

Output1 Introduction

Evaluation techniques and geographical maps of field nutrient fertility characteristics are developed

Output 1 aimed to develop a simple analytical method to evaluate nutrient characteristics and distribution of rice and responses to fertilizer inputs. For this, we set three targets:

1. Establish soil and spectral analytical methods applicable to the rice fields in the target area.
2. Develop inter-field scale maps for the fertilizer application.
3. Transfer the developed evaluation methods to at least 20% of extension services in the target regions of Madagascar.

Output 1 produced several key achievements starting from soil characterization in the target regions, development of nutrient evaluation techniques enabling effective fertilizer management for rice production, and technical dissemination through manuals and workshops. Chemical analysis of soils revealed the low availability and high spatial variation in soil nutrient status, particularly phosphorus (P), at the field-scale level. We developed new techniques to rapidly and accurately assess such heterogeneous soil nutrient status and the responses to fertilizer management.

Firstly, we developed models using the indoor spectral information of soils and partial least squares (PLS) regression approach to estimate carbon (C) and nutrient (N and P) contents in soils collected from the rice fields. The soil P estimation model was further updated to apply to various ecosystems in Madagascar using a deep learning of one-dimensional convolutional neural networks (1D-CNN) approach. Additionally, color and magnetic susceptibility sensors were applied alone or in combination to estimate soil C and nutrient contents in the rice fields. These sensors are portable and inexpensive and thus can provide quick and easy multi-location measurements of soil properties in the field. Through repeated pot and multi-field trials, we found that P fertilizer application can efficiently increase P uptake and yield of rice in soils with lower soil P retention capacity. Moreover, we revealed that soil P retention capacity can be simply and accurately estimated from air-dried soil moisture content without chemical analysis. Three of the above achievements: “Quick assessment of bioavailable P in soils: Method and algorithm with FieldSpec”, “Estimation of rice response to P fertilizer application using soil P retention capacity”, and “Estimating soil organic matter content by visual assessment of soil color”, were compiled as a technical manual and made available for everyone on the LRI website. At the end of the project, a workshop was conducted to hand over these techniques to the beneficiaries,

namely agricultural extension officers, research technicians of agricultural research institutions, and farmers.

The information obtained in Output 1 should contribute to the quantitative assessment of nutrient-poor fields for developing efficient fertilizer management based on the soil status in the target regions and elsewhere in SSA. The equipment, technology, and knowledge introduced in the project have already been used by researchers and students in Madagascar. The publications and manuals developed in this output have been compiled.

A memory of Output1 activities



Quick assessment of bioavailable P in soils: Method and algorithm with FieldSpec

Keywords | Deep learning, Soil fertility, Spectroscopy

Summary

Quick soil P evaluation methods and the algorithms with FieldSpec provide a rapid and accurate soil P prediction. This supports farmers to perform efficient fertilizer design to improve rice yield in the paddy fields with phosphorus-deficient soil.

Background & Objective

Phosphorus (P) deficiency is a major constraint to rice production in highly weathered soils of tropical agroecosystems. Quick soil P evaluation method with FieldSpec assess the P deficiency level for making appropriate fertilizer management.

Contents & Characteristics

- The P availability (oxalate-extractable P) can be predicted from laboratory spectroscopy (FieldSpec) with partial least squares (PLS) regression model (Fig. 1a) using selected wavebands (Fig. 2)
- The algorithm using one-dimensional convolutional neural network (1D-CNN) provides the improved prediction across different land use systems (Fig. 1b)

Application

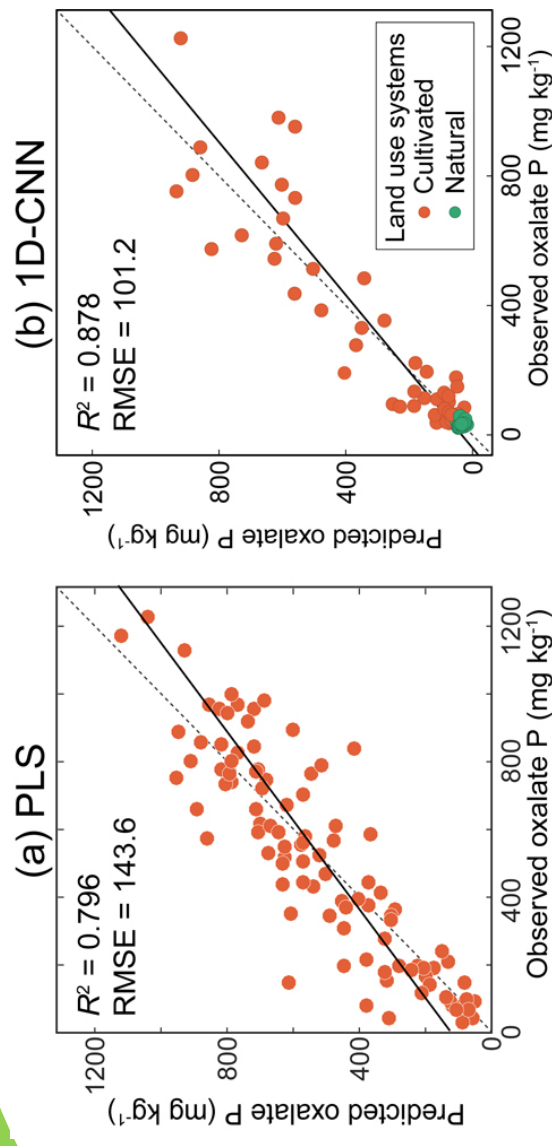
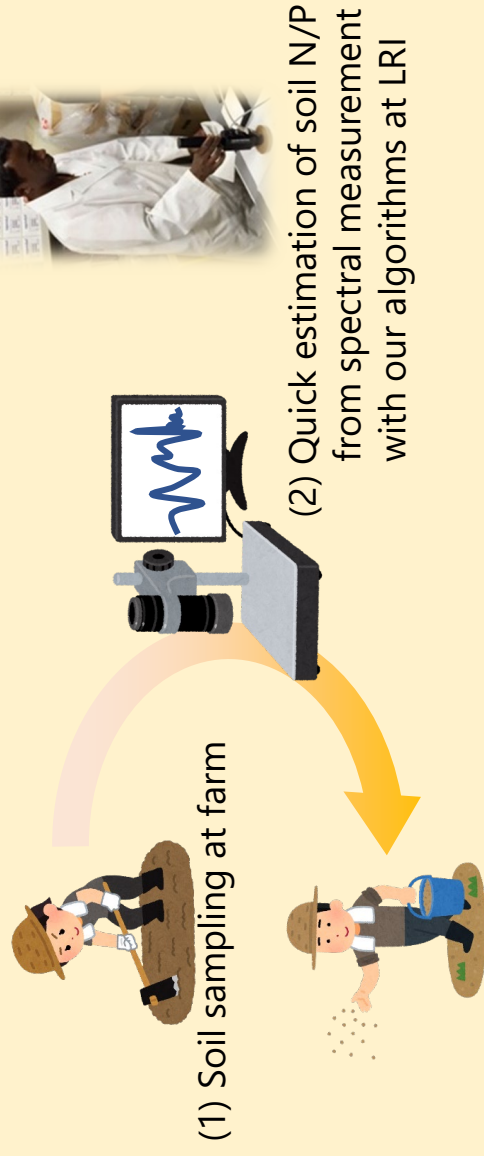


Fig. 1: Validation plots between observed and predicted values of soil P using (a) PLS and (b) 1D-CNN models.

RMSE: root mean squared errors

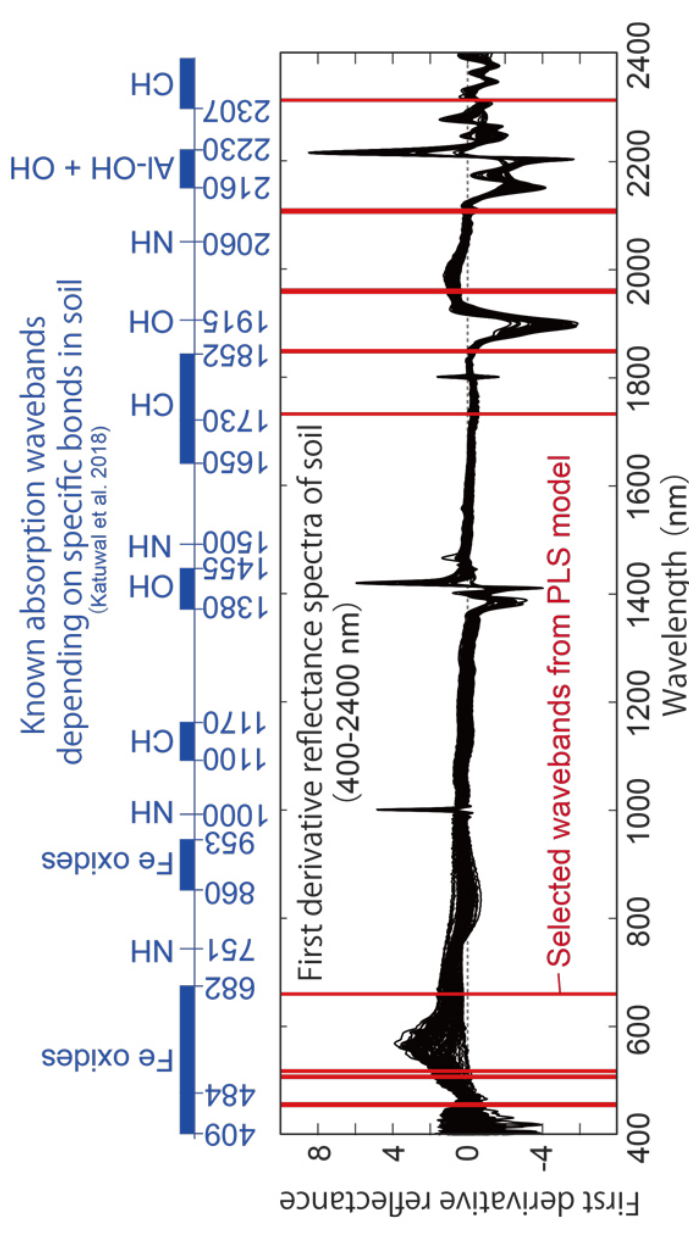


Fig. 2: Selected wavebands (red bars) in the PLS model as important wavelength for soil P assessment.

Authors | Kensuke Kawamura¹, Tomohiro Nishigaki¹, Andry Andriamananjara² and Hobimiarantsoa Rakotonindrina²

Affiliation | 1: Japan International Research Center for Agricultural Sciences (IRCAS),
2: Laboratoire des Radio-Isotopes, Université d'Antananarivo

References | Remote Sens (2019) 10.3390/rs11050506

Agriculture (2020) 10.3390/agriculture10050177

Remote Sens (2021) 10.3390/rs13081519

Estimation of rice response to P fertilizer application using soil P retention capacity

Keywords | Soil diagnosis, Fertilizer management

Summary

Rice response to P fertilizer application can be estimated by soil P retention capacity. This enables farmers to apply fertilizer efficiently as following the estimated rice response to P fertilizer based on soil P retention capacity.

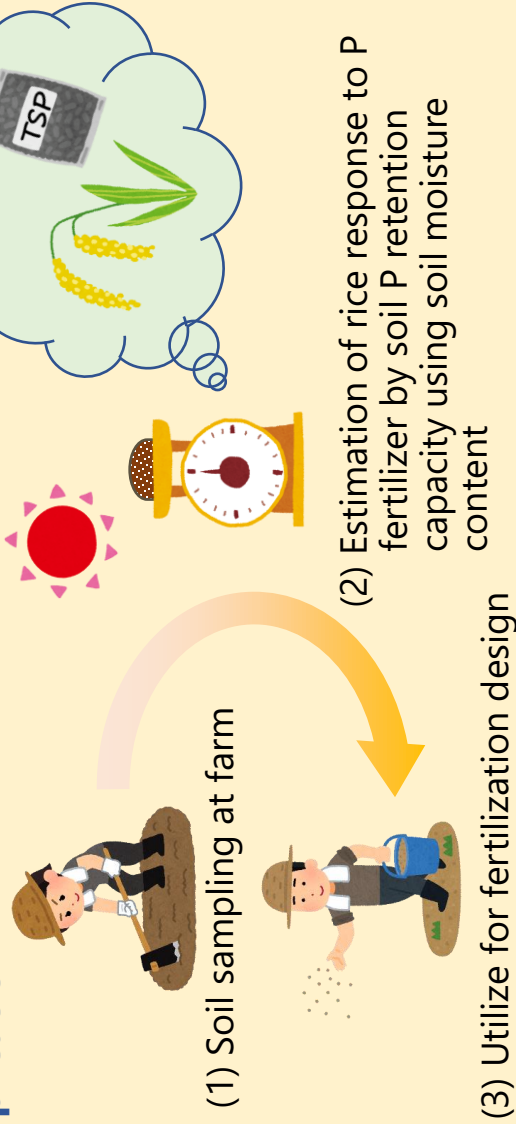
Background & Objective

Farmers can efficiently apply fertilizer if they can know the response of rice to fertilizer application beforehand. Here, the relationship between rice growth and soil characteristic is presented. Further, a simple soil test to estimate P retention is proposed.

Contents & Characteristics

- The increase of P uptake and grain yield of rice in response to P-fertilizer application negatively correlated with soil P retention (Figure 1).
- Soil P retention can be easily estimated by the moisture content of air-dried soils (Figure 2).

Application



Supporting Data

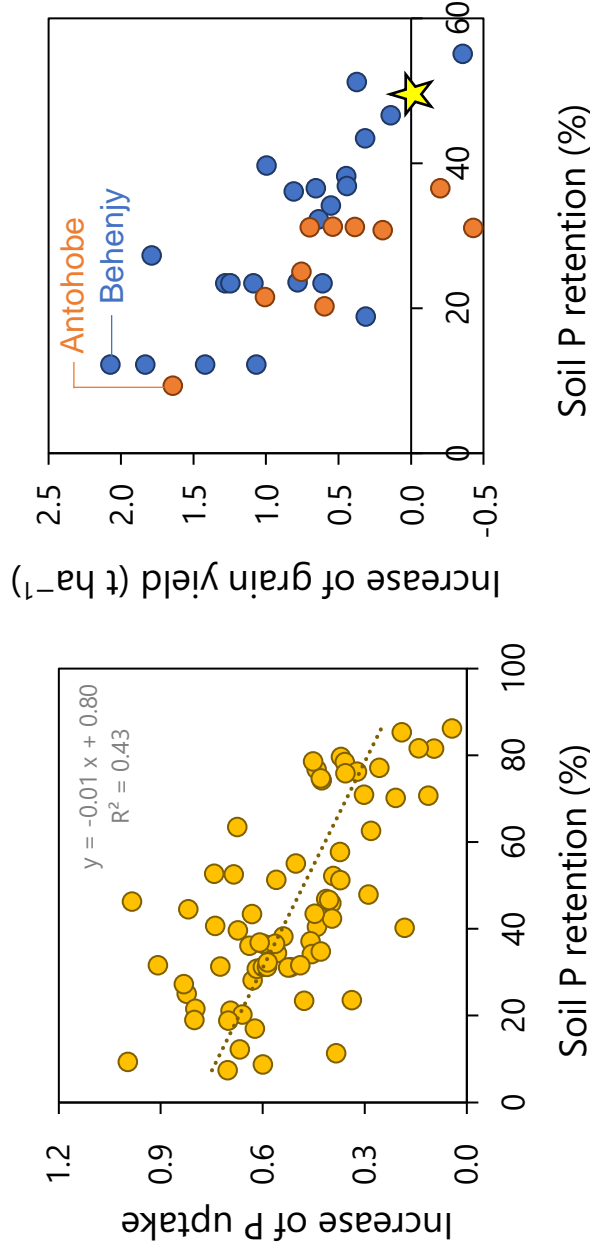


Figure 1. Relationship between soil P retention and the increase of P uptake (left) and grain yield (right) of rice plants in response to P fertilizer application.

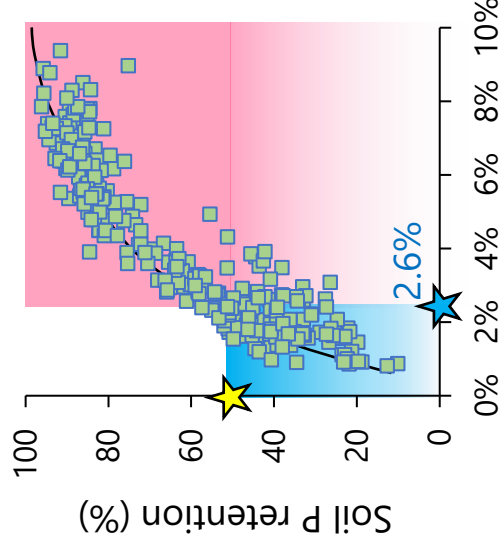


Figure 2. Relationship between moisture content of air-dried soil and soil P retention capacity.

Little or no effect of P fertilizer application on rice yield is expected due to high P retention capacity of soils.

Fertilizer application is still effective to improve rice production due to low P retention capacity of soils.

If your field soil has 2.6% or less of air-dried soil moisture content, P fertilizer application is recommended for improving rice production.

Authors | Nishigaki, T.¹, Tsujimoto, Y.¹, Andriamananjara, A.², Rakotoson, T.²

Affiliation | ¹Japan International Research Center for Agricultural Sciences, Japan;

²Laboratoire des Radio-Isotopes, Université d'Antananarivo, Madagascar

References | Nishigaki et al. (2021) *Geoderma*

DOI: 10.1016/j.geoderma.2021.115326

Estimating soil organic matter content by visual assessment of soil color

Keywords | Munsell soil color chart, Soil organic matter

Summary

Rice soils in Madagascar are diverse in color, reflecting organic matter and iron mineralogy. Visual observation of soil color can be useful for roughly estimating soil organic matter content.

Background & Objective

Soil organic matter plays a key role for enhancing soil productivity. Soil color originates mainly from organic matter, iron mineralogy and moisture content. Here, a visual assessment method for estimating soil organic matter content from soil color is presented.

Contents & Characteristics

- Lowland and upland rice soils in Madagascar have a wide range of color (Figure 1), reflecting organic matter and iron mineralogy.
- Soil total carbon content can be roughly estimated from the product of Munsell value and Munsell chroma (Figure 2).



Figure 1. Air-dried, 2-mm sieved soils from rice fields in Madagascar. Yellow flamed samples show upland rice soils.

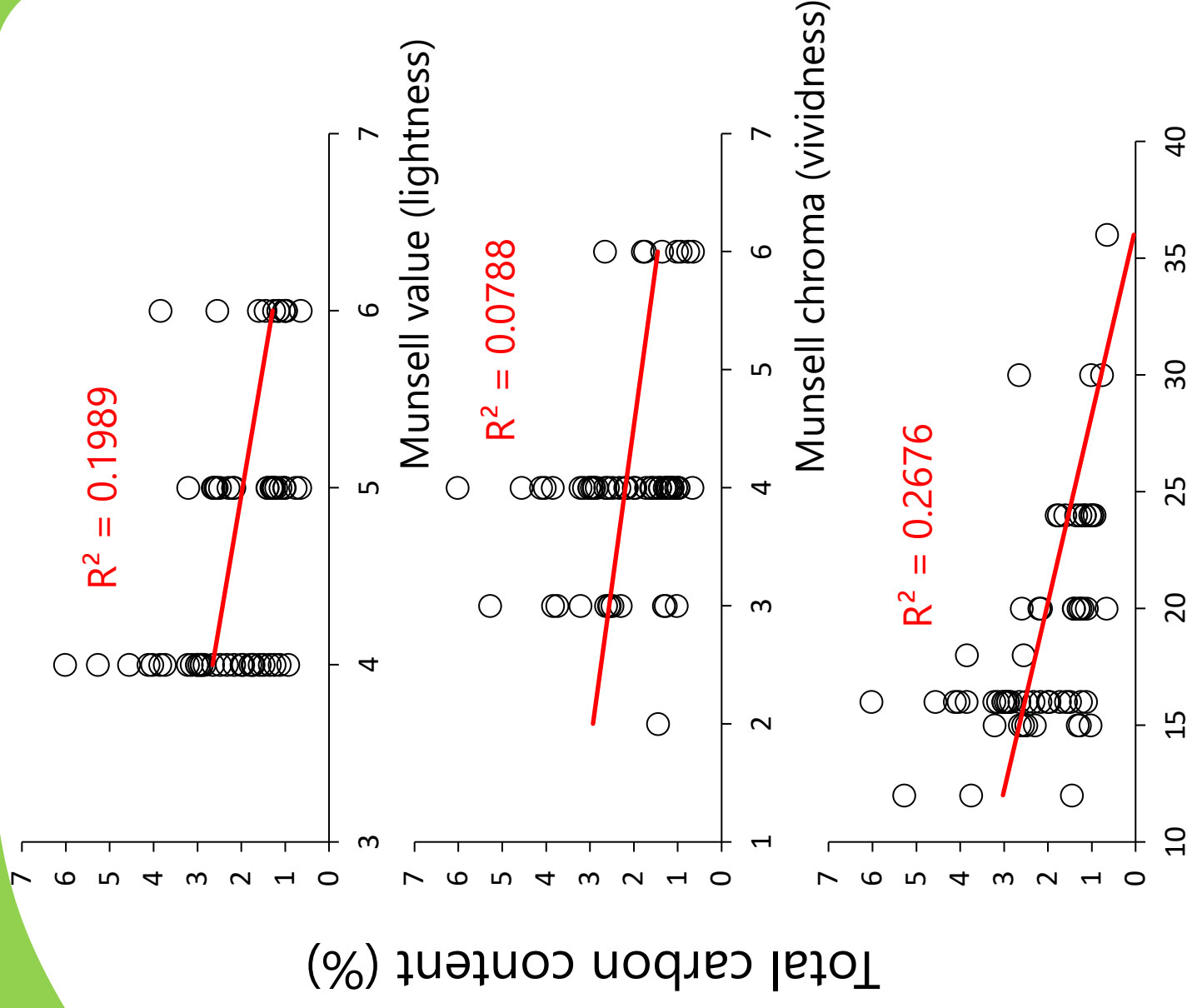


Figure 2. Visual assessment of soil total carbon content using a Munsell soil color chart (tested samples are depicted in Figure 1).

Authors | Naoki Moritsuka¹, Kensuke Kawamura², Yasuhiro Tsujimoto², Michel Rabenarivo³, Andry Andriamananjara³, Tovohery Rakotoson³

Affiliation | ¹ Kochi University, Japan;

² Japan International Research Center for Agricultural Sciences, Japan;

³ Laboratory of Radioisotopes, University of Antananarivo, Madagascar

References | Moritsuka et al. (2019) *Soil Science and Plant Nutrition*

DOI: 10.1080/00380768.2019.1676624

ORIGINAL ARTICLE

Soil survey of the east coast and the central highlands indicates need to update Madagascar soil map

Tomohiro Nishigaki ^a, Kenta Ikazaki^a, Yasuhiro Tsujimoto ^a, Andry Andriamananjara ^b, Tovohery Rakotoson^b and Tantely Razafimbelo^b

^aCrop, Livestock and Environment Division, Japan International Research Center for Agricultural Sciences, Tsukuba, Japan; ^bLaboratoire des Radioisotopes, Université d'Antananarivo, Antananarivo, Madagascar

ABSTRACT

The soil map of Madagascar has not been substantially updated since the 1960s. The number of reported soil profiles that meet the criteria of the Reference Pedon Description – the highest level of data reliability used to ascertain FAO soil unit classification – is also limited for a country-scale soil map. Therefore, we conducted a soil survey in the east coast and the central highlands, the most populated and major food-producing areas in Madagascar, to assess need for reconsideration of the current soil map. As compared to the latest soil map, our results indicate that Ferralsols cover less area on the flat terrains from the east coast to the central highlands and that Geric Ferralsol is more frequently observed soil type rather than Haplic Ferralsol on acidic rocks. The soil type of paddy fields in the southwestern plain of Lake Alaotra – the country's largest rice-producing region – was Vertisol with a hydric horizon rather than Histosol. This is attributed to long-term paddy cultivation in the region or the difference of the definition of peat soils between the current and old soil classification systems. The volcanic-origin soils that sporadically exist in the central highlands can be classified as Eutric Andosol rather than Eutric Cambisol. These results show the need for updating the current soil map based on the quantitative data and also have important implications regarding the land history and for appropriate land management in agriculturally intensive areas in the east coast and central highlands of Madagascar.

ARTICLE HISTORY

Received 16 August 2019

Accepted 11 May 2020

KEY WORDS

Effective base saturation; Reference Pedon Description; soil classification; topography

1. Introduction

Soil maps based on quality-assessed soil data and interpreted soil information provide fundamental information for appropriate agricultural management. The latest soil map of Madagascar – the Soil Atlas of Africa (EU 2013) – was developed based on the Soil Map of the World (FAO-UNESCO 1977) with little modification (Figure 1). In addition, the Madagascar part of the Soil Map of the World was created from information in an atlas (Roederer and Bourgeat 1969) and a pedological map (Riquier 1968) of Madagascar. Therefore, the latest soil map of Madagascar basically relies on an old French soil classification system used prior to the Commission de Pédologie et de Cartographie des Sols (CPCS) (1967), and there has been no significant update since the 1960s. In addition, Batjes (2008) reported that only 20 of 54 soil profiles in Madagascar in Version 3.1 of the ISRIC-WISE database meet the Reference Pedon Description criteria, which is the highest data reliability level used to ascertain FAO soil unit classification. This amount is clearly insufficient to update a country-scale soil map of Madagascar, particularly with its diverse environment. Thus, collecting new soil data and classifying the soils according to the latest soil classification system based on those data would be an important first step in updating the soil map in this country.

The east coast and the central highlands of Madagascar are characterized as the most populated and major food-

producing area of the country. According to the Köppen climate classification system, the climate greatly varies from tropical with high temperatures and precipitation (Af) to temperate with a dry winter and hot-warm summer (Cwa, Cwb) (Figure 2). Regardless of the various climatic conditions, Haplic Ferralsols (FR-ha) are reported to be predominant on acidic rocks widely distributed from the east coast to the central highlands (EU 2013). Haplic Ferralsols are characterized as soils rich in kaolinite and oxides and are generally considered unfertile. Because the classification of soils as FR-ha is mainly based on information in the old atlas (Roederer and Bourgeat 1969) and pedological map (Riquier 1968), there is a considerable doubt whether these classifications are currently applicable.

Lake Alaotra and its surrounding wetlands are the most intensive rice-producing area in Madagascar. The fertile sedimentary soils in the vast southwestern plain of Lake Alaotra have been utilized as paddy fields with irrigation canals and a constant water supply from Lake Alaotra. According to the latest soil map (EU 2013), the lake's surrounding wetlands are covered by Dystric Histosols (HS-dy), soils with thick organic layers and a low pH. Although this soil type may be correct for the marsh and reed bed, it is doubtful that HS-dy is the correct soil type for the paddy fields that have been cultivated for many years considering the inadequate characteristics of HS-dy for agricultural practices.

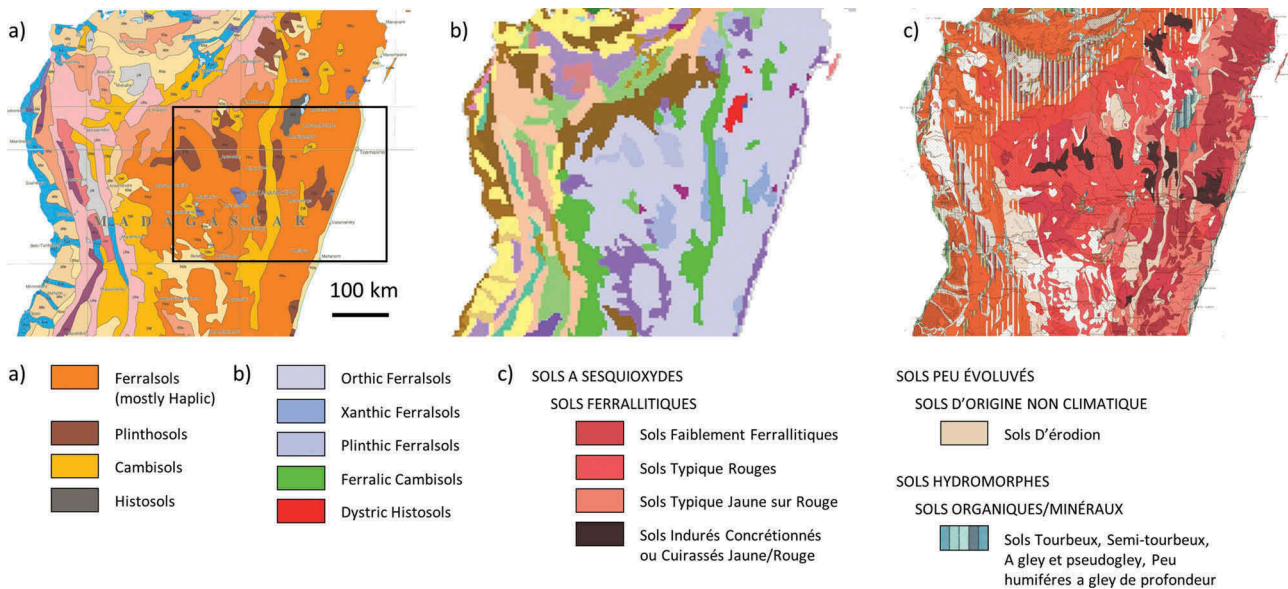


Figure 1. Changes in the soil map of Madagascar: (a) EU (2013), (b) FAO-UNESCO (1977) digitalized in FAO (1991), and (c) Riquier (1968). Only soils mentioned in the main text are shown. A rectangle in (a) represents the area shown in Figure 3

The high mountains that sporadically appear in the central highlands of Madagascar are mostly of volcanic origin, whereas the rest of the highlands are essentially formed by schists, migmatites, and gneiss of the basement complex. The Quaternary volcanic fields overlie Precambrian basement rocks in the central highlands of Madagascar (Rufer et al. 2014), and the soils of volcanic origin can have substantially different properties and higher crop productivity as compared with the strongly weathered unfertile soils surrounding them (Nishigaki et al. 2018). Therefore, the areas with productive soils should be clearly characterized. Currently, however, volcanic soils in the central highlands are classified as Eutric Cambisols (CA-eu) (EU 2013), which offers little information for appropriate land management, particularly for volcanic soils with a high phosphate fixation.

Given the aforementioned obsolete and incomplete nature of the soil database in Madagascar, we established the following objectives in this study: (1) to increase the number of soil profiles meeting the criteria of Reference Pedon Description in Madagascar, (2) to validate the distribution of FR-ha from the east coast to the central highlands, and (3) to obtain representative soil profile information in areas with intensive agriculture in the central highlands to assess need for reconsideration of the current soil types (HS-dy in Lake Alaotra and CA-eu in volcanic soils) in terms of agricultural management.

2. Materials and methods

2.1. General description of the east coast and central highlands

The east coast and central highlands of Madagascar consist of three flat terrains and two slopes (Figure 2). The upper flat terrain (FT1) is a constituent of the central highlands and has an altitude between 1200 and 1500 m a.s.l. (data from U.S. Geological Survey EROS Data Center (1997)), a mean annual temperature of

13–19°C, and a mean annual rainfall of 1200–1800 mm (data from Fick and Hijmans (2017)). On the east, FT1 ends in an escarpment (S1) connecting FT1 and the middle flat terrain (FT2). The FT2 is also a constituent of the central highlands and has an altitude between 600 and 900 m a.s.l., a mean annual temperature of 19–21°C, and a mean annual rainfall of 1500–2400 mm. Slope 2 (S2) is relatively gentle and connects FT2 and the lower flat terrain (FT3). The FT3 is a coastal plain and a constituent of the east coast. It is rarely more than 30 km wide and has an altitude of less than 100 m a.s.l., a mean annual temperature of 23–25°C, and a mean annual rainfall of 2400–3000 mm. According to the current soil map (EU 2013) and geological map (Schlüter 2008), a predominant soil type on acidic rocks on FT1–3 is FR-ha (Figure 1), and the dominant soils on basic rocks (i.e., the greenstone belt) in S2 are Plinthosols, while Cambisols are also found on S1 and S2 between flat terrains.

To estimate the slope distribution within a rectangle of the surveyed area in Figure 1 (top, 17°45'S; bottom, 20°05'S; left, 46°48'E; right, 49°57'E; totally 75,610 km²), ALOS World 3D-30 m images (JAXA 2018) were used. The slope for each pixel having a resolution of 32 m was calculated by slope function of a GIS software (ArcGIS version 10.4, ESRI).

2.2. Soil survey and physicochemical analysis

Ten soil pits were made to obtain a detailed description of the soil profile and to collect soil samples (Figures 2, 3 and A, excluding ID 4). Each pit was 1 m × 1 m in area and generally up to about 1.2 m deep. The locations were carefully determined from the viewpoint of regional representativeness using geological maps (Basairie 1968; Riquier 1968; Schlüter 2008), soil maps (Riquier 1968; EU 2013), and satellite images (e.g., Google Earth) (Table 1). The specific location was selected based on field observation such that each pit was on a summit or ridge to the extent possible to avoid

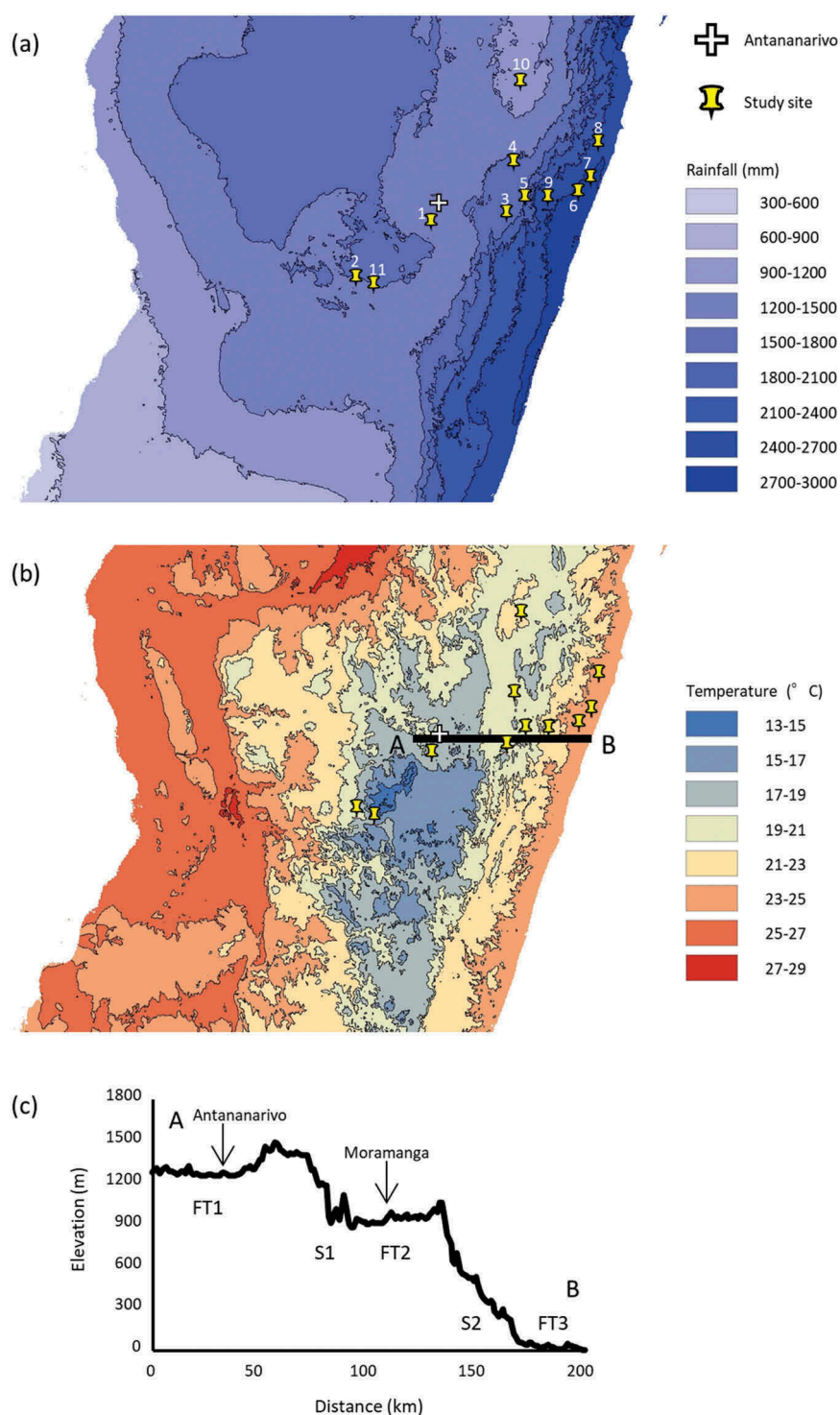


Figure 2. Mean annual rainfall (a), mean annual temperature (b) and topographic cross section along the line AB (c). Maps were created using the WorldClim 2 dataset (Fick and Hijmans 2017). Topographic data were derived from GTOPO30 (U.S. Geological Survey EROS Data Center 1997).

the effect of recent colluvial deposits. Eight pits (Figure 3; IDs 1–8) were targeted in areas with FR-ha on FT1–3, which is the predominant soil type in this area based on the latest soil map (EU 2013). Due to bad road conditions caused by a landslide, we could not reach the selected location for one soil pit (ID 4), and therefore no sample was obtained for ID 4. A soil pit (ID 9) was targeted on the greenstone belt on S2 where Plinthosols were expected according to EU

(2013). One pit (ID 10) was made in the intensive paddy fields in the south-western plain of Lake Alaotra on FT2, and another (ID 11) was made in a cropland in Betafo where volcanic soils are distributed on FT1. In EU (2013), soil types in these agricultural areas were HS-dy and CA-eu, respectively. Soil profiles in the pits were described according to terms in FAO (2006). Soil samples were taken from each horizon using a trowel, air-dried, passed through a 2-mm

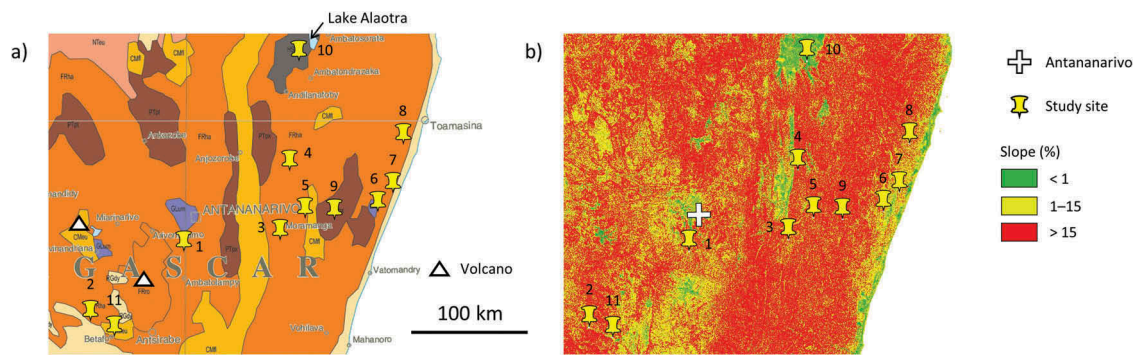


Figure 3. Location of soil pits on (a) soil map (EU 2013) and (b) slope map.

sieve, and stored at a room temperature until subsequent physical and chemical analyses in the laboratory. Cylinder core samples were also taken for measuring bulk density.

Soil samples were mixed in a soil to solution (H_2O and 1 M KCl) ratio of 1:5, and the pH was measured with a pH electrode (LAQUA F-72, HORIBA) after being shaken for 2 h. Total carbon and total nitrogen were determined using the dry combustion method with an elemental analyzer (SUMIGRAPH NC-220F, Sumika Chemical Analysis Service). Exchangeable bases, exchangeable acidity, and cation exchangeable capacity (CEC) were generally measured according to van Reeuwijk (2002), but inductively coupled plasma atomic emission spectroscopy (ICP-AES) with a spectrometer (ICPE-9000, SHIMADZU) was used in this study. Effective base saturation (EBS), as defined by the IUSS Working Group WRB (2015), was calculated by dividing exchangeable ($\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^+ + \text{Na}^+$) by exchangeable ($\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^+ + \text{Na}^+ + \text{Al}^{3+}$). Available phosphorus was determined by the Bray-I method (Bray and Kurtz 1945; van Reeuwijk 2002) using a Shimadzu UV-1800 spectrophotometer. Particle size distribution was measured as described by Gee and Bauder (1986). After chemical degradation of organic matter by H_2O_2 and sample dispersion with sodium hexametaphosphate, coarse- and fine-sand fractions (0.2–2 and 0.02–0.2 mm, respectively) were measured by sieving, and silt (0.002–0.02 mm) and clay (<0.002 mm) fractions were measured with the pipette method. Bulk density was determined by measuring the masses of 100-ml core samples after drying at 105°C for 48 h. To test the diagnostic criteria of Andic properties, phosphorus retention (Soil Survey Staff 2014a) and oxalate-extractable Al and Fe (Courchesne and Turmel 2008) were measured for soils from ID 11.

Based on soil profile descriptions and soil physical and chemical data, soil profiles were classified to the level of principal qualifiers using the current World Reference Base (WRB) system (IUSS Working Group WRB 2015) as well as to the level of subgroups using Keys to Soil Taxonomy (Soil Survey Staff 2014b).

3. Results and discussion

3.1. Distribution of Ferralsols and Plinthosols

Among IDs 1–8 targeted in areas with FR-ha on FT1–3, Ferralsols were not found on steep land (> 15% slope) (Table 1; IDs 3, 7). Rather, soils in steep land on FT1–3 (IDs

3, 7) were classified as Cambisols due to the lack of clay accumulation and a ferralic horizon (Tables 2 and 3), which is similar to the distribution of Cambisols on S1 and a part of S2 (EU 2013). Since the area with steep slope (> 15%) was estimated to cover 62% of the whole surveyed area (Figure 3; the average slope was 21%), it was suggested that Ferralsols cover less area than reported by EU (2013). Even on flat to sloping land (< 15% slope) on FT1–3 (Table 1), we found both Ferralsols (IDs 1, 6, 8) and Lixisols (IDs 2, 5), the latter of which has an argic horizon but not a ferralic horizon. According to the legacy soil data of Africa Soil Profiles Database (Leenaars, van Oostrum, and Ruiperez Gonzalez 2014), the 20 soil profiles with reliability level 1 in Madagascar (same as 20 profiles in Version 3.1 of the ISRIC-WISE database meeting the Reference Pedon Description criteria) were located within the area of 1,130 km² covered mainly by steep slope (74% of total area) on FT2. Although those 20 soil profiles were expected to be FR-ha (EU 2013), only three profiles were classified as Ferralsols, while the most of others were classified as Cambisols or Fluvisols. This together with our results indicates that, due to the extensive steep land, Ferralsols cover less area on FT1–3 than reported in EU (2013); Cambisols could be more common than Ferralsols.

The area covered by FR-ha in EU (2013) basically corresponds to the range of '*Sols Ferrallitiques*' in Roederer and Bourgeat (1969) and Riquier (1968), both of which followed the old French classification system (Aubert 1965). It is worth noting that neither Aubert (1965) nor CPCS (1967) used the concept of a diagnostic horizon, which has been adopted in many current soil classification systems (e.g., Obara et al. 2011; Soil Survey Staff 2014b; IUSS Working Group WRB 2015). Moreover, the criteria for '*Sols Ferrallitiques*' have no critical value for thickness and soil texture (Aubert 1965). Latham (1981) noted the absence of precise diagnostic features as a problem of the old French classification systems, and that sometimes too much was left open to personal interpretation. In contrast, the ferralic horizon is used as a diagnostic horizon for Ferralsols and has more specific criteria, for example, certain values must be met for soil texture class, CEC, and exchangeable cations ($\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^+ + \text{Na}^+ + \text{Al}^{3+}$) (IUSS Working Group WRB 2015). Our results indicate that relying on the old classification system resulted in an overestimation of FR-ha in the targeted region in EU (2013).

Table 1. List of candidate sites for soil survey.

ID	Latitude	Longitude	Geology			Soil			Land use	Slope ¹
			Schlüter (2008)	Basaire (1968)	Riquier (1968)	Riquier (1968)	EU (2013)			
1	19°12'S	47°28'E	Migmatites (Archean)	Basaire (1968)	Riquier (1968)	Riquier (1968)	Haplic Ferralsols	Pasture	Gently sloping	
2	19°46'S	46°41'E	Migmatites (Archean)	Gneiss et micaschistes	Granite	Sols à Sesquioxides, Sols Typiques, Rouges, Phase érodée	Haplic Ferralsols	Cultivated	Gently sloping	
3	19° 7'S	48°14'E	Migmatites (Archean)	Migmatites	Gneiss (acide)	Sols à Sesquioxides, Sols Typiques, Rouges, Phase érodée	Haplic Ferralsols	Secondary forest	Steep	
4	18°35'S	48°18'E	Granitoid migmatites (Archean)	Migmatites	Gneiss (acide)	Sols à sesquioxides, Sols Typiques, Rouges, Phase humifère locale	Haplic Ferralsols	-	-	
5	18°57'S	48°25'E	Granitoid migmatites (Archean)	Migmatites granitoïdes, Granites migmatitiques	Gneiss (acide)	Sols à Sesquioxides, Sols Typiques, Rouges, Phase humifère locale	Haplic Ferralsols	Secondary forest	Strongly sloping	
6	18°54'S	48°58'E	Migmatites (Archean)	Migmatites granitoïdes, Granites migmatitiques	Gneiss (roches diverses acides)	Sols à Sesquioxides, Sols indurés concrétionnés ou cuirassés, Jaune/Rouge (en général)	Haplic Ferralsols	Long-term fallow	Strongly sloping	
7	18°45'S	49° 5'E	Migmatites (Archean)	Migmatites	Gneiss (acide)	Sols à Sesquioxides, Sols Typiques, Jaune sur Rouge (parfois tendance hydromorphe)	Haplic Ferralsols	Long-term fallow	Moderately steep	
8	18°24'S	49°10'E	Granitoid migmatites (Archean)	Migmatites	Gneiss (acide)	Sols à Sesquioxides, Sols Typiques, Jaune sur Rouge (parfois tendance hydromorphe)	Haplic Ferralsols	5-year fallow	Strongly sloping	
9	18°58'S	48°39'E	Greenstone belt (Paleoproterozoic)	Migmatites à amphibole	Gneiss (roches diverses acides)	Sols à Sesquioxides, Sols indurés concrétionnés ou cuirassés, Jaune/Rouge (en général)	Pisoplinthic Plinthosols	Long-term fallow	Steep	
10	17°46'S	48°22'E	Greenstone belt (Paleoproterozoic)	Alluvions, sables	-	Sols Hydromorphes, Sols Organiques/Minéraux, Sols Tourbeux, Semi-tourbeux, A gley et pseudogley, Peu humifères a gley de profondeur	Dystric Histosols	Cultivated	Nearly level	
11	19°51'S	46°52'E	Migmatites (Archean)	Basanites, Basanitoïdes, Quaternary	Cendres volcaniques	Sols des pays tropicaux, Sols Bruns Eutrophes, Peu évolués	Eutric Cambisols	Cultivated	Very gently sloping	

¹Flat: 0–0.2%, Level: 0.2–0.5%, Nearly level: 0.5–1.0%, Very gently sloping: 1.0–2.0%, Gently sloping: 2–5%, Sloping: 5–10%, Strongly sloping: 10–15%, Moderately steep: 15–30%, Steep: 30–60%, Very steep: > 60% (FAO 2006).

Table 2. Soil classification and physical properties.

ID	Depth	Rock fragment	Hardness	Bulk density	Coarse sand	Fine sand	Coarse silt	Fine silt	Clay
Soil classification based on WRB and Soil Taxonomy	cm	%	mm	Mg m ⁻³	% fine earth				
1	0–12	0.3	29.0	1.32	22.9	16.9	7.2	11.1	41.9
Geric Ferralsol	12–35	3.7	31.7	1.27	26.3	14.6	5.9	12.6	40.6
(Eutric Chromic Sideralic Cambisol over Geric Ferralsol),	35–73/80	3.6	29.0	1.52	24.6	14.0	7.0	14.8	39.7
Typic Dystrudept (Typic Dystrudept)	73/80–125+	0.3	28.7	1.50	26.3	15.1	10.6	22.2	25.8
2	0–13	-	-	-	19.7	13.8	9.6	26.3	30.7
Rhodic Lixisol,	13–33	-	-	-	18.0	10.5	4.9	16.1	50.6
Typic Rhodudalf	33–47	-	-	-	17.4	12.4	8.5	17.8	44.1
	47–72	-	-	-	14.0	8.0	4.5	9.7	63.8
	72–83+	-	-	-	13.0	11.5	5.9	11.4	58.2
3	0–20	0.1	20.0	0.96	31.9	10.6	2.5	8.9	46.0
Dystric Chromic Sideralic Cambisol,	20–40	0.1	26.7	1.23	26.6	10.6	3.2	14.8	44.8
Typic Dystrudept	40–78	0.3	27.7	1.43	25.6	16.3	5.3	22.4	30.4
	78–104	0.1	24.7	1.41	25.9	24.8	5.5	20.7	23.1
	104–125+	0.1	23.0	1.34	28.6	24.9	6.1	20.5	19.8
5	0–5	9.1	8.0	1.28	57.6	8.0	3.1	7.1	24.1
Ferric Lixisol	5–39	6.6	26.7	1.57	41.2	10.9	4.2	11.0	32.7
(Dystric Sideralic Cambisol over Eutric Regosol),	39–80	5.7	28.3	1.57	21.4	4.6	2.1	12.2	59.7
Typic Hapludalf (Inceptic Hapludalf)	80–125+	9.1	27.3	1.58	24.1	4.2	1.9	10.5	59.4
6	0–22	3.7	20.7	1.24	33.6	11.3	3.7	13.2	38.2
Geric Xanthic Ferralsol,	22–56	4.8	25.7	1.53	35.4	9.9	3.3	12.3	39.1
Anionic Acrudox	56–85	9.9	26.0	1.54	34.4	8.8	3.6	11.2	42.0
	85–130+	3.8	24.7	-	37.4	8.1	5.5	14.3	34.7
7	0–26	0.6	17.3	1.27	55.1	11.6	2.5	6.8	23.9
Dystric Chromic Cambisol	26–55	7.2	23.0	1.63	51.4	11.5	2.4	6.2	28.4
(Dystric Cambisol over Dystric Chromic Cambisol),	55–100	0.5	26.0	1.40	23.6	8.3	6.9	23.1	38.1
Typic Paleudult (Ruptic-Ultic Dystrudept)	100–130+	0.9	24.0	1.43	19.1	11.8	8.3	24.3	36.4
8	0–18	0.2	24.7	1.23	42.7	8.9	2.7	7.3	38.3
Geric Xanthic Ferralsol,	18–36	5.0	26.3	1.51	41.8	8.6	3.2	8.8	37.6
Oxic Dystrudept	36–59	0.7	24.7	1.55	39.0	7.8	3.1	10.1	39.9
	59–93	1.5	24.7	1.45	26.6	9.0	6.5	20.5	37.4
	93–130+	0.1	23.7	1.49	18.7	11.4	10.0	28.3	31.6
9	0–26	5.8	19.3	1.0	43.1	17.3	3.2	10.5	25.8
Geric Plinthosol,	26–55	1.7	25.0	1.6	36.0	15.1	3.5	8.4	37.0
Plinthic Kandiodalf	55–81	3.5	21.0	1.5	30.8	12.4	2.9	10.1	43.7
	81–125+	3.1	20.7	1.3	21.5	9.4	3.2	13.5	52.5
10	0–18	0.0	34.7	1.4	2.3	4.8	6.9	17.2	68.9
Hydragric Vertisol,	18–34	0.0	29.3	1.4	1.2	4.6	5.5	13.8	74.8
Ustic Epiaquert	34–43	0.0	24.7	1.1	9.7	8.0	10.0	20.1	52.3
	43–82	0.0	22.7	1.6	1.5	12.4	12.9	16.7	56.5
	82–130+	1.0	13.3	1.6	41.1	30.1	6.5	3.6	18.7
11	0–10	-	-	-	19.7	23.2	16.8	28.2	12.0
Eutric Umbric Silandic Andosol,	10–25	-	-	-	19.4	21.4	15.8	31.6	11.8
Typic Hapludand	25–37	-	-	-	14.1	16.7	24.9	32.5	11.7
	37–60	-	-	-	18.2	18.1	19.8	34.5	9.5
	60–70	-	-	-	-	-	-	-	-
	70–80+	-	-	-	13.6	18.2	18.4	35.7	14.2

Soil classification in parenthesis shows the case of soil profile developed from base materials and overlying redeposited materials.

In addition, Ferralsols (IDs 1, 6, 8) generally have geric properties, i.e., low effective CEC (sum of exchangeable cations) and/or high anion exchangeable capacity ($\text{pH}_{\text{KCl}} - \text{pH}_{\text{water}} \geq +0.1$) (Table 3), and thus Geric Ferralsols (FR-gr) should be predominant in the targeted region rather than FR-ha. It is difficult to interpret the reason of this mismatch because we have no means of knowing how '*Sols Faiblement Ferrallitiques*,' '*Sols Typique Rouges*,' and '*Sols Typique Jaune sur Rouge*' in '*Sols Ferrallitiques*' were translated into Orthic Ferralsols in FAO-UNESCO (1977) and then FR-ha in EU (2013). It could be due to the lack of enough soil data. Note that Geric properties were not used in FAO-UNESCO (1977) and were firstly introduced in Revised Legends (FAO 1988), which could have influenced the mismatch.

Translation of a soil order in WRB system to a similar soil order in Soil Taxonomy and vice versa is a simple method often applied to make a soil map. Ferralsols are considered to overlap with Oxisols – highly weathered tropical soils with low natural fertility in Soil Taxonomy classification system (Soil Survey Staff 2014b). It's worth noting that, however, Oxisols were rarely found in the targeted region. This is attributed to the stricter condition of the oxic horizon than the ferralic horizon: an oxic horizon requires < 5% of rock structure and a clay increase with increasing depth, whereas a ferralic horizon accepts < 80% coarse fragment and a clay decrease with increasing depth like IDs 1 and 8. This difference has caused the mismatch of Ferralsols and Oxisols in the targeted region.

Table 4. Description of soil profiles.

ID	Depth (cm)	Horizon ⁴	Color	Boundary (distinctness/ ¹ / topography ²)	Rock Fragment (abundance/ ³ / weathering ⁴ /shape ⁵ /size ⁶ /nature ⁷)	Structure (grade ⁸ /type ⁹ / size ¹⁰)	Consistence ¹¹	Moisture ¹²	Porosity (size ¹³ /type ¹⁴ / abundance ³)
1	0–12	A	5YR6/6	C/S	none	M/SB/FI	SHA	D	VF/C/M, F/C/C, M/C/C, C/C/V
	12–35	Bw1	5YR5/6	C/W	C/F/A/F/QU	M/SB/ME	FI	SM	VF/C/M, F/C/C, M/C/F
	35–73/80	Bw2	5YR4/6	C/S	F/F/A/F/QU, M/F/A/C/QU	M/SB/CO	FI	SM	VF/C/C, F/C/F, M/C/F
	73/80–125+	Bw3	2.5YR4/5, 7.5R3/6		M/S/S/C/MI	W/SB/ME	FR	SM	VF/C/C, F/C/F
2	0–13	Ap	5YR3/4	C/S	none	M/SB/FI	SO	D	VF/C/F, F/C/C
	13–33	Bw1	2.5YR3/6	G/S	none	S/SB/CO	FI	SM	VF/C/F, F/C/F
	33–47	Bw2	2.5YR3/6	G/S	none	S/SB/CO	FR	SM	VF/C/F, F/C/V
	47–72	Btw	10R3/6	G/S	none	S/SB/CO	FR	M	VF/C/F
3	72–83+	Bw3	10R3/6		none	S/SB/CO	FR	M	VF/C/F
	0–20	A	5YR6/2, 5YR7/6	C/W	none	M/GR/CO	FR	SM	VF/C/M, F/C/M, M/C/C, C/C/V
	20–40	Bw1	5YR6/6	G/W	none	M/SB/ME	FR	M	VF/C/C, F/C/C, M/C/F
	40–78	Bw2	2.5YR7/6	D/S	none	M/SB/CO	FR	M	VF/C/C, F/C/C
5	78–104	Bw3	2.5YR7/6	C/W	F/F/S/M/QU	W/SB/CO	FR	M	VF/C/C
	104–125+	BC	2.5YR7/6		C/F/S/C/QU	W/SB/CO	FR	M	VF/C/C
	0–5	A	10YR6/6	C/W	C/F/S/F/QU	W/GR/ME	FR	SM	VF/C/M, F/C/M, M/C/C, C/C/F
	5–39	Bw	10YR5/8	G/S	F/F/S/F/QU, F/F/A/S/QU	W/SB/CO	FR	M	VF/C/M, F/C/C
6	39–80	Btc	5YR5/8	G/W	none	M/SB/CO	FR	M	VF/C/C, F/C/F
	80–125+	Bwc2	5YR5/8		none	M/SB/CO	FR	M	VF/C/C, F/C/F
	0–22	A	10YR6/4	G/W	F/F/A/M/QU	M/GR/CO	FR	SM	VF/C/M, F/C/M, M/C/C
	22–56	Bw1	7.5YR6/8	D/S	F/F/A/M/QU	W/SB/CO	FR	SM	VF/C/M, F/C/C, M/C/F
7	56–85	Bw2	7.5YR5/8	D/W	F/F/A/M/QU	W/SB/CO	FR	M	VF/C/M, F/C/C, M/C/V
	85–130+	BC	7.5YR5/8		F/F/A/F/QU	W/SB/ME	FR	M	VF/C/M, F/C/C
	0–26	A1	2.5YR4/1	G/W	C/F/A/F/QU	W/GR/CO	FR	M	VF/C/M, F/C/M, M/C/C
	26–55	A2	7.5YR6/6	G/W	F/F/A/F/QU	W/SB/CO	FR	M	VF/C/M, F/C/C, M/C/V
8	55–100	Bw1	7.5YR6.5/8, 2.5YR4/8	D/S	V/F/A/F/QU	M/SB/CO	FR	M	VF/C/M, F/C/F
	100–130+	Bw2	2.5YR4/8, 7.5YR6.5/8		V/F/A/F/QU	M/SB/CO	FR	M	VF/C/M, F/C/F
	0–18	A1	10YR5/3	C/S	none	M/GR/CO	FR	M	VF/C/M, F/C/M, M/C/F, C/C/V
	18–36	A2	7.5YR4/2, 7.5YR6/8	G/W	F/F/A/F/QU	M/SB/CO	FR	M	VF/C/M, F/C/C, M/C/V
9	36–59	Bw1	7.5YR5/8, 7.5YR4/2	G/W	V/F/A/F/QU	M/SB/CO	FR	M	VF/C/M, F/C/F
	59–93	Bw2	2.5YR5/8, 7.5YR5/8	D/W	V/F/A/F/QU	W/SB/CO	FR	M	VF/C/C, F/C/V
	93–130+	Bw3	10R4/6		V/F/A/F/QU	W/SB/CO	FR	M	VF/C/C, F/C/V
	0–26	Ac	10YR4/4	G/I	C/F/A/C/QU	W/GR/CO	FR	SM	VF/C/M, F/C/M, M/C/M
10	26–55	Bwc1	7.5YR5/8	G/S	F/F/A/C/QU	M/SB/CO	FR	M	VF/C/M, F/C/C
	55–81	Bwc2	7.5YR5/8	D/S	C/F/A/S/QU	M/SB/CO	FR	M	VF/C/C, F/C/C
	81–125+	Bwc3	5YR5/8		C/F/A/S/QU	M/SB/CO	FR	M	VF/C/C, F/C/C
	0–18	Alpg	10YR5/1	A/W	none	S/SP/VC	EFI	SM	VF/C/M, F/C/F, C/I/V
17	18–34	Alg	10YR5/1	C/W	none	S/SP/VC	VFI	SM	VF/I/M, VF/C/M, F/C/V
	34–43	Bwc	10YR5/1, 10YR7/6	C/W	none	M/SB/FI	FR	SM	VF/C/M, F/C/M
	43–82	Bwl	7.5YR6.5/1, 10YR6/1	A/S	none	S/MA	VFI	SM	VF/C/M, F/C/F
	82–130+	Cl	10YR5/2		C/F/S/F/QU	no	LO	W	VF/C/M, F/C/M

(Continued)

Table 4. (Continued).

ID	Depth	Horizon ⁵	Color	Boundary	Rock Fragment	Structure	Consistence ¹¹	Moisture ¹²	Porosity
11	0–10	Ap	5YR2.5/1	G/S	none	W/GR/FI	VFR	SM	VF/B/C, F/C/C
	10–25	A	5YR2.5/1	C/W	none	W/GR/FI	VFR	SM	VF/B/C, F/C/F
	25–37	2Ab1	10YR2/1	G/S	none	M/SB/FI	FR	M	VF/B/C, F/C/V
	37–60	2Bwb	5YR2.5/1	A/B	none	M/SB/ME	FR	M	VF/B/F, F/C/V
	60–70	2C	5Y2/1	A/B	M/F/A/F/FE, A/F/A/M/FE	no	'	-	-
	70–80+	3Ab2	10YR3/4		none	no	LO	SM	VF/C/V

¹Horizon in parenthesis shows the case of soil profile to be buried.

¹A: abrupt, C: clear, G: gradual, D: diffuse.

²S: smooth, W: wavy, I: irregular, B: broken.

³none, V: very few, F: few, C: common, M: many, A: abundant, D: dominant.

⁴F: fresh or slightly weathered, W: weathered, S: strongly weathered.

⁵F: flat, A: angular, S: subrounded, R: rounded.

⁶F: fine gravel, M: medium gravel, C: coarse gravel, S: stones, B: boulders, L: large boulders.

⁷QU: quartz, MI: mica, FE: feldspar.

⁸W: weak, M: moderate, S: strong.

⁹RS: rock structure, ST: stratified structure, SG: single grain, MA: massive, PM: porous massive, AB: angular blocky, AS: angular and subangular blocky, SA: subangular and angular blocky, SB: subangular blocky, NS: nutty subangular blocky, PR: prismatic, SP: subangular prismatic, WE: wedge-shaped, CO: columnar, GR: granular, WC: worm cast, PL: platy, CL: cloddy, CR: crumbly, LU: lumpy

¹⁰VF: very fine, FI: fine, ME: medium, CO: coarse, VC: very coarse, EC: extremely coarse.

¹¹LO: loose, VFR: very friable, FR: friable, FI: firm, VFI: very firm, EFI: extremely firm, SO: soft, SHA: slightly hard, HA: hard, VHA: very hard, EHA: extremely hard

¹²VD: very dry, D: dry, SM: slightly moist, M: moist, W: wet.

¹³VF: very fine, F: fine, M: medium, C: coarse, VC: very coarse.

¹⁴I: interstitial, B: vesicular, V: vughs, C: channels, P: planes.

In the former version of the WRB system (IUSS Working Group WRB 2006), Lixisols for IDs 2 and 5 were classified as Acrisols, but the diagnostic criteria to key out Acrisols was changed in the latest WRB system (IUSS Working Group WRB 2015). Base saturation defined as exchangeable ($\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^+ + \text{Na}^+$) divided by CEC (by 1 M NH_4OAc) was used in the former system, whereas effective base saturation defined as exchangeable ($\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^+ + \text{Na}^+$) divided by exchangeable ($\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^+ + \text{Na}^+ + \text{Al}^{3+}$) is used in the latest system. Owing to this modification, the soil profiles with low exchangeable Al^{3+} tend to be classified as Lixisols. This is in contrast to that Ferralsols used to be considered to occur alongside of Acrisols on acidic rocks (FAO 2001) because the principal difference between those two soil types was just the presence or absence of clay accumulation. Therefore, the introduction of effective base saturation needs special attention when we describe soil maps of tropical area.

A soil profile (ID 9) located on basic rocks (amphibole migmatite) was classified as Geric Plinthosol (PT-gr); the profile had a plinthic horizon, but not a pisoplinthic horizon. This is basically in agreement with the soil type previously reported as '*Sols indurés*' (indurated soils) (Riquier 1968) and Pisoplinthic Plinthosol (PT-px) in EU (2013), although the principal qualifier was different (Figure 1). The cover range of PT-gr is expected to follow that of PT-px in the latest soil map.

It is worth noting that, because of the large gap in soil texture, color, or content of rock fragment between a pair of adjacent layers within a soil profile, some pedons (IDs 1, 5, 7) can be considered as not being evolved from a base parent material but developed from different parent materials; that is, there are base materials and overlying redeposited materials. This can lead to classifying these pedons as Cambisols due to the absence of a ferralic horizon in overlying soils (IDs 1, 5, 7; Tables 2–4). Consequently, Ferralsols were rarely found in the central highlands, and Cambisols were the predominant soil type. Roederer and Bourgeat (1969) noted that the pronounced relief of Madagascar has led to intense soil erosion and stripping of old soil formations. Therefore, there might be various soils that recently developed from the base parent materials or from the overlying redeposited materials. More work is required to reveal the extent to which soils developed from secondary materials existing in the east coast and central highlands of Madagascar.

3.2. Paddy soils around Lake Alaotra and sporadic volcanic soils in the central highlands

Riquier (1968) classified the soil type of southwestern plain of Lake Alaotra as '*Sols Hydromorphes, Sols Organiques/Minéraux*' (hydromorphic, organic/mineral soils), and this was translated into Histosols in the latest soil map (EU 2013). Irrigated rice cultivation in this area doubled between 1959 and 1989, and about 820 km² of irrigated paddy fields currently cover the Alaotra topographic basin, or about 45% of the total basin (Mietton et al. 2018). The soil profile observed in the southwestern plain of Lake Alaotra (ID 10)

Table 5. Description of soil profiles.

ID	Depth (cm)	Mottle (color/abundance ¹ /contrast ² /size ³ / boundary ⁴)	Concentration		Root (size ³ / abundance ¹)	Other biological feature (abundance ¹ /type ⁹)
			(abundance ¹ /kind ⁵ /size ³ /shape ⁶ /hardness ⁷ /nature ⁸ / color)			
1	0–12	none	none	none	VF/M, F/C, M/V	C/T
	12–35	none	none	none	VF/M, F/F, M/V	F/T
	35–73/80	none	none	none	VF/C, F/V	none
	73/80–125	none	none	none	VF/F	none
2	0–13	none	none	none	VF/M, F/V	none
	13–33	none	none	none	VF/C	none
	33–47	none	none	none	VF/F	none
	47–72	none	none	none	VF/V	none
	72–83+	none	none	none	none	none
3	0–20	none	none	none	VF/M, F/C, M/F	C/E
	20–40	none	none	none	VF/C, F/F, M/V	none
	40–78	none	none	none	VF/F, F/V	none
	78–104	none	none	none	VF/V	none
5	104–125+	none	none	none	VF/V	none
	0–5	none	none	none	VF/M, F/C, M/F, C/ V	F/E
	5–39	none	none	none	VF/C, F/F	none
	39–80	none	M/C/F/E/H/F/5YR5/8, C/SC/F/E/B/F/7.5YR6/8	none	VF/F	none
6	80–125+	none	M/C/F/E/H/F/5YR5/8, C/SC/M/E/B/F/7.5YR6/8	none	none	none
	0–22	none	C/R/M/I/H/F/7.5R4/6	VF/M, F/C, M/C	C/E	
	22–56	none	C/R/C/I/H/F/7.5R4/6	VF/C, F/F, M/F	F/T	
	56–85	none	C/R/C/I/H/F/7.5R4/6	VF/C, F/F	F/T	
	85–130+	none	A/R/C/I/H/F/7.5R4/6	VF/F, F/V	none	
7	0–26	none	none	none	VF/M, F/C, M/C, C/ V	C/E
	26–55	none	M/R/C/I/H/F/7.5R4/6	VF/C, F/F, M/V	F/T	
	55–100	none	C/R/C/I/H/F/7.5R4/6	VF/F, F/F	none	
	100–130+	none	C/R/C/I/H/F/7.5R4/6	VF/V, F/V	none	
8	0–18	none	none	none	VF/M, F/C, M/V, C/ V	C/E
	18–36	none	F/C/M/I/B/F/10R4/8, 7.5YR5/8	VF/C, F/F, M/V	F/C	
	36–59	none	none	VF/F, F/V, M/V	F/C	
	59–93	none	C/R/C/I/B/F/7.5R4/6	VF/F, F/V, M/V	none	
9	93–130+	none	C/R/C/I/B/F/7.5R4/6	VF/V, F/V	none	
	0–26	none	M/C/C/I/H/FM/10YR2/2, 10YR5/6	VF/M, F/C, M/F	F/T, C/E	
	26–55	none	C/C/C/I/H/FM/10YR2/2, 10YR5/6	VF/C, F/F	none	
	55–81	none	C/C/C/I/H/FM/10YR2/2, 10YR5/6	VF/F, F/V	none	
	81–125+	none	M/C/C/I/H/FM/10YR2/2, 10YR5/6	VF/V, F/V	none	
10	0–18	2.5YR4/8/M/P/VF/S, 2.5YR4/8/V/P/F/S	none	none	VF/M, F/F	none
	18–34	2.5YR4/8/M/P/VF/S, 2.5YR4/8/M/P/F/S	none	none	VF/M, F/F	none
	34–43	none	A/C/M/E/H/F/10YR7/8	VF/C, F/F	none	
	43–82	2.5YR4/8/C/P/VF/S	none	none	VF/C, F/F	none
	82–130+	2.5YR4/8/N/P/VF/S	none	none	VF/V	none
11	0–10	none	none	none	VF/C, F/C	none
	10–25	none	none	none	VF/C, F/F	none
	25–37	none	none	none	VF/F, F/V	none
	37–60	none	none	none	VF/F	none
	60–70	none	none	none	none	none
	70–80+	none	none	none	none	none

¹none, V: very few, F: few, C: common, M: many, A: abundant, D: dominant.

²F: faint, D: distinct, P: prominent.

³VF: very fine, F: fine, M: medium, C: coarse, VC: very coarse.

⁴S: sharp, C: clear, D: diffuse.

⁵T: crystal, C: concretion, SC: soft concretion, S: soft segregation, N: nodule, IP: pore infillings, IC: crack infillings, R: residual rock fragment, O: other.

⁶R: rounded, E: elongated, F: flat, I: irregular, A: angular.

⁷H: hard, S: soft, B: both hard and soft.

⁸K: carbonates, KQ: carbonates-silica, C: clay, CS: clay-sesquioxides, GY: gypsum, SA: salt, GB: gibbsite, JA: jarosite, S: sulfur, Q: silica, F: iron, FM: iron-manganese, M: manganese, NK: not known.

⁹A: artifacts, B: burrows, BO: open large burrows, BI: infilled large burrows, C: charcoal, E: earthworm channels, P: pedotubules, T: termite or ant channels and nests, I: other insect activity.

was classified as Hydragic Vertisol (VR-hg), which had high nutritional status in the top three layers presumably due to constant sedimentation through irrigation water (Tables 2 and 3). In addition, there were redoximorphic features in the top two layers and Fe and Mn accumulation in the underlying horizon, suggesting the effect of irrigated rice

cultivation (Tables 4 and 5). The combined thickness of anthraquic horizon and hydragic horizon can vary depending on the distance from canals and farmers' cultivation practice. Therefore, other places within the paddy fields in southwestern plain of Lake Alaotra could be classified as Anthrosols which can be keyed out prior to Vertisols if the

Table 6. Oxalate-extractable Al and Fe and P retention at ID 11.

ID	Depth cm	Alox	Feox	Alox + 0.5 Feox	
				Feox	P retention
Soil classification based on WRB and Soil Taxonomy				%	
11	0–10	3.7	2.4	4.8	92.2
Eutric Umbric	10–25	3.5	2.2	4.6	93.3
Silandic Andosol,	25–37	5.0	2.6	6.3	98.2
Typic Hapludand	37–60	5.3	2.5	6.5	98.4
	60–70	-	-	-	-
	70–80+	1.7	2.0	2.7	88.4

Alox, oxalate-extractable Al; Feox, oxalate-extractable Fe. Alox + 0.5 Feox value of $\geq 2\%$ and P retention of $\geq 85\%$ are required for Andic properties (IUSS Working Group 2015).

combined thickness of anthraquic horizon and an underlying hydragric horizon is > 50 cm.

It is indicated that long-term paddy-rice cultivation (> 50 years) associated with a drainage system generated redoximorphic features within soil profiles (Tables 4 and 5) and accelerated the oxidation of the organic layer of Histosols, thereby changing Histosols into Vertisols. The previous studies also reported the changes in soil order by the rapid oxidation of the organic layer. Veenstra and Burras (2012) showed that 32% of the pedons originally classified as 'Black soils' (Phaeozem and Chernozem) (IUSS Working Group WRB 2006) in Iowa, United States, have changed enough to result in reclassification at the WRB Reference Group level, mainly because of degradation of organic molic horizons after more than 50 years of cultivation. Everett (1983) suggested that a few decades of oxidation after the drainage can make the thickness or organic matter content in the organic layer insufficient for the soil to meet the requirements for the order of Histosols. Kempen et al. (2009) reported that accelerated oxidation due to the intensive tillage and lowering of groundwater levels largely changed peat soils into other soil groups in North-East Netherlands. We should also mention that the old soil classification system (Aubert 1965) can classify the soils having much less than 20% of soil organic carbon, which are not classified as Histosols in the current WRB system, as "*Sols Hydromorphes*." The difference of the definition of peat soils between old and current soil classification systems could also be attributed to the change of soil type.

One soil profile (ID 11) was classified as Andosol, and it had a layer with andic properties (Table 6). This result is in contrast to the latest soil map (EU 2013), which reported CA-eu in this area, but was similar to Riquier (1968), who classified the soils in this area as '*Sols bruns eutrophes peu évolués cendres volcanique*' (little-evolved eutrophic brown soils on volcanic ash). Therefore, future soil survey should focus on investigating to what extent the area of the '*Sols bruns eutrophes peu évolués cendres volcanique*' in Betafo can be translated as Andosols in the current classification system. This update at the WRB Reference Group level would lead us to expect that the soils in this area have favorable properties for cultivation, such as a high water-holding capacity and organic matter content, but also unfavorable properties, such as a high phosphorus fixing capacity caused by active Al and Fe. These findings on the

paddy soils around Lake Alaotra and sporadic volcanic soils in the central highlands have important implications about the land history and for appropriate land management in the agriculturally intensive areas in the central highlands of Madagascar.

4. Conclusions

There was a considerable gap between our results and the latest country-scale soil map of Madagascar, which was put together from an atlas and pedological map published more than 50 years ago. The current soil classification systems have more distinct diagnostic criteria, and intensive cultivation has changed soil conditions in some areas. Clearly, there is a need to reconsider the country-scale soil map of Madagascar. Further soil survey toward updating soil map in Madagascar should focus on how to link the topographical information and distribution of Ferralsols and accumulation of soil profile information of overlooked soil types in the current soil map to determine their occurrence.

Acknowledgments

The authors would like to thank the Ministry of Mines and Strategic Resources of Madagascar, regional officers, and farmers for their kind support for our soil survey. The authors would also like to thank Dr. Makoto Shibata for his valuable comments on data of the soil profiles. The soil samples were legally imported to Japan for soil analysis with permission of 30Y1175.

Disclosure statement

The authors report no conflicts of interest.

Funding

This work was financially supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA) [Grant No. JPMJSA1608].

ORCID

Tomohiro Nishigaki  <http://orcid.org/0000-0002-6669-803X>
 Yasuhiro Tsujimoto  <http://orcid.org/0000-0001-7738-9913>
 Andry Andriamananjara  <http://orcid.org/0000-0001-5372-7359>

References

- Aubert, G. 1965. "La classification pedologique utilisée en France." *Pédologie (Grand.) Symposium International*, 3. *Classification des Sols*, 25–26. http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_5/b_fdi_10-11/12533.pdf
- Basairie, H. 1968. "Carte géologique 1/500000." Madagascar: Service Géologique.
- Batjes, N. H. 2008. "ISRIC-WISE Harmonized Global Soil Profile Dataset (Ver. 3.1)." *Report 2008/02*. Wageningen: ISRIC – World Soil Information.
- Bray, R. H., and L. T. Kurtz. 1945. "Determination of Total, Organic, and Available Forms of Phosphorus in Soils." *Soil Science* 59: 39–45. doi:10.1097/00010694-194501000-00006.

- Courchesne, F., and M. C. Turmel. 2008. "Extractable Al, Fe, Mn, and Si." In *Soil Sampling and Methods of Analysis*, edited by M. R. Carter and E. G. Gregorich, 307–315. 2nd ed. Boca Raton, FL: Canadian Society of Soil Science, CRC Press.
- CPCS. 1967. "Classification des Sols." *Ecole Nationale Supérieure d'Agronomie, Grignon*. http://library.wur.nl/isric/fulltext/isricu_i33696_001.pdf
- EU. 2013. *Soil Atlas of Africa*. Edited by A. Jones, H. Breuning-Madsen, M. Brossard, A. Dampha, J. Deckers, O. Dewitte, T. Gallali et al. Luxembourg: European Commission, Publications Office of the European Union.
- Everett, K. R. 1983. "Histosols." In *Pedogenesis and Soil Taxonomy: II. The Soil Orders*, edited by L. P. Wilding, N. E. Smeck, and G. F. Hall, 1–53. Amsterdam: Elsevier Science Publishers.
- FAO. 1988. "FAO-UNESCO Soil Map of the World Revised Legend." Rome: FAO.
- FAO. 1991. "Digital Soil Map of the World, VI Africa." http://www.fao.org/fileadmin/user_upload/soils/docs/Soil_map_FAOUNESCO/acrobat/Africa_VI.pdf
- FAO. 2001. "Lecture Notes on the Major Soil." <http://www.fao.org/3/a-y1899e.pdf>
- FAO. 2006. *Guidelines for Soil Description*. 4th ed. Rome: FAO and UN.
- FAO-UNESCO. 1977. "Soil Map of the World, 1:5000000." *Volume VI Africa*. Paris: Unesco.
- Fick, S. E., and R. J. Hijmans. 2017. "Worldclim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37: 4302–4315. doi:10.1002/joc.5086.
- Gee, G. W., and J. W. Bauder. 1986. "Particle-Size Analysis." In *Methods of Soil Analysis: Part 1. Physical and Mineralogical Methods*, edited by A. Klute, 383–411. Madison, WI: ASA-SSSA publisher.
- IUSS Working Group WRB. 2006. "World Reference Base for Soil Resources 2006." *World Soil Resources Reports No. 103*. Rome: FAO.
- IUSS Working Group WRB. 2015. "World Reference Base for Soil Resources 2014, Update 2015, International Soil Classification System for Naming Soils and Creating Legends for Soil Maps." *World Soil Resources Reports No. 106*. Rome: FAO.
- JAXA. 2018. "ALOS Global Digital Surface Model "ALOS World 3D - 30m (AW3D30)" Version 2.1." Accessed 8 January 2020. <https://www.eorc.jaxa.jp/ALOS/en/aw3d30/index.htm>
- 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. doi:10.1016/j.geoderma.2009.04.023.
- Latham, M. 1981. "French Soil Classifications and Their Application in the South Pacific Islands." In *Proceedings of the South Pacific Regional Forum on Soil Taxonomy*, edited by R. J. Morrison and D. M. Leslie, 185–199. Suva, Fiji: Institute of Natural Resources, The university of the South Pacific.
- Leenaars, J. G. B., A. J. M. van Oostrum, and M. Ruiperez Gonzalez. 2014. "Africa Soil Profiles Database, Version 1.2. A Compilation of Georeferenced and Standardised Legacy Soil Profile Data for Sub-Saharan Africa (With Dataset)." *ISRIC Report 2014/01*. Wageningen: Africa Soil Information Service (AFSIS) project and ISRIC - World Soil Information.
- Mietton, M., Y. Gunnell, G. Nicoud, L. Ferry, R. Razafimahefa, and P. Grandjean. 2018. "'Lake' Alaotra, Madagascar: A Late Quaternary Wetland Regulated by the Tectonic Regime." *Catena* 165: 22–41. doi:10.1016/j.catena.2018.01.021.
- Nishigaki, T., S. Sugihara, K. Kobayashi, Y. Hashimoto, M. Kilasara, H. Tanaka, T. Watanabe, and S. Funakawa. 2018. "Fractionation of Phosphorus in Soils with Different Geological and Soil Physicochemical Properties in Southern Tanzania." *Soil Science and Plant Nutrition* 64: 291–299. doi:10.1080/00380768.2018.1436406.
- Obara, H., T. Ohkura, Y. Takata, K. Kohyama, Y. Maejima, and T. Hamazaki. 2011. "Comprehensive Soil Classification System of Japan-First Approximation." *Bulletin of National Institute for Agro-Environmental Sciences* 29: 1–73.
- Riquier, J. 1968. "Carte pédologique de Madagascar, 1/1000000." Tananarive: ORSTOM.
- Roederer, P., and F. Bourgeat. 1969. "Planche 16 Pedologie." *Atlas de Madagascar*. Tananarive: ORSTOM.
- Rufer, D., F. Preusser, G. Schreurs, E. Gnos, and A. Berger. 2014. "Late Quaternary History of the Vakinankaratra Volcanic Field (Central Madagascar): Insights from Luminescence Dating of Phreatomagmatic Eruption Deposits." *Bulletin of Volcanology* 76: 817. doi:10.1007/s00445-014-0817-7.
- Schlüter, T. 2008. *Geological Atlas of Africa. With Notes on Stratigraphy, Tectonics, Economic Geology, Geohazards, Geosites and Geoscientific Education of Each Country*. 2nd four-coloured revised and enlarged edition. Berlin Heidelberg: Springer Verlag.
- Soil Survey Staff. 2014a. "Kellogg Soil Survey Laboratory Methods Manual." *Soil Survey Investigations Report No. 42, Version 5.0*, Edited by Burt R. and Soil Survey Staff, U.S. Department of Agriculture, Natural Resources Conservation Service.
- Soil Survey Staff. 2014b. *Keys to Soil Taxonomy*. 12th ed. U.S. Department of Agriculture, Natural Resources Conservation Service, Washington, DC.
- U.S. Geological Survey EROS Data Center. 1997. "USGS 30 ARC-second Global Elevation Data, GTOPO30." *Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory*. doi:10.5065/A1Z4-EE71.
- van Reeuwijk, L. P. 2002. "Procedures for Soil Analysis." *Technical Paper, No. 9*. Wageningen: ISRIC. http://www.isric.org/sites/default/files/ISRIC_TechPap09.pdf
- Veenstra, J. J., and C. L. Burras. 2012. "Effects of Agriculture on the Classification of Black Soils in the Midwestern United States." *Canadian Journal of Soil Science* 92: 403–411. doi:10.4141/CJSS2010-018.

Letter

Vis-NIR Spectroscopy and PLS Regression with Waveband Selection for Estimating the Total C and N of Paddy Soils in Madagascar

Kensuke Kawamura ^{1,*} , Yasuhiro Tsujimoto ¹, Michel Rabenarivo ², Hidetoshi Asai ¹, Andry Andriamananjara ² and Tovohera Rakotoson ²

¹ Japan International Research Center for Agricultural Sciences (JIRCAS), 1-1 Ohwashi, Tsukuba, Ibaraki 305-8686, Japan; tsjmt@affrc.go.jp (Y.T.); asai0817@affrc.go.jp (H.A.)

² Laboratoire des Radio-Isotopes, Université d'Antananarivo, BP 3383, Route d'Andraisoro, 101 Antananarivo, Madagascar; miarabenarivo@yahoo.fr (M.R.); njaraandry1@gmail.com (A.A.); tovohera.r@gmail.com (T.R.)

* Correspondence: kamuken@affrc.go.jp; Tel.: +81-29-838-6614

Received: 20 September 2017; Accepted: 20 October 2017; Published: 23 October 2017

Abstract: Visible and near-infrared (Vis-NIR) diffuse reflectance spectroscopy with partial least squares (PLS) regression is a quick, cost-effective, and promising technology for predicting soil properties. The advantage of PLS regression is that all available wavebands can be incorporated in the model, while earlier studies indicate that PLS models include redundant wavelengths, and selecting specific wavebands can refine PLS analyses. This study evaluated the performance of PLS regression with waveband selection using Vis-NIR reflectance spectra to estimate the total carbon (TC) and total nitrogen (TN) in soils collected mainly from the surface of upland and lowland rice fields in Madagascar ($n = 59$; after outliers were removed). We used iterative stepwise elimination-based PLS (ISE-PLS) to estimate soil TC and TN and compared the predictive ability with standard full-spectrum PLS (FS-PLS). The predictive abilities were assessed using the coefficient of determination (R^2), the root mean squared error of cross-validation (RMSECV), and the residual predictive deviation (RPD). Overall, ISE-PLS using first derivative reflectance (FDR) showed a better predictive accuracy than ISE-PLS for both TC ($R^2 = 0.972$, RMSECV = 0.194, RPD = 5.995) and TN ($R^2 = 0.949$, RMSECV = 0.019, RPD = 4.416) in the soil of Madagascar. The important wavebands for estimating TC (12.59% of all wavebands) and TN (3.55% of all wavebands) were selected from all 2001 wavebands over the 400–2400 nm range using ISE-PLS. These findings suggest that ISE-PLS based on Vis-NIR diffuse reflectance spectra can be used to estimate soil TC and TN contents in Madagascar with an improved predictive accuracy.

Keywords: Acrisols; calibration; Ferralsols; first derivative reflectance; Oxisols; partial least squares regression; spectral assessments; surface paddy soil

1. Introduction

Carbon (C) and nitrogen (N) contents in soils are two key parameters for sustaining soil and environmental quality, as well as for improving crop productivity because of their involvement in a number of natural processes related to soil health and fertility [1]. Moreover, monitoring C levels in soils is increasingly needed because the depleted C levels, particularly in croplands, present an opportunity for carbon sequestration through adequate management practices [2]. To efficiently manage C and N in soils, a large number of soil samples must be evaluated for soil spatial variability [3]. However, standard procedures for assessing the state of C and N in soils are costly and time consuming [4,5] and require experienced operators. Thus, possible alternatives such as visible (Vis, 400–700 nm) and near-infrared (NIR, 700–2500 nm) spectroscopy are gaining attention; both of these alternatives

have been widely accepted as fast and non-destructive methods for estimating soil properties [6,7]. These techniques measure the radiation absorbed by various bonds of O-H, C-H, N-H, C=O, C-N, N-H, or C=C, resulting in bending, twisting, stretching, or scissoring [8,9]. Diffusely reflected NIR radiation is then correlated to measure material properties using various multivariate calibration techniques [10]. Among linear multivariate analyses, partial least squares (PLS) regression is the most commonly used approach for soil spectral analyses. Using PLS regression analyses, many calibrations have been conducted in recent decades to predict soil properties from Vis-NIR spectral data [11,12]. The infra-red PLS method of soil property predictions was shown to be well suited for the characterization of soils [13].

However, waveband selection can also refine the performance of PLS analysis not only for the prediction of soil properties [14,15], but also for other chemical and physical properties, such as forage in paddy fields [16], forest [17], and grassland [18,19], or for water quality in irrigation ponds [20], food [21], and fuel [22]. The PLS regression method combines the most useful information from hundreds of wavebands into the first several PLS factors (or latent variables), whereas the less important factors might include background effects [17,23]. Thus, many approaches for selecting wavebands or wavelength regions have been developed to eliminate useless (or to select useful) wavebands/wavelength regions in PLS analyses; these approaches include iterative stepwise elimination PLS (ISE-PLS) [24], uninformative variable elimination PLS (UVE-PLS) [25], competitive adaptive reweighted sampling (CARS) [26], interval PLS (iPLS) [27], moving window PLS (MW-PLS) [28], and genetic algorithm PLS (GA-PLS) [29]. Much of the literature has reported that more accurate calibration models may be achieved by selecting the most informative spectral variables instead of using the standard full-spectrum PLS (FS-PLS). In addition, waveband selection attempts to reduce the complexity and thus improve the robustness of a calibration model [23,30,31]. For example, Kawamura et al. [23] reported that removal of the redundant wavebands by ISE-PLS greatly improved the estimation accuracy of herbage mass and forage chemical properties in pasture. The results also suggested that ISE-PLS has the advantage of tuning the optimum bands for PLS regression with a better predictive ability in pastures, although this method has not been applied to soil spectra and soil properties.

In Madagascar, rice is important not only as the country's staple food, but also as the major rural income-generating resource. However, rice yield has been stagnant at less than 3 t ha⁻¹ in recent decades despite relatively favorable water conditions, with 70% of rice-cropping areas categorized as irrigated in this country [32]. In a survey of several rice fields in Madagascar's central highland, Tsujimoto et al. [33] showed a significant and linear response of rice yield against the soil organic carbon (SOC) content in relation to the N-supplying capacity of soils, which strongly indicates the importance of soil fertility management for increasing regional rice yields. Extensive research on SOC has been conducted using standard procedures, but most studies have focused on forest carbon stocks in the context of carbon dynamics, global warming, and environmental degradation in Madagascar [34–38]. Extensive and field-based soil C and N evaluations concerning the development of appropriate soil and nutrient management recommendations for the rice-cropping system, the country's major land use, are limited.

The aim of this study was to evaluate whether waveband selection by ISE-PLS would improve the predictive ability of calibrations using laboratory Vis-NIR spectroscopy when predicting soil total C (TC) and total N (TN) contents in Madagascar. The study compares the performance of ISE-PLS with FS-PLS using a set of 59 soil samples collected from upland and lowland rice fields in the central highland of Madagascar.

2. Materials and Methods

2.1. Study Site and Soil Sampling and Chemical Analyses

The field survey was conducted in the central highland of Madagascar (Figure 1). This region belongs to a subtropical climate with an altitude of 1000–1500 m above sea level. The mean temperature is 14–17 °C in winter and 20–22 °C in summer. The average annual rainfall is 1100 mm (>80% occurs in November–March) [33]. The area is dominated by inherently nutrient-poor soil types that are mainly classified into Ferralsols and Acrisols [39] or into Oxisols of semiarid to humid climates [40].

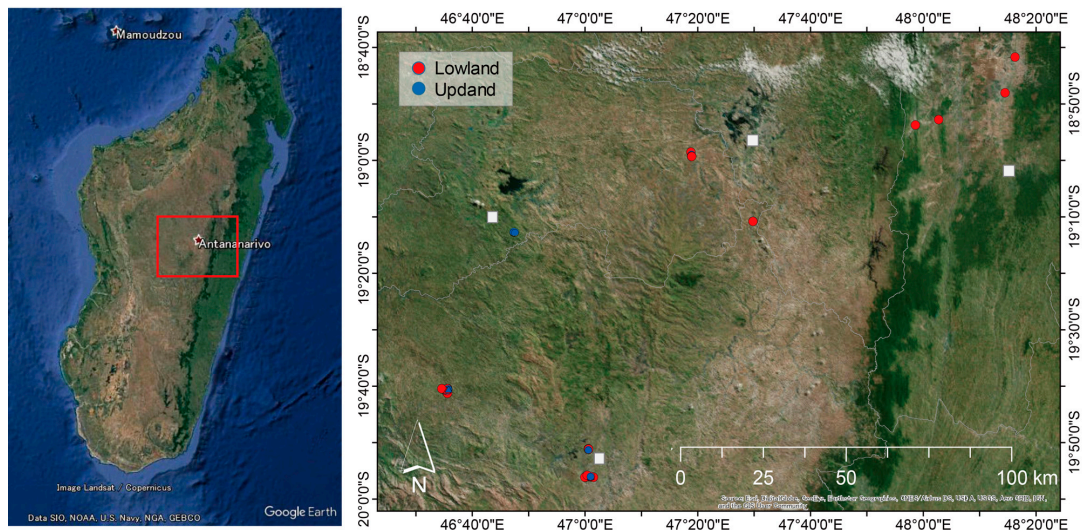


Figure 1. Locations of studied regions and soil sampling points.

Soil sampling was conducted in 55 rice fields from August to November in 2016, consisting of eight upland and 47 lowland fields under various cropping systems (Figure 1). The sampling positions were recorded with a handy GPS (Colorado300, Garmin, Ltd., Kansas, TX, USA). Surface soil samples were collected from a 0–10 cm depth as composites of three to four cores in each field. Within three fields, sub-surface samples (10–20 cm depth in a field; 10–20, 20–30, and 30–40 cm depth in two fields) were also collected. Thus, 62 soil samples were obtained.

2.2. Soil Chemical Analyses

In the laboratory, soil samples were sieved to <2 mm and air dried for seven days. Earlier studies compared the effect of samples sieved to 2 mm and ground to 200 μ m and did not obtain highly significant differences with respect to accuracy [41]. Thus, we worked with 2 mm crushed and sieved soil samples (0.6 g) in this study.

The TC and TN contents of soils were determined using an automatic NC analyzer, the SUMIGRAPH NC-220F (Sumika Chemical Analysis Service, Ltd., Osaka, Japan).

2.3. Vis-NIR Diffuse Reflectance Measurement

Laboratory soil reflectance measurements were conducted in a dark room at the Graduate School of Agriculture, Kyoto University, Japan, on 12–13 December 2016, using a portable spectro-radiometer (ASD FieldSpec 4 Hi-Res, ASD Inc., Longmont, CO, USA) and an ASD contact-probe (Figure 2). The ASD FieldSpec measures spectral reflectance in the 350–2500 nm wavelength region with spectral sampling of 1.4 nm in the 350–1000 nm range and 2 nm in the 1000–2500 nm range. The spectral resolution (full-width-half-maximum; FWHM) was 3 nm in the 350–1000 nm range and 6 nm in the 1000–2500 nm range, which were calculated to 1 nm resolution wavelengths for output data using the cubic spline interpolation function in ASD software (RS3 for Windows; ASD). The contact probe

light source (halogen lamp) was aligned at 12° to the probe body, ensuring illumination at a fixed angle without the influence of ambient light. The fiber optic cable of the ASD FieldSpec was attached to the contact probe at a fixed measurement angle of 35°. The sensed spot area had a diameter of ~1.1 cm with a field of view of 1.33 cm². A Spectralon (Labsphere, Inc., Sutton, NH, USA) reference panel (white reference) was used to optimize the ASD instrument prior to taking Vis-NIR reflectance measurements for each sample.

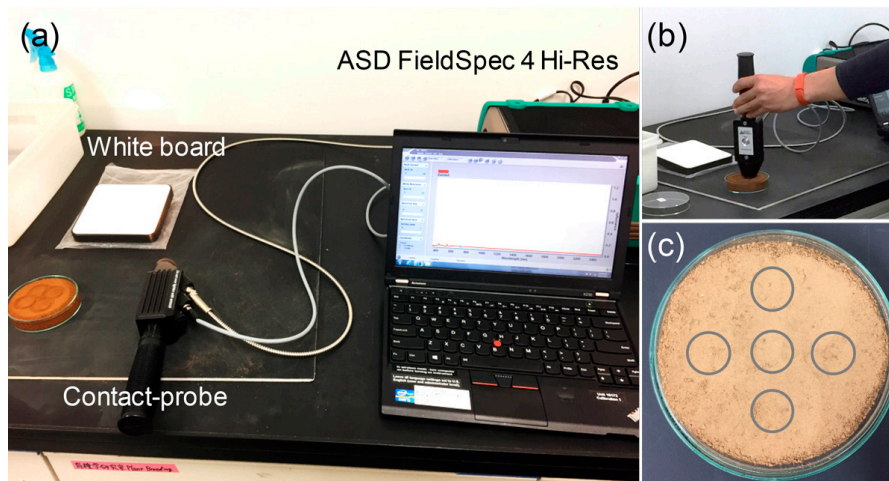


Figure 2. (a) The setup used to measure the soil reflectance in a dark room; (b) the use of a contact probe that touches the surface of the soil sample; and (c) the five measuring spots on a soil sample.

Bulk soil samples were spread in optical-glass Petri dishes 85 mm in diameter and pressed to form a layer ~19 mm thick. The soil surfaces were scanned 25 times with five replications for the soil samples (see Figure 2c), and the spectral readings were averaged.

2.4. Preprocessing of Spectral Data

Spectral data in both edge wavelength regions (350–399 nm and 2401–2500 nm) were eliminated because of low signal-to-noise ratios in the instrument. Thus, a total of 2001 spectral bands between 400 nm and 2400 nm were used for analyses.

First derivative reflectance (FDR) spectra were used to reduce baseline variation and enhance spectral features [42]. The FDR was calculated using the Savitzky-Golay smoothing filter [43]. A third-order, 15-band moving polynomial was fitted according to the original reflectance signatures. The parameters of this polynomial were subsequently used to calculate the derivative at the center waveband of the moving spline window. In addition, a standard normal variate transform (SNV) was employed to reduce the particle size effect [41].

To detect outliers, a principal component analysis was performed on spectral data for calculating the Mahalanobis distance H , and samples with $H > 3$ were eliminated as outliers. As a result, three samples were considered outliers, leaving 59 samples for further analyses.

2.5. Standard Full-Spectrum Partial Least Squares (FS-PLS) Regression

PLS regression analyses were performed to estimate soil parameters using reflectance and FDR datasets ($n = 59$). The standard FS-PLS regression equation is as follows:

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon \quad (1)$$

where the response variable y is a vector of the soil parameters (TN and TC); the predictor variables x_1 to x_i are the surface reflectance or FDR values for spectral bands 1 to i (400, 401, ..., 2400 nm),

respectively; β_1 to β_i are the estimated weighted regression coefficients; and ε is the error vector. The latent variables were introduced to simplify the relationship between the response variables and predictor variables. To determine the optimal number of latent variables (NLV), leave-one-out (LOO) cross-validation was performed to avoid over-fitting of the model, which was based on the minimum value of the root mean squared error of cross-validation (RMSECV) (see in Supplementary Materials: Figure S1). The RMSECV was calculated as follows:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^n (y_i - y_p)^2}{n}} \quad (2)$$

where y_i and y_p represent the measured and predicted soil parameters for sample i , respectively, and n is the number of samples in the data sets ($n = 59$).

2.6. Iterative Stepwise Elimination Partial Least Squares (ISE-PLS) Regression

ISE-PLS is a PLS model that incorporates a waveband elimination algorithm. The ISE method eliminates noisy variables and selects useful predictors. When PLS models include large numbers of redundant variables or outliers, the models' predictive abilities may perform poorly, while the ISE method can overcome such problems. Performance depends on the importance of predictors (z_i), described as follows:

$$z_i = \frac{|\beta_i|s_i}{\sum_{i=1}^I |\beta_i|s_i} \quad (3)$$

where s_i is the standard deviation and β_i is the regression coefficient; both s_i and β_i correspond to the predictor variable of the waveband i .

Initially, all available wavebands (2001 bands, 400–2400 nm) are used to develop the PLS regression model. Then, to create a scope in which useless predictor variables are removed and the predictive ability is improved, each predictor z_i is evaluated, and the minimum values are eliminated as less informative wavebands. Subsequently, the PLS model is re-calibrated with the remaining predictors [44]. The model-building procedure is repeated until the final model is calibrated with the maximum predictive ability.

2.7. Predictive Ability of the PLS Models

The predictive abilities of the FS-PLS and ISE-PLS models were assessed by calculating the coefficient of determination (R^2), RMSECV, and the residual predictive deviation (RPD) using LOO cross-validation. High R^2 and low RMSECV values indicate the best model for predicting the soil parameters. The RPD has been defined as the ratio of standard deviation (SD) of reference data for predicting RMSECV [45]. For the performance ability of calibration models, RPD was suggested to be at least 3 for agriculture applications, while RPD values between 2 and 3 indicate a model with a good prediction ability, $1.5 < \text{RPD} < 2$ is an intermediate model needing some improvement, and an $\text{RPD} < 1.5$ indicates that the model has a poor prediction ability [13].

To determine the significant wavelengths used in FS-PLS calibrations, the variable importance in the projection (VIP) [46,47] was used and referred to the selected wavelength regions from ISE-PLS models. The VIP score gives a summary of the importance of an x -variable (waveband) for an observed y -variable and is calculated using the following equation:

$$\text{VIP}_k(a) = m \sum_a W_{ak}^2 \left(\frac{\text{SSY}_a}{\text{SSY}_t} \right) \quad (4)$$

where $\text{VIP}_k(a)$ is the importance of the k th predictor variable based on a model with a factors, W_{ak} is the corresponding loading weight of the k th variable in the a th PLS regression factor, SSY_a is the explained sum of squares of y obtained from a PLS regression model with a factors, SSY_t is the total

sum of squares of y , and m is the total number of predictor variables. A high VIP score indicates an important x -variable (waveband) [46,48].

All the data handling and linear regression analyses were performed using MATLAB software ver. 9.0 (MathWorks, Sherborn, MA, USA).

3. Results and Discussion

3.1. Soil Properties (TC and TN) and Their Correlations with Each Waveband

Table 1 shows the descriptive analysis for soil TC and TN in the 59 samples. The mean (and SD) values of TC and TN were 2.18% ($\pm 1.16\%$) and 0.17% ($\pm 0.08\%$), respectively. The soil samples yielded a wide range of TC (coefficients of variation [CV] = 53.35) and TN values (CV = 48.08). The SD and range of sample affect the accuracy of soil property predictions using Vis-NIR spectroscopy [11]. In the present study, the ranges in soil TC and TN were considered sufficiently large to develop the calibration models using PLS regression analyses.

Table 1. Descriptive statistics of soil sample data.

Soil Parameters	n	Min	Max	Mean	SD	CV
TC (%)	59	0.65	6.02	2.18	1.16	53.35
TN (%)	59	0.06	0.44	0.17	0.08	48.08

n , number of samples; SD, standard deviation; CV, coefficient of variation ($=\text{Mean}/\text{SD} \times 100\%$).

A significant correlation coefficient ($r = 0.977$, $p < 0.001$) was found between TC and TN in the soil samples. The results revealed that the soil TC and TN showed a similar shape of correlation using Vis-NIR reflectance and FDR spectra (see in Supplementary Materials: Figure S2). In the reflectance data, reflectance values at 1413 and 2207 nm were highly correlated with the soil TC and TN contents. A peak of negative correlation at 598 nm was also obtained in the Vis wavelength region. In a previous study [49], soil reflectance in the NIR wavelength region was characterized by well-defined absorption features associated with overtones of O-H and H-O-H stretch vibrations in free water (1455 and 1915 nm) and overtones and combinations of O-H stretch and metal-OH bends in a clay lattice (1415 and 2207 nm).

3.2. Comparison between FS-PLS and ISE-PLS Models

Figure 3 shows changes in the RMSECV and R^2 values with iterative stepwise elimination procedures of redundant wavebands in the prediction of TC and TN using FDR. The RMSECV decreased as wavebands were removed but increased rapidly after more than 1749 and 1930 wavebands had been removed for TC and TN, respectively. Similarly, the R^2 value tended to increase slowly until the maximum value was obtained when 1749 and 1930 wavebands had been removed. The remaining 252 ($=2001 - 1749$) and 71 ($=2001 - 1930$) wavebands were considered useful wavelengths for estimating TC and TN, respectively. The selected number of wavebands (NW) and the selected NW as a percentage of the full spectrum ($\text{NW}\% = \text{NW}/\text{whole waveband}$ [$N = 2001$]) are presented in Table 2, with the values of NLV, R^2 , RMSEC/CV, and RPD from the FS-PLS and ISE-PLS models using the FDR dataset. The optimum NLV ranged between 7 and 15, determined as the lowest RMSECV values calculated from LOO cross-validation to avoid over-fitting of the model.

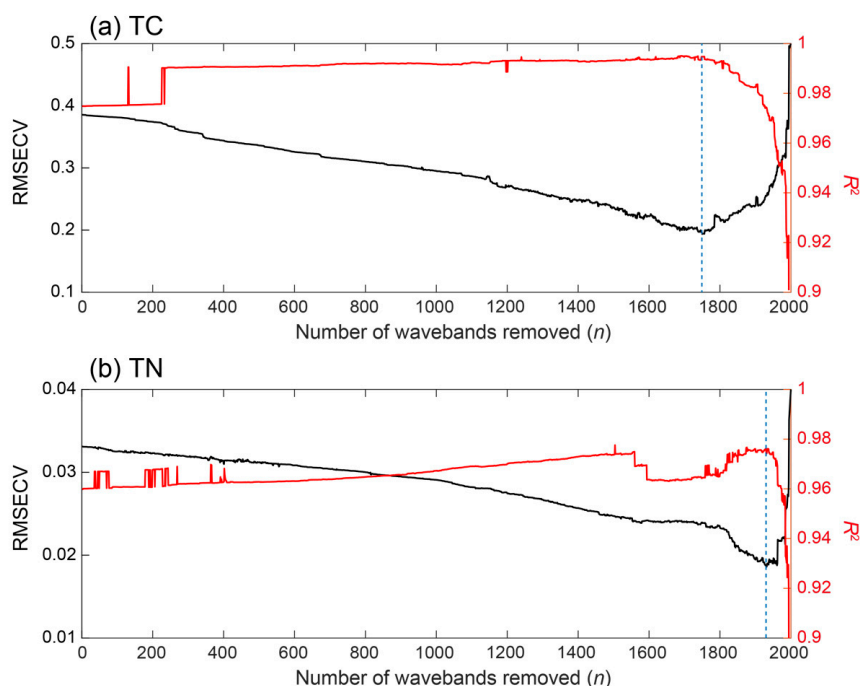


Figure 3. Changes in RMSECV (black line) and R^2 values (red line) in models to estimate total carbon (TC) (a) and total nitrogen (TN) (b) with the stepwise removal of redundant wavebands. The minimum value of the root mean squared error of cross-validation (RMSECV) (blue dotted line) was obtained when 1749 and 1930 wavebands were removed for TC and TN, respectively.

Table 2. Optimum number of latent variables (NLV), coefficient of determination (R^2), root mean squared errors of calibration (RMSEC) and cross-validation (RMSECV), and residual predictive values (RPD) from full-spectrum PLS (FS-PLS) and iterative stepwise elimination PLS (ISE-PLS) models with a selected number of wavebands (NW) and their percentages of the full spectrum (NW%).

Soil Parameter	Regression Method	Calibration			Cross-validation			NW	NW%
		NLV	R^2	RMSEC	R^2	RMSECV	RPD		
Total carbon (TC, %)	FS-PLS	14	0.996	0.076	0.893	0.379	3.064	252	12.59
	ISE-PLS	12	0.995	0.084	0.972	0.194	5.995		
Total nitrogen (TN, %)	FS-PLS	9	0.960	0.016	0.837	0.033	2.480	71	3.55
	ISE-PLS	7	0.974	0.013	0.949	0.019	4.416		

FS-PLS, full-spectrum partial least squares; ISE-PLS, iterative stepwise elimination PLS; NLV, number of latent variables; RMSEC (or RMSECV), root mean squared error of calibration (or cross-validation); NW, number of wavebands; NW%, number of waveband percentages of all available bands ($=NW/2001 \text{ bands} \times 100\%$).

Considering the difference in model accuracies between the FS-PLS and ISE-PLS (Table 2), better predictive accuracies were obtained in ISE-PLS than FS-PLS for both soil TC ($R^2 = 0.972$, RMSECV = 0.194) and TN ($R^2 = 0.949$, RMSECV = 0.019), with RPDs of 5.995 and 4.416, respectively. Figure 4 shows the relationships between the observed and cross-validated predicted values of soil TC and TN from ISE-PLS using FDR data. These results indicate that the soil TC and TN can be rapidly and accurately predicted from Vis-NIR diffuse reflectance spectroscopy using PLS regression. Selecting a subset of wavebands related to soil chemical properties and removing unrelated wavebands further improved the PLS regression results. Moreover, based on $RPD > 3$, the quality and future applicability of our results could be considered to have an excellent predictive ability. The remaining NW (NW%) of TC and TN was 252 (12.59%) and 71 (3.55%), respectively, suggesting that over 87% of the waveband information from the soil reflectance spectrum was redundant and did not contribute to or disturb the prediction of soil TC and TN.

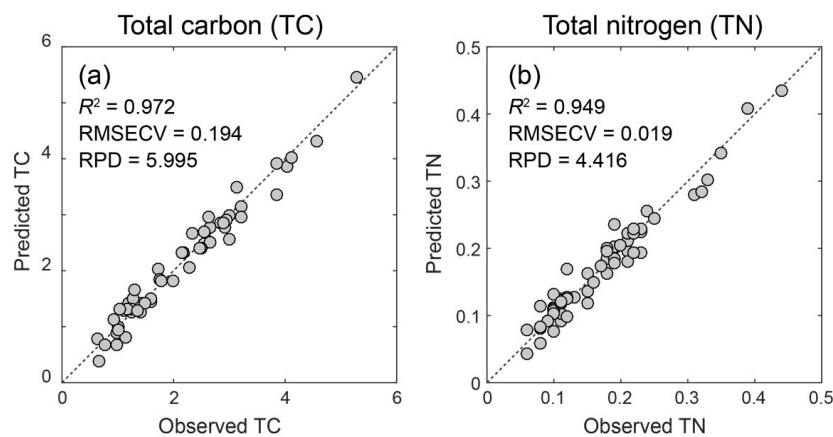


Figure 4. Observed and predicted values of soil total carbon (TC) and soil total nitrogen (TN) contents using ISE-PLS models with first derivative reflectance (FDR) data ($n = 59$). The coefficient of determination (R^2), root mean squared error of cross-validation (RMSECV), and residual predicted value (RPD) are cross-validated (leave-one-out cross-validation method) coefficient of determination, root mean squared error, and residual predictive values, respectively (see Table 2).

These results agree with previous results indicating that the most useful information in the Vis-NIR region (400–2400 nm) was less than 20% for predicting forage [18,19] and water parameters [20]. These findings also support previous results showing that the performance of PLS models can be improved through waveband selection. Yang et al. [14] suggested that reducing large spectral datasets is valuable for more efficient storage, computation, and transmission, as well as for the ease of spectral analysis [50]. In addition, when fewer wavebands are used, simpler and cheaper spectro-radiometer processes can be developed.

3.3. Selected Wavebands from ISE-PLS Models

The selected wavebands from ISE-PLS using FDR spectra to estimate soil TC and TN are shown in Figure 5, with VIP score values from FS-PLS. Based on the VIP score (>1), the wavelengths centered near 418, 470, 760, 1408, 1912, 2255, 2314, and 2339 nm were identified as common important wavelengths for estimating soil TC and TN. Most of the VIP peak regions were selected in the final ISE-PLS models. Although they did not perfectly fit with previously known absorption wavelength regions, some of the wavelengths were revealed within 30 nm of known absorption features. For soil TC prediction, the final model included Vis wavelength regions (400–480 and 640–700 nm), which are associated with soil color and had a huge influence on model calibration. Soil becomes darker as soil organic matter (SOM) increases; thus, several researchers have tried to use soil color information to estimate SOM [9,51]. However, soil darkness is only a useful discriminator within limited geological variation. In general, soil reflectance decreases with increasing organic matter content [49] and water content [52]. Absorptions of approximately 400, 450, 510, 550, 700, 870, and 1000 nm are characterized by the presence of ferrous and ferric iron oxides and are due to the electronic transitions of the iron cations [53]. A spectral band of 2100–2500 nm contributes to the model calibration of C and N in soils [54].

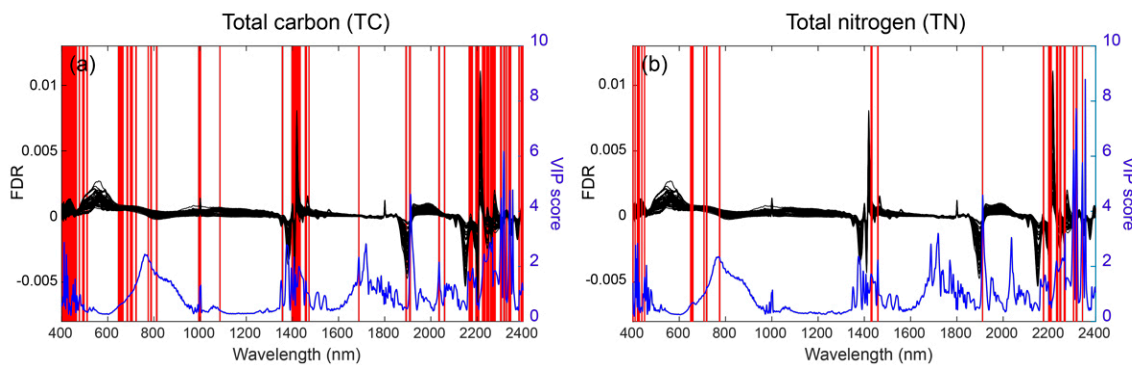


Figure 5. Soil reflectance and its first derivative reflectance (FDR) spectra for the total carbon (TC; **a**) and total nitrogen (TN; **b**) datasets and selected waveband (red bar) in iterative stepwise elimination of partial least squares (ISE-PLS) with variable importance in the prediction (VIP) score (blue line) from full-spectrum PLS (FSPLS) models.

Martin et al. [55] reported that the NIR spectroscopy-based prediction of TN may be indirect due to a close correlation with TC, and that the calibration accuracy is higher for TC than for TN. Chang and Laird [56] confirmed that the NIR spectroscopy determination of TN does not always rely on a strong correlation with TC and can determine TN directly. Brunet et al. [41] hypothesized that, depending on the studied dataset, TN can be predicted based on its correlation with TC when the correlation is high; otherwise, it can be predicted directly. In our result, soil TC data showed a high correlation with soil TN data ($r = 0.977$), and calibrations obtained a better predictive accuracy for TC ($R^2 = 0.972$, RMSECV = 0.194) than for TN ($R^2 = 0.949$, RMSECV = 0.019). Within the selected wavebands of soil TN (Figure 5), 90.1% ($=64/71 \text{ bands} \times 100\%$) overlapped with the selected wavebands of soil TC, whereas different wavebands in TC calibration were revealed mainly in the NIR region (707, 717–719, 774 nm). These results indicated that TN prediction using our dataset was affected by strong correlations with TC data but might be directly estimated.

Lastly, we note that this study was carried out on heterogeneous sample data sets, which were collected at upland and lowland soils under various rice-based cropping systems, including wide ranges of soil types in Madagascar. However, several researchers consider the reliability of the prediction questionable when studying heterogeneous sample sets [41]. Particle size and arrangement might also affect the calibration due to the light transmission path [57]. Moreover, to map the carbon stock at a larger spatial scale in Madagascar, evaluating an appropriate spatial scale with a larger data set is required [58]. In future study, thus, more information concerning the effect of a heterogeneous data set on the accuracy of NIRS predictions at different scales is needed in order to apply the methodology to soil characterization of the whole island of Madagascar.

4. Conclusions

We investigated the performance of waveband selection in the spectral estimation of soil TC and TN using Vis-NIR reflectance data. The results indicated that soil TC and TN in Madagascar can be more accurately estimated by ISE-PLS than by standard FS-PLS using laboratory Vis-NIR spectroscopy. ISE-based wavelength selection in PLS calibration suggested that the important wavebands for estimating soil TC and TN were, respectively, 12.59% and 3.55% of all 2001 wavebands in the 400–2400 nm range. Based on selected FDR wavelengths in the ISE-PLS model, soil TC and TN were determined to provide excellent predictions ($RPD > 3$), with 0.194% and 0.019% error, respectively. The use of PLS with ISE waveband selection in Vis-NIR reflectance spectra is promising for the spectral assessment of soil TC and TN in Madagascar. Furthermore, the waveband selection procedure refined the predictive ability expected by optimizing the wavelength subset using ISE-PLS.

Such timely and accurate soil TC and TN predictions might efficiently provide useful insights into fertilizer management.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/9/10/1081/s1, Figure S1: Changes in RMSE (grey circle/line) and RMSECV (black circle/line) based on the number of latent variables (NLV) in models to estimate soil total carbon (TC) (a,c) and total nitrogen (TN) (b,d) using FS-PLS and ISE-PLS regressions. The optimal NLV (red vertical line) was determined the minimum value of RMSECV, Figure S2: Correlation coefficients (r) between soil chemical parameters (total carbon (C) and total nitrogen (TN)) at each wavelength: (a) reflectance and (b) first derivative reflectance (FDR).

Acknowledgments: This research was supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA). We would like to give our special thanks to Naoki Moritsuka, Graduate School of Agriculture, Kyoto University in Japan, for his support in soil spectral measurement and for valuable comments on this manuscript.

Author Contributions: Kensuke Kawamura, Yasuhiro Tsujimoto, and Tovohery Rakotoson designed this study and the field work; Yasuhiro Tsujimoto, Michel Rabenarivo, Hidetoshi Asai, and Andry Andriamananjara performed the fieldwork and carried out the soil chemical analyses; Kensuke Kawamura performed laboratory spectral measurements and the data processing, and wrote the manuscript; and all the authors revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Weil, R.; Magdoff, F. Significance of soil organic matter to soil quality and health. In *Soil Organic Matter in Sustainable Agriculture*; Magdoff, E., Weil, R.R., Eds.; CRC Press: Boca Raton, FL, USA, 2004; p. 412.
2. Lal, R. Soil carbon sequestration to mitigate climate change. *Geoderma* **2004**, *123*, 1–22. [[CrossRef](#)]
3. Mouazen, A.M.; Maleki, M.R.; De Baerdemaeker, J.; Ramon, H. On-line measurement of some selected soil properties using a vis–nir sensor. *Soil Tillage Res.* **2007**, *93*, 13–27. [[CrossRef](#)]
4. Conant, R.T.; Ogle, S.M.; Paul, E.A.; Paustian, K. Measuring and monitoring soil organic carbon stocks in agricultural lands for climate mitigation. *Front. Ecol. Environ.* **2011**, *9*, 169–173. [[CrossRef](#)]
5. Sinfield, J.V.; Fagerman, D.; Colic, O. Evaluation of sensing technologies for on-the-go detection of macro-nutrients in cultivated soils. *Comput. Electron. Agric.* **2010**, *70*, 1–18. [[CrossRef](#)]
6. Conforti, M.; Castrignanò, A.; Robustelli, G.; Scarciglia, F.; Stelluti, M.; Buttafuoco, G. Laboratory-based vis–NIR spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content. *Catena* **2015**, *124*, 60–67. [[CrossRef](#)]
7. Islam, K.; Singh, B.; McBratney, A. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Soil Res.* **2003**, *41*, 1101–1114. [[CrossRef](#)]
8. Miller, C.E. Chemical principles of near-infrared technology. In *Near Infrared Technology in the Agricultural and Food Industries*, 2nd ed.; Williams, P.C., Horris, K.H., Eds.; American Association of Cereal Chemists: St. Paul, MN, USA, 2001; pp. 19–37.
9. Viscarra Rossel, R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [[CrossRef](#)]
10. Mouazen, A.M.; Kuang, B.; De Baerdemaeker, J.; Ramon, H. Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma* **2010**, *158*, 23–31. [[CrossRef](#)]
11. Kuang, B.; Mouazen, A.M. Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three european farms. *Eur. J. Soil Sci.* **2011**, *62*, 629–636. [[CrossRef](#)]
12. Fystro, G. The prediction of C and N content and their potential mineralisation in heterogeneous soil samples using vis–nir spectroscopy and comparative methods. *Plant Soil* **2002**, *246*, 139–149. [[CrossRef](#)]
13. D'Acqui, L.P.; Pucci, A.; Janik, L.J. Soil properties prediction of western mediterranean islands with similar climatic environments by means of mid-infrared diffuse reflectance spectroscopy. *Eur. J. Soil Sci.* **2010**, *61*, 865–876. [[CrossRef](#)]
14. Yang, H.; Kuang, B.; Mouazen, A.M. Quantitative analysis of soil nitrogen and carbon at a farm scale using visible and near infrared spectroscopy coupled with wavelength reduction. *Eur. J. Soil Sci.* **2012**, *63*, 410–420. [[CrossRef](#)]

15. Vohland, M.; Ludwig, M.; Thiele-Bruhn, S.; Ludwig, B. Determination of soil properties with visible to near- and mid-infrared spectroscopy: Effects of spectral variable selection. *Geoderma* **2014**, *223*, 88–96. [[CrossRef](#)]
16. Inoue, Y.; Sakaiya, E.; Zhu, Y.; Takahashi, W. Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sens. Environ.* **2012**, *126*, 210–221. [[CrossRef](#)]
17. Bolster, K.L.; Martin, M.E.; Aber, J.D. Determination of carbon fraction and nitrogen concentration in tree foliage by near infrared reflectance: A comparison of statistical methods. *Can. J. For. Res.* **1996**, *26*, 590–600. [[CrossRef](#)]
18. Kawamura, K.; Watanabe, N.; Sakanoue, S.; Lee, H.-J.; Inoue, Y.; Odagawa, S. Testing genetic algorithm as a tool to select relevant wavebands from field hyperspectral data for estimating pasture mass and quality in a mixed sown pasture using partial least squares regression. *Grassl. Sci.* **2010**, *56*, 205–216. [[CrossRef](#)]
19. Kawamura, K.; Watanabe, N.; Sakanoue, S.; Lee, H.-J.; Lim, J.; Yoshitoshi, R. Genetic algorithm-based partial least squares regression for estimating legume content in a grass-legume mixture using field hyperspectral measurements. *Grassl. Sci.* **2013**, *59*, 166–172. [[CrossRef](#)]
20. Wang, Z.; Kawamura, K.; Sakuno, Y.; Fan, X.; Gong, Z.; Lim, J. Retrieval of chlorophyll-a and total suspended solids using iterative stepwise elimination partial least squares (ISE-PLS) regression based on field hyperspectral measurements in irrigation ponds in higashihiroshima, Japan. *Remote Sens.* **2017**, *9*, 264. [[CrossRef](#)]
21. Fan, W.; Shan, Y.; Li, G.; Lv, H.; Li, H.; Liang, Y. Application of competitive adaptive reweighted sampling method to determine effective wavelengths for prediction of total acid of vinegar. *Food Anal. Meth.* **2012**, *5*, 585–590. [[CrossRef](#)]
22. Cramer, J.A.; Kramer, K.E.; Johnson, K.J.; Morris, R.E.; Rose-Pehrsson, S.L. Automated wavelength selection for spectroscopic fuel models by symmetrically contracting repeated unmoving window partial least squares. *Chemom. Intell. Lab. Syst.* **2008**, *92*, 13–21. [[CrossRef](#)]
23. Kawamura, K.; Watanabe, N.; Sakanoue, S.; Inoue, Y. Estimating forage biomass and quality in a mixed sown pasture based on pls regression with waveband selection. *Grassl. Sci.* **2008**, *54*, 131–146. [[CrossRef](#)]
24. Boggia, R.; Forina, M.; Fossa, P.; Mosti, L. Chemometric study and validation strategies in the structure-activity relationships of new cardiotoxic agents. *Quant. Struct.-Act. Relatsh.* **1997**, *16*, 201–213. [[CrossRef](#)]
25. Centner, V.; Massart, D.L.; de Noord, O.E.; de Jong, S.; Vandeginste, B.M.; Sterna, C. Elimination of uninformative variables for multivariate calibration. *Anal. Chem.* **1996**, *68*, 3851–3858. [[CrossRef](#)] [[PubMed](#)]
26. Li, H.; Liang, Y.; Xu, Q.; Cao, D. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal. Chim. Acta* **2009**, *648*, 77–84. [[CrossRef](#)] [[PubMed](#)]
27. Nørgaard, L.; Saudland, A.; Wagner, J.; Nielsen, J.P.; Munck, L.; Engelsen, S.B. Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Appl. Spectrosc.* **2000**, *54*, 413–419. [[CrossRef](#)]
28. Jiang, J.H.; Berry, R.J.; Siesler, H.W.; Ozaki, Y. Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data. *Anal. Chem.* **2002**, *74*, 3555–3565. [[CrossRef](#)] [[PubMed](#)]
29. Leardi, R. Application of a genetic algorithm to feature selection under full validation conditions and to outlier detection. *J. Chemom.* **1994**, *8*, 65–79. [[CrossRef](#)]
30. Yoshida, H.; Leardi, R.; Funatsu, K.; Varmuza, K. Feature selection by genetic algorithms for mass spectral classifiers. *Anal. Chim. Acta* **2001**, *446*, 483–492. [[CrossRef](#)]
31. Xiaobo, Z.; Jiewen, Z.; Povey, M.J.W.; Holmes, M.; Hanpin, M. Variables selection methods in near-infrared spectroscopy. *Anal. Chim. Acta* **2010**, *667*, 14–32. [[CrossRef](#)] [[PubMed](#)]
32. GriSP (Global Rice Science Partnership). *Rice Almanac*, 4th ed.; International Rice Research Institute: Los Banos, Philippines, 2013; p. 298.
33. Tsujimoto, Y.; Horie, T.; Randriamihary, H.; Shiraiwa, T.; Homma, K. Soil management: The key factors for higher productivity in the fields utilizing the system of rice intensification (SRI) in the central highland of madagascar. *Agric. Syst.* **2009**, *100*, 61–71. [[CrossRef](#)]
34. Grinand, C.; Maire, G.L.; Vieilledent, G.; Razakamanarivo, H.; Razafimbelo, T.; Bernoux, M. Estimating temporal changes in soil carbon stocks at ecoregional scale in madagascar using remote-sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *54*, 1–14. [[CrossRef](#)]

35. Ramifehiarivo, N.; Brossard, M.; Grinand, C.; Andriamananjara, A.; Razafimbelo, T.; Rasolohery, A.; Razafimahatratra, H.; Seyler, F.; Ranaivoson, N.; Rabenarivo, M.; et al. Mapping soil organic carbon on a national scale: Towards an improved and updated map of madagascar. *Geoderma Reg.* **2017**, *9*, 29–38. [[CrossRef](#)]
36. Razakamanarivo, R.H.; Grinand, C.; Razafindrakoto, M.A.; Bernoux, M.; Albrecht, A. Mapping organic carbon stocks in eucalyptus plantations of the Central Highlands of Madagascar: A multiple regression approach. *Geoderma* **2011**, *162*, 335–346. [[CrossRef](#)]
37. Andriamananjara, A.; Hewson, J.; Razakamanarivo, H.; Andrisoa, R.H.; Ranaivoson, N.; Ramboatiana, N.; Razafindrakoto, M.; Ramifehiarivo, N.; Razafimanantsoa, M.-P.; Rabeharisoa, L.; et al. Land cover impacts on aboveground and soil carbon stocks in malagasy rainforest. *Agric. Ecosyst. Environ.* **2016**, *233*, 1–15. [[CrossRef](#)]
38. Asner, G.P.; Mascaro, J.; Muller-Landau, H.C.; Vieilledent, G.; Vaudry, R.; Rasamoelina, M.; Hall, J.S.; van Breugel, M. A universal airborne lidar approach for tropical forest carbon mapping. *Oecologia* **2012**, *168*, 1147–1160. [[CrossRef](#)] [[PubMed](#)]
39. IUSS Working Group WRB. *World Reference Base for Soil Resources 2014, Update 2015 International Soil Classification System for Naming Soils and Creating Legends for Soil Maps*; World Soil Resources Reports No. 106; FAO: Rome, Italy, 2015.
40. Soil Survey Staff. *Keys to Soil Taxonomy*, 12th ed.; USDA-Natural Resources Conservation Service: Washington, DC, USA, 2014.
41. Brunet, D.; Barthès, B.G.; Chotte, J.-L.; Feller, C. Determination of carbon and nitrogen contents in alfisols, oxisols and ultisols from africa and brazil using nirs analysis: Effects of sample grinding and set heterogeneity. *Geoderma* **2007**, *139*, 106–117. [[CrossRef](#)]
42. Reeves, J.; McCarty, G.; Mimmo, T. The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environ. Pollut.* **2002**, *116*, S277–S284. [[CrossRef](#)]
43. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
44. Forina, M.; Lanteri, S.; Oliveros, M.C.C.; Millan, C.P. Selection of useful predictors in multivariate calibration. *Anal. Bioanal. Chem.* **2004**, *380*, 397–418. [[CrossRef](#)] [[PubMed](#)]
45. Williams, P.C. Implementation of near-infrared technology. In *Near-Infrared Technology in the Agricultural and Food Industries*, 2nd ed.; Williams, P.C., Norris, K.H., Eds.; American Association of Cereal Chemists Inc.: St. Paul, MN, USA, 2001; pp. 145–169.
46. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
47. Chong, I.-G.; Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112. [[CrossRef](#)]
48. Li, B.; Liew, O.W.; Asundi, A.K. Pre-visual detection of iron and phosphorus deficiency by transformed reflectance spectra. *J. Photochem. Photobiol. B* **2006**, *85*, 131–139. [[CrossRef](#)] [[PubMed](#)]
49. Ben-Dor, E.; Inbar, Y.; Chen, Y. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sens. Environ.* **1997**, *61*, 1–15. [[CrossRef](#)]
50. Viscarra Rossel, R.A.; Lark, R.M. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. *Eur. J. Soil Sci.* **2009**, *60*, 453–464. [[CrossRef](#)]
51. Viscarra Rossel, R.A.; Fouad, Y.; Walter, C. Using a digital camera to measure soil organic carbon and iron contents. *Biosyst. Eng.* **2008**, *100*, 149–159. [[CrossRef](#)]
52. Whiting, M.L.; Li, L.; Ustin, S.L. Predicting water content using gaussian model on soil spectra. *Remote Sens. Environ.* **2004**, *89*, 535–552. [[CrossRef](#)]
53. Ben Dor, E.; Irons, J.R.; Epema, J.F. Soil reflectance. In *Manual of Remote Sensing: Remote Sensing for the Earth Sciences*; John Wiley & Sons: New York, NY, USA, 1999; Volume 3, pp. 111–188.
54. Yang, H. Spectroscopic calibration for soil N and C measurement at a farm scale. *Proc. Environ. Sci.* **2011**, *10*, 672–677. [[CrossRef](#)]
55. Martin, P.D.; Malley, D.F.; Manning, G.; Fuller, L. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Can. J. Soil Sci.* **2002**, *82*, 413–422. [[CrossRef](#)]

56. Chang, C.-W.; Laird, D.A. Near-infrared reflectance spectroscopic analysis of soil C and N. *Soil Sci.* **2002**, *167*, 110–116. [[CrossRef](#)]
57. Chang, C.W.; Laird, D.; Mausbach, M.J.; Hurburgh, C.R.J. Nearinfrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [[CrossRef](#)]
58. Saiano, F.; Oddo, G.; Scalenghe, R.; La Mantia, T.; Ajmone-Marsan, F. DRIFTS sensor: Soil carbon validation at large scale (Pantelleria, Italy). *Sensors* **2013**, *13*, 5603–5613. [[CrossRef](#)] [[PubMed](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Exploring relevant wavelength regions for estimating soil total carbon contents of rice fields in Madagascar from Vis-NIR spectra with sequential application of backward interval PLS

Kensuke Kawamura^{a*}, Tomohiro Nishigaki^{a*}, Yasuhiro Tsujimoto^a, Andry Andriamananjara^b, Michel Rabenaribo^b, Hidetoshi Asai^a, Tovohery Rakotoson^b and Tantely Razafimbelo^b

^aJapan International Research Center for Agricultural Sciences (JIRCAS), Tsukuba, Japan; ^bLaboratoire des Radio-Isotopes, Université d'Antananarivo, Antananarivo, Madagascar

ABSTRACT

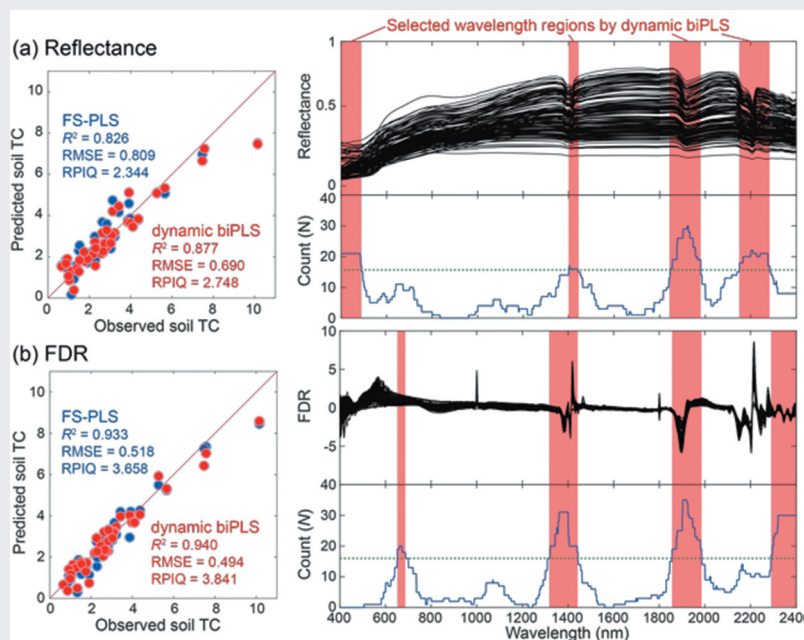
Laboratory visible and near-infrared (Vis-NIR) spectroscopy with partial least squares (PLS) regression can be used to determine the soil carbon (C) content, and the waveband selection procedures can refine the predictive ability. However, individually selected wavebands are not always the same depending on the location, scale, and approach. To simplify the variable selection issue, some methods for selecting wavelength regions instead of individual wavebands have been proposed. In this study, we explore relevant wavelength regions for predicting the total carbon (TC) content of lowland and upland soils in Madagascar from Vis-NIR spectroscopy using a dynamic version of backward interval PLS (biPLS) regression. The predictive ability of dynamic biPLS was compared with that of standard full-spectrum PLS (FS-PLS) using the cross-validated coefficient of determination (R^2), root mean squared error (RMSE), and ratio of performance to interquartile distance (RPIQ). The biPLS models using reflectance ($R^2 = 0.877$, RMSE = 0.690) and first derivative reflectance (FDR) ($R^2 = 0.940$, RMSE = 0.494) data sets showed better predictive accuracy than the FS-PLS models using reflectance ($R^2 = 0.826$, RMSE = 0.809) and FDR ($R^2 = 0.933$, RMSE = 0.518) data sets, the spectral efficiency was improved. By using biPLS to predict soil TC, the model was simplified by using only four selected wavelength regions in the reflectance (400–490, 1402–1440, 1846–1980 and 2151–2283 nm) and FDR (652–687, 1322–1443, 1856–1985, and 2290–2400 nm) data sets, which yielded reliable (RPIQ > 2.5) predictions.

ARTICLE HISTORY

Received 7 November 2019
Revised 15 April 2020
Accepted 8 June 2020

KEYWORDS

Madagascar; soil carbon; soil fertility; spectral assessments; wavelength region selection



CONTACT Kensuke Kawamura  kamuken@affrc.go.jp; Tomohiro Nishigaki  nishigaki@affrc.go.jp

*These authors contributed equally to this work.

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The soil carbon (C) content is one of the most important properties in assessments of general soil fertility. Then, timely assessments of soil C can be used for effective and sustainable fertilizer management practices, which is particularly true in Sub-Saharan Africa (SSA), where smallholder farmers still rely on indigenous nutrient supplies from soils for crop production. Rice cultivation in Madagascar is typical of farming systems in SSA in which smallholder farmers are impoverished by stagnant yields resulting from infertile soil conditions with minimal external inputs (Tsujimoto et al., 2019). In Madagascar, rice is uniquely important not only as the staple food of the country but also as the major income resource for rural livelihoods. However, rice yields have been stagnant at less than 3 t ha^{-1} for the last decades despite relatively favorable water conditions, with 70% of rice cropping areas categorized as irrigated in this country (Partnership, 2013). In a survey of several rice fields in the central highland of Madagascar, Tsujimoto et al. (2009) showed a significant and linear relationship between rice yield and the soil organic carbon (SOC) content in relation to the N-supplying capacity of soils, strongly indicating that soil fertility management is critical to improving rice yields in the region.

Visible and near-infrared (Vis-NIR) spectroscopy, as a rapid and non-destructive technology, has been widely used to perform quantitative analyses of complex samples in agricultural sciences. The reflectance spectra of Vis-NIR (400–2500 nm) obtained by laboratory spectral measurements include wavebands that have been related to physical and chemical properties of soil. The prediction of soil properties requires the development of a spectral library relating to spectra with reference data. To date, soil spectral libraries have been developed by the Vis-NIR enthusiasm around the world at country (Li et al., 2015; Romero et al., 2018), continental (Johnson et al., 2019; Stevens et al., 2013), and global (Viscarra Rossel et al., 2016) scales. These libraries can be used to develop calibration models for the prediction of soil properties. In Madagascar, however, only a small number of qualified data set was recorded in the soil libraries; continent scale ($n = 82$) (Johnson et al., 2019) and global scale ($n = 18$) (Viscarra Rossel et al., 2016).

Partial least squares (PLS) regression is the most widely used multivariate calibration method because it can extract information on the target component from a spectral matrix with hundreds or even thousands of wavebands (Conforti et al., 2013, 2015). However, as a linear multivariate calibration, the accuracy of PLS analysis tends to decrease due to the non-linear nature of the relationship between spectral data and the

dependent variable (Araújo et al., 2014). As data-mining approaches, machine learning techniques, such as artificial neural network (ANN) (Kuang et al., 2015), support-vector machines (SVM) (Morellos et al., 2016), and random forest (Cipullo et al., 2019; Douglas et al., 2018; De Santana et al., 2018) outperformed the PLS analysis for predicting soil properties as it is able to account for the non-linearity associated with the soil spectral responses. More recently, deep learning is a rapidly developing frontier in machine learning that has also been tested for calibrating soil spectra (Ng et al., 2019; Padarian et al., 2019). Although the literatures reported the machine learning outperformed PLS regression, they did not suggest that it would be suitable for everyone, because the machine learning, and especially deep learning, is a very data-hungry approach that requires a lot of data to be able to make a good prediction (Ng et al., 2019).

In Madagascar, there is no spectral library (data set) to perform machine learning for estimating soil properties. In the present study, we focused on the development of a robust PLS model using waveband selection approach based on local data set that we collected in the central highland of Madagascar. In general, Vis-NIR spectra contains thousands of wavebands, and such large number of spectra variables often contribute to collinearity, and redundancies rather than relevant information. Waveband selection is an important step not only for developing a robust calibration model and also for better understanding the relationship between spectral and soil properties. Indeed, it is increasingly evident that the inclusion of uninformative or redundant spectra in the Vis-NIR spectrum degrade the PLS model and lead to inaccurate predictions (Andersen & Bro, 2010; H. D. Li et al., 2012). Thus, additional waveband selection methods based on the PLS regression are necessary for NIR spectral analysis to refine the predictive ability.

To date, a large number of waveband selection methods for PLS analysis have been proposed, including individual waveband selection and wavelength region selection. For individual waveband selection, many methods have been developed, such as iterative stepwise elimination PLS (ISE-PLS) (Boggia et al., 1997), uninformative variable elimination PLS (UVE-PLS) (Centner et al., 1996), and genetic algorithm PLS (GA-PLS) (Leardi et al., 1992). Among the waveband selection methods, GA-PLS has been used as a suitable method in chemometrics (Leardi, 2000). Earlier studies reported that, after suitable modifications, GA-PLS shows better predictive performance and yields more interpretable results because the selected wavelengths are less dispersed than those with other methods (Kawamura et al.,

2010; Leardi & González, 1998; Lucasius & Kateman, 1994). In our previous research (Kawamura et al., 2017), we developed a PLS model to estimate the total carbon (TC) contents of paddy soils from laboratory Vis-NIR measurements of soil samples collected from various rice fields in the central highland of Madagascar. The results indicated improvements in the predictive ability by applying individual waveband selection with ISE-PLS. Additionally, our previous study (Kawamura et al., 2019) indicated that GA-PLS obtained better solutions than ISE-PLS when estimating the oxalate-extractable P content of paddy soils in Madagascar. However, the computational cost of GA-PLS is very high when the number of wavebands is large. Another considerable issue with GA-PLS is over-fitting when using a large number of wavebands (>200) (Leardi & Nørgaard, 2004).

One solution to simplifying the problem of variable selection is to reduce the number of variables involved in the optimization (Zhang et al., 2017). Some methods for selecting wavelength regions instead of individual wavebands have been proposed, such as moving window PLS (MWPLS) (Jiang et al., 2002; Kasemsumran et al., 2004), interval PLS (iPLS) (Nørgaard et al., 2000), and backward interval PLS (biPLS) (Leardi & Nørgaard, 2004). MWPLS (Jiang et al., 2002; Kasemsumran et al., 2004) searches for informative spectral regions using a moving window, which moves over the whole spectral region to identify useful spectral intervals. However, the sub-region selected by the window with a fixed size does not always supply the best predictions. Therefore, Du et al. (2004) developed changeable size moving window PLS (CSMWPLS) to optimize the informative regions and their combinations to further improve the predictive ability of the PLS models. Both iPLS (Nørgaard et al., 2000) and biPLS (Leardi & Nørgaard, 2004) calculate local PLS models using equally sized subintervals of the full spectrum region and identify the optimal combinations of regions by forward and backward selection, respectively. Similar to MWPLS, however, they encounter problems when the border between two contiguous intervals of equally spaced spectral regions falls inside the same spectral feature, such as when the main part of a reflectance peak is in one interval and its tail in the next interval. The solution can be found by running biPLS several times, with a different number of intervals with different interval sizes each time. To overcome these problems, a dynamic version of biPLS has been developed by Leardi and Nørgaard (2004). Dynamic biPLS runs several times using a different composition of the deletion groups (determined by randomizing the order of the samples) and with a different number of intervals (e.g. from 16 to 25).

Here, we adopted this dynamic biPLS approach to explore relevant wavelength regions for prediction of

the TC content of upland and lowland soils in Madagascar. To evaluate the performance of the selected wavelength regions for TC calibration, this study compares the predictive ability of dynamic biPLS with that of standard full-spectrum PLS (FS-PLS).

2. Materials and methods

2.1. Data set

This study used the same data set used in our previous study (Kawamura et al., 2019, 2017); the data set was generated based on laboratory Vis-NIR spectroscopy using soil samples ($n = 162$) collected from upland and lowland rice fields in the central highland of Madagascar (Figure 1). This area is dominated by inherently nutrient-poor soil types that are mainly classified into Ferralsols and Acrisols (IUSS Working Group, WRB, 2014) or into Oxisols of semiarid to humid climates (Soil Survey Staff, 2014). In 2015 and 2016, soil sampling was conducted in 158 rice fields. Surface soil samples were collected from a 0–10 cm depth as composites of three to four cores in each field. Within three fields, sub-surface samples (10–20 cm depth in a field; 10–20, 20–30, and 30–40 cm depth in two fields) were also collected to evaluate the effect of the depth of soil layers. Thus, 165 soil samples were obtained. The soil samples were sieved to <2 mm and air-dried for 7 days in the laboratory. The TC contents of soils were determined using the dry combustion method with an automatic NC analyser, SUMIGRAPH NC-220 F (Sumika Chemical Analysis Service, Ltd., Osaka, Japan).

The Vis-NIR spectra for dry soil samples were recorded using an ASD FieldSpec 4 Hi-Res spectroradiometer (ASD Inc., Longmont, CO, USA) and an ASD contact-probe in a dark room. Preprocessing, including noise reduction by standard normal variate (SNV) and outlier detection, was performed on the reflectance and first derivative reflectance (FDR) spectra over a wavelength range from 400 to 2400 nm (2001 bands). The outliers were detected based on the Mahalanobis distance $H > 3$ from principal component analysis (Kawamura et al., 2017). As a result, three samples were considered outliers, leaving 162 samples for further analyses.

2.2. PLS calibrations

PLS calibrations were performed based on reflectance and FDR spectra data sets using 'PLS_Toolbox' in MATLAB software ver. 9.3 (MathWorks Herborn, MA, USA). The dynamic biPLS (Leardi & Nørgaard, 2004) was computed using the 'iToolbox' revision released in March 2013 (<http://www.models.kvl.dk/iToolbox>).

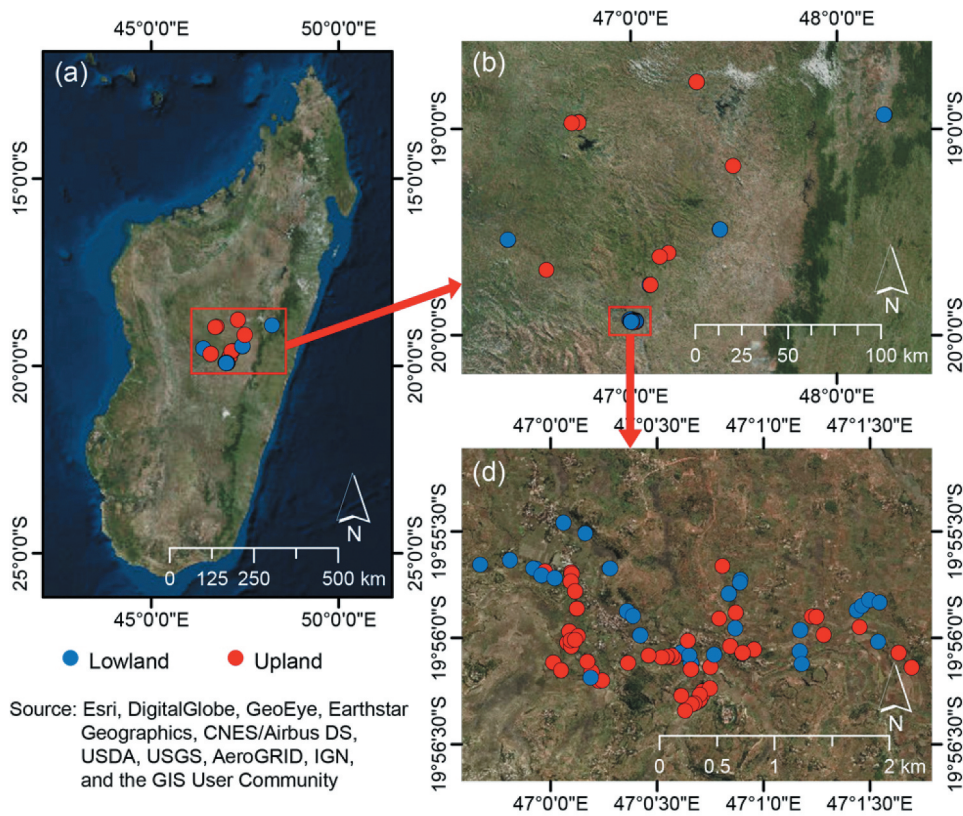


Figure 1. Location of study area and soil sampling points. Source for background images in (a), (b) and (d): Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community.

The dynamic biPLS was performed in 50 runs with the number of intervals varying from 16 to 25 and with three runs (each with a different composition of the deletion groups) for each number of intervals. After 50 runs of biPLS, the wavelength regions were selected by a backward stepwise selection procedure based on the frequency of selections. The final output is a plot showing how many times each waveband was retained after 50 runs with the threshold value (Leardi & Nørgaard, 2004).

2.3. Predictive ability of the PLS models

To evaluate the predictive ability of the FS-PLS and dynamic biPLS, a k -fold cross-validation procedure based on independent training and test data sets was performed (Emmert-Streib & Dehmer, 2019). Initially, the data were divided randomly into training ($n = 120$) and test ($n = 42$) data sets. Next, the training data were split randomly into k -folds. Here, we used $k = 5$; therefore, each k -fold has $n = 24$ samples. The PLS model was built on $k - 1$ folds of the training data set ($n = 96$), and then the error of the k th fold was recorded as validation data ($n = 24$). This process was repeated until each of the k -folds served as the validation data set. The coefficient

of determination (R^2) and root mean squared error (RMSE) values were used to assess model accuracy. Finally, the model was applied to the test data set, and then the predictive ability was evaluated from the R^2 , the RMSE and the ratio of performance to interquartile range (RPIQ) (Bellon-Maurel et al., 2010) in the test data set.

The R^2 is calculated as:

$$R^2 = \frac{RSS}{TSS} \quad (1)$$

where RSS is the residual sum of squares and TSS is the total sum of squares. The RMSE is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y_p)^2}{n}} \quad (2)$$

where y_i and y_p represent the measured and predicted soil TC contents for sample i , respectively, and n is the number of samples in the test data sets ($n = 42$). R^2 is a measure of how well the variation in one variable explains the variation in an other variable and is presented as the percentage of the variation explained by a best-fit regression line. RMSE indicates the total prediction error of the model. In general, high R^2 and low RMSE values reflect models that can better predict the soil TC content (Kusumo et al., 2008).

The RPIQ is defined as follow:

$$\text{RPIQ} = \frac{\text{IQ}}{\text{RMSEP}} \quad (3)$$

where IQ is the inter quantile distance between Q3 and Q1 of the observed values. In terms of the performance ability of the calibration model, RPIQ values >2.5 are considered to reflect excellent models, RPIQ values between 2.0 and 2.5 indicate a very good model with predictive ability, RPIQ values between 1.7 and 2.0 indicate a good model, RPIQ values between 1.4 and 1.7 indicate a fair model in need of some improvement, and RPIQ values <1.4 indicate that a model has a very poor predictive ability (Nawar & Mouazen, 2017).

3. Results and discussions

3.1. Soil TC statistics

The descriptive statistics of the soil TC (%) in the 162 samples (Kawamura et al., 2019, 2017) are shown in Table 1, and the data distribution is illustrated in Figure 2. The mean (and standard deviation (SD)) value was 3.05% (1.72%), with a range of 0.65–10.15%. The TC content was left-skewed, with a higher mean value (3.05%) than the median value (2.70%). The coefficients of variation (CVs) were relatively high (56.40%), indicating a rather high degree of variation, and the distribution was heterogeneous. The SD and range of the sample affect the accuracy of soil property predictions using Vis-NIR spectroscopy, and the wide range of variability indicated that this site is a reasonably optimal case study site (Kuang & Mouazen, 2011). In the present study, the range of soil TC values was considered sufficiently large to develop calibration models using PLS regression analyses. Soil TC was generally higher in

Table 1. Descriptive statistics of soil TC data.

Soil parameter	<i>n</i>	Min	Max	Median	Mean	SD	CV
TC (%)	162	0.65	10.15	2.70	3.05	1.72	56.40

surface soils than in sub-surface soils at the respective fields where sub-surface samples were collected. Meanwhile, the variation in TC content between surface and sub-surface layers within each field was much smaller than that among surface soils of all the 158 fields.

3.2. Soil reflectance and FDR spectra

Figure 3 shows the soil reflectance and FDR spectra. Large variations in the reflectance spectra were observed in the heterogeneous soil samples, which were collected from upland and lowland fields under various rice-based cropping systems. In the reflectance spectra, three strong absorption features were found in the 1400-, 1900- and 2200-nm wavelength regions (Figure 3(a)). The FDR spectra also showed some peaks in the same regions and in visible regions (Figure 3(b)). The Vis-NIR spectra are general characteristics of absorption wavebands associated with color (400–700 nm), the bending (1413 nm) and stretching (1916) of the O-H bonds of free water, and lattice minerals (approximately 2210 nm) (Viscarra Rossel et al., 2006; Ben-Dor, 2002; Knadel et al., 2013; Stenberg et al., 2010).

In general, soil reflectance decreases with increasing organic matter (Ben-Dor et al., 1997) and water content (Whiting et al., 2004). Wavelengths centered at approximately 400, 450, 510, 550, 700, 870 and 1000 nm are characteristics of the presence of ferrous and ferric iron oxides and are due to the electronic transitions of the iron cations (Ben-Dor et al., 1999). In

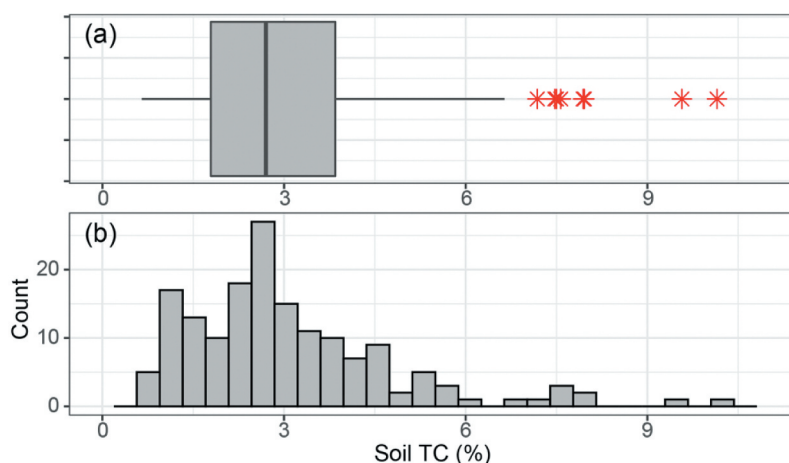


Figure 2. (a) Box-and-whiskers plot with outliers and (b) a histogram of soil TC.

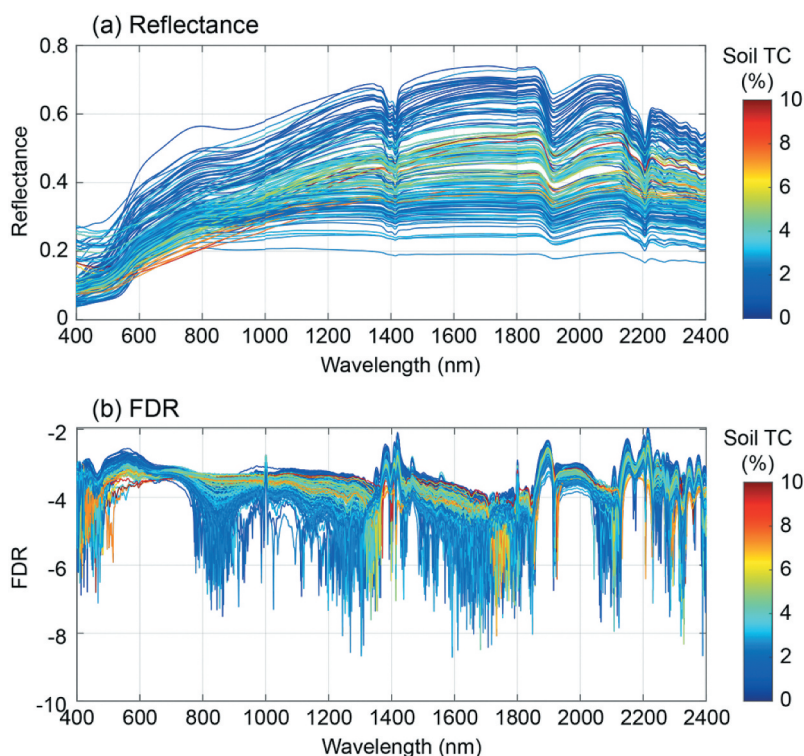


Figure 3. Raw reflectance spectra (a) and FDR spectra on a log₁₀ scale (b) of the soil samples.

addition to soil components, physical soil properties, such as particle size distribution and aggregate size and density, also affect both the reflectance intensity and shape of the soil spectra through the phenomena of light scattering and reflection (Bellon-Maurel & McBratney, 2011; Conforti et al., 2018). Thus, the soil spectral behavior can be considered as combination of chemical and physical properties of soil (Clark, 1999). Conforti et al. (2018) reported that soil reflectance showed relatively high value for loamy sand soils due to the high amount of quartz in the sand fraction, while reflectance decreased when clay content dominated from phyllosilicates increased, and consequently, when SOC concentration increased.

3.3. Selected wavelength regions from dynamic biPLS

The selected wavelength regions and the frequency of selections after 50 runs of dynamic biPLS using reflectance and FDR spectra to estimate soil TC are shown in Figure 4, and Table 2 summarizes the selected wavelength regions with previously known wavebands related to soil components to assist in considering the importance of the selected wavelengths.

In the reflectance data set, four regions of 400–490, 1402–1440, 1846–1980 and 2151–2283 nm were

selected in the model. Judging from the high selection frequency, the 1846–1980 nm region was considered the most important region in the reflectance data sets for soil TC predictions. The regions of 1402–1440 and 1846–1980 nm include several wavebands known to be relevant to soil free water and to vary with the soil organic matter content (Knadel et al., 2013). Our current data set also showed a significant correlation between the air-dried soil water content and TC content ($p < 0.1\%$, $r = 0.625$). Since soil organic matter increases soil water retention (Rawls et al., 2003), we assumed that the increase in soil water reflected a co-occurrence relation with the increase in soil organic matter. The wavelength region of 2151–2283 nm mainly consisted of the wavebands associated with organic matter and Al hydroxides. Meanwhile, the wavelength region of 400–490 nm is related to Fe oxides, which are mainly associated with soil color as well as soil organic carbon. Soils in the tropics are rich in Fe and Al (hydr)oxides because of intensive weathering and leaching (Ramaroson et al., 2018), and Fe and Al (hydr)oxides are well known to increase the stability of organic matter in soils through the formation of organo-metal complexes (Van De Vreken et al., 2016). Organic matter is spectrally active in large regions of the Vis-NIR spectrum due to overtones and combinations of NH, CH, and CO groups (Bendor et al., 1997).

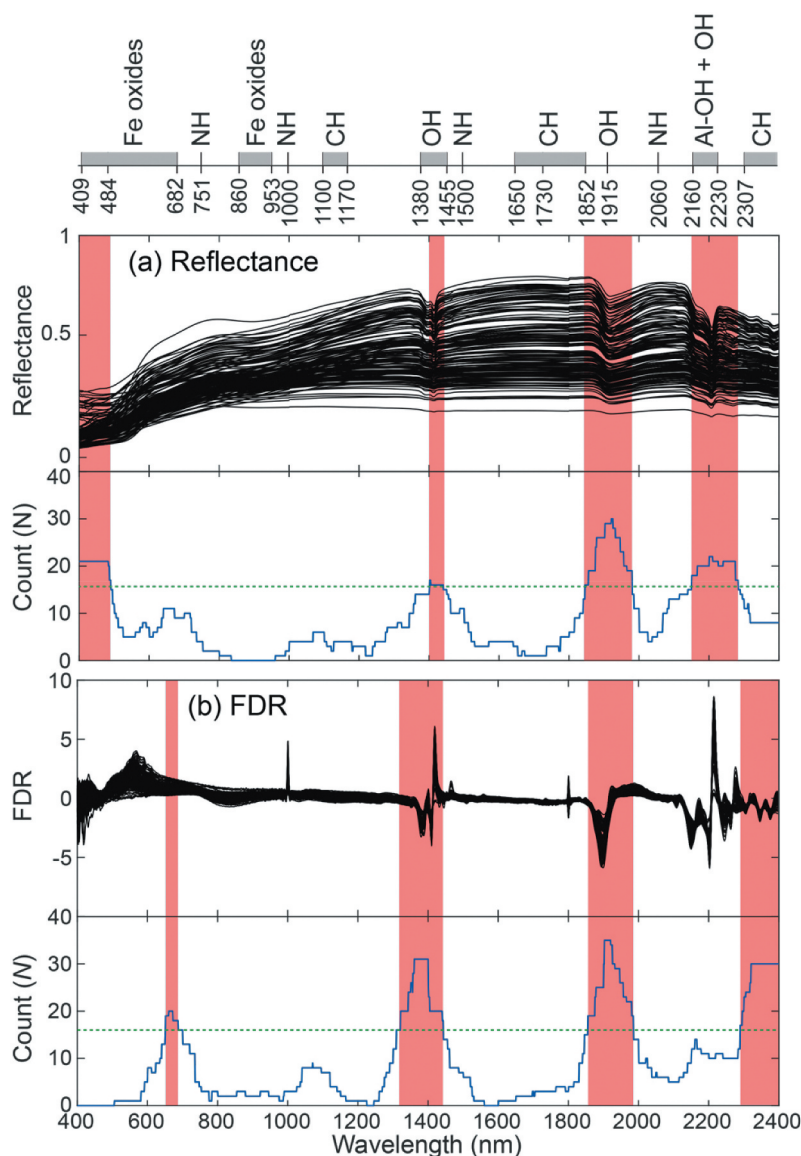


Figure 4. Selected wavelength regions (red bars) from dynamic biPLS for estimating the TC content of paddy soils using reflectance (a) and FDR (b) spectra with the frequencies (count number (N); blue line) of the selected wavebands in dynamic biPLS. Specific absorption bands for the different bonds in soil are specified in the top x-axis (modified by Katuwal et al. (2018)).

In the FDR data set, four different wavelength regions in terms of wavelength and width (652–687, 1322–1443, 1856–1985, and 2290–2400 nm) were selected in the dynamic biPLS model. However, the two wavelength regions (1322–1443 and 1856–1985 nm) overlap with those selected in the reflectance data set (1402–1440 and 1846–1980 nm), suggesting the constant significance of soil free water for soil TC predictions. The 2290–2400 nm region contains many wavebands related to soil components, such as organic matter and Fe hydroxides. The wavelengths associated with organic matter composition were selected in the higher wavelength regions with precise information on the structural and

functional groups in this data set rather than the reflectance data set because the FDR process enhances the narrow absorption features of organic matter. The wavelength region of 652–687 nm includes the wavebands related to Fe oxides (goethite and hematite), which are important for the stabilization of organic matter in soils (Saidy et al., 2012). These selected wavelengths were also identified as potentially important wavebands for soil TC prediction in our previous study using part of the current sample set (Kawamura et al., 2017).

Those selected wavelength regions did not substantially vary among the surface and sub-surface samples, while the spectral response of reflectance and FDR

Table 2. Selected wavelength regions from dynamic biPLS for estimating the soil TC content using reflectance and FDR data sets and possible soil components.

Selected wave-length region (nm)	Previously known wave-bands related to soil components	Waveband (nm)	Soil component	Reference
400–490		420, 427, 480	Goethite	Scheinost et al. (1998), Sherman and Waite (1985)
	652–687	570–700	Fe ³⁺ , ferric oxide Organic matter	Hunt (1977) Galvao and Vitorello (1998)
		660	Goethite	Stenberg et al. (2010)
		665	Organic matter, chlorophyll pigment	Ben-Dor et al. (1997)
		682	Hematite	Scheinost et al. (1998)
1402–1440	1322–1443	1367, 1358	OH in water	Ben-Dor et al. (1997)
		1400	Molecular water, first overtone of O–H stretch, Al–OH or Mg–OH	Hunt (1977)
		1400 (1395, 1415)	Kaolinite, overtones of the O–H stretch vibration near 2778 nm, smectite (structural O–H stretching), combination vibrations of water in interlayer lattice	Bishop et al. (1994)
		1415	OH, overtones of O–H stretch vibrations near 2778 nm	Oinuma and Hayashi (1965)
		1449	Carboxylic acid (C=O)	Clark et al. (1990), Clark (1999)
1846–1980	1856–1985	1870	Carbonates	Hunt (1977), White (1971)
		1875	Water	Hunt (1977)
		1900	Smectite, combination vibrations of water in interlayer lattice	Bishop et al. (1994)
		1900, 1906	Molecular H ₂ O	Hunt (1977)
2151–2283		2160	Al–OH, kaolin	Oinuma and Hayashi (1965)
		2169	Starch, lignin, wax, tannins	Ben-Dor et al. (1997)
		2180	Organic matter	Daniel et al. (2003)
		2193	Amid II, protein	Ben-Dor et al. (1997)
		2199, 2206	Illite	Post and Noble (1993)
		2200	Al–OH bend with O–H stretch combinations (kaolinite, montmorillonite, illite, smectite)	Clark et al. (1990), Post and Noble (1993)
		2206	O–H stretch (smectite, illite)	Oinuma and Hayashi (1965)
		2240	protein	Fourty et al. (1996)
		2279	CH ₂ , CH ₃	Ben-Dor et al. (1997)
		2275	Aliphatics (C–H)	Clark (1999)
	2290–2400	2290	Fe–OH	Clark et al. (1990), Post and Noble (1993)
		2300	Mg–OH	Clark et al. (1990), Post and Noble (1993)
		2300, 2350	C–H stretch fundamentals	Clark et al. (1990)
		2307–2469	Methyls	Clark (1999)
		2347	Aliphatic C–H, cellulose, lignin, glucan	Eyal Ben-Dor et al. (1997)
		2380	Lignin	Fourty et al. (1996)
		2380	Muscovite	Post and Noble (1993)
		2381	Carbohydrates (C–O)	White (1971)

simply followed soil TC gradient of the samples. This suggested that the important wavelengths to estimate soil TC content did not change along the soil depth and that the developed model could be applicable to a wide range of soil samples in the rice fields in the central highlands of Madagascar. Nevertheless, in the current study, only a limited number of soil samples were collected from both surface and sub-surface layers, and the difference in soil TC content among surface and sub-surface soil samples varied 0.76–2.95%, which is much narrower than the variation among the all the 165 sam-

ples (Table 1). Further study is needed to confirm the effect of soil depth on the associated wavelengths to soil TC content.

3.4. Evaluation of the predictive ability

Table 3 summarizes the number of selected wavebands (NW), the NW percentages of all available bands (NW % = NW/2001 bands × 100%), the optimal number of latent variables (NLVs) and the cross-validated mean R^2 and RMSE in the validation data set ($n = 24$) and the R^2 ,

Table 3. The mean R^2 and RMSE from the 5-fold cross-validation using the training data sets based on PLS analyses and the R^2 , RMSE and RPIQ based on the model applied to the test data sets.

Spectral data type	Regression method	Training data ($n = 120$)			Test data ($n = 42$)			NW ^b	NW% ^c
		NLV ^a	R^2	RMSE	R^2	RMSE	RPIQ		
Reflectance	FS-PLS	10	0.863	0.606	0.826	0.809	2.344	2001	100.0
	biPLS	12	0.872	0.585	0.877	0.690	2.748	398	19.9
FDR	FS-PLS	8	0.932	0.425	0.933	0.518	3.658	2001	100.0
	biPLS	10	0.947	0.377	0.940	0.494	3.841	399	19.9

^aNumber of latent variable.

^bNumber of wavebands.

^cNumber of wavebands percentage of all available bands.

RMSE and RPIQ values from the model on the independent test data set ($n = 42$). The optimal NLVs, which were determined as the lowest RMSE values calculated from LOO-CV to avoid over-fitting of the model, were lower with FS-PLS (10 in the reflectance data set and 8 in the FDR data set) than with biPLS (12 and 10, respectively). The NW (NW%) remaining after 50 runs of dynamic biPLS was 398 (19.9%) for the reflectance data set and 399 (19.9%) for the FDR data set, suggesting that more than 80% of the waveband information from the soil reflectance spectrum was redundant and did not contribute to the prediction or disturbed the prediction.

These results are consistent with previous findings suggesting that the spectral efficiency of PLS models can be improved through waveband selection and that the most useful information in the Vis-NIR region (400–2400 nm) predicted less than 20% of spectrum data (Kawamura et al., 2017, 2010; Wang et al., 2017).

Figure 5 shows the relationship between observed and predicted soil TC contents in the test data set ($n = 42$) from the FS-PLS and dynamic biPLS models using the reflectance and FDR spectra data sets. Clearly, the FDR data sets subjected to FS-PLS ($R^2 = 0.933$, RMSE = 0.518) and biPLS ($R^2 = 0.940$, RMSE = 0.494) yielded better predictive

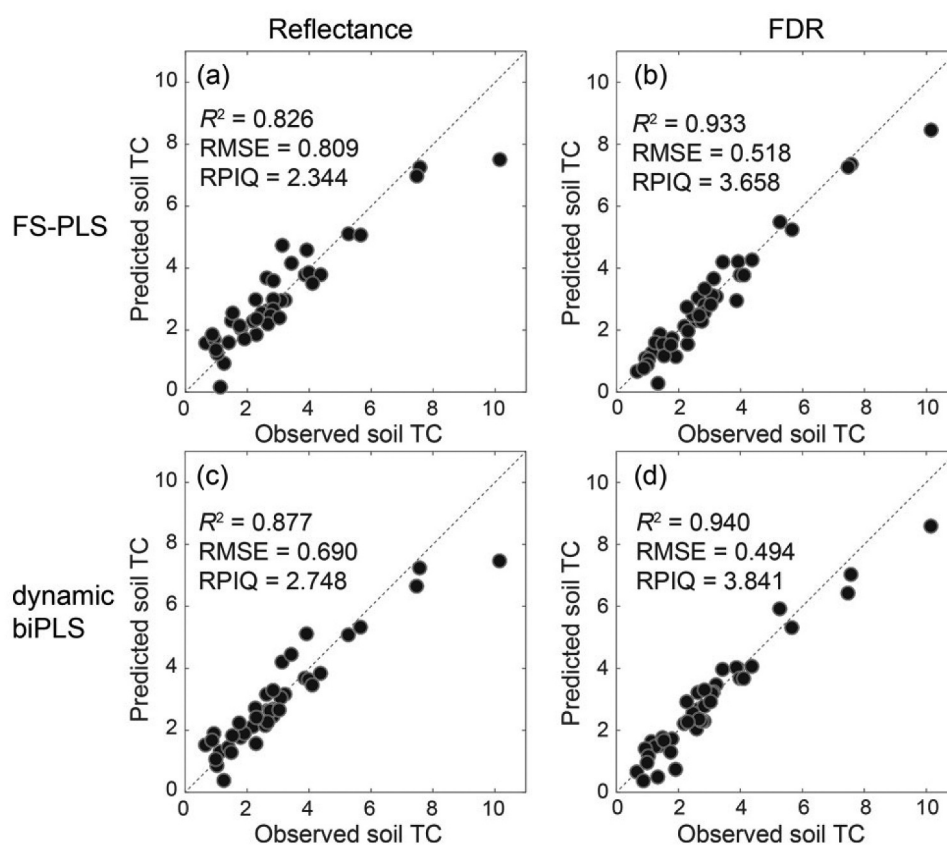


Figure 5. Observed and predicted soil TC contents from the FS-PLS (blue) and dynamic biPLS (red) models using original reflectance (a) and FDR (b) data.

accuracies than the reflectance data sets subjected to FS-PLS ($R^2 = 0.826$, RMSE = 0.809) and biPLS ($R^2 = 0.877$, RMSE = 0.690). Based on the RPIQ values from the FDR data set (RPIQ > 2.5), the quality and future applicability of our results can be considered to reflect excellent predictive ability. First derivative processing is a key preprocessing step in analytical chemistry for reducing the background signal (e.g. soil color or particle size) and enhancing the narrow absorption features related to soil properties (Brunet et al., 2007; Reeves et al., 2002). Thus, many researchers have used FDR spectra to estimate soil C contents (Chang et al., 2001; Kawamura et al., 2017; Reeves et al., 2002; Russell, 2003; Shepherd & Walsh, 2002).

The dynamic biPLS models showed better predictive ability than the FS-PLS models with fewer variables (selected wavelength regions), and a simpler and cheaper spectrophotometer can be used. These findings confirm previous results showing that the performance of PLS models can be improved through wavelength selection (Cramer et al., 2008; Du et al., 2004; Goicoechea & Olivieri, 2003; Jiang et al., 2002; Kasemsumran et al., 2004). Moreover, previous researchers have suggested that reducing large spectral datasets is valuable for more efficient storage, computation, and transmission (Yang et al., 2012) and for increasing the ease of spectral analysis (Viscarra Rossel & Lark, 2009).

In the present study, our results confirmed that soil TC could be rapidly predicted using selected wavelength regions with better predictive accuracy than FS-PLS, and sequential application of dynamic biPLS may be a feasible strategy for local assessments of soil TC. However, we note that our results were derived from heterogeneous and small numbers of soil samples ($n = 162$), which were collected from upland and lowland soils under various rice-based cropping systems, including wide ranges of soil types, in the central highland of Madagascar. Several researchers consider the reliability of predictions questionable when studying heterogeneous sample sets (Brunet et al., 2007). Particle size and arrangement may also affect the calibration due to the light transmission path (Chang et al., 2001). Moreover, if the calibration model is to be widely implemented, then large-scale data sets with regional or global content must be considered because the target function's nature strongly affects the performance of the different prediction approaches, and, different studies therefore provide different results (Gholizadeh et al., 2018). Compare to previous researches using large spectral data set (Hermansen et al., 2016; Stevens et al., 2013), our data set was relatively small diversity on soil carbon content, which is difficult to assess the robustness and applicability at larger spatial scale. To map and assess the spatial

distributions of the carbon stock at a larger spatial scale in Madagascar, evaluating an appropriate spatial scale with a larger data set is required (Ramifehiarivo et al., 2017; Saiano et al., 2013). Meanwhile, a study by Stevens et al. (2013) using a large-scale EU soil survey data set ($n = 20,000$) reported that predictive ability of SOC calibrations related to the different levels of SOC and variations in other soil properties (sand and clay content). They also suggest that large spectral data set can be valuable to build local and more accurate models that are specific to given geographical entity or soil type. Therefore, to apply the methodology to soil characterization of the whole island of Madagascar, the calibration models should be evaluated for the effect of the heterogeneous data set and updated using a larger dataset collected from various regions in Madagascar in the future.

4. Conclusions

Wavelength region selection rather than individual waveband selection is one approach to simplifying variable selection complexity. In this study, we explored relevant wavelength regions for prediction of the TC content in paddy soils in Madagascar using dynamic biPLS. Our results confirmed that a large range of soil TC (0.65–10.15%) can be rapidly and non-destructively predicted by Vis-NIR spectroscopy, and that the predictive ability was improved by wavelength region selection with dynamic biPLS. Rapid estimations of soil TC can be used to assess soil fertility and supports farmers in implementing suitable fertilizer management practices to improve crop production. Sequential application of biPLS suggested that the important wavelength regions for estimating soil TC were 400–490, 1402–1440, 1846–1980 and 2151–2283 nm in the reflectance data sets (398 bands, 19.9%) and 652–687, 1322–1443, 1856–1985, and 2290–2400 nm in the FDR data sets (399 bands, 19.9%). The selected wavelength regions were considered to be associated with organic matter, Fe and Al oxides, which are common in tropical soils and effective for sorbing and stabilizing soil organic matter. These findings are consistent with previously known soil TC-related absorption features. Thus, the selected wavelength regions should be considered informative wavelength regions for estimating soil TC. Based on the selected FDR wavelength regions in the biPLS model, the soil TC predictions were considered to be excellent (RPIQ > 2.5), with an RMSE of 0.494% in the independent test data set. These findings indicated that sequential application of biPLS was a feasible approach for optimizing wavelength region selection and the combinations for soil TC prediction with Vis-

NIR spectroscopy. To up-scale the soil TC calibrations, future analyses will be expanded to examine the effects of heterogeneous samples and extended to the whole island of Madagascar.

Acknowledgments

We would like to especially thank Dr. Naoki Moritsuka, Graduate School of Agriculture, Kyoto University in Japan, for his valuable comments on this manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA) (Grant No. JPMJSA1608).

ORCID

Kensuke Kawamura  <http://orcid.org/0000-0002-2824-1266>
Hidetoshi Asai  <http://orcid.org/0000-0003-0125-1234>

References

- Andersen, C. M., & Bro, R. (2010). Variable selection in regression – A tutorial. *Journal of Chemometrics*, 24(11–12), 728–737. <https://doi.org/10.1002/cem.1360>
- Araújo, S. R., Wetterlind, J., Demattê, J. A. M., & Stenberg, B. (2014). Improving the prediction performance of a large tropical vis-NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. *European Journal of Soil Science*, 65(5), 718–729. <https://doi.org/10.1111/ejss.12165>
- Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M., & McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends in Analytical Chemistry*, 29(9), 1073–1081. <https://doi.org/10.1016/j.trac.2010.05.006>
- Bellon-Maurel, V., & McBratney, A. (2011). Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – Critical review and research perspectives. *Soil Biology and Biochemistry*, 43(7), 1398–1410. <https://doi.org/10.1016/j.soilbio.2011.02.019>
- Ben-Dor, E. (2002). *Quantitative remote sensing of soil properties* (Vol. 75). Academic Press. [https://doi.org/10.1016/S0065-2113\(02\)75005-0](https://doi.org/10.1016/S0065-2113(02)75005-0)
- Ben-Dor, E., Irons, J. R., & Epema, J. F. (1999). Soil reflectance. In A. N. Rencz (ed.), *Manual of remote sensing: Remote sensing for the earth sciences* (Vol. 3, pp. 111–188). John Wiley & Sons.
- Ben-Dor, E., Inbar, Y., & Chen, Y. (1997). The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sensing of Environment*, 61(1), 1–15. [https://doi.org/10.1016/S0034-4257\(96\)00120-4](https://doi.org/10.1016/S0034-4257(96)00120-4)
- Bishop, J. L., Pieters, C. M., & Edwards, J. O. (1994). Infrared spectroscopic analyses on the nature of water in montmorillonite. *Clays and Clay Minerals*, 42(6), 702–716. <https://doi.org/10.1346/CCMN.1994.0420606>
- Boggia, R., Forina, M., Fossa, P., & Mosti, L. (1997). Chemometric study and validation strategies in the structure-activity relationships of new cardiotoxic agents. *Quantitative Structure-Activity Relationships*, 16(3), 201–213. <https://doi.org/10.1002/qsar.19970160303>
- Brunet, D., Barthès, B. G., Chotte, J.-L., & Feller, C. (2007). Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma*, 139(1), 106–117. <https://doi.org/org/http://dx.doi.org/10.1016/j.geoderma.2007.01.007>
- Centner, V., Massart, D. L., de Noord, O. E., de Jong, S., Vandeginste, B. M., & Sterna, C. (1996). Elimination of uninformative variables for multivariate calibration. *Analytical Chemistry*, 68(21), 3851–3858. <https://doi.org/10.1021/ac960321m>
- Chang, C. W., Laird, D., Mausbach, M. J., & Hurburgh, C. R. J. (2001). Nearinfrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65(2), 480–490. <https://doi.org/10.2136/sssaj2001.652480x>
- Cipullo, S., Nawar, S., Mouazen, A. M., Campo-Moreno, P., & Coulon, F. (2019). Predicting bioavailability change of complex chemical mixtures in contaminated soils using visible and near-infrared spectroscopy and random forest regression. *Scientific Reports*, 9(1), 4492. <https://doi.org/10.1038/s41598-019-41161-w>
- Clark, R. N., King, T. V. V., Klejwa, M., Swayze, G. A., & Vergo, N. (1990). High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research*, 95(B8), 12653–12680. <https://doi.org/10.1029/JB095iB08p12653>
- Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy. In A. N. Rencz (ed.), *Manual of remote sensing*. (pp. 3–58). John Wiley and Sons, Inc. <https://speclab.cr.usgs.gov/PAPERS.refl-mrs/refl4.html>
- Conforti, M., Buttafuoco, G., Leone, A. P., Aucelli, P. P. C., Robustelli, G., & Scarciglia, F. (2013). Studying the relationship between water-induced soil erosion and soil organic matter using Vis-NIR spectroscopy and geomorphological analysis: A case study in southern Italy. *Catena*, 110, 44–58. <https://doi.org/10.1016/j.catena.2013.06.013>
- Conforti, M., Castrignanò, A., Robustelli, G., Scarciglia, F., Stelluti, M., & Buttafuoco, G. (2015). Laboratory-based Vis-NIR spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content. *Catena*, 124, 60–67. <https://doi.org/10.1016/j.catena.2014.09.004>
- Conforti, M., Matteucci, G., & Buttafuoco, G. (2018). Using laboratory Vis-NIR spectroscopy for monitoring some forest soil properties. *Journal of Soils and Sediments*, 18(3), 1009–1019. <https://doi.org/10.1007/s11368-017-1766-5>
- Cramer, J. A., Kramer, K. E., Johnson, K. J., Morris, R. E., & Rose-Pehrsson, S. L. (2008). Automated wavelength selection for

- spectroscopic fuel models by symmetrically contracting repeated unmoving window partial least squares. *Chemometrics and Intelligent Laboratory Systems*, 92(1), 13–21. <https://doi.org/DOI10.1016/j.chemolab.2007.11.007>
- Daniel, K. W., Tripathi, N. K., & Honda, K. (2003). Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Australian Journal of Soil Research*, 41(1), 47–59. <https://doi.org/10.1071/SR02027>
- De Santana, F. B., de Souza, A. M., & Poppi, R. J. (2018). Visible and near infrared spectroscopy coupled to random forest to quantify some soil quality parameters. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 191, 454–462. <https://doi.org/10.1016/j.saa.2017.10.052>
- Douglas, R. K., Nawar, S., Alamar, M. C., Mouazen, A. M., & Coulon, F. (2018). Rapid prediction of total petroleum hydrocarbons concentration in contaminated soil using vis-NIR spectroscopy and regression techniques. *Science of the Total Environment*, 616–617, 147–155. <https://doi.org/10.1016/j.scitotenv.2017.10.323>
- Du, Y. P., Liang, Y. Z., Jiang, J. H., Berry, R. J., & Ozaki, Y. (2004). Spectral regions selection to improve prediction ability of PLS models by changeable size moving window partial least squares and searching combination moving window partial least squares. *Analytica Chimica Acta*, 501(2), 183–191. <https://doi.org/10.1016/j.aca.2003.09.041>
- Emmert-Streib, F., & Dehmer, M. (2019). Evaluation of regression models: Model assessment, model selection and generalization error. *Machine Learning and Knowledge Extraction*, 1(1), 521–551. <https://doi.org/10.3390/make1010032>
- Fourty, T., Baret, F., Jacquemoud, S., Schmuck, G., & Verdebout, J. (1996). Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems. *Remote Sensing of Environment*, 56(2), 104–117. [https://doi.org/10.1016/0034-4257\(95\)00234-0](https://doi.org/10.1016/0034-4257(95)00234-0)
- Galvao, L. S., & Vitorello, I. (1998). Role of organic matter in obliterating the effects of iron on spectral reflectance and colour of Brazilian tropical soils. *International Journal of Remote Sensing*, 19(10), 1969–1979. <https://doi.org/10.1080/014311698215090>
- Gholizadeh, A., Saberioon, M., Carmon, N., Boruvka, L., & Bendor, E. (2018). Examining the performance of PARACUDA-II data-mining engine versus selected techniques to model soil carbon from reflectance spectra. *Remote Sensing*, 10(8), 8. <https://doi.org/10.3390/rs10081172>
- Goicoechea, H. C., & Olivieri, A. C. (2003). A new family of genetic algorithms for wavelength interval selection in multivariate analytical spectroscopy. *Journal of Chemometrics*, 17(6), 338–345. <https://doi.org/10.1002/Cem.812>
- Hermansen, C., Knadel, M., Moldrup, P., Greve, M. H., Gislum, R., & de Jonge, L. W. (2016). Visible-near-infrared spectroscopy can predict the clay/organic carbon and mineral fines/organic carbon ratios. *Soil Science Society of America Journal*, 80(6), 1486–1495. <https://doi.org/10.2136/sssaj2016.05.0159>
- Hunt, G. R. (1977). Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics*, 42(3), 501–513. <https://doi.org/10.1190/1.1440721>
- IUSS Working Group, WRB. (2014). *World reference base for soil resources 2014, Update 2015 international soil classification system for naming soils and creating legends for soil maps* (World Soil Resources Reports No. 106). Food and Agriculture Organization of the United Nations.
- Jiang, J. H., James, R., Siesler, B. H. W., & Ozaki, Y. (2002). Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data. *Analytical Chemistry*, 74(14), 3555–3565. <https://doi.org/10.1021/ac011177u>
- Johnson, J.-M., Vandamme, E., Senthilkumar, K., Sila, A., Shepherd, K. D., & Saito, K. (2019). Near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for assessing soil fertility in rice fields in sub-Saharan Africa. *Geoderma*, 354, 113840. <https://doi.org/10.1016/j.geoderma.2019.06.043>
- Kasemsumran, S., Du, Y. P., Murayama, K., Huehne, M., & Ozaki, Y. (2004). Near-infrared spectroscopic determination of human serum albumin, [gamma]-globulin, and glucose in a control serum solution with searching combination moving window partial least squares. *Analytica Chimica Acta*, 512(2), 223–230. <http://www.sciencedirect.com/science/article/B6TF4-4C605NF-3/2/f9907924f85a5243807fc7afc37aca67>
- Katuwal, S., Knadel, M., Moldrup, P., Norgaard, T., Greve, M. H., & de Jonge, L. W. (2018). Visible-near-infrared spectroscopy can predict mass transport of dissolved chemicals through intact soil. *Scientific Reports*, 8(1), 11188. <https://doi.org/10.1038/s41598-018-29306-9>
- Kawamura, K., Tsujimoto, Y., Nishigaki, T., Andriamananjara, A., Rabenarivo, M., Asai, H., Rakotoson, T., & Razafimbelo, T. (2019). Laboratory visible and near-infrared spectroscopy with genetic algorithm-based partial least squares regression for assessing the soil phosphorus content of upland and lowland rice fields in Madagascar. *Remote Sensing*, 11(5), 506. <https://doi.org/10.3390/rs11050506>
- Kawamura, K., Tsujimoto, Y., Rabenarivo, M., Asai, H., Andriamananjara, A., & Rakotoson, T. (2017). Vis-NIR spectroscopy and PLS regression with waveband selection for estimating the total C and N of paddy soils in Madagascar. *Remote Sensing*, 9(10), 10. <https://doi.org/10.3390/rs9101081>
- Kawamura, K., Watanabe, N., Sakanoue, S., Lee, H. J., Inoue, Y., & Odagawa, S. (2010). Testing genetic algorithm as a tool to select relevant wavebands from field hyperspectral data for estimating pasture mass and quality in a mixed sown pasture using partial least squares regression. *Grassland Science*, 56(4), 205–216. <https://doi.org/10.1111/j.1744-697X.2010.00196.x>
- Knadel, M., Viscarra Rossel, R. A., Deng, F., Thomsen, A., & Greve, M. H. (2013). Visible-near infrared spectra as a proxy for topsoil texture and glacial boundaries. *Soil Science Society of America Journal*, 77(2), 568. <https://doi.org/10.2136/sssaj2012.0093>
- Kuang, B., & Mouazen, A. M. (2011). Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *European Journal of Soil Science*, 62(4), 629–636. <https://doi.org/10.1111/j.1365-2389.2011.01358.x>
- Kuang, B., Tekin, Y., & Mouazen, A. M. (2015). Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content. *Soil & Tillage Research*, 146(Part B), 243–252. <https://doi.org/10.1016/j.still.2014.11.002>
- Kusumo, B. H., Hedley, C. B., Hedley, M. J., Hueni, A., Tuohy, M. P., & Arnold, G. C. (2008). The use of diffuse reflectance spectroscopy for in situ carbon and nitrogen

- analysis of pastoral soils. *Australian Journal of Soil Research*, 46(6–7), 623–635. <https://doi.org/10.1071/SR08118>
- Leardi, R. (2000). Application of genetic algorithm-PLS for feature selection in spectral data sets. *Journal of Chemometrics*, 14(5–6), 643–655. [https://doi.org/10.1002/1099-128X\(200009/12\)14:5/6%3C643::AID-CEM621%3E3.0.CO;2-E](https://doi.org/10.1002/1099-128X(200009/12)14:5/6%3C643::AID-CEM621%3E3.0.CO;2-E)
- Leardi, R., Boggia, R., & Terrile, M. (1992). Genetic algorithms as a strategy for feature selection. *Journal of Chemometrics*, 6(5), 267–281. <https://doi.org/10.1002/cem.1180060506>
- Leardi, R., & González, A. L. (1998). Genetic algorithms applied to feature selection in PLS regression: How and when to use them. *Chemometrics and Intelligent Laboratory Systems*, 41(2), 195–207. [https://doi.org/10.1016/S0169-7439\(98\)00051-3](https://doi.org/10.1016/S0169-7439(98)00051-3)
- Leardi, R., & Nørgaard, L. Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions. (2004). *Journal of Chemometrics*, 18(11), 486–497. <https://doi.org/10.1002/cem.893>
- Li, H. D., Liang, Y. Z., Cao, D. S., & Xu, Q. S. (2012). Model-population analysis and its applications in chemical and biological modeling. *TrAC – Trends in Analytical Chemistry*, 38, 154–162. <https://doi.org/10.1016/j.trac.2011.11.007>
- Li, S., Ji, W., Chen, S., Peng, J., Zhou, Y., & Shi, Z. (2015). Potential of VIS-NIR-SWIR spectroscopy from the Chinese soil spectral library for assessment of nitrogen fertilization rates in the paddy-rice region, China. *Remote Sensing*, 7(6), 7029–7043. <https://doi.org/10.3390/rs70607029>
- Lucasius, C. B., & Kateman, G. (1994). Gates towards evolutionary large-scale optimization: A software-oriented approach to genetic algorithms-I. General perspective. *Computers & Chemistry*, 18(2), 127–136. [https://doi.org/10.1016/0097-8485\(94\)85007-0](https://doi.org/10.1016/0097-8485(94)85007-0)
- Morellos, A., Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R., Tziotziou, G., Wiebensohn, J., Bill, R., & Mouazen, A. M. (2016). Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering*, 152, 104–116. <https://doi.org/10.1016/j.biosystemseng.2016.04.018>
- Nawar, S., & Mouazen, A. M. (2017). Predictive performance of mobile vis-near infrared spectroscopy for key soil properties at different geographical scales by using spiking and data mining techniques. *Catena*, 151, 118–129. <https://doi.org/10.1016/j.catena.2016.12.014>
- Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., & McBratney, A. B. (2019). Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma*, 352, 251–267. <https://doi.org/10.1016/j.geoderma.2019.06.016>
- Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J. P., Munck, L., & Engelsen, S. B. (2000). Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Applied Spectroscopy*, 54(3), 413–419. <https://doi.org/10.1366/0003702001949500>
- Oinuma, K., & Hayashi, H. (1965). Infrared study of mixed-layer clay minerals. *American Mineralogist*, 50(1958), 1213–1227. <https://doi.org/10.1177/1356336X09345226>
- Padarian, J., Minasny, B., & McBratney, A. B. (2019). Using deep learning to predict soil properties from regional spectral data. *Geoderma Regional*, 16, e00198. <https://doi.org/10.1016/j.geodrs.2018.e00198>
- Partnership, G. R. S. (2013). *Rice almanac* (4th ed.). International Rice Research Institute.
- Post, J. L., & Noble, P. N. (1993). The near-infrared combination band frequencies of dioctahedral smectites, micas, and illites. *Clays and Clay Minerals*, 41(6), 639–644. <https://doi.org/10.1346/CCMN.1993.0410601>
- Ramaroson, V. H., Becquer, T., Sá, S. O., Razafimahatratra, H., Delarivière, J. L., Blavet, D., Vendrame, P. R. S., Rabeharisoa, L., & Rakotondrazafy, A. F. M. (2018). Mineralogical analysis of ferralitic soils in Madagascar using NIR spectroscopy. *Catena*, 168, 102–109. <https://doi.org/10.1016/j.catena.2017.07.016>
- Ramifiehiarivo, N., Brossard, M., Grinand, C., Andriamananjara, A., Razafimbelo, T., Rasolohery, A., Razafimahatratra, H., Seyler, F., Ranaivoson, N., Rabenarivo, M., Albrecht, A., Razafindrabe, F., & Razakamanarivo, H. (2017). Mapping soil organic carbon on a national scale: Towards an improved and updated map of Madagascar. *Geoderma Regional*, 9, 29–38. <https://doi.org/http://dx.doi.org/10.1016/j.geodrs.2016.12.002>
- Rawls, W. J., Pachepsky, Y. A., Ritchie, J. C., Sobecki, T. M., & Bloodworth, H. (2003). Effect of soil organic carbon on soil water retention. *Geoderma*, 116(1–2), 61–76. [https://doi.org/10.1016/S0016-7061\(03\)00094-6](https://doi.org/10.1016/S0016-7061(03)00094-6)
- Reeves, J., McCarty, G., & Mimmo, T. (2002). The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environmental Pollution*, 116 (Supplement 1), S277–S284. [https://doi.org/10.1016/S0269-7491\(01\)00259-7](https://doi.org/10.1016/S0269-7491(01)00259-7)
- Romero, D. J., Ben-Dor, E., Demattê, J. A. M., Souza, A. B. E., Vicente, L. E., Tavares, T. R., ... Eitelwein, M. T. (2018). *Internal soil standard method for the Brazilian soil spectral library: Performance and proximate analysis*. *Geoderma* 312, 95–103. <https://doi.org/10.1016/j.geoderma.2017.09.014>
- Russell, C. A. (2003). Sample preparation and prediction of soil organic matter properties by near infra-red reflectance spectroscopy. *Communications in Soil Science and Plant Analysis*, 34(11–12), 1557–1572. <https://doi.org/10.1081/CSS-120021297>
- Saiano, F., Oddo, G., Scalenghe, R., La Mantia, T., & Ajmone-Marsan, F. (2013). DRIFTS sensor: Soil carbon validation at large scale (Pantelleria, Italy). *Sensors*, 13(5), 5603. <http://www.mdpi.com/1424-8220/13/5/5603>.
- Saidy, A. R., Smernik, R. J., Baldock, J. A., Kaiser, K., Sanderman, J., & Macdonald, L. M. (2012). Effects of clay mineralogy and hydrous iron oxides on labile organic carbon stabilization. *Geoderma*, 173–174, 104–110. <https://doi.org/10.1016/j.geoderma.2011.12.030>
- Scheinost, A. C., Chavernas, A., Barrón, V., & Torrent, J. (1998). Use and limitations of second-derivative diffuse reflectance spectroscopy in the visible to near-infrared range to identify and quantify Fe oxide minerals in soils. *Clays and Clay Minerals*, 46(5), 528–536. <https://doi.org/10.1346/CCMN.1998.0460506>
- Shepherd, D. K., & Walsh, G. M. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of American Journal*, 66(3), 988–998. <https://doi.org/10.2136/sssaj2002.9880>
- Sherman, D. M., & Waite, D. T. (1985). Electronic spectra of Fe³⁺ oxides and oxide hydroxides in the near IR to near UV. *American Mineralogist*, 70, 1262–1269.

- Soil Survey Staff. (2014). *Keys to soil taxonomy* (12th ed.). USDA-Natural Resources Conservation Service.
- Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, 107(C), 163–215. [https://doi.org/10.1016/S0065-2113\(10\)07005-7](https://doi.org/10.1016/S0065-2113(10)07005-7)
- Stevens, A., Nocita, M., Tóth, G., Montanarella, L., & van Wesemael, B. (2013). Prediction of soil organic carbon at the European scale by visible and near infrared reflectance spectroscopy. *PlosOne*, 8(6), e66409. <https://doi.org/10.1371/journal.pone.0066409>
- Tsujimoto, Y., Horie, T., Randriamihary, H., Shiraiwa, T., & Homma, K. (2009). Soil management: The key factors for higher productivity in the fields utilizing the system of rice intensification (SRI) in the central highland of Madagascar. *Agricultural Systems*, 100(1–3), 61–71. <https://doi.org/10.1016/j.agsy.2009.01.001>
- Tsujimoto, Y., Rakotoson, T., Tanaka, A., & Saito, K. (2019). Challenges and opportunities for improving N use efficiency for rice production in sub-Saharan Africa. *Plant Production Science*, 22(4), 1–15. <https://doi.org/10.1080/1343943x.2019.1617638>
- Van De Vreken, P., Gobin, A., Baken, S., Van Holm, L., Verhasselt, A., Smolders, E., & Merckx, R. (2016). Crop residue management and oxalate-extractable iron and aluminium explain long-term soil organic carbon sequestration and dynamics. *European Journal of Soil Science*, 67(3), 332–340. <https://doi.org/10.1111/ejss.12343>
- Viscarra Rossel, R. A., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aichi, H., Barthès, B. G., Bartholomeus, H. M., Bayer, A. D., Bernoux, M., Böttcher, K., Brodský, L., Du, C. W., Chappell, A., Genot, V., ... Ji, W. (2016). A global spectral library to characterize the world's soil. *Earth-Science Review*, 155, 198–230. <http://dx.doi.org/10.1016/j.earscirev.2016.01.012>
- Viscarra Rossel, R. A., & Lark, R. M. (2009). Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. *European Journal of Soil Science*, 60(3), 453–464. <https://doi.org/10.1111/j.1365-2389.2009.01121.x>
- Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1–2), 59–75. <https://doi.org/10.1016/j.geoderma.2005.03.007>
- Wang, Z., Kawamura, K., Sakuno, Y., Fan, X., Gong, Z., & Lim, J. (2017). Retrieval of chlorophyll-a and total suspended solids using iterative stepwise elimination partial least squares (ISE-PLS) regression based on field hyperspectral measurements in irrigation ponds in Higashiroshima, Japan. *Remote Sensing*, 9(3), 264. <https://doi.org/10.3390/rs9030264>
- White, W. B. (1971). Infrared characterization of water and hydroxyl ion in the basic magnesium carbonate minerals. *American Mineralogist*, 56(October), 46–53. <https://doi.org/10.1128/IAI.00496-08>
- Whiting, M. L., Li, L., & Ustin, S. L. (2004). Predicting water content using Gaussian model on soil spectra. *Remote Sensing of Environment*, 89(4), 535–552. <https://doi.org/10.1016/j.rse.2003.11.009>
- Yang, H., Kuang, B., & Mouazen, A. M. (2012). Quantitative analysis of soil nitrogen and carbon at a farm scale using visible and near infrared spectroscopy coupled with wavelength reduction. *European Journal of Soil Science*, 63(3), 410–420. <https://doi.org/10.1111/j.1365-2389.2012.01443.x>
- Zhang, J., Cui, X., Cai, W., & Shao, X. (2017). Combination of heuristic optimal partner bands for variable selection in near-infrared spectral analysis. *Journal of Chemometrics*, 32(11), e2971. <https://doi.org/10.1002/cem.2971>

Article

Farm-Scale Estimation of Total Nitrogen Content in Surface Paddy Soils by Extraction with Commercially Available Hydrogen Peroxide

Naoki Moritsuka ^{1,*} , Hiroki Saito ², Ryosuke Tajima ³, Yukitsugu Takahashi ⁴ and Hideaki Hirai ⁵

¹ Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

² Tropical Agriculture Research Front, Japan International Research Center for Agricultural Sciences, Okinawa 907-0002, Japan; hirokisaito@affrc.go.jp

³ Field Science Center, Tohoku University, Miyagi 989-6711, Japan; tazy@m.tohoku.ac.jp

⁴ University Farm, Utsunomiya University, Tochigi 321-4415, Japan; takahashi@cc.utsunomiya-u.ac.jp

⁵ School of Agriculture, Utsunomiya University, Tochigi 321-8505, Japan; hirai@cc.utsunomiya-u.ac.jp

* Correspondence: morituka@kais.kyoto-u.ac.jp

Received: 17 November 2019; Accepted: 21 December 2019; Published: 26 December 2019



Abstract: We recently proposed a simple method for estimating total nitrogen content in paddy soil. In this method, soil is extracted with a commercial 3% hydrogen peroxide (H₂O₂) solution at 25 °C for 40 h, and electrical conductivity (EC (H₂O₂)) of the extract is measured. This study aimed to further evaluate the method's applicability to soil samples collected at the farm scale by using the original and six additional H₂O₂ solutions that are locally and commercially available. The results obtained with the original solution indicated that the determination coefficients between EC (H₂O₂) and total N were statistically significant at all farms examined: Moka, 0.78 ($n = 13$); Kyoto, 0.50 ($n = 16$); Kizu, 0.43 ($n = 89$); and Kawatabi, 0.25 ($n = 18$). The EC of the tested H₂O₂ solutions varied from less than 0.05 to 1.4 mS cm⁻¹ because of the addition of different stabilizers. EC (H₂O₂) values obtained with the less stabilized H₂O₂ solutions (one from Japan, one from USA, and the analytical grade 6% solution) agreed well with those obtained with the original solution. Thus, the proposed method can be useful for estimating the farm-scale variation in soil total N, provided a H₂O₂ solution with a low EC (<0.2 mS cm⁻¹) is used for the extraction.

Keywords: electrical conductivity; farm-scale variation; hydrogen peroxide; nitrogen; paddy soil; simple method

1. Introduction

In irrigated paddy fields, mineralization of organic nitrogen (N) in soil during a cropping period often limits rice growth and yield. Nutrient omission trials conducted in more than a thousand paddy fields throughout Japan have demonstrated that lowland rice responds to the application of N to a greater extent than to the application of phosphorus or potassium [1]. To optimize the rate of application of N fertilizer, many researchers have proposed laboratory methods to evaluate soil N availability suitable for routine use, and some of them have been validated by comparing the results with the content of potentially mineralizable organic N estimated by the long-term laboratory incubation [2]. Several biological and chemical methods such as short-term anaerobic incubation and extraction with hot KCl have proved useful for rapid estimation of soil N availability [2,3].

In addition to analysis of the labile fraction of N in soil, analysis of total N is also a widely accepted method for estimating soil N availability [2,4]. For example, in 31 surface paddy soils in Japan classified as non-volcanic ash soils, the content of potentially mineralizable organic N evaluated by long-term aerobic incubation was significantly correlated ($r = 0.53$, $p < 0.05$) with the total N content [5]. By reviewing previous papers, Sahrawat [6] also reported that the total N content in paddy soil usually showed a significant positive correlation with the amount of ammonium ion (NH_4^+) produced under anaerobic incubation.

The total N content in soil has been measured by two methods: the Kjeldahl method which is a wet oxidation procedure, and the Dumas method which is fundamentally a dry oxidation (combustion) procedure [7]. These methods were originally developed during the 19th century and the revised versions are still in use. Total N and total carbon (C) contents can be measured simultaneously by the dry combustion method. However, both methods require laboratory facilities, so total N analysis are typically carried out by specialists in well-equipped laboratories.

To reduce the time and cost for measurement, Sharifi et al. [4] proposed the sodium hydroxide direct distillation method. In this method, the digestion step was fully eliminated, and the amount of N liberated by steam distillation of soil with a strong alkaline reagent for about 10 min was determined. The amount of N obtained from this method had a significant positive correlation with the content of soil total N measured by the dry combustion method. Christianson and Holt [8] also proposed a rapid digestion procedure with H_2SO_4 and H_2O_2 (the peroxy reagent) without addition of metal catalysts such as Se, Hg, and Cu. The total time for digestion was only 38 min, and the rate of N recovery from six soils ranged from 89% to 98% compared to Kjeldahl digestion. However, these methods are still limited to the laboratory use, as they require hazardous reagents.

To develop a simple method for estimating soil total N usable by non-specialists, we recently proposed a method in which a dilute hydrogen peroxide (H_2O_2) solution is used as the soil extractant [9]. In this method, surface paddy soil is extracted with a commercially available 3% H_2O_2 solution at 25 °C for 40 h. Then, the electrical conductivity (EC (H_2O_2)) of the soil extract is measured with an EC electrode. Because 3% H_2O_2 commercially available in Japan is called “oxydol” (a Japanese pharmacopoeia term), this method was called the oxydol method in the previous paper [9].

The concept of this method is based on previous findings on the decomposition of soil organic matter with H_2O_2 and the detection of NH_4^+ by EC. Robinson [10] first introduced H_2O_2 digestion as a pretreatment of soil texture analysis. Considering the volatile nature of both H_2O_2 and CO_2 produced by the reaction of soil organic matter with H_2O_2 , another Robinson [11] proposed a gravimetric method for determining the content of soil organic matter by the loss in weight caused by H_2O_2 digestion. Robinson [11] also reported that practically all of the soil nitrogen was transformed to NH_4^+ during H_2O_2 digestion. This finding was confirmed by Harada and Inoko [12], who reported that, for the eight soil samples tested, 59–100% of the total N in the original soil was present in the digested solution and that most of the water-soluble N was present as NH_4^+ . Guan et al. [13] also reported that, for the eight samples including six volcanic ash soils, more than 90% of organic C and less than 11% of N in the original soil were lost by H_2O_2 treatment. The concentration of NH_4^+ in a solution can be measured quantitatively by non-selective EC detection after its separation from other cations by ion chromatography [14]. If NH_4^+ is the dominant cation in a sample solution, the NH_4^+ concentration can be roughly estimated from the solution EC without a separation pretreatment.

To validate the oxydol method, we applied it to 83 surface soils collected from paddy fields throughout Japan [9]. Then, there was a significant positive correlation between EC (H_2O_2) and total N content ($r^2 = 0.70$, $p < 0.01$) when 11 volcanic ash soils were excluded from the analysis. When we applied the method to two sets of paddy soil samples collected at farm scale, the correlation between EC (H_2O_2) and total N became different between the two farms ($r^2 = 0.23$ and 0.68). The low accuracy of the estimation for volcanic ash soils and some farm-scale samples was mainly due to the low rate of mineralization of soil organic N by the H_2O_2 treatment. This means that the estimation accuracy was decreased by the uneven decomposability of soil N within a sample set. The variations in both soil

total N content and its decomposability at farm scale were smaller than those observed at national scale [9]. At any scale of investigation, the estimation accuracy is expected to increase with an increase of the variation in total N content and also with a decrease of the variation in its decomposability. However, the number of sample sets collected at farm scale was too small to evaluate whether or not the oxydol method can be useful for estimating the farm-scale variation in soil total N.

Moreover, only one type of H_2O_2 sold in Japan was used in previous experiments [9]. It is uncertain whether other types of H_2O_2 solutions can produce results comparable to those obtained previously. This is because commercial H_2O_2 solutions contain various stabilizers to minimize decomposition of H_2O_2 into water and oxygen under normal storage conditions. For example, the H_2O_2 used in the original method [9] contained phenacetin as a stabilizer. Other stabilizers typically used in commercial H_2O_2 are colloidal stannate, sodium pyrophosphate, organo-phosphates, and colloidal silica [15]. These stabilizers are expected to increase the EC of the commercial H_2O_2 solution and consequently the EC (H_2O_2) value; the estimation accuracy of soil total N from EC (H_2O_2) might be significantly affected.

The objective of this study was to further evaluate the oxydol method in terms of (1) the applicability to farm-scale samples and (2) the possibility of using different types of commercially available H_2O_2 solutions.

2. Materials and Methods

2.1. Soil Samples

We used 136 surface paddy soils collected from four university farms in Japan; 89 samples from Kizu (Kyoto University), 16 samples from Kyoto (Kyoto University), 18 samples from Kawatabi (Tohoku University), and 13 samples from Moka (Utsunomiya University). Kizu and Kyoto were located on non-volcanic ash soils, and Kawatabi and Moka were located on volcanic ash soils (Andosols) (Figure 1). According to the World Reference Base for Soil Resources, the soils were classified as Fluvisols at Kizu and Kyoto, Aluandic Andosol at Kawatabi, and Silandic Andosol at Moka.

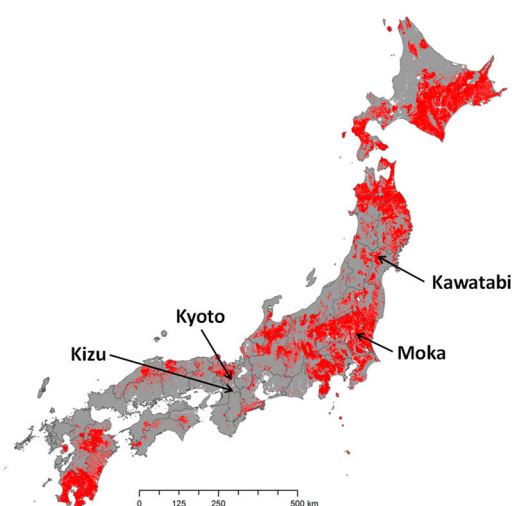


Figure 1. Map of the Japan mainland showing the distribution of Andosols (red) and the locations of the sampled farms. The distribution of Andosols is based on map data in the Digital Soil Map of Japan [16], available at the website of the National Agriculture and Food Research Organization (NARO).

The areas of the fields from which the samples were collected was approximately 2.9, 0.75, 7.5, and 7.9 ha at Kizu, Kyoto, Kawatabi, and Moka, respectively. All fields had a rectangular shape, and many of them were contiguous. At all farms except Kizu, soil samples were collected from individual fields. One representative soil sample was obtained from each field by mixing surface (<15 cm) soil samples collected from five points: the center and each of the four corners of the field. At Kizu, a more systematic grid sampling was carried out by dividing the five contiguous fields, each with an area of about 50 × 100 m, into 89 plots, each with an area of 16.7 × 16.7 m. Then, a representative soil sample was obtained from each plot by mixing surface (<10 cm) soil samples collected from five points; the center and four additional points, each located about 5 m from the center. Soil sampling was carried out in April and May, 2017, before puddling and transplanting of rice (*Oryza sativa* L.) seedlings. The samples were air-dried and passed through a 2-mm sieve before analyses. Selected properties of the soil samples used in this study are listed in Table 1.

Table 1. Selected properties of the samples used in this study.

Site (No. of Samples)	EC (1:5 <i>w/v</i>) (mS cm ⁻¹)		Total C (g kg ⁻¹)		C/N Ratio		Mineralized N (mg kg ⁻¹)		Sand (g kg ⁻¹)	
	Ave.	CV	Ave.	CV	Ave.	CV	Ave.	CV	Ave.	CV
Kizu (89)	0.17	30.3	16.1	18.1	14.7	10.6	68.7	13.2	674	1.9
Kyoto (16)	0.05	30.0	23.6	15.5	11.8	4.0	74.7	27.5	682	3.7
Kawatabi (18)	0.06	20.9	43.2	18.1	14.1	8.8	147.2	17.7	618	11.0
Moka (13)	0.08	31.6	89.2	5.7	14.3	4.3	111.2	32.6	441	2.4

EC, electrical conductivity; Ave., average; CV, coefficient of variation (%). EC was measured with an EC sensor (ES-51, Horiba, Ltd., Kyoto, Japan) after extracting samples with water at 1:5 (*w/v*). Total C was measured by the dry combustion method (Sumigraph NC-95A, Sumika Chem. Anal. Service, Osaka, Japan). Mineralized N was determined by measuring the content of NH₄⁺ after an anaerobic incubation at 40 °C for 7 days [17]. Sand content was measured by the nylon mesh sieving method [18].

2.2. Measurement of EC (H₂O₂) by the Oxydol Method

Figure 2 shows the procedures of the original oxydol method [9]. Prior to the experiment, 180-mL glass bottles with an inner diameter of 51 mm and a depth of 101 mm were washed with tap water. Then, 5.00 g of air-dried, 2-mm sieved soil was placed in each bottle, 100 mL of oxydol (Showa) was added to the sample, and the bottle was left standing at 25 °C for 40 h. After extraction, the supernatant was gently swirled by hand, and the EC of the soil extract (EC (H₂O₂)) was measured with a stick-type EC sensor (HI98331 Soil Test, Hanna Instruments, Woonsocket, RI, USA). An incubator was used to keep the temperature at 25 °C during the extraction. The extraction period of 40 h was set rather arbitrarily by measuring the increase of EC of six different paddy soils every 30 min after adding oxydol and then evaluating the time at which the EC reached a plateau [9].

In this study, we modified the original method slightly by using the Showa oxydol and six additional types of H₂O₂ solution (Figure 3). Except for one analytical grade reagent (Wako), the other H₂O₂ solutions are locally and commercially available (Table 2). The analytical grade solution contained H₂O₂ at about 30%, and was diluted with distilled water to 3% or 6% before use. The other solutions were used without dilution because the H₂O₂ concentration of each was about 3%. In addition, EC (H₂O₂) was measured with a stick-type EC sensor (HI98331 Gro Line Soil Test, Hanna Instruments, Woonsocket, RI, USA), which was an updated model of the EC sensor that we had used in the original method.

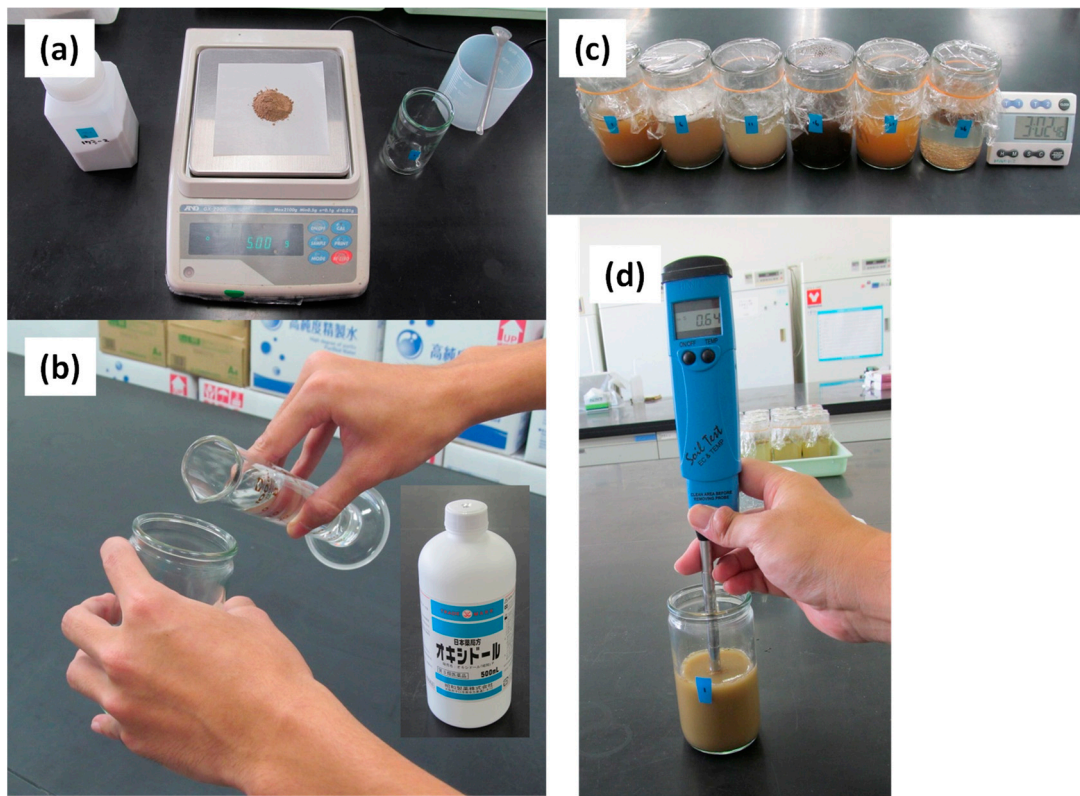


Figure 2. The original oxydol method procedure: (a) sub-sample measurement, (b) addition of 3% H₂O₂ solution (Showa), (c) extraction at 25 °C for 40 h, and (d) measurement of EC (H₂O₂), reproduced from [9] with permission from the Japanese Society of Soil Science and Plant Nutrition.



Figure 3. The analytical grade (Wako) and commercially available H₂O₂ solutions used in this study. From left to right, Wako, Showa, Kenei, Kozakai, Swan, Essential Oxygen, and Álcool e Géneros Alimentares (AGA).

Table 2. Hydrogen peroxide solutions used in this study.

Type of H ₂ O ₂	Country	Type of Stabilizers
Wako (analytical grade, 30%)	Japan	none
Showa	Japan	phenacetin
Kenei	Japan	phenacetin, phosphoric acid
Kozakai	Japan	ethanol, phenacetin
Swan	USA	unknown
Essential Oxygen	USA	unknown
AGA	Portugal	sodium benzoate

2.3. Other Analyses

The EC of the H₂O₂ solutions used as extractants was measured with an EC sensor attached with a graphite electrode (Handy SC meter TCX-999i attached with CX90CS, Toko Chemical Laboratories, Co., Ltd., Tokyo, Japan). The concentration of sodium (Na) was also measured by atomic absorption spectrometry (AA-7000, Shimadzu Corporation, Kyoto, Japan) because Na⁺ was the only cation detected in any of the solutions by ion chromatography (PIA-1000, Shimadzu Corporation, Kyoto, Japan).

For the soil samples, the total N content was measured by the dry combustion method (Sumigraph NC-95A, Sumika Chem. Anal. Service, Osaka, Japan) using the finely ground samples [7].

The amount of soil organic N that was mineralized by H₂O₂ extraction was evaluated by measuring the NH₄⁺ concentration in the soil extracts with the Showa solution. The NH₄⁺ concentration was determined by filtering the extracts, converting NH₄⁺ to NH₃ by steam distillation of the extracts with 10 mol L⁻¹ NaOH, collecting the liberated NH₃ with 2% H₃BO₃, and measuring the NH₄⁺ concentration colorimetrically by the indophenol method [19]. Because the NH₄⁺ in the H₂O₂ extracts was recovered by steam distillation with strong alkali, the recovered NH₄⁺ would include inorganic NH₄⁺ present before the H₂O₂ treatment as well as soluble organic N hydrolyzed with alkali during the steam distillation.

The oxydol method was compared with two other simple methods that may allow estimation of soil total N; measurement of sample lightness (L* value) based on the CIE 1976 (L*, a*, b*) color system and of the 0.02 mol L⁻¹ potassium permanganate-oxidizable organic C (POXC) content. The L* value of the finely ground samples was measured with a Soil Color Reader (SPAD-503, Konica Minolta, Tokyo, Japan) according to the method described by Moritsuka et al. [20]. Sample lightness was measured because soil lightness, evaluated visually by using a Munsell soil color chart, has been used for estimating the organic matter content [21,22]. The POXC concentration was measured according to the method described by Weil et al. [23] and Culman et al. [24]. The POXC method proposed by Weil et al. [23] is based on the oxidation reaction of soil organic C with potassium permanganate. As the reaction proceeds with time, the intensity of the purple permanganate color decreases in proportion to the amount of organic C oxidized. To determine the content of biologically active C in soil, the period of sample reaction was set at 12 min (shaking for 2 min and settling for 10 min) [24]. Then, after diluting the supernatant, the change in permanganate color was measured by the absorbance at 550 nm, and the amount of POXC was calculated as described by Weil et al. [23].

2.4. Statistical Analysis

The accuracy of estimation of soil total N was evaluated by the magnitude of determination coefficient obtained by linear regression analysis. The regression line obtained from each farm was compared with that obtained from the nationwide analysis [9] by the analysis of covariance (ANCOVA). Comparability of EC (H₂O₂) with the original method using the Showa oxydol was evaluated by the Student's *t*-test and the correlation analysis using the data of all samples (*n* = 136). These statistical analyses were performed using Microsoft Office Excel 2010 and Ekuseru-Toukei 2012 (Social Survey Research Information Co., Ltd., Tokyo, Japan).

3. Results

3.1. Farm-Scale Variations in Total N in Surface Paddy Soils

Table 3 shows the descriptive statistics of the total N content in surface paddy soils. Compared to the national averages in Japan (about 2.3 g kg⁻¹ [9,25]), the average total N content was lower at farms on non-volcanic ash soils (Kizu and Kyoto) and higher at farms on volcanic ash soils (Kawatabi and Moka). The content at Kizu was lowest, probably because the paddy fields in Kizu were constructed in 2016 and had not been used for rice production before we collected samples in April 2017. On the other hand, the high total N contents in Kawatabi and Moka would be due to the predominance of dark-colored organic matter stabilized by active Al and Fe. This is one of the most striking features of volcanic ash soils, by which the content of organic matter in volcanic ash soils has been estimated visually from the Munsell value (lightness) and chroma (vividness) [20]. The coefficients of variation (CV) observed at farm scale were smaller than the national-scale CV values reported previously (45–50% in Table 3). Nevertheless, farm-scale variations in soil total N were detectable at all sites, suggesting that farm-scale evaluation of soil total N may be useful for site-specific fertilizer management [26].

Table 3. Descriptive statistics of the soil total N content (g kg⁻¹) at each farm.

Site (No. of Samples)	Average	Maximum	Minimum	CV (%)
Kizu (89)	1.09	1.41	0.76	11.8
Kyoto (16)	2.00	2.72	1.58	15.5
Kawatabi (18)	3.10	4.61	2.21	21.3
Moka (13)	6.24	7.08	5.45	8.0
Nationwide (83) ^a	2.24	6.32	0.75	49.0
Nationwide (2750) ^b	2.39	no data	no data	47.3

^a, Surface paddy soils collected throughout Japan [9]. ^b, Surface paddy soils collected throughout Japan [25].

3.2. Accuracy of Estimation of Total N Content

Figure 4 shows the relationship between EC (H₂O₂) and total N content at farm scale. The determination coefficient was statistically significant when all farm-scale samples were included ($r^2 = 0.70$). However, the Kizu samples were plotted below the regression line fitted to the nationwide data, whereas the samples from the other three farms, especially Moka, were plotted above the regression line. The Moka samples were plotted outside of the prediction interval at 95% for non-volcanic ash soils at national scale. Thus, the y -intercept of the regression line fitted to the data from each farm, especially Moka, differed significantly from that obtained for the non-volcanic ash soils at national scale (Table 4). The slope of the regression line also varied considerably from 2.28 (Kizu) to 4.51 (Moka) (Table 4). Within each farm, however, the relationship was significantly positive at all sites, including Kawatabi and Moka that were classified as volcanic ash soils (Figure 4). This result was contrasting to our previous results, where we found no relationship for 11 volcanic ash soils collected from different farms ($r^2 = 0.00$; Moritsuka et al. [9]).

These results suggest that the relationship between EC (H₂O₂) and total N content is influenced by several factors, including the presence of cations other than NH₄⁺ in the H₂O₂ extracts and the degree of mineralization of soil organic N during H₂O₂ extraction. We therefore evaluated the NH₄⁺ concentration in the H₂O₂ extracts and calculated the H₂O₂-mineralizable organic N content in the samples. Then, EC (H₂O₂) values were significantly and positively correlated with the H₂O₂-mineralizable organic N content at all farms (Figure 5), suggesting that NH₄⁺ was the dominant cation in the H₂O₂ extracts.

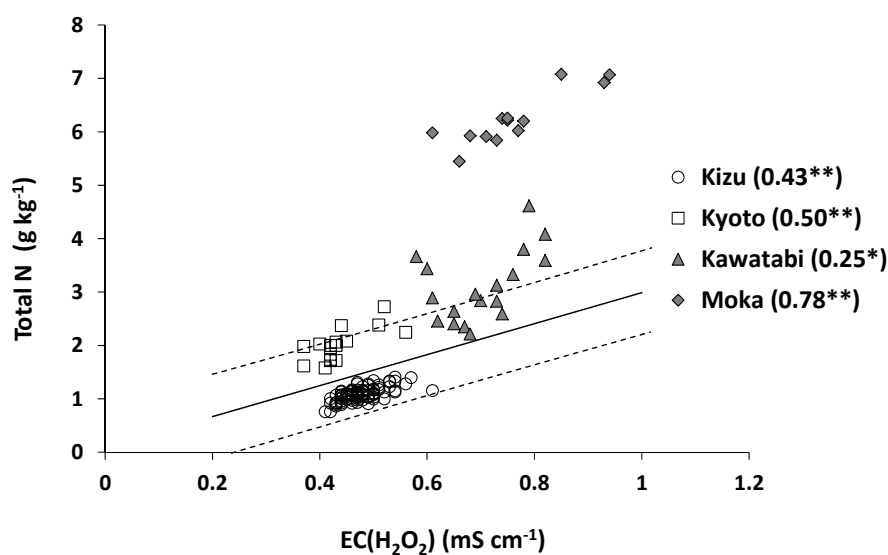


Figure 4. Relationship between EC (H_2O_2) and total N content at farm scale. The solid line in the figure shows the regression line obtained from the nationwide samples excluding volcanic ash soils ($n = 72$), and the broken lines indicate the prediction interval at 95% confidence [9]. ** and * indicate statistical significance of the determination coefficients at $p < 0.01$ and $p < 0.05$, respectively.

Table 4. Slope and y -intercept of the regression lines obtained for the relationship between EC (H_2O_2) and total N content at each farm and at national scale (Figure 4).

Site	Slope	y -Intercept
Kizu	2.28	0.00 **
Kyoto	4.29	0.13 **
Kawatabi	4.43	−0.01 **
Moka	4.51	2.80 **
Japan (nationwide)	2.90	0.09

The slopes at farm scale were not significantly different from the slope of the nationwide samples excluding volcanic ash soils (Japan (nationwide)). ** indicates the significant difference from Japan (nationwide) at $p < 0.01$ (ANCOVA).

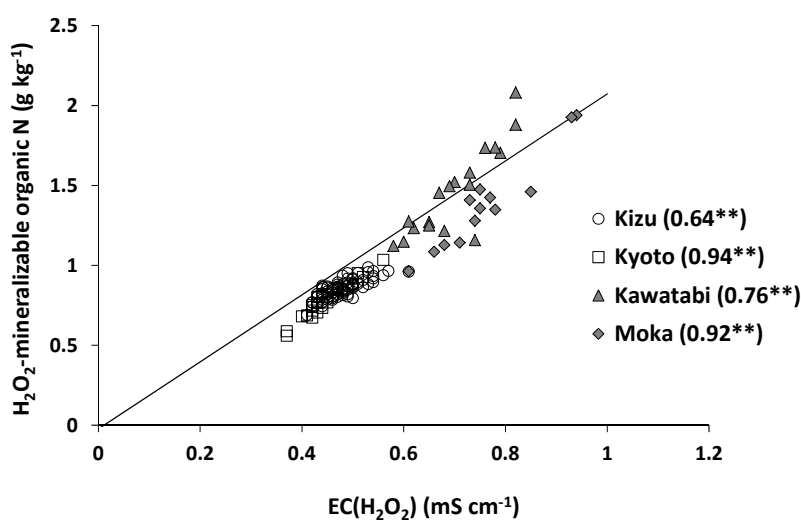


Figure 5. Relationship between EC (H_2O_2) and H_2O_2 -mineralizable organic N at farm scale. The line in the figure shows the regression line obtained from the nationwide samples ($n = 83$) [9]. ** indicates statistical significance of the determination coefficients at $p < 0.01$.

On the other hand, the proportion of H₂O₂-mineralizable organic N in total N varied considerably among the farms (Figure 6). The average percentage was 78.8% at Kizu but was less than 50% at the other farms (Kawatabi, 48.1%; Kyoto, 37.9%; and Moka, 22.0%). Compared to the percentages obtained from the nationwide samples, the average percentage of Kizu was similar to the average of the non-volcanic ash soils, whereas the average percentages of the other farms were similar to the average of the volcanic ash soils. Although the percentages differed across farms, within each farm, CV values were less than 20% and lower than the CV for nationwide samples. Thus, the proportion of soil organic N mineralized was relatively constant within each farm. This within-farm consistency must have contributed to the significant positive relationships between EC (H₂O₂) and total N content as well as to the different slopes of the regression line at farm scale (Figure 4, Table 4).

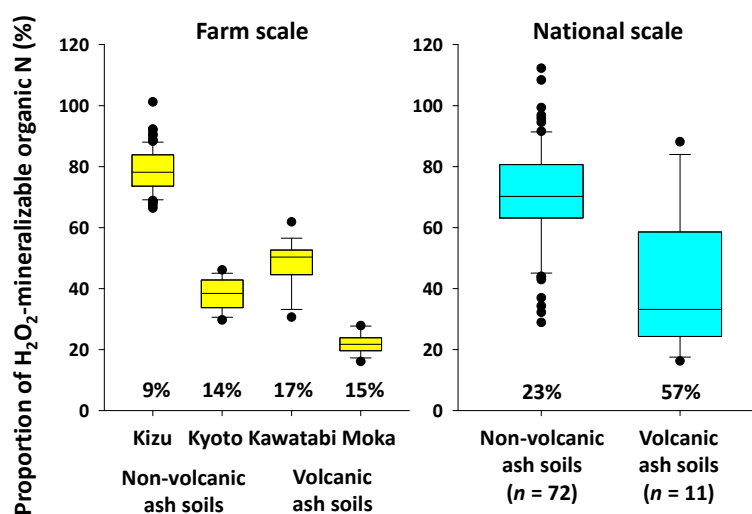


Figure 6. Box plot of the proportion of H₂O₂-mineralizable organic N in total N. Data for the national-scale samples are cited from [9]. The horizontal line in each box represents the median. The lower and upper limits of each box are the 25th and 75th percentiles, and the lower and upper whiskers represent the 10th and 90th percentiles. The values below the box plot indicate the CV values for each group.

3.3. Comparability of Commercially Available H₂O₂ Solutions

The data reported in the above sections were obtained with the original method using the Showa oxydol containing phenacetin as a stabilizer (Table 2). The EC values of the tested H₂O₂ solutions varied widely (Table 5). EC of the analytical grade H₂O₂ (Wako, 30%) was very low (0.007 mS cm⁻¹), close to that of distilled water, whereas the EC values of the six commercial H₂O₂ ranged from less than 0.05 to 1.4 mS cm⁻¹. The Na concentration of the solutions was positively correlated with EC, suggesting that Na-bearing stabilizers such as sodium benzoate (AGA, Table 2) and sodium phosphate increased the EC values of some commercial products.

Table 6 summarizes the comparability of EC (H₂O₂) values obtained with the H₂O₂ solutions tested. The average EC (H₂O₂) values of soil extracts obtained with Kozakai, Swan, and Wako (6%) did not differ significantly from the Showa value. On the other hand, the average EC (H₂O₂) values obtained with Essential Oxygen and AGA were much higher than the Showa value, reflecting the high EC of the original solutions (Table 5). The average EC (H₂O₂) obtained with Essential Oxygen was lower than the EC of the solution before soil extraction (Tables 5 and 6). The correlation coefficients with Showa were significantly positive for all the H₂O₂ solutions except Essential Oxygen.

Taken together, the results obtained with Kozakai, Swan, and Wako (6%) were considered most comparable to the Showa results. The results obtained with Kenei and Wako (3%) were also close to the Showa results, although their average EC (H₂O₂) values were slightly lower than that of Showa.

Table 5. EC values and Na concentrations of the different H₂O₂ solutions used.

Type of H ₂ O ₂	EC (mS cm ⁻¹)	Na (mg L ⁻¹)
Wako (30%)	0.010	0.10
Showa	0.029	0.51
Kenei	0.050	0.29
Kozakai	0.026	0.54
Swan	0.19	33.2
Essential Oxygen	1.40	321.6
AGA	0.55	165.5

Table 6. Comparability of EC (H₂O₂) values obtained by using different H₂O₂ solutions.

Type of H ₂ O ₂	Average EC (H ₂ O ₂) (mS cm ⁻¹)	R with Showa	Comparability with Showa
Showa	0.53	-	-
Kenei	0.50 *	0.99	high
Kozakai	0.55	0.92	very high
Swan	0.53	0.95	very high
Essential Oxygen	1.20 **	-0.73	low
AGA	0.98 **	0.95	medium
Wako (3%)	0.49 *	0.94	high
Wako (6%)	0.51	0.90	very high

* and ** indicate significant differences from the value obtained with Showa at $p < 0.05$ and $p < 0.01$, respectively (t -test). The correlation coefficient significantly positive at 1% level was 0.23 ($n = 136$).

3.4. Comparing the Estimation Accuracy between the Oxydol Method and Other Methods

Figure 7 shows the relationship between the L* value (lightness) of finely-ground samples and total N content at farm scale. Almost all samples (132 out of 136) were within the prediction interval of non-volcanic ash soils at national scale, and the determination coefficient obtained from all samples was high ($r^2 = 0.82$). At farm scale, however, the relationship was statistically significant only at Kizu.

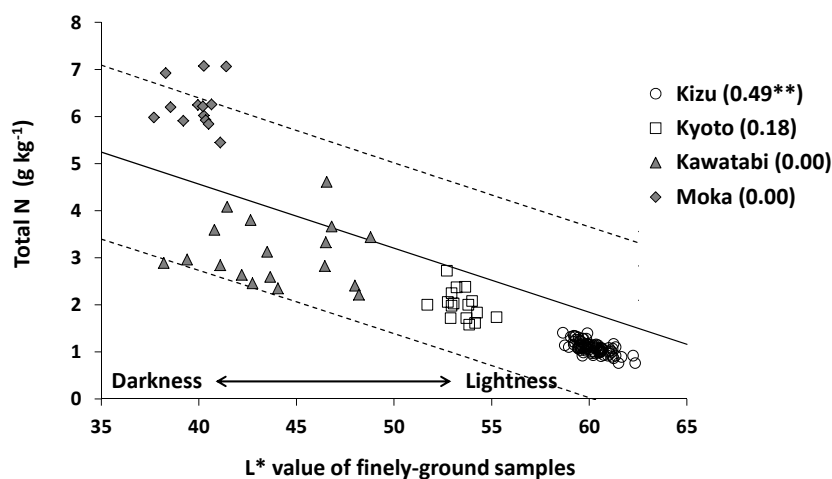


Figure 7. Relationship between the L* value of finely-ground samples and total N content at farm scale. The solid line in the figure shows the regression line obtained from the nationwide samples ($n = 147$) [20], and the broken lines indicate the prediction interval at 95% confidence. ** indicates statistical significance of the determination coefficient at $p < 0.01$.

Similar results were obtained in the relationship between POXC content and total N content. The determination coefficient was very high when all samples were included ($r^2 = 0.90$), whereas it ranged from 0.25 to 0.62 at farm scale (Figure 8). The percentage of POXC in total C ranged from 1.2% to 5.1%, suggesting that the amount of soil organic matter decomposed by the POXC method was

much smaller than the amount decomposed by the oxydol method (Figure 6). These results imply that measurements of soil lightness (L^* value) and POXC can be useful for estimating the total N content at national scale but not so much at farm scale.

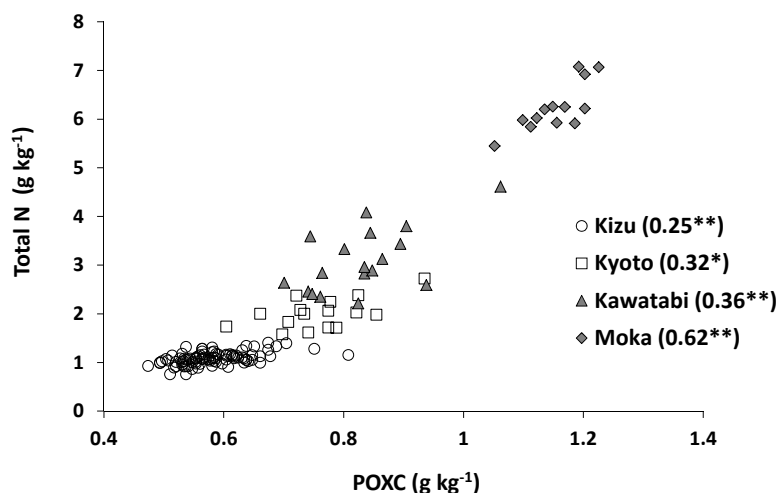


Figure 8. Relationship between POXC (potassium permanganate-oxidizable organic C) content and total N content at farm scale. ** and * indicate statistical significance of the determination coefficients at $p < 0.01$ and $p < 0.05$, respectively.

4. Discussion

4.1. Merits and Limitations of the Oxydol Method

The overall merit of the oxydol method lies in its simplicity because it was designed for non-specialists having no access to laboratory facilities. For example, it is easy to purchase a 3% H_2O_2 solution used for disinfecting minor cuts and abrasions. Filtration of the soil extracts prior to their measurement is not necessary. A stick-type EC sensor (HI98331 Gro Line Soil Test, Hanna Instruments) costs less than 200 USD. Except the commercial H_2O_2 solution, reagents and distilled water are not necessary for any part of the procedure. It takes only 1–2 min per sample to measure EC (H_2O_2) [9], and many samples (e.g., 136 samples in this study) can be handled in a single batch by one person. The soil extracts can be safely disposed of without special treatment. These features are advantages over the conventional methods such as the POXC method. Thus, the oxydol method can be used by farmers as one of the effective decision-support tools for site-specific application of N fertilizer to paddy fields. Moreover, the method is potentially useful for scientific investigation, especially for an initial screening of samples to identify appropriate samples prior to detailed analysis by a conventional method.

On the other hand, the biggest weakness of the oxydol method lies in its limited applicability to some soils. Mikutta et al. [27] reviewed previous publications and reported that H_2O_2 is less effective than sodium hypochlorite and disodium peroxodisulfate for decomposition of soil organic matter. For example, according to Leifeld and Kögel-Knabner [28], the H_2O_2 treatment used for soil texture analysis could not remove 15–30% of the N from the solid phase of 20- μ m sieved agricultural soils due to the presence of organic C resistant to the H_2O_2 treatment (6–17%) as well as NH_4^+ fixed strongly in 2:1 type clay minerals. In our case, the percentage of H_2O_2 -mineralizable organic N in total N was relatively low for volcanic ash soils (Figure 6), and this is probably because of the stabilization of soil organic matter with active Al derived from non-crystalline clay minerals, such as allophane and imogolite, in volcanic ash soil. Furthermore, for soils containing manganese oxides and alkaline soils containing carbonates, less than 30% of the organic matter might be decomposed because the H_2O_2 itself would be rapidly decomposed to water and oxygen [27]. Thus, the applicability of the oxydol

method is limited to neutral to acidic soils that are not rich in manganese oxides and highly stabilized organic matter.

The EC (H_2O_2) value can also be influenced by the presence of cations other than NH_4^+ . In a previous study, we measured the concentration of cations in the H_2O_2 extracts of 83 samples collected at national scale. Then, the determination coefficient between EC (H_2O_2) and NH_4^+ ($r^2 = 0.83$) was much higher than that between EC (H_2O_2) and other cations (Ca^{2+} , Mg^{2+} , K^+ , or Na^+) ($r^2 < 0.4$) [9]. For the samples analyzed in this study, the EC (1:5 *w/v*) values of the Kizu samples were relatively high (Table 1), but the EC (1:20 *w/v*) measured at the beginning of the H_2O_2 extraction (0.04 mS cm^{-1} on average) was much lower than the EC (H_2O_2) value (0.48 mS cm^{-1} on average). These results suggest that, as far as surface paddy soils in Japan are concerned, the NH_4^+ produced by the H_2O_2 extraction is a primary factor affecting EC (H_2O_2) (Figure 5).

From a technical aspect, the oxydol method requires an incubator or an air conditioner to keep the temperature at 25 °C during soil extraction, and the extraction period (40 h) is much longer than the POXC method [24]. In addition, it is important to select a suitable EC electrode. The EC sensors used in this study (HI98331 Soil Test EC tester; Handy SC meter TCX-999i with CX90CS) did not cause the H_2O_2 to decompose. In contrast, a platinum-black electrode attached to a conventional EC sensor causes rapid decomposition of H_2O_2 into water and oxygen, and the oxygen bubbles on the electrode surface may decrease the measured EC (H_2O_2) values. In this study, when the EC of the H_2O_2 solutions was measured with a conventional EC sensor (ES-51, Horiba, Ltd., Kyoto, Japan), the EC values were decreased by the presence of oxygen bubbles to 52–86% of those listed in Table 5.

4.2. Applicability of the Method to Farm-Scale Samples

By examining the similarities and differences between scientists' and farmers' evaluation of soil fertility, Yageta et al. [29] pointed out that scientists and farmers observe soil at different spatial scales, reflecting different direction of interests. Scientists tend to evaluate soil fertility at regional, national and global scales based on quantitative soil data, whereas farmers tend to evaluate soil fertility at farm, field and plot scales based on their local soil experience [29]. As the oxydol method was designed for non-specialists including farmers, we evaluated the method's applicability to soil samples collected at the farm scale.

At all farms examined, soil total N content varied considerably (Table 3), and the EC (H_2O_2) was positively correlated with the total N content (Figure 4). These results suggested that the oxydol method is useful for estimating farm-scale variations in soil total N. The slopes and intercepts of the regression lines fitted to the farm-scale data were different from those fitted to nationwide data (Table 4). Thus, site-specific calibration is desirable to improve the estimation accuracy at a particular farm, especially at farms having volcanic ash soil such as Moka. The POXC method was equally effective at farm scale (Figure 8), whereas soil lightness (L^* value) was significantly correlated with the total N content only at Kizu (Figure 7). At Kizu, the maximum difference of the L^* value among the samples was only 3.7, which is less than half of the L^* value difference between the neighboring color chips in a Munsell color chart (about 10). Thus, the soil color variations at Kizu were too small to evaluate visually by using a Munsell color chart.

At farm scale, the determination coefficients between EC (H_2O_2) and mineralized N were 0.23, 0.40, 0.18, and 0.53 at Kizu, Kyoto, Kawatabi, and Moka, respectively. Although these values were statistically significant at $p < 0.01$ except Kawatabi, they were consistently lower than the corresponding values between EC (H_2O_2) and total N (Figure 4). This result suggests that the oxydol method is most suitable for estimating the total N content in soil. Similar results were observed for the samples collected at national scale. Some of the soil samples analyzed by Moritsuka et al. [9] were also analyzed by Sano et al. [5]. For such samples collected throughout Japan ($n = 31$), the determination coefficient between the EC (H_2O_2) and soil N properties decreased from 0.64 (for total N) to 0.57 (for 0.33 mol L^{-1} potassium permanganate-oxidizable organic N) and 0.20 (for 0.067 mol L^{-1} phosphate

buffer-extractable organic N). Among the various forms of soil N, total N showed the highest correlation with EC (H_2O_2).

4.3. Possibility of Using Different Types of H_2O_2 Solutions

The sensitivity of the oxydol method depends on how much soil EC increases during the extraction with 3% H_2O_2 . The minimum scale of the EC sensor used in this study was 0.01 mS cm^{-1} , and the soil to extractant ratio for oxydol method was set at 1:20 (*w/v*). Because the EC of ammonium sulfate or ammonium chloride containing NH_4^+ at 100 mgN L^{-1} is slightly greater than 1 mS cm^{-1} , the minimum scale for detecting the total N content is estimated to be larger than 0.02 g kg^{-1} , under the assumption that soil N is fully mineralized to NH_4^+ by the H_2O_2 treatment. In the case of surface paddy soils in Japan, only 5 of the 83 samples had EC (H_2O_2) values higher than 1 mS cm^{-1} [9]. EC (H_2O_2) values of the samples analyzed in this study were always lower than 1 mS cm^{-1} . In contrast, the EC of the 3% H_2O_2 tested was higher than 0.2 mS cm^{-1} for Essential Oxygen and AGA (Table 5). Accordingly, EC (H_2O_2) may not reflect the concentration of NH_4^+ in the extracts, if the background EC is increased by the H_2O_2 stabilizer (Table 6).

The accuracy of estimation of soil total N is another important issue to consider. The color of soil extracts differed among the type of H_2O_2 solution used (Figure S1). In particular, the soil extracts obtained with Essential Oxygen were much darker than the others, which implies that more soil organic matter remained undecomposed in the extracts with Essential Oxygen. Except for Essential Oxygen, EC (H_2O_2), values obtained with other different solutions were relatively comparable with each other (Table 6).

Overall, these results suggest that an H_2O_2 solution with a relatively low EC ($<0.2 \text{ mS cm}^{-1}$) can be effectively used for the oxydol method.

5. Conclusions

The oxydol method was designed to be as simple as possible, so it can be used by non-specialists having no access to laboratory facilities. In this study, we demonstrated that the EC of commercially available 3% H_2O_2 solutions varied widely with the amount and type of stabilizers added, and that the oxydol method can be useful for estimating the farm-scale variation in soil total N, provided a less stabilized H_2O_2 is used for soil extraction. We have so far evaluated the method's applicability to paddy soils sampled from Japan, and it would be worth evaluating its applicability to paddy soils in other countries and to soils under different land use.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4395/10/1/40/s1>, Figure S1: Color differences among extracts of the same soil sample treated with different H_2O_2 solutions. From left to right, Wako, Showa, Kozakai, Kenei, Swan, Essential Oxygen, and AGA.

Author Contributions: N.M. designed and carried out the experiments, and wrote the initial draft; H.S., R.T., Y.T., and H.H. managed their farm resources, supported soil sampling, and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by JSPS KAKENHI Grant No. JP15K00634 and SATREPS (JST/JICA) Grant No. JPMJSA1608.

Acknowledgments: We are grateful to the staff of Laboratory of Soil Chemistry, Kyoto Prefectural University, and the staff of Laboratory of Soil Science, Kyoto University, for allowing us to use an NC analyzer and an atomic absorption spectrometer, respectively, and to Hisashi Kagata, Kyoto University Farm, for providing us the information of land reclamation and field management at Kizu.

Conflicts of Interest: The authors declare that there is no conflict of interest.

References

1. Yoshida, S. *Fundamentals of Rice Crop Science*; The International Rice Research Institute: Los Baños, Philippines, 1981.
2. Griffin, T.S. Nitrogen Availability. In *Nitrogen in Agricultural Systems, Agronomy Monograph*; no. 49; Schepers, J.S., Raun, W.R., Eds.; ASA-CSSA-SSSA: Madison, WI, USA, 2008; pp. 613–646.
3. Gianello, C.; Bremner, J.M. Comparison of chemical methods of assessing potentially available organic nitrogen in soil. *Commun. Soil Sci. Plant Anal.* **1986**, *17*, 215–236. [[CrossRef](#)]
4. Sharifi, M.; Zebarth, B.J.; Burton, D.L.; Grant, C.A.; Hajabbasi, M.A.; Abbassi-Kalo, G. Sodium hydroxide direct distillation: A method for estimating total nitrogen in soil. *Commun. Soil Sci. Plant Anal.* **2009**, *40*, 2505–2520. [[CrossRef](#)]
5. Sano, S.; Yanai, J.; Kosaki, T. Relationships between labile organic matter and nitrogen mineralization in Japanese agricultural soils with reference to land use and soil type. *Soil Sci. Plant Nutr.* **2006**, *52*, 49–60. [[CrossRef](#)]
6. Sahrawat, K.L. Nitrogen mineralization in lowland rice soils: The role of organic matter quantity and quality. *Arch. Agron. Soil Sci.* **2010**, *56*, 337–353. [[CrossRef](#)]
7. Bremner, J.M. Nitrogen-Total (Chapter 37). In *Methods of Soil Analysis, Part 3, Chemical Methods*; SSSA Book Series no. 5; Sparks, D.L., Ed.; SSSA-ASA: Madison, WI, USA, 1996; pp. 1085–1121.
8. Christianson, C.B.; Holt, L.S. Rapid digestion procedure for the determination of total nitrogen and nitrogen-15 content of soils. *Soil Sci. Soc. Am. J.* **1989**, *53*, 1917–1919. [[CrossRef](#)]
9. Moritsuka, N.; Matsuoka, K.; Katsura, K.; Sano, S.; Yanai, J. Estimation of total nitrogen content in surface paddy soils by measuring their electrical conductivity after hydrogen peroxide treatment. *Jpn. J. Soil Sci. Plant Nutr.* **2017**, *88*, 327–335. (In Japanese)
10. Robinson, G.W. Note on the mechanical analysis of humus soils. *J. Agric. Sci.* **1922**, *12*, 287–291. [[CrossRef](#)]
11. Robinson, W.O. The determination of organic matter in soils by means of hydrogen peroxide. *J. Agric. Res.* **1927**, *34*, 339–356.
12. Harada, Y.; Inoko, A. The oxidation products formed from soil organic matter by hydrogen peroxide treatment. *Soil Sci. Plant Nutr.* **1977**, *23*, 513–521. [[CrossRef](#)]
13. Guan, G.; Marumoto, T.; Nishiyama, M. Influence of iron and aluminium oxide on the amount of microbial biomass and decomposition of rice straw in H₂O₂-treated soil. *Jpn. J. Soil Sci. Plant Nutr.* **1999**, *70*, 171–176. (In Japanese)
14. Eith, C.; Kolb, M.; Seubert, A.; Viehweger, K.H. *Practical Ion Chromatography. An Introduction*; Metrohm Monograph, Metrohm Ltd.: Herisau, Switzerland, 2001.
15. Schmuicker-Castner, J.; Desai, D. Rheology modification of hydrogen peroxide-based applications using cross-linked polyacrylic acid polymers. *Int. J. Cosmet. Sci.* **1999**, *21*, 313–325. [[CrossRef](#)] [[PubMed](#)]
16. NARO. Digital Soil Map of Japan. 2019. Available online: <https://soil-inventory.dc.affrc.go.jp/> (accessed on 15 November 2019). (In Japanese)
17. Bundy, L.G.; Meisinger, J.J. Nitrogen Availability Indices (Chapter 41). In *Methods of Soil Analysis, Part 2, Microbiological and Biochemical Properties*; SSSA Book Series no. 5; Weaver, R.W., Ed.; SSSA: Madison, WI, USA, 1994; pp. 951–984.
18. Moritsuka, N.; Izawa, G.; Katsura, K.; Matsui, N. Simple method for measuring soil sand content by nylon mesh sieving. *Soil Sci. Plant Nutr.* **2015**, *61*, 501–505. [[CrossRef](#)]
19. Mulvaney, R.L. Nitrogen—Inorganic forms (Chapter 38). In *Methods of Soil Analysis, Part 3, Chemical Methods*; SSSA Book Series no. 5; Sparks, D.L., Ed.; SSSA-ASA: Madison, WI, USA, 1996; pp. 1123–1184.
20. Moritsuka, N.; Matsuoka, K.; Katsura, K.; Sano, S.; Yanai, J. Soil color analysis for statistically estimating—Total carbon, total nitrogen and active iron contents in Japanese agricultural soils. *Soil Sci. Plant Nutr.* **2014**, *60*, 475–485. [[CrossRef](#)]
21. Japanese Society of Pedology. *A Handbook for Soil Survey. Revised Edition*; Japanese Society of Pedology: Hakuyusha, Tokyo, 1997; p. 169. (In Japanese)
22. van Leeuwen, M.M.W.J.; Heuvelink, G.B.M.; Wallinga, J.; de Boer, I.J.M.; van Dam, J.C.; van Essen, E.A.; Moolenaar, S.W.; Verhoeven, F.P.M.; Stoorvogel, J.J.; Stoof, C.R. Visual soil evaluation: Reproducibility and correlation with standard measurements. *Soil Tillage Res.* **2018**, *178*, 167–178. [[CrossRef](#)]

23. Weil, R.R.; Islam, K.R.; Stine, M.A.; Gruver, J.B.; Samson-Liebig, S.E. Estimating active carbon for soil quality assessment: A simplified method for laboratory and field use. *Am. J. Altern. Agric.* **2003**, *18*, 3–17. [[CrossRef](#)]
24. Culman, S.W.; Freeman, M.; Snapp, S. *Procedure for the Determination of Permanganate Oxidizable Carbon*; Kellogg Biological Station, Michigan State University: Hickory Corners, MI, USA, 2014; Available online: <https://lter.kbs.msu.edu/protocols/133> (accessed on 5 June 2019).
25. MAFF Agricultural Production Bureau. *Summary on the National Soil Survey Programs in Japanese Agricultural Land from 1979 to 2003*; Agricultural Production Bureau, Ministry of Agriculture, Forestry and Fisheries: Tokyo, Japan, 2008; p. 485. (In Japanese)
26. Moritsuka, N.; Matsuoka, K.; Katsura, K.; Yanai, J. Farm-scale variations in soil color as influenced by organic matter and iron oxides in Japanese paddy fields. *Soil Sci. Plant Nutr.* **2019**, *65*, 166–175. [[CrossRef](#)]
27. Mikutta, R.; Kleber, M.; Kaiser, K.; Jahn, R. Organic matter removal from soils using hydrogen peroxide, sodium hypochlorite, and disodium peroxodisulfate. *Soil Sci. Soc. Am. J.* **2005**, *69*, 120–135. [[CrossRef](#)]
28. Leifeld, J.; Kögel-Knabner, I. Organic carbon and nitrogen in fine soil fractions after treatment with hydrogen peroxide. *Soil Biol. Biochem.* **2001**, *33*, 2155–2158. [[CrossRef](#)]
29. Yageta, Y.; Osbahr, H.; Morimoto, Y.; Clark, J. Comparing farmers' qualitative evaluation of soil fertility with quantitative soil fertility indicators in Kitui County, Kenya. *Geoderma* **2019**, *344*, 153–163. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Phosphorus uptake of rice plants is affected by phosphorus forms and physicochemical properties of tropical weathered soils

Tomohiro Nishigaki · Yasuhiro Tsujimoto ·
Seheno Rinasoa · Tovohery Rakotoson ·
Andry Andriamananjara · Tantely Razafimbelo

Received: 15 August 2018 / Accepted: 31 October 2018 / Published online: 12 November 2018
© The Author(s) 2018

Abstract

Aims Phosphorus (P) deficiency is a major constraint for rice production in the tropics. Field-specific P management is key for resource-limited farmers to increase yields with minimal inputs. We used soil P fractionation analysis to identify the relevant factors controlling P uptake and the responses to P fertilization of rice in flooded and highly weathered soils.

Methods Phytometric pot-based experiments and a modified Hedley fractionation analysis were repeated for soils from extensive regions and from geographical adjacent fields in Madagascar.

Results Large field-to-field variations in indigenous P supply from soils (total P uptake of rice when P is

omitted) and fertilizer-P recovery efficiencies (increased P uptake when P is applied) were observed not only for soils with various geological backgrounds but also for soils from adjacent fields. Regression models indicated that the indigenous P supply in soils was largely controlled by readily available inorganic and organic P pools ($r^2 = 0.72$), whereas fertilizer-P recovery efficiencies were controlled by the abundance of oxalate-extractable aluminum and iron in soils ($r^2 = 0.81$).

Conclusions Spatial heterogeneity even within adjacent fields leads to benefits from field-specific fertilizer management based on indigenous P supply from soils and fertilizer-P recovery efficiencies evaluated by different soil properties.

Responsible Editor: Terry James Rose.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11104-018-3869-1>) contains supplementary material, which is available to authorized users.

T. Nishigaki · Y. Tsujimoto (✉)
Crop, Livestock and Environment Division, Japan International
Research Center for Agricultural Sciences, 1-1 Owashi, Tsukuba,
Ibaraki 305-8686, Japan
e-mail: tsjmt@affrc.go.jp

T. Nishigaki
e-mail: nishigaki@affrc.go.jp

S. Rinasoa · T. Rakotoson · A. Andriamananjara ·
T. Razafimbelo
Laboratoire des Radio-Isotopes, Université d'Antananarivo, BP
3383, Route d'Andraisoro 101, Antananarivo, Madagascar

S. Rinasoa
e-mail: seheno.rinasoa@gmail.com

T. Rakotoson
e-mail: tovohery.rakotoson@gmail.com

A. Andriamananjara
e-mail: njaraandry1@gmail.com

T. Razafimbelo
e-mail: tantely.razafimbelo@gmail.com

Keywords Hedley fractionation · Nutrient omission trial · Phosphorus uptake · Rice · Sub-Saharan Africa

Abbreviations

Alo	oxalate-extractable Al
ANOVA	Analysis Of Variance
Δ Biomass	the difference in biomass of rice plant between the +NP and +N treatments
Feo	oxalate-extractable Fe
ICP	Inductively Coupled Plasma
Pi	inorganic P in each fraction
Po	organic P in each fraction
P _{i+o}	total P in each fraction
Δ P uptake	the difference in P uptake of rice plant between the +NP and +N treatments
SSA	Sub-Saharan Africa
TEB	Total Exchangeable Bases

Introduction

Reduced availability of phosphorus (P) in soils results in reduced biomass and grain yield of rice grown in highly weathered soils (Dogbe et al. 2015; Fageria and Baligar 1997; Koné et al. 2011, 2013). Therefore, improvement of P availability in tropical weathered soils and enhancement of P uptake in rice plants is crucial to meet the rising demand for rice in Sub-Saharan Africa (SSA) (Nziguheba et al. 2016; Saito et al. 2015b; van Oort et al. 2015). Tropical weathered soils are typically acidic and rich in iron (Fe) and aluminum (Al) (hydr)oxides, which result in a substantial capacity to sorb phosphate (Balemi and Negisho 2012). These features may partly explain why the increase of P availability due to the reduced condition is less expected in highly weathered soils, while P availability generally tends to be less-limiting for paddy rice than for upland crops in equivalent soils (Fageria et al. 2011). There is also concern regarding the non-renewable use of global rock-phosphate reserves (Van Vuuren et al. 2010), and fertilization is usually neither economical nor readily available to subsistence farmers in SSA (Nziguheba et al. 2016). Hence, under such circumstances, it is necessary for farmers to adopt strategies to improve the P uptake of rice plants, through 1) selecting a field with a high indigenous P supply capacity, and/or 2) increasing fertilizer-P recovery efficiency.

In rice fields, in which soil fertility varies spatially, site-specific soil management is expected to benefit rice

production (Saito et al. 2015a; Schmitter et al. 2010; Yanai et al. 2012). This may be particularly true for the fields of smallholder farmers in SSA, where the response to applied nutrients varies over small distances and is governed mostly by the influence of past management practices (Kihara et al. 2016; Schut et al. 2018; Zingore et al. 2011). In addition, our previous studies indicated that the forms of soil P were important for determining P availability in the tropics and that soil P forms can be affected by land management practices as well as soil properties (Nishigaki et al. 2018; Sugihara et al. 2012). Therefore, it is necessary to evaluate the spatial variation of soil P forms in rice fields on a regional or community scale and to elucidate the response of P uptake in rice to the P forms in soils.

The P fractionation method of Hedley et al. (1982) has been widely used to characterize soil P forms based on their availability, with the fundamental assumption being that extractants of varying strength estimate P fractions of differing availability. Nevertheless, P fractions separated by the same sequential method are not of equal availability to plants in all soils (Guo et al. 2000), and rice plants draw P from a continuum of chemically extracted fractions that are assumed to have different plant P availability (Zhang et al. 2006). In low-input systems, in which fertilizer-P additions are very low or absent, pools of P that are less available seem to act as a buffering pool for labile inorganic P, particularly in highly weathered soils (Beck and Sanchez 1994; Guo et al. 2000). In addition, previous studies argued that organic P may play an important role in supplying available P in unfertilized soils (George et al. 2018; Guo et al. 2000; Tiessen et al. 1992). However, little is known about the significance of these different P pools in the supply of P to rice plants grown on weathered soils.

The focus for low-P soils in the tropics is on increasing fertilizer-P recovery efficiency and preventing the accumulation of recalcitrant soil P (Menezes-Blackburn et al. 2017). In the normal pH range of agricultural soils of the tropics, P is mainly bound to Fe- and Al-(hydr)oxides, with the sorption reactions including the precipitation of metal phosphates (Haynes and Mokolobate 2001). Alternatively, it is widely reported that soluble organic constituents derived from the application of organic amendments can enhance P solubility and mobility, as well as compete with P for sorption sites

(Chassé and Ohno 2016; Guppy et al. 2005; Yan et al. 2013). It is, therefore, necessary to take into account not only the soil P forms, but also other soil physicochemical properties that inhibit or enhance P absorption by rice plants. Thus far, few studies have shown how rice P uptake is regulated by soil P forms and soil physicochemical properties in highly weathered soils.

Relative biomass or nutrient uptake (ratio with the omission of a target nutrient relative to its full application) has been used in previous omission trials as an index to evaluate nutrient deficiency (Kihara et al. 2016; Shehu et al. 2018). Still, there are debates about this index and whether it can clearly identify the influential factors with respect to high P-fixing capacity, the originally high P-supplying capacity, and deficiencies of other nutrients. In this study, we aimed to distinguish the indigenous soil P-supplying capacity and the fertilizer-P recovery efficiency using different parameters, i.e., total P uptake of rice plants in the P-omitted treatment and increased P uptake with P application, respectively. Our objective was, therefore, to evaluate P uptake of rice plants with special reference to soil P forms and soil physicochemical properties in tropical weathered soils. Of particular interest was to examine the variations in P forms in a range of rice-field soils and to determine the factors that have a substantial effect on P uptake in rice plants.

Materials and methods

Soil sampling and pot-based experiments

Two sets of pot-based experiments were conducted with lowland and upland field soils collected from a wide region of the central highland of Madagascar ($n = 35$, Exp. 1) and those collected within a relatively small area of one village ($n = 16$, Exp. 2) (Fig. 1). These two sets of experiments were to confirm the applicability of the results at the landscape level in soils that have been rather affected by geological changes and at the field-scale level in soils rather affected by individual farmers' management practices. Experimental soils were taken from 0 to 15 cm depth as composites of four to five cores in each field and divided into two parts for the pot-based experiments and soil analysis, respectively. Each of Exp. 1 and Exp. 2 was conducted in a greenhouse at the Laboratoire des Radio-Isotopes, University of Antananarivo, in Madagascar (18°53'56.0"S, 47°33'01.2"E, 1222 m alt.) during September to October 2016 (Exp. 1) and December 2016 to

January 2017 (Exp. 2). The daily mean temperatures throughout the growing periods ranged from 18.2 to 25.6 °C and from 21.6 to 25.9 °C in Exp. 1 and Exp. 2, respectively (Watchdog 2475, Spectrum Technologies Ltd.).

Each of the collected soils was put into a 1-l plastic pot (13 cm diameter, 15 cm height). Each pot contained 1 kg of air-dried and sieved (4 mm) soil. In each experiment, three sets of different fertilizer treatments were established with two replicates: 1, *Cont* (no fertilizer application); 2, +*N* (0.2 g pot⁻¹ and 0.3 g pot⁻¹ of N as NH₄NO₃ in Exp. 1 and Exp. 2, respectively); 3, +*NP* (0.2 g pot⁻¹ and 0.3 g pot⁻¹ of both N as NH₄NO₃ and P₂O₅ as KH₂PO₄ in Exp. 1 and Exp. 2, respectively). Potassium was applied to all the pots, including the *Cont* treatment at the rate of 0.2 g pot⁻¹ and 0.3 g pot⁻¹ of K₂O as KCl in Exp. 1 and Exp. 2, respectively. Each nutrient was uniformly incorporated into soils one day after being flooded and one day prior to transplanting. Then, two 20-day-old seedlings of a local rice cultivar, X265, grown in free-nutrient sand were transplanted to each pot. The pots were continuously flooded at the depth of 2–5 cm with distilled water throughout the growing periods. The weeds were removed manually, and no specific pest management was required.

Plant analysis

The plants were harvested at the soil surface, 34 days after transplanting for both Exp. 1 and Exp. 2. Above-ground biomass was determined after oven drying at 70 °C to a constant weight. Each plant sample was ground into a fine powder using a high-speed vibrating sample mill (Model T1–100, Heiko Co. Ltd., Fukushima, Japan). Then, the plant P concentration was determined with the molybdate blue method (Murphy and Riley 1962) after dry-ashing at 550 °C for 2 h and digestion with 0.5 M HCl. The plant P uptake (mg P pot⁻¹) was calculated as the product of aboveground biomass and P concentration of plants (total P content).

Soil analysis

The collected soils were air-dried and sieved to 2 mm for subsequent soil analysis. Soil particle size distribution was determined with the wet-sieving and pipet method (Gee and Bauder 1986). Soil pH was determined in deionized water at a soil-to-solution ratio of 1:2.5. Exchangeable cations (K⁺, Na⁺, Ca²⁺, and Mg²⁺) were

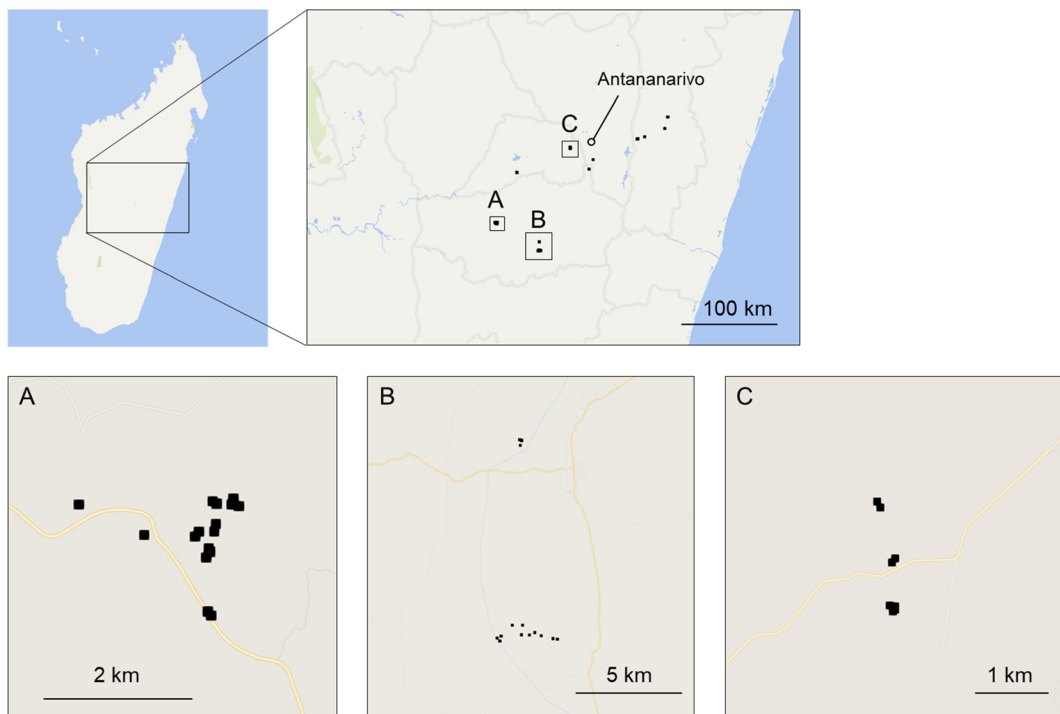


Fig. 1 Map of soil sampling sites. Soils for Exp. 2 were collected from the sites in sub-map A, whereas soils for Exp. 1 were from the other sites

measured following ISRIC protocol (2002), and total exchangeable base was calculated as the sum of those four cations. Total carbon (C) and N were quantified using the dry combustion method with an NC analyzer (Sumigraph NC-220F, SCAS, Japan). Soil samples were digested in 60% HClO₄ following Kuo (1996), and total P was determined with the molybdovanadate method (Kitson and Mellon 1944). Extractable Al, Fe, and P contents were determined with the acid ammonium oxalate method (Alo, Feo, and Oxalate P, respectively) as described by Courchesne and Turmel (2008). The concentrations of Al, Fe, and P in the oxalate extraction were measured with an inductively coupled plasma mass spectrometer (ICPE-9000, Shimadzu, Japan).

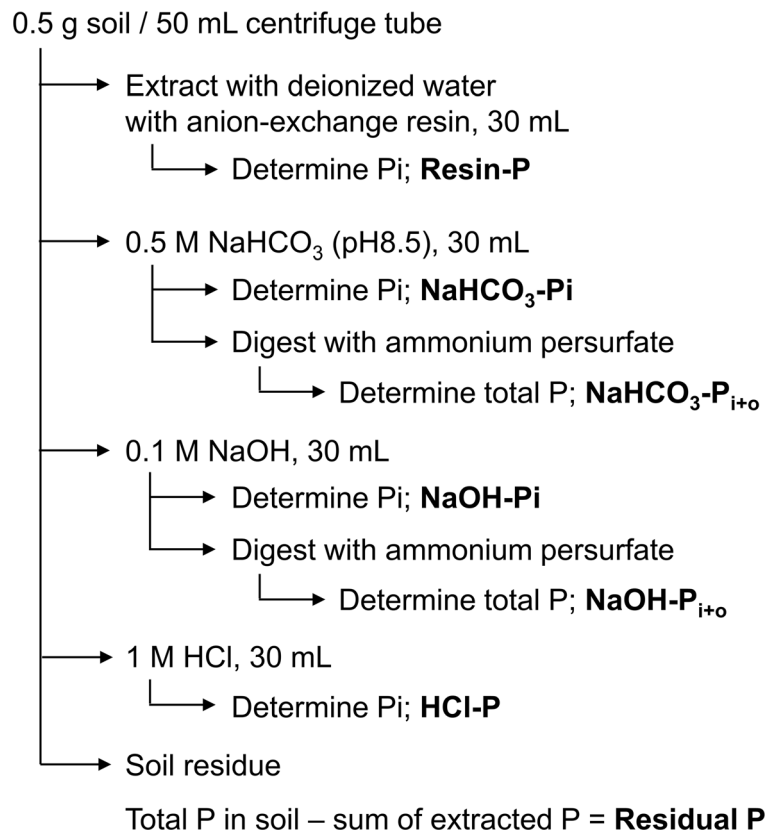
Soil P was sequentially fractionated using a modification of the Hedley method (Tiessen and Moir 2007; Sugihara et al. 2012) (Fig. 2). Briefly, 0.5 g of soil was placed in a 50 mL centrifuge tube and sequentially extracted with 30 mL of each of the extractant solutions, which were added in the following order: deionized water (deionized water with two anion-exchange resins in the bicarbonate form; Resin-P), 0.5 M NaHCO₃ (pH 8.5) (NaHCO₃-P), 0.1 M NaOH (NaOH-P), and 1 M HCl (HCl-P). Each extraction was performed for 16 h using a horizontal shaker followed by centrifugation at 2500×g for

20 min at room temperature, and the supernatant was filtered (5C, ADVANTEC) and the filtrate collected to measure the concentration of inorganic P (Pi) using the molybdate-ascorbic acid method (Murphy and Riley 1962) after pH adjustment using *p*-nitrophenol. Each total P (P_{i+o}) in NaHCO₃ and NaOH extracts was determined after digestion with ammonium persulfate in an autoclave at 103.5 kPa and 120 °C for 60 min, and the concentration was determined with the method of Murphy and Riley (1962). The organic P (Po) was calculated as the difference between P_{i+o} and Pi. The difference between total P in soils and the sum of Pi and Po in all the extracted fractions was defined as Residual P.

Statistics

Statistical analyses were performed using JMP11 software (JMP11.0 Windows, SAS Institute Inc.). First, two-way analysis of variance (ANOVA) was conducted to determine the individual and interaction effects of soils (S) and fertilizer treatments (T) on the above-ground biomass and total P uptake of rice plants. Second, the obtained above-ground biomass and total P uptake data for both Exp. 1 and Exp. 2 were standardized into the mean value at 0 with the sample variation at

Fig. 2 Simple flow chart of the sequential P extraction according to a modified Hedley method (Tiessen and Moir 2007; Sugihara et al. 2012)



1 to simultaneously handle the results of two experiments grown under different environmental conditions. Then, a stepwise regression analysis was repeated by applying the K-fold ($K = 5$) cross validation to develop robust models over two experiments that explained the variations in: 1) total P uptake in the $+N$ treatment as an index of indigenous soil P-supplying capacity, and 2) differences in P uptake between the $+N$ and $+NP$ treatments as an index of fertilizer-P recovery efficiency, respectively. The candidate explanatory factors included clay content, total C, N, and P contents, soil pH, Alo, Feo, Alo + 0.5 Feo, Resin-P, $\text{NaHCO}_3\text{-Pi}$, $\text{NaHCO}_3\text{-Po}$, NaOH-Pi , NaOH-Po , and HCl-P . In the stepwise process, the ‘selection’ and ‘removal’ of factors was controlled with an F-value of $P < 0.10$ at each step. The step ended when no remaining candidate factors produced an eligible F-value. Then, a multiple linear regression model was developed using the selected variables. At the end, standardized partial regression coefficients were calculated to assess the effect size of each selected variable. In addition, a *t*-test for the simple regression coefficient was conducted to identify the interrelationship of these soil properties.

Results

Plant biomass and P uptake

Compared with Exp. 1, Exp. 2 yielded higher values for both aboveground biomass and total P uptake; this was attributable to the higher temperature and rate of nutrient application as well as greater amount of solar radiation during the growing period (Table 1). ANOVA demonstrated highly significant effects of experimental soil, fertilizer treatment, and most importantly of their interactions on aboveground biomass and P uptake for both Exps. 1 and 2. Consequently, there were large variations in biomass and P uptake within each experiment. In a comparison of F-values, the magnitude of variations as affected by different soils and treatments was almost equivalent between the two experiments even though the soils for Exp. 2 were collected from adjacent fields in a small area of one village.

The aboveground biomass in the $+N$ treatment ranged from 0.20 to 2.01 g pot^{-1} and 1.38 to 6.40 g pot^{-1} in Exp. 1 and Exp. 2, respectively (Fig. 3a). The effect of P application, as measured by the difference in biomass between the $+NP$ and $+N$ treatments ($\Delta\text{Biomass}$), was

Table 1 Aboveground biomass and P uptake of rice plants as affected by different soils and fertilizer treatments

		Aboveground biomass (g pot ⁻¹)			Plant P uptake (mg P pot ⁻¹)		
		mean	s.d.		mean	s.d.	
Exp. 1	Cont	0.48 c	0.33		0.63 b	0.69	
	+N	0.75 b	0.48		0.81 b	0.80	
	+NP	1.89 a	0.57		3.29 a	1.34	
Exp. 2	Cont	1.36 c	0.53		1.89 c	1.78	
	+N	3.37 b	1.25		3.20 b	2.84	
	+NP	5.60 a	0.64		10.36 a	3.59	
ANOVA summary		Aboveground biomass			Plant P uptake		
		df	SS	F-value	df	SS	F-value
Exp. 1	Soil (S)	35	26.1	12.5 ^{***}	16	134.2	14.6 ^{***}
	Treatment (T)	2	66.7	556.7 ^{***}	2	283.4	538.6 ^{***}
	S × T	70	14.1	3.4 ^{***}	32	69.8	3.8 ^{***}
Exp. 2	Soil (S)	16	31.9	8.7 ^{***}	16	428.1	43.5 ^{***}
	Treatment (T)	2	296.3	645.9 ^{***}	2	1275.3	1036.2 ^{***}
	S × T	32	23	3.1 ^{***}	32	108.3	5.5 ^{***}

Mean values with different letters indicate significant differences based on Tukey HSD at $P < 0.05$

s.d., standard deviation ($n = 35$ and $n = 16$ for Exp. 1 and Exp. 2, respectively); df, degrees of freedom; SS, sum of squares

*** $P < 0.001$

positive for all soils whereas the increased rate varied greatly among soils, ranging from 0.08 to 2.33 g pot⁻¹ (Exp. 1) and from 0.40 to 4.72 g pot⁻¹ (Exp. 2). The same soil-to-soil or field-to-field variations were observed for the P uptake data in the +N treatment, and P uptake increased when P was applied (Fig. 3b). The P uptake without P application (+N treatment) ranged from 0.10 to 2.77 mg P pot⁻¹ (Exp. 1) and 0.95 to 12.97 mg P pot⁻¹ (Exp. 2). The increased P uptake when P was applied (ΔP uptake; difference between the +NP and +N treatments) ranged from 0.47 to 5.02 mg P pot⁻¹ (Exp. 1) and from 2.67 to 10.14 mg P pot⁻¹ (Exp. 2). Interestingly, there were large differences in the response to P application, i.e., Δ Biomass and ΔP uptake, even among the

soils that consistently had low biomass production or P uptake under +N treatment (P omitted).

Soil characteristics, total P and fractionated P

The original soils in Exp. 2 had poor nutrient status in terms of total C, N, and P compared with the soils in Exp. 1 (Table 2). Because the soils for Exp. 2 were collected from a small area within a single community, the standard deviation values for most soil properties were smaller in Exp. 2 than Exp. 1. The Alo and Alo + 0.5 Feo values were significantly greater for the soils of Exp. 1 than those of Exp. 2, whereas Feo was comparable between the experiments.

Fig. 3 Biomass (a) and P uptake (b) in the +N treatment and their responses to P application. Δ Biomass, the difference in biomass of rice plant between the +NP and +N treatments; ΔP uptake, the difference in P uptake of rice plant between the +NP and +N treatments

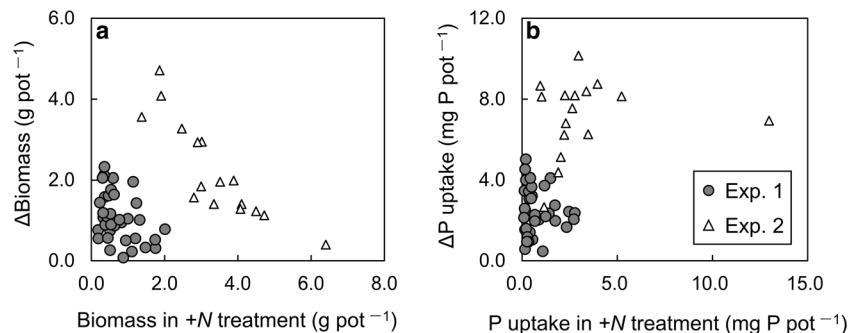


Table 2 Physicochemical properties of soils for pot-based experiments

		Exp. 1			Exp. 2			Total	
		(n = 35)			(n = 16)			(n = 51)	
Total C	(g C kg ⁻¹)	26	a	(11.0)	14	b	(5.2)	22	(11.1)
Total N	(g N kg ⁻¹)	2.0	a	(0.8)	1.2	b	(0.4)	1.7	(0.8)
Total P	(mg P kg ⁻¹)	1200	a	(720)	540	b	(190)	960	(670)
TEB	(cmol _c kg ⁻¹)	3.4		(3.2)	2.8		(2.4)	3.2	(3.0)
pH (H ₂ O)		5.6		(0.4)	5.6		(0.3)	5.6	(0.4)
Clay	(%)	32		(8.2)	31		(10.5)	32	(9.0)
Alo	(g Al kg ⁻¹)	4.3	a	(2.6)	1.4	b	(0.5)	3.4	(2.6)
Feo	(g Fe kg ⁻¹)	8.6		(5.0)	5.8		(4.2)	7.7	(5.0)
Alo + 0.5 Feo	(g kg ⁻¹)	8.6	a	(4.5)	4.3	b	(1.8)	7.2	(4.3)

The values given in parentheses are standard deviation. Different letters indicate a significant difference between the experiments (Tukey test, *P* < 0.05)

TEB, total exchangeable bases (sum of exchangeable Ca²⁺, Mg²⁺, Na⁺, and K⁺); Alo, oxalate-extractable Al; Feo, oxalate-extractable Fe

Total P concentration in all soils ranged from 200 to 3200 mg P kg⁻¹, with a mean value of 960 mg P kg⁻¹ (Fig. 4). The amount of P varied substantially among the fractions, with the NaOH-Pi value being greatest (260 mg P kg⁻¹ on average) followed by NaOH-Po (81 mg P kg⁻¹ on average). The HCl-P was generally low in most soils (median 8.0 mg P kg⁻¹), but some soils had quite high HCl-P values that increased with soil pH. The Resin-P, NaHCO₃-Pi, and NaHCO₃-Po values were lower than for the other fractions, namely 12, 19, and 18 mg P kg soil⁻¹ on average, respectively. The Oxalate P in all soils ranged from 49 to 1200 mg P kg soil⁻¹ (mean, 330 mg P kg soil⁻¹). The ratio of the sum of labile P (Resin-P, NaHCO₃-Pi, and NaHCO₃-Po) to total P was smaller in Exp. 2 (4.9%) than Exp. 1 (6.1%). Except for

the NaHCO₃-Po fraction, all fractions had a significant positive correlation with total P (Table 3).

The Alo values correlated significantly with Resin-P, NaHCO₃-Pi, and NaOH-Pi, whereas the Feo values correlated only relatively weakly in this respect (Table 3). The Alo + 0.5 Feo values also correlated significantly with Resin-P, NaHCO₃-Pi, and NaOH-Pi (Table 3). The amount of Oxalate P clearly corresponded to the total amount of extractable Pi in all the fractions (TPi) (Fig. 5).

Relationship between responses to fertilizer application and soil properties

Table 4 presents the results of step-wise regression analysis for P uptake in the +N treatment and for the

Fig. 4 Variation of total phosphorus and fractionated phosphorus among soil samples (n = 51). The cross in each box represents the mean, the central vertical bar shows the median, the box represents the interquartile range, the whiskers show the location of the most extreme data points that are still within a factor of 1.5 of the upper or lower quartiles, and the points are values that fall outside the whiskers

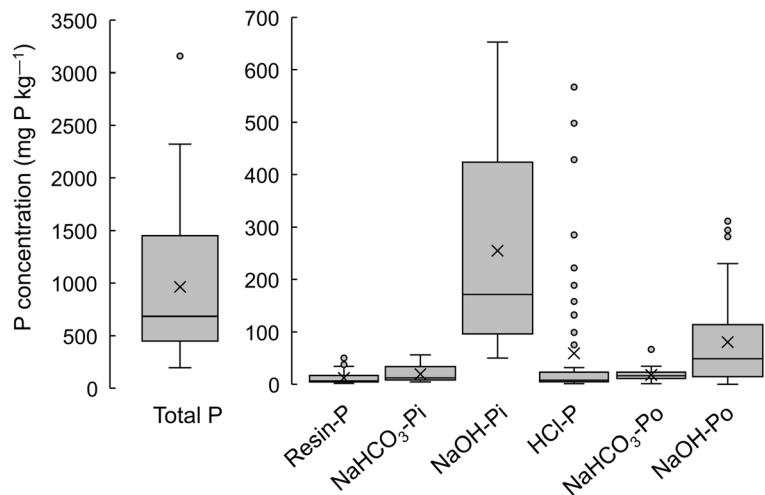


Table 3 Correlation matrix among soil physicochemical properties and fractionated phosphorus

	Total P	Oxalate P	Resin-P	NaHCO ₃ -Pi	NaOH-Pi	HCl-P	TPi	NaHCO ₃ -Po	NaOH-Po	TPo	Residual P
Alo	0.703 **	0.764 **	0.619 **	0.719 **	0.857 **	0.438	0.783 **	0.494	0.714 **	0.724 **	0.476
Feo	0.552 *	0.700 **	0.463	0.599 **	0.502	0.496	0.572 *	-0.059	0.425	0.305	0.473
Alo + 0.5Feo	0.733 **	0.854 **	0.632 **	0.769 **	0.796 **	0.544 *	0.792 **	0.259	0.655 **	0.604 **	0.553 *
pH (H ₂ O)	0.374	0.456	0.342	0.357	0.201	0.618 **	0.418	-0.447	-0.093	-0.090	0.368
Clay	-0.145	-0.256	-0.246	-0.188	-0.071	-0.413	-0.237	0.284	0.121	0.139	-0.106
Total C	0.367	0.351	0.227	0.385	0.603 **	-0.049	0.387	0.536 *	0.765 **	0.766	0.172
Total P	1.000 **	0.874 **	0.782 **	0.894 **	0.864 **	0.742 **	0.930 **	0.189	0.604 **	0.575 *	0.939 **
Oxalate P		1.000 **	0.890 **	0.915 **	0.823 **	0.837 **	0.950 **	0.194	0.478	0.439	0.724 **
Resin-P			1.000 **	0.888 **	0.780 **	0.772 **	0.898 **	0.251	0.331	0.309	0.628 **
NaHCO ₃ -Pi				1.000 **	0.907 **	0.710 **	0.953 **	0.276	0.567 *	0.503	0.743 **
NaOH-Pi					1.000 **	0.510	0.919 **	0.498	0.727 **	0.704 **	0.666 **
HCl-P						1.000 **	0.807 **	-0.130	0.097	0.060	0.689 **
TPi							1.000 **	0.282	0.542 **	0.505	0.770 **
NaHCO ₃ -Po								1.000 **	0.371	0.468	0.003
NaOH-Po									1.000 **	0.994 **	0.425
TPo										1.000 **	0.403
Residual P											1.000 **

Alo, oxalate-extractable Al; Feo, oxalate-extractable Fe; Oxalate P, oxalate-extractable P; TPi, sum of extracted inorganic P; TPo, sum of extracted organic P; Residual P, difference between total P in soils and the sum of inorganic and organic P in all the extracted fractions

* $P < 0.01$, ** $P < 0.001$

difference in P uptake (ΔP uptake) between the +N and +NP treatments using various soil properties. The Resin-P and NaHCO₃-Po were selected as the positive explanatory variables for P uptake in the +N treatment, whereas Alo and total P were selected as negative explanatory variables. Using these selected parameters, the regression model explained 72% of the variation in P uptake of the +N treatment across Exp. 1 and Exp. 2. The standardized partial regression coefficients indicated that the effect size of selected variables decreased in the order of Resin-P, Alo, NaHCO₃-Po, and total P. Using the same procedure, Alo + 0.5 Feo was selected as a strong and negative explanatory variable for ΔP uptake. In addition, total C was selected as a weak and negative explanatory variable ($P = 0.051$). The multiple linear regression analysis of these two variables

explained 81% of the variation in ΔP uptake across Exp. 1 and Exp. 2.

Discussion

Large field-to-field variations in soil P status and response of rice to nutrient inputs

Two sets of pot-based experiments confirmed that the production of irrigated rice—at least in the early growth stage—is limited by P deficiency in most of the soils in the central highland of Madagascar. In addition, there were large field-to-field variations in responses of rice production or P uptake when P was omitted (+N treatment) and when P was applied (+NP treatment)x

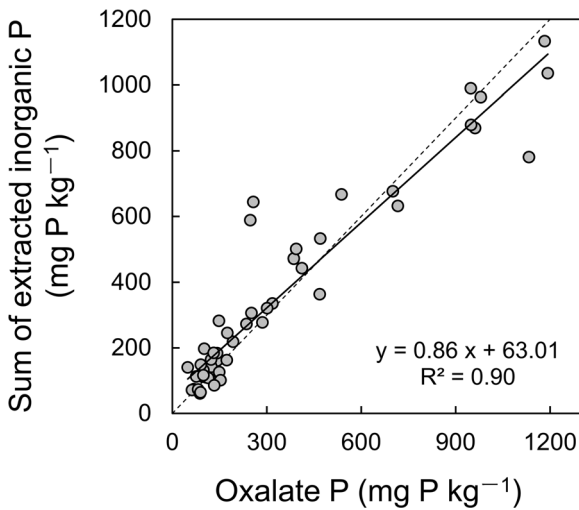


Fig. 5 Relationship between oxalate-extractable P and sum of extracted inorganic P. **footer:** Solid line shows best fit from regression analysis, and the broken line shows 1:1 ratio

(Table 1, Fig. 3); these large variations in the indigenous soil P-supplying capacity and fertilizer-P recovery efficiencies were confirmed not only for those soils differing in geological backgrounds within a particular region (Exp. 1) but also for soils in adjacent fields in a small area (Exp. 2) (Table 1). The results corroborate recent assertion that better information is needed concerning

Table 4 Multiple regression analysis for plant P uptake in the +N treatment and difference in P uptake between the +NP and +N treatments

	Selected variable	Partial regression coefficient	Standardized partial regression coefficient
P uptake in +N treatment	Intercept	-1.170 ^{ns}	-0.471 ^{ns}
	Resin-P	0.034 ^{***}	0.419 ^{***}
	NaHCO ₃ -Po	0.021 ^{***}	0.222 ^{***}
	Alo	-0.100 ^{***}	-0.286 ^{***}
	Total P	-0.0003 [*]	0.197 [*]
	Adjusted R square = 0.72		
ΔP uptake ^a	Intercept	1.010 ^{ns}	0.252 ^{ns}
	Alo + 0.5 Feo	-0.056 ^{**}	-0.244 ^{**}
	Total C	-0.158 [†]	-0.177 [†]
	Adjusted R square = 0.81		

^a Difference in phosphorus uptake between +NP and +N treatments

*P < 0.05; **P < 0.01; ***P < 0.001; †P = 0.051; ns, not significant

Alo, oxalate-extractable Al; Feo, oxalate-extractable Fe

differences in the response to applied nutrients within and between fields to effectively increase yields and returns of inputs for smallholder farmers in SSA (e.g. Schut et al. 2018).

Phosphorus is generally considered to have less mobility compared with other plant nutrients because of its strong fixation to the high levels of Fe and Al oxides in tropical soils. Therefore, soil P status is relatively consistent within a geological background at the landscape level (Nishigaki et al. 2018). However, our results for soil P index revealed a substantial variation in total P and fractionated P despite the inclusion of samples collected over small distances. Total P was generally high in the lowland soils compared with the soils collected from adjacent upland fields supposedly due to the replenishment of sediments in runoff water in the lowland fields. Yet, large variations in total P were observed even within lowland soils in a small area (see [Supplementary data sheet](#)). Schmitter et al. (2010) also reported a large spatial variation in soil physicochemical properties in rice paddy terraces in Northwest Vietnam. These results support the importance of site-specific fertilizer management practices based on differences in soil P status and responses of rice to P fertilizer. The large variation in soil P status that we observed in Madagascar could be attributable to land management practices and pervasive soil erosion on the sloping topography.

P uptake of rice plants in relation to soil P supply capacity

The result of the regression analysis revealed that P uptake of rice without P application was increased by Resin-P and NaHCO₃-Po, indicating that these fractions constitute a readily available P pool in soils for rice grown under flooded conditions (Table 4). In other words, the indigenous P supply for lowland rice largely relies on Resin-P, which accounts for merely 1% of total P in soils (Fig. 4). In addition, our results revealed a quantitative effect of labile organic P (NaHCO₃-Po) on the P uptake values of rice plants in highly weathered soils. Indeed, previous studies have emphasized the potential role of organic P pool as a source of available P in tropical weathered soils (George et al. 2018; Turner 2006).

The total organic P pool (NaHCO₃-Po and NaOH-Po) in this study was 97 mg P kg⁻¹, representing 10% of total P. This is comparable with results of a previous report on soil organic P in rice fields in Madagascar, i.e., the Po pool ranged from 6.7% to 29% (Turner 2006).

Randriamanantsoa et al. (2015) reported that gross organic P mineralization rates were 0.8 ± 0.5 mg P $\text{kg}^{-1} \text{day}^{-1}$ and 1.7 ± 0.2 mg P $\text{kg}^{-1} \text{day}^{-1}$ in non-amended and residue-amended Ferralsols, respectively. Our results and those of these previous studies also support the potential role for soil organic P in rice nutrition at the initial growth stage in tropical soils that typically have low available inorganic P. On the other hand, Alo in soils had a negative effect on P uptake of rice. This is likely attributable to the high sorption capacity of active Al in soils (Agbenin 2003; Nwoke et al. 2003). We therefore conclude that the P-supplying capacity of strongly weathered rice-fields is mainly controlled by the size of the labile inorganic and organic P pools and the sorption capacity of the soil matrix. However, the potential contribution of less-labile P fractions through the root-induced chemical reactions in rhizosphere soils (Hinsinger 2001) should not be completely excluded when we consider relatively low P mobility in highly weathered soils and the mass balance between P uptakes of rice plants and labile P pools in rhizosphere soils.

Effect of oxalate-extractable Al, Fe, and P on fertilizer-P recovery efficiency and P cycling

The result of the regression analysis revealed that the difference in P uptake between the +NP and +N treatments was strongly controlled by the Alo + 0.5 Feo (Table 4). This underscores the importance of Al and Fe (hydr)oxides in regulating fertilizer-P recovery efficiency for lowland rice production in highly weathered soils. This suggests that the applied P was mainly sorbed to the active Al and Fe, and hence P availability for plants decreased. Sugihara et al. (2012) reported that P in fertilizer was predominantly fixed by Al and Fe oxides in a clay-rich maize cropland in Tanzania. Nwoke et al. (2003) also suggested that oxalate-extractable Al and Fe can regulate the standard P requirement and P availability in soils of the West African savanna.

We found that the Alo + 0.5 Feo correlated positively with each inorganic P fraction, and most highly with $\text{NaHCO}_3\text{-Pi}$ and NaOH-Pi , suggesting that P in these labile and less-labile fractions is mainly associated with the active Al and Fe. Despite the large proportion of NaOH-Pi in total P (Fig. 4), NaOH-Pi did not contribute to the P uptake of rice in +N treatment owing to the low availability of P strongly fixed by the active Al and Fe. Considering the higher regression coefficients for Alo

than Feo with $\text{NaHCO}_3\text{-Pi}$ and NaOH-Pi , Alo is likely to be the main controlling factor for those fractions. Nwoke et al. (2003) suggested that Al plays a greater role than Fe in removing P from soil solution of the surface layer because the pH values of soils are in the range in which P is more likely to react with Al than with Fe.

It has been reported that the application of organic matter, such as animal manure, can decrease P sorption and increase P availability in highly weathered soils (Azeez and Averbeke 2011; Guppy et al. 2005), owing to competition between organic matter molecules and P for the available sorbing sites on soil particles (Chassé and Ohno 2016). Hence, farmyard manure application, which is widely used for agricultural lands in the central highlands of Madagascar, is indeed an effective management option for improving the phytoavailability of applied phosphate in weathered soils. Andriamananjara et al. (2018) reported that farmyard manure application to weathered upland soils of Madagascar increases P fertilizer recovery efficiency for upland rice.

Our results reveal that the amount of Oxalate P clearly corresponded to sum of extracted inorganic P (Fig. 5). This fact leads us to surmise that Oxalate P represents all the labile and less-labile inorganic P pools in soils. Most of the P sorbed by active Fe and Al compounds is slowly released back to the soil solution, providing service flows for 5 to 10 years (Sanchez et al. 1997). Agbenin and Goladi (1998) indicated that Resin-P is attributable to the less-labile P sources, namely NaOH-Pi and HCl-P , which act as sinks for fertilizer-P. These results suggest that Oxalate P can be an index for the evaluation of both labile and less-labile P pools, namely the moderately active P pool, which may be involved in the subsequent slow P cycling in rice croplands (Lookman et al. 1995). This could account for why many previous studies, which were mostly conducted using soils with an abundant labile P pool, have shown that Oxalate P is poorly effective at predicting P uptake by crops (Eichler-Löbermann et al. 2007; Nawara et al. 2017). Interestingly, however, we found a significant correlation between P uptake and Oxalate P in both Exp. 1 ($r=0.75$, $P<0.001$ in the *Cont* treatment) and Exp. 2 ($r=0.72$, $P<0.01$ in the *Cont* treatment). This is likely attributable to the fact that less-labile P in tropical weathered soils is more abundant than in temperate soils and, therefore, the overall effect of less-labile P is greater in tropical weathered soils.

Conclusion

Our experiments successfully extracted the most relevant soil components to determine the indigenous soil P-supplying capacity and fertilizer-P recovery efficiencies for rice production. Phosphorus uptake of rice in the early growth stage is tightly controlled by labile inorganic and organic P pools and active Al and Fe in P-deficient weathered soils. In addition, a high degree of variability in rice P uptake in response to fertilizer treatment and soil characteristics is observed even within a small area. These results provide fundamental information for developing more effective nutrient management practices and improving financial returns for individual fields. Further studies are required to determine whether the results from our pot-based experiments are applicable to the responses of rice in the field. To achieve practical field-specific fertilizer management, further studies are also needed to reveal how to increase available P—in both inorganic and organic forms—and to reduce the activity of Al and Fe oxides in highly weathered soils.

Acknowledgments The authors are grateful to Dr. Patrick E. Hayes for his valuable comments on this manuscript. We also thank Ms. Mayumi Yonemura for assisting with the laboratory analysis. This research was financially supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



References

- Agbenin JO (2003) Extractable iron and aluminum effects on phosphate sorption in a savanna Alfisol. *Soil Sci Soc Am J* 67:589–595
- Agbenin JO, Goladi JT (1998) Dynamics of phosphorus fractions in a savanna Alfisol under continuous cultivation. *Soil Use Manag* 14:59–64
- Andriamananjara A, Rakotoson T, Razanakoto OR, Razafimanantsoa M-P, Rabeharisoa L, Smolders E (2018) Farmyard manure application in weathered upland soils of Madagascar sharply increase phosphate fertilizer use efficiency for upland rice. *Field Crop Res* 222:94–100
- Azeez JO, Averbek WV (2011) Effect of manure types and period of incubation on phosphorus-sorption indices of a weathered tropical soil. *Comm Soil Sci Plant Anal* 42:2200–2218
- Balemi T, Negisho K (2012) Management of soil phosphorus and plant adaptation mechanisms to phosphorus stress for sustainable crop production: a review. *J Soil Sci Plant Nutr* 12: 547–561
- Beck MA, Sanchez PA (1994) Soil phosphorus fraction dynamics during 18 years of cultivation on a Typic Paleudult. *Soil Sci Soc Am J* 34:1424–1431
- Chassé AW, Ohno T (2016) Higher molecular mass organic matter molecules compete with orthophosphate for adsorption to iron (oxy)hydroxide. *Env Sci Tech* 50:7461–7469
- Courchesne F, Turmel MC (2008) Extractable Al, Fe, Mn, and Si. In: Carter MR, Gregorich EG (eds) *In Soil sampling and methods of analysis*, 2nd edn. Canadian Society of Soil Science, CRC Press, Boca Raton
- Dogbe W, Sogbedji JM, Buah SSJ (2015) Site-specific nutrient management for lowland rice in the northern Savannah zones of Ghana. *Curr Agr Res* 3:109–117
- Eichler-Löbermann B, Köhne S, Köppen D (2007) Effect of organic, inorganic, and combined organic and inorganic P fertilization on plant P uptake and soil P pools. *J Plant Nutr Soil Sci* 170:623–628
- Fageria NK, Baligar VC (1997) Response of common bean, upland rice, corn, wheat, and soybean to soil fertility of an Oxisol. *J Plant Nutr* 20:1279–1289
- Fageria NK, Carvalho GD, Santos AB, Ferreira EPB, Knupp AM (2011) Chemistry of lowland rice soils and nutrient availability. *Comm Soil Sci Plant Anal* 42:1913–1933
- Gee GW, Bauder JW (1986) Particle-size analysis. In: Klute A (ed) *Method of soil analysis part 1. Physical and mineralogical methods*. Agronomy monograph no. 9, second ed. American Society of Agronomy and Soil Science of America, Wisconsin, pp 383–411
- George TS, Giles CD, Menezes-Blackburn D et al (2018) Organic phosphorus in the terrestrial environment: a perspective on the state of the art and future priorities. *Plant Soil* 427:191–208
- Guo F, Yost RS, Hue NV, Evensen CI, Silva JA (2000) Changes in phosphorus fractions in soils under intensive plant growth. *Soil Sci Soc Am J* 64:1681–1689
- Guppy CN, Menzies NW, Moody PW, Blamey FPC (2005) Competitive sorption reactions between phosphorus and organic matter in soil: a review. *Soil Res* 43:189–202
- Haynes RJ, Mokolobate MS (2001) Amelioration of Al toxicity and P deficiency in acid soils by additions of organic residues: a critical review of the phenomenon and the mechanisms involved. *Nutr Cycl Agroecosyst* 59:47–63
- Hedley MJ, Stewart JWB, Chauhan BS (1982) Changes in inorganic and organic soil phosphorus fractions induced by cultivation practices and by laboratory incubations. *Soil Sci Soc Am J* 46:970–976
- Hinsinger P (2001) Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: a review. *Plant Soil* 237:173–195
- ISRIC (2002) Cation exchange capacity (CEC) and exchangeable bases (ammonium acetate method). In: van Reeuwijk LP (ed)

- Procedures for soil analysis. International Soil Reference and Information Centre, Wageningen, p 9–1–9–14
- Kihara J, Nziguheba G, Zingore S, Coulibaly A, Esilaba A, Kabambe V, Njoroge S, Palm C, Huising J (2016) Understanding variability in crop response to fertilizer and amendments in sub-Saharan Africa. *Agric Ecosyst Environ* 229:1–12
- Kitson RE, Mellon MG (1944) Colorimetric determination of phosphorus as molybdivanadophosphoric acid. *Ind Eng Chem Anal Ed* 16:379–383
- Koné B, Amadji GL, Aliou S, Diatta S, Akakpo C (2011) Nutrient constraint and yield potential of rice on upland soil in the south of the Dahoumey gap of West Africa. *Arch Agron Soil Sci* 57:763–774
- Koné B, Fofana M, Sorho F, Diatta S, Ogunbayo A, Moussa S (2013) Nutrient constraint of rainfed rice production in foot slope soil of Guinea Forest in Côte d'Ivoire. *Arch Agron Soil Sci* 60:735–746
- Kuo S (1996) Phosphorus. In: Sparks DL (ed) *Methods of Soil Analysis*. Agronomy 9. ASA-SSA, Madison, Wisconsin, USA, pp 869–919
- Lookman R, Freese D, Merckx R, Vlassak K, van Riemsdijk WH (1995) Long-term kinetics of phosphate release from soil. *Environ Sci Technol* 29:1569–1575
- Menezes-Blackburn D, Giles C, Darch T, George TS, Blackwell M, Stutter M, Shand C, Lumsdon d CP, Wendler R, Brown L, Almeida DS, Wearing C, Zhang H, Haygarth PM (2017) Opportunities for mobilizing recalcitrant phosphorus from agricultural soils: a review. *Plant Soil* 427:5–16
- Murphy J, Riley JP (1962) A modified single solution method for the determination of phosphate in natural waters. *Anal Chem Acta* 27:31–36
- Nawara S, Van Dael T, Merckx R, Amery F, Elsen A, Odeurs W, Vandendriessche H, Mcgrath S, Roisin C, Jouany C, Pellerin S, Denoroy P, Eichler-Löbermann B, Börjesson G, Goos P, Akkermans W, Smolders E (2017) A comparison of soil tests for available phosphorus in long-term field experiments in Europe. *Eur J Soil Sci* 68:873–885
- Nishigaki T, Sugihara S, Kobayashi K, Hashimoto Y, Kilasara M, Tanaka H, Watanabe T, Funakawa S (2018) Fractionation of phosphorus in soils with different geological and soil physicochemical properties in southern Tanzania. *Soil Sci Plant Nutr* 64:291–299
- Nwoke OC, Vanlauwe B, Diels J, Sanginga N, Osonubi O, Merckx R (2003) Assessment of labile phosphorus fractions and adsorption characteristics in relation to soil properties of west African savanna soils. *Agric Ecosyst Environ* 100:285–294
- Nziguheba G, Zingore S, Kihara J, Merckx R, Njoroge S, Otinga A, Vandamme E, Vanlauwe B (2016) Phosphorus in smallholder farming systems of sub-Saharan Africa: implications for agricultural intensification. *Nutr Cycl Agroecosyst* 104:321–340
- Randriamanantsoa L, Frossard E, Oberson A, Bünemann EK (2015) Gross organic phosphorus mineralization rates can be assessed in a Ferralsol using an isotopic dilution method. *Geoderma* 257–258:86–93
- Saito K, Diack S, Dieng I, N'Diaye MK (2015a) On-farm testing of a nutrient management decision-support tool for rice in the Senegal River valley. *Comput Electron Agr* 116:36–44
- Saito K, Dieng I, Toure AA, Somado EA, Wopereis MSC (2015b) Rice yield growth analysis for 24 African countries over 1960–2012. *Glob Food Secur* 5:62–69
- Sanchez PA, Shepherd KD, Soule MJ, Place FM, Buresh RJ, Izac AN, Mokwunye AU, Kwesiga FR, Ndiritu CG, Woomer PL (1997) Soil fertility replenishment in Africa: an investment in natural resource capital. In: Buresh RJ, Sanchez PA, Calhoun F (eds) *Replenishing soil fertility in Africa*. SSSA Special Publication Number, vol 51, pp 1–46
- Schmitter P, Dercon G, Hilger T, Thi Le Ha T, Huu Thanh N, Lam N, Duc Vien T, Cadisch G (2010) Sediment induced soil spatial variation in paddy fields of Northwest Vietnam. *Geoderma* 155:298–307
- Schut AGT, Traore PCS, Blaes X, de By RA (2018) Assessing yield and fertilizer response in heterogeneous smallholder fields with UAVs and satellites. *Field Crop Res* 221:98–107
- Shehu BM, Merckx R, Jibrin JM, Kamara AY, Rurinda J (2018) Quantifying variability in maize yield response to nutrient applications in the northern Nigerian savanna. *Agronomy* 8:18
- Sugihara S, Funakawa S, Nishigaki T, Kilasara M, Kosaki T (2012) Dynamics of fractionated P and P budget in soil under different land management in two Tanzanian croplands with contrasting soil textures. *Agric Ecosyst Environ* 162:101–107
- Tiessen H, Moir JO (2007) Characterization of available phosphorus by sequential extraction. In: Carter MR, Gregorich EG (eds) *Soil sampling and methods of analysis*, second edition, Canadian Society of Soil Science. Lewis Publishers, Boca Raton, pp 293–306
- Tiessen H, Salcedo IH, Sampaio EVSB (1992) Nutrient and soil organic matter dynamics under shifting cultivation in semi-arid northeastern Brazil. *Agric Ecosyst Environ* 38:139–151
- Turner BL (2006) Organic phosphorus in Madagascan rice soils. *Geoderma* 136:279–288
- Van Oort PAJ, Saito K, Tanaka A, Amovin-Assagba E, Van Bussel LGJ, Van Wart J, De Groot H, Van Ittersum MK, Cassman KG, Wopereis MCS (2015) Assessment of rice self-sufficiency in 2025 in eight African countries. *Glob Food Secur* 5:39–49
- Van Vuuren DP, Bouwman AF, Beusen AHW (2010) Phosphorus demand for the 1970–2100 period: a scenario analysis of resource depletion. *Glob Environ Change* 20:428–439
- Yan X, Wang D, Zhang H, Zhang G, Wei Z (2013) Organic amendments affect phosphorus sorption characteristics in a paddy soil. *Agric Ecosyst Environ* 175:47–53
- Yanai J, Lee CK, Kaho T, Iida M, Matsui T, Umeda M, Kosaki T (2012) Geostatistical analysis of soil chemical properties and rice yield in a paddy field and application to the analysis of yield-determining factors. *Soil Sci Plant Nutr* 47:291–301
- Zhang Q, Wang GH, Feng YK, Sun QZ, Witt C, Dobermann A (2006) Changes in soil phosphorus fractions in a calcareous paddy soil under intensive rice cropping. *Plant Soil* 288:141–154
- Zingore S, Tittonell P, Corbeels M, van Wijk MT, Giller KE (2011) Managing soil fertility diversity to enhance resource use efficiencies in smallholder farming systems: a case from Murewa District, Zimbabwe. *Nutr Cycl Agroecosyst* 90:87–103

Article

Laboratory Visible and Near-Infrared Spectroscopy with Genetic Algorithm-Based Partial Least Squares Regression for Assessing the Soil Phosphorus Content of Upland and Lowland Rice Fields in Madagascar

Kensuke Kawamura ^{1,*} , Yasuhiro Tsujimoto ¹, Tomohiro Nishigaki ¹, Andry Andriamananjara ², Michel Rabenarivo ², Hidetoshi Asai ¹, Tovohera Rakotoson ² and Tantely Razafimbelo ² 

¹ Japan International Research Center for Agricultural Sciences (JIRCAS), 1-1 Ohwashi, Tsukuba, Ibaraki 305-8686, Japan; tsjmt@affrc.go.jp (Y.T.); nishigaki@affrc.go.jp (T.N.); asai0817@affrc.go.jp (H.A.)

² Laboratoire des Radio-Isotopes, Université d'Antananarivo, BP 3383, Route d'Andraisoro, 101 Antananarivo, Madagascar; njaraandry1@gmail.com (A.A.); miarabenarivo@yahoo.fr (M.R.); tovohera.rakotoson@gmail.com (T.R.); tantely.razafimbelo@gmail.com (T.R.)

* Correspondence: kamuken@affrc.go.jp; Tel.: +81-29-838-6624

Received: 1 February 2019; Accepted: 22 February 2019; Published: 2 March 2019



Abstract: As a laboratory proximal sensing technique, the capability of visible and near-infrared (Vis-NIR) diffused reflectance spectroscopy with partial least squares (PLS) regression to determine soil properties has previously been demonstrated. However, the evaluation of the soil phosphorus (P) content—a major nutrient constraint for crop production in the tropics—is still a challenging task. PLS regression with waveband selection can improve the predictive ability of a calibration model, and a genetic algorithm (GA) has been widely applied as a suitable method for selecting wavebands in laboratory calibrations. To develop a laboratory-based proximal sensing method, this study investigated the potential to use GA-PLS regression analyses to estimate oxalate-extractable P in upland and lowland soils from laboratory Vis-NIR reflectance data. In terms of predictive ability, GA-PLS regression was compared with iterative stepwise elimination PLS (ISE-PLS) regression and standard full-spectrum PLS (FS-PLS) regression using soil samples collected in 2015 and 2016 from the surface of upland and lowland rice fields in Madagascar ($n = 103$). Overall, the GA-PLS model using first derivative reflectance (FDR) had the best predictive accuracy ($R^2 = 0.796$) with a good prediction ability (residual predictive deviation (RPD) = 2.211). Selected wavebands in the GA-PLS model did not perfectly match wavelengths of previously known absorption features of soil nutrients, but in most cases, the selected wavebands were within 20 nm of previously known wavelength regions. Bootstrap procedures ($N = 10,000$ times) using selected wavebands also confirmed the improvements in accuracy and robustness of the GA-PLS model compared to those of the ISE-PLS and FS-PLS models. These results suggest that soil oxalate-extractable P can be predicted from Vis-NIR spectroscopy and that GA-PLS regression has the advantage of tuning optimum bands for PLS regression, contributing to a better predictive ability.

Keywords: Madagascar; oxalate-extractable soil P; partial least squares regression; soil fertility; spectral assessments; waveband selection

1. Introduction

Phosphorus (P) deficiency is a major constraint for rice production in the tropics [1] because strongly weathered soils, which cover vast regions of the tropics, contain low concentrations of readily exchangeable inorganic phosphate [2,3]. In tropical soils, available P can generally be even less due to strong sorption to aluminium (Al) and iron (Fe) oxides and often limits crop production in low input agricultural systems [4,5].

In our previous study [6], we examined highly weathered and typical P-deficient soils in the central highland of Madagascar and found that the P uptake of rice under flooded conditions is related not only to easily soluble P content but also to the amounts of active Fe and Al, which are bound to inalcitrant P fractions. Acid ammonium oxalate extraction is a powerful extraction method and covers both soluble and inalcitrant Fe- and Al-bound P fractions, unlike certain conventional extraction methods (e.g., the Olsen method) that normally consider only easily soluble P [6,7]. In addition, recently, Helfenstein et al. [8] revealed that the inalcitrant P fraction (NaOH-extractable inorganic P pool) turns over in weeks to months, suggesting that the inalcitrant P fraction would potentially play a significant role as a P source within a cropping season. Rabeharisoa et al. [9] also found the amount of oxalate-extractable P in soils had a significant correlation with the P concentrations of rice leaves in farmers' fields in Madagascar. Therefore, we assumed that oxalate-extractable P reflects bioavailable P for rice production in the region and applied this assumption in the current study.

Soil P occurs in a variety of chemical forms that differ markedly in their behavior and bioavailability in the soil environment [6,10]. Our previous study also revealed that these soil P contents and forms largely varied among neighboring fields. These observations indicate that P nutrient management for rice production can be further improved by understanding field-to-field variations in bioavailable P (i.e., oxalate-extractable P) in the tropics. Thus, the development of a rapid and accurate methodology for evaluating bioavailable P in soils is needed. However, a quantitative assessment of soil P using standard procedures (e.g., wet chemistry) can often be difficult, especially in spatially heterogeneous assessments that require numerous soil samples, a process that is costly and time consuming.

To overcome the issues with the standard procedure, laboratory visible and near-infrared (Vis-NIR) spectroscopy has been widely adopted for soil studies as a non-destructive, rapid and reproducible analytical method and has been used for the simultaneous prediction of a variety of primary and secondary soil attributes [11]. Vis-NIR spectroscopy is an analytical technique that characterizes materials according to their reflectance at light absorption in the visible (400–700 nm) and NIR (700–2,500 nm) regions. These techniques measure the radiation absorbed by various bonds of O-H, C-H, N-H, C=O, C-N, N-H, or C=C, resulting in bending, twisting, stretching, or scissoring [11,12]. Spectroscopy has been used in conjunction with chemometric (multivariate regression) analyses to relate soil spectra to soil attributes, such as carbon content, clay and iron oxide [13–15].

Although partial least squares (PLS) regression is the most commonly used approach for soil spectral analyses, waveband selection can refine the performance of a PLS analysis [16–18]. The PLS regression method combines the most useful information from hundreds of wavebands into the first several PLS factors (or latent variables), whereas the less important factors might include background effects [19,20]. Thus, many techniques for selecting wavebands or wavelength regions have been developed, such as iterative stepwise elimination-PLS (ISE-PLS) regression [21], uninformative variable elimination-PLS (UVE-PLS) regression [22], competitive adaptive reweighted sampling (CARS) regression [23], interval PLS (iPLS) regression [24], moving window-PLS (MW-PLS) regression [25], and genetic algorithm-PLS (GA-PLS) regression [26]. In our previous study [18], removal of the redundant wavebands by ISE-PLS regression greatly improved the estimation of total carbon (TC) and total nitrogen (TN) in paddy soils. Among the waveband selection methods, GA-PLS regression has been used as a suitable method in chemometrics [27]. Leardi and González [28] demonstrated that the GA-PLS method, after suitable modifications, produces more interpretable results because the selected

wavelengths are less dispersed in this method than in other methods. Several studies have reported that the GA-PLS method obtained a better solution than did the ISE-PLS method [29–31].

To date, there have been several attempts to predict soil P using Vis-NIR spectroscopy at field and laboratory scales with the standard full-spectrum PLS (FS-PLS) method [32–34]. However, the predictive accuracy was relatively low compared with that for other macro-nutrients (e.g., nitrogen, carbon), and waveband selection coupled with PLS regression analysis has not been evaluated. The objective of this study was to develop a laboratory-based proximal sensing method based on an empirical relationship between soil P and Vis-NIR spectral characteristics using PLS analyses. To improve the predictive ability, we investigated the potential to use GA-PLS regression analyses to estimate the soil P status of upland and lowland soils from laboratory Vis-NIR reflectance data. Here, we targeted amounts of oxalate-extractable P based on the above-noted field observations regarding its relative importance for rice production and on P uptake as noted by Rabeharisoa et al. [9] and Nishigaki et al. [6]. The predictive ability of the GA-PLS method was compared with the predictive abilities of ISE-PLS and FS-PLS methods using first derivative reflectance (FDR) spectra data. Rapid measurements of soil P status at low cost and with less soil sample preparation could be an application of the present study instead of the routine chemical methods.

2. Materials and Methods

2.1. Study Site and Soil Sampling and Chemical Analyses

The field survey was conducted in the central highland of Madagascar (Figure 1). This region has a subtropical climate with an altitude of 1000–1500 m above sea level. The mean temperature is 14–17°C in winter and 20–22°C in summer. The average annual rainfall is 1100 mm (>80% occurs in November–March) [35]. The area is dominated by inherently nutrient-poor soil types that are mainly classified into Ferralsols and Acrisols [36] or into Oxisols of semiarid to humid climates [37].

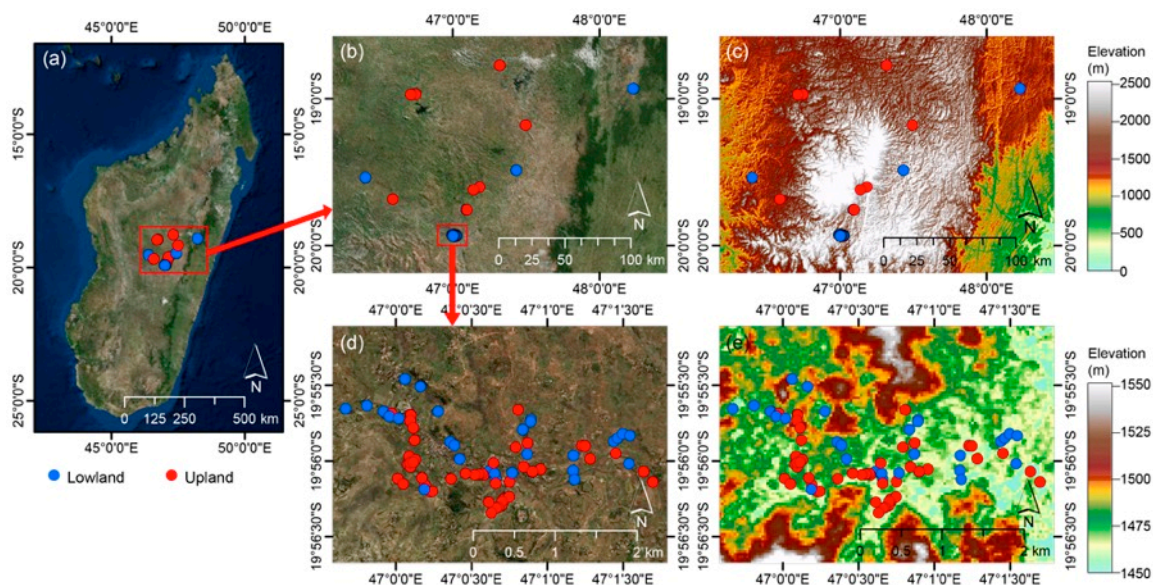


Figure 1. Location of studied regions and soil sampling points. Source in (a), (b) and (d): Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community. Source in (c) and (e): The ASTER GDEM version 2 data were downloaded via EarthExplorer (<https://earthexplorer.usgs.gov/>).

In 2015 and 2016, soil sampling was conducted in 103 rice fields—the major production system in the region — with 63 upland and 40 lowland fields under various management practices. The sampling positions were recorded with a handheld GPS (Colorado 300, Garmin, Ltd., Kansas, TX, USA). Surface

soil samples were collected at a 0–10 cm depth as composites of three to four cores in each field. The spatial distributions of soil sampling points were plotted on the maps using satellite images (Figure 1a,b,d) and ASTER global digital elevation model (GDEM) version 2 data (Figure 1c,e), which is a product of METI and NASA.

In the laboratory, soil samples were air dried for 14 days and sieved to <2 mm. Soil P was extracted using the acid ammonium oxalate method as described by Schwertmann [38], and the concentration of P in the oxalate extraction was analyzed using the malachite green colorimetric method [39].

2.2. Vis-NIR Diffuse Reflectance Measurement

Laboratory soil reflectance measurements were conducted in a dark room at the Japan International Research Center for Agricultural Sciences (JIRCAS), Japan, on July 31–August 1, 2017. Soil samples were scanned by a portable spectroradiometer (FieldSpec 4 Hi-Res, Analytical Spectral Devices (ASD) Inc., Longmont, CO, USA) and an ASD contact probe. The ASD FieldSpec measures spectral reflectance in the 350–2500 nm wavelength region, which has one silicon array (350–1000 nm) and two indium gallium arsenide (InGaAs) detectors (1000–1800 and 1800–2500 nm). The spectral sampling interval was 1.4 nm in the 350–1000 nm range and 1.1 nm in the 1001–2500 nm range. The spectral resolution (full-width-half-maximum; FWHM) was 3 nm in the 350–1000 nm range and 6 nm in the 1000–2500 nm range, which were calculated to 1 nm resolution wavelengths for output data using the cubic spline interpolation function in ASD software (RS3 for Windows; ASD Inc., Longmont, CO, USA).

The contact probe light source (halogen lamp) was aligned at 12° to the probe body, ensuring illumination at a fixed angle without the influence of ambient light. The fiber optic cable of the ASD FieldSpec was attached to the contact probe at a fixed measurement angle of 35°. The sensed spot area had a diameter of ~1.1 cm with a field of view of 1.33 cm². A Spectralon (Labsphere Inc., Sutton, NH, USA) reference panel (white reference) was used to optimize the ASD instrument prior to taking Vis-NIR reflectance measurements for each sample.

Bulk soil samples were spread in optical-glass Petri dishes that were 85 mm in diameter and pressed to form a layer ~19 mm thick. The soil surfaces were scanned 25 times with five replications for the soil samples, and the spectral readings were averaged.

2.3. Overview of Data Processing

In this section, an overview of the data processing process is described using a flowchart in Figure 2 that shows a general overview of the methodology. In this study, two types of validations were performed for the models: (i) a leave-one-out cross-validation (LOO-CV) procedure based on whole data sets ($n = 103$) and (ii) a modified bootstrap procedure based on an independent test data set, which was similar to our previous study [30]. Here, the LOO-CV procedure included waveband selection in ISE-PLS and GA-PLS regression analyses, while the bootstrap procedure was performed using the selected wavebands; the best GA-PLS model and final wavebands from five GA runs were justified by the residual predictive values (RPD). More details on the predictive abilities are described in Section 2.9.

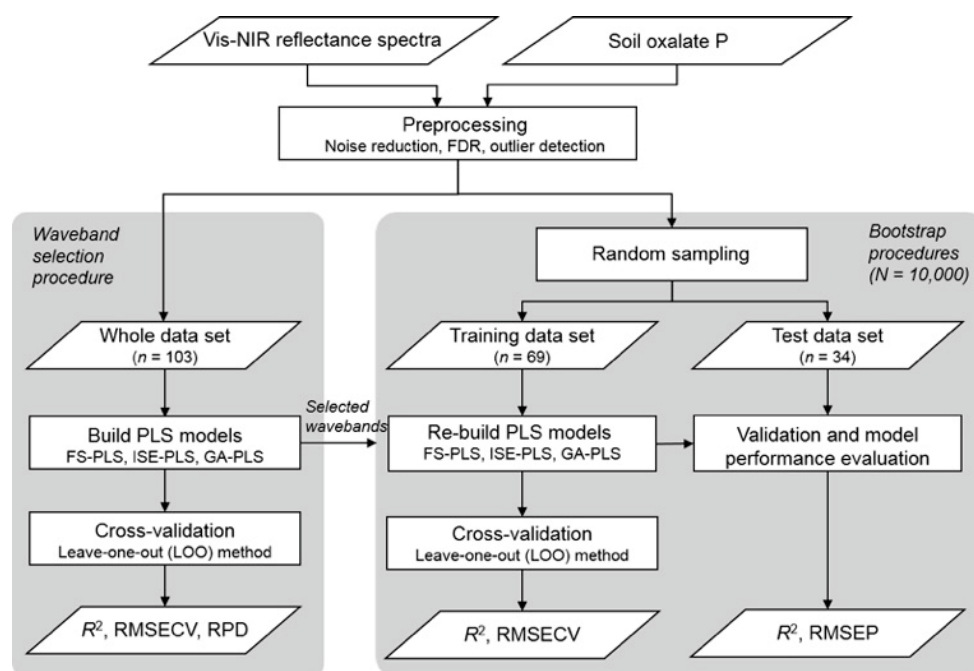


Figure 2. Flowchart depicting a general overview of the methodology.

All data handling and linear regression analyses were performed using PLS_Toolbox version 8.6 (Eigenvector Research Inc., Manson, WA, USA) in MATLAB software ver. 9.3 (MathWorks Sherborn, MA, USA).

2.4. Preprocessing of Spectral Data

Spectral data in both edge wavelength regions (350–399 nm and 2401–2500 nm) were eliminated because of low signal-to-noise ratios in the instrument. Thus, a total of 2001 spectral bands between 400 nm and 2400 nm were used for the analyses.

FDR spectra were used to reduce baseline variation and enhance spectral features [40]. The FDR was calculated using the Savitzky-Golay smoothing filter [41]. A third-order, 15-band moving polynomial was fitted according to the original reflectance signatures. The parameters of this polynomial were subsequently used to calculate the derivative at the center waveband of the moving spline window. In addition, a standard normal variate transform (SNV) was employed to reduce the particle size effect [42].

To detect outliers, a principal component analysis was performed on spectral data for calculating the Mahalanobis distance H , and samples with $H > 3$ were eliminated as outliers. As a result, three samples were considered outliers, leaving 103 samples for further analyses.

2.5. Standard Full-Spectrum Partial Least Squares (FS-PLS) Regression

PLS regression analyses were performed to estimate soil parameters using reflectance and FDR data sets ($n = 103$). The standard FS-PLS regression equation was calculated as follows:

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon \quad (1)$$

where the response variable y is a vector of the soil oxalate-extractable P; the predictor variables x_1 to x_i are surface reflectance or FDR values for spectral bands 1 to i (400, 401, ..., 2400 nm), respectively; β_1 to β_i are the estimated weighted regression coefficients; and ε is the error vector. The latent variables were introduced to simplify the relationship between response variables and predictor variables. To determine the optimal number of latent variables (NLV), a LOO-CV was performed to avoid

over-fitting of the model and was based on the minimum value of the root mean squared error of cross-validation (RMSECV). The RMSECV was calculated as follows:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^n (y_i - y_p)^2}{n}} \quad (2)$$

where y_i and y_p represent the respective measured and predicted soil parameters for sample i , and n is the number of samples in the data sets ($n = 103$).

2.6. Iterative Stepwise Elimination Partial Least Squares (ISE-PLS) Regression

ISE-PLS is a PLS model that incorporates a waveband elimination algorithm. The ISE method eliminates noisy variables and selects useful predictors. When PLS models include large numbers of redundant variables or outliers, the models' predictive abilities may perform poorly, while the ISE method can overcome such problems. Performance depends on the importance of predictors (z_i), described as follows:

$$z_i = \frac{|\beta_i|s_i}{\sum_{i=1}^I |\beta_i|s_i} \quad (3)$$

where s_i is the standard deviation, and β_i is the regression coefficient; both s_i and β_i correspond to the predictor variable of the waveband i .

Initially, all available wavebands (2001 bands, 400–2400 nm) are used to develop the PLS regression model. Then, to create a scope in which useless predictor variables are removed and predictive ability is improved, each predictor z_i is evaluated, and the less informative wavebands are eliminated. Subsequently, the PLS model is re-calibrated with the remaining predictors [43]. The model-building procedure is repeated until the final model is calibrated with the maximum predictive ability.

2.7. Genetic Algorithm Partial Least Squares (GA-PLS) Regression

The GA is an efficient numerical optimization method based on genetic principles and natural selection [44]. In a GA, a population of individuals (or chromosomes) is created automatically and typically stored as binary strings in a computer memory, which means that a binary integer “zero” or “one” represent one gene. Then, each chromosome consists of sequences of a “gene” or “bits.” During this evolutionary computation, one or more bits are swapped within or between individuals by computer operations using mechanisms of natural variation, selection and inheritance. Briefly, selection, crossover and mutation form the core of GA, and these three operations are applied to the initial populations to generate a new population. This process is repeated until a pre-defined number of generations is propagated.

Although GA is well suited for solving variable subset selection problems [45], the major risk associated with using GA-PLS regression is over-fitting due to the large number of variables (wavebands) used in the Vis-NIR spectroscopy data set. To minimize the risk of over-fitting, Leardi [27] developed the GA program used in the present study. The GA program was designed to contain the following features: (i) The parameters are set with the highest possible elitism, a very limited population size and a relatively high mutation rate to ensure a rapid response increase and to find a good solution very early in the process. Here, elitism means to encourage the propagation of the best band repressors between generations without being disrupted by crossover or mutation so that the search speed of the program can be improved. (ii) The final model is determined via 100 independent and short GA runs. (iii) A weighted average of the selection frequency of the variables from the starting run to the previous runs is used to describe the current frequency of selection of the variables. Thus, each run is able to ‘learn’ information from the previous runs. (iv) A moving average (window size 3) is applied to the selection of the variables to take into account high spectral correlations and to ensure that highly correlated spectral bands are selected together.

Before the GA run, the suitability of our data set for applying GA band selection was assessed by a fitness function [27,46]. In the present study, five GA runs were performed on the FDR data set because each GA-PLS gave a slightly different model. The parameters and their conditions (Table 1) were taken from previous studies [27,30,31,46].

Table 1. Parameter conditions of the genetic algorithms-partial least squares (GA-PLS) regression.

Parameter	Condition
Population size	30 chromosomes
Regression method	PLS
Response	Cross-validated percent explained variance (five deletion group; the number of components is determined by cross-validation)
Maximum number of variables selected for the same chromosome	30
Probability of mutation	1%
Maximum number of latent variables	15
Number of runs	100
Window size for smoothing	3

2.8. Predictive Ability of the PLS Models

To evaluate the predictive ability of the FS-PLS, ISE-PLS and GA-PLS models, two types of validation were used for the models (see Figure 2): (i) a waveband selection with a LOO-CV procedure based on whole data sets ($n = 103$), including waveband selection in ISE-PLS and GA-PLS models, and (ii) a modified bootstrap procedure based on an independent test data set using the selected wavebands from (i).

In waveband selections with the LOO-CV procedure, each sample is estimated using the remaining samples. This process means that for each variant, we developed 103 individual models, which were constructed with data from 102 observations. The calibration model was then used to predict the observation that was left out. As the predicted samples were not the same as the samples used to establish the models, the RMSECV was used as the accuracy indicator of the model in predicting unknown samples. The predictive abilities of the PLS models were assessed by calculating the coefficient of determination (R^2), RMSECV and the residual predictive deviation (RPD) using a LOO-CV. High R^2 and low RMSECV values indicate the best model for predicting the soil parameters. The RPD has been defined as the ratio of standard deviation (SD) of reference data for predicting RMSECV. For the performance ability of calibration models, an RPD of 3 has been suggested for agriculture applications, while RPD values between 2 and 3 indicate a model with good prediction ability; $1.5 < \text{RPD} < 2$ is an intermediate model needing some improvement; and an $\text{RPD} < 1.5$ indicates that the model has poor prediction ability.

In the bootstrap procedure, the data were divided randomly into training ($n = 69$) and test ($n = 34$) data sets with replacement for $N = 10,000$ times. In each process, a PLS regression model was developed using the training data set. Here, ISE-PLS and GA-PLS were developed using selected wavebands in the LOO-CV procedure. The PLS model was then used to predict soil oxalate-extractable P in the test data set. The robustness of the calibration models was evaluated by the mean (\pm SD) values of R^2 and the root mean squares error of prediction (RMSEP) from 10,000 runs in the test data set. The RMSEP was calculated as follows:

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i^v - y_p^v)^2}{n}} \quad (4)$$

where y_i^v and y_p^v are the measured and predicted soil parameters, respectively, for sample i in the test data set.

2.9. Assessing Significant Wavelengths

To assess the importance of the wavelengths in the FS-PLS calibration, the variable importance in the projection (VIP) [47,48] was used and referred to the selected wavelength regions from the ISE-PLS and GA-PLS models. The VIP score provides a summary of the importance of an x variable (waveband) for an observed y variable and is calculated using the following equation:

$$VIP_k(a) = m \sum_a W_{ak}^2 \left(\frac{SSY_a}{SSY_t} \right) \quad (5)$$

where $VIP_k(a)$ is the importance of the k th predictor variable based on a model with a factors, W_{ak} is the corresponding loading weight of the k th variable in the a th PLS regression factor, SSY_a is the explained sum of squares of y obtained from a PLS regression model with a factors, SSY_t is the total sum of squares of y , and m is the total number of predictor variables. A high VIP score (>1) indicates an important x variable (waveband) [47,49].

3. Results and Discussion

3.1. A Wide Range of Soil Oxalate-Extractable P Contents in Upland and Lowland Rice Fields

The descriptive statistics of soil oxalate-extractable P in the whole ($n = 103$) upland ($n = 63$) and lowland ($n = 40$) data sets are shown in Figure 3, and Table 2 summarizes the minimum, maximum, median, mean, SD and coefficients of variation (CV) values. The soil oxalate-extractable P values in the upland and lowland data sets ranged between 30.73–1225.16 mg kg⁻¹ and 30.73–826.64 mg kg⁻¹, respectively. The mean value of the upland data set (588.74 mg kg⁻¹) showed significantly higher values than that of the lowland data set (319.41 mg P kg⁻¹) ($p < 0.001$, two sample t -test). The lower values in soil oxalate-extractable P is probably because of little fertilizer input in lowland fields compared to upland fields in the central highlands of Madagascar. Similarly, the soil TC was significantly higher ($p < 0.05$) in lowland soils due to the anaerobic condition, while there was no significant difference in soil clay contents (Figure S1). It is, therefore, suggested that soil physicochemical properties were inherently not different between upland and lowland soils as they were collected nearby fields, and have been changed by the agricultural practices, i.e., fertilization and flooding.

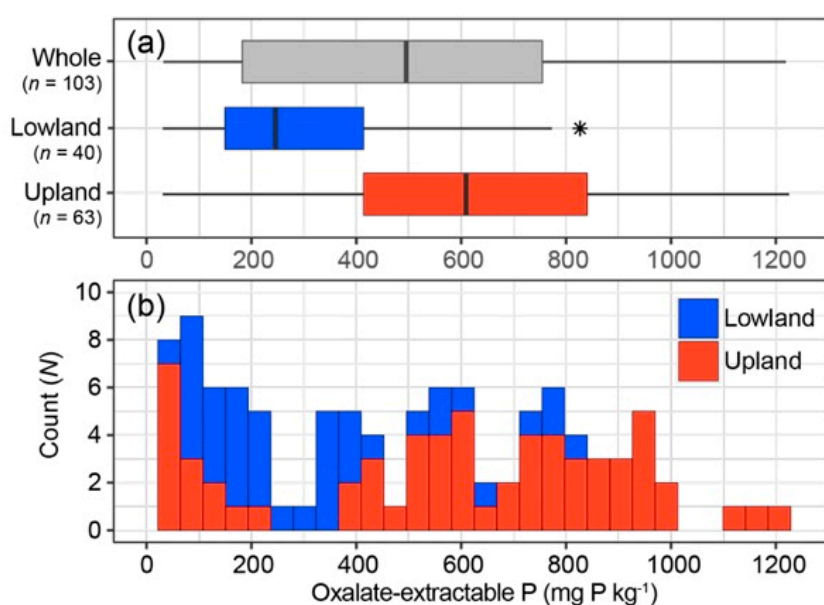


Figure 3. Box plot (a) and histogram (b) of soil oxalate-extractable P in the whole, lowland and upland data sets.

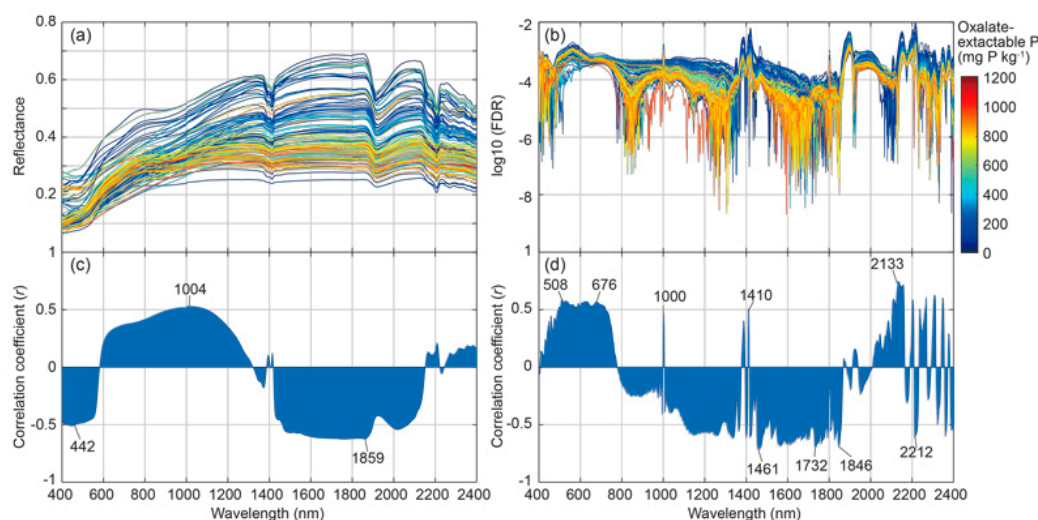
Table 2. Descriptive statistics of soil oxalate-extractable P data. *n*, number of samples; SD, standard deviation; CV, coefficient of variation (SD/mean × 100%).

Data Set	<i>n</i>	Min	Max	Median	Mean	SD	CV
Whole	103	30.73	1225.16	496.84	484.15	319.10	65.91
Upland	63	30.78	1225.16	609.23	588.74	324.93	55.70
Lowland	40	30.73	826.64	245.79	319.41	223.26	69.89

Our whole data set (upland + lowland) covered a wide range of variations in oxalate-extractable P content. The mean (and SD) values of soil oxalate-extractable P were 484.15 mg kg⁻¹ (±319.10 mg kg⁻¹), with a range of 30.73–1225.16 mg kg⁻¹, and CV = 65.91%. The SD and range of the sample affect the accuracy of soil property predictions using Vis-NIR spectroscopy [34]. In the present study, the range of soil oxalate-extractable P values was considered sufficiently large to develop the calibration models using PLS regression analyses. Our data set also demonstrated that the oxalate-extractable P content had a good correlation with the total P content in soils [6].

3.2. Soil Spectral Response and Its Correlation to Oxalate-Extractable P in Soil

Figure 4 shows the original reflectance spectra, FDR and Pearson's correlation coefficient (*r*) values between soil oxalate-extractable P content and reflectance and FDR spectra at each waveband. Large variations in the reflectance spectra were obtained from heterogeneous soil samples, which were collected from upland and lowland soils under various rice-based cropping systems. In general, soils from different fields show variations in the absorbance at wavelengths associated with iron oxides (400–500 nm), clay minerals (OH bond: 1400 and 1900 nm, Al-OH bond: 2200 nm) and organic matter (CH bond: 2300–2400 nm) in soil [50].

**Figure 4.** Raw reflectance spectra (a) and first derivative reflectance (FDR) spectra on a log10 scale (b) of the soil samples and their correlation of coefficients to soil oxalate-extractable P in each waveband (c,d).

Soil reflectance in visible regions (400–700 nm) is primarily associated with absorption in minerals containing Fe [51–53] and organic matter [54,55]. NIR regions (700–2500 nm) are dominated by absorption related to water (1400 and 1900 nm), minerals (1300–1400, 1800–1900, and 2200–2500 nm) and organic matter (1100, 1600, 1700–1800, 2000, and 2200–2400 nm) [56]. Carbonates also have weak absorption peaks in the NIR region [57]. These absorptions in the NIR region are due to overtone and combination bands primarily of C-H, N-H and O-H groups with fundamental bands related to molecular stretching that occurs in the mid-infrared (MIR) spectral region. However, there is no specific absorption by P in the Vis-NIR region, and thus, differences in the shape of reflectance

due to P are still in the process of being determined [33]. In the reflectance spectra of this study, oxalate-extractable P had a positive correlation with the wavelength between the red and shorter range NIR regions (600–1300 nm), while oxalate-extractable P had negative correlations in the blue–green region (400–580 nm) and NIR regions (1420–2100 nm). In the FDR spectra, positive correlations were found in the green–red region (508–676 nm) with some peaks in the NIR region (1000, 1410 and 2133 nm), while negative correlations were observed in the 800–1850 nm region with some peaks in the 2200–2400 nm region.

3.3. Selected Wavebands from ISE-PLS and GA-PLS Models

Selected wavebands from ISE-PLS analysis and five GA-PLS runs using FDR spectra to estimate soil oxalate-extractable P contents are shown in Figure 5, with the regression coefficient and VIP score in the FS-PLS model as information to assist in considering the importance of the selected wavelengths. In comparison to the ISE-PLS model, the GA-PLS model selected a wider range of spectral wavelength regions from within the visible (400–699 nm) and NIR (700–2400 nm) spectra, with slightly different regions for the five runs. The commonly selected regions from the five GA-PLS runs (red bar in Figure 5) were 454–457, 506–508, 517, 518, 660, 1732, 1847–1849, 1957–1961, 2105, 2107, 2109, and 2312 nm. These wavelengths did not match those identified for soil characteristics in previous studies (Table 3); in most cases, the wavelengths were found within 20 nm of previously known wavebands. As P is not spectrally active in the Vis-NIR region, the wavelengths detected in this study are potentially important spectral bands in the FDR spectra to indirectly estimate the soil oxalate-extractable P content via a link to other soil components with spectral properties.

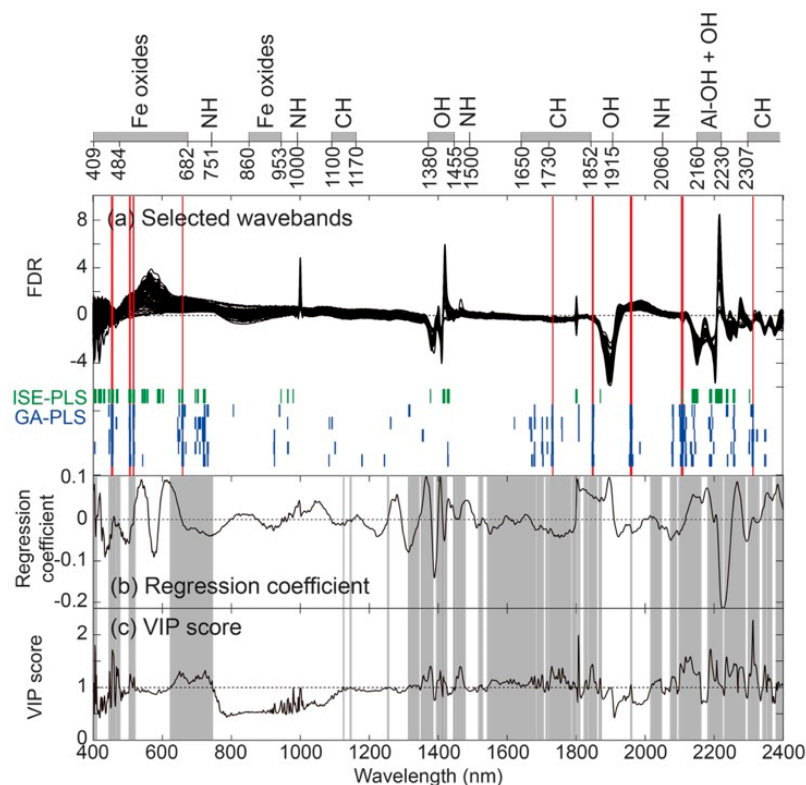


Figure 5. (a) Selected wavebands in ISE-PLS analysis (green bars) and GA-PLS (blue bars in each run) using the FDR data set ($n = 103$) to estimate oxalate-extractable P contents of upland and lowland soils, with commonly selected wavebands from five GA-PLS runs (red bars), (b) regression coefficients in the FS-PLS model, and (c) VIP score (>1 , grey bars). Specific absorption wavebands for the different bonds present in soil are specified on the top x-axis (modified by [58]).

When comparing the GA-PLS model-selected wavebands with the ISE-PLS models, the selected wavebands were different but the GA-PLS model selected wavebands in the visible region (22-bands in 454–457, 506–508, 517, 518, 660 nm) were all selected with the ISE-PLS model. In contrast, no overlapping wavebands in the NIR region were found between the GA-PLS and ISE-PLS models. ISE-PLS regression analysis is a model-wise elimination technique, while GA-PLS regression analysis is a numerical optimization method based on genetic principles and natural selection, which are slightly different models. Thus, we performed five GA-PLS runs and used the commonly selected wavebands. However, the GA-PLS models also contained the wavebands selected by ISE-PLS model plus additional bands, especially in the NIR region. In some cases, these extra regions could be readily interpreted, but in other cases, they could not [27]. Nevertheless, previous studies have reported that the wavebands selected by GA-PLS models clearly contain relevant information since the waveband subset decreases the RMSECV, and the wavebands were consistently selected in independent GA runs [27,30,31].

The commonly selected wavebands in the visible region (454–457, 506–508, 517, 518 and 660 nm) seem to be closely relevant to Fe oxide minerals (Table 3). In addition, the waveband at approximately 2270 nm, which is relevant to gibbsite (Al oxide mineral) [59], had a high regression coefficient and VIP score, although it was not always selected. These results are supported by our previous findings [6], in which oxalate-extractable P was significantly and positively correlated with active Al and Fe, respectively. The other selected wavebands likely overlapped or corresponded to previously known wavebands, which are most likely related to soil organic matter. Based on the study by Knadel et al. [56], the selected wavebands were considered to be associated with organic matter (C-H bond: 1720, 2111 and 2300 nm) [55,60,61], methyls (C-H bond: 1730–1852 nm) [61,62] and phenolics (C-OH bond: 1961 nm) [61,62]. Our previous study also found that active Al and Fe had a positive correlation with soil TC and organic phosphorus content, respectively [6]. This finding suggested that active Al and Fe play a significant role as sorbents for both oxalate-extractable P mainly in inorganic forms and organic matter in the studied soils. Turner [63] investigated the chemical nature of P in a range of rice field soils in Madagascar and reported that a considerable proportion of the TP extracted by NaOH-EDTA occurred in organic forms (19–44%), mostly as phosphate monoesters. Since the acid ammonium oxalate method can partly extract organically bound P, our results also indirectly indicated the importance of organic compounds, probably containing organic P, in oxalate-extractable P. Thus, the selected wavelengths in our data set should also be relevant to chemical associations of oxalate-extractable P in soils.

Table 3. Commonly selected wavebands from five GA-PLS runs to estimate soil oxalate-extractable P using the FDR data set ($n = 103$) and possible soil components.

Selected Waveband (nm)	Previously Known Waveband and Related Soil Component		
	Waveband (nm)	Soil Component	Reference
454–457	400–700	organic matter (color)	[11,64]
	470	Fe ³⁺ , ferric oxide	[65]
506–508, 517, 518	488–499	ferrihydrite	[52]
	495, 510	hematite	[66]
		goethite	[67]
660	660	schwertmannite	[52]
	655	organic matter	[55]
1732	1720	aliphatic C-H stretch, cellulose, lignin, starch, pectin, wax, humic acid	[60]
	1726	protein, cellulose, aliphatic C-H stretch, lignin, starch, pectin, wax,	[68]
	1730	humic acid	
1847–1849	1730–1852	methyl (C-H)	[61,62]
1957–1961	1950	sugar, starch, cellulose, lignin, protein	[68]
	1961	phenolics (C-OH)	[61,62]
	1970	smectite, shoulder due to absorbed water	[69]
2105, 2107, 2109	2111	organic matter, cellulose, glucan, pectin	[60]
	2300	C-H stretch fundamentals	[61]
2312	2307–2469	methyl	[62]
		aliphatic C-H, aromatic stretch, humic acid wax, starch	[60]
	2309	oil	[54]
	2310		

3.4. Waveband Selection with Cross-Validated Calibration Results

Figure 6 shows the relationships among the cross-validated calibration results between the FDR spectra and soil oxalate-extractable P using FS-PLS, ISE-PLS and GA-PLS regression analyses, with the selected number of wavebands (NW) and the selected NW as a percentage of the full-spectrum ($NW\% = NW / \text{whole waveband} [n = 2001] \times 100$) (Table 4). The optimum NLVs were 7, 7 and 6 using FS-PLS, ISE-PLS and GA-PLS methods, respectively, and they were determined as the lowest RMSECV values calculated from LOO-CV to avoid over-fitting of the model. Overall, the best R^2 and lowest RMSECV values were obtained with the GA-PLS model for estimating the soil oxalate-extractable P content ($R^2 = 0.796$ and $RMSECV = 143.625$). Based on $RPD > 2$ in the GA-PLS and ISE-PLS models, the quality and future applicability of our results could be considered to have a good predictive ability.

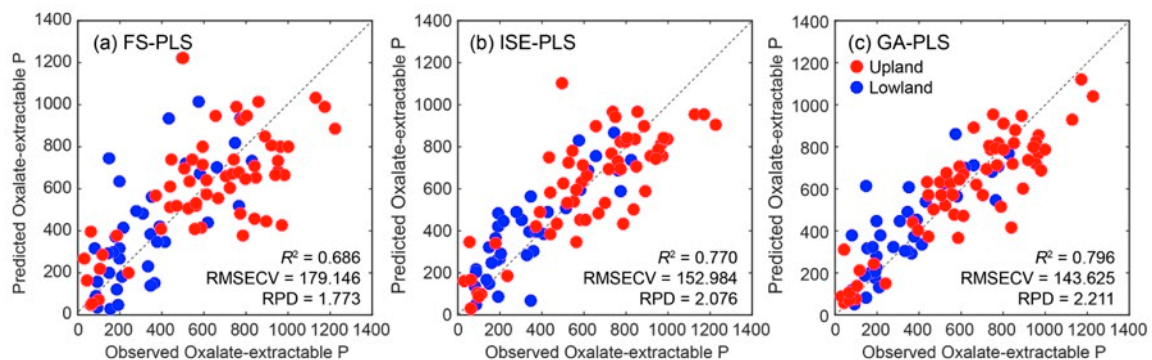


Figure 6. Relationships between observed and predicted values of soil oxalate-extractable P contents using (a) FS-PLS, (b) ISE-PLS and (c) GA-PLS regressions.

Table 4. Optimum number of latent variables (NLV), coefficient of determination (R^2), root mean squared errors of cross-validation (RMSECV), and residual predictive values (RPD) from FS-PLS, ISE-PLS and GA-PLS models with selected number of wavebands (NW) and their percentages of the full spectrum ($NW\%$). $NW\% = NW / 2001 \text{ bands} \times 100\%$.

Regression Method	Cross-Validation for Whole Data Set (n = 103)					
	NLV	R ² CV	RMSECV	RPD	NW	NW%
FS-PLS	7	0.686	179.146	1.773		
ISE-PLS	7	0.770	152.984	2.076	158	7.9
GA-PLS	6	0.796	143.625	2.211	94	4.7

The NW ($NW\%$) remaining after waveband selection was 158 (7.9%) in ISE-PLS and 94 (4.7%) in GA-PLS, which were considered useful wavelengths for estimating the soil oxalate-extractable P content. These results also suggested that over 92% of the waveband information from the soil reflectance spectrum was redundant and did not contribute to or disturb the prediction. These findings support previous findings that the performance of PLS models can be improved through waveband selection, and the most useful information in the Vis-NIR region (400–2400 nm) predicted less than 20% of the forage [30,70], water [71] and soil parameters [18]. Moreover, the spectral data efficiency is also expected to improve by the optimization of the waveband subset using the GA-PLS model [30].

3.5. Evaluation of Predictive Ability Using Modified Bootstrapping

To evaluate the predictive ability of the PLS models, a modified bootstrap procedure ($N = 10,000$ times) was conducted using selected wavebands in the FDR data set. Table 5 summarizes the mean values of NLV, R^2 and RMSECV in the training data set ($n = 69$) and R^2 , RMSEP and the percent difference of RMSEP ($\Delta RMSEP$) between FS-PLS and ISE-PLS or GA-PLS models in the test data set ($n = 34$). In addition, Figure 7 demonstrates the distribution of R^2 values in the test data set. The mean

optimum NLV ranged from 5.364 in the GA-PLS model to 7.285 in the FS-PLS model. In the training data set, GA-PLS obtained the best mean R^2 (0.782) and the lowest mean RMSECV (148.930 mg P kg⁻¹) values, and ISE-PLS performed better than FS-PLS. Similarly, in the test data set, the GA-PLS model obtained the best mean R^2 (0.787) and the lowest mean RMSEP (149.013 mg P kg⁻¹) values for estimating soil oxalate-extractable P. In comparison with the FS-PLS model and GA-PLS, the Δ RMSEP showed greater predictive accuracies in ISE-PLS (−16.21%) and GA-PLS (−24.69%) models, respectively.

Table 5. Mean values of NLV, R^2 and RMSECV/RMSEP from $N = 10,000$ evaluations using independent training and test data sets with FS-PLS, ISE-PLS and GA-PLS.

Regression Method	Training Data Set ($n = 69$)			Test Data Set ($n = 34$)		
	Mean NLV	Mean R^2	Mean RMSECV	Mean R^2	Mean RMSEP	Δ RMSEP ¹
FS-PLS	7.285	0.659	188.560	0.638	197.860	
ISE-PLS	6.419	0.751	160.180	0.742	165.786	−16.21
GA-PLS	5.364	0.782	148.930	0.787	149.013	−24.69

¹ Δ RMSEP, percent difference in the RMSEP to FS-PLS.

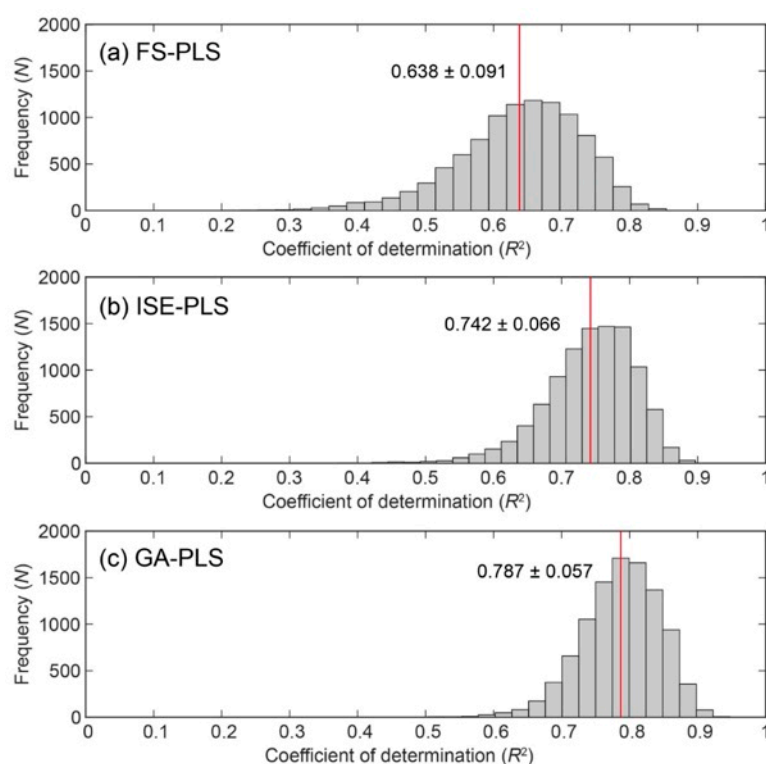


Figure 7. Comparisons of the frequency distributions of R^2 values in the test data ($n = 34$) from (a) FS-PLS, (b) ISE-PLS and (c) GA-PLS using FDR, with mean (red line) \pm standard deviation (SD) values.

These findings confirm previous results that showed that the performance of PLS models can be improved through wavelength selection [20,70,72] and the predictive ability of GA-PLS can overcome the ISE-PLS [29–31]. Yang et al. [16] suggested that reducing large spectral data sets is valuable for more efficient storage, computation, and transmission as well as for the ease of spectral analysis.

Although spectral data efficiency could be improved by the optimization of a wavelength subset in the PLS model, the over-fitting problems still remained in the GA-PLS method. Leardi and Nørgaard [73] addressed a limitation of GA-PLS, which was that a greater number of variables (> 200) would result in over-fitting and reduce the capability of obtaining a solution with good predictive

ability. To solve this problem, they proposed a sequential application of backward-interval PLS (biPLS) and GA for the selection of relevant spectral regions. The biPLS removes the non-informative regions prior to GA runs, thereby reducing the number of variables. Future work needs to examine using such sequential application of biPLS and GA-PLS with a larger number of data sets collected in different fields with a range of various soil properties.

4. Conclusions

A timely and accurate assessment of the soil P content is crucial for resource-limited farmers in Madagascar to improve rice production by site-specific fertilizer management. In this study, we investigated the performance of GA-PLS regression analysis in laboratory Vis-NIR reflectance spectroscopy for estimating soil oxalate-extractable P contents as a diagnostic indicator of soil P status in rice fields of Madagascar. Our results showed that a large range of soil oxalate-extractable P (30.73 to 1225.16 mg P kg⁻¹) can be rapidly and non-destructively predicted by Vis-NIR spectroscopy for rice fields irrespective of different cropping systems and geographical locations and that the predictive ability was improved by GA-based waveband selection coupled with PLS regression analysis. GA-based waveband selection in the PLS calibration suggested that the important wavebands for estimating soil oxalate-extractable P were 4.7% of all 2001 wavebands in the 400–2400 nm range. The selected wavebands were different from previously published absorption peaks of specific materials. However, most of the peaks were within the 20 nm vicinity of such a peak and apparently relevant to chemical associations of oxalate-extractable P in soils bound to Al and Fe oxides and organic compounds. Thus, the selected wavelength in our study should be considered informative for estimating soil oxalate-extractable P contents. Based on the selected FDR wavebands in the GA-PLS model, soil oxalate-extractable P was determined to provide a good prediction (RPD = 2.211), with 20.4% and 21.3% of errors when cross-validating and testing, respectively, the independent data set. Such timely P sensing in soils might allow Madagascar's farmers to implement better fertilizer management.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/11/5/506/s1>, Figure S1: Box plot and histogram of soil total carbon (TC) and clay contents in the lowland and 2 upland data sets.

Author Contributions: K.K., Y.T., H.A., T.R. and T.R. designed this study and the field work; Y.T., T.N., A.A., M.R. and H.A. performed the field work and carried out the soil chemical analyses; K.K., A.A. and M.R. performed the laboratory spectral measurements and data processing; K.K., T.N. and Y.T. wrote the manuscript; and all the authors revised the paper.

Funding: This research was supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA).

Acknowledgments: We are grateful to Ms. Marie-Paule Razafimanantsoa and her laboratory staff members of Laboratoire des Radio-Isotopes, Université d'Antananarivo, for their support in conducting the oxalate-extractable P analysis. We would like to especially thank Dr. Naoki Moritsuka, Graduate School of Agriculture, Kyoto University in Japan, for his valuable comments on this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dogbe, W.; Sogbedji, J.M.; Buah, S.S.J. Site-specific Nutrient Management for Lowland Rice in the Northern Savannah Zones of Ghana. *Curr. Agric. Res. J.* **2015**, *3*, 109–117. [[CrossRef](#)]
2. Kone, B.; Fofana, M.; Sorho, F.; Diatta, S.; Ogunbayo, A.; Sie, M. Nutrient constraint of rainfed rice production in foot slope soil of Guinea Forest in Côte d'Ivoire. *Arch. Agron. Soil Sci.* **2014**, *60*, 735–746. [[CrossRef](#)]
3. Koné, B.; Amadji, G.L.; Aliou, S.; Diatta, S.; Akakpo, C. Nutrient constraint and yield potential of rice on upland soil in the south of the Dahoumey gap of West Africa. *Arch. Agron. Soil Sci.* **2011**, *57*, 763–774. [[CrossRef](#)]

4. Tamburini, F.; Bernasconi, S.M.; Paytan, A. Phosphorus in the environment. In *Eos*; White, P.J., Hammond, J.P., Eds.; Springer Netherlands: Dordrecht, The Netherlands, 2012; Volume 93, p. 405. ISBN1 978-1-4020-8434-8. ISBN2 978-1-4020-8435-5.
5. Balemi, T.; Negisho, K. Management of soil phosphorus and plant adaptation mechanisms to phosphorus stress for sustainable crop production: A review. *J. Soil Sci. Plant Nutr.* **2012**, *12*, 547–562. [[CrossRef](#)]
6. Nishigaki, T.; Tsujimoto, Y.; Rinasoa, S.; Rakotoson, T.; Andriamananjara, A.; Razafimbelo, T. Phosphorus uptake of rice plants is affected by phosphorus forms and physicochemical properties of tropical weathered soils. *Plant Soil* **2018**, *435*, 27–38. [[CrossRef](#)]
7. Wuenschel, R.; Unterfrauner, H.; Peticzka, R.; Zehetner, F. A comparison of 14 soil phosphorus extraction methods applied to 50 agricultural soils from Central Europe. *Plant Soil Environ.* **2015**, *61*, 86–96. [[CrossRef](#)]
8. Helfenstein, J.; Tamburini, F.; von Sperber, C.; Massey, M.S.; Pistocchi, C.; Chadwick, O.A.; Vitousek, P.M.; Kretzschmar, R.; Frossard, E. Combining spectroscopic and isotopic techniques gives a dynamic view of phosphorus cycling in soil. *Nat. Commun.* **2018**, *9*, 3226. [[CrossRef](#)] [[PubMed](#)]
9. Rabeharisoa, L.; Razanakoto, O.R.; Razafimanantsoa, M.P.; Rakotoson, T.; Amery, F.; Smolders, E. Larger bioavailability of soil phosphorus for irrigated rice compared with rainfed rice in Madagascar: Results from a soil and plant survey. *Soil Use Manag.* **2012**, *28*, 448–456. [[CrossRef](#)]
10. Sims, J.T.; Sharpley, A.N.; Condron, L.M.; Turner, B.L.; Cade-Menun, B.J. Chemistry and Dynamics of Soil Organic Phosphorus. *Phosphorus Agric. Environ.* **2005**, 87–121. [[CrossRef](#)]
11. Viscarra Rossel, R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [[CrossRef](#)]
12. Miller, C.E. Chemical principles of near-infrared technology. In *Near Infrared Technology in the Agricultural and Food Industries*; Williams, P.C., Horris, K.H., Eds.; American Association of Cereal Chemists: St. Paul, MN, USA, 2001; pp. 19–37.
13. Cañasveras, J.C.; Barrón, V.; del Campillo, M.C.; Torrent, J.; Gómez, J.A. Estimation of aggregate stability indices in Mediterranean soils by diffuse reflectance spectroscopy. *Geoderma* **2010**, *158*, 78–84. [[CrossRef](#)]
14. Cañasveras Sánchez, J.C.; Barrón, V.; del Campillo, M.C.; Viscarra Rossel, R.A. Reflectance spectroscopy: A tool for predicting soil properties related to the incidence of Fe chlorosis. *Span. J. Agric. Res.* **2012**, *10*, 10.
15. Viscarra Rossel, R.A.; Behrens, T.; Ben-Dor, E.; Brown, D.J.; Demattê, J.A.M.; Shepherd, K.D.; Shi, Z.; Stenberg, B.; Stevens, A.; Adamchuk, V.; et al. A global spectral library to characterize the world's soil. *Earth-Sci. Rev.* **2016**, *155*, 198–230. [[CrossRef](#)]
16. Yang, H.; Kuang, B.; Mouazen, A.M. Quantitative analysis of soil nitrogen and carbon at a farm scale using visible and near infrared spectroscopy coupled with wavelength reduction. *Eur. J. Soil Sci.* **2012**, *63*, 410–420. [[CrossRef](#)]
17. Vohland, M.; Ludwig, M.; Thiele-Bruhn, S.; Ludwig, B. Determination of soil properties with visible to near- and mid-infrared spectroscopy: Effects of spectral variable selection. *Geoderma* **2014**, *223*, 88–96. [[CrossRef](#)]
18. Kawamura, K.; Tsujimoto, Y.; Rabenarivo, M.; Asai, H.; Andriamananjara, A.; Rakotoson, T. Vis-NIR spectroscopy and PLS regression with waveband selection for estimating the total C and N of paddy soils in Madagascar. *Remote Sens.* **2017**, *9*, 1081. [[CrossRef](#)]
19. Bolster, K.L.; Martin, M.E.; Aber, J.D. Determination of carbon fraction and nitrogen concentration in tree foliage by near infrared reflectance: A comparison of statistical methods. *Can. J. For. Res.* **1996**, *26*, 590–600. [[CrossRef](#)]
20. Kawamura, K.; Watanabe, N.; Sakanoue, S.; Inoue, Y. Estimating forage biomass and quality in a mixed sown pasture based on PLS regression with waveband selection. *Grassl. Sci.* **2008**, *54*, 131–146. [[CrossRef](#)]
21. Boggia, R.; Forina, M.; Fossa, P.; Mosti, L. Chemometric study and validation strategies in the structure-activity relationships of new cardiotoxic agents. *Quant. Struct. Relatsh.* **1997**, *16*, 201–213. [[CrossRef](#)]
22. Centner, V.; Massart, D.L.; de Noord, O.E.; de Jong, S.; Vandeginste, B.M.; Sterna, C. Elimination of uninformative variables for multivariate calibration. *Anal. Chem.* **1996**, *68*, 3851–3858. [[CrossRef](#)] [[PubMed](#)]
23. Li, H.; Liang, Y.; Xu, Q.; Cao, D. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal. Chim. Acta* **2009**, *648*, 77–84. [[CrossRef](#)] [[PubMed](#)]

24. Nørgaard, L.; Saudland, A.; Wagner, J.; Nielsen, J.P.; Munck, L.; Engelsen, S.B. Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Appl. Spectrosc.* **2000**, *54*, 413–419. [[CrossRef](#)]
25. Jiang, J.H.; James, R.; Siesler, B.H.W.; Ozaki, Y. Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data. *Anal. Chem.* **2002**, *74*, 3555–3565. [[CrossRef](#)] [[PubMed](#)]
26. Leardi, R. Application of a genetic algorithm to feature selection under full validation conditions and to outlier detection. *J. Chemom.* **1994**, *8*, 65–79. [[CrossRef](#)]
27. Leardi, R. Application of genetic algorithm-PLS for feature selection in spectral data sets. *J. Chemom.* **2000**, *14*, 643–655. [[CrossRef](#)]
28. Leardi, R.; González, A.L. Genetic algorithms applied to feature selection in PLS regression: How and when to use them. *Chemom. Intell. Lab. Syst.* **1998**, *41*, 195–207. [[CrossRef](#)]
29. Lucasius, C.B.; Kateman, G. Understanding and using genetic algorithms Part 2. Representation, configuration and hybridization. *Chemom. Intell. Lab. Syst.* **1994**, *25*, 99–145. [[CrossRef](#)]
30. Kawamura, K.; Watanabe, N.; Sakanoue, S.; Lee, H.J.; Inoue, Y.; Odagawa, S. Testing genetic algorithm as a tool to select relevant wavebands from field hyperspectral data for estimating pasture mass and quality in a mixed sown pasture using partial least squares regression. *Grassl. Sci.* **2010**, *56*, 205–216. [[CrossRef](#)]
31. Kawamura, K.; Watanabe, N.; Sakanoue, S.; Lee, H.J.; Lim, J.; Yoshitoshi, R. Genetic algorithm-based partial least squares regression for estimating legume content in a grass-legume mixture using field hyperspectral measurements. *Grassl. Sci.* **2013**, *59*, 166–172. [[CrossRef](#)]
32. Bogrekcı, I.; Lee, W.S. Spectral soil signatures and sensing phosphorus. *Biosyst. Eng.* **2005**, *92*, 527–533. [[CrossRef](#)]
33. Maleki, M.R.; Van Holm, L.; Ramon, H.; Merckx, R.; De Baerdemaeker, J.; Mouazen, A.M. Phosphorus Sensing for Fresh Soils using Visible and Near Infrared Spectroscopy. *Biosyst. Eng.* **2006**, *95*, 425–436. [[CrossRef](#)]
34. Kuang, B.; Mouazen, A.M. Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *Eur. J. Soil Sci.* **2011**, *62*, 629–636. [[CrossRef](#)]
35. Tsujimoto, Y.; Horie, T.; Randriamihary, H.; Shiraiwa, T.; Homma, K. Soil management: The key factors for higher productivity in the fields utilizing the system of rice intensification (SRI) in the central highland of Madagascar. *Agric. Syst.* **2009**, *100*, 61–71. [[CrossRef](#)]
36. IUSS Working Group, WRB. *World Reference Base for Soil Resources 2014, Update 2015 International Soil Classification System for Naming Soils and Creating Legends for Soil Maps*; World Soil Resources Reports No. 106; Food and Agriculture Organization of the United Nations: Rome, Italy, 2015.
37. Soil Survey Staff. *Keys to Soil Taxonomy*, 12th ed.; USDA-Natural Resources Conservation Service: Washington, DC, USA, 2014.
38. Schwertmann, U. The differentiation of iron oxides in soils by extraction with ammonium oxalate solution. *Z. Pflanz. Bodenkd.* **1964**, *105*, 194–202. [[CrossRef](#)]
39. van Veldhoven, P.P.; Mannaerts, G.P. Inorganic and organic phosphate measurements in the nanomolar range. *Anal. Biochem.* **1987**, *161*, 45–48. [[CrossRef](#)]
40. Reeves, J.; McCarty, G.; Mimmo, T. The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environ. Pollut.* **2002**, *116*, S277–S284. [[CrossRef](#)]
41. Savitzky, A.; Golay, E.J.M. Smoothing and difference of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
42. Brunet, D.; Barthès, B.G.; Chotte, J.-L.; Feller, C. Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma* **2007**, *139*, 106–117. [[CrossRef](#)]
43. Forina, M.; Lanteri, S.; Oliveros, M.C.C.; Millan, C.P. Selection of useful predictors in multivariate calibration. *Anal. Bioanal. Chem.* **2004**, *380*, 397–418. [[CrossRef](#)] [[PubMed](#)]
44. Leardi, R.; Boggia, R.; Terrile, M. Genetic Algorithms as a strategy for feature selection. *J. Chemom.* **1992**, *6*, 267–281. [[CrossRef](#)]
45. Ding, Q.; Small, G.W.; Arnold, M.A. Genetic algorithm-based wavelength selection for the near-infrared determination of glucose in biological matrixes: Initialization strategies and effects of spectral resolution. *Anal. Chem.* **1998**, *70*, 4472–4479. [[CrossRef](#)] [[PubMed](#)]

46. Leardi, R.; Seasholtz, M.B.; Pell, R.J. Variable selection for multivariate calibration using a genetic algorithm: Prediction of additive concentrations in polymer films from Fourier transform-infrared spectral data. *Anal. Chim. Acta* **2002**, *461*, 189–200. [[CrossRef](#)]
47. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
48. Chong, I.-G.; Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112. [[CrossRef](#)]
49. Li, B.; Liew, O.W.; Asundi, A.K. Pre-visual detection of iron and phosphorus deficiency by transformed reflectance spectra. *J. Photochem. Photobiol. B Biol.* **2006**, *85*, 131–139. [[CrossRef](#)] [[PubMed](#)]
50. Ben-Dor, E. Quantitative remote sensing of soil properties. In *Advances in Agronomy*; Academic Press: New York, NY, USA, 2002; Volume 75, pp. 173–243. ISBN 9780120007936.
51. Drăguț, L.; Dornik, A. Land-surface segmentation as a method to create strata for spatial sampling and its potential for digital soil mapping. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 1359–1376. [[CrossRef](#)]
52. Scheinost, A.C.; Chavernas, A.; Barrón, V.; Torrent, J. Use and limitations of second-derivative diffuse reflectance spectroscopy in the visible to near-infrared range to identify and quantify Fe oxide minerals in soils. *Clays Clay Miner.* **1998**, *46*, 528–536. [[CrossRef](#)]
53. Mortimore, J.L.; Marshall, L.-J.R.; Almond, J.M.; Hollins, P.; Matthews, W. Analysis of red and yellow ochre samples from Clearwell Caves and Çatalhöyük by vibrational spectroscopy and other techniques. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2004**, *60*, 1179–1188. [[CrossRef](#)] [[PubMed](#)]
54. Shonk, J.L.; Gaultney, L.D.; Schulze, D.G.; Van Scoyoc, G.E. Spectroscopic sensing of soil organic-matter content. *Trans. ASAE* **1991**, *34*, 1978–1984. [[CrossRef](#)]
55. Daniel, K.W.; Tripathi, N.K.; Honda, K. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Aust. J. Soil Res.* **2003**, *41*, 47–59. [[CrossRef](#)]
56. Knadel, M.; Viscarra Rossel, R.A.; Deng, F.; Thomsen, A.; Greve, M.H. Visible–Near Infrared Spectra as a Proxy for Topsoil Texture and Glacial Boundaries. *Soil Sci. Soc. Am. J.* **2013**, *77*, 568. [[CrossRef](#)]
57. Hunt, G.H.; Salisbury, J.W. Visible and Near Infrared Spectra of Minerals and Rocks: XI. Sedimentary Rocks. *Mod. Geol.* **1976**, *5*, 211–217.
58. Katuwal, S.; Knadel, M.; Moldrup, P.; Norgaard, T.; Greve, M.H.; de Jonge, L.W. Visible–Near-Infrared Spectroscopy can predict Mass Transport of Dissolved Chemicals through Intact Soil. *Sci. Rep.* **2018**, *8*, 11188. [[CrossRef](#)] [[PubMed](#)]
59. Ramarason, V.H.; Becquer, T.; Sá, S.O.; Razafimahatratra, H.; Delarivière, J.L.; Blavet, D.; Vendrame, P.R.S.; Rabeharisoa, L.; Rakotondrazafy, A.F.M. Mineralogical analysis of ferralitic soils in Madagascar using NIR spectroscopy. *CATENA* **2018**, *168*, 102–109. [[CrossRef](#)]
60. Ben-Dor, E.; Inbar, Y.; Chen, Y. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sens. Environ.* **1997**, *61*, 1–15. [[CrossRef](#)]
61. Clark, R.N.; King, T.V.V.; Klejwa, M.; Swayze, G.A.; Vergo, N. High spectral resolution reflectance spectroscopy of minerals. *J. Geophys. Res.* **1990**, *95*, 12653–12680. [[CrossRef](#)]
62. Clark, R.N. Spectroscopy of rocks and minerals, and principles of spectroscopy. In *Manual of Remote Sensing*; John Wiley and Sons, Inc.: Chichester, UK, 1999; pp. 3–58.
63. Turner, B.L. Organic phosphorus in Madagascan rice soils. *Geoderma* **2006**, *136*, 279–288. [[CrossRef](#)]
64. Viscarra Rossel, R.A.; Fouad, Y.; Walter, C. Using a digital camera to measure soil organic carbon and iron contents. *Biosyst. Eng.* **2008**, *100*, 149–159. [[CrossRef](#)]
65. Hunt, G.R. Spectral Signatures of Particulate Minerals in the Visible and Near Infrared. *Geophysics* **1977**, *42*, 501–513. [[CrossRef](#)]
66. Sherman, D.M.; Waite, D.T. Electronic spectra of Fe³⁺ oxides and oxide hydroxides in the near IR to near UV. *Am. Mineral.* **1985**, *70*, 1262–1269.
67. Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind, J. Visible and Near Infrared Spectroscopy in Soil Science. *Adv. Agron.* **2010**, *107*, 163–215.
68. Fourty, T.; Baret, F.; Jacquemoud, S.; Schmuck, G.; Verdebout, J. Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems. *Remote Sens. Environ.* **1996**, *56*, 104–117. [[CrossRef](#)]




69. Bishop, J.L.; Pieters, C.M.; Edwards, J.O. Infrared spectroscopic analyses on the nature of water in montmorillonite. *Clays Clay Miner.* **1994**, *42*, 702–716. [[CrossRef](#)]
70. Darvishzadeh, R.; Skidmore, A.; Atzberger, C.; van Wieren, S. Estimation of vegetation LAI from hyperspectral reflectance data: Effects of soil type and plant architecture. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *in press*, 358–373. [[CrossRef](#)]
71. Wang, Z.; Kawamura, K.; Sakuno, Y.; Fan, X.; Gong, Z.; Lim, J. Retrieval of chlorophyll-a and total suspended solids using iterative stepwise elimination partial least squares (ISE-PLS) regression based on field hyperspectral measurements in irrigation ponds in Higashihiroshima, Japan. *Remote Sens.* **2017**, *9*, 264. [[CrossRef](#)]
72. Holst, J.; Liu, C.; Yao, Z.; Brüggemann, N.; Zheng, X.; Han, X.; Butterbach-Bahl, K. Importance of point sources on regional nitrous oxide fluxes in semi-arid steppe of Inner Mongolia, China. *Plant Soil* **2007**, *296*, 209–226. [[CrossRef](#)]
73. Leardi, R.; Nørgaard, L. Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions. *J. Chemom.* **2004**, *18*, 486–497. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

Prediction of Soil Oxalate Phosphorus using Visible and Near-Infrared Spectroscopy in Natural and Cultivated System Soils of Madagascar

Hobimiarantsoa Rakotonindrina ¹, Kensuke Kawamura ² , Yasuhiro Tsujimoto ² ,
Tomohiro Nishigaki ², Herintsitohaina Razakamanarivo ¹, Bruce Haja Andrianary ¹ and
Andry Andriamananjara ^{1,*} 

¹ Laboratoire des Radiosotopes, Université d'Antananarivo, BP 3383, Route d'Andraisoro, 101 Antananarivo, Madagascar; hobimiarantsoa@gmail.com (H.R.); herintsitohaina.razakamanarivo@gmail.com (H.R.); hajabruce@yahoo.fr (B.H.A.)

² Japan International Research Center for Agricultural Sciences (JIRCAS), 1-1 Ohwashi, Tsukuba, Ibaraki 305-8686, Japan; kamuken@affrc.go.jp (K.K.); tsjmt@affrc.go.jp (Y.T.); nishigaki@affrc.go.jp (T.N.)

* Correspondence: njaraandry1@gmail.com; Tel.: +261-34-12-569-62

Received: 23 March 2020; Accepted: 15 May 2020; Published: 16 May 2020



Abstract: Phosphorus is among the main limiting nutrients for plant growth and productivity in both agricultural and natural ecosystems in the tropics, which are characterized by weathered soil. Soil bioavailable P measurement is necessary to predict the potential growth of plant biomass in these ecosystems. Visible and near-infrared reflectance spectroscopy (Vis-NIRS) is widely used to predict soil chemical and biological parameters as an alternative to time-consuming conventional laboratory analyses. However, quantitative spectroscopic prediction of soil P remains a challenge owing to the difficulty of direct detection of orthophosphate. This study tested the performance of Vis-NIRS with partial least square regression to predict oxalate-extractable P (Pox) content, representing available P for plants in natural (forest and non-forest including fallows and degraded land) and cultivated (upland and flooded rice fields) soils in Madagascar. Model predictive accuracy was assessed based on the coefficient of determination (R^2), the root mean squared error of cross-validation (RMSECV), and the residual predictive deviation (RPD). The results demonstrated successful Pox prediction accuracy in natural ($n = 74$, $R^2 = 0.90$, RMSECV = 2.39, and RPD = 3.22), and cultivated systems ($n = 142$, $R^2 = 0.90$, RMSECV = 48.57, and RPD = 3.15) and moderate usefulness at the regional scale incorporating both system types ($R^2 = 0.70$, RMSECV = 71.87 and RPD = 1.81). These results were also confirmed with modified bootstrap procedures (N = 10,000 times) using selected wavebands on iterative stepwise elimination–partial least square (ISE–PLS) models. The wavebands relevant to soil organic matter content and Fe content were identified as important components for the prediction of soil Pox. This predictive accuracy for the cultivated system was related to the variability of some samples with high Pox values. However, the use of “pseudo-independent” validation can overestimate the prediction accuracy when applied at site scale suggesting the use of larger and dispersed geographical cover sample sets to build a robust model. Our study offers new opportunities for P quantification in a wide range of ecosystems in the tropics.

Keywords: Madagascar; partial least square (PLS) regression; precision farming; soil oxalate phosphorus; spectroscopy

1. Introduction

Phosphorus is an essential plant nutrient. The low P availability of strongly weathered soil can seriously affect plant growth and limit crop yields [1] while soil P limitation can cause a decline in

climax ecosystems by decreasing biomass productivity [2]. The limitation of net primary production in terrestrial ecosystems with low soil P, such as tropical forests, leads to a carbon balance that tends to increase CO₂ release [3]. Ferralsols are characterized by very low available soil P, mainly due to sorption on and in Fe and Al oxyhydroxides [4,5]. The availability of P for plants has been evaluated from soil P tests and calibrated with field and pot experiments [6,7]. Oxalate-extractable P (Pox) is reported to accurately predict the availability of P in highly weathered soil [8] because of oxalate's potential to extract the active reductant-soluble P fraction [9]. The quantification of Pox is based on P extraction with ammonium oxalate and oxalic acid [10]. Acidified ammonium oxalate extractant dissolves amorphous, poorly crystalline oxides, and hydroxides of Fe and Al, and consequently released P [4,9,11]. Oxalate P is highly correlated with rice plant P uptake in lowland and upland fields [4,12]. It also extracts more P than other chemical methods [13,14]. Pox is thus the best indicator of P availability for both fertilizer management in agricultural systems and natural ecosystem management.

There is a need for more reliable, rapid, and accurate soil P assessment as an alternative to time-consuming conventional laboratory analyses. The more rapid, cost-effective alternative approaches of spectrometry analysis and chemometric techniques have been widely used to estimate soil and plant compositions [15–17]. Spectrometry in visible and near infra-red (Vis-NIR) is based on the absorption of radiation at a specific wavelength by certain molecular bonds in the near-infrared (NIR) region [18]. Spectral data are calibrated using the specific soil properties of samples, determined by conventional methods. The Vis-NIR region (400–2500 nm) is dominated by weaker and broader signals from vibration overtones and combination bands [19]. The absorption coefficients are much lower, which allows for better penetration of light into the material [20]. The Vis-NIRS approach has been successfully applied to predict soil chemical and biological parameters [21–23].

Challenges, however, still remain. P is not spectrally active in the Vis-NIR region, and thus, it can only be detected indirectly, in an organically bound form [24]. However, previous studies have shown the potential of this approach to predict soil P, such as that of Kawamura et al. [25], who reported that Vis-NIRS coupled with partial least square (PLS) regression can predict soil Pox in Malagasy lowland and upland rice fields with moderate accuracy ($R^2 = 0.78$). These authors also suggested that the performance of PLS models could be improved through wavelength selection. However, the PLS models were developed only for rice fields, with a dataset (106 samples) obtained in central Madagascar, and their applicability to non-farm soils is still unknown. Thus, further analysis using a larger number of datasets including other land uses is needed. In Madagascar, almost 45% of the cultivated area is occupied by rice fields, as rice is the staple food crop for Malagasy people and is cultivated by 85% of farmers [26,27]. Moreover, the so-called “natural system”, as found in eastern Madagascar, was characterized by traditional farming practices in which forest and fallow land are subject to slash and burn agriculture. Land-use change affects soil properties [28] and assessment of these is in turn required to inform land management practices. Therefore, the present study evaluated the potential of the PLS model to predict Pox across different Malagasy land-use systems, including those where the natural system has been converted into a cultivated system.

This study aimed to investigate the usefulness of the Vis-NIRS approach with PLS modeling in predicting soil Pox in cultivated and natural systems and its applicability as a rapid method to assess soil properties at the ecosystem scale. To improve its predictive accuracy, we applied wavelength selection in the PLS procedures and compared its performance against standard full-spectrum PLS (FS-PLS) in cultivated and natural systems.

2. Materials and Methods

2.1. Study Area and Soil Sample Dataset

The soil samples used for this study were collected from areas of central and eastern Madagascar (Figure 1). The central sites, located in the Vakinankaratra region, were characterized by a humid climate with an annual mean precipitation of 1381 mm and a mean annual temperature of 16.9 °C.

They are dominated by ferritic soils (FAO soil classification) which are generally acid with low available phosphorus [29,30]. The Vakinankaratra region is also among the rice-growing areas of Madagascar. The eastern sites are characterized by perhumid and humid climates with a mean annual rainfall of 2500 mm and a mean annual temperature of 18–24 °C [31,32]. This region is characterized by red and yellow ferralsols [33].

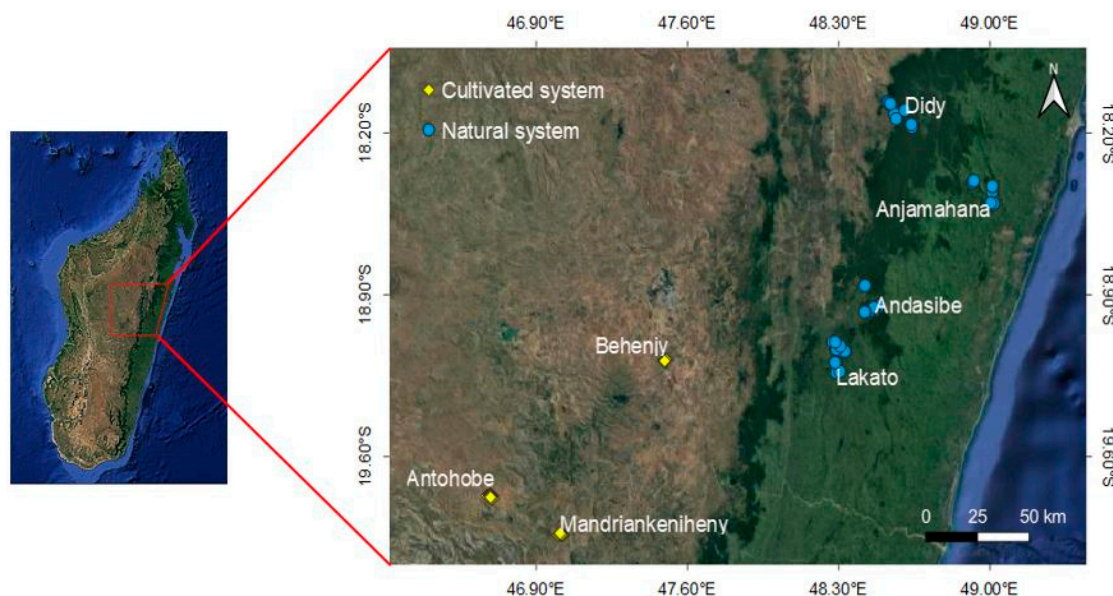


Figure 1. Spatial distribution of site for sampling. Cultivated system samples were from the Vakinankaratra region and natural system samples were from eastern Madagascar.

In the Vakinankaratra area, soil sampling at 15 cm depth was conducted in 142 farmer field plots under irrigated and upland rice systems during 2018 and 2019 (Table 1). In eastern Madagascar, soil samples were collected similarly during 2014 and 2015 from 74 forest and non-forest plots, the latter including fallow and degraded land systems [34]. The descriptive statistics of soil parameters for each studied site are reported in Table 2.

Table 1. Spatial characterization of the soil sample used for the study.

Regions	System	Land Uses	Altitude (m)	MAT (°C)	MAP (mm)	Sampling Year	Number of Samples
Central (Vakinankaratra)	Cultivated systems	Upland rice	1247–1481	16.9	1381	2017–2018	8
		Lowland rice	1237–1481			2017–2018	134
Eastern	Natural systems	Forest	134–1200	18–24	2500	2014–2015	16
		Non-Forest	94–1101			2014–2015	58

MAT, Mean annual temperature (°C), MAP, Mean annual precipitation (mm).

Table 2. Soil parameters description of the study sites. The values in parentheses show the range.

Soil Parameters	Cultivated System Area	Natural System Area
Sand (%)	34.6 [10.4–72.5]	53.6 [30.8–80.6]
Silt (%)	32.8 [7.92–63.7]	14.4 [4.72–23.6]
Clay (%)	32.6 [4.30–52.0]	32.0 [9.45–53.6]
SOC (mg kg ⁻¹)	25.5 [9.47–94.9]	37.9 [7.29–75.4]
Feox (g kg ⁻¹)	7.44 [1.03–19.1]	2.38 [0.32–9.45]
Pox (mg kg ⁻¹)	115.1 [22.3–856.8]	35.1 [21.9–57.9]

SOC—Soil organic carbon, Feox—oxalate-extractable Fe, Pox—oxalate-extractable P.

2.2. Laboratory Analyses

Soils were air-dried, ground, and sieved through 2 mm and 0.2 mm mesh prior to chemical analysis. All soil samples were analyzed for texture and for phosphorus and organic carbon contents. Oxalate-extractable P and Fe were determined following Schwertmann [10]. Soil organic carbon was determined by wet combustion using dichromate oxidation [35]. The separation of soil fractions for the soil texture analysis was carried out with the pipetting method in which soil samples pretreated with heat and H₂O₂ (35%) to remove organic matter are dispersed into clay, silt and sand fractions using NaOH.

2.3. Spectral Data Acquisition Using Vis-NIRS

Spectral data were recorded in a dark room at the Laboratoire des Radioisotopes, Antananarivo University using a Vis-NIR portable spectro-radiometer with 350–2500 nm range (ASD FieldSpec 4 Hi-Res, ASD Inc., Longmont, CO, USA). The recorded spectral resolution was 3 nm between 350 nm and 1000 nm and 6 nm between 1000 nm and 2500 nm. The output data were generated at 1 nm resolution using the cubic spline interpolation function in the ASD software (RS3 for Windows; ASD). Before each measurement, the spectrometer was calibrated using a white reference spectrum [17]. Soil samples were previously spread and leveled in optical-glass Petri dishes 85 mm in diameter. Five measurements were carried out at different positions for each soil sample. For each measurement, the instrument made 25 internal scans to optimize the signal-to-noise ratio. The generated spectra were averaged into one spectrum for each sample. Further details can be found in Kawamura et al. [36].

2.4. Spectral Analyses and Modeling Approaches

Prior to the modeling of Pox using PLS regression, data pre-processing was applied. Spectral data were reduced to 400–2400 nm by removing the spectral regions of 350–399 nm and 2401–2500 nm, in order to eliminate the influence of noise [36–38]. The reflectance spectra (R) were transformed into apparent absorbance ($A = \log(1/R)$). To reduce noise and enhance the signals, first derivative reflectance (FDR) using a Savitzky–Golay smoothing filter [39] was used with an order 3 polynomial. The generated Vis-NIR spectra were mean-centered. Scatter correction using a standard normal variate transform (SNV) was applied to all spectra to reduce the particle size effect.

The modeling approach consisted of testing whether these reflectance spectra could be used to predict chemical data and identifying which spectral regions contribute to the prediction [40]. The PLS model incorporated the algorithms that extract a small number of latent factors as the independent variables relating to reflectance spectra, then used these factors in regression analysis with the chemical data as the dependent variables. The PLS regression model describing the relationship between soil spectra and measured soil Pox was built from the spectroscopic modeling. Leave-one-out cross-validation was used to select the best latent variable number and to avoid over-fitting of the PLS regression model [36,38,41]. The optimum number of latent variables was chosen by minimizing both the root mean squared error (RMSE) and the number of factors or latent vectors.

Two PLS regression approaches were performed to estimate soil parameters: FS-PLS and iterative stepwise elimination regression (ISE-PLS) [36]. The FS-PLS is a standard PLS model using FDR datasets. ISE-PLS is a PLS model using a waveband elimination algorithm to remove noisy variables and to select those able to improve predictive performance.

The prediction accuracies were evaluated using the coefficient of determination (R^2), the root mean squared error of cross-validation (RMSECV), and the residual predictive deviation (RPD). The RPD is the ratio of standard deviation (SD) of the measured data to the standard error of prediction [42]. The model with the larger R^2 and RPD, and the smaller RMSE was considered the best model to predict soil Pox. It is generally accepted that an RPD value greater than 3 indicates an excellent predictive model for agricultural applications, and values between 2 and 3 indicate good predictive ability;

values between 1.5 and 2 indicate an acceptable model requiring some improvement, and those below 1.5 indicate a poor predictive model [36,40].

To assess the predictive ability and reliability of the PLS models, a modified bootstrap procedure was performed [25]; the data was divided randomly into training (70%) and test (30%) data sets with a replacement for $N = 10,000$ times. In each process, a PLS model was developed using the training data set. Here, FS-PLS and ISE-PLS were developed using selected wavebands, and then the models were used to predict Pox in the test data set. The robustness of the prediction models was evaluated by the mean (\pm SD) values of R^2 and the root mean squared error of prediction (RMSEP) from 10,000 runs in the test data sets.

All data handling and statistical analysis were performed using MATLAB software (Version 9.3; The MathWorks, Sherborn, MA, USA) and R software version 3.1.3 [43] (R Core Team 2015).

3. Results and Discussion

3.1. Soil Characteristics by Chemical Analysis

The descriptive statistics for soil Pox as measured by chemical analysis for all data and by the system are summarized in Table 3. The coefficient of variation (CV) for Pox when all data were combined data indicated large Pox variability (148.57%) with a heterogeneous distribution. The Pox content averaged $87.66 \text{ mg}\cdot\text{kg}^{-1}$ across all data, ranging from 21.89 to $856.84 \text{ mg}\cdot\text{kg}^{-1}$. As illustrated in Figure 2, the Pox level varied markedly within the cultivated rice systems, much more so than in the natural systems. Indeed, the highest Pox value recorded from natural systems was $57.93 \text{ mg}\cdot\text{kg}^{-1}$ with a CV of 22.23%, in contrast to that of the cultivated system, which was $856.84 \text{ mg}\cdot\text{kg}^{-1}$ with a CV of 133.56%. The third quartile cutoff, containing 75% of the data was $38.73 \text{ mg}\cdot\text{kg}^{-1}$ for all the natural systems and $106.62 \text{ mg}\cdot\text{kg}^{-1}$ for all the cultivated systems. The variation in P level seen in the substantial dispersion of the cultivated system data probably results from the different levels of fertilizer application to farmers' plots. Based on the study by Dardenne et al. [44], such wide variation ($\text{CV} > 50\%$) is recommended to achieve good NIRS calibration accuracy, indicating that our soil data were suitable for developing the spectroscopy model.

Table 3. Summary statistics for soil oxalate-extractable P content (mg P kg^{-1} soil) obtained by chemical analysis by system.

System	n	Min	Max	Mean	SD	CV (%)
All systems	216	21.89	856.84	87.66	130.23	148.57
Cultivated system	142	22.25	856.84	115.07	153.69	133.56
Natural system	74	21.89	57.93	35.05	7.79	22.23

n, number of samples; SD, standard deviation; CV, coefficient of variation ($\text{SD}/\text{mean} \times 100\%$).

The difference in soil characteristics, including soil texture and the level of Pox in each system, can explain the high accuracy of prediction for each specific model. The correlation matrix between the Pox, SOC, and their related soil parameters are shown in Table 4. In the ensemble of the data, no significant correlation was observed for Pox and SOC. Among the significant relationships observed, soil parameters which could affect the Pox were SOC, sand, clay, and Fe contents. In the natural system, the Pox was positively correlated with SOC, clay, and Feox while negative relations were observed between Pox and Feox with sand content suggesting a direct effect of soil organic matter and texture on Pox contents. In the cultivated system, Pox is more affected by Feox than the SOC. Principal component (Figure 3) and texture triangle (Figure 4) analyses showed the contrasting properties of cultivated and natural soils. Natural system soils with a coarse texture were marked by low Pox and Feox content compared to the cultivated soils. Cultivated soils with a clayey loam texture had high Pox and lower SOC compared to natural soils, probably due to the soil management techniques applied.

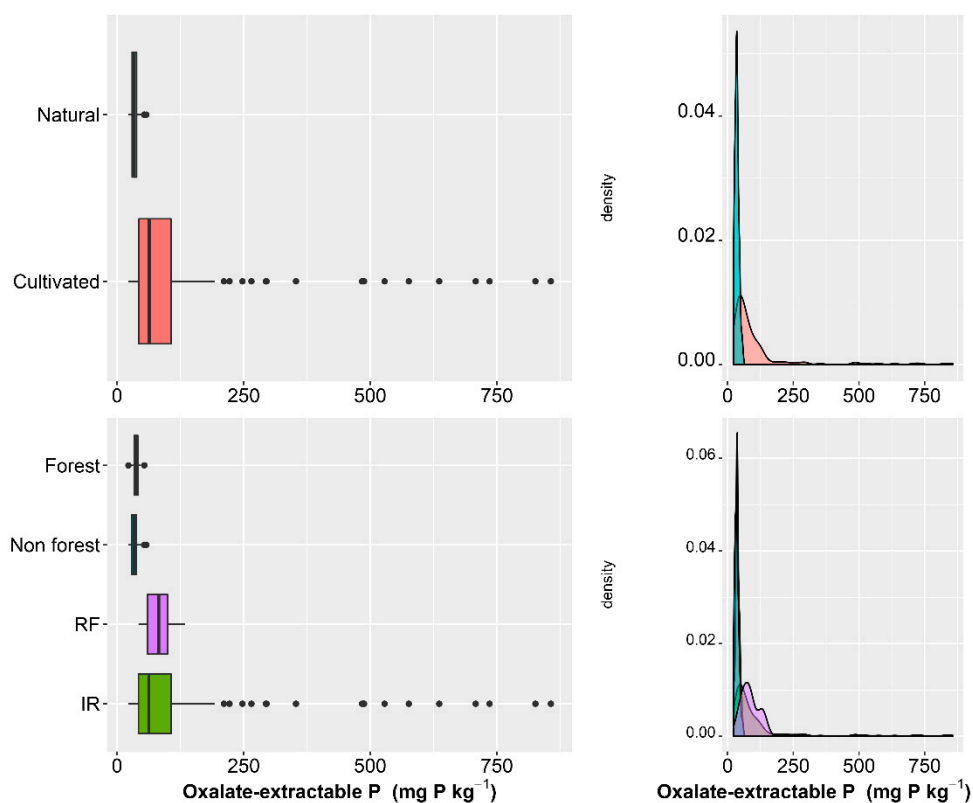


Figure 2. Distribution (left side) and kernel density plots (right side) for oxalate-extractable P content by the land-use system. RF: Upland rice field, IR: Lowland rice field.

Table 4. Pearson correlation coefficients between soil properties for each system.

	Pox	SOC	Sand	Silt	Clay	Feox
All Systems						
Pox	1.00	0.10	−0.20	0.23	−0.00	0.55
SOC		1.00	0.15	−0.32	0.24	−0.06
Sand			1.00	−0.82	−0.49	−0.41
Silt				1.00	−0.09	0.35
Clay					1.00	0.18
Feox						1.00
Natural						
Pox	1.00	0.61	−0.29	−0.06	0.37	0.45
SOC		1.00	−0.44	−0.00	0.53	0.32
Sand			1.00	−0.67	−0.96	−0.48
Silt				1.00	0.45	0.22
Clay					1.00	0.50
Feox						1.00
Cultivated						
Pox	1.00	0.30	−0.03	0.05	−0.03	0.51
SOC		1.00	−0.05	−0.05	0.15	0.33
Sand			1.00	−0.79	−0.32	0.06
Silt				1.00	−0.33	−0.15
Clay					1.00	0.14
Feox						1.00

Values in bold are significant at $P < 0.05$.

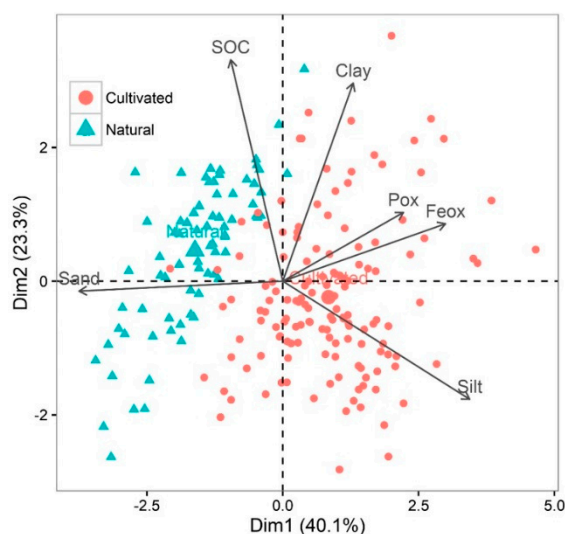


Figure 3. Principal component analysis incorporating measured variables by soil system. Pox, oxalate-extractable P; Feox, oxalate Fe; SOC, soil organic carbon.

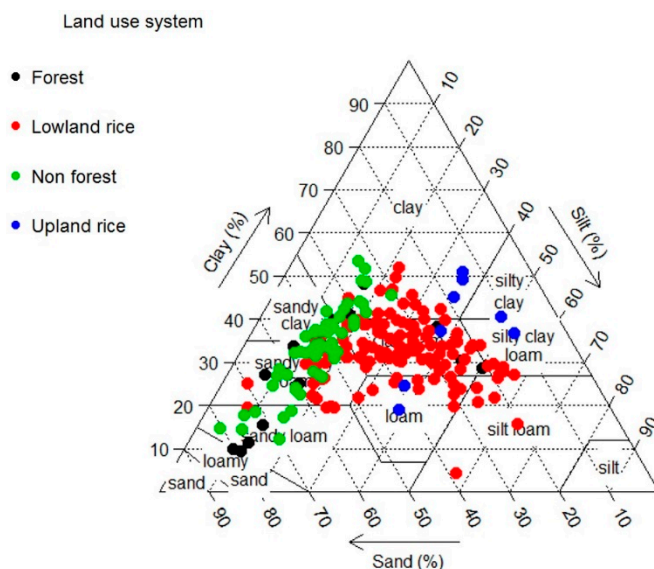


Figure 4. Soil texture triangle for different land-use systems based on the USDA classification.

The mineral properties of soil are strongly related to their NIR-spectra absorption patterns [45]. Mouazen et al. [46] confirmed that soil texture affected the reflectance of the soil surface during NIR spectral measurement. Light scattering increased with increasing sand content due to a large amount of quartz in the sand fraction, which increases the intensity of spectral reflectance [47]. The spectral absorption related to some soil components (O-H and metal O-H, O-H in water) increased with increasing clay content [48].

Soil preparation, specifically tillage, could break up soil particles and aggregates and thereby accelerate the mineralization of soil organic matter, resulting in lower SOC compared to that of natural systems [49,50]. The level of Pox in the cultivated systems is due to fertilizer input and high mineralization rates, which released the soil nutrients (including phosphorus).

3.2. Model Prediction Accuracy for Oxalate-Extractable P under Different Land-Use Systems

Predictions of Pox content were made using standard FS-PLS and ISE-PLS regressions for all combined systems and for each system individually. The PLS regression model predictions of Pox levels

are shown in Table 5 and Figure 5. ISE-PLS regression always improved Pox prediction compared to FS-PLS regardless of the land-use system.

Table 5. Comparison of different soil oxalate-extractable P prediction models for all land-use systems.

Processing	Systems	n	NLV	R ²	RMSECV	RPD
FS-PLS	All systems	216	13	0.48	96.58	1.34
	Cultivated	142	15	0.70	83.72	1.82
	Natural	74	2	0.18	7.10	1.08
ISE-PLS	All systems	216	13	0.70	71.87	1.81
	Cultivated	142	15	0.90	48.57	3.15
	Natural	74	14	0.90	2.39	3.22

n, number of samples, NLV, number of latent variables; FS-PLS, full-spectrum partial least square regression; ISE-PLS, iterative stepwise elimination-partial least square regression.

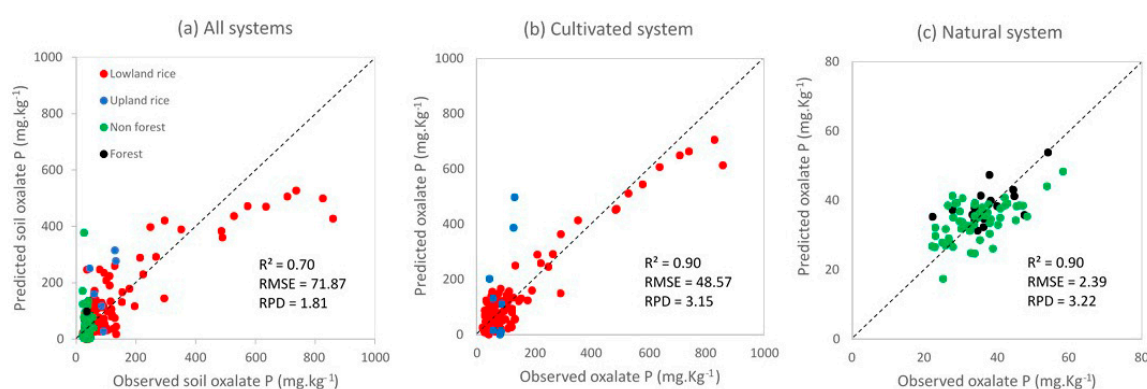


Figure 5. Relationships between observed and predicted values of soil oxalate-extractable P contents using ISE-PLS (iterative stepwise elimination-partial least square) regression with first derivative reflectance data for (a) natural systems ($n = 74$), (b) cultivated systems ($n = 142$), and (c) all systems combined ($n = 216$). RMSE, root mean square error; RPD, residual predictive deviation.

ISE-PLS regression performed well; we attributed this to the importance of waveband selection for Pox prediction. The percentage of wavebands (NW%) used in the model was the ratio of the number of selected wavebands (NW) to the total wavebands for a full-spectrum ($NW\% = NW / 2001 \text{ bands} \times 100$). The NW% results were 20.6% and 7.5% for cultivated and natural systems, respectively. In other words, fewer than 21% of available wavelengths contributed to the prediction of Pox for the cultivated system, with over 79% neither contributing to nor disturbing the predictions [51]. Selecting wavebands related to soil Pox and eliminating unusable wavebands improved the predictive ability of ISE-PLS for Pox compared to FS-PLS. This finding was in agreement with previous studies, in which fewer than 20% of wavelengths contained information relevant to the prediction of soil properties [25,36]. ISE-PLS produced excellent predictions of Pox in natural and cultivated systems, with RPD values greater than three and an R^2 of 0.90 (Table 5). Although the performance of model prediction is better for the cultivated system than the natural system, this prediction model accuracy seems to be associated with the large distribution of Pox values, which were characterized here by some samples with high Pox value. A high variation of the data set could affect the accuracy of NIRS calibration and predictive performance [52]. The performance of ISE-PLS models was better for individual land-use systems than for the combined data ($R^2 = 0.70$, $RMSE = 71.9$, $RPD = 1.81$). Stevens et al. [48] highlighted the importance of building local, more accurate models that are specific to a given geographical entity or soil type, suggesting that this feature is a strength, rather than a weakness, of this model.

The results of a modified bootstrap procedure were reported in Table 6 and Figure 6. Table 6 gives the mean values of R^2 and RMSEP between FS-PLS and ISE-PLS models for each system in the test data set (30%). Figure 6 illustrated the distribution of R^2 values in the test data set for each

system. The accuracy of the model prediction with validation data showed that the ISE-PLS models predicted soil oxalate-extractable P more accurately than FS-PLS in terms of R^2 and RMSEP for all systems. The ISE-PLS resolved 70% to 88% of the variation in Pox whereas total variance explained with FS-PLS was from 14% to 50%. The best mean R^2 and the lowest RMSEP values were obtained from the natural system. The predictive ability and reliability of the ISE-PLS models were confirmed by this modified bootstrap procedure.

Table 6. Mean and standard deviation (SD) values of R^2 and RMSEP from $N = 10,000$ evaluations with FS-PLS and ISE-PLS in test data sets (30%).

Processing	Systems	n	R^2	RMSEP
FS-PLS	All systems	64	0.502 ± 0.124	89.01 ± 9.21
	Cultivated	42	0.678 ± 0.079	79.13 ± 8.30
	Natural	22	0.141 ± 0.096	7.15 ± 1.62
ISE-PLS	All systems	64	0.703 ± 0.115	60.48 ± 5.94
	Cultivated	42	0.883 ± 0.038	57.42 ± 5.57
	Natural	22	0.822 ± 0.051	3.26 ± 0.59

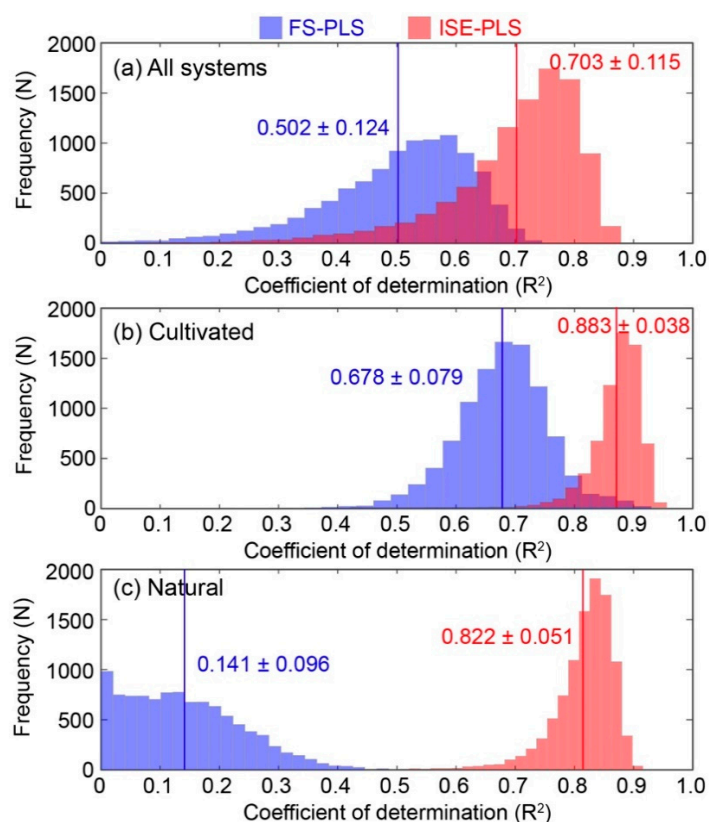


Figure 6. Comparisons of the frequency distributions of R^2 values in the test data for each system: (a) all systems; (b) cultivated; (c) natural using FS-PLS and ISE-PLS models, with mean (red and blue line) \pm standard deviation (SD) values.

3.3. Properties of the Prediction-Relevant Wavebands

Figure 7 shows the selected wavebands used for the PLS regression modeling and prediction of Pox resulting from the preprocessing of the spectra using first derivative data. All samples showed similar spectral absorption features, characteristic of mineral and organic spectra as reported by several authors [15,53]. The most influential wavelengths in terms of the Pox prediction model were recorded in the visible light range (around 500 nm) and in the NIR range (at 1400 nm and from 2000 nm). The spectral absorption peaks in the Vis-NIRS region are related to iron oxides, clay minerals, and some

functional groups of soil organic matter (SOM) [37]. In our study, the selected wavebands in the visible region common to both natural and cultivated systems (409, 430, 431, 443, 444, 591, and 592 nm) were associated with Fe-containing minerals (hematite, goethite) and dark-colored organic matter [54,55]. Residual minerals like hematite and goethite have an effect on the organic matter sorption of soil nutrients such as phosphorus [56].

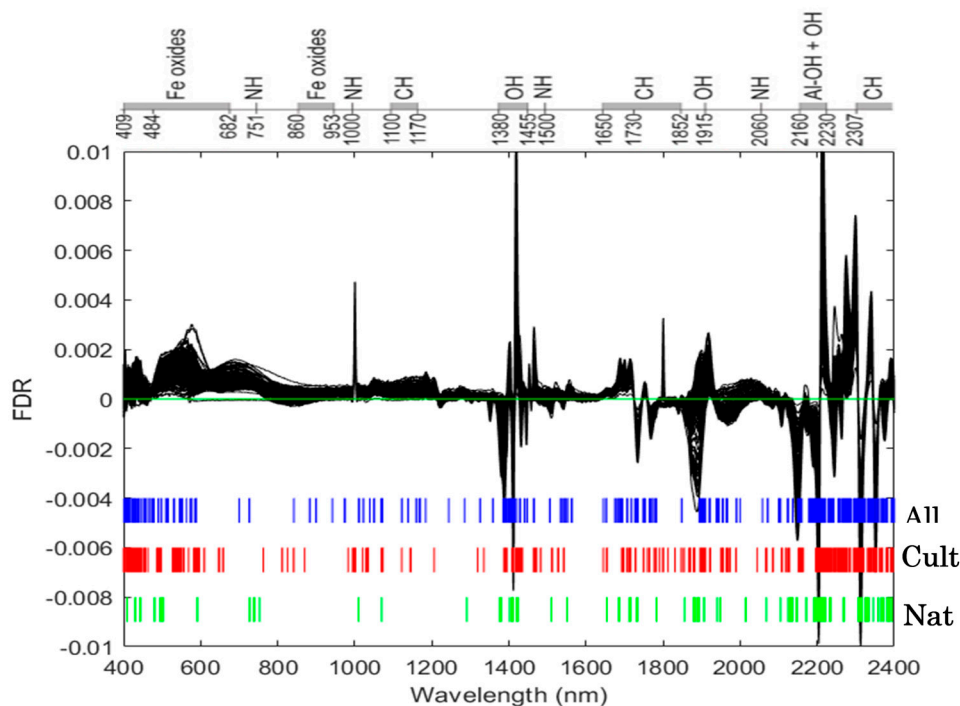


Figure 7. Wavebands used in the ISE-PLS analysis for all combined data (blue bars, All), cultivated systems (red bars, Cult), and natural systems (green bars, Nat) using the first derivative reflectance (FDR) dataset to estimate oxalate-extractable phosphorus. Specific absorption wavebands for the different bonds present in soil are specified on the top x-axis (modified from Kawamura et al. [25]).

The bands in the NIR range usually attributed to O-H chemical bonds at 1400 nm, to C-H stretch at 1700 nm, to water (H-O-H) at 1900 nm, and metal-OH bending and O-H stretching modes near 2000 nm, 2300 nm, and 2400 nm are often associated with clay mineral types (Table 7) [15,53,57]. The spectral bands at 1906–1907 nm, 2200–2235 nm, and 306–2400 nm, related to minerals and water [15,58], and that at 2270 nm, corresponding to gibbsite (an Al oxide mineral) [56,59], contribute to Pox prediction. The detection of the mineral and organic compounds in soils allow soil spectroscopy to predict Pox because of the potential relation between phosphorus and carbon content [22].

The number of selected wavelengths for Pox prediction is higher for cultivated systems than natural systems (Figure 7). The specific selected visible wavelengths for cultivated areas were 527–590 nm, associated with hematite and organic matter; and 763–870 nm, related to amine N-H, aromatic C-H, Fe³⁺, and ferric oxide [58–60]. The regions related to amine N-H at 1000 nm; aromatic C-H at 1100 nm; alkyl C-H at 1170 nm; O-H in water, CH₂, lignin, and cellulose at 1464–1483 nm [61]; and Al-OH and kaolin at 2160–2164 nm [62] contributed to Pox prediction in the NIR regions. In contrast, the specific selected wavelengths for natural systems were 738–740 nm and 753 nm (amine N-H); and 1291 nm, related to lignin, starch, and protein [59–61]. These specific bands for each system demonstrated the variation in SOM and absorbents contributing to Pox prediction, which may explain the low accuracy of prediction when all data were combined.

Table 7. Selected visible and near infra-red (NIR) wavelengths related to soil components and functional groups as reported in the literature, and common and specific selected wavelengths observed in our study.

Spectra Regions (nm)	Common Selected Wavelength (nm)	Specific Wavelength (nm) for Cultivated	Specific Wavelength (nm) for Natural	Functional Groups	References
400–700	409, 430, 431, 443, 444, 591, 592	527–590		Associated to mineral with Fe (hematite, goethite) SOM: chromophores and darkness of OM	[54,55]
550		550–557		Chromophore FeOOH in goethite	[55]
Near Infra-Red (NIR)					
751, 825		763, 826	738–740, 753	Amine C-H, aromatic C-H	[59]
860		870		Ferric oxide, Fe ³⁺	[58]
1000		1000		Amine N-H	[59]
1100		1122–1144		Aromatic C-H	[59]
1170		1170		Alkyl asymmetric-symmetric doublet (C-H)	[59,60]
1260			1291	Lignin, starch, protein, OH in water, CH ₂ , cellulose, lignin, starch, pectin	[61]
1465, 1470		1464–1483		Al-OH, Kaolin	[61]
2160		2160–2164		Metal-OH, O-H	[62]
2200–2300	2200–2270			C-H stretch fundamentals	[15]
2300, 2350		2302–2306/2350–2355		Carbonates	[59,60]
2335	2330–2334				[58]
850, 1200, 1400, 1900		1950–1956		H ₂ O	[15,58]
2200, 2300	2200–2270			Al-OH, O-H	[59]
1900	1906–1907			H-O-H	[59]
Visible-NIR					
450, 900	453–457			Fe+3	[54]

3.4. Factors Influencing the Prediction Model Accuracy for Oxalate-Extractable P

According to our results, the main soil components which contributed to the prediction of Pox were organic matter and iron oxides, in both natural and cultivated systems. This is consistent with the study of Sørensen and Dalsgaard, which suggested that indirect relationships between soil P and organic components would be useful in soil P prediction using spectrographic methods [63], and that of Ludwig et al., in which a useful calibration of soil P, measured using the Olsen method, was found to positively correlate with SOC [22]. The present study showed that Pox is significantly correlated with SOC in natural and cultivated systems with coefficients of correlation (r) of 0.61 ($P < 0.001$) and 0.30 ($P < 0.001$), respectively, but not when all data are combined ($r = 0.10$, $P = 0.15$). Abdi et al. confirmed that successful prediction of soil total P is related to its significant correlation with soil carbon [42]. Soil P is obtainable by NIRS through covariation with other soil properties but this relation may vary between datasets [16], possibly explaining the lack of correlation between soil carbon and Pox for all combined data. The high correlation between Pox and SOC in natural systems may have resulted from the accumulation of P in the surface layer through litter input, while in the cultivated system P is lost with the harvested crops.

Phosphorus in soil was mainly fixed and in solid phase with Fe, Al in acidic soil, and Ca in alkaline soil. These elements are the main adsorbing agents for phosphate [64]. Khalid et al. [65] found that higher P availability under flooded soil was related to ammonium oxalate Fe. In our study, Pox and oxalate Fe (Feox) were significantly and positively correlated for cultivated, natural and all combined systems with correlation coefficients of 0.51 ($P < 0.001$), 0.45 ($P < 0.001$), and 0.55 ($P < 0.001$), respectively. In addition to the selected wavebands for Pox prediction in the Vis-NIRS regions associated with iron oxides, this result is in agreement with previous studies confirming the primary role of Fe in P sorption [7,25]. This highlights the importance of Fe to Pox prediction model development.

The high correlation between Pox, SOC, and Feox observed mainly under the natural system can be associated with the related properties of this system such as fallow without fertilization, justifying here the high accuracy of the model. As the high performance of model prediction in the cultivated system could be related to some samples with high Pox content ($n = 15$), a low prediction accuracy was obtained with selected samples excluding these high Pox samples (data not shown) suggesting that in the cultivated system under varying fertilization and other management practices may interfere and disturb the correlation of Pox with organic matter and iron oxides. The correlations of Pox with SOC and Feox are very weak for the selected samples (without the high Pox samples), $r = 0.22$, $P < 0.05$ and $r = 0.03$, $P = 0.69$, respectively. Application of ISE-PLS model in a large sample with a large geographical cover can help to understand the main drivers of Pox in the cultivated and the natural system in order to build more robust models.

In this study, the “pseudo-independent” approach of using a randomly selected sample (30%) for a validation in the modified bootstrap procedure or LOOCV, which provide more accurate PLS models in Pox prediction, presents a limitation. A previous study on SOC prediction using the first derivative Vis-NIRS PLS approach reports a stable model accuracy from a “pseudo-independent” validation (random selection of non-independent test samples), but the prediction models failed when applied for each site through site-hold validation (using samples from one site for validation and the samples from the remaining sites for model calibration) [66]. We attempted to perform the FS-PLS based on the site-hold cross-validation by considering the seven studied sites and found very poor results (data not shown). This may be due to the mixture of sites and land-use systems using a small number of samples. This suggests building models using a large geographical cover and relatively dispersed sample sets for a regional application.

4. Conclusions

Soil P is an important limiting nutrient for plant growth. An accurate assessment of available P is essential for effective fertilizer management in agriculture and sustainable management of ecosystems. Vis-NIRS is a simple and nondestructive method that can be used to predict several soil

properties. This study demonstrates that Vis-NIRS models, in combination with ISE-PLS regression, can successfully predict soil oxalate-extractable phosphorus (Pox) in soil samples from natural and cultivated systems in Madagascar. Together, these methods were able to estimate soil Pox in both systems with high accuracy ($R^2 = 0.90$, RPD > 3) using fewer than 21% of wavelengths in the Vis-NIRS region. ISE-PLS regression outperformed FS-PLS regression. However, model accuracy for cultivated systems was affected by some samples with high Pox value. The effective wavebands for the two land-use systems were associated with Fe and Al oxides, and organic components. The accuracy of Pox prediction was related to its significant correlation with soil organic carbon and iron content. The use of “pseudo-independent” validation in the current study can also overestimate the prediction accuracy when applied at site scale suggesting the use of larger and dispersed geographical cover sample sets to build a robust model in the future. The Vis-NIRS approach has potential as a tool for rapid soil P evaluation and may be useful for soil management. Further investigations using large numbers of soil samples for external validation of the Vis-NIRS approach are required to enable application at regional and national scales.

Author Contributions: Conceptualization, K.K., Y.T., H.R. (Hobimiarantsoa Rakotonindrina), and A.A.; methodology, H.R. (Hobimiarantsoa Rakotonindrina), K.K., and A.A.; software, K.K.; validation, A.A., K.K. and Y.T.; formal analysis, A.A.; investigation, H.R. (Hobimiarantsoa Rakotonindrina) and B.H.A.; resources, A.A., K.K., and H.R. (Herintsitohaina Razakamanarivo); data curation, H.R. (Hobimiarantsoa Rakotonindrina); writing—original draft preparation, H.R. (Hobimiarantsoa Rakotonindrina) and A.A.; writing—review and editing, K.K., Y.T., T.N., H.R. (Herintsitohaina Razakamanarivo), and A.A.; visualization, K.K.; supervision, A.A. and H.R. (Herintsitohaina Razakamanarivo); project administration, Y.T.; funding acquisition, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA) (Grant No. JPMJSA1608).

Acknowledgments: This work is financed by FY VARY Project. We acknowledge the project team and all staff of Laboratoire des Radioisotopes (LRI) including DFID-ESPA p4GES program for help with fieldwork and lab analyses.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Andriamananjara, A.; Rakotoson, T.; Razafimbelo, T.; Rabeharisoa, L.; Razafimanantsoa, M.P.; Masse, D. Farmyard manure improves phosphorus use efficiency in weathered P deficient soil. *Nutr. Cycl. Agroecosyst.* **2019**, *115*, 407–427. [[CrossRef](#)]
2. Turner, B.L.; Engelbrecht, B. Soil organic phosphorus in lowland tropical rain forests. *Biogeochemistry* **2011**, *103*, 297–315. [[CrossRef](#)]
3. Liu, X.; Meng, W.; Liang, G.; Li, K.; Xu, W.; Huang, L.; Yan, J. Available phosphorus in forest soil increases with soil nitrogen but not total phosphorus: Evidence from subtropical forests and a pot experiment. *PLoS ONE* **2014**, *9*, e88070. [[CrossRef](#)]
4. Rabeharisoa, L.; Razanakoto, O.R.; Razafimanantsoa, M.P.; Rakotoson, T.; Amery, F.; Smolders, E. Larger bioavailability of soil phosphorus for irrigated rice compared with rainfed rice in Madagascar: Results from a soil and plant survey. *Soil Use Manag.* **2012**, *28*, 448–456. [[CrossRef](#)]
5. Bollyn, J.; Castelein, L.; Smolder, E. Fate and bioavailability of phosphorus loaded to iron oxyhydroxide nanoparticles added to weathered soils. *Plant Soil* **2019**, *438*, 297. [[CrossRef](#)]
6. Nawara, S.; Van Dael, T.; Merckx, R.; Amery, F.; Elsen, A.; Odeurs, W.; Vandendriessche, H.; Mcgrath, S.; Roisin, C.; Jouany, C.; et al. A comparison of soil tests for available phosphorus in long-term field experiments in Europe. *Eur. J. Soil Sci.* **2017**, *68*, 873–885. [[CrossRef](#)]
7. Nishigaki, T.; Tsujimoto, Y.; Rinasoa, S.; Rakotoson, T.; Andriamananjara, A.; Razafimbelo, T. Phosphorus uptake of rice plants is affected by phosphorus forms and physicochemical properties of tropical weathered soils. *Plant Soil* **2019**, *435*, 27–38. [[CrossRef](#)]
8. Guo, F.; Yost, R.S. Quantifying the available soil phosphorus pool with the acid ammonium oxalate method. *Soil Sci. Soc. Am. J.* **1999**, *63*, 651–656. [[CrossRef](#)]

9. Shahandeh, H.; Hossner, L.R.; Turner, F.T. Phosphorus Relationships in Flooded Rice Soils with Low Extractable Phosphorus. *Soil Sci. Soc. Am. J.* **1994**, *58*, 1184–1189. [[CrossRef](#)]
10. Schwertmann, U. The differentiation of iron oxides in soils by extraction with ammonium oxalate solution. *Z. Pflanz. Bodenkd.* **1964**, *105*, 194–202. [[CrossRef](#)]
11. Schwertmann, U. Use of oxalate for Fe extraction from soils. *Can. J. Soil Sci.* **1973**, *53*, 244–246. [[CrossRef](#)]
12. Narteh, L.T.; Sahrawat, K.L. Oxalate and EDTA extractable soil phosphorus and iron in relation to P availability in lowland rice soils of West Africa. *Ghana J. Agric. Sci.* **1999**, *32*, 189–198. [[CrossRef](#)]
13. Six, L.; Smolders, E.; Merckx, R. The performance of DGT versus conventional soil phosphorus tests in tropical soils—Maize and rice responses to P application. *Plant Soil* **2013**, *366*, 49–66. [[CrossRef](#)]
14. Neyroud, J.A.; Lischer, P. Do different methods used to estimate soil phosphorus availability across Europe give comparable results? *J. Soil Sci. Plant Nutr.* **2003**, *166*, 422–431. [[CrossRef](#)]
15. Viscarra Rossel, R.V.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [[CrossRef](#)]
16. Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind, J. Visible and near infrared spectroscopy in soil science. *Adv. Agron.* **2010**, *107*, 163–215.
17. Nawar, S.; Buddenbaum, H.; Hill, J.; Kozak, J.; Mouazen, A.M. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil Tillage Res.* **2016**, *155*, 510–522. [[CrossRef](#)]
18. Recena, R.; Fernández-Cabanás, V.M.; Delgado, A. Soil fertility assessment by Vis-NIR spectroscopy: Predicting soil functioning rather than availability indices. *Geoderma* **2019**, *337*, 368–374. [[CrossRef](#)]
19. McCarty, G.W.; Reeves, J.B., III. Comparison of near infrared and mid infrared diffuse reflectance spectroscopy for field-scale measurement of soil fertility parameters. *Soil Sci.* **2006**, *171*, 94–102. [[CrossRef](#)]
20. Bellon-Maurel, V.; McBratney, A. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils—Critical review and research perspectives. *Soil Biol. Biochem.* **2011**, *43*, 1398–1410. [[CrossRef](#)]
21. Chang, C.-W.; Laird, D.A.; Mausbach, M.J.; Hurburgh, C.R. Near-infrared reflectance spectroscopy—Principal components regression analyses of soil properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [[CrossRef](#)]
22. Ludwig, B.; Khanna, P.K.; Bauhus, J.; Hopmans, P. Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. *For. Ecol. Manag.* **2002**, *171*, 121–132. [[CrossRef](#)]
23. Zornoza, R.; Guerrero, C.; Mataix-Solera, J.; Scow, K.M.; Arcenegui, V.; Mataix-Beneyto, J. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biol. Biochem.* **2008**, *40*, 1923–1930. [[CrossRef](#)]
24. Kruse, J.; Abraham, M.; Amelung, W.; Baum, C.; Bol, R.; Kühn, O.; Lewandowski, H.; Niederberger, J.; Oelmann, Y.; Rüger, C.; et al. Innovative methods in soil phosphorus research: A review. *J. Plant Nutr. Soil Sci.* **2015**, *178*, 43–88. [[CrossRef](#)] [[PubMed](#)]
25. Kawamura, K.; Tsujimoto, Y.; Nishigaki, T.; Andriamananjara, A.; Rabenarivo, M.; Asai, H.; Rakotoson, T.; Razafimbelo, T. Laboratory visible and near-infrared spectroscopy with genetic algorithm-based partial least squares regression for assessing the soil phosphorus content of upland and lowland rice fields in Madagascar. *Remote Sens.* **2019**, *11*, 506. [[CrossRef](#)]
26. Badjeck, B.; Ibrahima, N.C.; Slaviero, F. *Evaluation de la Sécurité Alimentaire à Madagascar*; FAO: Rome, Italy, 2013.
27. Penot, E.; Domas, R.; Paulin, H.; Durand, C. *Rôle et Place du Riz Pluvial Dans les Exploitations Agricoles à Madagascar. Le Cas du Lac Alaotra et du Vakinankaratra*; Conference paper; Académie d’Agriculture: Antananarivo, Madagascar, 2011.
28. Lal, R. Soil carbon sequestration to mitigate climate change. *Geoderma* **2004**, *123*, 1–22. [[CrossRef](#)]
29. Andriamaniraka, H. Le Phosphore et la Fertilisation Phosphatée Dans les Sols Ferrallitiques à Madagascar: Amélioration de la fertilité des sols. In *Mémoire D’habilitation à Diriger des Recherches*; Université d’Antananarivo: Antananarivo, Madagascar, 2016.
30. Nishigaki, T.; Ikazaki, K.; Tsujimoto, Y.; Andriamananjara, A.; Rakotoson, T.; Razafimbelo, T. Soil survey of the east coast and the central highlands indicates need to update Madagascar soil map. *Soil Sci. Plant Nutr.* **2020**. Accepted. [[CrossRef](#)]

31. Cornet, A. *Essai de Cartographie Bioclimatique à Madagascar*; Notice Explicative No. 55; ORSTOM: Paris, France, 1974.
32. Schatz, G.E. Endemism in the Malagasy tree flora. In *Diversity and Endemism in Madagascar*; Lourenço, W.R., Goodman, S.M., Eds.; Mémoires de la Société de Biogéographie: Paris, France, 2000; pp. 1–9.
33. Andriamananjara, A.; Hewson, J.; Razakamanarivo, H.; Andrisoa, R.H.; Ranaivoson, N.; Ramboatiana, N.; Razafindrakoto, M.; Ramifehiarivo, N.; Razafimanantsoa, M.P.; Rabeharisoa, L.; et al. Land cover impacts on aboveground and soil carbon stocks in Malagasy rainforest. *Agric. Ecosyst. Environ.* **2016**, *233*, 1–15. [[CrossRef](#)]
34. Andriamananjara, A.; Ranaivoson, N.; Razafimbelo, T.; Hewson, J.; Ramifehiarivo, N.; Rasolohery, A.; Andrisoa, R.H.; Razafindrakoto, M.A.; Razafimanantsoa, M.-P.; Rabetokotany, N.; et al. Towards a better understanding of soil organic carbon variation in Madagascar. *Eur. J. Soil Sci.* **2017**, *68*, 6. [[CrossRef](#)]
35. Walkley, A.; Black, I.A. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Sci.* **1934**, *37*, 29–38. [[CrossRef](#)]
36. Kawamura, K.; Tsujimoto, Y.; Rabenarivo, M.; Asai, H.; Andriamananjara, A.; Rakotoson, T. Vis-NIR spectroscopy and PLS regression with waveband selection for estimating the total C and N of paddy soils in Madagascar. *Remote Sens.* **2017**, *9*, 1081. [[CrossRef](#)]
37. Pätzold, S.; Leenen, M.; Frizen, P.; Heggemann, T.; Wagner, P.; Rodionov, A. Predicting plant available phosphorus using infrared spectroscopy with consideration for future mobile sensing applications in precision farming. *Prec. Agric.* **2019**, 1–25. [[CrossRef](#)]
38. Xu, D.; Ma, W.; Chen, S.; Jiang, Q.; He, K.; Shi, Z. Assessment of important soil properties related to Chinese Soil Taxonomy based on vis–NIR reflectance spectroscopy. *Comput. Electron. Agric.* **2018**, *144*, 1–8. [[CrossRef](#)]
39. Savitzky, A.; Golay, E.J.M. Smoothing and difference of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
40. Summers, D.; Lewis, M.; Ostendorf, B.; Chittleborough, D. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecol. Indic.* **2011**, *11*, 123–131. [[CrossRef](#)]
41. Viscarra Rossel, R.A.; Fouad, Y.; Walter, C. Using a digital camera to measure soil organic carbon and iron contents. *Biosyst. Eng.* **2008**, *100*, 149–159. [[CrossRef](#)]
42. Abdi, D.; Tremblay, G.F.; Ziadi, N.; Bélanger, G.; Parent, L.É. Predicting soil phosphorus-related properties using reflectance spectroscopy. *Soil Sci. Soc. Am. J.* **2012**, *76*, 2318–2326. [[CrossRef](#)]
43. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015.
44. Dardenne, P.; Sinnaeve, G.; Baeten, V. Multivariate calibration and chemometrics for near infrared spectroscopy: Which method? *J. Near Infrared Spectrosc.* **2000**, *8*, 229–237. [[CrossRef](#)]
45. Mohamed, E.S.; Saleh, A.M.; Belal, A.B.; Abd-Allah, G. Application of near-infrared reflectance for quantitative assessment of soil properties. *Egypt. J. Remote Sens. Space Sci.* **2018**, *21*, 1–14. [[CrossRef](#)]
46. Mouazen, A.M.; Karoui, R.; De Baerdemaeker, J.; Ramon, H. Classification of soil texture classes by using soil visual near infrared spectroscopy and factorial discriminant analysis techniques. *J. Near Infrared Spectrosc.* **2005**, *13*, 231–240. [[CrossRef](#)]
47. Conforti, M.; Matteucci, G.; Buttafuoco, G. Using laboratory Vis-NIR spectroscopy for monitoring some forest soil properties. *J. Soils Sediments* **2018**, *18*, 1009–1019. [[CrossRef](#)]
48. Stevens, A.; Nocita, M.; Toth, G.; Montanarella, L.; van Wesemael, B. Prediction of soil organic carbon at the European scale by visible and near infrared reflectance spectroscopy. *PLoS ONE* **2013**, *8*, e66409. [[CrossRef](#)] [[PubMed](#)]
49. Segda, Z.; Bonzi, M.; Gnankambary, Z.; Lompo, F.; Sedogo, M.P. Influence of soil fertility management on organic carbon mineralization in irrigated rice. *J. Agric. Crop Res.* **2014**, *2*, 32–43.
50. Balesdent, J.; Chenu, C.; Balabane, M. Relationship of soil organic matter dynamics to physical protection and tillage. *Soil Tillage Res.* **2000**, *53*, 215–230. [[CrossRef](#)]
51. Wang, Z.; Kawamura, K.; Sakuno, Y.; Fan, X.; Gong, Z.; Lim, J. Retrieval of chlorophyll-a and total suspended solids using iterative stepwise elimination partial least squares (ISE-PLS) regression based on field hyperspectral measurements in irrigation ponds in Higashi hiroshima, Japan. *Remote Sens.* **2017**, *9*, 264. [[CrossRef](#)]
52. Nduwamungu, C.; Ziadi, N.; Tremblay, G.F.; Parent, L.-É. Near- infrared reflectance spectroscopy prediction of soil properties: Effects of sample cups and preparation. *Soil Sci. Soc. Am. J.* **2009**, *73*, 1896–1903. [[CrossRef](#)]

53. Liu, Y.; Boss, E.; Chase, A.P.; Xi, H.; Zhang, X.; Röttgers, R.; Pan, Y.; Bracher, A. Spectral particulate absorption coefficients and their standard deviation derived from underway AC-S measurements during POLARSTERN cruise PS99.2. *PANGAEA* **2019**. [[CrossRef](#)]
54. Sherman, D.M.; Waite, D.T. Electronic spectra of Fe³⁺ oxides and oxide hydroxides in the near IR to near UV. *Am. Mineral.* **1985**, *70*, 1262–1269.
55. Mortimore, J.L.; Marshall, L.J.R.; Almond, M.J.; Hollins, P.; Matthews, W. Analysis of red and yellow ochre samples from Clearwell Caves and Catalhoyuk by vibrational spectroscopy and other techniques. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2004**, *60*, 1179–1188. [[CrossRef](#)]
56. Ramarason, V.H.; Becquer, T.; Sá, S.O.; Razafimahatratra, H.; Delarivière, J.L.; Blavet, D.; Vendrame, P.R.S.; Rabeharisoa, L.; Rakotondrazafy, A.F.M. Mineralogical analysis of ferralitic soils in Madagascar using NIR spectroscopy. *Catena* **2018**, *168*, 102–109. [[CrossRef](#)]
57. Blaschek, M.; Roudier, P.; Poggio, M.; Hedley, C.B. Prediction of soil available water-holding capacity from visible near-infrared reflectance spectra. *Sci. Rep.* **2019**, *9*, 12833. [[CrossRef](#)] [[PubMed](#)]
58. Hunt, G.R. Spectral signatures of particulate minerals in visible and near-infrared. *Trans. Am. Geophys. Union* **1977**, *58*, 553. [[CrossRef](#)]
59. Clark, R.N.; King, T.V.V.; Klejwa, M.; Swayze, G.A.; Vergo, N. High spectral resolution reflectance spectroscopy of minerals. *J. Geophys. Res.* **1990**, *95*, 12653–12680. [[CrossRef](#)]
60. Clark, R.N. Spectroscopy of rocks and minerals and principles of spectroscopy. In *Remote Sensing for the Earth Sciences: Manual of Remote Sensing*; Rencz, A.N., Ed.; John Wiley & Sons: Chichester, UK, 1999; pp. 3–58.
61. Ben-Dor, E.; Inbar, Y.; Chen, Y. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a control decomposition process. *Remote Sens. Environ.* **1997**, *61*, 1–15. [[CrossRef](#)]
62. Knadel, M.; Viscarra Rossel, R.A.; Deng, F.; Thomsen, A.; Greve, M.H. Visible–near infrared spectra as a proxy for topsoil texture and Glacial boundaries. *Soil Sci. Soc. Am. J.* **2013**, *77*, 568. [[CrossRef](#)]
63. Sørensen, L.K.; Dalsgaard, S. Determination of clay and other soil properties by near infrared spectroscopy. *Soil Sci. Soc. Am. J.* **2005**, *69*, 159–167. [[CrossRef](#)]
64. Jiang, X.; Bol, R.; Willbold, S.; Vereecken, H.; Klumpp, E. Speciation and distribution of P associated with Fe and Al oxides in aggregate-sized fraction of an arable soil. *Biogeosciences* **2015**, *12*, 6443–6452. [[CrossRef](#)]
65. Khalid, R.A.; Patrick, W.H.; Delaune, R.D. Phosphorus sorption characteristics of flooded soils. *Soil Sci. Soc. Am. Proc.* **1977**, *41*, 305–310. [[CrossRef](#)]
66. Brown, D.J.; Brickleyer, R.S.; Miller, P.R. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma* **2005**, *129*, 251–267. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Article

Using a One-Dimensional Convolutional Neural Network on Visible and Near-Infrared Spectroscopy to Improve Soil Phosphorus Prediction in Madagascar

Kensuke Kawamura ^{1,*}, Tomohiro Nishigaki ¹, Andry Andriamananjara ², Hobimiarantsoa Rakotonindrina ², Yasuhiro Tsujimoto ¹, Naoki Moritsuka ³, Michel Rabenarivo ² and Tantely Razafimbelo ²

¹ Japan International Research Center for Agricultural Sciences (JIRCAS), 1-1 Ohwashi, Tsukuba, Ibaraki 305-8686, Japan; nishigaki@affrc.go.jp (T.N.); tsjmt@affrc.go.jp (Y.T.)

² Laboratoire des Radio-Isotopes, Université d'Antananarivo, BP 3383, Route d'Andraisoro, Antananarivo 101, Madagascar; njaraandry1@gmail.com (A.A.); hobimiarantsoa@gmail.com (H.R.); miarabenarivo@yahoo.fr (M.R.); tantely.razafimbelo@gmail.com (T.R.)

³ Faculty of Agriculture and Marine Science, Kochi University, Nankoku, Kochi 783-8502, Japan; moritsuka@kochi-u.ac.jp

* Correspondence: kamuken@affrc.go.jp; Tel.: +81-29-868-6628



Citation: Kawamura, K.; Nishigaki, T.; Andriamananjara, A.; Rakotonindrina, H.; Tsujimoto, Y.; Moritsuka, N.; Rabenarivo, M.; Razafimbelo, T. Using a One-Dimensional Convolutional Neural Network on Visible and Near-Infrared Spectroscopy to Improve Soil Phosphorus Prediction in Madagascar. *Remote Sens.* **2021**, *13*, 1519. <https://doi.org/10.3390/rs13081519>

Academic Editor: Frédéric Cointault

Received: 5 March 2021

Accepted: 8 April 2021

Published: 15 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: As a proximal soil sensing technique, laboratory visible and near-infrared (Vis-NIR) spectroscopy is a promising tool for the quantitative estimation of soil properties. However, there remain challenges for predicting soil phosphorus (P) content and availability, which requires a reliable model applicable for different land-use systems to upscale. Recently, a one-dimensional convolutional neural network (1D-CNN) corresponding to the spectral information of soil was developed to considerably improve the accuracy of soil property predictions. The present study investigated the predictive ability of a 1D-CNN model to estimate soil available P (oxalate-extractable P; P_{ox}) content in soils by comparing it with partial least squares (PLS) and random forest (RF) regressions using soil samples ($n = 318$) collected from natural (forest and non-forest) and cultivated (upland and flooded rice fields) systems in Madagascar. Overall, the 1D-CNN model showed the best predictive accuracy ($R^2 = 0.878$) with a highly accurate prediction ability (ratio of performance to the interquartile range = 2.492). Compared to the PLS model, the RF and 1D-CNN models indicated 4.37% and 23.77% relative improvement in root mean squared error values, respectively. Based on a sensitivity analysis, the important wavebands for predicting soil P_{ox} were associated with iron (Fe) oxide, organic matter (OM), and water absorption, which were previously known wavelength regions for estimating P in soil. These results suggest that 1D-CNN corresponding spectral signatures can be expected to significantly improve the predictive ability for estimating soil available P (P_{ox}) from Vis-NIR spectral data. Rapid and accurate estimation of available P content in soils using our results can be expected to contribute to effective fertilizer management in agriculture and the sustainable management of ecosystems. However, the 1D-CNN model will require a large dataset to extend its applicability to other regions of Madagascar. Thus, further updates should be tested in future studies using larger datasets from a wide range of ecosystems in the tropics.

Keywords: deep learning; Madagascar; oxalate-extractable soil P; visible and near-infrared spectroscopy

1. Introduction

Phosphorus (P) deficiency is a major constraint for crop production in low-input agricultural systems in the tropics [1], and stems from the predominance of strongly weathered soils in which the availability of P is lowered by strong sorption to aluminum (Al) and iron (Fe) (hydr)oxides [2,3]. Even in natural ecosystems, limited soil available P can lead to a decline in the climax ecosystem by reducing biomass productivity [4]. Hence, rapid and quantitative information on soil fertility status is essential for improving

biomass production in agricultural and natural ecosystems and developing sustainable land management. Among many extraction methods for evaluating available P, the method using acid ammonium oxalate solution is known to be suited to tropical weathered soils because it can solubilize the active reductant-soluble P, which is the dominant P pool for P cycling in tropical ecosystems [5,6].

Among the proximal soil-sensing techniques, laboratory visible and near-infrared (Vis-NIR) spectroscopy has long been adopted as a rapid, cost-effective, and quantitative analytical method for predicting soil properties [7]. Efforts have been made to characterize the chemical, physical, and mineralogical composition of soil using Vis-NIR spectra (400–2500 nm) [8]. However, making reliable predictions at larger scales requires a large spectral and reference dataset collected from various soil conditions, and this requires the development of a spectral library [9,10]. With the increasing number of data sources, a standard approach based on a spectral library is required [11]. Currently, soil spectral libraries are available at the country [11,12], continental [13,14], and global [15] scales. These spectral libraries allow the development of calibration models for predicting soil properties.

For the calibration method, the partial least squares (PLS) regression [16] has long been used as a standard approach in Vis-NIR spectroscopy because it can extract information on target soil properties from a spectral matrix with hundreds or even thousands of wavebands [17,18]. The predictive accuracy is improved by waveband selection in PLS analysis by removing redundant wavebands [19,20]. Our previous papers showed that the application of waveband selection improves the performance of PLS analysis for estimating the total carbon (TC) content of paddy soils in Madagascar [21–23]. However, as a linear multivariate calibration, the accuracy of PLS analysis tends to decrease because of the nonlinear nature of the relationship between spectral data and the dependent variable [24]. To overcome this issue, machine learning techniques have been increasingly adopted because they can account for the nonlinearity associated with soil spectral responses. The major machine learning approaches are artificial neural networks (ANNs) [25], support-vector machines (SVM) [26], and random forest (RF) [27–29].

More recently, deep learning approaches have rapidly evolved in machine learning techniques with promising results for data analysis in nature [30] and have also been tested for soil spectral calibration [31,32]. Among deep learning approaches, convolutional neural networks (CNNs) [33] are one of the most popular learning architectures. In the remote sensing of agriculture, CNN-based approaches have been applied to various images with different research objectives, such as land cover classification [34], weed mapping [35], and crop yield prediction [36]. CNN-based deep learning was originally suitable for 2D image data, but its applicability has been extended to one-dimensional (1D) spectral data in recent years. For example, Padarian et al. [32] demonstrated that the CNN model, by converting the 1D soil spectra into a 2D spectrogram as input, can predict the soil TC, cation exchange capacity (CEC), clay, sand, and pH with better accuracy than conventional methods (PLS and Cubist). Ng et al. [31] developed a 1D-CNN model and compared it with a 2D-CNN model to estimate the major soil properties (TC, organic carbon (OC), CEC, clay, sand, and pH) based on the Kellogg Soil Survey Laboratory (KSSL) database. The results indicated that the 1D-CNN model was more effective than the 2D-CNN model. For areas not included in the soil spectral libraries, a local model can be developed or improved by transfer learning from the global model [37]. However, the prediction accuracy for P is lower than that for other soil components, and thus, improvements in predictive accuracy are required. Tsakiridis et al. [38] reported that the R^2 for soluble P in sodium hydrogen carbonate was 0.42, while the R^2 for clay content, soil organic carbon, and total nitrogen was 0.86, 0.86, and 0.83, respectively, using the visible, near-infrared and shortwave-infrared (VNIR-SWIR) spectra of the Land Use and Coverage Area Frame Survey (LUCAS) soil spectral library in combination with a localized multi-channel 1-D CNN model.

As mentioned above, state-of-the-art deep learning approaches with soil spectral libraries are currently available to construct models for predicting soil properties at local, regional, and global scales. However, soil P predictions remain challenging despite their importance to the sustainable management of agricultural systems aimed at addressing both soil degradation in the tropics and the environmental impacts of its excess use [39]. This is because soil P has no specific absorption in the Vis-NIR wavelength region [40]. Our previous studies in a range of rice cultivation soils in Madagascar identified the important wavelength for estimation and developed a reliable prediction model for oxalate-extractable P in soils, which is a suitable indicator of P availability for rice in the region [41], using PLS analysis with waveband selection [23,42]. Furthermore, different P cycling between natural and cultivated systems resulted in different soil P forms, which consequently decreased the accuracy of a common model in both land-use systems (cultivated vs. natural) [42]. The estimation accuracy should be improved by applying a comprehensive model to soil diagnosis in farm fields and other ecosystems for data compatibility and usability.

Therefore, the purpose of the present study was to develop a model based on the 1D-CNN architecture for estimating the soil available P content (Pox) in Madagascar. We used the dataset from our previous reports [23,42] because in Madagascar, only a small number of qualified datasets have been recorded in the soil spectral libraries: continent scale ($n = 82$) [13] and global scale ($n = 18$) [15]. Furthermore, because most soil spectral libraries use total P (TP) as reference data, the transfer learning approach cannot be applied to localize the model to Madagascar. Therefore, we compared the predictive ability of the 1D-CNN model to the previously used regression approaches, PLS and RF. We also performed a sensitivity analysis to identify the important wavebands used by the CNN model to predict soil Pox, and then evaluated the importance of the wavebands showing high sensitivity compared to PLS and RF.

2. Materials and Methods

2.1. Study Site and Dataset

The combined dataset included soil Pox content and spectral measurement data from 318 sampling points from the central highlands and eastern forest regions in Madagascar (Figure 1). Soil sampling was carried out in cultivated ($n = 244$) and natural ($n = 74$) systems. The climate in the area is sub-humid to humid tropics with an annual rainfall of 1381–2500 mm and a mean annual temperature of 16.9–24.0°C, depending on the elevation (110–1667 m above sea level). The area is dominated by Geric Ferralsols, which are generally acidic with low available P [43,44].

In the central highlands, soil sampling was performed in farmers' fields under irrigated ($n = 173$) and upland ($n = 71$) rice systems in 2015–2016 and 2018–2019. Surface soil samples were collected at a depth of 0–15 cm as composites of three to four cores in each field. In the eastern forest region, 74 soil samples were collected in 2014 and 2015 from forest ($n = 16$) and non-forest ($n = 58$) sites. The non-forest sites included fallow and degraded land. The soil samples were air-dried for 14 days and sieved to <2 mm. Soil P was extracted using the acid ammonium oxalate method [45], and the concentration of P in the oxalate extraction was analyzed using the malachite green colorimetric method [46].

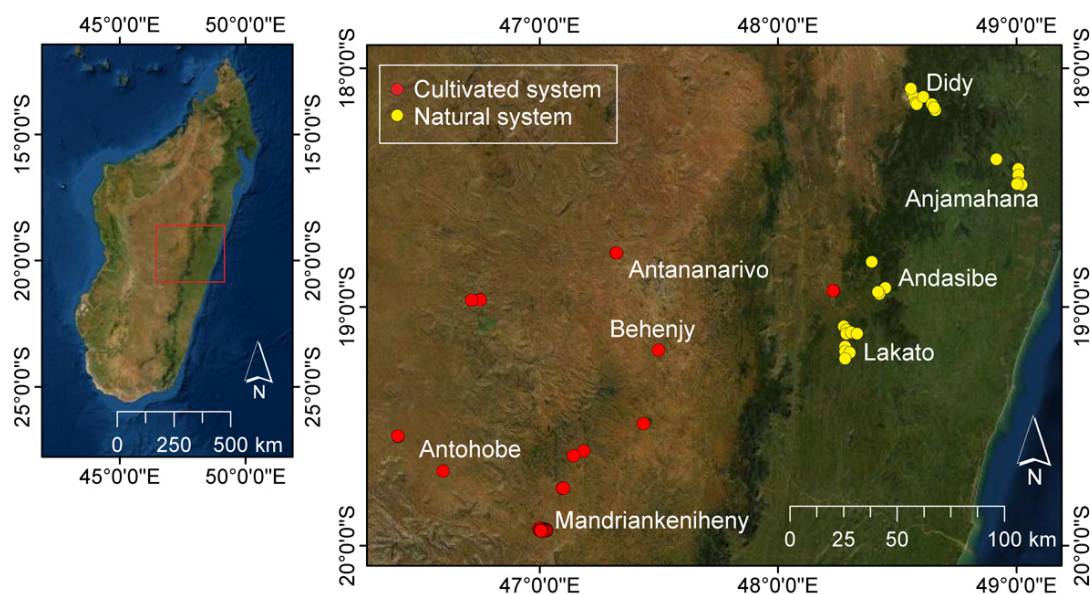


Figure 1. Studied regions and soil sampling points. Sources of background image: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, AeroGRID, IGN, and the GIS User Community.

2.2. Spectral Measurements and Preprocessing

Spectral measurements of soil samples were taken in dark rooms at the Japan International Research Center for Agricultural Science (JIRCAS) and the Laboratoire des Radioisotopes, Antananarivo University, Madagascar, using ASD FieldSpec 4 Hi-Res (ASD Inc., Longmont, CO, USA) and an ASD contact probe. Compared to multispectral sensors, the hyperspectral soil reflectance data in laboratory Vis-NIR spectroscopy has some advantages such as waveband richness, sharpness of wavebands, and spectral continuity [47]. The ASD FieldSpec measures the spectral reflectance in the 350–2500 nm wavelength region. The spectral resolution (full width at half maximum; FWHM) was 3 nm in the 350–2500 nm range and 6 nm in the 1000–2500 nm range, which were resampled to 1 nm resolution wavelengths for output data using the cubic spline interpolation function in ASD software (RS3 for Windows; ASD Inc.). A Spectralon (Labsphere Inc., Sutton, NH, USA) reference panel (white reference) was used to optimize the ASD instrument prior to taking Vis-NIR reflectance measurements for each sample.

Bulk soil samples (<2 mm) were filled into an optical-glass Petri dish with a diameter of 85 mm and pressed to form a layer about 19 mm thick. The soil surface reflectance measurements were done on the surface for 25 times with five replications for each soil sample, and the spectral readings were averaged.

During preprocessing, spectral data were initially converted from reflectance (R) to absorbance ($A = \log(1/R)$). Then, the absorbance spectra were converted to the first derivative absorbance (FDA) spectra using a Savitzky-Golay smoothing filter [48] with a third-order, 15-band moving polynomial. In addition, FDA spectra were standardized using a standard normal variate transform (SNV) to reduce the particle size effect [49]. Finally, the FDA data of both edge wavelength regions (350–419 nm and 2401–2500 nm) were removed because of the low signal-to-noise ratios in the instruments. Thus, the remaining 1981 spectral bands between 420 and 2400 nm were used in the analyses.

2.3. Model Development

2.3.1. Partial Least Squares (PLS) Regression

PLS is one of the most widely used linear regression approaches to analyze high-dimensional datasets, such as hyperspectral and Vis-NIR data, because it combines all available waveband information without multicollinearity issues, unlike standard multiple linear regression analysis. The PLS regression treats each waveband as an independent

explanatory variable for estimating the response variable of the target component (soil Pox in the present study). The regression coefficient of PLS is computed by weighting the optimal number of PLS factors; it is also called a weighted coefficient (β_w). Here, a leave-one-out cross-validation (CV) method was used to select the optimal number of PLS factors to be included in the regression model. The important wavebands can be determined by the high absolute value of β_w .

2.3.2. Random Forest (RF) Regression

RF is a machine-learning approach that uses a non-linear statistical ensemble method that builds a large number of decision trees (*ntree*) for classification or regression and then averages them [50]. This approach evaluates the relationship between explanatory and response variables using a set of decision rules constructed by recursively dividing the input space into smaller regions. In the procedure, the model is developed with a set of trees by selecting a subset based on a bagging approach. Each tree grows until it reaches a certain minimum number of nodes (*node size*). Then, the remaining subset (called the out-of-bag (OOB) sample) is used for internal cross-validation to assess the average accuracy and error rate over all predictions [51]. In addition, the OOB is used to estimate the variable (feature) importance. Finally, the output class is calculated using the maximum votes from the *ntree* in the forest. In this study, we set *ntree* = 5000 and *node size* = 5 to construct the RF model. The importance of spectral wavebands in the model was evaluated by the Gini impurity, which was calculated by summing all decreases in Gini impurity at each tree node split, normalized by the number of trees.

2.3.3. One-Dimensional Convolutional Neural Network (1D-CNN)

CNNs are the most widely used architectures in deep learning approaches. CNN architectures are generally composed of an input layer, several hidden layers (convolution layers, pooling, and fully connected), and an output layer. Among the CNNs, 1D-CNN has an input layer and 1D filters on the convolution layers suitable for one-dimensional spectral data [52]. In the present study, we used the 1D-CNN architecture proposed by Ng et al. [31] (Figure 2) because the structure is simple and well described, and there are existing reports on other soil properties. This is good to compare the accuracy of our newly performed soil oxalate-P estimation with the results of other soil properties. The architecture included 10 hidden layers (Table 1): four convolutional layers, four max-pooling layers, and two fully connected layers. The activation function used a rectified linear unit (ReLU) for all hidden layers [53]. Two dropout rates of 0.4 and 0.2 were used to avoid overfitting [31].

Table 1. Architecture of one-dimensional convolutional neural network (CNN).

Type	Filter Size	No. of Filters	Activation
Convolutional	20	32	ReLU
Max-pooling	2	-	-
Convolutional	20	64	ReLU
Max-pooling	5	-	-
Convolutional	20	128	ReLU
Max-pooling	5	-	-
Convolutional	20	256	ReLU
Max-pooling	5	-	-
Dropout (0.4)	-	-	-
Flatten	-	-	-
Fully-connected	-	100	ReLU
Dropout (0.2)	-	-	-
Fully-connected	-	1	Linear

ReLU, rectified linear unit.

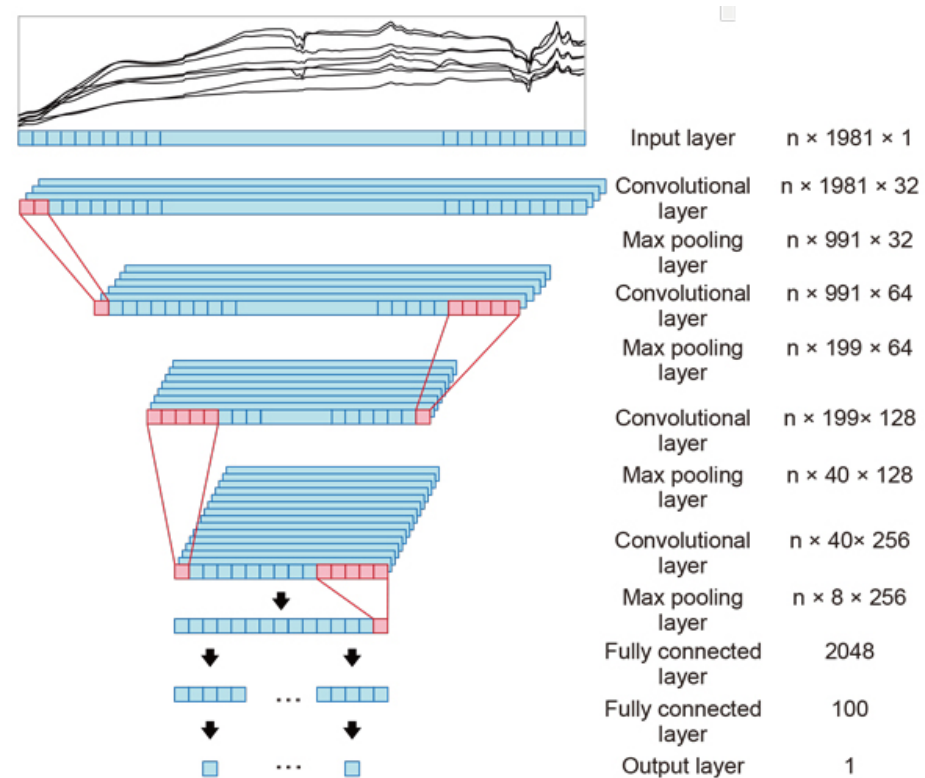


Figure 2. One-dimensional neural convolutional network architecture (revised from the Figure 1 in Ng et al. [31]).

2.4. Data Handling and Implementation

Our whole dataset ($n = 318$) was split into training ($n = 238$) and test ($n = 80$) subsets. The training data were used to develop the PLS, RF, and 1D-CNN models. Then, the models were applied to the test dataset to assess the predictive abilities of the models for soil Pox prediction. Table 2 summarizes the minimum, maximum, median, mean, and standard deviation (SD) values of Pox (mg P kg^{-1}) for the training ($n = 238$) and test ($n = 80$) datasets, including data collected from different land-use systems (natural and cultivated). Soils in cultivated systems showed higher Pox values and wider ranges compared with soils in natural systems, probably due to the effect of fertilization on cultivated land [42].

Table 2. Summary statistics of soil oxalate-extractable P data.

Dataset	System	n	Min	Max	Median	Mean	SD
Training	All	238	21.9	1172.0	67.7	214.7	278.0
	Cultivated	183	23.7	1172.0	106.0	268.7	296.5
	Natural	55	21.9	53.9	34.8	35.1	7.2
Test	All	80	22.3	1225.2	68.5	220.9	290.0
	Cultivated	62	22.3	1225.2	106.2	274.8	309.6
	Natural	18	22.9	57.9	33.8	35.5	9.5

n , number of samples; SD, standard deviation

The data distributions of soil Pox for all systems in the training and test datasets are shown in Figure 3. The data range of soil Pox content was similar for the training ($21.9\text{--}1172.0 \text{ mg P kg}^{-1}$) and test ($22.3\text{--}1225.2 \text{ mg P kg}^{-1}$) datasets. Moreover, the SD (278.0 and $290.0 \text{ mg P kg}^{-1}$) showed similar values. This is important for correct validation by applying the model created with the training data to the test data. The data distribution was left-skewed because of the predominance of strongly weathered soils in the region [43,44], with higher mean (214.7 and $220.9 \text{ mg P kg}^{-1}$ for training and test datasets, respectively)

than median (67.7 and 68.5 mg P kg⁻¹) values. Theoretically, the SD, data range, and distribution pattern affect the accuracy of the regression analysis. Compared to previous studies (121–991 mg P kg⁻¹) [54], our data set included a larger variance of Pox values, which can be expected to improve the predictive ability of a model.

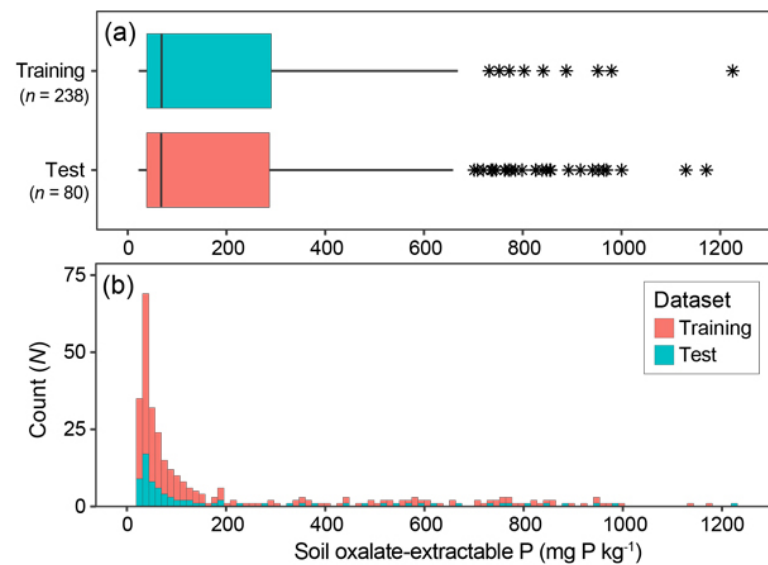


Figure 3. Box-and-whiskers plot (a) and histogram (b) of soil oxalate-extractable P in the training and test datasets.

In the present study, PLS and RF models were implemented using R ver. 3.6.3 [55] with the packages ‘pls’ ver. 2.7-2 [56] and ‘randomForest’ ver. 4.6-14 [57]. The CNN was implemented using Python ver. 3.6.9 [58] with the Keras library ver. 2.2.4 [59], and TensorFlow ver. 1.14.0 [60] backend. All computations were performed on a desktop PC with Intel® Core™ i9-9900X 3.5 GHz processor, 64 GB RAM, and the NVIDIA GeForce RTX 2080 Ti.

2.5. Predictive Accuracy Evaluation

To assess the predictive accuracy of the PLS, RF, and 1D-CNN models, a 5-fold cross-validation was performed in the training dataset ($n = 238$). In the procedure, the training data were split randomly into 5-folds. Each model was built on 4 (=5 – 1) folds, and then the error of the remaining 1-fold was recorded as validation data. The process was repeated until each of the 5-folds served as a validation data set. The mean values of the coefficient of determination (R^2), root mean squared error (RMSE), and bias were used to assess model accuracy. The RMSE and bias were defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}, \quad (1)$$

$$\text{bias} = \frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)}{n}, \quad (2)$$

where Y and \hat{Y} are the observed and predicted soil Pox values, respectively, and n is the total number of observations i .

Moreover, the models were applied to the test dataset ($n = 80$), and then the predictive ability was evaluated from the R^2 , RMSE, bias, and the ratio of performance to the interquartile range (RPIQ) [61]. The RPIQ is an auxiliary indicator of the model, and the performance of predictive ability and future reliability can be judged as follows: (i) >2.5 : excellent model, (ii) 2.0–2.5: very good model with predictive ability, (iii) 1.7–2.0: good model, (iv) 1.4–1.7: fair model in need of some improvement, and (v) <1.4 : model with very poor predictive ability [62].

2.6. Sensitivity Analysis of 1D-CNN Model for Evaluating Important Wavebands

In the PLS and RF models, the importance of wavebands for estimating soil Pox can be assessed using PLS regression coefficients and RF importance (Gini purity index). However, it is difficult to directly evaluate the important wavebands in the 1D-CNN model. To assess the importance of wavebands in the 1D-CNN model, a sensitivity analysis was performed using the method proposed by Ng et al. [31], in which the sensitivity is calculated as a function of the variance of the model for each window of spectra [63]:

$$S_i = \frac{V(f(X_1, \dots, X_2, \dots, X_n) - f(\bar{X}))}{V(Y)}, \quad (3)$$

where V is the variation calculation, $f(X_1, \dots, X_2, \dots, X_n)$ is the prediction of spectra due to variation in waveband i with the other wavebands held constant on average, $f(\bar{X})$ is the prediction value using the mean values of the spectra, and Y is the observed value of the soil Pox. In essence, this approach compares the sensitivity of the wavebands and calculates how the model changes [31].

3. Results

3.1. Comparison of Predictive Abilities in PLS, RF, and 1D-CNN Models

To compare the predictive accuracy in the PLS, RF, and 1D-CNN models using FDA spectral datasets, the mean values of R^2 , RMSE, bias, and the relative improvement (%RMSE) from 5-fold cross-validation were compared in the training dataset ($n = 238$) (Table 3). Computational times for developing the PLS, RF, and 1D-CNN models were 3.9, 264.1, and 5908.45 s (CPU time), respectively. Clearly, the 1D-CNN required a lot of learning time to develop the model, but it showed much better predictive accuracy ($R^2 = 0.989$, RMSE = 35.636) than the RF ($R^2 = 0.842$, RMSE = 108.820) and PLS ($R^2 = 0.827$, RMSE = 114.854) models. Compared with the RMSE values of the PLS model, the RF model showed a slight improvement in accuracy (%RMSE = 5.254%), while the 1D-CNN model showed a significant improvement (68.973%).

Table 3. Mean values of R^2 , RMSE, bias, and relative improvement (%RMSE) from 5-fold cross validation in the training dataset for developing the partial least squares (PLS), random forest (RF), and one dimensional convolutional neural network (1D-CNN) models.

Model	R^2	RMSE	Bias	%RMSE ¹
PLS	0.827	114.854	16.577	-
RF	0.842	108.820	13.517	5.254
1D-CNN	0.989	35.636	-2.202	68.973

¹ Relative improvement (%RMSE) of the RF and CNN models in comparison to PLS model. RMSE, root mean squared error.

In addition, to confirm the reliability of the models, predictive ability was assessed from the relationship between the observed and predicted soil Pox content from the PLS, RF, and 1D-CNN models in the individual test dataset ($n = 80$) (Figure 4). The R^2 , RMSE, bias, RPIQ, and the relative improvement in RMSE (%RMSE) are summarized in Table 4. Similarly to the cross-validation results in the training dataset, the 1D-CNN showed higher predictive accuracy ($R^2 = 0.878$, RMSE = 101.154) than the RF ($R^2 = 0.808$, RMSE = 126.894) and PLS ($R^2 = 0.792$, RMSE = 132.694) models. Compared with the PLS model, the RF

and 1D-CNN models indicated a 4.37% and 23.77% relative improvement in %RMSE, respectively. Based on the RPIQ values, the PLS and RF models (RPIQ < 2.0) could be judged as good models, but they required further work to improve quality and increase future applicability, while the 1D-CNN model (RPIQ = 2.492) was considered to be a very good model with high predictive ability.

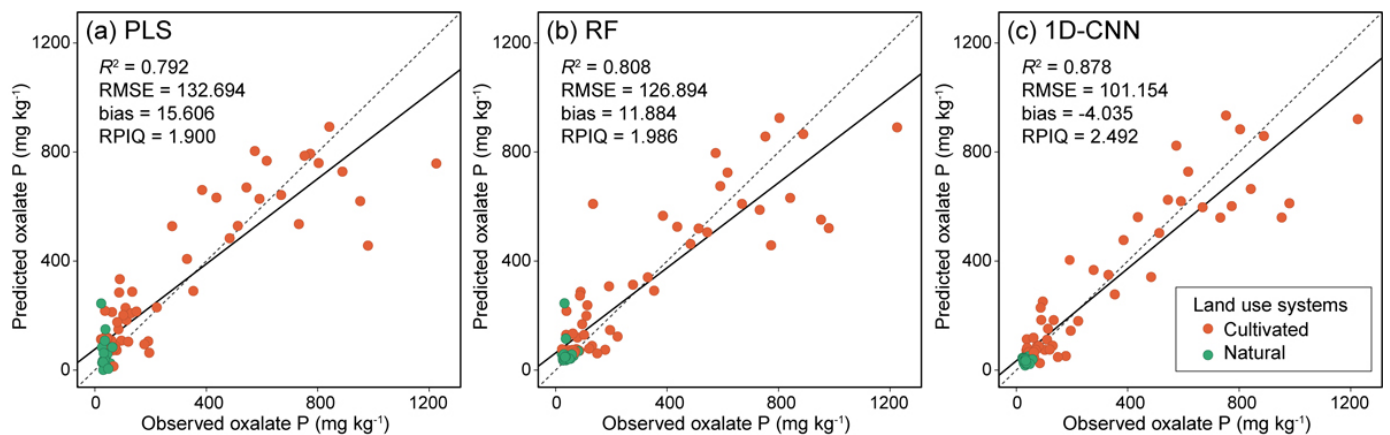


Figure 4. Relationships between observed and predicted values of soil oxalate-extractable P contents with (a) partial least squares (PLS), (b) random forest (RF), and (c) one-dimensional convolutional neural network (1D-CNN) models. RMSE, root mean squared error; RPIQ, ratio of performance to the interquartile range

Table 4. Comparison of predictive abilities in the partial least squares (PLS), random forest (RF), and one dimensional convolutional neural network (1D-CNN) models.

Model	R^2	RMSE	Bias	RPIQ	%RMSE ¹
PLS	0.792	132.694	15.606	1.900	-
RF	0.808	126.894	11.884	1.986	4.371
1D-CNN	0.878	101.154	-4.035	2.492	23.769

¹ Relative improvement (%RMSE) of the RF and CNN models in comparison to PLS model. RMSE, root mean squared error; RPIQ, ratio of performance to the interquartile range.

3.2. Important Wavelengths

The importance of each waveband in the models for soil Pox prediction was assessed. Figure 5 shows the FDA spectra in logarithmic form to easily refer to the wavelength characteristics depending on the soil P values, and the absolute values of PLS coefficients, RF importance (Gini purity index), and 1D-CNN sensitivity (s) as indicators of the importance of each waveband in the soil Pox estimation models.

The important wavebands for the PLS model were revealed in the visible (433, 446, 509, 550, and 590 nm) and NIR (1001, 1412, 1904, and 2219 nm) wavelength regions. The RF model showed constant values over the entire wavelength range, but the 446, 1336, 1366, 2365, and 2397 nm were particularly important. In the 1D-CNN model, high sensitivity was observed in the narrow and limited wavelength regions, with peaks at 432, 590, 1433, and 2274 nm. The three peaks (432, 590, and 1433 nm) did not exactly match but also showed a high value in PLS coefficients, and thus were considered to be important for estimating soil Pox.

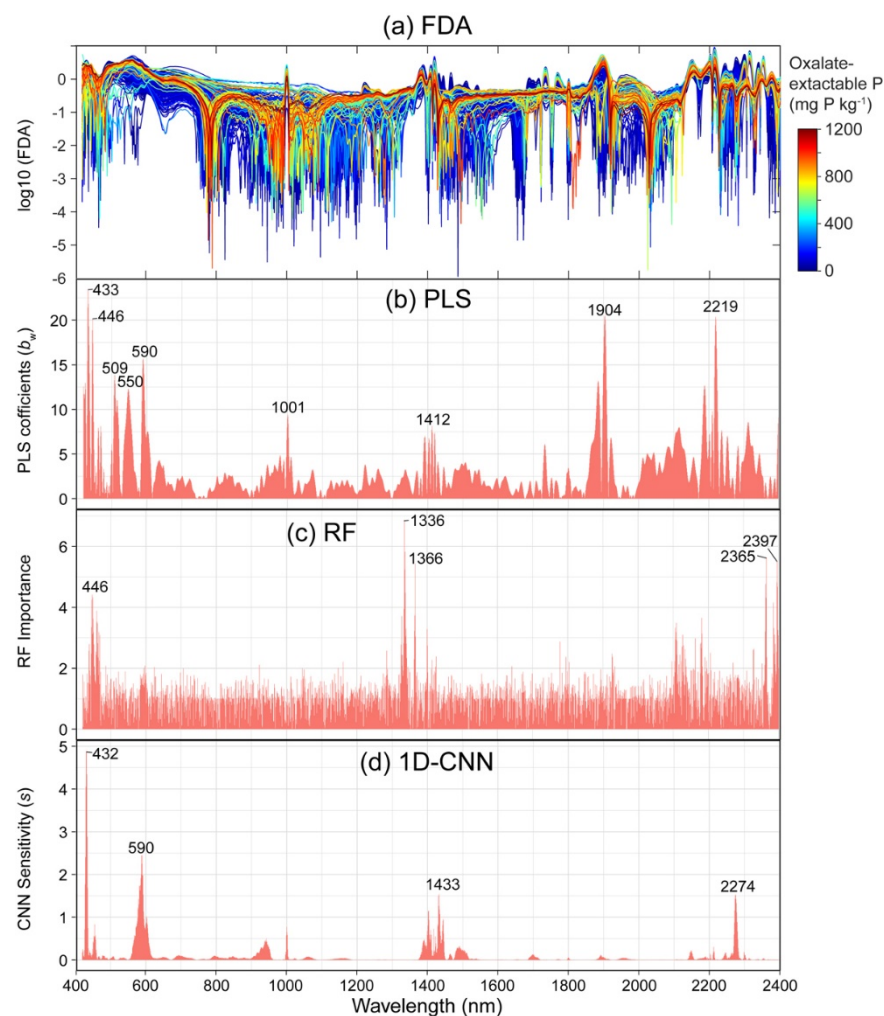


Figure 5. (a) First derivative absorbance (FDA) spectra on \log_{10} scale of soil oxalate-extractable P (Pox) content and importance of wavelength to (b) partial least squares (PLS), (c) random forest (RF), and (d) one dimensional convolutional neural network (1D-CNN) models in soil Pox estimation.

4. Discussion

As a deep learning approach suitable for 1D spectral data, 1D-CNN with Vis-NIR spectral data was applied to predict soil Pox content in Madagascar, and its predictive ability was compared with that of PLS and RF models. The 5-fold cross-validation in the training dataset demonstrated that the 1D-CNN model can estimate soil Pox from Vis-NIR spectral data with improved accuracy compared to the other common methods of PLS and RF. Even though our soil samples were collected from different land-use systems, the best predictive ability in the 1D-CNN model was also confirmed in the individual test dataset. These results indicated that a comprehensive model to predict Pox in soils with high accuracy could be developed irrespective of land use systems using a deep learning approach with a 1D-CNN model rather than PLS (conventional method) and RF (another machine learning approach). Similarly, the lowest bias was observed in the 1D-CNN model. Based on the RPIQ value, our 1D-CNN model could be considered to have a very good predictive ability, and therefore could provide high-quality results in future applications. These results confirm that the CNN-based model can outperform PLS and other machine learning approaches for estimating soil properties, as suggested by previous studies [31,32]. Another update from our previous report using PLS regression analysis [42] was that the 1D-CNN model enabled the prediction of soil P as a single model, even based on the dataset collected from different land-use systems in Madagascar. These improvements are considered to be advances in the holistic understanding of soil P dynamics and their

rational management in agriculture and natural ecosystems in Madagascar. Furthermore, the timely and accurate prediction of available P content in soils using our 1D-CNN model can be expected to contribute to effective fertilizer management in agriculture and the sustainable management of ecosystems.

We also assessed the important wavebands of the 1D-CNN model using a sensitivity analysis method developed by Ng et al. [63]. Assessing the important specific spectral features and their regions may help clarify the relationship between spectral signatures and underlying molecular activity within the spectral wavelength regions used in the calibration [64]. In the Vis-NIR wavelength region, theoretically, there is no specific absorption by Pox; thus, differences in spectral shape due to P content cannot be clearly determined [65–67]. However, indirect correlations between P forms and spectrally active soil properties can be studied. For example, Pätzold et al. [68] reported that the most relevant variables for the estimation of soil P by PLS were selected based on the variable importance in projection (VIP) method, and that PLS regression coefficients were located around 500 nm and 2200–2400 nm. Our previous study in Madagascar identified important Pox regions at 454–660 nm and 1732–2312 nm based on five runs of the variable selection method genetic algorithm followed by PLS regression [23].

In the present study, our 1D-CNN model showed high sensitivity in the narrow and limited wavelength regions with peaks at 432, 590, 1433, and 2274 nm. Since these wavebands correspond to the regions reported previously [23,68], it can be suggested that soil Pox tends to be primarily associated with the visible (400–660 nm) and longer NIR (1700–2400 nm) wavelength regions. The visible part of the spectral region was dominated by absorption due to organic matter (OM) and electronic transitions of Fe, and the selected wavebands at 432 and 590 nm were attributed to absorption by ferric oxide and OM, respectively [40,69–71]. Our previous study also reported that oxalate-extractable Al and Fe were positively correlated with soil TC and organic P content, respectively [5]. The longer NIR wavelength region (1700–2400 nm) was dominated by a combination of bands from intense fundamental vibrations of O-H, C-H bonds, and Al metal-OH groups [7,72]. The waveband at 1433 nm was attributed to sorptive water, and that at 2274 nm was associated with gibbsite (Al-oxide mineral) [73]. The significance of the sorptive water in air-dried soils for predicting Pox was corroborated by our previous finding that active Al was clearly correlated with air-dried soil moisture content (Nishigaki et al., under review).

Overall, our findings were in agreement with those of earlier studies that reported that machine learning outperformed PLS regression. However, such approaches may not be suitable for all situations because the deep learning approach requires large, qualified datasets to develop good prediction models [31]. Therefore, many CNN-based deep learning approaches have been performed using a large topsoil spectral dataset, such as the KSSL database ($n = 14,594$) [31] or LUCAS soil spectral libraries ($n = 19,036$) [32,37,38,74]. Our dataset collected from the central and eastern regions of Madagascar was relatively small ($n = 318$). This is a limitation of our study for developing a robust model on a larger spatial scale. However, the Pox targeted in our study was more highly correlated with other soil properties (e.g., oxalate-extractable Al and Fe, and soil pH) than total P or easily soluble P [5]. This possibly resulted in a higher prediction accuracy using spectral data for Pox than that achieved in previous reports. Further evaluation by updating the model with a larger dataset will be required in future studies to map and assess the spatial distribution of soil Pox status on a larger spatial scale in Madagascar.

5. Conclusions

The present study investigated the performance of a 1D-CNN model through comparison with two conventional methods, PLS and RF, for estimating soil Pox content with Vis-NIR spectral data in soil samples collected from cultivated (upland and flooded rice fields) and natural (forest and non-forest) systems in Madagascar. The main conclusions are as follows:

- With the potential to provide high predictive ability and performance in deep learning approaches, Vis-NIR spectroscopy with 1D-CNN is a promising method for predicting soil Pox content.
- Our 1D-CNN model provided the best predictive ability to estimate soil Pox content compared with the PLS and RF models.
- The RPIQ value from the 1D-CNN is suggested to be a very good model with high predictive ability for future applicability.
- The important wavebands from the sensitivity analysis of the 1D-CNN model were revealed in the visible region (432 and 590 nm) associated with Fe-oxides and diverse functional groups in soil OM; at 1433 nm, associated with water absorption; and at around 2270 nm with gibbsite (Al oxide mineral). These wavelength regions are known to be of high importance in the PLS model, and are in line with previous studies.
- The 1D-CNN model we developed allowed soil P prediction based on a single model, even using data from different land-use systems.

Our findings illustrate the potential of deep learning approaches for predicting soil P availability using a proximal sensing technique. The main contribution from our study is the demonstration of a 1D-CNN model that can be applied to soil Pox prediction in different ecosystems from the central to eastern regions of Madagascar. Moreover, rapid and accurate prediction of soil Pox using our results can be expected to contribute to effective fertilizer management in agriculture and the sustainable management of ecosystems. Nevertheless, the 1D-CNN approach employed in this study should be updated with a larger dataset and further evaluated on a larger spatial scale in Madagascar.

Author Contributions: K.K., T.N., Y.T., N.M., and T.R. designed this study; Y.T., T.N., A.A., H.R., and M.R. performed the field work and carried out the soil chemical analyses; K.K., A.A., H.R., and M.R. performed the laboratory spectral measurements and data processing; K.K., T.N., and Y.T. wrote the manuscript; and all authors revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA) (Grant No. JPMJSA1608).

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Acknowledgments: The authors would like to express their gratitude to Marie-Paule Razafimanantsoa and laboratory staff members of Laboratoire des Radio-Isotopes, Université d'Antananarivo, for their support in conducting the oxalate-extractable P analysis. We are grateful to the editor and anonymous reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Koné, B.; Amadji, G.L.; Aliou, S.; Diatta, S.; Akakpo, C. Nutrient constraint and yield potential of rice on upland soil in the south of the Dahoumey gap of West Africa. *Arch. Agron. Soil Sci.* **2011**, *57*, 763–774. [[CrossRef](#)]
2. Andriamananjara, A.; Rakotoson, T.; Razafimbelo, T.; Rabeharisoa, L.; Razafimanantsoa, M.P.; Masse, D. Farmyard manure improves phosphorus use efficiency in weathered P deficient soil. *Nutr. Cycl. Agroecosyst.* **2019**, *115*, 407–425. [[CrossRef](#)]
3. Balemi, T.; Negisho, K. Management of soil phosphorus and plant adaptation mechanisms to phosphorus stress for sustainable crop production: A review. *J. Soil Sci. Plant. Nutr.* **2012**, *12*, 547–562. [[CrossRef](#)]
4. Turner, B.L.; Engelbrecht, B.M.J. Soil organic phosphorus in lowland tropical rain forests. *Biogeochemistry* **2011**, *103*, 297–315. [[CrossRef](#)]
5. Nishigaki, T.; Tsujimoto, Y.; Rinasoa, S.; Rakotoson, T.; Andriamananjara, A.; Razafimbelo, T. Phosphorus uptake of rice plants is affected by phosphorus forms and physicochemical properties of tropical weathered soils. *Plant. Soil* **2019**, *435*, 27–38. [[CrossRef](#)]
6. Guo, F.; Yost, R.S. Quantifying the available soil phosphorus pool with the acid ammonium oxalate method. *Soil Sci. Soc. Am. J.* **1999**, *63*, 651–656. [[CrossRef](#)]

7. Ben-Dor, E. Quantitative remote sensing of soil properties. In *Advances in Agronomy*; Academic Press: Cambridge, MA, USA, 2002; Volume 75, pp. 173–243, ISBN 9780120007936.
8. Ben-Dor, E.; Banin, A. Visible and near-infrared (0.4–1.1 μm) analysis of arid and semiarid soils. *Remote Sens. Environ.* **1994**, *274*, 261–274. [[CrossRef](#)]
9. Viscarra Rossel, R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [[CrossRef](#)]
10. Nocita, M.; Stevens, A.; van Wesemael, B.; Brown, D.J.; Shepherd, K.D.; Towett, E.; Vargas, R.; Montanarella, L. Soil spectroscopy: An opportunity to be seized. *Glob. Chang. Biol.* **2015**, *21*, 10–11. [[CrossRef](#)] [[PubMed](#)]
11. Romero, D.J.; Ben-Dor, E.; Demattê, J.A.M.; Souza, A.B.E.; Vicente, L.E.; Tavares, T.R.; Martello, M.; Strabeli, T.F.; da Silva Barros, P.P.; Fiorio, P.R.; et al. Internal soil standard method for the Brazilian soil spectral library: Performance and proximate analysis. *Geoderma* **2018**, *312*, 95–103. [[CrossRef](#)]
12. Li, S.; Ji, W.; Chen, S.; Peng, J.; Zhou, Y.; Shi, Z. Potential of VIS-NIR-SWIR spectroscopy from the Chinese Soil Spectral Library for assessment of nitrogen fertilization rates in the paddy-rice region, China. *Remote Sens.* **2015**, *7*, 7029–7043. [[CrossRef](#)]
13. Johnson, J.-M.; Vandamme, E.; Senthikumar, K.; Sila, A.; Shepherd, K.D.; Saito, K. Near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for assessing soil fertility in rice fields in sub-Saharan Africa. *Geoderma* **2019**, *354*, 113840. [[CrossRef](#)]
14. Stevens, A.; Nocita, M.; Tóth, G.; Montanarella, L.; van Wesemael, B. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *PLoS ONE* **2013**, *8*, e66409. [[CrossRef](#)]
15. Viscarra Rossel, R.A.; Behrens, T.; Ben-Dor, E.; Brown, D.J.; Demattê, J.A.M.; Shepherd, K.D.; Shi, Z.; Stenberg, B.; Stevens, A.; Adamchuk, V.; et al. A global spectral library to characterize the world's soil. *Earth Sci. Rev.* **2016**, *155*, 198–230. [[CrossRef](#)]
16. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
17. Conforti, M.; Buttafuoco, G.; Leone, A.P.; Aucelli, P.P.C.; Robustelli, G.; Scarciglia, F. Studying the relationship between water-induced soil erosion and soil organic matter using Vis-NIR spectroscopy and geomorphological analysis: A case study in southern Italy. *Catena* **2013**, *110*, 44–58. [[CrossRef](#)]
18. Conforti, M.; Castrignanò, A.; Robustelli, G.; Scarciglia, F.; Stelluti, M.; Buttafuoco, G. Laboratory-based Vis-NIR spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content. *Catena* **2015**, *124*, 60–67. [[CrossRef](#)]
19. Yang, H.; Kuang, B.; Mouazen, A.M. Quantitative analysis of soil nitrogen and carbon at a farm scale using visible and near infrared spectroscopy coupled with wavelength reduction. *Eur. J. Soil Sci.* **2012**, *63*, 410–420. [[CrossRef](#)]
20. Vohland, M.; Ludwig, M.; Thiele-Bruhn, S.; Ludwig, B. Determination of soil properties with visible to near- and mid-infrared spectroscopy: Effects of spectral variable selection. *Geoderma* **2014**, *223*, 88–96. [[CrossRef](#)]
21. Kawamura, K.; Tsujimoto, Y.; Rabenarivo, M.; Asai, H.; Andriamananjara, A.; Rakotoson, T. Vis-NIR spectroscopy and PLS regression with waveband selection for estimating the total C and N of paddy soils in Madagascar. *Remote Sens.* **2017**, *9*, 1081. [[CrossRef](#)]
22. Kawamura, K.; Nishigaki, T.; Tsujimoto, Y.; Andriamananjara, A.; Rabenarivo, M.; Asai, H.; Rakotoson, T.; Razafimbelo, T. Exploring relevant wavelength regions for estimating soil total carbon contents of rice fields in Madagascar from Vis-NIR spectra with sequential application of backward interval PLS. *Plant. Prod. Sci.* **2021**, *24*, 1–14. [[CrossRef](#)]
23. Kawamura, K.; Tsujimoto, Y.; Nishigaki, T.; Andriamananjara, A.; Rabenarivo, M.; Asai, H.; Rakotoson, T.; Razafimbelo, T. Laboratory visible and near-infrared spectroscopy with genetic algorithm-based partial least squares regression for assessing the soil phosphorus content of upland and lowland rice fields in Madagascar. *Remote Sens.* **2019**, *11*, 506. [[CrossRef](#)]
24. Araújo, S.R.; Wetterlind, J.; Demattê, J.A.M.; Stenberg, B. Improving the prediction performance of a large tropical vis-NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. *Eur. J. Soil Sci.* **2014**, *65*, 718–729. [[CrossRef](#)]
25. Kuang, B.; Tekin, Y.; Mouazen, A.M. Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content. *Soil Tillage Res.* **2015**, *146*, 243–252. [[CrossRef](#)]
26. Morellos, A.; Pantazi, X.E.; Moshou, D.; Alexandridis, T.; Whetton, R.; Tziotzios, G.; Wiebensohn, J.; Bill, R.; Mouazen, A.M. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosyst. Eng.* **2016**, *152*, 104–116. [[CrossRef](#)]
27. Cipullo, S.; Nawar, S.; Mouazen, A.M.; Campo-Moreno, P.; Coulon, F. Predicting bioavailability change of complex chemical mixtures in contaminated soils using visible and near-infrared spectroscopy and random forest regression. *Sci. Rep.* **2019**, *9*, 4492. [[CrossRef](#)] [[PubMed](#)]
28. De Santana, F.B.; de Souza, A.M.; Poppi, R.J. Visible and near infrared spectroscopy coupled to random forest to quantify some soil quality parameters. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2018**, *191*, 454–462. [[CrossRef](#)] [[PubMed](#)]
29. Douglas, R.K.; Nawar, S.; Alamar, M.C.; Mouazen, A.M.; Coulon, F. Rapid prediction of total petroleum hydrocarbons concentration in contaminated soil using vis-NIR spectroscopy and regression techniques. *Sci. Total Environ.* **2018**, *616–617*, 147–155. [[CrossRef](#)] [[PubMed](#)]
30. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]

31. Ng, W.; Minasny, B.; Montazerolghaem, M.; Padarian, J.; Ferguson, R.; Bailey, S.; McBratney, A.B. Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma* **2019**, *352*, 251–267. [[CrossRef](#)]
32. Padarian, J.; Minasny, B.; McBratney, A.B. Using deep learning to predict soil properties from regional spectral data. *Geoderma Reg.* **2019**, *16*, e00198. [[CrossRef](#)]
33. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2323. [[CrossRef](#)]
34. Hsieh, T.-H.; Kiang, J.-F. Comparison of CNN Algorithms on Hyperspectral Image Classification in Agricultural Lands. *Sensors* **2020**, *20*, 1734. [[CrossRef](#)] [[PubMed](#)]
35. Sa, I.; Popović, M.; Khanna, R.; Chen, Z.; Lottes, P.; Liebisch, F.; Nieto, J.; Stachniss, C.; Walter, A.; Siegwart, R. WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming. *Remote Sens.* **2018**, *10*, 1423. [[CrossRef](#)]
36. Nevavuori, P.; Narra, N.; Lipping, T. Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.* **2019**, *163*, 104859. [[CrossRef](#)]
37. Padarian, J.; Minasny, B.; McBratney, A.B. Transfer learning to localise a continental soil vis-NIR calibration model. *Geoderma* **2019**, *340*, 279–288. [[CrossRef](#)]
38. Tsakiridis, N.L.; Keramaris, K.D.; Theocharis, J.B.; Zalidis, G.C. Simultaneous prediction of soil properties from VNIR-SWIR spectra using a localized multi-channel 1-D convolutional neural network. *Geoderma* **2020**, *367*, 114208. [[CrossRef](#)]
39. MacDonald, G.K.; Bennett, E.M.; Potter, P.A.; Ramankutty, N. Agronomic phosphorus imbalances across the world's croplands. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3086–3091. [[CrossRef](#)] [[PubMed](#)]
40. Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind, J. Visible and Near Infrared Spectroscopy in Soil Science. In *Advances in Agronomy*; Academic Press: Cambridge, MA, USA, 2010; Volume 107, pp. 163–215.
41. Rabeharisoa, L.; Razanakoto, O.R.; Razafimanantsoa, M.-P.; Rakotoson, T.; Amery, F.; Smolders, E. Larger bioavailability of soil phosphorus for irrigated rice compared with rainfed rice in Madagascar: Results from a soil and plant survey. *Soil Use Manag.* **2012**, *28*, 448–456. [[CrossRef](#)]
42. Rakotonindrina, H.; Kawamura, K.; Tsujimoto, Y.; Nishigaki, T.; Razakamanarivo, H.; Andrianary, B.H.; Andriamananjara, A. Prediction of soil oxalate phosphorus using visible and near-infrared spectroscopy in natural and cultivated system soils of madagascar. *Agriculture* **2020**, *10*, 177. [[CrossRef](#)]
43. Andriamaniraka, H. *Le Phosphore et la Fertilisation Phosphatée Dans les sols Ferrallitiques à Madagascar: Amélioration de la Fertilité des Sols*; Université d'Antananarivo: Antananarivo, Madagascar, 2016.
44. Nishigaki, T.; Ikazaki, K.; Tsujimoto, Y.; Andriamananjara, A.; Rakotoson, T.; Razafimbelo, T. Soil survey of the east coast and the central highlands indicates need to update Madagascar soil map. *Soil Sci. Plant. Nutr.* **2020**, *66*, 469–480. [[CrossRef](#)]
45. Schwertmann, U. Differenzierung der Eisenoxide des Bodens durch Extraktion mit Ammoniumoxalat-Lösung. *Zeitschrift für Pflanzenernährung Düngung Bodenkd.* **1964**, *105*, 194–202. [[CrossRef](#)]
46. Van Veldhoven, P.P.; Mannaerts, G.P. Inorganic and organic phosphate measurements in the nanomolar range. *Anal. Biochem.* **1987**, *161*, 45–48. [[CrossRef](#)]
47. Inoue, Y.; Miah, G.; Sakaiya, E.; Kaneko, K.; Kawamura, K. NDSI map and IPLS using hyperspectral data for assessment of plant and ecosystem variables: With a case study on remote sensing of grain protein content, chloro. *J. Remote Sens. Soc. Jpn.* **2008**, *28*, 317–330.
48. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
49. Brunet, D.; Barthès, B.G.; Chotte, J.-L.; Feller, C. Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma* **2007**, *139*, 106–117. [[CrossRef](#)]
50. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
51. Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [[CrossRef](#)] [[PubMed](#)]
52. Malek, S.; Melgani, F.; Bazi, Y. One-dimensional convolutional neural networks for spectroscopic signal regression. *J. Chemom.* **2018**, *32*, e2977. [[CrossRef](#)]
53. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *2*, 1097–1105. [[CrossRef](#)]
54. Sánchez-Esteva, S.; Knadel, M.; Kucheryavskiy, S.; de Jonge, L.W.; Rubæk, G.H.; Hermansen, C.; Heckrath, G. Combining Laser-Induced Breakdown Spectroscopy (LIBS) and Visible Near-Infrared Spectroscopy (Vis-NIRS) for Soil Phosphorus Determination. *Sensors* **2020**, *20*, 5419. [[CrossRef](#)]
55. R Core Team. *R: A Language and Environment for Statistical Computing*; The R Foundation: Vienna, Austria, 2020.
56. Mevik, B.-H.; Wehrens, R.; Liland, K.H. pls: Partial Least Squares and Principal Component Regression. *J. Stat. Soft.* **2019**, *18*. [[CrossRef](#)]
57. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.
58. Van Rossum, G.; Drake, F.L. *Python 3 Reference Manual*; ACM: Scotts Valley, CA, USA, 2009; ISBN 1441412697.

59. Chollet, F. Keras 2015. Available online: <https://github.com/keras-team/keras> (accessed on 5 March 2021).
60. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv* **2016**, arXiv:1603.04467.
61. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.M.; McBratney, A. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends Anal. Chem.* **2010**, *29*, 1073–1081. [[CrossRef](#)]
62. Nawar, S.; Mouazen, A.M. Predictive performance of mobile vis-near infrared spectroscopy for key soil properties at different geographical scales by using spiking and data mining techniques. *Catena* **2017**, *151*, 118–129. [[CrossRef](#)]
63. Ng, W.; Minasny, B.; de Mendes, W.S.; Demattê, J.A.M. Estimation of effective calibration sample size using visible near infrared spectroscopy: Deep learning vs machine learning. *SOIL Discuss.* **2019**, 1–21. [[CrossRef](#)]
64. Ben-Dor, E.; Inbar, Y.; Chen, Y. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sens. Environ.* **1997**, *61*, 1–15. [[CrossRef](#)]
65. Bogrekci, I.; Lee, W.S. Spectral soil signatures and sensing phosphorus. *Biosyst. Eng.* **2005**, *92*, 527–533. [[CrossRef](#)]
66. Mouazen, A.M.; Kuang, B. On-line visible and near infrared spectroscopy for in-field phosphorous management. *Soil Tillage Res.* **2016**, *155*, 471–477. [[CrossRef](#)]
67. Maleki, M.R.; van Holm, L.; Ramon, H.; Merckx, R.; de Baerdemaeker, J.; Mouazen, A.M. Phosphorus Sensing for Fresh Soils using Visible and Near Infrared Spectroscopy. *Biosyst. Eng.* **2006**, *95*, 425–436. [[CrossRef](#)]
68. Pätzold, S.; Leenen, M.; Frizen, P.; Heggemann, T.; Wagner, P.; Rodionov, A. Predicting plant available phosphorus using infrared spectroscopy with consideration for future mobile sensing applications in precision farming. *Precis. Agric.* **2019**, *21*, 737–761. [[CrossRef](#)]
69. Sherman, D.M.; Waite, T.D. Electronic spectra of Fe³⁺ oxides and oxide hydroxides in the near IR to near UV. *Am. Mineral.* **1985**, *70*, 1262–1269.
70. Scheinost, A.C.; Chavernas, A.; Barrón, V.; Torrent, J. Use and limitations of second-derivative diffuse reflectance spectroscopy in the visible to near-infrared range to identify and quantify Fe oxide minerals in soils. *Clays Clay Miner.* **1998**, *46*, 528–536. [[CrossRef](#)]
71. Viscarra Rossel, R.A.; Fouad, Y.; Walter, C. Using a digital camera to measure soil organic carbon and iron contents. *Biosyst. Eng.* **2008**, *100*, 149–159. [[CrossRef](#)]
72. Daniel, K.W.; Tripathi, N.K.; Honda, K. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Aust. J. Soil Res.* **2003**, *41*, 47–59. [[CrossRef](#)]
73. Ramaroson, V.H.; Becquer, T.; Sá, S.O.; Razafimahatratra, H.; Delarivière, J.L.; Blavet, D.; Vendrame, P.R.S.; Rabeharisoa, L.; Rakotondrazafy, A.F.M. Mineralogical analysis of ferralitic soils in Madagascar using NIR spectroscopy. *Catena* **2018**, *168*, 102–109. [[CrossRef](#)]
74. Li, R.; Yin, B.; Cong, Y.; Du, Z. Simultaneous prediction of soil properties using multi_cnn model. *Sensors* **2020**, *20*, 6271. [[CrossRef](#)]



Soil phosphorus retention can predict responses of phosphorus uptake and yield of rice plants to P fertilizer application in flooded weathered soils in the central highlands of Madagascar

Tomohiro Nishigaki^{a,*}, Yasuhiro Tsujimoto^a, Tovohery Rakotoson^b, Michel Rabenarivo^b, Andry Andriamananjara^b, Hidetoshi Asai^a, Haja Bruce Andrianary^b, Hobimiarantsoa Rakotonindrina^b, Tantely Razafimbelo^b

^a Crop, Livestock and Environment Division, Japan International Research Center for Agricultural Sciences, 1-1 Owashi, Tsukuba, Ibaraki 305-8686, Japan

^b Laboratoire des Radio-Isotopes, Université d'Antananarivo, BP 3383, Route d'Andraisoro, 101 Antananarivo, Madagascar

ARTICLE INFO

Handling Editor: Andrew Margenot

Keywords:

Fertilizer management
Lowland rice
Nutrient omission trial
Sub-Saharan Africa

ABSTRACT

Efficient phosphorus (P) fertilizer management is highly required for resource-limited farmers to achieve higher yields in typical P-deficient farmlands in sub-Saharan Africa. Although soil P retention has been conventionally used to estimate the possible responses to P fertilizer applied to soils, its applicability has not been clearly confirmed at the field level. In this study, the applicability of P retention to predict the response of rice plants to P application was investigated using pot experiments with soils collected from various locations (N = 62) and on-farm experiments at two nutrient-poor sites (N = 38) in the central highlands of Madagascar. Further, we explored a simple prediction method of P retention for multiple location assessment in local farmers' fields using an alternative property of soils collected from rice fields in the same region (N = 213). P retention was negatively correlated with the increase in rice P uptake (ΔP_{uptake} , $r = -0.550$) in the pot experiment and with the increase in yield (ΔYield , $r = -0.697$) in the on-farm field experiment as a response to P application. Path analysis revealed that oxalate-extractable aluminum (Alox) content was the most important factor of P retention across all the soils, indicating that rice plants grown on soils with higher Alox and P retention are less sensitive to P application. Given its high correlation with P retention ($r = 0.642$) and its simplicity in measurement, we proposed the use of moisture content of air-dried soils (ω_{air}) as a parameter to predict soil P retention. It was further confirmed that ω_{air} had a significant negative correlation with ΔP_{uptake} based on the pot experiment across a wide range of soil statuses ($r = -0.518$). However, ω_{air} could not clearly explain ΔYield within a small range of soil properties in the field experiment. Overall, soil P retention can predict the responses of rice plants to P application in the typical P-deficient and low-yielding lowlands in the central highlands of Madagascar, and this simple evaluation technique using air-dried soil moisture content will be helpful for the assessment of multiple rice fields on a broad scale.

1. Introduction

Phosphorus (P) deficiency is widely found in tropical agroecosystems in sub-Saharan Africa (SSA) (Kihara et al., 2016; Nziguheba et al., 2016; Saito et al., 2019; Shehu et al., 2018). This is due to the inherent properties of tropical weathered soils, which are characterized by low available P content, low pH, and large amounts of P-fixing aluminum (Al) and iron (Fe) oxides (Kirk et al., 1998; Nishigaki et al., 2019;

Nziguheba et al., 2016). The crop response to P application is low in high P-fixing soils in the tropics; thus, large amounts of P have been applied to achieve high grain yields. Flooding condition in paddy soils can partly increase P availability for rice due to increase in P transport and reductive dissolution of ferric oxides (Zhang et al., 2003). However, the effect of flooding is limited due to the high P fixation potential but low P content of soils rich in crystalline oxides, e.g., gibbsite and goethite, which are typically observed in tropical agroecosystems (Batjes, 2011).

* Corresponding author.

E-mail addresses: nishigaki@affrc.go.jp (T. Nishigaki), tsjmt@affrc.go.jp (Y. Tsujimoto), tovohery.rakotoson@gmail.com (T. Rakotoson), miarabenarivo@yahoo.fr (M. Rabenarivo), njaraandry1@gmail.com (A. Andriamananjara), asai0817@affrc.go.jp (H. Asai), hajabruce@yahoo.fr (H.B. Andrianary), hobimiarantsoa@gmail.com (H. Rakotonindrina), tantely.razafimbelo@gmail.com (T. Razafimbelo).

<https://doi.org/10.1016/j.geoderma.2021.115326>

Received 20 October 2020; Received in revised form 25 June 2021; Accepted 26 June 2021

Available online 7 July 2021

0016-7061/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Furthermore, considering the finite nature of P fertilizer resources and limited access to chemical fertilizer by local farmers in SSA, improved P fertilizer management is highly required for sustainable crop production (Alewell et al., 2020; Nziguheba et al., 2016).

One practical approach to improve P fertilizer use efficiency is site-specific P fertilizer management based on a spatial assessment of soil P deficiency (Tsujiimoto et al., 2019). This approach should be particularly beneficial in low-input and nutrient-poor soils in the tropics where field-to-field variations in P deficiency status are large (Nishigaki et al., 2019). Relative biomass or nutrient uptake (ratio with the omission of a target nutrient relative to its full application) has been used in previous omission trials as an index to evaluate the spatial variation of indigenous nutrient deficiency and the responses of crops to fertilizer inputs (Haeffele and Wopereis, 2005; Kihara et al., 2016; Saito et al., 2019; Shehu et al., 2018). Conversely, soil testing by chemical extraction is the most common method for assessing the amount of readily extractable P in soils and determining the need for P additions to meet crop needs in field trials (Jordan-Meille et al., 2012; Nawara et al., 2017; Rosen et al., 2014). However, readily extractable P contents determined in soil tests alone are often poor predictors of yield response or plant P uptake response to P application in both uplands and lowlands (Dobermann et al., 2003; Nawara et al., 2017). This discrepancy between the results of the soil P test and the crop responses to P application could be partly explained by the fact that the response of crops to P application is associated not only with readily extractable P contents in soils but also with soil P-sorption capacity (Andriamananjara et al., 2016; Nziguheba et al., 2016). Nevertheless, few attempts have been made to reveal the effect of soil P retention and crop response to fertilizer input.

It was reported that intensity-based P tests (e.g., CaCl₂ extraction or diffusive gradient in thin-films test) were superior to quantity-based P tests (Olsen or Oxalate) for predicting the relative yield of crops in typical tropical soils with low P availability and large P sorption (Nawara et al., 2017). This was supported by the strong negative correlation between intensity-based P and the P sorption index of tropical soils (Six et al., 2013) and by the negative relationship between Alox + 0.5 Feox, but not the available P content, and fertilizer-P recovery efficiency of rice plants grown in the P-deficient soils of Madagascar rice fields (Nishigaki et al., 2019; Oo et al., 2020). These results suggest that P retention or its relevant P sorbents could be a more important factor than conventional soil P tests to determine rice responses to P application in tropical P-deficient soils. Therefore, we hypothesized that the P retention capacity of soils is a reliable parameter to predict crop responses to P fertilizer application for tropical lowland rice production. P retention is a simple measurable parameter that has been conventionally used to estimate the possible responses to P application of soils at a broad scale (Johnston et al., 2014). However, in contrast to soil P tests, the applicability of P retention for predicting the crop response to P application has not been well studied, particularly at the field level.

The target area of this study was the central highlands of Madagascar, which is mostly covered by soils categorized with very high to high P retention due to their low pH and weathered soils represented by Ferralsols (Andriamananjara et al., 2016; Batjes, 2011; Nishigaki et al., 2020). The optimal application rate of P fertilizer based on the evaluation of soil P retention may be more economical in the tropics than in temperate regions due to the predominance of such weathering soils. This study aimed to assess the applicability of soil P retention to predict the response of rice to P fertilizer application in flooded weathered soils. In addition, we explored a simple prediction method of soil P retention using correlation analysis with various soil properties.

2. Materials and methods

2.1. Soil sampling and analysis

Soil samples were collected from 213 widely distributed lowland rice fields in the central highlands of Madagascar. The fields containing the

sampling points were located at an altitude of 900–1600 m a.s.l. The area is mostly covered by Ferralsols according to the national soil map (EU, 2013), while the soils in the rice fields are most likely to be classified as Fluvisols or Cambisols based on previous field observations (Nishigaki et al., 2020).

In each field, a surface layer of 0–15 cm depth was collected as four to five composites using a core sampler and a shovel for soil analysis and the pot experiments, respectively. The collected soil samples were air-dried, sieved to pass through a 2-mm mesh screen, and kept in plastic bags until the analysis of P retention and relevant soil properties.

Soil P retention was measured using the method described by Soil Survey Staff (2014). Briefly, 5.0 g of soil was shaken in a 25-mL aliquot of a 1000 mg L⁻¹P solution, which was a mixture of KH₂PO₄, CH₃COONa, and CH₃COOH with pH 4.6, for 24 h. The aliquot was centrifuged and filtered, and the P concentration in the aliquot was quantified by a colorimetric method with nitric vanadomolybdate acid reagent at an absorbance of 466 nm. This concentration correlated with the concentration of the non-adsorbed P remaining in the sample solution. The P retention was then determined as the initial P concentration minus the P remaining in the sample solution and was shown as percent P retained. Soil particle size distribution was determined using the wet-sieving and pipet method (Gee and Bauder, 1986). Soil pH was determined in deionized water at a soil-to-solution ratio of 1:2.5. Exchangeable cations (K⁺, Na⁺, Ca²⁺, and Mg²⁺) were measured following Soil Survey Staff (2014), and the total exchangeable cation (TEC) was calculated as the sum of these four cations. Cation exchange capacity (CEC) was determined by the ammonium acetate method at pH 7.0 (Soil Survey Staff, 2014). Base saturation was calculated as the percentage of TEC to CEC (TEC/CEC). Total carbon (TC) and nitrogen (TN) were quantified using the dry combustion method with an NC analyzer (Sumigraph NC-220F; Sumika Chemical Analysis Service, Ltd., Osaka, Japan). Extractable Al, Fe, and P contents were determined using the acid ammonium oxalate method (Alox, Feox, and Pox, respectively) as described by Courchesne and Turmel (2008). The concentrations of Al, Fe, and P in the oxalate extraction were measured using an inductively coupled plasma mass spectrometer (ICPE-9000; Shimadzu Corp., Kyoto, Japan).

The air-dried soil moisture content (ω_{air}) was measured as follows: 2.0 g of soil was fully wet by ~ 2 mL of deionized water and kept for 12 h at room temperature. Then, the wet soil was dried at 60 °C for 72 h in an oven to a constant weight, and the weight of the air-dried soil was measured (A). The soil was subsequently dried at 105 °C for 24 h in an oven, and then the weight of the oven-dried soil was measured (B). Finally, ω_{air} (%) was calculated as follows:

$$\omega_{\text{air}} = \frac{A - B}{B} \times 100$$

2.2. Pot experiment

Out of the 213 soil samples for soil analysis mentioned above, 62 soil samples were selected to evenly represent a wide range of soil status in the central highlands of Madagascar, to minimize the potential biasing of the samples from being skewed toward low fertility. The soil samples were used for three sets of pot experiments (Exps. 1–3) to analyze the effect of P application on rice P uptake under flooded conditions. The soil samples for the pot experiments were air-dried and sieved (< 4 mm), and 1 kg of each soil was put into a 1-L plastic pot (13 cm diameter, 15 cm height). The pot experiments were all conducted in a greenhouse at the Laboratoire des Radio-Isotopes, University of Antananarivo, Madagascar (18° 53' 57" S, 47° 33' 01" E, 1,222 m alt.) from September to October 2016 (Exp. 1, N = 24), October to November 2018 (Exp. 2, N = 26), and September to October 2019 (Exp. 3, N = 12). The daily mean temperatures throughout the growing periods ranged 18–26 °C, 20–29 °C, and 15–26 °C in Exps. 1–3, respectively (Watchdog 2475, Spectrum Technologies Ltd.). The data from Exp. 1 was presented in our

previous study (Nishigaki et al., 2019). In each experiment, three fertilizer treatments were established with two replicates: 1, *Cont.* (no fertilizer application); 2, +N (0.2 g pot⁻¹ of N as NH₄NO₃); and 3, +NP (0.2 g pot⁻¹ of both N as NH₄NO₃ and P₂O₅ as KH₂PO₄). The application rate of N was adequate in the flooded condition despite the form of NH₄NO₃ since clear positive response in plant biomass was observed for some pots under the +N treatment compared with the *Cont.* treatment (Supplemental table). Potassium was applied to all the pots, including the *Cont.* treatment, at a rate of 0.2 g pot⁻¹ of K₂O as KCl. Each nutrient was uniformly incorporated into the soils one day prior to transplanting. Then, two 20-day-old seedlings of a local rice cultivar, X265, grown in free-nutrient sand were transplanted into each pot. The pots were continuously flooded with distilled water throughout the growing period. Weeds were manually removed. No specific pest management was required.

The plant samples were harvested at the soil surface 33–41 days after transplanting for Exps. 1–3. Aboveground biomass was determined after oven-drying at 70 °C to a constant weight. Each plant sample was ground into a fine powder using a high-speed vibrating sample mill (Model T1-100; Heiko Co. Ltd., Fukushima, Japan). Then, the plant P concentration was determined using the molybdate blue method (Murphy and Riley, 1962) after dry-ashing at 550 °C for 2 h and digestion with 0.5 M HCl. The plant P uptake (mg P pot⁻¹) was calculated as the product of the aboveground biomass and P concentration of plants (total P content).

2.3. Field experiment

We conducted field trials to assess whether P retention can successfully estimate the response of rice plants to P application under on-farm flooded conditions. The on-farm trials were conducted in a total of 38 lowlands, which were selected from the soils for the pot experiment, in the communes of Antohobe (19° 46' S, 46° 41' E, 1270 m asl.) and Behenjy (19° 13' S, 47° 28' E, 1380 m asl.) during the wet seasons of 2017–2018 and 2018–2019. Due to the higher elevation in Behenjy, the air temperature is generally lower in Behenjy (20.8 °C on average during the cropping season from November 2018 to April 2019) than that in Antohobe (22.2 °C during the same period). In the same manner as in the pot experiments, three fertilizer treatments were allocated with two to three replicates in each field: 1, *Cont.* (no fertilizer application); 2, +N (80 kg ha⁻¹ of N as urea with 50 kg N ha⁻¹ broadcasted at the period of transplanting and 30 kg N ha⁻¹ top-dressed at the pre-heading period); and 3, +NP (50 kg ha⁻¹ of P as triple super phosphate at the period of transplanting along with the +N treatment). No fertilizers were applied other than N and P, which are the most limiting nutrients for rice in SSA (Saito et al., 2019). The plot size ranged from 5.1 m² to 8.0 m².

The variety X265 was commonly used in all the fields. Either farmers' transplanting patterns or random transplanting with a density of 23.6 to 67.0 hills m⁻² was adopted in 28 fields. In the remaining 10 fields, two seedlings per hill were transplanted at a density of either 25 hills m⁻² (0.20 m × 0.20 m) or 50 hills m⁻² (0.20 m × 0.10 m). Weeds and water were adequately controlled throughout the rice growing periods. No marked damage from pests or disease was observed at any site. At maturity, grain yields (t ha⁻¹) were determined by harvesting plants from an area of 2.6 m² to 8.0 m² in the center of each plot. Grain yield was expressed based on filled grain weight and corrected to 14% moisture content using a grain moisture sensor (Riceter F; Kett Electric Laboratory, Tokyo, Japan).

2.4. Statistical analyses

Statistical analyses were performed using Excel (Microsoft Office 365 ProPlus; Microsoft Corp., Redmond, WA, USA) and JMP14 software (SAS Institute Inc., Cary, NC, USA). Forward stepwise regression analysis was performed for P retention and ω_{air} using all the other soil properties as explanatory variables with the selection level at p < 0.005.

Then, a multiple linear regression model was developed using the selected variables. Path analysis was performed to examine the causal path of the soil properties, which were selected in the regression analysis, to P retention and ω_{air}, separately. The direct and indirect effects in the path analysis were derived from (i) multiple linear regression of soil properties on P retention or ω_{air} and (ii) simple correlation coefficients between soil properties. The direct effects of the soil properties on P retention or ω_{air} were termed path coefficients (P) and were standardized partial regression coefficients for each of the soil properties in the multiple linear regression against P retention or ω_{air} (Williams et al., 1990). The indirect effects of soil properties on P retention or ω_{air} were determined from the product of the simple correlation coefficient between the soil properties and the path coefficient. The correlation between the P retention/ω_{air} and each soil property was the sum of the entire path connecting the two variables. Path analysis results can be presented in a concise table that consists of a matrix with the main diagonal representing direct effects and off-diagonal elements representing indirect effects (Williams et al., 1990).

The plant P uptake data of each pot experiment (Exps. 1–3) were normalized to the sample variation from 0 to 1 to compare the results of the three pot experiments grown under different environmental conditions. Then, the P uptake response of rice plants to P application was expressed as the difference in normalized P uptake between the +NP and +N treatments (ΔP_{uptake}) in the pot experiment. Similarly, the response of rice yield to P application was expressed as the difference in yield, which was not normalized due to the limited sample number of each experiment, between the +NP and +N treatments (ΔYield) in the field experiment. A Pearson correlation analysis was performed among the soil properties and between ΔP_{uptake} and the soil properties at p < 0.01. After a significant correlation was confirmed, a linear regression was conducted between P retention and ΔP_{uptake}. Fields with grain yields of > 3.0 t ha⁻¹ in the *Cont.* plots (n = 7) were excluded from the correlation analysis to avoid high-yielding fields causing little response of rice to P fertilizer application in the field experiment. A Pearson correlation analysis was also performed between ΔYield and P retention. After a significant correlation was confirmed, samples were divided by sites into the Behenjy site (N = 10) and Antohobe site (N = 22).

3. Results

3.1. Soil characteristics

The soil properties of all soils and sub-groups for the pot and field experiments are summarized in Table 1. Since soil samples were collected from a range of environments, the sample set of all soils had a variety of soil characteristics. P retention had a large variation, ranging from 0.02% to 85.3%, in all the soil samples. TEC and Alox also showed large variations among the samples. In contrast, soil pH had little variation, and was slightly acidic with a mean value of 5.58 for all the soils. The mean values and coefficient of variation (CV) of soil properties were generally lower in the soils in the field experiment than those in the pot experiment (Table 1); the sites for the field experiment had relatively low fertility and little variation.

The correlation matrix of the soil properties is shown in Table 2. The P retention was poorly correlated with the soil pH and silt content. As a single variable, P retention had the highest correlation with Alox. Stepwise regression analysis indicated that 72.9% of the variability of P retention was explained by Alox, clay, Feox, and TEC/CEC (Table 3). The path coefficient (P, underlined in Table 3) indicated that Alox ($\underline{P} = 0.646$) was the most important direct causal factor in predicting P retention. An examination of the indirect effects revealed that the correlations of P retention with Feox, and TEC/CEC were mostly partitioned to the indirect effect of Alox.

ω_{air} was highly correlated with P retention for all the soil samples (Table 2). However, the correlation became weak in soil samples with low P retention (<50%) (r = 0.250, N = 197), while it was strong in soil

Table 1
Mean values of the soil properties of all soils and sub-groups for the pot and field experiments.

		All soils (N = 213)		Soils for pot experiment (N = 62)		Soils for field experiment (N = 38)	
TEC	($\text{cmol}_c \text{ kg}^{-1}$)	2.75	(0.96)	3.45	(1.15)	2.12	(0.58)
CEC	($\text{cmol}_c \text{ kg}^{-1}$)	11.8	(0.41)	12.0	(0.50)	11.0	(0.48)
CEC/Clay	($\text{cmol}_c \text{ kg}^{-1}$)	36.1	(0.34)	33.6	(0.45)	33.1	(0.47)
TEC/CEC	(%)	22.9	(0.64)	26.2	(0.65)	20.8	(0.59)
TC	(%)	1.77	(0.43)	1.97	(0.46)	1.70	(0.31)
TN	(%)	0.14	(0.37)	0.16	(0.41)	0.14	(0.26)
C/N		12.3	(0.09)	12.4	(0.08)	12.2	(0.07)
pH (H ₂ O)		5.58	(0.05)	5.60	(0.06)	5.60	(0.04)
Alox	(g Al kg ⁻¹)	1.56	(0.75)	2.15	(0.86)	1.42	(0.53)
Feox	(g Fe kg ⁻¹)	6.36	(0.47)	6.84	(0.56)	6.26	(0.54)
Pox	(mg P kg ⁻¹)	102.3	(1.50)	187.4	(1.39)	79.8	(0.33)
ω_{air}	(%)	0.40	(0.44)	0.46	(0.63)	0.33	(0.26)
P retention	(%)	25.0	(0.66)	39.3	(0.46)	31.3	(0.40)
Clay	(%)	32.1	(0.22)	33.6	(0.23)	33.9	(0.19)
Silt	(%)	25.8	(0.33)	22.6	(0.36)	20.6	(0.30)
Sand	(%)	41.7	(0.26)	42.5	(0.25)	45.5	(0.18)

The values given in parentheses are the coefficients of variation. TEC, total exchangeable cations (sum of exchangeable Ca²⁺, Mg²⁺, Na⁺, and K⁺); CEC, cation exchangeable capacity; TC, total carbon; TN, total nitrogen; ω_{air} , air-dried soil moisture content; Alox, oxalate-extractable Al; Feox, oxalate-extractable Fe; Pox, oxalate-extractable P.

samples with high P retention ($r = 0.831$, $N = 16$). Stepwise regression analysis indicated that 71.2% of the variability of ω_{air} was explained by Alox, CEC, and Pox (Table 4). The path coefficient indicated that Alox ($P = 0.404$) was the most important direct causal factor in predicting ω_{air} . An examination of the indirect effects revealed that the correlations of ω_{air} with Pox and CEC were mostly partitioned to the indirect effect of Alox.

3.2. Pot experiment

The P uptake of rice plants in the + N treatment ranged from 0.05 to 2.46 mg P pot⁻¹ (CV = 90.1%), 0.11 to 1.05 mg P pot⁻¹ (CV = 69.9%), and 0.11 to 1.08 mg P pot⁻¹ (CV = 105.3%) in Exps. 1–3, respectively (Supplemental table). The ΔP_{uptake} ranged from 0.10 to 0.91 (CV = 44.2%), 0.37 to 0.98 (CV = 25.4%), and 0.33 to 1.00 (CV = 27.2%) in Exps. 1–3, respectively.

As a single variable, ΔP_{uptake} had the highest correlation with P retention (Table 5, Fig. 1) and had significant correlations with the other soil properties, except CEC/Clay, TEC/CEC, C/N, pH, and clay. As a single variable, P retention explained 30.2% of the variance of ΔP_{uptake} ($p < 0.001$). P retention was weakly correlated with rice P uptake in the Cont. plot ($r = 0.37$, $p = 0.003$) and the + N plot ($r = 0.43$, $p < 0.001$).

3.3. Field experiment

The crop yield in the Cont. plot ranged from 0.47 to 4.58 t ha⁻¹ (Supplemental table). The ΔYield ranged from - 0.50 to 2.07 t ha⁻¹. High-yielding fields (>3.0 t ha⁻¹ in the Cont. plots) were excluded from the following analysis to avoid high soil fertility causing little response of rice to P fertilizer application. The ΔYield was significantly negatively correlated with soil P retention at both sites (Fig. 2). The correlation coefficient between P retention and ΔYield was - 0.697 for all fields except the high-yielding ones. The average temperature at the Behenjy site is generally lower than that at the Antohobe site because of the higher elevation; the mean temperature during the cropping season from November 2018 to April 2019 was 20.8 °C at Behenjy and 22.2 °C at Antohobe. The slope of the regression model was steeper at the Antohobe site than that at the Behenjy site, showing the higher ΔYield at the Behenjy site than that at the Antohobe site at the same P retention. ΔYield reduced by 35.3 kg ha⁻¹ and 62.2 kg ha⁻¹ per unit increase in P retention at Behenjy and Antohobe, respectively. The value of P retention when ΔYield was 0 (no effect of P fertilizer) was greater at Behenjy (54.0%) than that at Antohobe (35.2%). Meanwhile, the ΔYield was not correlated with ω_{air} for either of the two sites.

4. Discussion

4.1. Response of P uptake and yield of rice plants to P application as affected by soil P retention

It is advantageous for farmers to know the potential response of rice growth prior to the application of fertilizer. Although the importance of soil P retention for crop responses to P fertilizer was addressed in a previous study (Nziguheba et al., 2016), to the best of our knowledge, there have been no studies on the quantitative relationship between soil P retention capacity and grain yields at a field scale. It is noteworthy that increasing soil P retention resulted in a decrease in the yield response of rice to the application of P fertilizer in the typical P-deficient and low-yielding lowlands in the central highlands of Madagascar (Fig. 2). The results of the pot experiment using a wide range of soils corroborated those of the field experiments (Fig. 1). Therefore, it is suggested that P retention can be a promising index for estimating the variability of the responses of rice plants to mineral P fertilizer application in flooded weathered soils.

The variation in the response of rice P uptake to P application has been generally explained by the combination of available P content and other influential soil parameters involved in the P sorption process, such as soil pH, clay content, and active Al (Andriamananjara et al., 2016; Batjes, 2011). Thus, it is often difficult to extrapolate the interpretation from one site to another. However, the current study showed that P retention is a significant parameter for predicting rice responses to P fertilizer application in tropical rice fields at both broad (pot experiment) and commune scales (field experiment) in the central highlands of Madagascar (Figs. 1 and 2). This was supported by previous observations that intensity-based P tests, which are negatively correlated with P sorption, were superior at predicting the relative yield of crops in tropical soils (Nawara et al., 2017; Six et al., 2013). Nwoke et al. (2003) also suggested that Alox and Feox can regulate the standard P requirement and P availability in soils of the West African savanna. Therefore, P retention can be used as a comprehensive index to predict the response of rice P uptake and yield to P application in P-deficient tropical soils, as it represents the relevant properties involved in soil P supply capacity and the P sorption process.

The site differences in the yield response to P application can be ascribed to the higher elevation and lower temperature conditions in Behenjy than those in Antohobe. P deficiency is known to delay phenological development and thus prolong the days to heading and maturity (Dobermann and Fairhurst, 2000). Our recent field experiment revealed that the effect of P application on rice grain yield is greater in

Table 2
Correlation matrix among the soil properties of all samples.

	TEC	CEC	CEC/Clay	TEC/CEC	TC	TN	C/N	pH (H ₂ O)	Alox	Feox	Pox	ω _{air}	P retention	Clay	Silt
CEC	0.511														
CEC/Clay	0.418	0.850													
TEC/CEC	0.819	0.052	-0.069												
TC	0.187	0.529	0.291	-0.054											
TN	0.191	0.528	0.290	-0.047	0.986										
C/N	0.152	0.377	0.207	-0.039	0.655	0.533									
pH (H ₂ O)	0.436	0.165	0.116	0.376	-0.252	-0.237	-0.199								
Alox	0.557	0.521	0.277	0.205	0.626	0.605	0.446	0.108							
Feox	0.472	0.505	0.327	0.253	0.461	0.501	0.191	0.223	0.331						
Pox	0.808	0.418	0.216	0.490	0.213	0.233	0.081	0.357	0.749	0.445					
ω _{air}	0.644	0.638	0.468	0.239	0.529	0.537	0.283	0.245	0.826	0.473	0.787				
P retention	0.517	0.498	0.175	0.267	0.688	0.690	0.450	0.078	0.781	0.517	0.615	0.642			
Clay	-0.150	0.311	-0.048	-0.294	0.551	0.563	0.343	-0.247	0.158	0.233	-0.137	0.105	0.387		
Silt	0.261	0.387	0.443	0.068	0.102	0.099	0.087	0.213	0.233	0.113	0.161	0.303	0.098	-0.055	
Sand	-0.179	-0.543	-0.344	0.092	-0.480	-0.491	-0.294	-0.022	-0.370	-0.278	-0.135	-0.396	-0.405	-0.596	-0.757

Bold values indicate a significant correlation at $p < 0.01$.
TEC, total exchangeable cations (sum of exchangeable Ca²⁺, Mg²⁺, Na⁺, and K⁺); CEC, cation exchangeable capacity; TC, total carbon; TN, total nitrogen; ω_{air}, air-dried soil moisture content; Alox, oxalate-extractable Al; Feox, oxalate-extractable Fe; Pox, oxalate-extractable P.

Table 3

Direct effects (diagonal, underlined) and indirect effects (off-diagonal) of soil properties on P retention.

	Alox	Clay	Feox	TEC/CEC	r	R ²
Alox	<u>0.646</u>	0.048	0.061	0.036	0.781	0.729
Clay	0.102	<u>0.305</u>	0.043	-0.051	0.387	
Feox	0.214	0.071	<u>0.184</u>	0.044	0.517	
TEC/CEC	0.133	-0.090	0.047	<u>0.173</u>	0.267	

r shows the Pearson correlation coefficient between P retention and each parameter.

Alox, oxalate-extractable Al; Feox, oxalate-extractable Fe; TEC, total exchangeable cations (sum of exchangeable Ca²⁺, Mg²⁺, Na⁺, and K⁺); CEC, cation exchangeable capacity.

Table 4

Direct effects (diagonal, underlined) and indirect effects (off-diagonal) of soil properties on air-dried soil moisture content.

	Alox	CEC	Pox	r	R ²
Alox	<u>0.404</u>	0.144	0.300	0.826	0.712
CEC	0.210	<u>0.277</u>	0.168	0.638	
Pox	0.303	0.116	<u>0.401</u>	0.787	

r shows the Pearson correlation coefficient between the air-dried soil moisture content and each parameter.

Alox, oxalate-extractable Al; CEC, cation exchangeable capacity; Pox, oxalate-extractable P.

high-elevation and late-transplanted sites because P application alleviates not only P deficiency stress but also cold-induced sterility later in the growth stage by shortening growth durations (Rakotoarisoa et al., 2020). On the other hand, if the risk of lowering temperature is not a problem, the prolonged growth period under P deficiency favors for additional biomass production and eventually nullifies the difference in rice growth between the plots with and without P application. In this way, rice yield was more responsive to P application in Behenjy than Antohobe at a same level of P retention. These results imply that efficient site-specific P fertilizer management should be dependent on both static soil P status and climatic growth conditions. It should be noted that this assessment is most likely not applicable to relatively high-yielding fields where rice plants display little response to P application irrespective of P retention capacity of soils, which is probably because P availability is not a main limiting nutrient for rice growth. However, the assessment of P retention should be a useful approach to increase the return from minimal fertilizer inputs in low-input and low-yielding fields, such as those in many lowlands in Madagascar and likely in other areas with similar rice cropping systems (Saito et al., 2019; Tsujimoto et al., 2019).

4.2. Air-dried soil moisture content as a potential index of P retention

To apply the findings in the present study in an economically feasible manner, a simplified prediction of P retention without chemical lab analysis is ideal (Dobermann et al., 2003). Given its high correlation with P retention and its simplicity for measurement, we selected air-dried soil moisture content (ω_{air}) as a parameter for predicting P retention. Recently, Kinoshita and Tani (2020) proposed a simple evaluation method for P retention using air-dried soil moisture content for Japanese cropland soils.

The applicability of ω_{air} was further tested to predict the response of rice plants to fertilizer application in the current study. Our pot experiment showed a significant correlation between ω_{air} and ΔP_{uptake} under flooded conditions (Fig. 1). However, we found that ω_{air} could not explain the variation in yield response to the P application in the field experiment. Compared with that in the soils in the pot experiment, the sites of the field experiments had low P retention (mostly < 50%) with narrow variation, and P retention and ω_{air} were weakly correlated. These

Table 5
Correlation coefficients between the response of rice P uptake to P application (ΔP_{uptake}) and the soil properties.

	P retention	ω_{air}	TEC	CEC	CEC/Clay	TEC/CEC	TC	TN	C/N	pH (H ₂ O)	Alox	Feox	Pox	Clay	Silt	Sand
ΔP_{uptake}	-0.550	-0.52	-0.388	-0.448	-0.256	-0.209	-0.436	-0.455	-0.106	-0.183	-0.519	-0.396	-0.450	-0.106	-0.401	0.473

Bold values indicate a significant correlation at $p < 0.01$.
TEC, total exchangeable cations (sum of exchangeable Ca²⁺, Mg²⁺, Na⁺, and K⁺); CEC, cation exchangeable capacity; TC, total carbon; TN, total nitrogen; ω_{air} , air-dried soil moisture content; Alox, oxalate-extractable Al; Feox, oxalate-extractable Fe; Pox, oxalate-extractable P.

results might affect the poor estimation of the response of rice plants to P fertilizer using ω_{air} within a small variation in the communes with low soil fertility. Hence, the current estimation model of P retention by ω_{air} should be used to test P retention across various soil types at a regional scale.

We explored the controlling factors that allow for the fine-tuning of the estimation model of P retention by ω_{air} for localized farmlands. Soil type was previously reported to be an influential factor in the relationship between P retention and ω_{air} (Kinoshita and Tani, 2020). In the case of the relationship between CEC and ω_{air} , Churchman and Burke (1991) reported that the determination coefficient was usually increased by separating soil groups according to the dominant clay mineral types ($R^2 = 0.34\text{--}0.88$) rather than using the whole sample set in New Zealand ($N = 1318$, $R^2 = 0.38$). However, in our study, dividing the samples by either soil type or clay mineralogy (CEC/Clay) could not significantly increase the determination coefficient of the estimation model of P retention by ω_{air} . Further studies should be conducted on the determinant factor of the site-specific model to improve the evaluation method of P retention using ω_{air} to apply to field-to-field measurements within a commune level.

4.3. Influential factors of P retention and air-dried soil moisture content

A mechanistic understanding of how soil P retention is related to rice responses to P application is important to further improve and apply the estimation model in wider agroecosystems. Our results showed that Alox was the most important direct causal factor of P retention in soils collected from rice fields in the central highlands of Madagascar (Table 3). The preferential P retention of Al minerals over Fe minerals was reported by Khare et al. (2005) and Nwoke et al. (2003), which is in agreement with our path analysis results. Hashimoto et al. (2012) reported that Alox had the most important direct effect on P retention capacity in allophanic Andisols using path analysis. In agreement with these previous reports and our soil data, the response of rice P uptake to fertilizer application (ΔP_{uptake}) also had a higher correlation with Alox than Feox in flooding condition (Table 5). Clay was also selected as a significant parameter for P retention, which was not partitioned to Alox (Table 3). This implies the possible contribution of clay minerals, such as kaolinite, to P sorption in weathered ferralitic soils (Gérard, 2016). Unexpectedly, base saturation (TEC/CEC) was found to be a direct causal factor of P retention (Table 3), indicating that the affinity for phosphate by soil was possibly enhanced by increasing base saturation. Amery and Smolders (2012) found that increasing soil CEC increased P solubility in tropical flooded soils, probably due to the decrease in Ca²⁺ and Fe²⁺ concentrations. However, the interaction between P retention and soil nutrient status is still unclear compared with that between P retention and Alox, and it requires further investigation. Overall, P retention is mainly controlled by active Al, clay (kaolinite), and base saturation in soils of lowland rice fields in the central highlands of Madagascar. To evaluate the applicability of this finding to lowland rice fields in different regions in SSA, further test is required with special reference to Alox and soil pH which had small variations in our sample set.

The air-dried soil moisture content (ω_{air}) was mainly regulated by Alox (Table 4). Water vapor adsorption used to be a measure of specific surface area in soils (Newman, 1983). The high correlation between ω_{air} and Alox was likely due to increasing volume of micro- and small mesopores (0.7–4 nm) with increasing Al oxides/hydroxides (Watanabe, 2017). CEC was also selected as a controlling factor of ω_{air} . This is also in agreement with previous studies which reported a close relationship between CEC and soil hygroscopic moisture (Churchman and Burke, 1991; Torrent et al., 2015), although CEC was mostly attributed to the indirect effect of Alox in our study. Given their high correlation, ω_{air} could be used to CEC in nutrient-poor soils in tropical lowland rice fields.

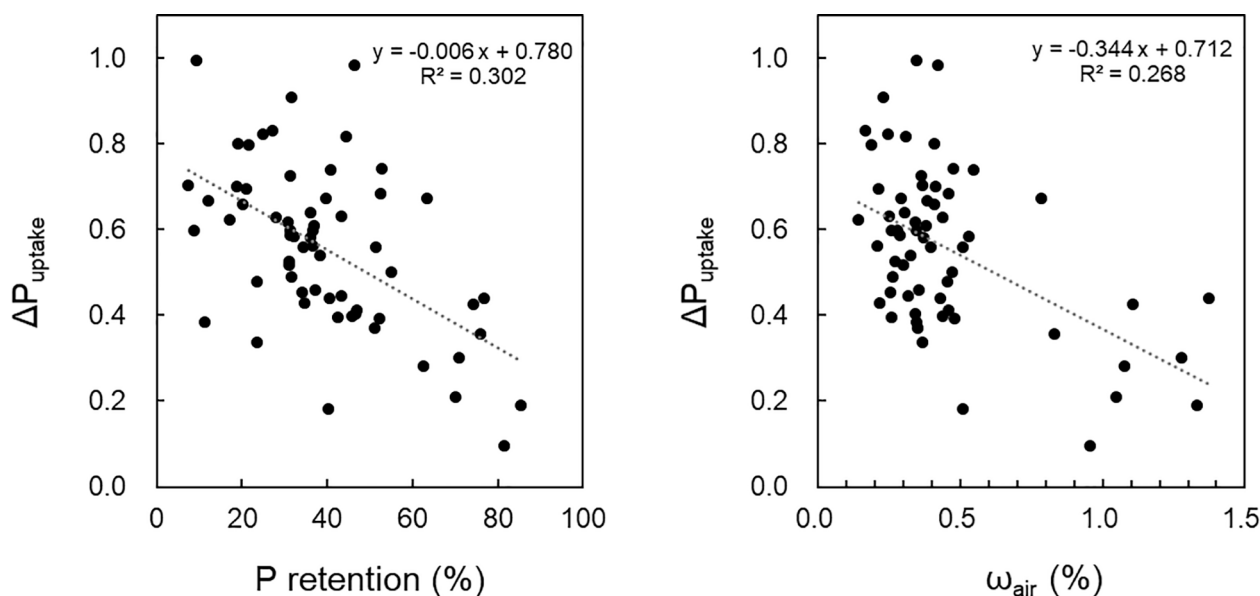


Fig. 1. Relationship between the response of rice phosphorus (P) uptake to P application (ΔP_{uptake}) and soil P retention (left) or air-dried soil moisture content (ω_{air}) in the pot experiment.

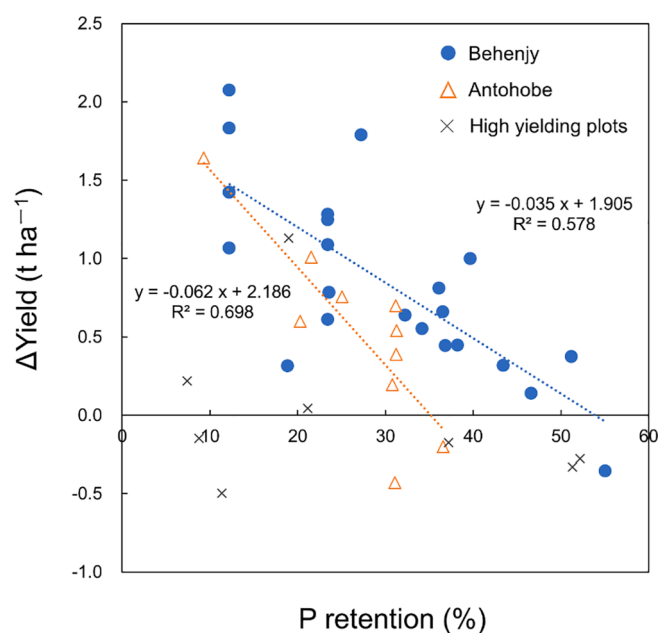


Fig. 2. Relationship between the response of rice yield to phosphorus (P) application (ΔYield) and soil P retention in the field experiment.

5. Conclusion

Soil P retention was shown to be a promising index for estimating the variability of responses of flooded rice plants to P application in the typical P-deficient and low-yielding lowlands in the central highlands of Madagascar. Additionally, the moisture content of air-dried soils was found to be a simple evaluation tool to predict soil P retention. Our findings will be useful to facilitate the use of evaluations of the P requirement of farmers' fields toward site-specific P fertilizer management and improvement of fertilizer-P use efficiency.

Declaration of Competing Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to Ms. Mayumi Yonemura for assisting with the laboratory analysis. We also thank Dr. Naoki Moritsuka, Kochi University, for kindly reviewing a draft of this paper and providing valuable comments. We would like to thank the Ministry of Mines and Strategic Resources of Madagascar, regional officers, and farmers for their kind support for our soil sampling. This research was financially supported by the Japan Prize Foundation and the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA) [Grant No. JPMJSA1608]. The soil samples were legally imported to Japan for soil analysis with permission from the Ministry of Agriculture, Forestry and Fisheries of Japan [30Y1175].

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.geoderma.2021.115326>.

References

- Alewell, C., Ringeval, B., Ballabio, C., Robinson, D.A., Panagos, P., Borrelli, P., 2020. Global phosphorus shortage will be aggravated by soil erosion. *Nat. Commun.* 11, 4546.
- Amery, F., Smolders, E., 2012. Unlocking fixed soil phosphorus upon waterlogging can be promoted by increasing soil cation exchange capacity. *Eur. J. Soil Sci.* 63 (6), 831–838.
- Andriamananjara, A., Rakotoson, T., Razanakoto, O.R., Razafimanantsoa, M.-P., Rabeharisoa, L., Smolders, E., 2016. Farmyard manure application has little effect on yield or phosphorus supply to irrigated rice growing on highly weathered soils. *Field Crops Research* 198, 61–69.
- Batjes, N.H., 2011. Global distribution of soil phosphorus retention potential. Wageningen, ISRIC.
- Churchman, G.J., Burke, C.M., 1991. Properties of sub soils in relation to various measures of surface area and water content. *J. Soil Sci.* 42 (3), 463–478.
- Courchesne, F., Turmel, M.-C., 2008. Extractable Al, Fe, Mn, and Si. In: Carter, M.R., Gregorich, E.G. (Eds.), *Soil sampling and methods of analysis*, 2nd edn. CRC Press, Boca Raton, pp. 307–316.
- Dobermann, A., Fairhurst, T.H., 2000. Rice: Nutrient disorders and nutrient management. International Rice Research Center (IRRI). Los Baños, Philippines.
- Dobermann, A., Witt, C., Abdulrachman, S., Gines, H.C., Nagarajan, R., Son, T.T., Tan, P. S., Wang, G.H., Chien, N.V., Thoa, V.T.K., Phung, C.V., Stalin, P., Muthukrishnan, P.,

- Ravi, V., Babu, M., Simbahan, G.C., Adviento, M.A.A., Bartolome, V., 2003. Estimating indigenous nutrient supplies for site-specific nutrient management in irrigated rice. *Agron. J.* 95 (4), 924–935.
- EU, 2013. Soil Atlas of Africa. Publications Office of the European Union, Luxembourg, European Commission.
- Gee, G.W., Bauder, J.W., 1986. Particle-size analysis. In: Klute, A. (Ed.), *Method of soil analysis part 1. Physical and mineralogical methods*, 2nd edition. Agronomy monograph. American Society of Agronomy and Soil Science Society of America, Wisconsin, pp. 383–411.
- Gérard, F., 2016. Clay minerals, iron/aluminum oxides, and their contribution to phosphate sorption in soils — A myth revisited. *Geoderma* 262, 213–226.
- Haefele, S.M., Wopereis, M.C.S., 2005. Spatial variability of indigenous supplies for N, P and K and its impact on fertilizer strategies for irrigated rice in West Africa. *Plant Soil* 270, 57–72.
- Hashimoto, Y., Kang, J., Matsuyama, N., Saigusa, M., 2012. Path analysis of phosphorus retention capacity in allophanic and non-allophanic Andisols. *Soil Sci. Soc. Am. J.* 76 (2), 441–448.
- Johnston, A.E., Poulton, P.R., Fixen, P.E., Curtin, D., 2014. Phosphorus: its efficient use in agriculture. *Advances in Agronomy*. Academic Press Inc., Cambridge, pp. 177–228.
- Jordan-Meille, L., Rubæk, G.H., Ehler, P.A.I., Genot, V., Hofman, G., Goulding, K., Recknagel, J., Provolò, G., Barraclough, P., 2012. An overview of fertilizer-P recommendations in Europe: soil testing, calibration and fertilizer recommendations. *Soil Use Manag.* 28 (4), 419–435.
- Khare, N., Hesterberg, D., Martin, J.D., 2005. XANES investigation of phosphate sorption in single and binary systems of iron and aluminum oxide minerals. *Environ. Sci. Technol.* 39 (7), 2152–2160.
- Kihara, J., Nziguheba, G., Zingore, S., Coulibaly, A., Esilaba, A., Kabambe, V., Njoroge, S., Palm, C., Huising, J., 2016. Understanding variability in crop response to fertilizer and amendments in sub-Saharan Africa. *Agric. Ecosyst. Environ.* 229, 1–12.
- Kinoshita, R., Tani, M., 2020. New estimation method of phosphate sorption coefficient using air-dry soil water content: application in arable soils of Tokachi district and Kamikawa district, Hokkaido. *Japanese Journal of Soil Science and Plant Nutrition* 91 (5), 385–388.
- Kirk, G.J.D., George, T., Courtois, B., Senadhira, D., 1998. Opportunities to improve phosphorus efficiency and soil fertility in rainfed lowland and upland rice ecosystems. *Field Crops Research* 56 (1–2), 73–92.
- Murphy, J., Riley, J.P., 1962. A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* 27, 31–36.
- Nawara, S., Van Dael, T., Merckx, R., Amery, F., Elsen, A., Odeurs, W., Vandendriessche, H., McGrath, S., Roisin, C., Jouany, C., Pellerin, S., Denoroy, P., Eichler-Löbermann, B., Börjesson, G., Goos, P., Akkermans, W., Smolders, E., 2017. A comparison of soil tests for available phosphorus in long-term field experiments in Europe. *Eur. J. Soil Sci.* 68 (6), 873–885.
- Newman, A.C.D., 1983. The specific surface of soils determined by water sorption. *J. Soil Sci.* 34, 23–32.
- Nishigaki, T., Ikazaki, K., Tsujimoto, Y., Andriamananjara, A., Rakotoson, T., Razafimbelo, T., 2020. Soil survey of the east coast and the central highlands indicates need to update Madagascar soil map. *Soil Science and Plant Nutrition* 66 (3), 469–480.
- Nishigaki, T., Tsujimoto, Y., Rinasoa, S., Rakotoson, T., Andriamananjara, A., Razafimbelo, T., 2019. Phosphorus uptake of rice plants is affected by phosphorus forms and physicochemical properties of tropical weathered soils. *Plant Soil* 435, 27–38.
- Nwoke, O.C., Vanlauwe, B., Diels, J., Sangina, N., Osonubi, O., Merckx, R., 2003. Assessment of labile phosphorus fractions and adsorption characteristics in relation to soil properties of West African savanna soils. *Agric. Ecosyst. Environ.* 100 (2–3), 285–294.
- Nziguheba, G., Zingore, S., Kihara, J., Merckx, R., Njoroge, S., Otinga, A., Vandamme, E., Vanlauwe, B., 2016. Phosphorus in smallholder farming systems of sub-Saharan Africa: implications for agricultural intensification. *Nutr. Cycl. Agroecosyst.* 104, 321–340.
- Oo, A.Z., Tsujimoto, Y., Rakotoarisoa, N.M., Kawamura, K., Nishigaki, T., 2020. P-dipping of rice seedlings increases applied P use efficiency in high P-fixing soils. *Sci. Rep.* 10, 11919.
- Rakotoarisoa, N.M., Tsujimoto, Y., Oo, A.Z., 2020. Dipping rice seedlings in P-enriched slurry increases grain yield and shortens days to heading on P-deficient lowlands in the central highlands of Madagascar. *Field Crops Research* 254, 107806.
- Rosen, C.J., Kelling, K.A., Stark, J.C., Porter, G.A., 2014. Optimizing phosphorus fertilizer management in potato production. *Am. J. Potato Res.* 91, 145–160.
- Saito, K., Vandamme, E., Johnson, J.-M., Tanaka, A., Senthilkumar, K., Dieng, I., Akakpo, C., Gbaguidi, F., Segda, Z., Bassoro, I., Lamare, D., Gbakatchetche, H., Abera, B.B., Jaiteh, F., Bam, R.K., Dogbe, W., Sékou, K., Rabeson, R., Kamissoko, N., Mossi, I.M., Tarfa, B.D., Bakare, S.O., Kalisa, A., Baggie, I., Kajiru, G.J., Ablede, K., Ayeva, T., Nanfumba, D., Wopereis, M.C.S., 2019. Yield-limiting macronutrients for rice in sub-Saharan Africa. *Geoderma* 338, 546–554.
- Shehu, B., Merckx, R., Jibrin, J., Kamara, A., Rurinda, J., 2018. Quantifying variability in maize yield response to nutrient applications in the northern Nigerian savanna. *Agronomy* 8 (2), 18.
- Six, L., Smolders, E., Merckx, R., 2013. The performance of DGT versus conventional soil phosphorus tests in tropical soils—maize and rice responses to P application. *Plant Soil* 366, 49–66.
- Soil Survey Staff, 2014. *Kellogg Soil Survey Laboratory Methods Manual*. In: R. Burt, S.S. Staff (Eds.), *Soil Survey Investigations Report No. 42 Version 5.0*. U.S. Department of Agriculture, Natural Resources Conservation Service, Washington, DC.
- Torrent, J., Del Campillo, M.C., Barrón, V., 2015. Predicting cation exchange capacity from hygroscopic moisture in agricultural soils of Western Europe. *Spanish Journal of Agricultural Research* 13 (4).
- Tsujimoto, Y., Rakotoson, T., Tanaka, A., Saito, K., 2019. Challenges and opportunities for improving N use efficiency for rice production in sub-Saharan Africa. *Plant Prod. Sci.* 22 (4), 413–427.
- Watanabe, T., 2017. Significance of active aluminum and iron on organic carbon preservation and phosphate sorption/release in tropical soils. In: Funakawa, S. (Ed.), *Soils, Ecosystem Processes, and Agricultural Development*. Springer Japan, Japan, pp. 103–125.
- Williams, W.A., Demment, M.W., Jones, M.B., 1990. A concise table for path analysis statistics. *Agron. J.* 82 (5), 1022–1024.
- Zhang, Y., Lin, X., Werner, W., 2003. The effect of soil flooding on the transformation on Fe oxides and the adsorption/desorption behavior of phosphate. *J. Plant Nutr. Soil Sci.* 166, 68–75.



ORIGINAL ARTICLE

Comparison of visual and instrumental measurements of soil color with different low-cost colorimeters

Naoki Moritsuka^a, Kensuke Kawamura^b, Yasuhiro Tsujimoto^b, Michel Rabenarivo^c, Andry Andriamananjara^c, Tovohery Rakotoson^c and Tantely Razafimbelo^c^aGraduate School of Agriculture, Kyoto University, Kyoto, Japan; ^bJapan International Research Center for Agricultural Sciences, Ibaraki, Japan;^cLaboratory of Radioisotopes, University of Antananarivo, Antananarivo, Madagascar**ABSTRACT**

Soil color has been conventionally measured using a Munsell soil color chart. Recently launched colorimeters can also measure the color of an object at a reasonable cost. This study was undertaken to evaluate to what extent such low-cost colorimeters (< 500 USD) can be useful for soil color analysis in the laboratory as compared with conventional colorimeters (> 3000 USD) and a Munsell soil color chart (about 200 USD). Sixty two air-dried soil samples collected from rice fields in Madagascar were subjected to two pretreatments for homogenization (2-mm sieving or additional hand-grinding) and instrumental analysis using four low-cost colorimeters (CS-10, Cube, Nix Pro, and Color Muse) and four conventional colorimeters (SPAD-503, CR-20, CR-400, and CR-410). The color of 2-mm sieved samples was also measured visually using a color chart. The effects of pretreatments and the analytical conditions were evaluated by the repeatability and stability of the measurement, the comparability of the soil color data obtained, and the time required for analysis. Overall, instrumental measurement was much more repeatable than visual observation. Both the repeatability and stability of the low-cost colorimeters tended to be lower than those of the conventional colorimeters. Among the low-cost colorimeters examined, soil color parameters (L^* , a^* , and b^*) obtained with Nix Pro were most comparable ($r > 0.97$ for all parameters of 2-mm sieved samples) with those obtained with SPAD-503, which is an instrument designed specifically for soil color analysis. Additional hand-grinding pretreatment improved the repeatability of the instrumental analysis and reduced the subsample weight to two grams. However, this pretreatment also increased the L^* value (lightness) of the samples, decreased the comparability with the data from 2-mm sieved samples, and prolonged the time required to complete the whole analysis. Among the various methods tested, 2-mm sieving of air-dried samples followed by the color measurement with Nix Pro several times per sample was considered the most cost-effective approach for measuring soil color in the laboratory.

ARTICLE HISTORYReceived 29 March 2019
Accepted 1 October 2019**KEY WORDS**

Cost effectiveness; grinding pretreatment; nondestructive analysis; repeatability; soil color

1. Introduction

Soil color has been used as one of the key attributes to differentiate horizons in a soil profile and to classify soil types in a region. Routine measurement of soil color is often performed visually by comparing a field moist sample with the chips in a Munsell soil color chart and selecting a chip with the nearest color. As visual color measurement can be performed rapidly without any special treatment, it may be the most appropriate initial screening test for soil discrimination (Sugita and Marumo 1996). For example, Schmidt and Ahn (2019) reviewed papers on the color of wetland soils published from 1960 to 2018 and found that 78 out of 96 papers (81%) utilized a Munsell color chart for soil color measurement. The number of color chips in a chart has been increased (Simonson 1993) to reach around 400 at present. In many cases, however, the sample color does not match well with any of the color chips, and the sensitivity of the visual method is limited by the number of chips. Furthermore, this method is prone to substantial errors due to several psychophysical and physical factors (Post et al. 1993; Marques et al. 2019).

Besides the visual method, several researchers have demonstrated the great potential of digital cameras including smartphone cameras for rapid measurement of soil color (Viscarra Rossel, Fouad, and Walter 2008; Fan et al. 2017; Aitkenhead et al. 2018; Schmidt and Ahn 2019). The number of smartphone users in the world was 1.86 billion in 2015 and is predicted to be approximately 2.9 billion by 2020 (Fan et al. 2017). However, the color of a soil picture taken with a digital camera is influenced by the ambient light conditions. Thus, additional techniques such as color correction with a reference gray card are needed to improve the accuracy of color measurements (Kirchner, Koeckhoven, and Sivakumar 2018).

More sensitive and objective measurement of soil color is possible by using a spectrophotometer or colorimeter in the laboratory, provided that appropriate white standards are used and the soil samples are carefully prepared (Torrent and Barrón 1993). Different from a Munsell color chart, a spectrophotometer or colorimeter can provide continuous quantitative data. A colorimeter designed specifically for soil color analysis (Soil Color Reader SPAD-503) was also launched by Konica Minolta in 1996. However, the prices of such instruments often exceed 3000

USD. This is more than 10 times of a soil color chart (about 200 USD), and the use of colorimeters for soil color analysis has been limited to soil scientists.

In the past decade, several new types of colorimeters have been developed and launched at prices below 500 USD (Kirchner, Koeckhoven, and Sivakumar 2019). Many of these colorimeters comprise a LED light source and a tristimulus color sensor, and are designed to be controlled wirelessly with a mobile smartphone or tablet via Bluetooth connection. Some of them are sold at less than 100 USD (Kirchner, Koeckhoven, and Sivakumar 2019). Owing to the rapid global expansion of the smartphone users, these low-cost colorimeters have enabled a wide range of users other than scientists to measure the color of an object at a reasonable cost. These tools allowed non-specialists such as farmers, gardeners, and students to measure soil color instrumentally, learn the relationship between soil color and soil components, and use the data for soil classification (Stiglitz et al. 2016a). Furthermore, Aitkenhead et al. (2017) developed a low-cost visible range spectrophotometer that is composed of a tungsten light source, several mirrors, a diffraction grating, and a digital camera to capture the visible-range spectrum of soil samples.

Nix Pro™ color sensor (Nix Sensor, Ontario, Canada), abbreviated as Nix Pro hereafter, is one of the low-cost colorimeters (Figure 1). Soon after its launch in 2015, Stiglitz et al. (2016b) used Nix Pro for soil color analysis and reported that CMYK (cyan-magenta-yellow-black) color codes of 31 soil samples (2-mm sieved, dry and moist conditions) measured with Nix Pro were significant positively correlated with those obtained from a conventional colorimeter (CR-400, Konica Minolta, Tokyo, Japan) and those by a visual method using a Munsell soil color chart. These results indicate that soil color analysis with Nix Pro is a promising alternative to the conventional methods. With a closer look of their data, however, the K% (blackness) measured by Nix Pro became lower than the corresponding K% measured by CR-400. This means that the soil color data from Nix Pro could not be compared directly with the data from CR-400.

Generally, soil color data from different instruments cannot be compared directly. This is partly because each instrument

has a unique design with respect to light source, illumination geometry, treatment of specular reflectance, and measurement area (Nishiyama et al. 2011; Gottenbos and van Biemen 2017). Furthermore, unlike the visual observation of soil color in the field, instrumental analysis is usually performed in the laboratory after sample pretreatment such as air-drying, crumbling, sieving, grinding, and remoistening (Sugita and Marumo 1996; Raeesi et al. 2019). But additional pretreatments would prolong the time required for the whole analysis. Because of the lack of thorough comparative experiments, it remains unclear how much soil color data and the required time are influenced by the combination of various pretreatments and measurement conditions.

The main objective of this study was to reveal the effectiveness of low-cost colorimeters for soil color analysis. Specific objectives were to 1) evaluate the repeatability and comparability of data obtained with several colorimeters or by visual measurement, 2) examine the effects of additional hand-grinding pretreatment of soil samples on instrumental color measurement as compared with the conventional method using 2-mm sieved samples, and 3) compare the required time among the pretreatments and measurements tested.

2. Materials and methods

2.1. Experimental design

Sixty two soil samples collected from rice fields in Madagascar were subjected to two pretreatments (2-mm sieving or additional hand-grinding) followed by color analysis with eight different colorimeters (four conventional and four low-cost) (Table 1). The color of 2-mm sieved samples was also measured visually using a Munsell soil color chart (Table 1). The effects of pretreatments and analytical conditions were evaluated by the repeatability and stability of the measurement, the comparability of the soil color data obtained, and the time required for the analysis. The term repeatability is the same as defined by Menditto, Patriarca, and Magnusson (2007), and the term comparability is defined here as the correlation of soil color data obtained with different methods.



Figure 1. Three types of low-cost colorimeters controllable with a smartphone or tablet. Cube (left), nix pro (middle), and color muse (right).

Table 1. Methods of soil color analysis used in this study.

Instruments or color chart	Sample pretreatment		Light source used	Color description
	2-mm sieved	Hand-ground		
Soil color reader (SPAD-503)	Yes	Yes	D65	CIELab
Color reader (CR-20)	Yes	Yes	D65	CIELab, Munsell
Color reader (CR-400)	Yes	Yes	D65	CIELab
Color reader (CR-410)	Yes	No	D65	CIELab
CS-10 colorimeter	Yes	Yes	D65	CIELab
Cube™	Yes	Yes	3 colored LEDs	CIELab
Nix Pro	Yes	Yes	D50	CIELab
Color Muse™	Yes	Yes	D50	CIELab
Munsell color chart	Yes	No	LED light	Munsell

To evaluate the differences in the soil color data obtained with the different instruments, the data obtained from SPAD-503 were regarded as control data.

2.2. Color description method

To describe sample color, CIE 1976 (L^* , a^* , b^*) color system (abbreviated as CIELab system hereafter) and Munsell color system were used (Table 1). The CIELab system was used for comparison of color data among the instruments, and the Munsell system was used for visual observation and comparison between visual and instrumental measurements.

The CIELab system can describe the color of an object as continuous quantitative values using three parameters (L^* , a^* , and b^*) (Viscarra Rossel et al. 2006). The L^* , a^* , and b^* axes represent whiteness-blackness (lightness-darkness), redness-greenness, and yellowness-blueness, respectively. The L^* value ranges from 0 (black) to 100 (white), whereas positive a^* and b^* values indicate the degree of redness and yellowness, respectively. On the other hand, the Munsell system can describe the color of an object qualitatively according to the visual perception using three parameters (hue, value, and chroma) (Viscarra Rossel et al. 2006). Hue is denoted categorically by the letter abbreviation of the color of the spectrum (e.g., R for red, YR for yellow-red, and Y for yellow), each followed by numbers from 0 to 10. Within each letter range, hue becomes more yellow and less red as the numbers increase. Value represents whiteness-blackness (lightness-darkness), and is described on a numerical scale from 0 (black) to 10 (white). It is almost equal to the division of the L^* value by 10. Chroma represents the degree of difference from neutral grays, and is also described on a numerical scale from 0 (neutral gray) to 20.

2.3. Soil samples tested

Sixty two soil samples with a diverse range of colors were used (Fig. S1). The sample set was the same as that described by

Kawamura et al. (2017). The samples were collected from 55 rice fields located in the Central Highlands of Madagascar. The fields contained 47 lowland and 8 upland rice fields, and their locations were mapped by Kawamura et al. (2017). Fifty-four of the samples were collected from the surface layer (< 15 cm) in different fields, and 8 samples were collected from the subsurface layer (10–50 cm from the surface) in 3 fields. Table 2 shows the descriptive statistics of the color parameters (L^* , a^* , and b^*) and chemical properties of the sample set.

2.4. Soil pretreatments and instrumental color analysis by the CIELab system

Table 3 shows the specifications of the colorimeters used. These instruments are tristimulus colorimeters, and they were arbitrarily classified into two types according to price. The production of SPAD-503 was stopped in 2016, and it was replaced by CR-20 which has almost the same specifications. Figure 1 shows three of the low-cost colorimeters examined, all of which are controllable with a smartphone. Among them, only Nix Pro was equipped with a protective cover between sensor and sample (Table 3).

Two pretreatments were applied to the samples prior to the soil color analysis with these instruments (Table 1). One pretreatment was 2-mm sieving after air-drying and crumbling. The sieved samples were packed in a plastic Petri dish (90 mm in diameter and 15 mm in depth) at the rate of about 60 cm³ to ensure the sample depth of about 1 cm (Viscarra Rossel et al. 2016). After the sample surface had been leveled with a spatula, the color was measured at five different points on the soil surface with the colorimeters (Kawamura et al. 2017), except for when CR-410 was used. For CS-10, Cube, and Color Muse, a thin plastic film was placed between the soil and the

Table 2. Color parameters and chemical properties of the soil sample set used ($n = 62$).

	L^*	a^*	b^*	Total C	Total N	Total Fe (g kg ⁻¹)	Fe _o	Fe _d
Average	44.0	11.8	21.4	21.8	1.7	61.0	8.8	41.2
Maximum	59.3	23.6	27.4	60.2	4.4	153.4	32.9	139.3
Minimum	29.5	3.9	10.8	6.5	0.6	11.1	0.3	4.3
CV (%)	17.2	33.8	15.2	53.3	48.1	43.6	67.0	51.7

The color parameters (L^* , a^* , and b^*) were obtained from the measurement of 2-mm sieved samples using SPAD-503.

Total C and N were measured by the dry combustion method using an automatic NC analyzer (Kawamura et al. 2017).

Total Fe was measured colorimetrically after digestion of soil samples with HF, HCl, and HNO₃.

Fe_o and Fe_d refer to acid oxalate-extractable and dithionite-citrate-extractable Fe, respectively, determined colorimetrically (Moritsuka et al. 2019).

Table 3. Specifications of the colorimeters examined.

Type of colorimeters	Name	Supplier (Country)	Market-launch year	Approximate retail price (USD)	Illumination geometry	Specular component	Light source	Observation angle (degrees)	Aperture diameter (mm)	Measurement diameter (mm)	Protective cover between sample and sensor
Conventional	SPAD-503	Konica Minolta (Japan)	1996	3000	8/d	SCE	D65	10	φ11	φ8	Yes
	CR-20	Konica Minolta (Japan)	2015	3000	8/d	SCE	D65	10	φ11	φ8	Yes
	CR-400	Konica Minolta (Japan)	2002	8000	d/o	SCI	D65/C	2	φ11	φ8	Yes
	CR-410	Konica Minolta (Japan)	2002	8000	d/o	SCI	D65/C	2	φ53	φ50	Yes
Low-cost	CS-10	CHNSpec (China)	2008	350	8/d	SCI	D65	10	φ7-8	unknown	No
	Cube	Palette (Australia)	2013	100	45/d	SCI?	3 colored LEDs	unknown	φ12	unknown	No
	Nix Pro	Nix Sensor (Canada)	2015	349	45/o	SCE?	D50	2	φ14	unknown	Yes
	Color Muse	Variable (USA)	2016	59	45/o	SCE?	D50	2	φ12	φ4	No

SCE: specular component excluded; SCI: specular component included. The question mark after these terms indicates our speculation from the illumination geometry.

sensor to avoid cross-contamination. For CR-410, which has a wide aperture (53 mm in diameter; Table 3), the samples were transferred from the Petri dish to a sealable plastic tub (85 mm in diameter and 45 mm in depth) and the color of the leveled surface was measured (Fig. S2); the samples in the tub were mixed thoroughly between measurements to renew the surface before every measurement, and measurements were repeated five times per sample.

The other pretreatment was additional hand-grinding of the 2-mm sieved samples (Moritsuka et al. 2014). Briefly, 2 g of 2-mm sieved samples were ground in a tungsten mortar for 3 min, and packed firmly in a disposable plastic cell with a light path length of 10 mm. The color of the sample as it appeared through the windows on both sides of the cell was measured with the colorimeters, except for when CR-410 was used. For the color measurements, all of the colorimeters, except for Nix Pro, were placed horizontally on the bench to ensure that the cell window was located in the center of the aperture of the instrument. It should be noted that the hand-grinding treatment was effective to decrease soil micro-aggregates and coarse sand particles without increasing the amount of silt plus clay particles (Moritsuka et al. 2014). It was a much weaker process than the ball-milling treatment, which significantly increases the amount of silt and clay particles (Matsuoka, Moritsuka, and Funakawa 2017).

To facilitate the comparison of the soil color data obtained from the different instruments, we selected the standard D65 light source for CR-400 and CR-410 (Table 1), and used the CIELab system for color description.

During the analysis of soil samples, the stability of the instruments was checked by measuring the white calibration plate (CR-A74) supplied with SPAD-503 once every few samples. The overall difference of the measured values from the corresponding values at time zero (ΔE_{ab}^*) was calculated by the CIELab color difference formula:

$$\Delta E_{ab}^* = ((\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2)^{1/2}$$

where ΔL^* , Δa^* , and Δb^* are the difference of the measured color parameters from the corresponding ones at time zero (Kirillova et al. 2018). When seemed necessary, recalibration was carried out for all instruments, except for Nix Pro which cannot be recalibrated by the end-user.

2.5. Soil color analysis by the Munsell system

The 2-mm sieved samples packed in the Petri dishes were also subjected to independent visual measurement by five soil scientists using a Munsell soil color chart with 389 color chips (Revised Standard Soil Color Charts, Fujihira Industry, Tokyo, Japan) (Table 1). The experiment was carried out on the same bench in the same laboratory under LED light conditions (RAD-458N, Endo Lighting Corp., Osaka, Japan). The light intensity on the bench was 1403 ± 289 lx (average \pm S.D., $n = 24$). The color chip that best matched the average sample color was selected by each observer. Munsell hue, value, and chroma of the selected chip were recorded, and the within-laboratory reproducibility of the method was evaluated. The sample

color was also reanalyzed visually by the same observer after an interval of 18 months to evaluate the method's repeatability.

To allow for comparison between the visual and instrumental measurements, the color of the same samples was measured with CR-20 and recorded by the Munsell system in the same way as we did by the CIELab system. Similarly, the color of the hand-ground samples was measured with CR-20 and recorded by the Munsell system. Munsell value and chroma were quantitative and used without adjustment. However, hue was not quantitative. For 2-mm sieved samples, the average values of hue measured by CR-20 ranged from 3.5YR to 9.8YR, except for one yellowish outlier with an average hue of 0.24Y. By the visual method, several samples were classified as 2.5Y. Accordingly, only the samples that were classified within a YR range by both instrumental and visual measurements ($n = 54$) were used for comparison.

2.6. Measurement of required time

The total time required for pretreatment and color measurement was recorded for each procedure by using a stopwatch. All pretreatments and instrumental analyses were performed by the same person to make a comparison possible.

2.7. Statistical analysis

The averages obtained from 62 samples were used for comparison. To evaluate the differences in the soil color data obtained with the different instruments, the data obtained from SPAD-503 were regarded as control data. The significant differences from the control data were identified by using Student's *t*-test. The effect of hand-grinding treatment was also evaluated by Student's *t*-test. It should be noted that the comparison of the averages can detect only systematic errors with an assumption that the tested method brings consistently higher or lower values than the reference method (Menditto, Patriarca, and Magnusson 2007). The coefficient of variation (CV) of the soil color data obtained from repeated measurements of each sample was averaged for all samples. The average CV was used as an index of measurement repeatability under each analytical condition. The comparability of the measurements obtained from different analytical conditions was analyzed by the correlation analysis and clustering using Ward's method. These analyses were performed using Microsoft Office Excel 2010 and Ekuseru-Toukei 2012 (Social Survey Research Information Co., Ltd., Tokyo, Japan).

3. Results and discussion

3.1. Average soil color parameters obtained with the different instruments under different analytical conditions

The average soil color parameters differed significantly among the instruments, when the data from SPAD-503 were compared with those from the other instruments (Table 4). For 2-mm sieved samples, the L^* value was significantly higher for CS-10, Cube, and Nix Pro. The a^* value was significantly lower for all instruments except for Nix Pro, and the b^* value was significantly higher for CR-400 and significantly lower for CR-410, CS-10, Cube, and Color Muse. Relatively similar results were obtained from hand-ground samples. All of the color parameters obtained with CS-10 and Cube were significantly different from those obtained with SPAD-503. These results reconfirm that soil color data obtained from low-cost colorimeters are not always quantitatively comparable with the data obtained from conventional colorimeters (Stiglitz et al. 2016b; Holman and Hopkins 2019).

Table 4 also shows the effect of hand-grinding pretreatment on the color measurements. It should be noted that the differences between 2-mm sieved and hand-ground samples may not have originated solely from sample color changes induced by the grinding treatment, since the sample color was measured through a plastic cell for the ground samples and through a thin plastic film for 2-mm sieved samples with CS-10, Cube, and Color Muse. The L^* value was significantly increased by the grinding treatment, except for when Nix Pro was used. The magnitude of increases did not differ among SPAD-503, CR-20, and CS-10 having similar specifications except for whether the specular component is included or excluded from the measurement. This suggests that specular reflection from the plastic cell was negligible. Increases in the L^* value upon grinding have been reported previously (Torrent and Barrón 1993; Nishiyama et al. 2011; Moritsuka et al. 2014). These observations can be explained by the grinding-induced reduction of sample surface roughness and the increase of reflection. The contrasting results obtained with Nix Pro may be due to the fact that the instrument's aperture was larger than the width of the plastic cell containing the sample. This likely caused leakage of LED light.

In contrast to the L^* values, the a^* and b^* values did not change or decreased significantly by the grinding treatment, except for when the Color Muse was used (Table 4). The changes were nonsignificant for SPAD-503, CR-20, and CS-10, which have 8/d illumination geometries, whereas a significant decrease of both parameters was observed when CR-400 having a d/0 geometry was used. This suggests that the effect of

Table 4. Average soil color parameter values obtained with different instruments with or without additional hand-grinding pretreatment ($n = 62$).

Sample pretreatment	Color parameter	SPAD-503	CR-20	CR-400	CR-410	CS-10	Cube	Nix Pro	Color Muse
2-mm sieved	L^*	44.0	44.3	46.2	44.8	53.0**	54.6**	48.5**	45.7
	a^*	11.8	10.1*	9.0**	8.2**	4.1**	10.1**	11.8	10.5*
	b^*	21.4	21.8	23.8**	19.1**	15.4**	15.3**	21.3	17.9**
Hand-ground	L^*	51.7	50.9	56.6**	no data	59.2**	64.2**	47.6**	57.0**
	a^*	11.8	9.9*	6.6**	no data	3.3**	7.4**	10.3	11.7
	b^*	21.8	21.5	17.4**	no data	14.6**	16.0**	18.9**	19.9**

** and * indicate significant difference at $p < 0.01$ and $p < 0.05$, respectively, from the corresponding value obtained with SPAD-503. Bold values in hand-ground samples indicate significant difference at $p < 0.05$ from the corresponding values in 2-mm sieved samples.

grinding pretreatment on the a^* and b^* values was influenced by the illumination geometry of the instrument. Contrary to our results, Nishiyama et al. (2011) reported that the b^* value of dark red soils measured with CR-200, which has the same specifications as CR-400, tended to increase by hand-grinding 0.15-mm sieved samples for 10 min. The reason for these contrasting results remains to be explained.

3.2. Stability of instrumental measurements

Instrumental differences were also observed when a white plate (CR-A74) was measured after the instruments were calibrated using a white calibration plate (Table 5). For example, the a^* and b^* values obtained with CS-10 were positive and negative, respectively, and were opposite to the results obtained with all of the other instruments. Furthermore, the color parameter values of the white plate fluctuated over time during the soil analysis (Figure 2). Except for CR-410, the data in this figure were obtained during the analysis of hand-ground samples. The fluctuations in the values tended to be larger for low-cost colorimeters, especially Color Muse. The a^* value of the white plate measured with SPAD-503 also decreased by 0.5 at 60 min after the start of the analysis. However, the overall fluctuations expressed as the difference from the measurement at time zero (ΔE_{ab}^*) were less than 2.0 for all instruments, suggesting that these drifts were barely recognizable visually (Kirillova et al. 2018). It should be noted that Color Muse was much more stable during the analysis of 2-mm sieved samples with keeping the upright position constantly ($\Delta E_{ab}^* < 0.3$), whereas the other low-cost colorimeters showed similar fluctuations ($0.5 < \Delta E_{ab}^* < 1.0$) after measuring 2-mm sieved samples for about 50 min.

3.3. Repeatability of instrumental measurements under different analytical conditions

The repeatability of soil color measurements differed significantly among the instruments, when the data from SPAD-503 were compared with those from the other instruments (Table 6). For 2-mm sieved samples, the average CV of the L^* value was significantly higher for Nix Pro and was significantly lower for CR-410, CS-10, and Cube. The average CVs of the a^* and b^* values were significantly higher for all of the low-cost instruments. Overall, increasing the diameter of measurement area from 8 mm (CR-400) to 50 mm (CR-410) (Table 3) did not considerably improve repeatability for the a^* and b^* values. The lower repeatability of the a^* and b^* values for the low-cost colorimeters may be due to the smaller diameter of measurement than the conventional instruments. Among the low-cost colorimeters, however, data on the measurement diameter were available only for

Color Muse (Table 3). The lowest repeatability of the a^* value for CS-10 was due to the smallest values (Table 4), some of which became below zero. Similar instrumental differences were observed for hand-ground samples.

The average CVs of hand-ground samples were comparable to, or significantly lower than, those of 2-mm sieved samples (Table 6). Thus, repeatability was improved considerably by hand-grinding treatment. The size to which a soil sample should be ground and homogenized to obtain repeatable results depends on the measurement area of the instrument to be used (subsample size) and the required level of repeatability. The required level will increase as sample color variation decreases, as we have delineated the spatial variation of soil color at the farm scale by using hand-grinding pretreatment (Moritsuka et al. 2019).

For soil pretreatment before chemical analysis, 2-mm sieving after air-drying and crumbling has become the global standard. When the subsample used for analysis is less than 2 g, further grinding of the 2-mm sieved sample is often carried out for sample homogenization (ISO 2006; Matsuoka, Moritsuka, and Funakawa 2017). Recently, the measurement of the 2-mm sieved sample was also proposed as the standard protocol for soil color analysis with colorimeters and spectrophotometers (Viscarra Rossel et al. 2016). However, our results clearly indicated that the hand-grinding of a 2-mm sieved sample was more effective to improve repeatability than increasing the measurement diameter from 8 to 50 mm (Table 6). To better interpret our result and improve the current protocol, more efforts are needed to understand the relationship among the sample heterogeneity, subsample size and repeatability with respect to soil color measurement.

3.4. Correlations among the instrumental measurements under different analytical conditions

The results shown in Table 4 reconfirmed that the soil color data obtained under a given analytical condition cannot be compared directly with data obtained under different analytical conditions. If the data are to be used for estimating the content of related soil properties, such as the amounts of organic matter and iron oxides in the samples, it is also necessary to evaluate the relationship between soil color measurements and the target variable and obtain a calibration equation for the sample set. Thus, the relative relationship of soil color data within the sample set becomes important, besides the absolute values of soil color measurements.

Overall, the correlation coefficients between the soil color data and the corresponding data obtained with SPAD-503 were higher than 0.97, 0.94, and 0.86 for the L^* , a^* and b^* values, respectively (Table 7). This indicates that the relationships between soil color parameters and their related properties can

Table 5. Color parameter values of a white calibration plate (CR-A74) measured with different instruments.

Color parameter	SPAD-503	CR-20	CR-400	CR-410	CS-10	Cube	Nix Pro	Color Muse
L^*	93.5 ± 0.00	94.7 ± 0.00	95.2 ± 0.10	94.7 ± 0.02	94.9 ± 0.11	99.7 ± 0.00	93.4 ± 1.19	97.2 ± 0.02
a^*	-0.5 ± 0.07	-0.6 ± 0.07	-0.4 ± 0.04	-0.6 ± 0.00	3.0 ± 0.00	-0.8 ± 0.07	-0.4 ± 0.06	-0.4 ± 0.06
b^*	3.4 ± 0.00	3.8 ± 0.07	2.8 ± 0.16	3.0 ± 0.00	-0.5 ± 0.04	2.4 ± 0.00	2.1 ± 0.35	3.6 ± 0.04

Data were average ± S.D. ($n = 2$) obtained soon after the instruments except for Nix Pro were calibrated with a white standard plate.

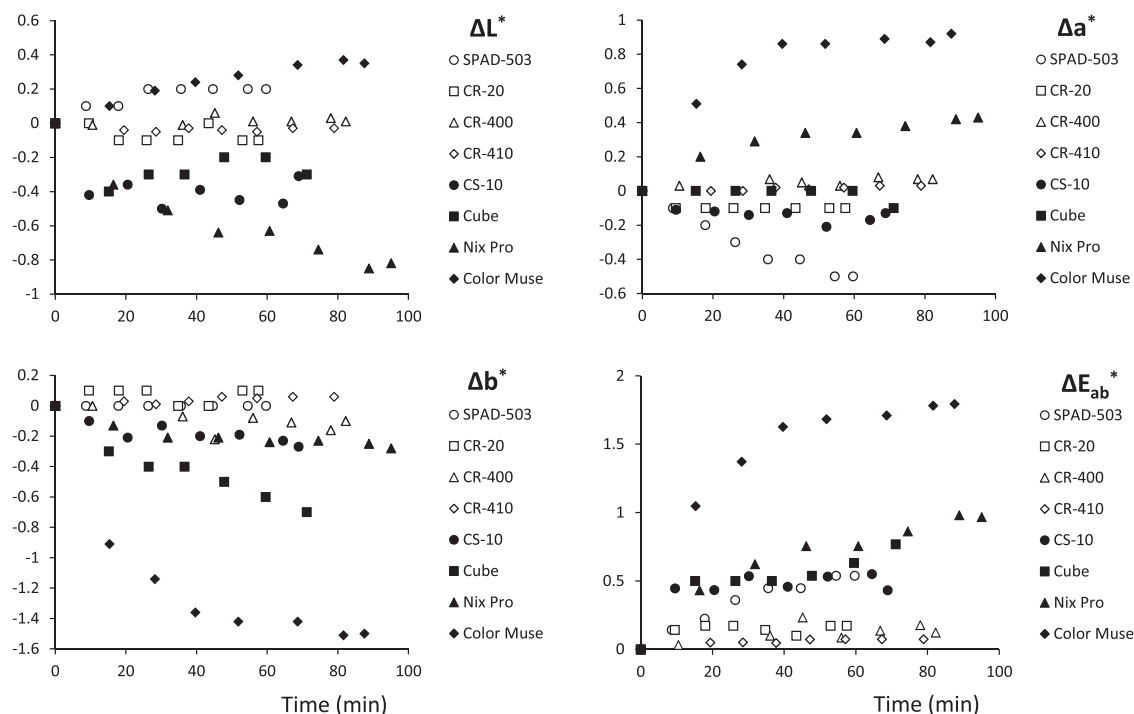


Figure 2. Temporal changes in the color parameter values of a white calibration plate (CR-A74), measured once every few samples. Values indicate the differences from the first measurement at time zero.

Table 6. Average CV (%) of repeated analysis of each sample obtained with different instruments with or without additional hand-grinding pretreatment.

Sample pretreatment	Color parameter	SPAD-503	CR-20	CR-400	CR-410	CS-10	Cube	Nix Pro	Color Muse
2-mm sieved	L*	1.73	1.67	1.73	1.09**	0.91**	1.37**	2.68**	1.59
	a*	1.61	1.52	1.98*	1.95*	10.61**	4.14**	2.11**	7.60**
	b*	1.41	1.33	1.27	1.45	1.84**	4.18**	2.06**	3.56**
Hand-ground	L*	0.32	0.28	0.39	no data	0.62**	0.24	1.79**	1.53**
	a*	0.52	0.52	0.79**	no data	9.77	3.97*	1.64**	2.35**
	b*	0.56	0.62	0.62	no data	1.16**	1.33**	2.00**	1.18**

** and * indicate significant difference at $p < 0.01$ and $p < 0.05$, respectively, from the corresponding value obtained with SPAD-503. Bold values in hand-ground samples indicate significant difference at $p < 0.05$ from the corresponding values in 2-mm sieved samples.

be evaluated properly regardless of the pretreatments and instrument used. This is analogous to the fact that a number of soil extraction methods are equally suitable for estimating the availability of a particular nutrient to crop plants. The correlation coefficients obtained in our study were generally higher than those reported previously ($r = 0.45\text{--}0.97$, $n = 31$), in which the CMYK color codes of dry 2-mm sieved soils were compared between Nix Pro and CR-400 (Stiglitz et al. 2016b).

The cluster analysis showed that the data obtained with SPAD-503 were correlated most closely with those with CR-20 or Nix Pro, irrespective of sample pretreatment (Figure 3). Among the low-cost colorimeters, Nix Pro was most comparable to SPAD-503. For the L* value, the data on 2-mm sieved samples were classified in a cluster separate from those of hand-ground samples. The effects of pretreatment were less conspicuous for the a* and b* values, and the data on 2-mm sieved samples obtained with SPAD-503 were distant from those of 2-mm sieved samples obtained with CS-10, Cube, and Color Muse. As shown from the correlation coefficients (Table 7) and the heights in the dendrogram, the b* value was less comparable among the treatments than the L* and a* values.

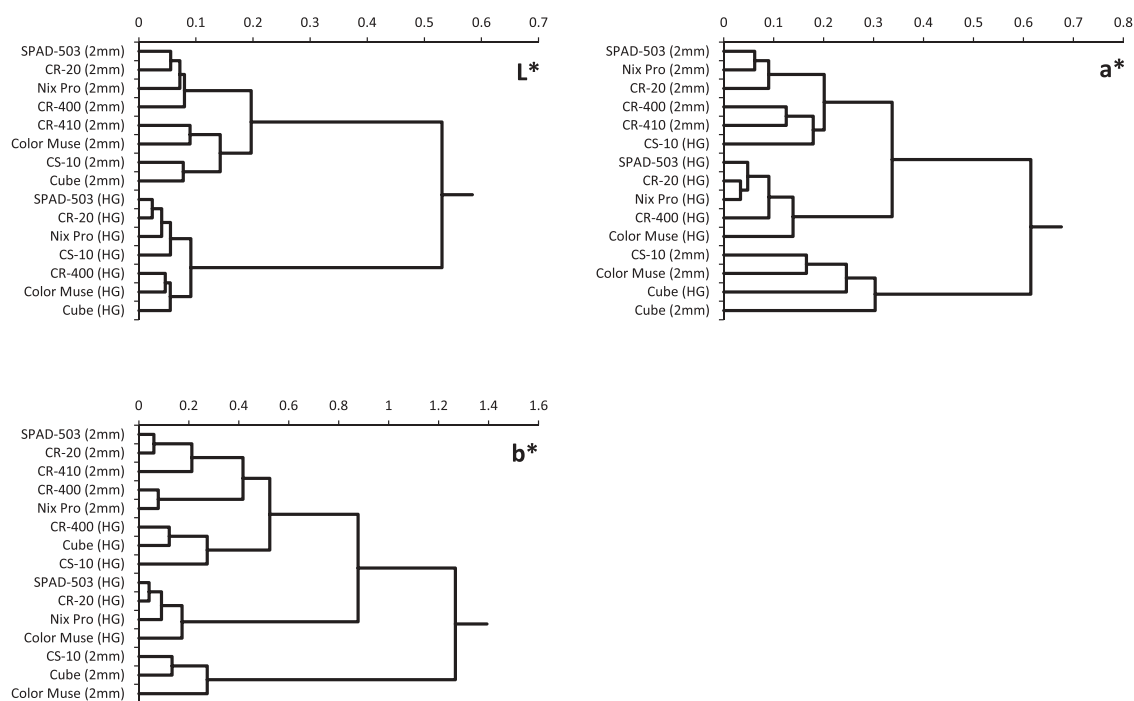
3.5. Comparison between visual and instrumental measurements

Visual observations of 2-mm sieved samples by five soil scientists produced average CVs for Munsell value and chroma of around 13% and 20%, respectively (Table 8). These CVs were much larger than those obtained from repeated measurements of the same samples using CR-20 (< 2%). It should be noted that these CVs obtained by visual and instrumental analysis denote within-laboratory reproducibility and repeatability, respectively, and may not be comparable directly (Menditto, Patriarca, and Magnusson 2007). When the sample color was reanalyzed visually by the same observer, the same color chip was selected again for only 26 out of 62 samples (42%). This repeatability was comparable with data from a similar experiment using 276 samples and four independent observers (22%, 31%, 49%, and 58%) (Marqués-Mateu et al. 2018). These results suggest that the sample color did not match exactly any color chip, thus preventing the observers from selecting the best-matching chip with consistency. By measuring the color of 161 soil samples and all of the color chips in a Munsell color chart with a portable spectrophotometer, Kirillova et al. (2018)

Table 7. Correlation coefficients for the color parameters obtained with different pretreatments and instruments.

Color parameter	Sample pretreatment	SPAD-503	CR-20	CR-400	CR-410	CS-10	Cube	Nix Pro	Color Muse
L*	2-mm sieved	control	1.00	1.00	0.99	0.99	0.99	1.00	0.99
	Hand-ground	0.98	0.98	0.97	no data	0.98	0.97	0.98	0.97
a*	2-mm sieved	control	1.00	0.99	0.99	0.98	0.94	1.00	0.98
	Hand-ground	0.98	0.98	0.98	no data	0.99	0.97	0.99	0.98
b*	2-mm sieved	control	1.00	0.96	0.98	0.88	0.89	0.97	0.95
	Hand-ground	0.86	0.86	0.94	no data	0.96	0.94	0.89	0.88

Color parameter values were compared with those obtained for 2-mm sieved samples with SPAD-503 (control).

**Figure 3.** Dendrogram showing clustering for two pretreatments and different instruments. 2 mm: 2-mm sieved sample; HG: hand-ground sample.**Table 8.** Averages and repeatability of Munsell value and chroma, as measured visually with a Munsell color chart or instrumentally with CR-20.

Color parameter		2-mm sieved		Hand-ground
		Color chart	CR-20	CR-20
Value	Average	4.96	4.31**	4.94
	Repeatability	12.53	1.77**	0.32**
Chroma	Average	4.12	3.90	3.89
	Repeatability	19.93	1.59**	0.56**

Repeatability refers to the average CV (%) of repeated analysis of each sample (same as Table 6).

** indicates significant difference at $p < 0.01$ from the corresponding value obtained using a color chart.

revealed that 48% of the soil samples were distinguishable visually ($\Delta E_{ab}^* > 3$) from their nearest color chip.

Despite the low reproducibility and repeatability of the visual method, the average Munsell value for each sample ($n = 5$) was significantly correlated ($r = 0.95$) with the instrumental measurement for both 2-mm sieved and hand-ground samples (Figure 4). The average value obtained from visual observation of 2-mm sieved samples was significantly higher than that obtained with CR-20 (2-mm sieved) and was close to that obtained with CR-20 (hand-ground) (Table 8, Figure 4). This trend was not observed for chroma. The average chroma (color chart) was more strongly correlated with that obtained with CR-20 (hand-ground) than

with that obtained with CR-20 (2-mm sieved) (Figure 4). As shown by the slopes of the regression equations exceeding 1.0, the variations of value and chroma obtained with the visual method were larger than those obtained with CR-20. Regression equations with a slope exceeding 1.0 were also observed for hue (Fig. S3), although the samples used for comparison were limited to those with a hue classified as YR ($n = 54$).

Similar results have been reported previously (Post et al. 1993; Fan et al. 2017; Marqués-Mateu et al. 2018). The observers of 2-mm sieved soil samples tended to assign a color chip with a higher value (in more than 75% of cases) and chroma (in 67% of cases) than the chip selected with a colorimeter (Konica-Minolta Chroma meter CS-100A) (Marqués-Mateu et al. 2018). By measuring 2-mm sieved soil samples, Fan et al. (2017) also found that visual observation of soil color using a Munsell color chart brought significantly higher L* value than the instrumental measurement using Konica-Minolta CM-600d with the average difference of 6.6 between the two methods. This difference corresponds to about 0.6 unit of Munsell value, and was similar to the difference observed in this study (0.65, Table 8).

These systematic discrepancies between visual and instrumental measurement have been explained by the differences in color sensing between the human eye and colorimeters (Post et al. 1993). From our results using hand-ground samples,

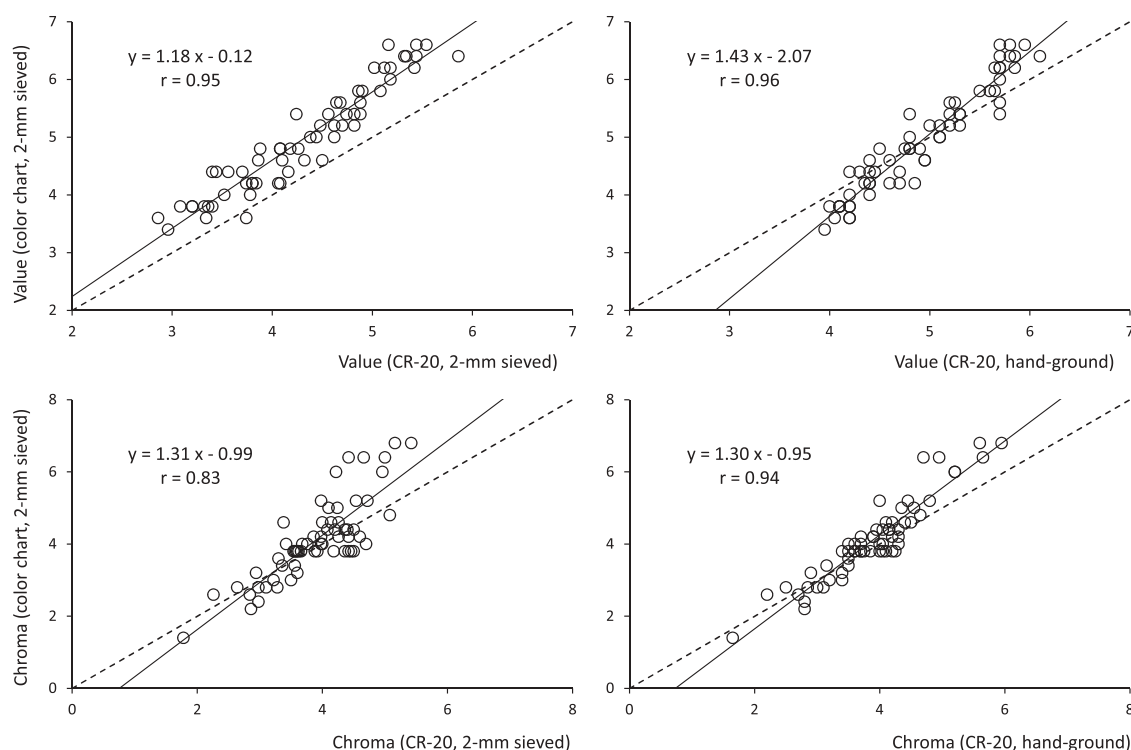


Figure 4. Relationship between visual and instrumental (CR-20) measurements of Munsell value and chroma for 2-mm sieved and hand-ground soil samples. The broken line in the figure shows a 1:1 line.

however, it can be postulated that a kind of optical illusion caused the following subsampling bias; observers tend to look for a flat, homogenous spot on the heterogeneous surface of the sample before comparing the sample color to the homogenous color chips (Fig. S4). If such a subsampling bias is present, a mismatch between visual and instrumental measurements will inevitably occur unless the soil surface is uniform in color. Further research is needed to identify and measure the subjective factors in each step of the visual method.

3.6. Time required for soil color analysis

The time required to complete soil color analysis can affect research design and associated labor cost, so time is as

important as instrument cost and data quality (Dobriyal et al. 2012; Bünemann et al. 2018). The x-axis of Figure 2 indicates that the time required for analysis of hand-ground samples (2 measurements per sample, $n = 62$) ranged from 58 to 95 min. Once the samples were ready for measurement, soil color could be measured within 2 min per sample with all of the instruments tested.

However, hand-grinding pretreatment required a much longer time than packing 2-mm sieved samples in a Petri dish (Table 9). The differences among the instruments, except for CR-410, were smaller than those between the pretreatments. As a result, when the total times for pretreatment and measurement were compared, the analysis

Table 9. Total time required for pretreatment and measurement of 62 soil samples.

Type of pretreatment	Type of measurement	Measurements per sample (times)	Pretreatment (min)	Measurement (min)	Total (min)
Petri dish packing ^a	SPAD-503	5	277	364	641
	CR-20	5	277	329	606
	CR-400	5	277	329	606
	CR-410	5	277	788	1065
	CS-10	5	277	250	527
	Cube	5	277	226	503
	Nix Pro	5	277	221	498
	Color Muse	5	277	246	523
	Color chart	1	277	132±18 ^b	409±18 ^b
	Hand-grinding and plastic cell packing	SPAD-503	2	841	60
CR-20		2	841	58	899
CR-400		2	841	82	923
CS-10		2	841	69	910
Cube		2	841	77	918
Nix Pro		2	841	95	936
Color Muse		2	841	88	936

^a Preparation of air-dried, 2-mm sieved samples is not included in this pretreatment.

^b Average±S.D. ($n = 5$). No replicates for other records.

All records except for the visual method using a color chart were obtained by the same person.

of hand-ground samples took longer than that of 2-mm sieved samples, irrespective of the instrument used. In spite of the differences among the methods tested, each analysis of the complete sample was finished within 2 working days (Table 9). This is a major advantage of soil color analysis over conventional destructive methods for soil analysis.

The visual method was faster than all of the instrumental analyses, thus reconfirming its usefulness for rapid screening without the need of electricity. However, a large variation of soil color within a sample set is needed for this method to be useful for soil quality assessment (van Leeuwen et al. 2018). In this study, the color variation was recognizable (Fig. S1), but many observers felt it difficult to assign the best-matching chip, probably due to the limited numbers of color chips to choose from. Soil color observation by more than one person is recommended to improve the accuracy of visual measurements (Melville and Atkinson 1985). In this case, the total man-hour will increase in proportion to the number of observers.

4. Conclusions

In this study, the repeatability and comparability of soil color data were compared among different methods. The comparison between conventional and low-cost instruments indicated that there was a trade-off between the instrumental cost and the repeatability of data to some extent. But the sensitivity of the low-cost colorimeters was much higher than that of the visual method. Therefore, the low-cost colorimeters can help to bridge the cost-effectiveness gap between visual observation with a Munsell color chart and instrumental measurement with a conventional colorimeter.

Among the various pretreatment and instrument combinations tested, 2-mm sieving of air-dried samples followed by color measurement with Nix Pro several times per sample was considered the most cost-effective approach for measuring soil color in the laboratory. Further investigations are needed to evaluate whether the low-cost colorimeters can be effective across various environments including field conditions.

Acknowledgments

We thank Ms. Sehen Rinasoa, Laboratory of Radioisotopes, University of Antananarivo, for acting as an observer during the visual soil color analysis, Dr. Tomohiro Nishigaki, JIRCAS, for providing valuable comments on an early version of our manuscript, and Mr. Kan Miyamoto, Ibaraki Agricultural Center, Ibaraki Prefecture, for providing useful information on the Cube colorimeter.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

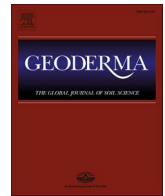
This research was supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and

Technology Agency (JST)/Japan International Cooperation Agency (JICA) [Grant no. JPMJSA1608].

References

- Aitkenhead, M., C. Cameron, G. Gaskin, B. Choisy, M. Coull, and H. Black. 2018. "Digital RGB Photography and Visible-range Spectroscopy for Soil Composition Analysis." *Geoderma* 313: 265–275. doi:10.1016/j.geoderma.2017.11.020.
- Aitkenhead, M. J., G. J. Gaskin, N. Lafouge, and C. Hawes. 2017. "PHYLIS: A Low-cost Portable Visible Range Spectrometer for Soil and Plants." *Sensors* 17: 99. doi:10.3390/s17010099.
- Bünemann, E. K., G. Bongiorno, Z. Bai, R. E. Creamer, G. D. Deyn, R. de Goede, L. Fleskens, et al. 2018. "Soil Quality – A Critical Review." *Soil Biology and Biochemistry* 120: 105–125. doi:10.1016/j.soilbio.2018.01.030.
- Dobriyal, P., A. Qureshi, R. Badola, and S. A. Hussain. 2012. "A Review of the Methods Available for Estimating Soil Moisture and Its Implications for Water Resource Management." *Journal of Hydrology* 458–459: 110–117. doi:10.1016/j.jhydrol.2012.06.021.
- Fan, Z., J. E. Herrick, R. Saltzman, C. Matteis, A. Yudina, N. Nocella, E. E. Crawford, R. Parker, and J. V. Zee. 2017. "Measurement of Soil Color: A Comparison between Smartphone Camera and the Munsell Color Charts." *Soil Science Society of America Journal* 81: 1139–1146. doi:10.2136/sssaj2017.01.0009.
- Gottenbos, R., and E. van Biemen. 2017. "Overview Low Cost Color Capture Devices." A post on the website of Color Technology Consultancy, Leiderdorp, The Netherlands. Accessed 22 August 2019. <https://www.coltechcon.com/publication/overview-low-cost-color-capture-devices/>
- Holman, B. W. B., and D. L. Hopkins. 2019. "A Comparison of the Nix Colour Sensor Pro™ and HunterLab MiniScan™ Colorimetric Instruments When Assessing Aged Beef Colour Stability over 72 h Display." *Meat Science* 147: 162–165. doi:10.1016/j.meatsci.2018.09.009.
- ISO. 2006. *Soil Quality – Pretreatment of Samples for Physico-chemical Analysis*. 2nd ISO 11464:2006. Geneva: ISO copyright office
- Kawamura, K. Y., T. M. Rabenarivo, H. Asai, A. Andriamananjara, and T. Rakotoson. 2017. "Vis-NIR Spectroscopy and PLS Regression with Waveband Selection for Estimating the Total C and N of Paddy Soils in Madagascar." *Remote Sensing* 9: 1081. doi:10.3390/rs9101081.
- Kirchner, E., P. Koeckhoven, and K. Sivakumar. 2018. "Improving Color Accuracy of Colorimetric Sensors." *Sensors* 18: 1252. doi:10.3390/s18041252.
- Kirchner, E., P. Koeckhoven, and K. Sivakumar. 2019. "Predicting the Performance of Low-cost Color Instruments for Color Identification." *Journal of the Optical Society of America A* 36: 368–376. doi:10.1364/JOSAA.36.000368.
- Kirillova, N. P., J. Grauer-Gray, A. E. Hartemink, T. M. Sileova, Z. S. Artemyeva, and E. K. Burova. 2018. "New Perspectives to Use Munsell Color Charts with Electronic Devices." *Computers and Electronics in Agriculture* 155: 378–385. doi:10.1016/j.compag.2018.10.028.
- Marques, K. P. P., R. Rizzo, A. C. Dotto, A. B. E. Souza, F. A. O. Mello, L. G. M. Neto, L. H. C. Dos Anjos, and J. A. M. Demattê. 2019. "How Qualitative Spectral Information Can Improve Soil Profile Classification?" *Journal of near Infrared Spectroscopy* 27: 156–174. doi:10.1177/0967033518821965.
- Marqués-Mateu, Á., H. Moreno-Ramón, S. Balasch, and S. Ibáñez-Asensio. 2018. "Quantifying the Uncertainty of Soil Colour Measurements with Munsell Charts Using a Modified Attribute Agreement Analysis." *Catena* 171: 44–53. doi:10.1016/j.catena.2018.06.027.
- Matsuoka, K., N. Moritsuka, and S. Funakawa. 2017. "Ball Milling Pretreatment Affects the Content of Fixed Ammonium in Soils in Response to the Content of Exchangeable Ammonium." *Soil Science and Plant Nutrition* 63: 321–328. doi:10.1080/00380768.2017.1366249.
- Melville, M. D., and G. Atkinson. 1985. "Soil Colour: Its Measurement and Its Designation in Models of Uniform Colour Space." *Journal of Soil Science* 36: 495–512. doi:10.1111/j.1365-2389.1985.tb00353.x.
- Menditto, A., M. Patriarca, and B. Magnusson. 2007. "Understanding the Meaning of Accuracy, Trueness and Precision." *Accreditation and Quality Assurance* 12: 45–47. doi:10.1007/s00769-006-0191-z.

- Moritsuka, N., K. Matsuoka, K. Katsura, and J. Yanai. 2019. "Farm-scale Variations in Soil Color as Influenced by Organic Matter and Iron Oxides in Japanese Paddy Fields." *Soil Science and Plant Nutrition* 65: 166–175. doi:10.1080/00380768.2019.1583542.
- Moritsuka, N. K., M. K. Katsura, S. Sano, and J. Yanai. 2014. "Soil Color Analysis for Statistically Estimating Total Carbon, Total Nitrogen and Active Iron Contents in Japanese Agricultural Soils." *Soil Science and Plant Nutrition* 60: 475–485. doi:10.1080/00380768.2014.906295.
- Nishiyama, K., T. Kimura, Y. Isono, and Y. Inoue. 2011. "Color Measurements of Rocks and Soils Using Colorimeters." *Journal of the Japan Society of Engineering Geology* 52: 62–71. in Japanese. doi:10.5110/jjseg.52.62.
- Post, D. F., R. B. Bryant, A. K. Batchily, A. R. Huete, S. J. Levine, M. D. Mays, and R. Escadafal. 1993. "Correlations between Field and Laboratory Measurements of Soil Color." In *Soil Color*, edited by J. M. Bigham and E. J. Ciolkosz, 35–49. Madison, WI: SSSA. doi:10.2136/sssaspecpub31.c3.
- Raeesi, M. A. A., M. R. Zolfaghari, Y. M. Gorji, and M. Sabetzade. 2019. "Prediction of Soil Organic Matter Using an Inexpensive Colour Sensor in Arid and Semiarid Areas of Iran." *Soil Research* 57: 276–286. doi:10.1071/SR18323.
- Schmidt, S. A., and C. Ahn. 2019. "A Comparative Review of Methods of Using Soil Colors and Their Patterns for Wetland Ecology and Management." *Communications in Soil Science and Plant Analysis* 50: 1293–1309. doi:10.1080/00103624.2019.1604737.
- Simonson, R. W. 1993. "Soil Color Standards and Terms for Field Use – History of Their Development." In *Soil Color*, edited by J. M. Bigham and E. J. Ciolkosz, 1–20. Madison, WI: SSSA. doi:10.2136/sssaspecpub31.c1.
- Stiglitz, R., E. Mikhailova, C. Post, M. Schlautman, and J. Sharp. 2016b. "Evaluation of an Inexpensive Sensor to Measure Soil Color." *Computers and Electronics in Agriculture* 121: 141–148. doi:10.1016/j.compag.2015.11.014.
- Stiglitz, R. Y., E. A. Mikhailova, C. J. Post, M. A. Schlautman, and J. L. Sharp. 2016a. "Teaching Soil Color Determination Using an Inexpensive Color Sensor." *Natural Sciences Education* 45: 1–7. doi:10.4195/nse2016.03.0005.
- Sugita, R., and Y. Marumo. 1996. "Validity of Color Examination for Forensic Soil Identification." *Forensic Science International* 83: 201–210. doi:10.1016/S0379-0738(96)02038-5.
- Torrent, J., and V. Barrón. 1993. "Laboratory Measurement of Soil Color: Theory and Practice." In *Soil Color*, edited by J. M. Bigham and E. J. Ciolkosz, 21–34. Madison, WI: SSSA. doi:10.2136/sssaspecpub31.c2.
- van Leeuwen, M. M. W. J., G. B. M. Heuvelink, J. Wallinga, I. J. M. de Boer, J. C. van Dam, E. A. van Essen, S. W. Moolenaar, F. P. M. Verhoeven, J. J. Stoorvogel, and C. R. Stoof. 2018. "Visual Soil Evaluation: Reproducibility and Correlation with Standard Measurements." *Soil and Tillage Research* 178: 167–178. doi:10.1016/j.still.2017.11.012.
- Viscarra Rossel, R. A., T. Behrens, E. Ben-Dor, D. J. Brown, J. A. M. Dematté, K. D. Shepherd, Z. Shi, et al. 2016. "A Global Spectral Library to Characterize the World's Soil." *Earth-Science Reviews* 155: 198–230. doi:10.1016/j.earscirev.2016.01.012.
- Viscarra Rossel, R. A., Y. Fouad, and C. Walter. 2008. "Using a Digital Camera to Measure Soil Organic Carbon and Iron Contents." *Biosystems Engineering* 100: 149–159. doi:10.1016/j.biosystemseng.2008.02.007.
- Viscarra Rossel, R. A., B. Minasny, P. Roudier, and A. B. McBratney. 2006. "Colour Space Models for Soil Science." *Geoderma* 133: 320–337. doi:10.1016/j.geoderma.2005.07.017.



Laboratory and field measurement of magnetic susceptibility of Japanese agricultural soils for rapid soil assessment

Naoki Moritsuka^{a,*}, Kaori Matsuoka^b, Keisuke Katsura^c, Shuji Sano^d, Junta Yanai^e

^a Faculty of Agriculture and Marine Science, Kochi University, Nankoku 783-8502, Japan

^b Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization, Tsukuba 305-8604, Japan

^c Graduate School of Agriculture, Tokyo University of Agriculture and Technology, Fuchu 183-8509, Japan

^d Faculty of Agriculture, Setsunan University, Hirakata 573-0101, Japan

^e Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Kyoto 606-8522, Japan

ARTICLE INFO

Handling Editor: Ingrid Kögel-Knabner

Keywords:

Agricultural soil
Proximal sensing
Repeatability
Scale dependence
Magnetic susceptibility

ABSTRACT

Measuring volume magnetic susceptibility (κ) of soil allows rapid soil assessment. In this study, we measured the κ value of agricultural surface soils in Japan at national and farm scales in order to evaluate the scale-dependent relationship between soil κ and other soil properties and to examine the repeatability of the κ value measured in the laboratory and in the field. A handheld field sensor was used to measure two sample sets: (1) 164 samples collected throughout Japan (national scale), and (2) 246 samples collected from paddy fields in Takatsuki and 117 (39 sites \times 3 times) samples collected from paddy fields in Mifune (farm scale). Laboratory measurements showed that the coefficients of variation of soil κ were 112% for the national scale samples, 13% for the Takatsuki samples, and 33–37% for the Mifune samples. The large variation at the national scale was due to several positive outliers, which were classified as Vitric or Silandic Andosols by the WRB classification. Repeated measurement of samples enabled estimation of the within-sample variation in κ , and the data was considered unreliable if the coefficient of variation exceeded 10%. The correlation analysis using the reliable data indicated that at the national scale, the κ value was most positively correlated with structural Fe (total Fe minus Fe_d) followed by Al_o and the andic properties ($\text{Al}_o + 1/2\text{Fe}_o$). In contrast, for the Mifune samples collected at the farm scale, the κ value was most positively correlated with sand content followed by K saturation percentage. Field measurements of soil κ at the 39 Mifune sites indicated that the coefficients of variation of κ at the same site were almost all below 10%. The within-site variations were due probably to the incomplete contact between the soil surface and the sensor. Our results suggest that measuring κ of agricultural surface soils in Japan can help to distinguish Vitric or Silandic Andosols from other soils, and to estimate the κ -related soil properties at the farm and national scales.

1. Introduction

The ratio of induced magnetization to an applied magnetic field is called the volume magnetic susceptibility (κ). The κ value of soil generally reflects the abundance, nature, and chemical composition of soil constituents possessing ferromagnetic, ferrimagnetic, canted anti-ferromagnetic, paramagnetic, or diamagnetic properties (Dearing, 1999). Ferrimagnetic materials such as magnetite (Fe_3O_4) and maghemite ($\gamma\text{-Fe}_2\text{O}_3$) have much higher κ than other materials. Canted anti-ferromagnetic materials (e.g., hematite and goethite) have moderate positive κ , whereas paramagnetic materials (e.g., biotite and olivine) and diamagnetic materials (e.g., water, quartz, and feldspars) have weak

positive and weak negative κ , respectively (Dearing, 1999).

As soil κ can be measured rapidly and nondestructively at a relatively low cost, the measurement of κ in the field can be a promising approach for rapid soil assessment. Although magnetite and maghemite are the two most important magnetic minerals in soil, soil κ depends not only on the concentration of magnetic minerals but also on their shape and size, and the method of measurement. Thus, field measurement of soil κ can be used for rapid soil assessment if measurements are reliable and soil κ is statistically correlated with the target soil variable.

To evaluate the reliability of field measurements, Boyko et al. (2004) compared field measurements of surface soil κ obtained with a field probe (Bartington MS2D, Bartington Instruments Ltd., UK) and

* Corresponding author.

E-mail address: moritsuka@kochi-u.ac.jp (N. Moritsuka).

<https://doi.org/10.1016/j.geoderma.2021.115013>

Received 29 October 2020; Received in revised form 12 February 2021; Accepted 13 February 2021

Available online 10 March 2021

0016-7061/© 2021 Elsevier B.V. All rights reserved.

Table 1
Numbers of soil samples in relation to soil type and land use for the national scale samples.

Soil type ^a	Relevant soil type ^b	Paddy	Upland	Grassland	Fallow	Total
Peat soils	Histosols	4	3	0	0	7
Sand-dune Regosols	Arenosols	0	1	0	0	1
Volcanogenous Regosols	Vitric Andosols	0	3	0	0	3
Wet Andosols	Gleyic Andosols	2	3	0	0	5
Non-allophanic Andosols	Aluandic Andosols	4	3	0	0	7
Andosols	Silandic Andosols	5	27	1	1	34
Lowland Paddy soils	Anthrosols	40	0	0	0	40
Gley Lowland soils	Fluvisols	21	3	0	0	24
Gray Lowland soils	Fluvisols	3	3	0	0	6
Brown Lowland soils	Fluvisols	0	16	1	1	18
Gray Upland soils	Planosols	0	4	1	0	5
Dark Red soils	Cambisols	0	2	0	0	2
Yellow soils	Alisols	3	7	0	0	10
Brown Forest soils	Cambisols	0	1	1	0	2
Total		82	76	4	2	164

^a By the Classification of Cultivated Soils in Japan (Third approximation).

^b By the WRB classification (only the dominant type is indicated).

laboratory measurements obtained with a laboratory sensor (Bartington MS2B, Bartington Instruments Ltd., UK). They tested forest soils that had an organic layer of varying depth above the mineral soil. The relationship between the two measurements was statistically significant ($r^2 = 0.73$, $n = 125$) but was not entirely satisfactory. The soil volume measured by the MS2D field probe was about 2300 cm³ (Lecoanet et al., 1999), whereas the soil volume sampled from each site for the laboratory measurement was about 15 cm³ (Boyko et al., 2004). It is likely that some of the scattered plots are due probably to the different scales of measurement, as well as to the microscale heterogeneity of the soils measured. Boyko et al. (2004) summarized the factors affecting the repeatability of field measurement of soil κ as follows; inhomogeneities within the place measured, positional precision of navigation, different equipment sets used, plant cover or plant litter, and anthropogenic activity. Maier et al. (2006) evaluated the repeatability of field measurements using the MS2D field probe and found that the influence of the diamagnetic contribution of water on soil κ was masked by other site-specific factors such as the degree of the contact between the measurement probe and the soil. Several researchers also reconfirmed the beneficial effects of removing plant litter before measuring surface soil κ in the field (Cervi et al., 2014; Declercq et al., 2019). Comparisons between field and laboratory measurements have been carried out with different instruments. The discrepancies between field and laboratory measurements are attributable to multiple factors. Thus, it is uncertain which of these is most important in obtaining repeatable data under field conditions.

In contrast with field measurements, laboratory measurements are more controllable and repeatable. The κ value is measured nondestructively, which allows the sample to be reused for additional laboratory analyses. Soil κ measured in the laboratory reflects the content of magnetic minerals such as magnetite and maghemite if the effects of particle size and shape on κ are negligibly small (Mullins, 1977; Lecoanet et al., 1999). Magnetic materials in soil originate from the parent material, by pedogenesis, and by human activities (Boyko et al., 2004; Grimley et al., 2004). If the contributions of the parent material and pedogenesis are small, measurement of κ can be useful for estimating anthropogenic soil pollution with heavy metals, as reviewed by Magiera et al. (2019). On the other hand, if the contributions are large, measurement of κ can be useful for classifying soils according to their parent material and pedogenesis (Dearing et al., 1996; Grimley et al., 2004; Hannam and Dearing, 2008; Blundell et al., 2009; Mileti et al., 2013; Cervi et al., 2014; Vingiani et al., 2014; Jordanova et al., 2016; De Mello et al., 2020). For example, soil magnetic susceptibility was significantly positively correlated with total Fe content in surface soils collected throughout England and Wales (Dearing et al., 1996; Blundell et al., 2009). Vingiani et al. (2014) reported that it was positively

correlated with $Al_0 + 1/2Fe_0$ (one of the requirements for andic soil properties, derived from acid-oxalate extractable Al and Fe) in eight soil profiles from southern Italy. However, there is no clear boundary between pedogenesis and human activities, because the pedogenesis is also caused by local field management such as irrigation of paddy fields that leads to the formation of man-made soils. Soil reduction due to anaerobic incubation in the laboratory may cause a sharp decrease of magnetic susceptibility within only a few months after the incubation begins, which is attributable to the dissolution of fine-grained pedogenic ferrimagnetic minerals under anaerobic conditions with influence from iron-reducing bacteria (Lu et al., 2012). Hannam and Dearing (2008) summarized the relationships between soil κ and the soil forming factors by dividing the spatial scales into national, regional, and local scales. They emphasized the need for evaluating soil-forming factors, like topography, that may cause soil κ variations at a local scale. This is because topographic factors can cause the small-scale variation of soil magnetic properties by redistributing soil magnetic particles from eroding sites to depositional sites (Quijano et al., 2014).

In Japan, researchers have measured soil κ to estimate the past and present environments such as aerial deposition and soil pollution (Torii and Fukuma, 1998; Torii, 2005; Kawasaki et al., 2017). On the aerial deposition, measurement of magnetic susceptibility of Lake Biwa sediment was useful to estimate the degree of aerial deposition of volcanic ash (Yoshikawa et al., 1993). As for soil pollution, Sunaga and Harada (2016) studied the contamination of forage by soil containing radiocesium from the Fukushima nuclear accident by measuring the content of magnetic minerals in forage samples and relating the magnetic content with the degree of contamination of forage with soil.

It is well known that agricultural soils in Japan developed under humid temperate climatic conditions, with variations due to local soil forming factors such as parent material, topography, and human activities (Yanai et al., 2012). Many of the soils in Japan have been influenced by volcanic activity to various degrees. Approximately 31% of the soils in Japan are classified as Andosols according to the World Reference Base for Soil Resources (WRB) (FAO, 2015; Kanda et al., 2018). Furthermore, about half of the agricultural land is used as irrigated paddy fields for staple rice production (Katayama et al., 2015), and 13% of the soils in Japan are classified as Fluvisols by the WRB (Kanda et al., 2018). Although Torii (2005) suggested that many soils in Japan are influenced by deposition of volcanic ash, which has high magnetic susceptibility, there is little data on the κ value of agricultural soils and the factors affecting it. Thus, it is unknown how useful soil κ measurements may be for estimating other soil properties, as far as agricultural surface soils in Japan are concerned. Especially unknown is the scale-dependent relationship between soil κ and other soil properties.

In this study, we measured the κ value of agricultural soils in Japan

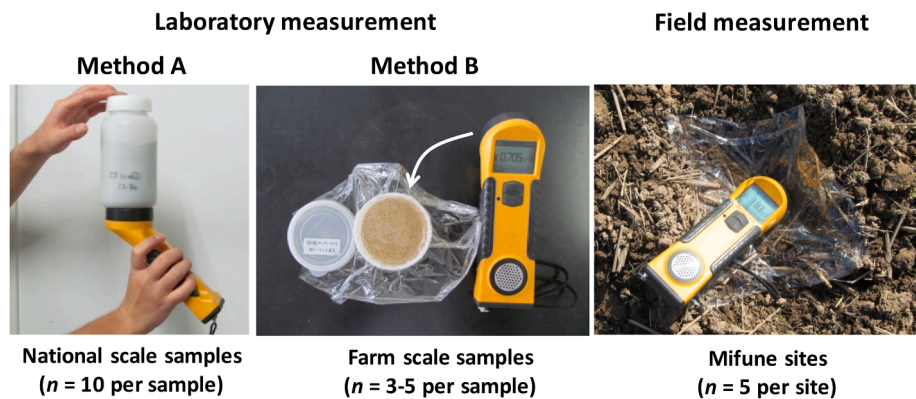


Fig. 1. Methods used in this study for measuring soil κ with a KT-10 m.

sampled at the national scale and the farm scale. The specific objectives were (1) to evaluate the repeatability of soil κ measured in the laboratory, (2) to examine the scale-dependent relationship between soil κ and other soil properties, and (3) to estimate the factors affecting the repeatability of soil κ measurements in the field.

2. Materials and methods

2.1. Soil samples

Surface soil samples were collected at the national scale and the farm scale. The national scale samples ($n = 180$) were collected during 1998–1999 (Sano et al., 2004; Yanai et al., 2012) from the plow layer (<15 cm) of fields throughout Japan. Among the samples analyzed by Yanai et al. (2012), samples with remaining weight greater than 100 g ($n = 164$) were selected for this study. The numbers of samples in relation to soil type and land use are listed in Table 1. According to the Classification of Cultivated Soils in Japan (Cultivated soil classification committee, 1995), the dominant soil type was Lowland Paddy soils ($n = 40$), followed by Andosols ($n = 34$), Gley Lowland soils ($n = 24$), and Brown Lowland soils ($n = 18$). Volcanogenous Regosols, Wet Andosols, Non-allophanic Andosols, and Andosols by this domestic classification correspond to Vitric Andosols, Gleyic Andosols, Aluandic Andosols, and Silandic Andosols, respectively, by the WRB (Table 1), and the WRB terms for Andosols are also used in this paper. The land uses of the selected samples were paddy fields ($n = 82$), upland fields ($n = 76$), grassland ($n = 4$) and fallow land ($n = 2$). Half of the sampling sites were in use as paddy fields at the time of sampling. These sites also included 23 paired paddy and upland fields that were located near each other, e. g., at the same agricultural experiment station. Although it is difficult to show the degree of representativeness of the national-scale samples, the percentage of samples classified as Vitric Andosols, Gleyic Andosols, Aluandic Andosols, and Silandic Andosols in our sample set (29.9%, Table 1) was close to the percentage of the distribution area of the four types of Andosols in the Japanese agricultural fields (28.1%) (Kanda et al., 2017).

Two farm-scale sample sets were used, referred to as the Takatsuki and Mifune samples. The Takatsuki samples were collected from five contiguous paddy fields at the former Kyoto University Farm, which had existed in Takatsuki city, Osaka prefecture, Japan (34°51'N, 135°37'E, about 7 m a.s.l.). The Takatsuki soils were classified as a Gley Lowland soil (Cultivated soil classification committee, 1995), one of the Fluvisols by the WRB. Each rectangular field was divided into 10 × 10 m plots, and soil samples were collected from the plow layer (<10 cm), for a total of 246 samples. Samples were collected in November 2013 after the rice crop had been harvested. Details of the sampling fields and sampling methods are described in Moritsuka et al. (2019). Forty samples were also collected from a nearby orchard in Kyoto University Farm for comparison with the paddy soils.

The Mifune samples were collected from eight adjacent paddy fields in Mifune town, Kamimashiki district, Kumamoto prefecture, Japan (32°43'N, 130°47'E, about 14 m a.s.l.). Sampling was carried out three times to evaluate the temporal stability of the results. The soil was classified as a Lowland Paddy soil (Cultivated soil classification committee, 1995), one of the Anthrosols by the WRB. All fields were located within an area of 700 × 700 m and were managed by the same farmer. Each field had an area of 0.15–0.3 ha. Many of the fields have been used for double cropping of rice and wheat. Five soil samples were collected from each field, from the center and the four surrounding corners, except for one triangular field (center and three corners), for a total of 39 soil samples per sampling. At each sampling site, four subsamples were collected from the plow layer (<15 cm) within a 5-m circular area and were combined into a single sample. Sampling was carried out three times, in May of 2014, 2016, and 2018, when the wheat crops were at the early ripening stage.

All soil samples were air-dried, crushed, and sieved through 2-mm mesh before measurement and analysis.

2.2. Laboratory measurement of magnetic susceptibility

The volume magnetic susceptibility κ was measured with a handheld magnetic susceptibility meter (KT-10 v2, Terraplus Inc., Ontario, Canada) designed for field measurement. This instrument uses a 10 kHz oscillator with a circular inductive coil with a diameter of 65 mm. The coil is placed on a flat soil surface and soil κ is measured to a depth of about 2 cm below the surface (De Mello et al., 2020). The effect of temperature on the frequency of the oscillator is minimized by measuring the κ value in the air before and after sample measurement. Before sample measurement, the accuracy of the instrument was checked by measuring a standard Mn-Zn ferrite sample with a known κ value (36×10^{-3} SI). The measured κ values were used as acquired, i.e., not converted to mass units ($\text{m}^3 \text{kg}^{-1}$) by dividing κ by the bulk density, to facilitate comparison between laboratory and field measurements. The bulk density of the farm-scale samples ranged from 1.05 to 1.30 g cm^{-3} for the Takatsuki samples and from 1.04 to 1.25 g cm^{-3} for the Mifune samples, and the κ values were strongly positively correlated with the mass magnetic susceptibility values with the r values of 0.96 and 1.00 for the Takatsuki and Mifune samples, respectively.

Two different methods were used for measuring soil κ in the laboratory. For the national scale samples, more than 100 g of a sample was packed into a 500-mL plastic bottle with an outer diameter of 78 mm (Product no. 02085, Sanplatec Co., Ltd., Osaka, Japan). The κ value was measured by placing the coil flat against the bottom of the bottle. This was done 10 times per sample, with mixing of the sample between measurements to renew the sample surface (Method A, Fig. 1).

For the farm scale samples, 100 g of a sample was packed into a 180-mL plastic jar with the outer and inner diameters of 85 and 74 mm, respectively (Product no. K142, Nakaya Kagaku Sangyo Co., Ltd, Osaka,

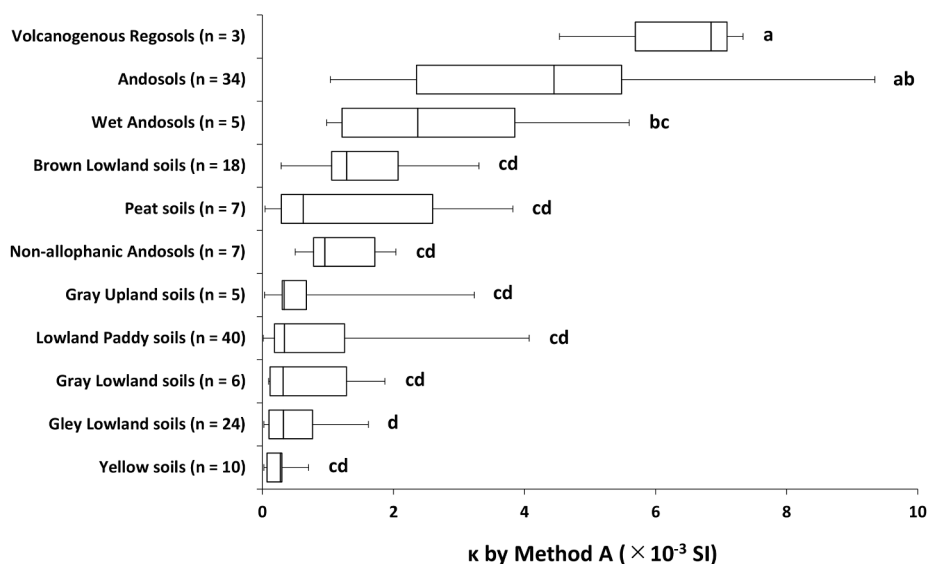


Fig. 2. Box plot of the κ values of the national scale samples. Among the 14 soil types, the box plots of κ of the 11 soil types including more than two samples are presented in the descending order of the mean κ values. The vertical line within each box indicates the median value. The left and right limits of each box are the 25th and 75th percentiles, respectively, and the left and right whiskers represent minimum and maximum values, respectively. Plots marked with the same letters are not significantly different at $p < 0.01$ (Tukey-Kramer test).

Japan). The κ value of the sample was measured by placing the coil of the KT-10 on the flat sample surface three or five times per sample, with mixing of the sample between measurements to renew the sample surface (Method B, Fig. 1). The number of measurements per sample was reduced to three for some of the Mifune samples, because the measurements were stable. A piece of plastic wrap was placed between the sample and the coil to prevent cross contamination. The amount of sample placed in the jar (100 g) was based on a preliminary experiment in which the κ value of a soil sample was measured repeatedly as the sample weight was raised in 10-g increments (Fig. S1).

To identify the source of magnetic susceptibility in soil, about 5 kg of a 2-mm sieved soil sample from one of the Mifune fields was subjected to sequential wet sieving using nylon mesh cloth with different openings (15, 48, 150, and 600 μm ; Fig. S2) as described by Moritsuka et al. (2015). Soil organic matter was not removed before the sieving treatments. The sample was divided into $< 15 \mu\text{m}$, 15–48 μm , 48–150 μm , 150–600 μm , and 600–2000 μm fractions by this method, and each fraction was collected by sedimentation and oven-dried at 60 °C. Sodium chloride was added to the finest fraction ($< 15 \mu\text{m}$) to accelerate sedimentation. The κ value of all but 600–2000 μm fractions were measured by Method B.

It should be noted here that the κ values obtained with Methods A and B were different. This is mainly because the samples measured by Method A were not in contact with the coil as closely as those measured by Method B. A preliminary experiment using 12 of the Mifune samples indicated that the κ values obtained from these two methods were strongly correlated with each other ($r^2 = 0.99$, $n = 12$) (Fig. S3). The values obtained from Method A were 58% of those from Method B on the average.

2.3. Field measurement of magnetic susceptibility

The same instrument used in the laboratory analyses was also used for the field measurements to facilitate comparison between the two sets of data. The κ value of the soil surface was measured at the same sites where the Mifune samples were collected. At these 39 sites, field measurement was carried out three times, in January of 2017, 2018, and 2020, when the surface soil had been plowed for wheat production. Winter was selected because the surface soil was homogeneous and weed growth was negligible (Fig. 1). The procedure was similar to Method B for laboratory measurement. After trampling the ground to make the measuring point flat and compacted, the κ value of the soil surface was measured five times per site within a 5-m circular area

(Fig. 1). Any plant residues present on the surface, especially rice residues, were removed from the measuring point before measurements were made. A piece of plastic wrap was placed between the soil surface and the coil to prevent cross contamination. Five measurements per site were recorded and their average was regarded as the κ value for that site.

To simulate the factors affecting the spatiotemporal variations of κ , a small artificial paddy field was prepared in an open top plastic box (W130 × D86 × H46 cm). The box was filled with the same Mifune soil as used for the nylon mesh sieving described previously and was placed in a greenhouse to protect the soil from rainfall. Rice was grown from May to September in the box as in the fields. Soon after the rice was harvested, the κ value of the soil surface was measured at 10 fixed points (Fig. S4). The soil κ was measured 41 times from October 3 to December 17, 2019, as the soil gradually dried. At every κ measurement date, the soil moisture content was also measured by the gravimetric method, using surface soil samples.

2.4. Analysis of other soil properties

Several other soil properties that may affect κ were also analyzed, in all except the Takatsuki samples. For the Takatsuki samples, we did not complete the whole analysis because of excessive within-sample variations in some of their κ values, as discussed in Section 3.2.

The other analyses focused on soil texture and the various forms of Fe and K. The properties measured in both national scale and Mifune samples included particle-size distribution (sand, silt, and clay), acid oxalate extractable Fe (Fe_o), citrate-dithionite extractable Fe (Fe_d), total Fe (Fe_t), exchangeable K (Ex-K), boiling HNO_3 extractable K ($\text{HNO}_3\text{-K}$), total K, acid oxalate extractable Al (Al_o), and cation exchange capacity (CEC). The content of structural Fe ($\text{Fe}_t - \text{Fe}_d$), which includes silicate-bound Fe (Torrent and Cabedo, 1986), was calculated by subtracting Fe_d from Fe_t . The content of nonexchangeable K (Nonex-K) was calculated by subtracting Ex-K from $\text{HNO}_3\text{-K}$. The K-saturation percentage (Ex-K/CEC) was calculated by dividing Ex-K by CEC, using the same unit ($\text{cmol}_c \text{kg}^{-1}$) for both properties. The content of $\text{Al}_o + 1/2\text{Fe}_o$, which is one of the requirements for andic soil properties (FAO, 2015), was also calculated.

For the national scale samples, we used published data for particle size distribution (sand, silt, and clay) based on the ISSS classification (Sano et al., 2004), Fe_t and total K (Yanai et al., 2012), Ex-K, $\text{HNO}_3\text{-K}$, Fe_o , and Al_o (Kitagawa et al., 2018). The analytical methods are described in the cited papers, and the same methods were used to analyze the Mifune samples. For the remaining properties, Fe_d was

Table 2
Descriptive statistics of the κ values of the soil samples measured with reference to the sampling scale and soil type.

	Sample type	Nos. of samples	Median (10^{-3} SI)	Mean	Maximum	Minimum	Coefficient of variation (%)
National scale (Method A)	All samples	164	1.07	1.77	9.34	0.01	112
	Peat soils ^a	7	0.63	1.47	3.82	0.04	107
	Sand-dune Regosols ^a	1	1.54	1.54	n.d.	n.d.	n.d.
	Volcanogenous Regosols ^a	3	6.85	6.24	7.33	4.53	24
	Wet Andosols ^a	5	2.37	2.80	5.60	0.98	69
	Non-allophanic Andosols ^a	7	0.95	1.21	2.04	0.50	51
	Andosols ^a	34	4.45	4.38	9.34	1.04	45
	Lowland Paddy soils ^a	40	0.34	0.82	4.07	0.01	119
	Gley Lowland soils ^a	24	0.32	0.49	1.62	0.02	102
	Gray Lowland soils ^a	6	0.32	0.71	1.87	0.10	112
	Brown Lowland soils ^a	18	1.29	1.54	3.30	0.29	54
	Gray Upland soils ^a	5	0.33	0.91	3.23	0.03	144
	Dark Red soils ^a	2	0.87	0.87	1.33	0.42	74
	Yellow soils ^a	10	0.27	0.24	0.71	0.02	83
Brown Forest soils ^a	2	1.13	1.13	2.22	0.05	136	
Farm scale (Method B)	Takatsuki samples	246	0.16	0.17	0.26	0.13	13
	Mifune samples in 2014	39	3.25	3.39	6.77	2.07	33
	Mifune samples in 2016	39	3.07	3.38	7.65	2.08	36
	Mifune samples in 2018	39	3.11	3.32	7.77	2.00	37

n.d.: no data.

^a By the Classification of Cultivated Soils in Japan (Third approximation).

Table 3
Correlation matrix of the soil κ measurements compared with several other soil properties of the national scale and Mifune samples.

	Sand	Silt	Clay	Fe _o	Fe _d	Fe _t	Fe _t - Fe _d	
National scale samples (n = 125)	0.27	-0.26	-0.21	0.46	0.24	0.64	0.74	
Mifune samples in 2014 (n = 39)	0.88	-0.80	-0.76	n.d.	n.d.	0.29	n.d.	
Mifune samples in 2016 (n = 39)	0.88	-0.82	-0.72	n.d.	n.d.	0.33	n.d.	
Mifune samples in 2018 (n = 39)	0.89	-0.87	-0.76	0.63	-0.04	0.22	0.33	
	Ex-K	HNO ₃ -K	Nonex-K	Total K	CEC	Ex-K/CEC	Al _o	Al _o + 1/2Fe _o
National scale samples (n = 125)	0.23	-0.19	-0.50	-0.59	0.34	0.06	0.69	0.68
Mifune samples in 2014 (n = 39)	0.61	0.66	0.64	0.08	-0.25	0.70	n.d.	n.d.
Mifune samples in 2016 (n = 39)	0.55	0.62	0.67	0.08	-0.29	0.66	n.d.	n.d.
Mifune samples in 2018 (n = 39)	0.61	0.63	0.54	0.02	-0.29	0.68	0.34	0.55

n.d.: no data.

Bold values indicate significant correlation at 1% level.

Values significant at 1% are 0.23 (national scale samples) and 0.41 (Mifune samples).

Fe_o: acid oxalate extractable Fe; Fe_d: dithionite-citrate extractable Fe; Fe_t: total Fe; Ex-K; exchangeable K; Nonex-K; nonexchangeable K.

CEC: cation exchange capacity; Al_o: acid oxalate extractable Al.

extracted from the finely ground samples with a mixture of sodium dithionite and sodium citrate (Loeppert and Inskeep, 1996). CEC was measured by the ammonium acetate (pH 7) method using the batch technique (Sumner and Miller, 1996). For the Mifune samples, after decomposition of organic matter with hydrogen peroxide, the contents of coarse sand (0.2–2 mm) and fine sand (0.02–0.2 mm) fractions were measured by the conventional sieving method.

2.5. Statistical analysis

The data were subjected to descriptive statistics, correlation analysis, regression analysis, Welch's *t*-test, and one-way ANOVA with Tukey-Kramer multiple comparison test using Microsoft Excel 365 for Windows and Excel Toukei (BellCurve for Excel) (Social Survey Research Information, Tokyo, Japan). Within-sample or within-site variation of the κ value was estimated from the coefficient of variation (CV) of the repeated measurements per sample or site, respectively. A CV below 10% was used as an allowable limit for assuring the reliability of κ in both laboratory and field measurements by referring to the CV applied to soil analysis by portable X-ray fluorescence (Ravansari et al., 2020).

3. Results and discussion

3.1. Descriptive statistics of the κ value of agricultural soils in Japan

There was considerable variability among the κ values of the samples (Table 1), especially among the national scale samples measured by Method A (Fig. 1). In the national scale samples, κ had a median of 1.07×10^{-3} SI and a mean of 1.77×10^{-3} SI. The CV for the κ values of all national scale samples was 112%. The lower median than the mean indicated that the distribution of κ was skewed to the right with the skewness of 1.53. Among the soil types, the CV for the κ values also exceeded 100% for 6 soil types that showed relatively low κ values. Box plots of the κ values of each soil type further indicated that Volcanogenous Regosols and Andosols, corresponding to Vitric Andosols and Silandic Andosols respectively by the WRB, had significantly higher κ values than other soil types (Fig. 2). Wet Andosols (Gleyic Andosols by the WRB) showed the third highest mean and median. However, the κ values were rarely significantly different among the soil types other than Volcanogenous Regosols and Andosols, which is due probably to the large CVs for κ of each soil type (Table 2). These results suggest that κ may be useful for rapidly discriminating Vitric or Silandic Andosols from other soils, as far as surface agricultural soils in Japan are concerned. Although the effects of land use on κ were mostly masked by the effects

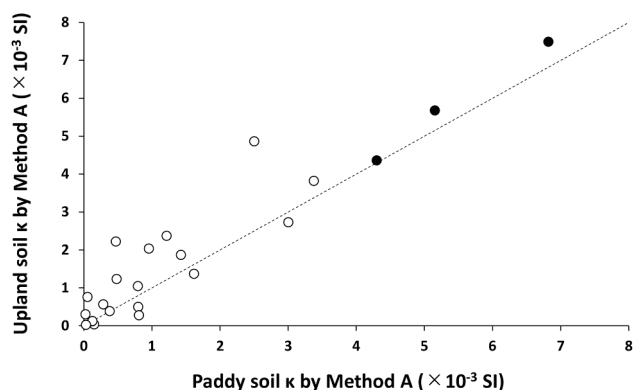


Fig. 3. Comparison of the κ values of paired paddy and upland soils collected from nearby fields ($n = 23$) as part of the national scale samples. The plots with a filled circle ($n = 3$) indicate soil pairs that were classified as Silandic Andosols. The dotted line in the figure is a 1:1 line.

of soil type, the κ value of the upland soil was higher than that of the neighboring paddy soil in 15 out of 23 sample pairs (Fig. 3). This implies that land use affected κ to some extent.

In the farm scale samples, the Takatsuki samples had much lower κ values than the Mifune samples. The relatively high κ of the Mifune samples would be due to volcanic ash deposition, as the soils on the hills near the sampling locations were classified as Silandic Andosols. The κ values of the Mifune samples were relatively stable over several years. The CVs of both sets of samples were smaller than that of the national scale samples (Table 1). Nevertheless, both sets had variations, suggesting that the measurement of soil κ may be helpful for classifying soils at the farm scale.

At the farm scale, multiple samples from each field were collected and analyzed separately, enabling calculation of within-field variations. The CVs of soil κ in each field were 7.1–16.7% in the Takatsuki fields, and 2.0–26.7% in the Mifune fields. Previous reports showed that the CVs of soil magnetic susceptibility within a field varied from 0.3 to 3.9% (Sunaga and Harada, 2016) to 10.9% (Sunaga and Harada, 2016) to

34.1% (Quijano et al., 2014). These results suggest that the within-field variation of soil magnetic susceptibility is not negligibly small in many cases.

3.2. Within-sample variation of soil κ measured in the laboratory

The KT-10 v2 user’s guide specifies that the calibration sample should be about 5 cm thick if it is a powdered material such as fine gravel or sand. This implies that the KT-10 can detect magnetic material within 5 cm below the surface of the sample. As shown in Fig. S1, however, soil κ did not increase linearly with an increase of soil weight, suggesting that the response of KT-10 to a magnetic material decreases exponentially with the distance between the coil and the material. In the case of KT-5 Kappameter (Geofizika, Brno, Czech Republic), which is an earlier version of KT-10 and has similar specifications, 90% of the signal detected by the instrument originated within 2.3 cm from the sample surface (Lecoanet et al., 1999). The similarly shallow penetration depth of the KT-10 enabled the evaluation of within-sample variations of soil κ here.

In this study, the repeatability of the laboratory measurement was evaluated by measuring the κ value several times per sample, with mixing between measurements to renew the sample surface. Fig. 4 shows the relationship between the average κ value for each sample and the repeatability of the measurement. It is useful to remember that regardless of the magnitude of the κ value, a high CV means high within-sample variation and therefore low reliability of the measurement.

In the national scale samples, the CV was usually <10% if the average κ value was higher than 0.2×10^{-3} SI (Fig. 4). There were 132 samples (80%) with average κ higher than 0.2×10^{-3} SI. Below this κ value, the CV increased with decreasing κ . This relationship agrees with the findings of Lee and Morris (2013), who measured the κ value of rock samples with KT-10 and found that the CV from repeated measurement of the same sample increased when the sample measurements became lower than 0.1×10^{-3} SI. This suggests that the variability of the results here may be more attributable to the limits of the sensitivity of the instrument than to actual within-sample heterogeneity.

In the farm scale samples, the results were different between the two

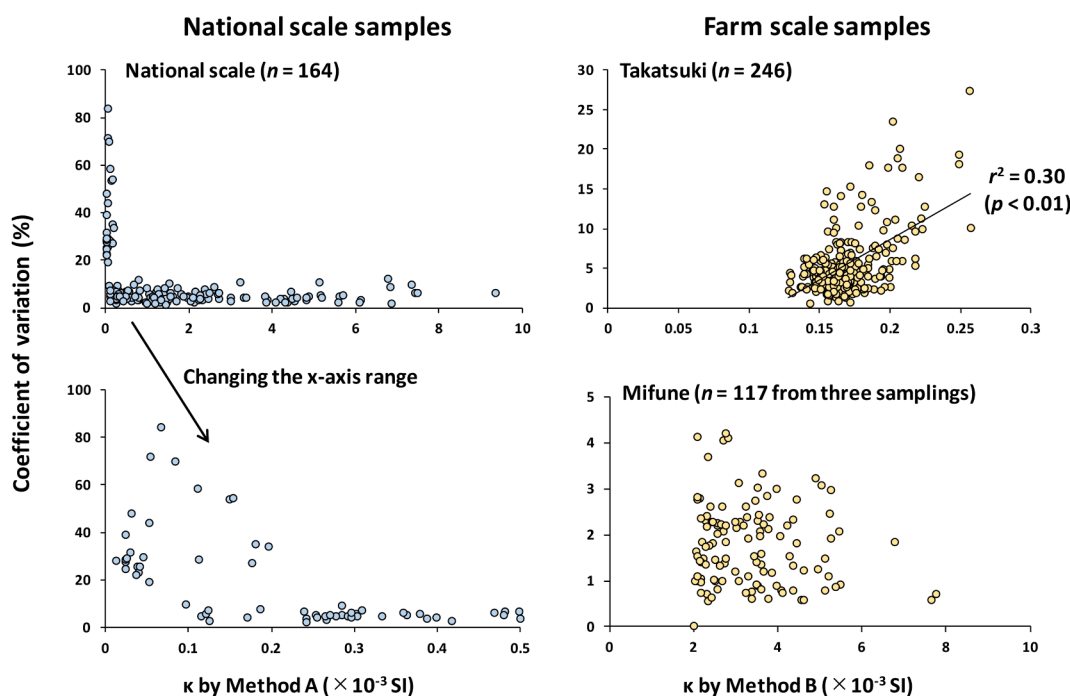


Fig. 4. Relationship between the average κ measurement of each sample and the repeatability of the sample measurements. A line in the figure of the Takatsuki samples indicates a linear regression line.

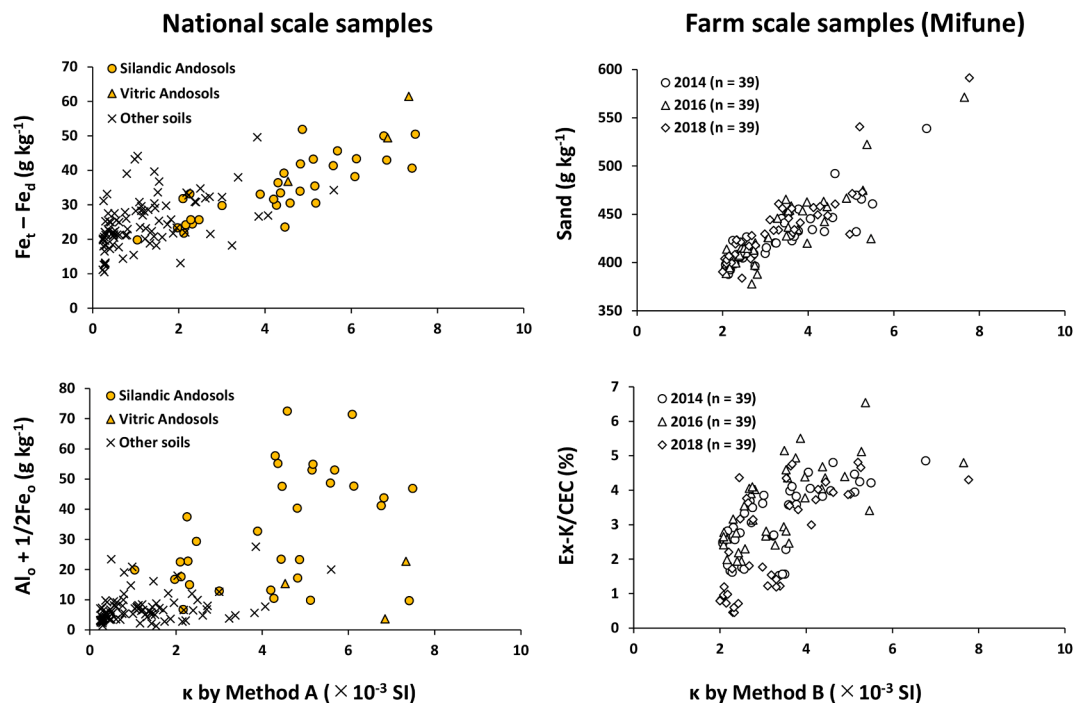


Fig. 5. Relationship between κ and the most strongly correlated soil properties. The national scale samples used included Silandic Andosols ($n = 32$), Vitric Andosols ($n = 3$), and other soils ($n = 90$). Abbreviations used for the soil properties are spelled out in the footnotes of Table 3.

sample sets (Fig. 4). In the Takatsuki samples, the κ values were between 0.1 and 0.3×10^{-3} SI, and the CV was higher than 10% in 28 of the 246 samples. Interestingly, the higher κ value had higher CVs ($r^2 = 0.30$, $p < 0.01$), opposite to the observations in the national scale samples. In addition, the κ values from the orchard in Kyoto University Farm were significantly higher ($p < 0.01$, Welch's t -test) than those of the Takatsuki samples collected nearby, and the orchard sample κ values were positively correlated with their CVs ($r^2 = 0.69$, $p < 0.01$) (Fig. S5). These results are attributable to the presence of highly magnetic particles that are sparsely distributed within a sample, creating magnetic hotspots. In contrast, in the Mifune samples, the CV was always $< 5\%$, and there was no significant relationship between soil κ values and their CVs (Fig. 4).

Repeated measurements of κ of the Takatsuki samples suggested that this is a useful technique for detecting magnetic hotspots in a sample. Further research is needed to reveal how magnetic hotspots are created at the within-sample scale.

3.3. Correlations among soil κ and other properties

The relationship between soil κ and other soil properties in the national scale and Mifune samples was evaluated. For the national scale samples, the correlation analysis was performed by using the data of only 125 samples, i.e., those that had average κ higher than 0.2×10^{-3} SI ($n = 132$) together with a complete dataset of the other soil properties. None of the Takatsuki samples were included, because the high κ values measured in some Takatsuki samples were considered unreliable.

Table 2 shows the correlation coefficients between soil κ and other soil properties. The results were different between the two sample sets. In the national scale samples, κ was positively correlated with sand, all Fe forms, Ex-K, CEC, Al_o , and $Al_o + 1/2Fe_o$, and it was negatively correlated with silt, Nonex-K, and total K. In the Mifune samples, κ was positively correlated with sand, Ex-K, HNO_3 -K, Nonex-K, and Ex-K/CEC, and it was negatively correlated with silt and clay. The κ value in the Mifune samples was also positively correlated with Fe_o and $Al_o + 1/2Fe_o$ in 2018.

At the national scale, structural Fe ($Fe_t - Fe_d$) was most positively correlated with κ (Table 2). Structural Fe includes silicate-bound Fe

(Torrent and Cabedo, 1986) and possibly coarse-grained lithogenic magnetite (Rennert, 2019). The positive correlation between structural Fe and κ at the national scale may be due to the relatively high concentration of magnetite in andesitic and dacitic volcanic ashes (1.2–6.5%) that are common in Japan (Shoji et al., 1987). In the Mifune samples collected at the farm scale, Fe_o rather than structural Fe was positively correlated with κ (Table 2). As Fe_o includes various types of short-range ordered secondary Fe minerals (Rennert, 2019), it was difficult to interpret why Fe_o was more strongly correlated with κ in the case of the Mifune samples collected at the farm scale. The relationship between soil κ and Fe_t content in the national scale samples agreed with those from England and Wales (Dearing et al., 1996; Blundell et al., 2009). According to Cervi et al. (2019), magnetic susceptibility of surface soils in Brazil usually increased by the removal of secondary Fe oxides using extraction with the dithionate-citrate-bicarbonate solution, suggesting that structural Fe rather than Fe_d was the main source of magnetic materials in their case.

Regarding the relationship between soil κ and $Al_o + 1/2Fe_o$, significantly positive correlations were observed in both sample sets (Table 2). In Italy, where soils with andic properties are distributed widely, soil magnetic susceptibility was positively correlated with $Al_o + 1/2Fe_o$ at the regional scale ($r = 0.79$, $n = 41$) (Vingiani et al., 2014). However, the correlation became lower at the national scale ($r = 0.31$), possibly due to distribution of soils developed over basaltic materials having high magnetic susceptibility regardless of their andic properties (Mileti et al., 2013). As the content of $Al_o + 1/2Fe_o$ in soil is positively correlated with the phosphorus adsorption capacity, these results further suggest that soil κ can predict the phosphorus adsorption capacity of soil in some circumstances. For example, Poggere et al. (2020) reported that soil magnetic susceptibility was weakly but positively correlated with the phosphorus adsorption capacity ($r = 0.50$, $p = 0.01$) in the B horizon of red soils from basic igneous rocks in Brazil.

Regarding the relationship between soil κ and the K-related properties, the results from the national scale samples were contrasting to those from the Mifune samples (Table 2). The national scale samples had a significantly negative correlation with total K. This may be due to the relatively low total K content of Vitric or Silandic Andosols (Kitagawa

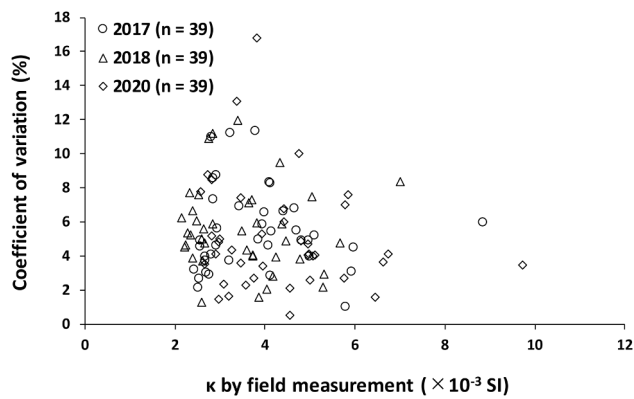


Fig. 6. Relationship between the measured κ values (averages of 5 measurements) at the 39 Mifune sites and the repeatability of these measurements.

et al., 2018) that had significantly higher κ values (Fig. 2). In the Mifune samples, the significantly positive correlations with the K saturation percentage may due to a significant positive correlation between sand content and the K saturation percentage ($r = 0.46\text{--}0.55$).

Some of the correlation coefficients listed in Table 2 were higher than 0.63 or lower than -0.63 , indicating that the soil κ could predict more than 40% of the variation of other related properties. Such properties include structural Fe ($\text{Fe}_t - \text{Fe}_d$) and $\text{Al}_o + 1/2\text{Fe}_o$ in the national scale samples, and sand and the percentage of K saturation (Ex-K/CEC) in the Mifune samples (Fig. 5). The high correlation with structural Fe owed to Vitric Andosols and Silandic Andosols that were relatively high in both structural Fe and κ . Furthermore, a positive y-intercept suggested that soil κ was originated from a part of the structural Fe components.

Among these relationships, the relationship between κ and the sand content at the Mifune samples was strongest. This relationship was examined by fractionating one of the Mifune samples by particle size and measuring the κ value of each fraction. The κ value of the original 2-mm sieved sample (4682 g) was $3.99 \times 10^{-3}\text{SI}$. The κ values of the fractions were $0.47 \times 10^{-3}\text{SI}$ for the $< 15 \mu\text{m}$ fraction (1787 g), $2.51 \times 10^{-3}\text{SI}$ for the $15\text{--}48 \mu\text{m}$ fraction (665 g), $7.55 \times 10^{-3}\text{SI}$ for the $48\text{--}150 \mu\text{m}$ fraction (776 g), and $11.93 \times 10^{-3}\text{SI}$ for the $150\text{--}600 \mu\text{m}$ fraction (605 g). Unfortunately, the κ value of the $600\text{--}2000 \mu\text{m}$ fraction (102 g) could not be measured due to severe contamination with light-fraction organic matter. Although the total weight recovered by the sequential sieving treatments was 84% of the original sample weight, the κ value estimated from the sum of the fractions ($3.96 \times 10^{-3}\text{SI}$) was almost equal to that of the original sample ($3.99 \times 10^{-3}\text{SI}$). The κ value of the measured fractions increased as the particle size increased. This suggests a causal relationship between the κ value and the sand content in this Mifune soil. To investigate this further, we determined the contents of coarse sand and fine sand fractions in all the Mifune samples. The κ values were correlated with the coarse sand content ($r = 0.76\text{--}0.81$) much more strongly than the fine sand content ($r = 0.17\text{--}0.32$). This suggests that the coarse sand fraction was the main source of magnetic materials in the Mifune samples.

It is unclear how far these local relationships between soil κ and texture can be extrapolated to other locations. The κ value of soils in Japan was weakly but positively correlated with sand content at the national scale (Table 2). Blundell et al. (2009) also reported that the log-transformed content of coarse sand ($500\text{--}2000 \mu\text{m}$) was significantly positively correlated with the log-transformed soil magnetic susceptibility in more than 5000 samples of surface soil collected throughout England and Wales. These results suggest that soil magnetic susceptibility was originated from primary minerals in soil. In contrast, the sand content of Brazilian soils collected from a relatively small area (<400 ha) was significantly negatively correlated with magnetic susceptibility (De Mello et al., 2020; Andrade et al., 2020), suggesting that magnetic

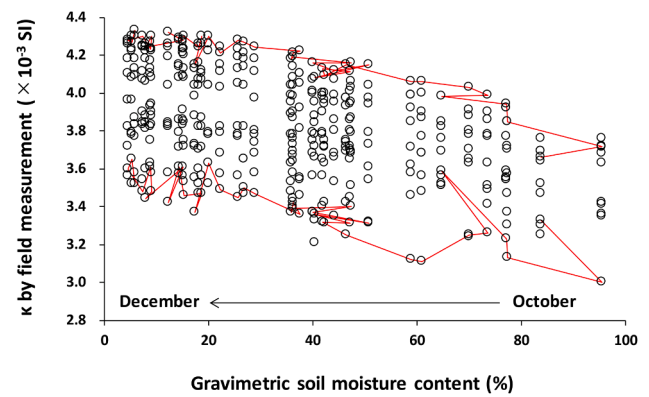


Fig. 7. The soil κ value at each 10 monitoring points in the artificial paddy field (Fig. S4) as the soil dried. The two red lines connect the time-series data of the two locations that usually produced the highest or lowest κ values.

susceptibility was originated from the secondary magnetic materials in the smaller particles.

3.4. Reliability of field measurement of soil κ

The instrument used in this study (KT-10) is designed for field measurement. It was used to measure the Mifune soils both in the field and in the laboratory to assess the reliability of field measurements. Fig. 6 shows the relationship between the average κ value at each site and the repeatability of the measurements. A high CV indicates high within-site variation. The CV of the field measurements exceeded 10% at 3 out of 39 sites every time κ was measured. The average CV was 5.5% in 2017, 5.6% in 2018, and 5.0% in 2020. Field measurements had lower repeatability (higher CV) than the laboratory measurements (Figs. 4 and 6). Interestingly, the CVs in the 2017 field measurements were weakly but positively correlated with those in 2018 ($r = 0.34$, $p < 0.05$). This implies that within-site variations in κ were influenced by the site-specific field environments. However, the coefficients of variation in 2020 were not significantly correlated with those in 2017 and 2018.

In our case, the CV below 10% was judged to be a feasible range to assure the reliability of field measurements. Schibler et al. (2002) evaluated the repeatability of surface soil κ by repeated measurement of κ at exactly the same place using a field probe (Bartington MS2D) and found that the coefficient of variation started to exceed 10% if the average κ value became $<0.1 \times 10^{-3}\text{SI}$. Their results suggest that the limit of quantification was around $0.1 \times 10^{-3}\text{SI}$ for the Bartington MS2D probe. Declercq et al. (2019) also evaluated the within-site variation of surface soil κ after removal of plant cover by measuring κ at 36 non-overlapping points within an area of about 1 m^2 using a Bartington MS2D probe according to the method of Maier et al. (2006). In their case, the coefficient of variation ranged from 11% (arable land and forest) to 14% (pasture), even if the measurements were higher than $0.1 \times 10^{-3}\text{SI}$.

To estimate the factors affecting the within-site variation, the κ value and the moisture content at the soil surface were monitored in a small artificial paddy field containing one of the Mifune soils. The κ value increased as the soil moisture content decreased (Fig. 7). The increase in κ (average of 10 measurement points) during drying from October to December was $0.58 \times 10^{-3}\text{SI}$. This increase may be due partly to drying-induced soil compaction. The κ value was also influenced by the location of the measurement points. The highest and lowest values were usually recorded at the same two points (Fig. 7). The average difference from highest to lowest κ among the 10 points on each measurement date was $0.78 \times 10^{-3}\text{SI}$. The average CV (7.1%) was comparable to that of the field measurements (Fig. 6). These two factors (moisture content and measurement location), especially the latter one, likely decreased the repeatability of the field measurement. As the soil was 2-mm sieved and

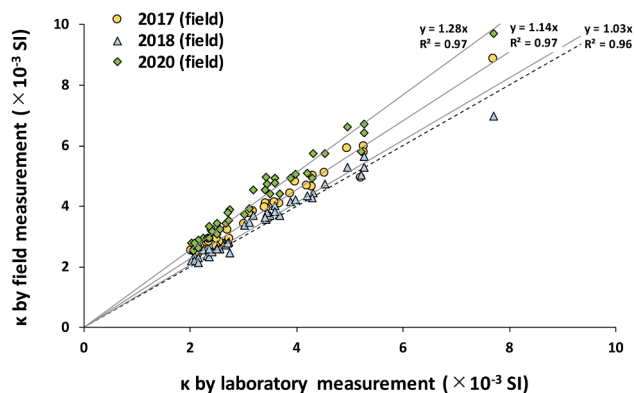


Fig. 8. Relationship between laboratory and field measurements of κ of the 39 Mifune soils. Three sets of field measurements were plotted against the laboratory-based measurements (averages of measurements from 2016 and 2018 samples). The dotted line in the figure represents a 1:1 line, and three solid lines are linear regression lines for each year's field data, with the y-intercepts forced to zero.

homogenized before placing in the box, the within-site variation may have been due to incomplete contact between the measurement surface and the coil of the instrument, as observed by Maier et al. (2006).

Despite the factors that may have decreased the reliability of field measurements, the determination coefficients between field and laboratory measurements were higher than 0.95 (Fig. 8). The correlation coefficients among the three sets of field measurements were also high ($r > 0.98$). However, the average κ values decreased in the following order; $4.32 \times 10^{-3}\text{SI}$ (2020), $3.84 \times 10^{-3}\text{SI}$ (2017), and $3.49 \times 10^{-3}\text{SI}$ (2018), although the κ values measured in the laboratory were stable over several years (Table 1). The difference between 2018 and 2020 ($0.83 \times 10^{-3}\text{SI}$) was larger than could be expected from the effects of soil moisture on the κ value (Fig. 7). These results suggest the presence of other factors causing a systematic error in the κ measurements of one year, which remain to be elucidated.

Nevertheless, the determination coefficients between field and laboratory measurements in our study ($r^2 > 0.95$) were much better than those reported previously (Boyko et al., 2004; Cervi et al., 2014; Quijano et al., 2014). This is due partly to the homogenization of surface soil by plowing and the absence of plant cover at the time of the field measurements. The correlation coefficients between the field-measured κ value at the Mifune sites in January 2018 and the laboratory-based data of soil samples collected in May 2018 were 0.89 for sand, -0.86 for silt, -0.76 for clay, and 0.67 for Ex-K/CEC. These correlation coefficients were very similar to those between the laboratory-measured κ values and these soil properties (Table 2). Accordingly, field measurement of soil κ was useful for rapidly estimating the related properties, at least at the Mifune sites. More comparative research is needed to confirm these results and to better establish field-based soil assessment.

4. Conclusions

The κ value of agricultural surface soils in Japan varied widely at the national scale. The soils classified as Vitric Andosols or Silandic Andosols had significantly higher κ values than other soils. Soil κ can be a useful proxy for the contents of structural Fe ($\text{Fe}_t - \text{Fe}_d$) and the andic soil properties ($\text{Al}_o + 1/2\text{Fe}_o$) at the national scale. Soil κ variations were also detectable at the much smaller scales. In some of the Takatsuki samples collected at the farm scale, large within-sample variations of soil κ suggested the presence of magnetic hotspots. In the Mifune samples collected at the farm scale, there was a causal relationship between sand content and soil κ . These results highlight the potential of soil κ measurement for rapid soil assessment. To ensure the validity of this method, the reliability of field measurements should be verified by

measuring the within-site variation. Use of the CV below 10% as an allowable limit for reliable measurements, as introduced in this study, should be tested in other sites and with other sensors to establish the field-based soil assessment method.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Dr. Tomohiro Nishigaki, JIRCAS, for providing valuable comments on an early version of the manuscript, and to Dr. Takashi Kawai, Kyoto University Farm (now at Okayama University) for allowing us to collect soil samples from an orchard in Kyoto University Farm.

Funding

This work was supported by JSPS KAKENHI Grant Numbers JP15K00634 and JP20K05766, and SATREPS (JST/JICA) Grant Number JPMJSA1608.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.geoderma.2021.115013>.

References

- Andrade, R., Silva, S.H.G., Faria, W.M., Poggere, G.C., Barbosa, J.Z., Guilherme, L.R.G., Curi, N., 2020. Proximal sensing applied to soil texture prediction and mapping in Brazil. *Geoderma Reg.* 23, e00321. <https://doi.org/10.1016/j.geodrs.2020.e00321>.
- Blundell, A., Dearing, J.A., Boyle, J.F., Hannam, J.A., 2009. Controlling factors for the spatial variability of soil magnetic susceptibility across England and Wales. *Earth Sci. Rev.* 95, 158–188. <https://doi.org/10.1016/j.earscirev.2009.05.001>.
- Boyko, T., Scholger, R., Stanjek, H., MAGPROX Team, 2004. Topsoil magnetic susceptibility mapping as a tool for pollution monitoring: repeatability of in situ measurements. *J. Appl. Geophys.* 55, 249–259. <https://doi.org/10.1016/j.jappgeo.2004.01.002>.
- Cervi, E.C., Costa, A.C.S., Souza Junior, I.G., 2014. Magnetic susceptibility and the spatial variability of heavy metals in soils developed on basalt. *J. Appl. Geophys.* 111, 377–383. <https://doi.org/10.1016/j.jappgeo.2014.10.024>.
- Cervi, E.C., Maher, B., Poloseli, P.C., Souza Junior, I.G., Costa, A.C.S., 2019. Magnetic susceptibility as a pedogenic proxy for grouping of geochemical transects in landscapes. *J. Appl. Geophys.* 169, 109–117. <https://doi.org/10.1016/j.jappgeo.2019.06.017>.
- Cultivated soil classification committee, 1995. Classification of cultivated soils in Japan. Third approximation. Miscellaneous publication of the National Institute of Agro-environmental Sciences (no. 17), Tsukuba, Japan.
- Dearing, J.A., Hay, K.L., Baban, S.M.J., Huddleston, A.S., Wellington, E.M.H., Loveland, P.J., 1996. Magnetic susceptibility of soil: an evaluation of conflicting theories using a national data set. *Geophys. J. Int.* 127, 728–734. <https://doi.org/10.1111/j.1365-246X.1996.tb04051.x>.
- Dearing, J., 1999. *Environmental Magnetic Susceptibility Using the Bartington MS2 System*, second edition. Bartington Instruments, Oxford.
- Declercq, Y., Samson, R., Castanheiro, A., Spassov, S., Tack, F.M.G., De Vijver, E.V., De Smedt, P., 2019. Evaluating the potential of topsoil magnetic pollution mapping across different land use classes. *Sci. Total Environ.* 685, 345–356. <https://doi.org/10.1016/j.scitotenv.2019.05.379>.
- De Mello, D.C., Demattè, J.A.M., Silvero, N.E.Q., Di Raimo, L.A.D.L., Poppiel, R.R., Mello, F.A.O., Souza, A.B., Safanelli, J.L., Resende, M.E.B., Rizzo, R., 2020. Soil magnetic susceptibility and its relationship with naturally occurring processes and soil attributes in pedosphere, in a tropical environment. *Geoderma* 372, 114364. <https://doi.org/10.1016/j.geoderma.2020.114364>.
- FAO, 2015. World reference base for soil resources 2014 International soil classification system for naming soils and creating legends for soil maps (Update 2015). World Soil Resources Reports 106, Food and Agriculture Organization of the United Nations, Rome, Italy. <http://www.fao.org/3/i3794en/i3794en.pdf>.
- Grimley, D.A., Arruda, N.K., Bramstedt, M.W., 2004. Using magnetic susceptibility to facilitate more rapid, reproducible and precise delineation of hydric soils in the midwestern USA. *Catena* 58, 183–213. <https://doi.org/10.1016/j.catena.2004.03.001>.

- Hannam, J.A., Dearing, J.A., 2008. Mapping soil magnetic properties in Bosnia and Herzegovina for landmine clearance operations. *Earth Planet. Sci. Lett.* 274, 285–294. <https://doi.org/10.1016/j.epsl.2008.05.006>.
- Jordanova, N., Jordanova, D., Petrov, P., 2016. Soil magnetic properties in Bulgaria at a national scale – Challenges and benefits. *Glob. Planet. Change* 137, 107–122. <https://doi.org/10.1016/j.gloplacha.2015.12.015>.
- Kanda, T., Takata, Y., Wakabayashi, S., Kohyama, K., Obara, H., 2017. Development of the 1:50,000 digital soil map of cultivated soil in Japan according to the comprehensive soil classification system of Japan – First approximation. *Jpn. J. Soil Sci. Plant Nutr.* 88, 29–34. https://www.jstage.jst.go.jp/article/dojo/88/1/88_880115/pdf-char/ja.
- Kanda, T., Takata, Y., Kohyama, K., Ohkura, T., Maejima, Y., Wakabayashi, S., Obara, H., 2018. New soil maps of Japan based on the Comprehensive Soil Classification Systems of Japan – First Approximation and its application to the World Reference Base for Soil Resources 2006. *Jpn. Agr. Res. Q.* 52, 285–292. <https://doi.org/10.6090/jarq.52.285>.
- Katayama, N., Baba, Y.G., Kusumoto, Y., Tanaka, K., 2015. A review of post-war changes in rice farming and biodiversity in Japan. *Agric. Syst.* 132, 73–84. <https://doi.org/10.1016/j.agry.2014.09.001>.
- Kawasaki, K., Horikawa, K., Sakai, H., 2017. Magnetic biomonitoring of roadside pollution in the restricted Midagahara area of Mt. Tateyama, Toyama, Japan. *Environ. Sci. Pollut. Res.* 24, 10313–10325. <https://doi.org/10.1007/s11356-017-8702-5>.
- Kitagawa, Y., Yanai, J., Nakao, A., 2018. Evaluation of nonexchangeable potassium content of agricultural soils in Japan by the boiling HNO₃ extraction method in comparison with exchangeable potassium. *Soil Sci. Plant Nutr.* 64, 116–122. <https://doi.org/10.1080/00380768.2017.1411168>.
- Lecoanet, H., Lévêque, F., Segura, S., 1999. Magnetic susceptibility in environmental applications: comparison of field probes. *Phys. Earth Planet. Inter.* 115, 191–204. [https://doi.org/10.1016/S0031-9201\(99\)00066-7](https://doi.org/10.1016/S0031-9201(99)00066-7).
- Lee, M. D., and Morris, W. A., 2013. Comparison of magnetic-susceptibility meters using rock samples from the Wopmay Orogen, Northwest Territories, Canada. *Geological Survey of Canada, Technical Note* 5. 10.4095/292739.
- Loeppert, R. H., Inskeep, W. P., 1996. Iron. In *Methods of Soil Analysis, Part 3, Chemical Methods*, ED. Sparks DL, et al., pp. 639–664. SSSA Book Series no. 5, WI, USA.
- Lu, S.G., Zhu, L., Yu, J.Y., 2012. Mineral magnetic properties of Chinese paddy soils and its pedogenic implications. *Catena* 93, 9–17. <https://doi.org/10.1016/j.catena.2012.01.002>.
- Magiera, T., Lukasik, A., Zawadzki, J., Rösler, W., 2019. Magnetic susceptibility as indicator of anthropogenic disturbances in forest topsoil: a review of magnetic studies carried out in Central European forests. *Ecol. Indic.* 106, 105518 <https://doi.org/10.1016/j.ecolind.2019.105518>.
- Maier, G., Scholger, R., Schön, J., 2006. The influence of soil moisture on magnetic susceptibility measurements. *J. Appl. Geophys.* 59, 162–175. <https://doi.org/10.1016/j.jappgeo.2005.10.001>.
- Mileti, F.A., Langella, G., Prins, M.A., Vingiani, S., Terribile, F., 2013. The hidden nature of parent material in soils of Italian mountain ecosystems. *Geoderma* 207–208, 291–309. <https://doi.org/10.1016/j.geoderma.2013.05.006>.
- Moritsuka, N., Izawa, G., Katsura, K., Matsui, N., 2015. Simple method for measuring soil sand content by nylon mesh sieving. *Soil Sci. Plant Nutr.* 61, 501–505. <https://doi.org/10.1080/00380768.2015.1016864>.
- Moritsuka, N., Matsuoka, K., Katsura, K., Yanai, J., 2019. Farm-scale variations in soil color as influenced by organic matter and iron oxides in Japanese paddy fields. *Soil Sci. Plant Nutr.* 65, 166–175. <https://doi.org/10.1080/00380768.2019.1583542>.
- Mullins, C.E., 1977. Magnetic susceptibility of the soil and its significance in soil science – A review. *J. Soil Sci.* 28, 223–246. <https://doi.org/10.1111/j.1365-2389.1977.tb02232.x>.
- Poggere, G.C., Barrón, V., Inda, A.V., Barbosa, J.Z., Brito, A.D.B., Curi, N., 2020. Linking phosphorus sorption and magnetic susceptibility in clays and tropical soils. *Soil Res.* 58, 430–440. <https://doi.org/10.1071/SR20099>.
- Quijano, L., Chaparro, M.A.E., Marié, D.C., Gaspar, L., Navas, A., 2014. Relevant magnetic and soil parameters as potential indicators of soil conservation status of Mediterranean agroecosystems. *Geophys. J. Int.* 198, 1805–1817. <https://doi.org/10.1093/gji/ggu239>.
- Ravansari, R., Wilson, S.C., Tighe, M., 2020. Portable X-ray fluorescence for environmental assessment of soils: Not just a point and shoot method. *Environ. Int.* 134, 105250 <https://doi.org/10.1016/j.envint.2019.105250>.
- Rennert, T., 2019. Wet-chemical extractions to characterise pedogenic Al and Fe species – a critical review. *Soil Res.* 57, 1–16. <https://doi.org/10.1071/SR18299>.
- Sano, S., Yanai, J., Kosaki, T., 2004. Evaluation of soil nitrogen status in Japanese agricultural lands with reference to land use and soil types. *Soil Sci. Plant Nutr.* 50, 501–510. <https://doi.org/10.1080/00380768.2004.10408506>.
- Schibler, L., Boyko, T., Ferdyn, M., Gajda, B., Höll, S., Jordanova, N., Rösler, W., MAGPROX Team, 2002. Topsoil magnetic susceptibility mapping: Data reproducibility and compatibility, measurement strategy. *Stud. Geophys. Geod.* 46, 43–57. <https://doi.org/10.1023/A:1019885532390>.
- Shoji, S., Ito, T., Saigusa, M., 1987. Andisol-Entisol transition problem. *Pedologist*, 31, 171–175. 10.18920/pedologist.31.2.171.
- Sumner, M. E., Miller, W. P., 1996. Cation exchange capacity and exchange coefficients. In *Methods of Soil Analysis, Part 3, Chemical Methods*, ED. Sparks DL, et al., pp. 1201–1229. SSSA Book Series no. 5, WI, USA.
- Sunaga, Y., Harada, H., 2016. Simple method for estimating soil mass loading onto plant surface using magnetic material content as a soil indicator – Influence of soil adhesion to vegetation on radioactive cesium concentration in forage. *J. Environ. Radioact.* 164, 125–132. <https://doi.org/10.1016/j.jenvrad.2016.07.016>.
- Torii, M., Fukuma, K., 1998. Initial magnetic susceptibility of the Chinese loess: a review. *The Quaternary Research (Daiyonki Kenkyu)* 37, 33–45. <https://doi.org/10.4116/jaqua.37.33>.
- Torii, M., 2005. Environmental magnetism: a brief review. *J. Geog. (Chigaku Zasshi)* 114, 284–295. <https://doi.org/10.5026/jgeography.114.2.284>.
- Torrent, J., Cabedo, A., 1986. Sources of iron oxides in reddish brown soil profiles from calcarenites in Southern Spain. *Geoderma* 37, 57–66. [https://doi.org/10.1016/0016-7061\(86\)90043-1](https://doi.org/10.1016/0016-7061(86)90043-1).
- Vingiani, S., Scarciglia, F., Mileti, F.A., Donato, P., Terribile, F., 2014. Occurrence and origin of soils with andic properties in Calabria (southern Italy). *Geoderma* 232–234, 500–516. <https://doi.org/10.1016/j.geoderma.2014.06.001>.
- Yanai, J., Okada, T., Yamada, H., 2012. Elemental composition of agricultural soils in Japan in relation to soil type, land use and region. *Soil Sci. Plant Nutr.* 58, 1–10. <https://doi.org/10.1080/00380768.2012.658349>.
- Yoshikawa, S., Kondo, Y., Inouchi, Y., 1993. Detection of invisible volcanic ashes by the magnetic susceptibility. *Jour. Geol. Soc. Jpn.* 99, 293–296. <https://doi.org/10.5575/geosoc.99.293>.