Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE 7 (2020) 294–306

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

Generating soil salinity, soil moisture, soil pH from satellite imagery and its analysis

Mochamad Firman Ghazali [a,c,e,*], *Ketut Wikantika* [a,b,c], *Agung Budi Harto* [a,b], *Akihiko Kondoh* [d]

[a] *Center for Remote Sensing, Bandung Institute of Technology (CRS-ITB), Bandung, Indonesia*
[b] *Remote Sensing and GIS Research Group, Faculty of Earth Science and Technology, Bandung Institute of Technology, Bandung, Indonesia*
[c] *ForMIND Institute, Indonesia*
[d] *Center for Environmental Remote Sensing (CeRES), Chiba University, Japan*
[e] *Department of Geodesy and Geomatics Engineering, Faculty of Engineering, University of Lampung, Indonesia*

ARTICLE INFO

Article history:
Received 4 June 2018
Received in revised form
5 July 2019
Accepted 20 August 2019
Available online 22 August 2019

Keywords:
Landsat 8
Soil moisture index
Soil salinity index
Soil pH
Paddy leaf model
Bare soil model

ABSTRACT

In an agricultural field, the water content and salt content are defined as soil moisture and soil salinity and have to be estimated precisely. The changing of these two factors can be assessed using remote sensing technology. This study was conducted by analysing the Landsat 8 satellite images, soil data of field surveys, laboratory analyses and statistical computations. Soil properties such as soil moisture and soil salinity were estimated using soil moisture index (SMI) and soil salinity index (SSI), respectively. The research combined and integrated the soil data from survey and laboratory with Landsat 8 satellite images to build two multiple regression equations model named the soil pH Index (SpHI). They are based on bare soil and paddy leaf models as the explanatory factors of soil moisture and soil salinity changes. All the computation processes were replicated three times using three different dates of Landsat 8 satellite images to produce the multi-temporal analysis. Soil moisture increased after 30 days, while the salt content was only trace amounts. Both proposed models detected 4.49–7.59 of soil pH, 4.66 in bare soil model and 6.62 in paddy leaf model. During the planting period, the soil pH in bare soil model decreased to 2.12–6.47 while the paddy leaf model increased to 4.49–7.59 with RMSE 1.40 and PRMSE 24% of accuracy. The spatial relationship between soil pH, soil salinity and soil moisture are linear but varied in correlation level from weak, moderate to strong. Based on the bare soil model, the relationship between soil pH and soil moisture shows a weak negative relationship with R2 8.37% and a strong positive relationship with R2 81.94% in paddy area and bare soil area respectively, as like as in paddy area based on the paddy leaf model with R2 100%. The relationship between soil temperature and soil pH shows a weak negative relationship for all models and a moderate negative relationship of soil salinity and soil pH in bare soil area based on the bare soil model with R2 34.89%.

* Corresponding author.
E-mail address: firman.ghazali@eng.unila.ac.id (M.F. Ghazali).
Peer review under responsibility of China Agricultural University.
https://doi.org/10.1016/j.inpa.2019.08.003
2214-3173

1. Introduction

The information of soil moisture and soil salinity play an important role in supporting the agriculture sustainability. Soil moisture and soil salinity are defined as the water and salt contents in soil. In tropical countries like Indonesia, understanding the soil moisture and salt concentration are important as rice becomes the major food source in Indonesia. A sufficient amount of rainfall is required to make the paddy fields receive enough water supply during the rainy season. During rainy season, salt content tends to be decreased while during the dry season it tends to increase. Most of the researchers discuss the importance of monitoring and assessing the soil moisture in agricultural land [1–4]. They state that soil moisture is very important to control the plant growth during the planting period and crop.

Paddy phenology is the other important factor that needs to be considered. Based on De Datta [5], paddy have three stages of phenology during planting period. The first stage is vegetative period, the second is reproductive period and the third is grain filling and maturation period.

The first stage is the vegetative period requires the sufficient water supply in the paddy field to create muddy soil that acts as the root support. Even though in the end the water will decrease naturally, the farmers have to maintain the water content in the soil from drying out during the grain filling and maturation period and harvesting period. The second stage is the reproductive period, where the paddy flowers are produced and slowly turn into grains of rice and ready to be filled until the maturation stage [6]. The planting period in all fields does not start at the same time due to several factors e.g. the difference in water supply, the beginning of the rainy season and the farmer decision. These factors cause different conditions in paddy phenology. As a result, there is a difference in the level of soil moisture in the entire paddy field and changes in soil pH and salinity.

The pH is a chemical term used to describe the condition of acidity, neutral and alkalinity in a solution. The terms acidity, neutral and alkalinity are based on the range of value of 0–6, 7 and 8–14 respectively. In this study, the solution means a soil solution in the paddy field. In the soil, the value of pH is controlled by soil colloids that are influenced by clay, organics matter and oxides contents [7]. Soil pH plays an important role in agricultural activity. It controls the amount and concentration of soil mineral required by plants to grow. The concentration of Boron, Manganese, Cooper, Zinc and Iron in the acidity condition are higher than in the alkaline condition [8]. Agreed with Jones (2002) [8], Smyth [9] also explains that the Hydrogen (H^+), Aluminium (Al^{3+}), and Manganese (Mn^{2+}) become barriers, disturb the plant to grow and poison the plant. At the same time, the concentration of Calcium (Ca), Magnesium (Mg), Potassium (P) and Molybdenum (Mo) in soil is decreased. On the other, in the alkalinity condition, the Iron, Manganese and Nitrogen concentrations are decreased [10]. Furthermore, according to Goto [11], soil pH gradually increases at different growth stages of the rice plants when soil pH is initially 5.5.

According to Al Khaeir [12], the salt content in the soil is the result of evaporation, rainfall, vegetation clearing, soils infiltration and irrigation [13,14]. There are many paddy fields in Indonesia relying on irrigation for water supply but in most cases the irrigation systems stop the water supply during dry season. In this season the temperature increases, rainfall decreases and the rest of the paddy plants are cleaned resulting in bare land. This happens just after the harvesting period, where the salt content exists in the paddy field. This means that there is a correlation between soil moisture and salt content in the paddy field and a change in the soil pH.

Paddy is cultivated in both low land and high land areas in most of the Asian countries. In West Java province, Indonesia, paddy is cultivated in the low land with alluvial soil [15] and in the terrace of high land area with volcanic soil (andosol), and reddish-brown to red soil (latosol) [16]. Although the terminology of paddy soil is used to describe a soil type that is used to grow paddy plant, but according to the soil taxonomy provided by USDA, there is no specific name to describe the paddy soil. However, the physical character of paddy soil is observed as a muddy, surface water level, high clay and silt contents soil. Volcanic and alluvial soils have a different type of source material. Naturally, a volcanic soil has lower pH [17,18] than alluvial soil [19,20].

In the study area, the paddy field is cultivated on alluvial soil. This soil is the fertile soil due to a high accumulation of minerals from erosion and sedimentation process. Most of this soil has a dark yellowish brown color, weak granular and sticky [16], high clay content that consists of montmorillonite and kaolinite [21]. The paddy soil with a high content of montmorillonite tends to crack and is dried and hardened during the dry season, then expands and becomes sticky in the rainy season [22]

2. Materials and methods

2.1. *Study location*

The study of soil pH and its relation with moisture and salinity was conducted in three areas of Majalaya, Rancaekek and Ciparay, East Bandung, Indonesia. This area is low land with an elevation of 600–800 m above sea level. All three areas are located near the industrial zone and flowed by Citarum River. These areas are the premium class as they produce paddy three times a year. Citarum River is a major water source for irrigation in the area despite the fact that it is known as the most polluted river in the world. The river is polluted by the liquid waste from the industry that causes the paddy field be polluted and decreased the quality. However the farmers have used the Citarum water for irrigation for a long time. Fig. 1 shows the paddy fields surrounded by the factory building that have the wastewater tunnel and potentially pollute the paddy plants.

It is important to consider the phenology stage of the paddy plant in the study area. There are two phenology stages including pre and post harvesting stages. Based on the phenology, the models for predicting soil pH were made by following these stages. Finally, pre-harvesting model is called paddy plant model and the pro-harvest model is called bare soil model.

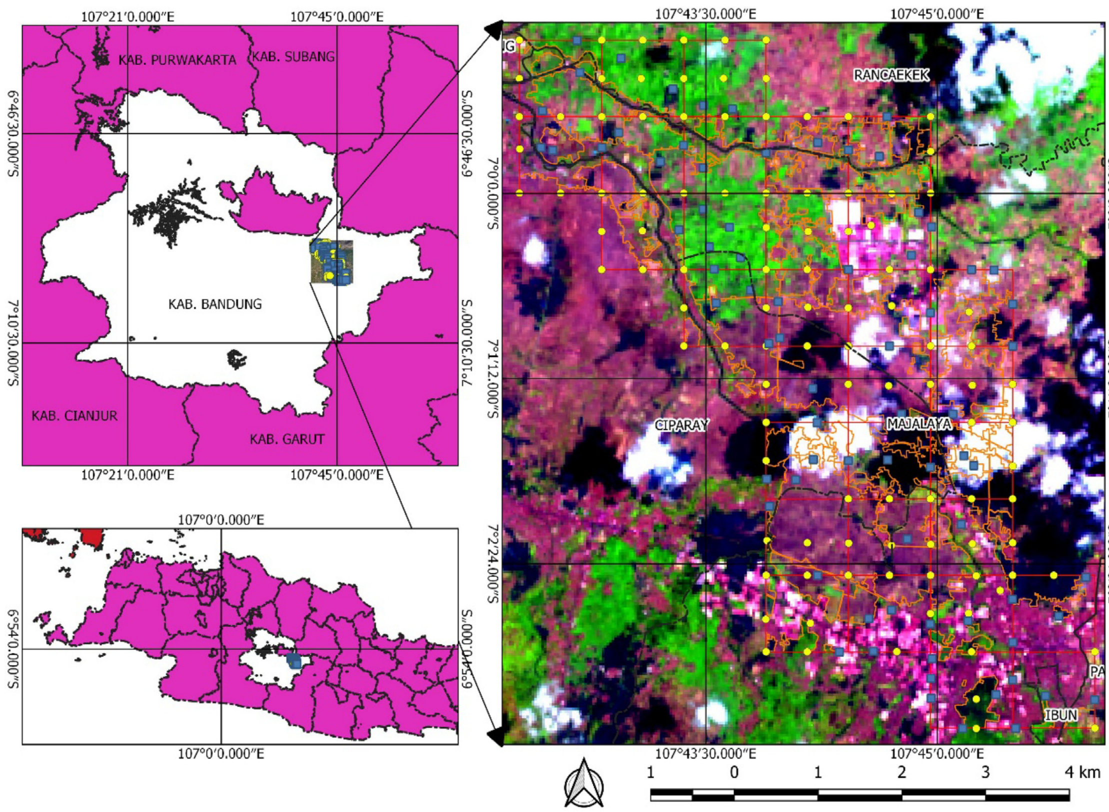


Fig. 1 – Study area is located in the three sub-districts, Majalaya, Rancaekek and Ciparay. The map (right) is overlayed with a false color of Landsat 8. The green color represents the paddy plant and the purple color represents bare soil in the paddy field. Soil pH is collected based on the distribution of point sampling used for collecting soil pH (Blue = ICP, Yellow = GCP). The white objects (left) are the clouds and large black spots are the cloud shadows.

2.2. *Data*

Satellite images and field surveys were required to understand the estimated soil pH and its relationship with other properties including soil moisture and salinity. Three Landsat 8 images were obtained in three different times that covered the study areas in 122 paths and 65 rows. The images were assessed freely from the USGS website (https://earthexplorer.usgs.gov/). The satellite data was required to perform a multi-temporal analyses including soil pH, moisture and salinity in the paddy field. All satellite data used in the study is discussed in Table 1.1.

The field survey was previously conducted to collect 100 soil pH. Soil pH was extracted directly from the soil using a 4 in 1 portable digital soil survey equipment. The tool worked by injecting the metal probe directly into the soil at 20 cm depth. In several locations, the soils are hard thus further analyses were taken to extract the soils. First, one Molar of Kalium Chloride (KCl) solution was taken to make a soil solution. Second, 250 g of soil samples from each location were taken into the laboratory and separated into 10, 100 and 140 g for the analyses of soil pH, moisture and salinity respectively [23]. The soil property was limited to observing only the soil pH, and other properties such as texture and structure were not observed. Some geographic information system (GIS) data, such as the vector map (shape file) of land cover and point sampling distribution was also required. They were constructed by performing a digitation on screen from a high-resolution image from Google Earth. To separate the paddy field from others (e.g. rivers, roads, residential and industrial areas) and to minimize the study area, the Landsat 8 images were used in this study.

The study area was divided into two groups. First is pre-harvesting paddy field with paddy plant and second is pro-harvesting period without paddy plant. They were recorded by Landsat 8 satellite, but some areas were covered by cloud

Table 1.1 – Three of Landsat 8 images used in this study.

No	Date	Path/Row	Function
1	September 29th, 2014	122/65	Estimate soil pH and calculate soil moisture and salinity
2	November 15th, 2014		
3	November 31th, 2014		

and shadow. This limitation caused the soil sampling to be impractical as the covered soil samples by the clouds were excluded from the estimation process. This step is important to conduct as the Landsat 8 images should be in the surface not covered by cloud or shadow. Therefore 8 out of 100 soil samples were excluded from the estimation process. This process involved 65 and 17 samples for evaluating the accuracy. The method is explained in Fig. 2.

2.3. *Paddy leaf model and bare soil model*

Both paddy leaf model (Eq. (7)) and bare soil model (Eq. (8)) are two different multiple regression equations that used to estimate soil pH value in the paddy field. Both models used the same corrected Landsat 8 band as input parameter and validated with soil pH from the survey. The soil pH was collected directly from the soil and laboratory. But in some areas (Fig. 1) paddy plant that covered the field might cause an error and influence the pH estimation. To minimize the error, the specific models were created to accommodate the existence of paddy plant. The models were built based on soil pH in the paddy plant area and bare area. Both models were separated from each other. On the other hand, the soil pH of bare areas was excluded from the model for paddy plant area.

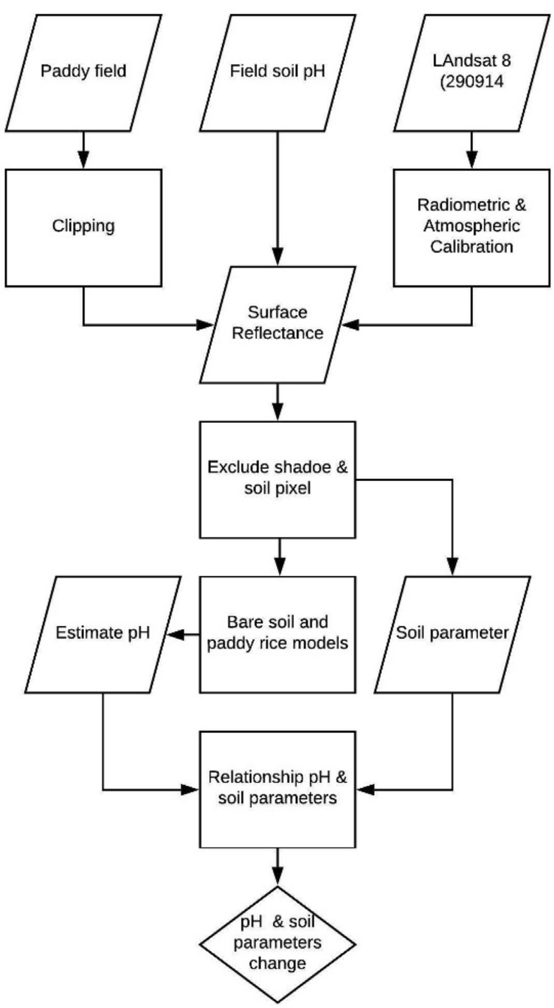


Fig. 2 – Research scheme in generating soil salinity, moisture and pH from satellite images, laboratory analysis and statistical calculation.

Remote sensing is a common perspective in distinguishing plant and soil through the reflectance. The plant reflectance increased from visible light to near infrared and decreased to short wave infrared [24], while bare soil reflectance is gradually increased from visible light to shortwave infrared [25]. This is a reasonable explanation to describe the origin of paddy plant and bare soil models. To predict soil pH using the regression model is easier if the area is all bare soil. However the harvesting times are not exactly the same, therefore the relationship between plant reflectance from Landsat 8 against the soil pH from the field are evaluated. Both models are examined for predicting soil pH in the paddy field. They were compared to understand the performance of each models.

2.4. *Estimation of soil moisture, soil salinity and soil pH*

All Landsat 8 data has to be calibrated prior to analysis. The calibration process consisted of radiometric and atmospheric corrections. The calibration process followed the method by Zanter [26], that aimed to convert the digital number to top of atmosphere (ToA) reflectance. The dark of subtraction (DOS) was also performed using the method by Chavez [27] to get the reflectance surface, called bottom of atmosphere (BoA). This procedure was applied only to visible and near-infrared bands (VisNIR), to minimize the atmospheric influence in Landsat 8 images especially from the aerosol and water vapoured soils and finally to improve the accuracy of estimation and classification [28–30]. For the Landsat 8 thermal band, the digital number had to be converted into radiance and brightness temperature in centigrade using Eq. (2) in USGS (2016) [26] and Eq. (3) in Planck equation [31] respectively

There are previous methods to estimate soil moisture based on satellite data. The study of soil moisture has been conducted by many researchers [3,32–33]. In this study, the soil moisture was estimated using the formula proposed by Pandolfo et al., (2013) [34]. Both Normalized Different Vegetation Index (NDVI) and Land Surface Temperature (LST) are required to determine the value of soil moisture

Normalized Different Vegetation Index (NDVI) was calculated by using a simple ratio of band Near Infrared (NIR) and Red. The first use of this formula (Eq. (1)) was reported by Rouse et al., (1974) [35]. The same formula used by previous authors [36–38]. The NDVI value ranges from −1 to +1. The value lower than 0 indicates the bare soils, rocks, clouds or snows, while the value higher than 0 indicates the vegetation. However, this situation depends on the characteristic of the area as the paddy field is a dynamic area where the change is more likely to occur over time due to its phenological phase of paddy plant. Therefore, in the same location, there would be a semi-bare field or paddy field in the different times. Semi-bare soil occurs after harvesting time where there is still paddy plant attached in the soil. The farmers use the field after for different plants including sweet potato and tomato. Therefore in this type of area is less likely to have a fully bare area like in the wheat field or desert.

The land surface temperature (LST) was calculated using Band 10 (thermal) of Landsat 8. To calculate LST, other parameters are provided in the Landsat 8 metadata [26]. Another study uses the SEBAL, the regression model of thermal and meteorological data and the integration between the normalized difference vegetation index (NDVI) and thermal data [39–41]. A study of paddy phenology characteristics using the NDVI was used to detect multi-temporal variation of paddy plant from Landsat 8 images. NDVI data was also used to detect phenology of vegetation especially for other crops [42–45]. An integration between LST and NDVI created two linear models named *TsMax* and *TsMin* (Eq. (5)), where *a* and *b* are represented as dry edge and wet edge respectively. After those parameters are clearly defined, the value of soil moisture could be calculated using the soil moisture index (Eq. (4)). The result of soil moisture index (SMI) gives a clear explanation that the water content always decreases at the end of the planting period in a paddy field. The alteration of soil moisture is used to understand the changing in soil pH and salinity.

To determine the salt content in the soil, an equation named soil salinity index was proposed by Al Khaeir [12] (Eq. (6)). This approach used a combination of near shortwave infrared (SWIR) of Landsat 8 satellite images. Unlike the reason stated before [12], in this study the salt content resulted from the natural interaction between soil alkalinity and acidity, irrigation, surface temperature that exist in paddy fields in the tropical region. Most recent studies related to the spatial distribution of salt [46,47]. The equations includes NDVI (Eq. (1)), LST (Eqs. (2) and (3)), SMI (Eqs. (4,5)) and SSI (Eq. (6)).

$$NDVI = \frac{(NIR - red)}{(NIR + red)} \tag{1}$$

$$L_\lambda = M_L * Qcal + A_L \tag{2}$$

$$T = \frac{K_2}{In\left(\frac{K_1}{L_\lambda} + 1\right)} \tag{3}$$

where:

L_λ = radiance value of band 10 (thermal) from Landsat 8
M_L = Radiance multiplicative scaling factor for the band
Q_{cal} = L1 pixel value in DN
AL = Radiance additive scaling factor for the band
K_2 and K_1 = Thermal conversion constant for the band

$$SMI = \frac{(TSMax - LST)}{(TSMax - TsMin)} \tag{4}$$

$$\begin{aligned} TSMax &= a1 * NDVI + b1 \\ TSMin &= a2 * NDVI + b2 \end{aligned} \tag{5}$$

$$SSI = \frac{(band6 - band7)}{(band6 + band7)} \tag{6}$$

To estimate soil pH, multiple regression techniques were used based on 100 point samples from surveys and pixel numbers of corrected Landsat 8 reflectance. According to Knudby [48], the modelling process using remote sensing data can be performed using the statistical correlation. In this study, statistics software was used to perform a regression calculation between the pixel value of Landsat 8 bands and field data. Based on their R^2 values, at least four models were created to estimate the value of soil pH in the study area. The models were used to estimate soil pH by performing the accuracy assessment as shown by previous researchers [49–53] utilising the RMSE (Eq. (9)) and PRMSE (Eq. (10)) [54]. Where *X1* represents an observed value, *X2* represents an estimated value, *n* represents the total data observed, $i = 1$represents the calculation from the first sample to the last sample and $\overline{X_2}$ represents the average value of the predicted value. The equation to calculate the root mean square error (RMSE) and percentage root mean square error (PRMSE) are shown as follow.

$$RMSE = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(X_1 - X_2)^2} \tag{9}$$

$$PRMSE = \left(\sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(X_1 - X_2)^2}/\overline{X_2}\right) \times 100\% \tag{10}$$

3. Results and discussion

3.1. Status of soil moisture in paddy field

The soil moisture can be estimated using a combination of NDVI and LST data. Based on the observation of NDVI from three Landsat 8 images (Fig. 3), the NDVI value changes over time showing the variations in 30 days. According to the distribution value of the NDVI, it has a minimum value of −0.1 and a maximum value of 0.7. In the Northern part of the study area on 29/09/2014, some of the paddy plant was still growing. The NDVI value of more than 0.5 indicated that the paddy existed and was probably at the end of the grain filling stage or maturation and harvesting stages. But in the Southern part of the paddy field, no paddy plant existed as indicated by NDVI of less than 0.4. But it did not mean that the whole area consisted of bare soil, as paddy trunk was still attached in the soil.

During 15/10/2014 to 31/10/2014 the whole area was harvested, but it was not entirely clean. The NDVI ranged from −0.1 to 0.4. The paddy field was in the dormant stage (resting period) and prepared for the next planting preparation. This situation changed gradually during 30 days.

During that period, the change of NDVI value was followed by the change of surface temperature. High NDVI value was followed by low temperature. This occurred as the canopy in the paddy plant contributed to reflect the sun radiation less than bare soil area. This hypothesis corresponded with the LST on 29/09/2014 and 31/10/2014, where the temperature recorded as 15–35 °C and the area with paddy plant on 29/09/2014 was 25 °C. But Fig. 3 shows another behaviour on 31/10/2014 when the temperature decreased in the whole area. It was recorded at 10–20 °C. This situation might be an anomaly since in the next observation on 31/10/2014, the temperature returned to normal, similar with 29/09/2014.

In general, based on the maps of estimated soil moisture shows that the water content in soil has slowly increased in

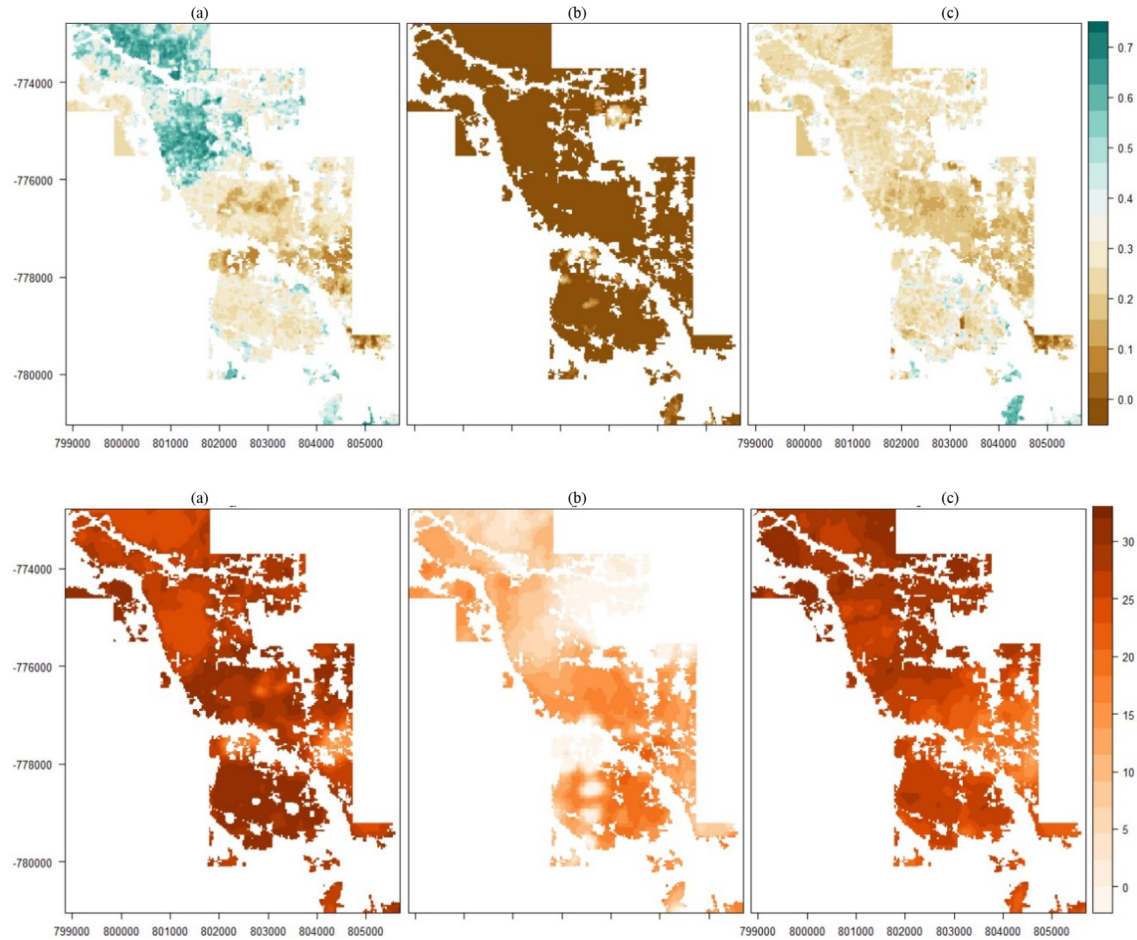


Fig. 3 – Normalized Different Vegetation Index (Upper Images) in range of value from 0 to 1 and Land Surface Temperature in paddy field (Lower images) in centigrade. All images were assessed from Landsat 8 images (a = 09/29/2014, b = 10/15/2014, c = 10/31/2014).

during 30 days. The spatial distribution of the estimated result of soil moisture (Fig. 4) on 29/09/2014 shows that the water content is probably as much as 4% per meters square and reach a maximum level at 8% per meters square on 31/10/2014. While taking a look into the map on 29/09/2014, when the paddy plant is still growing in the lower area of water content or it has a lower moisture. But in another location the area without paddy has a high moisture as much as 4% per pixel or equal to 30 m square. Therefore paddy does not require a large amount of water before harvested.

Even though the paddy plant is growing in a drier location (Fig. 4) this was the effect of the drought season in September when the rainfall is insufficient to provide the necessary water supply.

In the last observation (31/10/2014), the whole area of paddy field showed a higher soil moisture level at 8% per 30 m square of Landsat 8 pixel size. During that time, the water of Citarum river was utilized directly as a preparation of paddy field cultivation in the next period. Although, using water from the Citarum river might increase the soil moisture in paddy field. It possibly changes the chemical content in soil. However, understanding the relationship between chemical content (pollutant) in water that triggers the changing of the soil pH were important to gain. Misra [55] explained this phenomenon. Others factors e.g. the starting period of the rainy season might be able to raise the soil moisture twofold.

3.2. *Status of soil salinity in the paddy field*

The estimation result of soil salinity is showing an increase in trend (Fig. 5). During 30 days, the salinity level in the soil is going higher. Physically, salinity refers to a concentration of salt both in the soil and water. The number of salt content in the soil is different between the salt content in the water. It depends on rainfall and other factors such as vegetation clearing, soils infiltration and irrigation [13,14]. When the temperature was higher, the level of evapotranspiration and salt content would be higher. This is interesting to know the relationship between the information on soil moisture, soil salinity and land surface temperature, and the effect will lead into the soil pH change.

The salt content in the paddy field on 29/09/2014 was predicted at around −2 to 0 ds/m per 30 square meters. This occurred because during the last period of paddy cultivation, there was no salt content in the soil. However in the first fifteen days (10/15/2014), at least ~0.5 ds/m per 30 square

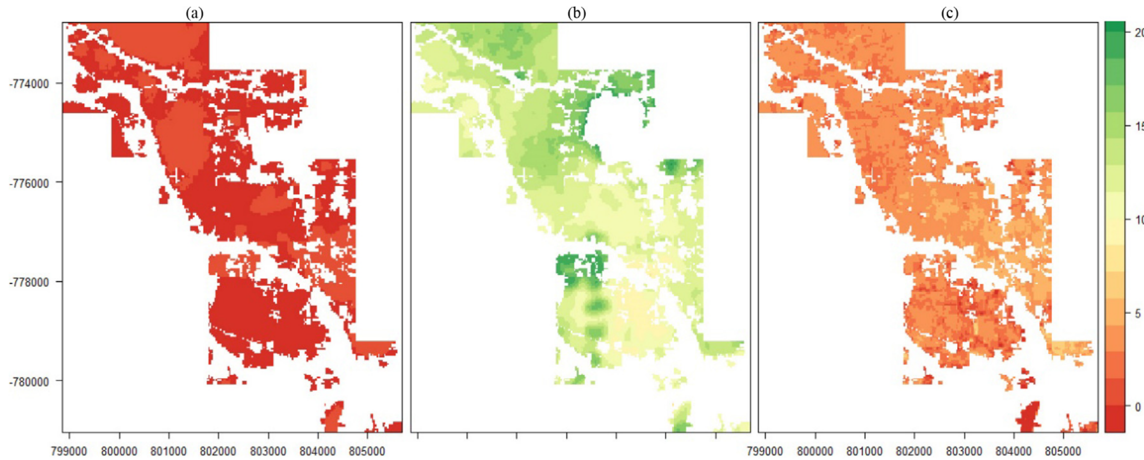


Fig. 4 – Soil Moisture condition in paddy field just before harvest Period (a), after harvesting (b), and before cultivation (c). Dark color and the high value indicate that the soil moisture is high. Bright color and low value indicate that the soil moisture is low. Both images were taken from Landsat 8 images (a = 09/29/2014, b = 10/15/2014, c = 10/31/2014).

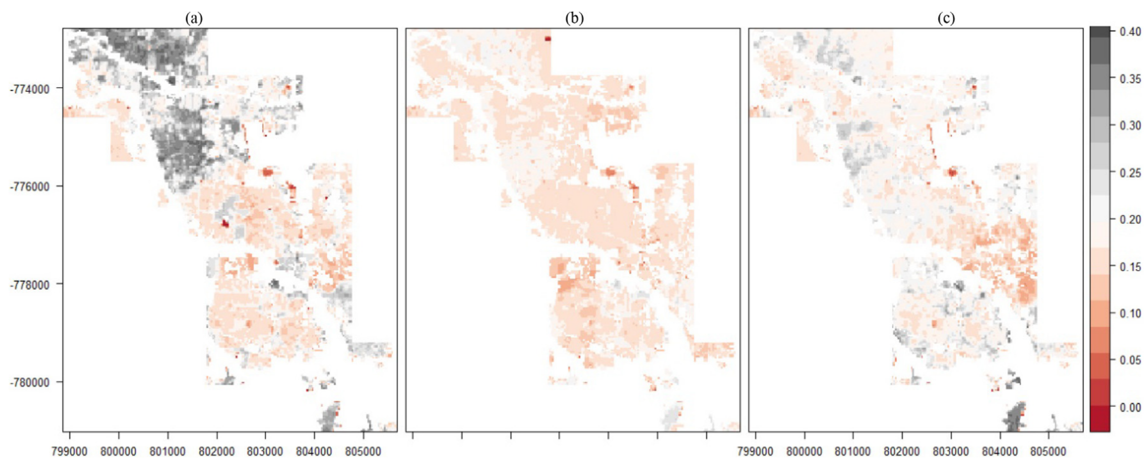


Fig. 5 – Soil Salinity Index (a = 09/29/2014, b = 10/15/2014, c = 10/31/2014). Bright color and high value indicate high salt contents while dark color and low value indicate low salt contents.

meters of salt content was detected in the paddy field and peaked on 10/31/2014 at ~0.5 ds/m per 30 square meters. The increasing number of salt content occurred during the start of the rainy season in October. Meanwhile the salt content in the soil did not exist during dry season in September.

3.3. *Status of soil pH in paddy field*

Statistical calculation from the pixel value of corrected Landsat 8 bands and field soil pH is shown in Table 1.2. All visible light and two short wave infrared bands show a weak correlation due to the low coefficient of determination or R^2 (less than 0.1), thus classified as the weak band. The short infrared band is absent from the statistical computation since it has no relationship with the soil pH. From this coefficient of determination, there are two multiple regression equations used to estimate the soil pH in paddy field, called paddy plant model (Eq. (7)) and bare soil model (Eq. (8)).

All soil pH samples collected from the canopied paddy field is classified as soil pH at paddy plant area then configured for paddy plant model. Besides that, the soil pH located in the bare soil area is recognised as bare soil model. Although all bands performance is weak or almost did not show a significant relationship, it might be useful to detect soil pH. This situation was the effect of the different spatial resolution between Landsat 8 and the number of soil samples obtained in the field, which was only 8 per 30 m. The reduced number of samples due to cloud cover and shadow were showed in Landsat 8 images.

$$pH = 6.493 - 35.152 * B2 - 52.380 * B3 + 1.099 * B4 + 30.040 * B6 - 8.181 * B7 \quad (7)$$

$$pH = 6.232 - 59.439 * B2 - 89.326 * B3 + 136.721 * B4 + 5.612 * B6 - 25.603 * B7 \quad (8)$$

Table 1.2 – Landsat 8 bands classification based on coefficient of determination (R^2) vs. soil pH in the paddy plant area (left) and bare soil area (right).

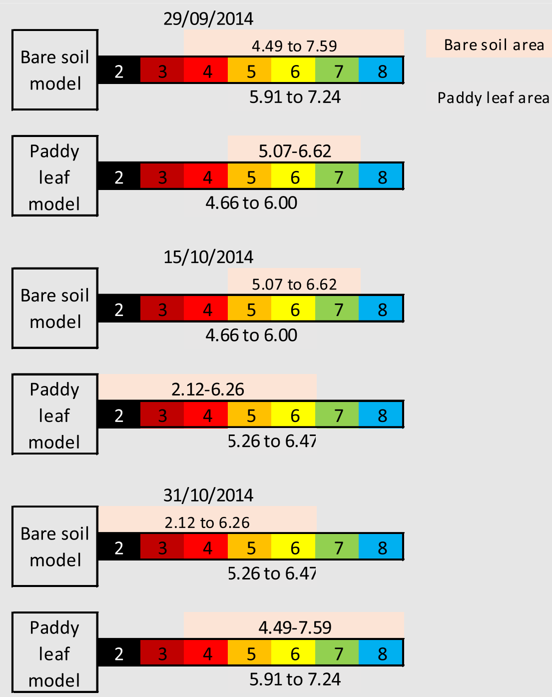
	Paddy leaf area					Bare soil area				
	B2	B3	B4	B6	B7	B2	B3	B4	B6	B7
pH Sensitivity	0.010 Weak	0.020	0.011	0.053	0.045	0.010 Weak	0.010	0.010	0.005	0.030

Table 1.3 – Comparison of estimated result of soil pH using three different data of Landsat 8 images.

Model	Paddy field without paddy	MIN	MAX	MEAN	RMSE	PRMSE
Bare soil model	29/09/2014	4.49	7.59	6.34	1.13	18%
	15/10/2014	5.07	6.62	5.76	1.00	17%
	31/11/2014	2.12	6.26	5.49	1.36	25%
	Average				1.17	20%
	Paddy field with paddy	**MIN**	**MAX**	**MEAN**	**RMSE**	**PRMSE**
	29/09/2014	5.91	7.24	6.60	2.12	32%
	15/10/2014	4.66	6.00	5.45	1.27	23%
	31/11/2014	5.26	6.47	5.81	1.52	26%
	Average				1.63	27%

Model	Paddy field without paddy	MIN	MAX	MEAN	RMSE	PRMSE
Paddy leaf model	29/09/2014	5.07	6.62	5.76	1.00	17%
	15/10/2014	2.12	6.26	5.49	1.36	25%
	31/11/2014	4.49	7.59	6.34	1.13	18%
	Average				1.17	20%
	Paddy field with paddy	**MIN**	**MAX**	**MEAN**	**RMSE**	**PRMSE**
	29/09/2014	4.66	6.00	5.45	1.27	23%
	15/10/2014	5.26	6.47	5.81	1.52	26%
	31/11/2014	5.91	7.24	6.60	2.12	32%
	Average				1.63	27%

Total accuracy	RMSE	PRMSE
Bare soil model	1.40	24%
Paddy leaf model	1.40	24%



The estimation result of soil pH produces by Eqs. (7) and (8) gave a good range value of soil pH. Basically the soil pH will be around 0–14, and those models performance are very impressive when it was applied to the first Landsat 8 image (29/09/2014) shows the soil pH in bare soil area around 4.49–7.59 and 5.07–6.62 based on bare soil model and paddy leaf model respectively. In paddy leaf area the soil pH is around 5.91–7.24 and 4.66–6.00 based on bare soil model and paddy leaf model respectively. The first impression is address into the model capability and performance since it successfully give a good estimation result, at least it located in a good range of pH measurement in the laboratory and in the field using measurement equipment. It's very reasonable by looking at the other result of the second also given soil pH in bare soil area around 5.07–6.62 and 2.12–6.26 based on bare soil model and paddy leaf model respectively. In paddy leaf area the soil pH is around 4.66–6.00 and 5.26–6.47 based on bare soil model and paddy leaf model respectively. At the third Landsat images in Table 1.3 in bare soil area around 2.12–6.26 and 4.49–7.59 based on bare soil model and paddy leaf model respectively. In paddy leaf area the soil pH is around 5.26–6.47 and 5.91–7.24 based on bare soil model and paddy leaf model respectively.

All models result the soil pH ranging from 2 to 8. All estimation produced from paddy plant model and bare soil model in paddy plant area is higher than in bare soil area. The soil pH in the bare soil area tends to decrease during 30 days while it decreases in paddy plant area (Fig. 6). As described above, the presence of paddy plant in paddy field might become a barrier for Landsat 8 reflectance to receive the soil characteristics. It makes a comparison between two values of soil pH. The difference between estimated soil pH in paddy plant area and in bare soil area was conducted by taking a sample from 29/09/14 and 31/10/14 resulting the difference of 0.7. This means that if the soil pH in the paddy plant area is 5.5, the soil pH in bare soil area is 4.8.

Completing the estimation result, the accuracy level is required. Based on the RMSE and PRMSE, the two models have accuracies of 1.40% and 24%. This result however still

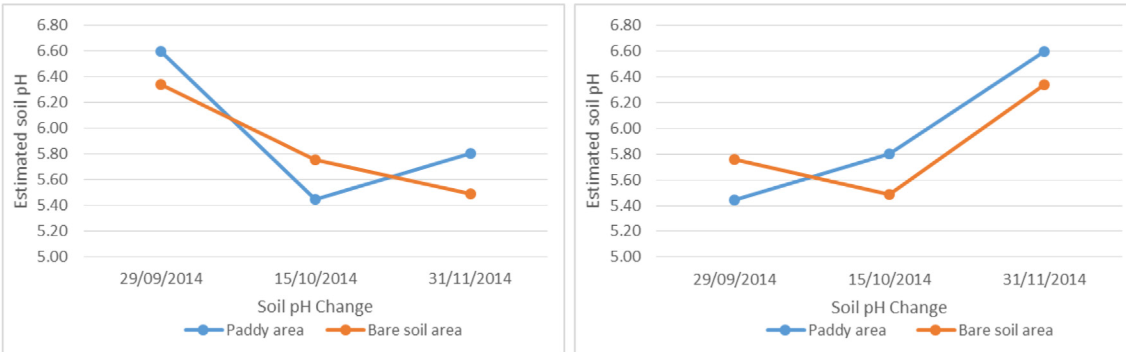


Fig. 6 – Soil pH change in 30 days. Soil pH in bare soil are and paddy plant area derived from bare soil model show decreasing patterns (left) and soil pH in bare soil and paddy plant area derived from the paddy plant model show increasing patterns (right).

Fig. 7 – Estimation result of soil pH from both bare soil model (top) and paddy leaf model (bottom). The solid red color in the map indicates cloud covered and cloud shadow. (a = 09/29/2014, b = 10/15/2014, c = 10/31/2014).

excludes the soil pH factors from the field surveys, cloud cover and shadow modelling.

Soil pH is spatially distributed in an irregular pattern (Fig. 7). Two main rivers are the primary sources that pollute the paddy fields. The soil pH in bare soil area decreased in 30 days. Fig. 1 shows that the middle to southern part of paddy field is dominated by bare soil while northern part is covered by paddy plant. This verifies that the soil pH decreases during the start of rainy season in bare soil area. For example, in the centre part of each image shows that the soil pH in the bare soil area decreases and is acidic. The blue color changes gradually, then turns to be yellow. This occurs in the southern part of each image where yellow turns into blue at first then changed to yellow. This also occurs in the upper side of paddy field (left image)

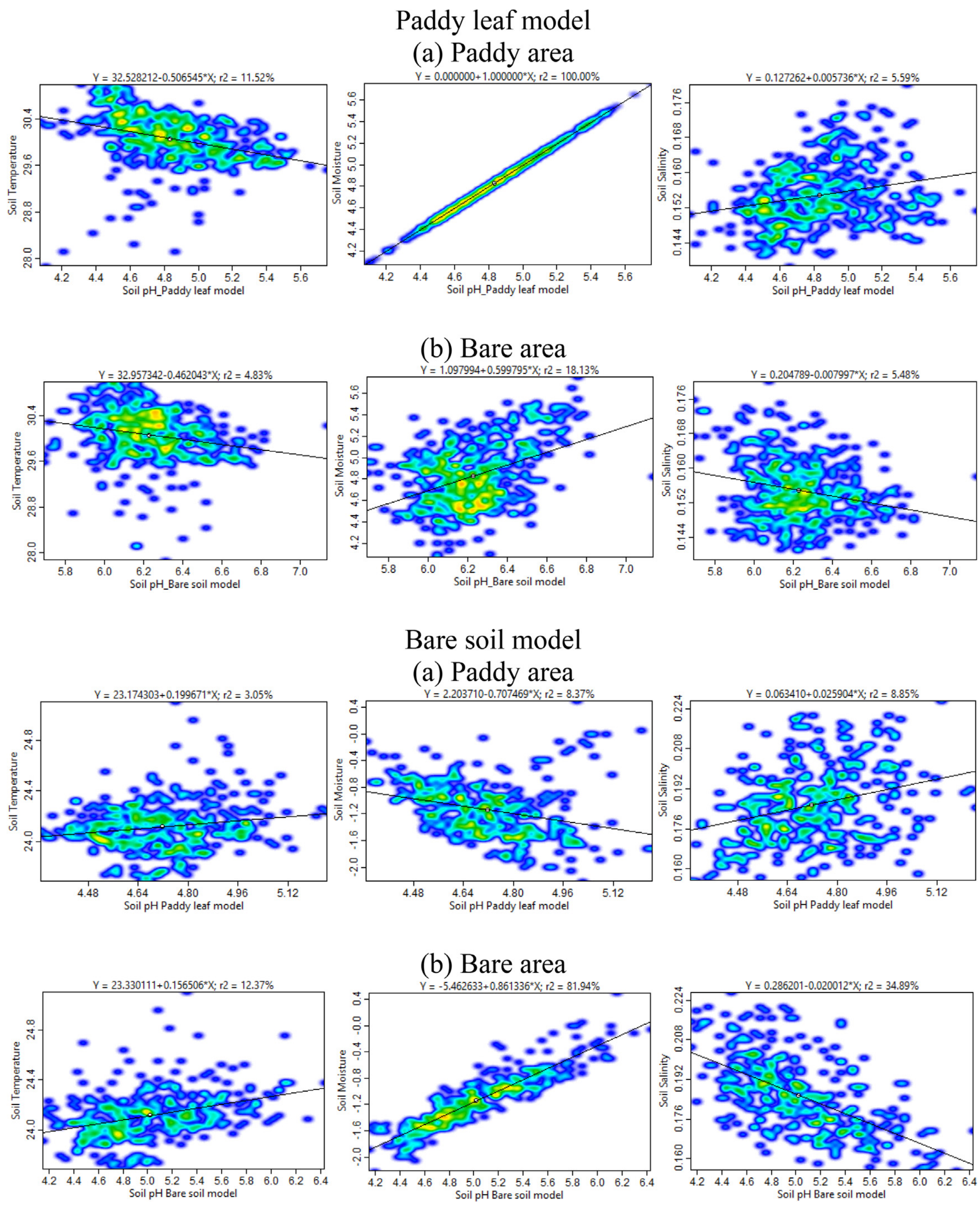


Fig. 8 – Regression model describes relationships between soil pH with temperature, moisture and salinity.

that is located in northern part of the river. Here, the soil pH is almost neutral (blue) then slowly decreased (centre image) and back to neutral (right image).

Other factors such as land surface temperature, moisture (soil moisture index) and soil salinity (soil salinity index) are taken into account in determining the soil pH dynamics [56], and the available amount of crops. Therefore the relationship between soil pH and crops can be recognised as well. When soil pH changes, the paddy field lose the capacity to produce the crop [5,57]. The crop production study in Bandung is needed to verify the crop annual report by the Central Bureau of Statistics (BPS) of West Java. The paddy production from 2013 to 2015 decreased from 11,538 ton to 10,508 ton. McDole [58] stated that the production can be increased if the soil pH is neutral (7).

3.4. *Relationship between soil moisture, soil pH and soil salinity in paddy field*

The relationships between soil temperature, moisture and salinity vs. soil pH were explained using the scattered plots.

Fig. 8 shows twelve pairs of scatter plots that explain the soil pH changing and the relationships in nature.

The plot of soil pH vs. soil temperature using the paddy plant model (Fig. 8 first column) shows weak negative relationships with R^2 11.52% and 4.83% in the paddy area and bare soil area respectively. The bare soil model results a weak positive relationship with R^2 3.05% and 12.37% in paddy area and bare soil area respectively. This shows that no more than 12% of the soil temperature triggers a changing in soil pH. This occurs as the temperature ranges from 20 to 25 °C, the soil pH reaches <4 and possibly affect the CO_2 supply for photosynthesis [59]. Considering the climatic factor for paddy plant, this range of temperature is suitable for the low land area (<750 m above sea level) [5]. In most cases in Indonesia, there are no paddy fields above 750 m although the hilly areas are also suitable for paddy field development.

In the second column of Fig. 8 (the second column), there are moderate to strong positive relationships between soil pH and soil temperature. Based on the paddy plant model, the soil pH in paddy area and bare soil area have R^2 100% and 18.13% respectively. Based on the bare soil model, soil moisture has a weak negative relationship against soil pH with R^2 8.37% in the paddy area. Furthermore the soil moisture has strong positive relationship with R^2 81.94% in bare soil area. In general, the soil moisture triggers the change in soil pH up to 100%.

The soil water content also can increase the soil pH. In some areas with a periodical flooding event, the flood turns the soil pH become alkaline [9,19–20,55]. This allows the root mass to increase and triggers various diseases in plants [60]. Although the physical properties of soil (e.g. soil texture and structure) were not observed in this study, the relation of soil pH and soil moisture was explained in Fig. 6. Agreed with Guriveth et al., [61] in McCauley [62] that the amount of water in soil is also observed by the variation of soil texture. The high clay soil retains more water than the sandy soil. Since the paddy plant in the study area is cultivated on alluvial soil, the clay content in the paddy field is very high.

Those relationships vary according to their regression line, but this is possible to explain other relationships. At the same time, soil pH is increased by increasing salt content. The temperature maintains the water content in the soil to be stable or gradually change and controls soil pH into an acceptable range to support the plant. Although the salt contents are increased, the water content stays in low level. The situation is applied in tropical areas where paddy fields on alluvial soil only. Another research on the same topic but in different soil type mention in Tan [16] is suggested to conduct.

Linear regression models result both positive and negative relationships between soil pH and soil salinity (Fig. 8 third column). Both paddy plant model and bare soil model in paddy area show a weak positive relationship with R^2 5.59% to 8.85% while in the bare soil area it has a weak to strong negative relationship R^2 5.48% to 34.89%. In paddy area, the soil pH decreases as the soil salinity increase at 34.89%. But in bare soil area, 8.85% of soil pH also changes by the increasing salt content in the soil. Salinity defined as the amount of salt content in the soil and produced by a chemical reaction between alkalinity and acidity. For instance, the NaCl (Natrium Chloride) has an acid natrium and an alkaline chloride. This paper does not address the source and type of salt. However, the hypothesis indicates that the soil undergoes as chemical reaction called cation exchange capacity (CEC), where cation always changes due to contamination by industrial waste.

As soil pH estimation and its relationship with other properties were conducted every 30 m^2 from Landsat 8, the usage was limited to paddy fields. The variations are found as the geographic characteristics, soil type and structure are variable between areas. Therefore, other studies have to be conducted using another size of pixel and/or using an aerial photograph from UAVs for better precision and accuracy. A basic configuration of UAV photo has a standard sensor with RGB band including a green band attached on it. As long as it is applied to estimate the soil pH, the result will be more precise than ones obtained from 30 m of Landsat 8 pixel.

4. Conclusion

Soil pH was successfully estimated by SpHI with an appropriate range of result (2–7.59) and an accepted estimation accuracy. The blue, green, red shortwave infrared bands of Landsat 8 satellite images are the most suitable bands to detect soil pH in paddy field. The soil pH based on bare soil model is useful to detect the soil pH in the preparation stage of the paddy field. The soil pH based on paddy plant model is useful to detect the soil pH in the planting season. The soil pH has linear relationship against soil temperature, soil moisture and soil salinity ranging from weak to strong relationships.

A further study is necessary to produce better estimation results by improving the extraction method and explaining the relationship uncertainties of soil pH vs. soil moisture, surface temperature and salinity.

As Landsat 8 satellite images have limited spectral and spatial resolutions, other remote sensing data with better spatial and spectral resolution have to be considered to solve this limitation. The same techniques have to be applied to other soil types with various physical properties. As such, the improvement of soil moisture (SMI), soil salinity (SSI) and soil pH (SpHI) and its relationship can be obtained with more detail. The negative correlation between short infrared reflectance of Landsat 8 satellite and the statistic calculation also need to be studied in the future and be a main focus.

Declaration of Competing Interest

The authors declared that there is no conflict of interest.

Acknowledgement

I would like to thank to the Center for Remote Sensing of Bandung Institute of Technology, and The Center for Environmental Remote Sensing of Chiba University for valuable guidance and support during the research work. Also for the LPDP, thank you for the scholarship and the full funding support for the research.

REFERENCES

[1] Shekhar R, Chandrasekar K, Sesha Sai MVR, Diwakar PG, Dadhwal VK. Automated system for generation of soil moisture products for agricultural drought assessment. Int Arch Photogramm Remote Sens Spat Inf Sci 2014;40:111–7.

[2] Parajuli PB, Duffy SE. Evaluation of soil organic carbon and soil moisture content from agricultural fields in Mississippi. Open J Soil Sci 2013;3:81–90.

[3] Narendra BH. Drought monitoring using rainfall data and spatial soil moisture modeling Master thesis. International Institute for Geo-Information Science and Earth Observation-Universitas Gadjah Mada; 2008.

[4] Arif C, Setiawan B, Mizoguchi M, Doi R. Estimation of soil moisture in paddy field using Artificial Neural Networks. Int J Adv Res Artif Intell 2012;1:17–21.

[5] De Datta SK. Germination, growth, and development of the rice plant. In: Principles and practices of rice production. Singapore: John Willey & Sons; 1981. p. 146–72.

[6] Mosleh MK, Hassan QK, Chowdhury EH. Application of remote sensors in mapping rice area and forecasting its production: a review. Sensors 2015;15:769–91.

[7] Delgado A, Gomez JA. The soil. Physical, chemical and biological properties. In: Villalobos FJ, Fereres E, editors. Principles of agronomy for sustainable agriculture. Springer International Publishing AG; 2016. p. 15–27.

[8] Jones JB. Soil pH, liming, and liming materials. In: Agronomic handbook management of crops, soils and their fertility. Washington DC: CRC Press; 2002. p. 237–51.

[9] Smyth JT. Soil acidity and liming. In: Handbook of soil sciences resource management and environmental impacts. Boca Raton: CRS Press; 2012. p. 373–9.

[10] The Potash Development Association. Soil analysis: key to nutrient management planning. Link: https://www.pda.org.uk/pda_leaflets/24-soil-analysis-key-to-nutrient-management-planning/pda-if24.pdf. 2011:12. [accessed September 23, 2018].

[11] Goto K. Relationships between soil pH, available calcium and prevalence of potato scab. Soil Sci Plant Nutr 1985;31:411–8.

[12] Al Khaeir F. Soil salinity detection using satellite remote sensing. Master thesis. International Institute for Geo-Information Science and Earth Observation-University of Twente; 2003.

[13] Lennard EB. The interaction between waterlogging and salinity in higher plants: causes, consequences and implications. Plant Soil 2003;253:35–54.

[14] Hingston F, Gailitis. The geographic variation of salt precipitated over Western Australia. Aust J Soil Res 1976;14:319–35.

[15] Kawaguchi K, Kvuma K. Description of fertility characteristics paddy soils in tropical Asia. Southeast Asian Stud 1974;12:3–24.

[16] Tan KH. The genesis and characteristics of paddy soils in Indonesia. Soil Sci Plant Nutr 1968;14:117–21.

[17] Bołzan BD. Effect of pH and soil environment. World News Nat Sci 2017;8:50–60.

[18] Page-dumroese D, Ferguson D, Mcdaniel P, Johnson-maynard J. Chemical changes induced by pH manipulations of volcanic ash-influenced soils. In: Page-dumroese D, Miller R, Mital J, McDaniel P, editors. Volcanic-ash-derived for soils inl northwest prop implic manag restor. Coeur d'Alene: U. S. Department of Agriculture; 2007. p. 185–202.

[19] Moore SH. Optimum soil applied nitrogen levels for cotton on a high pH alluvial soil. J Plant Nutr 1998;21:1139–44.

[20] Guo Z, Chai M, Wang J, Chen Z, Zhen-Shou Z, Wu-Ping Z, et al. Spatiotemporal variation of soil PH in the past 30 years of Guangdong Province, China. J Appl Ecol 2011;22:302–5.

[21] Yoothong K, Pukamphol M, Hutspardol A. Distribution of clay minerals-kaolinite, illite and montmorillonite in the coarse and the fine clay fractions in some alluvial soils. Kasetsart J Nat Sci 1986;20:285–99.

[22] Prasetio BH, Adiningsih JS, Subagyono K, Simanungkalit R. Mineralogi, kimia, fisika dan biologi tanah sawah. Lahan Sawah dan Teknol Pengolahannya, Bogor: Balai Penelitian Tanah, Kementerian Pertanian; 1996, p. 29–82.

[23] Motsara MR, Roy RN. Soil analysis. In: Guid to Lab Establ plant Nutr Anal. Rome: Food and Agriculture Organization of the United Nations; 2008. p. 17–76.

[24] Thenkabail PS, Lyon JG. Huete Alfredo. Hyperspectral vegetation indices. In: Hyperspectral Remote Sens Veg. Boca Raton: CRC Press-Taylor and Francis Group; 2012. p. 309–28.

[25] Maliki AAl, Owens G, Bruce D. Capabilities of remote sensing hyperspectral images for the detection of lead contamination. ISPRS Ann Photogramm Remote Sens Spat Inf Sci, Melbourne, vol. I–7. p. 55–60.

[26] U. S. Geological Survey. Conversion of DNs to physical units. In: Zanter K, editor. Landsat 8 data users Handb. 2nd ed., Sioux Falls, South Dakota: Department of the Interior U.S. Geological Survey; 2016, p. 60–1. doi:LSDS-1574.

[27] Chavez PS. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sens Environ 1988;24:459–79.

[28] Lin C, Wu C-C, Tsogt K, Ouyang Y-C, Chang C-I. Effects of atmospheric correction and pansharpening on LULC classification accuracy using WorldView-2 imagery. Inf Process Agric 2015;2:25–36.

[29] Matthew MW, Adler-golden SM, Berk A, Felde G, Anderson GP, Gorodetsky D, et al. Atmospheric correction of spectral imagery: evaluation of the FLAASH algorithm with AVIRIS data. In: Appl Imag Pattern Recognit Work, Washington, D.C.. p. 157–63.

[30] Maryantika N, Lin C. Exploring changes of land use and mangrove distribution in the economic area of Sidoarjo District, East Java using multi-temporal Landsat images. Inf Process Agric 2017;4:321–32.

[31] Weng Q, Lu D, Schubring J. Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. Remote Sens Environ 2004;89:467–83.

[32] Sandholt I, Rasmussen K, Andersen J. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sens Environ 2002;79:213–24.

[33] Wang X, Xie H, Guan H, Zhou X. Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions. J Hydrol 2007;340:12–24.

[34] Pandolfo A, Antunes P, Gong X. Using Landsat TM Data for soil moisture mapping. Leicester; 2013.

[35] Rouse JW, Haas RH, Scheel JA, Deering DW. Monitoring vegetation systems in the great plains with ERTS. 3rd Earth Resour Technol Satell Symp, vol. 1. p. 309–17.

[36] Engstrom R, Hope A, Kwon H, Stow D. The relationship between soil moisture and NDVI near barrow. Phys Geogr 2008;29:38–58.

[37] Esselink P, Gils H Van. Ground-based reflectance measurements for standing crop estimates. ITC J 1985;1:47–52. Link: https://www.researchgate.net/publication/294675955%0AGround-based [accessed April 20, 2018].

[38] Weier J, David H. Measuring vegetation (NDVI&EVI). Link: https://earthobservatory.nasa.gov/Features/MeasuringVegetation. 2000:2. [accessed August 31, 2018].

[39] Feizizadeh B, Blaschke T, Nazmfar H, Akbari E, Kohbanani HR. Monitoring land surface temperature relationship to land use/land cover from satellite imagery in Maraqeh county, Iran. J Environ Plan Manage 2012;56:1290–315.

[40] Ozelkan E, Bagis S, Ozelkan EC, Ustundag BB, Ormeci C. Land surface temperature retrieval for climate analysis and association with climate data. Eur J Remote Sens 2017;47:655–69.
[41] Bendib A, Dridi H, Kalla MI. Contribution of Landsat 8 data for the estimation of land surface temperature in Batna city, Eastern Algeria. Geocarto Int 2016;32:503–13.
[42] Wei W, Wu W, Li Z, Yang P, Zhou Q. Selecting the optimal NDVI time-series reconstruction technique for crop phenology detection. Intell Autom Soft Comput 2015;22:237–47.
[43] Hou X, Gao S, Niu Z, Xu Z. Extracting grassland vegetation phenology in North China based on cumulative SPOT-Vegetation NDVI data. Int J Remote Sens 2014;35:3316–30.
[44] Beck P, Jönsson P, Høgda K, Karlsen S, Eklundh L, Skidmore A. A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula. Int J Remote Sens 2007;28:4311–30.
[45] Zhang H, Li Q, Liu J, Shang J, Du X, Zhao L, et al. Crop classification and acreage estimation in North Korea using phenology features. GIScience Remote Sens 2017;54:381–406.
[46] Yu S, Jintong L, Liu H, Eneji AE, Han L. Spatial variability of soil salinity under subsurface drainage. Commun Soil Sci Plant Anal 2014;46:259–70.
[47] Swarajyalakshmi G, Gurumurthy P, Subbaiah G. Soil salinity in south India: problems and solutions. J Crop Prod 2008;7:247–75.
[48] Pittman SJ, Knudby A, Maina J, Rowlands G. Remote sensing and modeling of coral reef resilience. In: Remote sensing and modeling advances in coastal and marine resources. New York: Springer; 2014. p. 103–34.
[49] Chai T, Draxler R. Root mean square error (RMSE) or mean absolute error (MAE)? –Arguments against avoiding RMSE in the literature. Geosci Model Dev 2014;7:1247–50.
[50] Moses KP, Devadas MD. An approach to reduce root mean square error in toposheets. Eur J Sci Res 2012;91:268–74.
[51] Morad M, Chalmers A, O'Regan P. The role of root-mean-square error in the geo-transformation of images in GIS. Int J Geogr Inf Syst 2007;10:347–53.
[52] Ghufran R, Ramli MF. Geo-referencing the satellite image from google earth by relative and absolute positioning. Malasyian J Sci 2007;26:135–41.
[53] Hastie T, Tibshirani R, Gareth J, Witten D. Assesing model accuracy. In: Casella G, Fienberg S, Olkin S, editors. An introduction to statistical learning with application in R. New York: Springer; 2013. p. 29–37.
[54] Lin C, Thomson G, Popescu SC. An IPCC-compliant technique for forest carbon stock assessment using airborne lidar-derived tree metrics and competition index. Remote Sens 2016;8:1–19.
[55] Misra A, Tyler G. Influence of soil moisture on soil solution chemistry and concentrations of minerals in the Calcicoles phleum phleoides and Veronica spicata grown on a limestone soil. Ann Bot 1999;84:401–10.
[56] Kissel DE, Sonon L, Vendrell PF, Isaac RA. Salt concentration and measurement of soil pH. Commun Soil Sci Plant Anal 2009;40:179–87.
[57] Yoshida S. Mineral nutrition of rice. In: Fundamental rice crop science. Los Banos: The International Rice Research Institute; 1981. p. 135–46.
[58] Mahler R, McDole RE. Effect of soil ph on crop yield in northern Idaho. Agron J 1987;79:751–5.
[59] Reth S, Reichstein M, Falge E. The effect of soil water content, soil temperature, soil pH-value and the root mass on soil CO2 efflux – a modified model. Plant Soil 2005;268:21–33.
[60] Narisawa K, Shimura M, Usuki F, Fukuhara S, Hashiba T. Effects of pathogen density, soil moisture, and soil pH on biological control of clubroot in chinese cabbage by heteroconium chaetospira. Plant Dis 2005;89:285–90.
[61] Gurevitch J, Scheiner SM, Fox GA. Water relation and energy balance. In: The Ecology of plants. Sunderland, Massachusetts: Sinauer Associates, Inc.; 2006. p. 43–70.
[62] McCauley A, Jones C, Jacobsen J. Basic soil properties. In: Soil water Management Modul 1. Montana: Montana State University; 2005. p. 1–12.