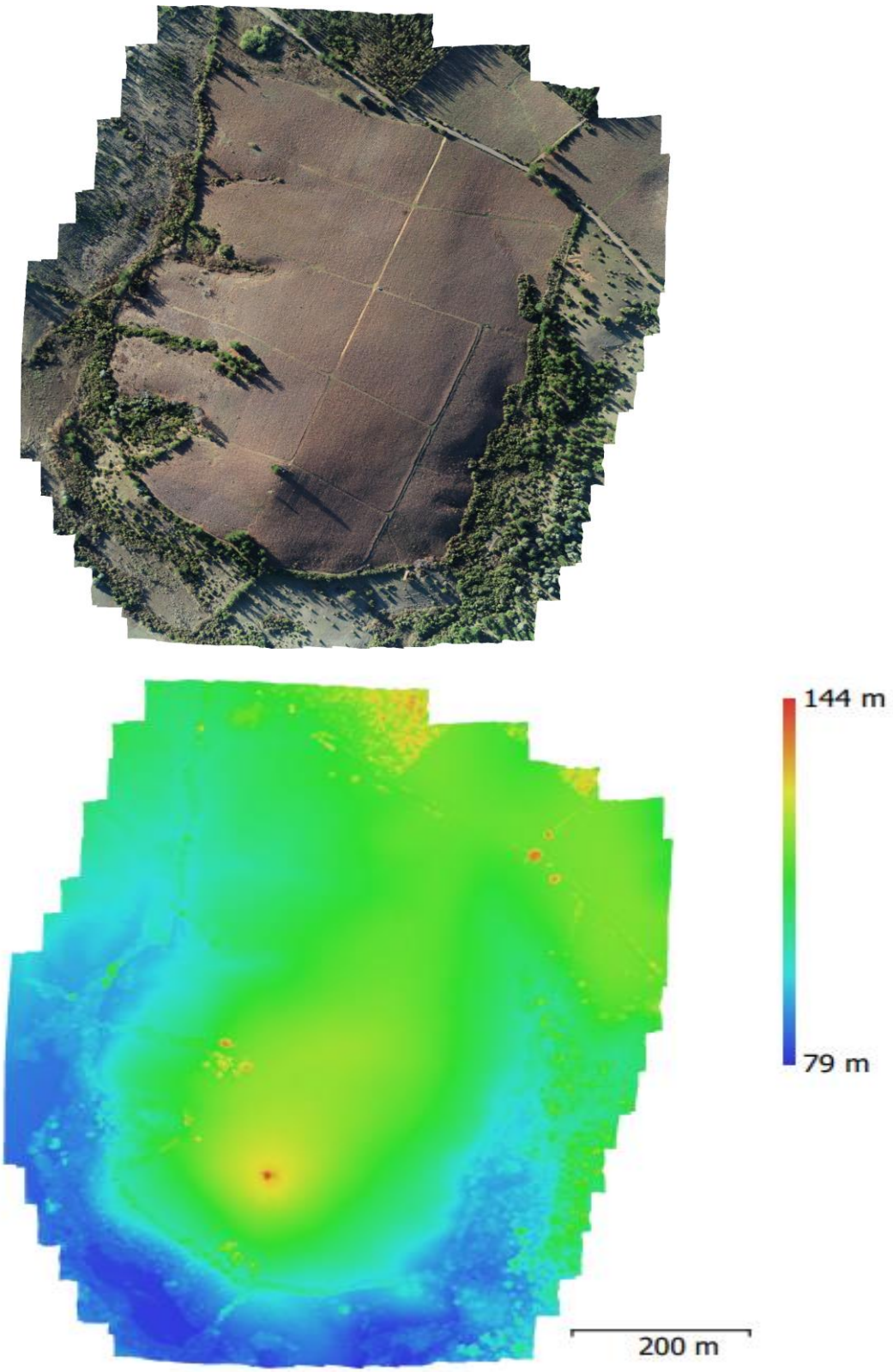
Yoon, Y., J. Kim and S. Hyun. 2007. Estimating soil water retention in a selected range of soil pores using tension disc infiltrometer data. *Soil and Tillage Research* 97: 107–116.

Zevenbergen, L.W., and C.R. Thorne. 1987. Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* 12(1): 47-56.

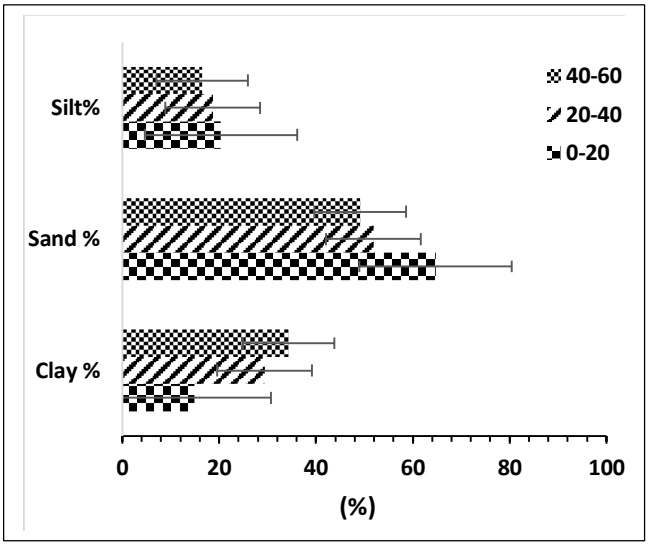
Zhao, K.G., S. Popescu, and R. Nelson. 2009. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment* 113: 182–196.

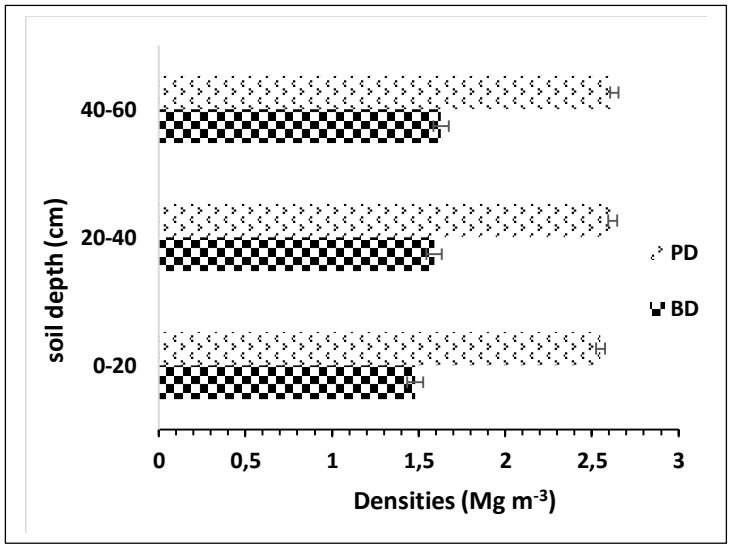
Zhang, S.W., Y.F. Huang, C.Y. Shen, H.H. Ye, and Y.C. Du. 2012. Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. *Geoderma* 171-172: 35–43.

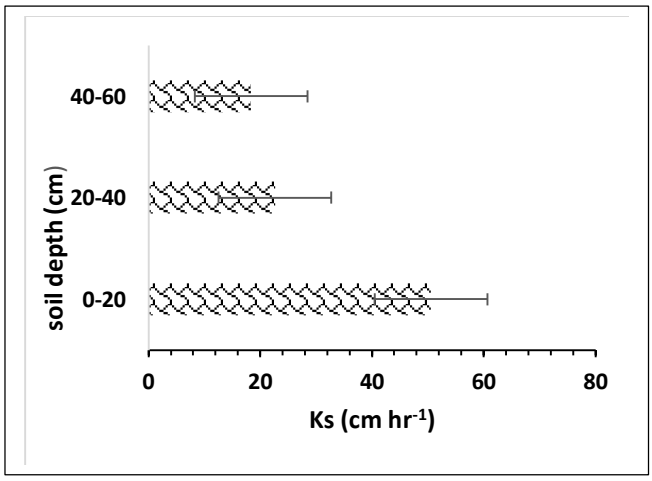
Appendix 1: Digital elevation model of the studied area

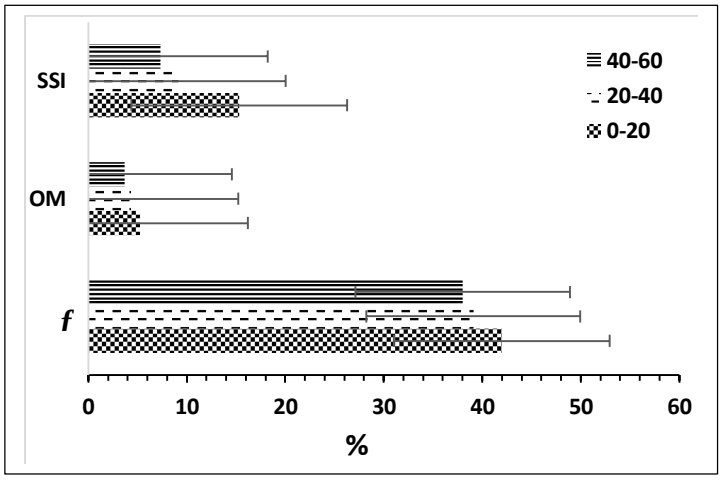


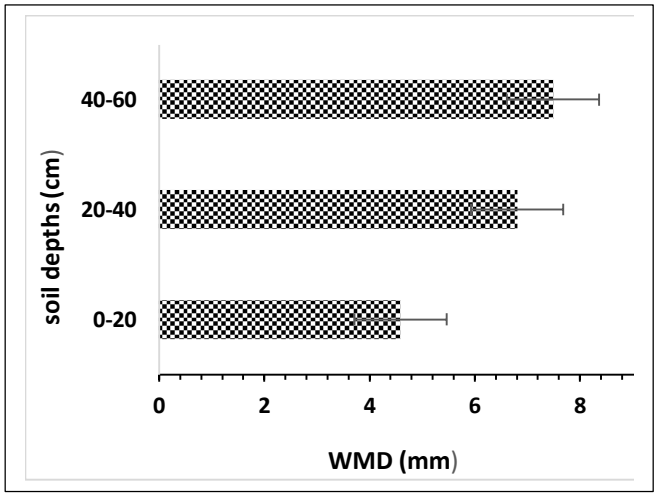
Appendix 2: Results

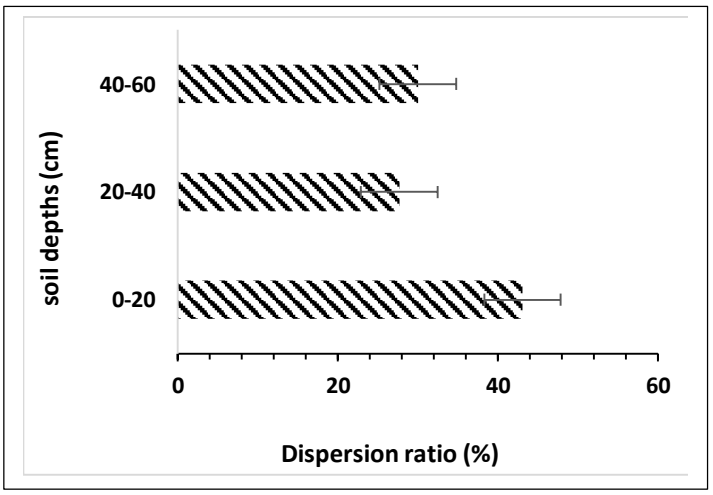


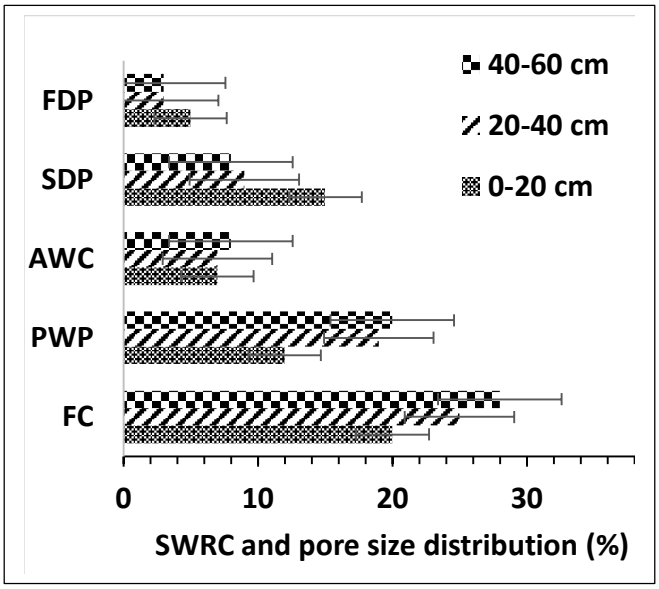


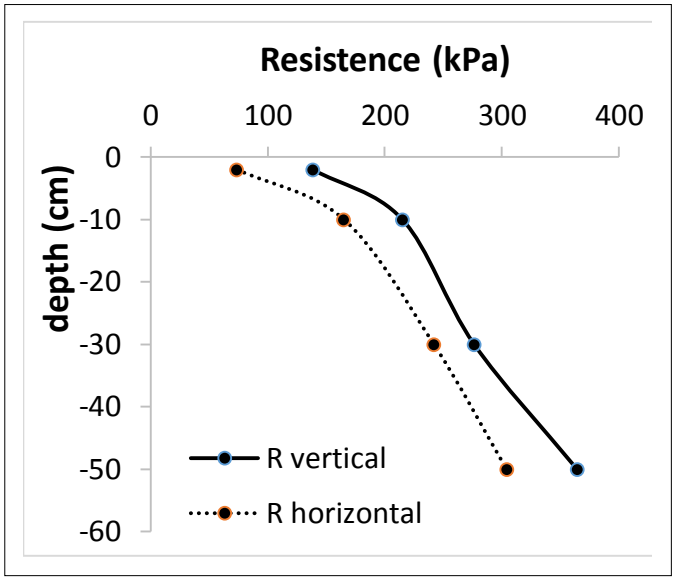









40-60 cm
20-40 cm
0-20 cm
FDP
SDP
AWC
PWP
FC
0
10
20
30
SWRC and pore size distribution (%)


Resistence (kPa)
0
100
200
300
400
depth (cm)
0
-10
-20
-30
-40
-50
-60
R vertical
R horizontal