Identifying geothermal steam spots based on binary logistic method to a high resolution imagery and field measurement

F. Z. Gandes\textsuperscript{1,2,*}, A. Saepuloh\textsuperscript{1,3}, and K. Wikantika\textsuperscript{1,2,4}

1Center for Remote Sensing, Insitut Teknologi Bandung
2Remote Sensing and GIS Research Group, Faculty of Earth Sciences and Technology, Institut Teknologi Bandung
3Geology Engineering Study Program, Faculty of Earth Sciences and Technology, Institut Teknologi Bandung
4Indonesian Young Researcher Forum (ForMIND)

e-mail: firdausizaharagandes@gmail.com

Abstract: The increasing population of a country has an impact on the number of energy needs that must be provided. Therefore, alternative sources of energy need to be developed to meet those needs. Indonesia is known as The Ring of Fire has very high geothermal resource potential. A study of geothermal mapping technology needs to be done to optimize the utilization of this energy. This study aims to recognize the relation between temperature and surface roughness on an up-flow zone of geothermal system and identify the steam spots at surface. We used a high resolution ALOS-2 PALSAR-2 images to obtain surface roughness and ASTER images to obtain a corrected Land Surface Temperature (LST). The study was conducted in an open area Papandayan Crater, Garut, West Java (Indonesia). Data pre-processing including atmospheric and geometric corrections were applied following an image processing to obtain the surface roughness and the corrected LST by estimating non-unique surface emissivity, pixel fitting between measured surface temperature and roughness at field, and identifying the steam spots of the geothermal system using binary logistic regression method. Equally good models of co-polarization and cross-polarization obtained because of the high correlation between co-polarization and cross-polarization surface roughness. There was a positive relationship between land surface temperature and surface roughness in the manifestation area of geothermal resources and surface roughness known more dominant in determining the manifestations of geothermal resources than surface temperature.

1. Introduction
Geothermal is renewable alternative energy that have some benefit such as provide stable energy resources and environmentally friendly with lower dioxide-carbon than other energy like fossil [1]. The remote sensing method would be an effective solution to do geothermal resources potential mapping remembered study area mostly have a wide area. One of those remote sensing technologies is thermal imagery and Synthetic Aperture Radar (SAR). Thermal imagery provides Land Surface Temperature (LST) of an area, while further processed SAR imagery could give surface roughness in formation. Both LST and surface roughness can be used as a geothermal manifestation indicator.

This research did geothermal manifestation identification use Advanced Spaceborn Thermal Emission and Reflection Radiometer – Thermal Infrared (ASTER-TIR) with 90 m spatial resolution, ASTER – Visible Near Infrared (VNIR) 15 m, and high-resolution SAR imagery from Advanced Land Observing Satellite 2 Phased Array type L-band Synthetic Aperture Radar 2 (ALOS-2 PALSAR-2). The combination and field correction of those data have been done to obtain the geothermal surface manifestation model. The binary regression method was used to obtain the model, this method is an
exponential approach regression for the dichotomous dependent variable. In this research, the dependent variable with 0 value means not a manifestation area and 1 manifestation area.

2. Data and Methods

2.1. Data
The study area is located at Papandayan Crater, West Java, Indonesia (Figure 1). There are 5 data used in this research. They are ASTER TIR, ASTER VNIR Day, ALOS-2 PALSAR-2, and field temperature and surface roughness measurement using fluke sensor and pin-meter, respectively. Field measurement was performed on 18-22 February 2019 with sunny weather and rainy rarely. There were 117 data collected and 10 data among them were collected just above the manifestation spots.

![Figure 1](image.png)

Figure 1. Study area located at West Java, Indonesia (a) and the field points distribution with the green points showing the steam spots and vice versa, showed by subset of the area in the Landsat 8 OLI with the composition of R, G, B = band 4, band 3, band 2 (b).

The ASTER imagery used in this research was acquired on March 3rd 2019, TIR day imagery chose because the available night imagery acquired at a time far enough from the time of field data acquisition (2 weeks adrift). ASTER orbits at an altitude of 795 km with sun-synchronous orbital system at an 98.3° inclination from equator, has 98.99 minute orbital period and temporal resolution 16 days. While ALOS-2 PALSAR-2 imagery used in this research was acquired on April 15th 2015. PALSAR-2 transmitted L-band radio wave (1.2 GHz) with 23.6 wavelength and received the reflected wave from earth surface to obtain information. SAR able to provide day and night imagery without depending on weather because it does not need sunlight for acquisition process. The L-band frequency is not much affected by clouds and rain, and able to penetrate up to the trunk. ALOS-2 PALSAR-2 has some data acquisition mode, stripmap full (quad.) polarimetry HBQ level 1.1 was used in this research. The backscattered value from SAR data is function of surface physical properties namely surface roughness and dialectric parameter, these two parameters is effective to distinguishing surface properties on the volcanic area such as piroclastic, lava, and lava flows [2]. Surface roughness is the most significant parameter to control backscattering value ([3], [4]).

2.2. Land Surface Temperature (LST) Extraction
ASTER imagery was geometry corrected, therefore only atmospheric correction needed to accomplished pre-processing. To achieve accurate LST from space, there arises a need to correct brightness temperatures with emissivity and other parameters. Emissivity is possible to differ from one
place to another, depending on surface roughness and moisture. Previous study show if there is a relationship between Normalized Difference Vegetation Index (NDVI) and emissivity [5]. The proportion of vegetation (sometimes known as fractional vegetation cover) needed to calculate the emissivity. The proportion of vegetation shows the percentage of vacant land and vegetation in determining the value of emissivity [6]:

$$P_v = \frac{(\text{NDVI}_v - \text{NDVI}_s)^2}{\text{NDVI}_v - \text{NDVI}_s}$$  \hspace{1cm} (1)

where NDVIv and NDVIv are the NDVI values of vegetation and soil, respectively. Emissivity is the ratio between radian flux coming out of the actual and blackbody at the same temperature, calculated by Equation (2) [7]:

$$\varepsilon_{13} = 0.968 + 0.022 P_v$$  \hspace{1cm} (2)

where $\varepsilon_{13}$ is emissivity value for band 13 and $P_v$ is the proportion of vegetation obtained from the previous process. Brightness temperature define radian temperature from hypothetical blackbody that emitted same radian at the same wavelength. After digital number converted to radian, the brightness temperature calculated by Normalization Emissivity tool in ENVI. This process obtain emissivity and temperature value, only temperature used in this research.

The Planck function is used to calculate the LST by inverse the Planck equation, this equation is a function of emissivity. Brightness temperature assuming that the earth's surface is a blackbody with emissivity value equal to 1. Planck equation can correct emissivity in the brightness temperature by using Equation (3) [8]:

$$T = \frac{BT}{[1 + (\lambda BT)^2 \times \ln \varepsilon]}$$  \hspace{1cm} (3)

where $T$ is the LST in Kelvin, $BT$ is the brightness temperatures, $\lambda$ is the effective wavelength (band 13= 10.659 μm), $\rho = 1.483 \times 10^{-2}$ mm/K dan $\varepsilon$ represents the emissivity. This correction process increasing the determination coefficient between LST calculated from ASTER imagery and temperature field measurement from 0.6189 to 0.6206. Data selection of field measurements data has done to achieve coefficient determination higher than 0.6, the regression equation used to correct the LST value obtained from the previous process.

2.3. Surface Roughness Extraction

Multilooked, speckle noise corrected, and geocoded and radiometric calibrated ALOS-2 PALSAR-2 imagery was geometrically corrected with GCP RMSE about 0.46 pixels and ICP RMSE 0.44 pixels. Surface roughness processing started with Normalized Radar Cross Section (NRCS) calculation. NRCS is a function of surface roughness and electrical properties [9], and affected by backscattered intensity (Shimada et al., 2007):

$$\sigma_{\eta \xi}^0 = 10 \times \log_{10}(\beta_{\eta \xi}^2) + cf$$  \hspace{1cm} (4)

where $\sigma^0$ is the NRCS value, $\beta$ is the backscatter intensity, and $cf$ is the conversion factor (ALOS PALSAR = -83). The polarization model is indicated by the value $\eta$ (transmitted polarization) and $\xi$ (received polarization). Initial surface roughness calculated after obtaining the NRCS value by using Equation (5) [10]:

$$h_0 (\lambda_{\eta \xi}) = \lambda \left[ -\frac{1}{60} \ln \left( 1 - \frac{\sigma_{\eta \xi}^0}{0.04 \cos \Phi} \right) \right]^{0.5}$$  \hspace{1cm} (5)
where $\sigma^o$ is the NRCS, $\phi$ is the incidence angle, and $\lambda$ is the wavelength (23.6 cm for ALOS 2 PALSAR 2). The crater area tends to have a low initial roughness.

Data selection with the correlation of determination higher than 0.6 for surface roughness field measurements also has been done to correct the imagery processing result. The regression equation applied to the initial roughness result. Combination of polarization has been done to increase the correlation of determination between initial roughness from ALOS-2 PALSAR-2 processing and field data, two combinations namely co-polarization (HH and VV polarization) and cross-polarization (HV and VH polarization) calculated by this equation:

$$h_c = \frac{(h_{o_{\eta\xi}} \times h_{o_{\eta\xi}})^3}{\cos(\phi)}$$

where $h_c$ is the surface roughness value after combination, $h_{o_{\eta\xi}}$ is the initial surface roughness depends on the polarization, and $\phi$ is the incidence angle. Co-polarization and cross-polarization of surface roughness do not make a significant difference.

2.4. Binary Logistic Regression Approach

Logistic regression is a classification algorithm that is used to predict the probability of a categorical dependent variable. In logistic regression, the dependent variable is a binary variable that contains data coded as 1 (yes, success) or 0 (no, failure). This regression used the exponential approach to predict the function and may be used to predict more than 1 independent variable (Equation 7). The assumption of normality is not needed for logistic regression.

$$\text{Logit}(Y) = \ln \frac{Y_i}{1-Y_i} = \alpha + B_1X_1 + B_2X_2 + \cdots + BX_n$$

(7.1)

$$\frac{Y_i}{1-Y_i} = \exp(\alpha + B_1X_1 + B_2X_2 + \cdots + B_nX_n)$$

(7.2)

$$Y(x) = \frac{\exp(\alpha + B_1X_1 + B_2X_2 + \cdots + B_nX_n)}{1+\exp(\alpha + B_1X_1 + B_2X_2 + \cdots + B_nX_n)}$$

(7.3)

where $Y$ is the dependent variable value, $\alpha$ is the regression coefficient, $B$ is the coefficient for each dependent variable, and $X$ is the independent variable (parameter value). Binary logistic regression started with correlation analysis between parameters. If the high correlation was found, it is recommended to select parameters or combine some parameters into a new parameter value. The $Y$ value will be in range 0 to 1, therefore we need to determine the threshold value which will be a boundary between to possible events. The dominant variable can also be determined using binary logistic regression by calculated the $B$ exponent value.

3. Result and Discussion

After processed the ASTER to obtain land surface temperature and ALOS-2 PALSAR-2 to obtain surface roughness, known positive relationship between surface roughness and temperature, proven by correlation coefficient 0.280 for co-polarization model and 0.297 for cross-polarization. The geothermal manifestation model was carried out using binary logistic regression with corrected LST and surface roughness as predictors. Binary logistic regression started with a correlation calculation between the parameters, showed by Table 1.
Table 1. Correlation between parameters for binary logistic regression

|                           | Co-polarization surface roughness | Cross-polarization surface roughness | Land Surface Temperature |
|---------------------------|----------------------------------|--------------------------------------|--------------------------|
| Co-polarization surface roughness | 1                                | 0.999                               | 0.280                    |
| Cross-polarization surface roughness | 0.999                            | 1                                   | 0.297                    |
| Land Surface Temperature   | 0.280                            | 0.297                               | 1                        |

The high correlation between surface roughnesses causes one of those parameters to need to be selected and not involved in the same equation. Obtained to model namely co-polarization and cross-polarization model. Those models expressed mathematically by the equations in showed in Table 2.

Table 2. Binary logistic regression equations

| Model                  | Equations                      |
|------------------------|-------------------------------|
| Co-polarization        | \( \text{Logit } Y = -36.183 + 3.516X_1 + 1.591X_2 \) |
| Cross-polarization     | \( \text{Logit } Y = -36.179 + 3.791X_1 + 1.588X_2 \) |

where \( X_1 \) is the surface roughness and \( X_2 \) is land surface temperature value. The equation obtained from binary logistic regression applied at surface roughness and LST processed before achieved new model showed in Figure 2. Co-polarization and cross-polarization model have same visualization allegedly due to high correlation their surface roughness.

The dependent variable (Y) that calculated by the binary logistic regression equation obtained in the range of 0.0000084 to 0.641. Threshold value determination done by involved-field measurement data (both surface roughness and temperature) to the binary logistic regression equation, manifestation area has Y value higher than 0.1. Hence, determined value of 0.1 as the threshold of geothermal manifestation determination (0 < Y < 0.1 categorized as not a steam spot area, while Y > categorized as steam spot area). The final geothermal manifestation model after threshold determination showed in Figure 3.

![Figure 2. Spatial model of geothermal manifestation model before threshold determination.](image1)

![Figure 3. Spatial model of geothermal manifestation after threshold determination.](image2)
Accuracy test of the final model done by comparing the processing result and field data. Temperature and surface roughness measurements done at 117 points and 10 among them measured just above the steam spot. As many as 8 out of 10 steam spots successfully detected correctly, showed in Figure 4.

The dominant parameter determined using binary logistic regression by calculating the exponent value of the dependent variable coefficient (B value) showed by Table 3. Surface roughness is known as a dominant parameter to determine geothermal potential areas.

Some points were measured in the area of dense vegetation. ALOS-2 PALSAR-2 imagery with 23.6 cm wavelength able to penetrate up to the trunk only, not to the root. That condition will affect the information of surface roughness obtained. Besides, the spatial resolution of Thermal ASTER data used is 90 m, while the field studied area is only 2 km x 2 km. The low spatial resolution of ASTER is certainly very influential on the results of extracting land surface temperature information.

Table 3. Exponent value of dependant variable (B)

| Surface Roughness | Temperature |
|-------------------|-------------|
| Co-polarization manifestation area | 33.637 | 4.906 |
| Co-polarization manifestation area | 44.290 | 4.892 |

4. Conclusion
In this study, the processing of ALOS-2 PALSAR-2 satellite image data to obtain surface roughness information and ASTER imagery to obtain land surface temperature has been carried out. The result of these satellite imagery processing has been corrected with field data measurement. The result is there is a positive relationship between surface roughness and land surface temperature in the geothermal manifestation area with a correlation coefficient of about 0.280 for the co-polarization model and 0.297 for cross-polarization model. Both of the models have the same result allegedly because of the high correlation between the surface roughness combination. As many as 8 out of 10 steam spots points successfully detected using the aforementioned method, the steam spots scattered in the form of fumaroles, hot springs, boiling springs, and steamy ground surface. Finally, the equation to determine potential geothermal manifestation area has been carried out and surface roughness known as a dominant parameter.

Figure 4. Ground checkpoints distribution showing the not steam spots (red) and steam spots successfully detected (green), and the wrong detection (yellow).
5. Acknowledgments

The authors would like to thank LPPM ITB for providing research funding under Multidiscipline Research Scheme and Papandayan Survey Team 2019. The ALOS-2 PALSAR-2 and ASTER data were provided by JAXA under collaborative research with Center for Volcanology and Geological Hazard (CVGHM) Indonesia and Earth Data DAAC using https://search.earthdata.nasa.gov, respectively.

6. References
[1]. Mock, J.E., J.W. Tester, and P.M. Wright, Geothermal energy from the earth: its potential impact as an environmentally sustainable resource. Annual review of Energy and the Environment, 1997. 22(1): p. 305-356.
[2]. Saepuloh, A., et al., SAR-and gravity change-based characterization of the distribution pattern of pyroclastic flow deposits at Mt. Merapi during the past 10 years. Bulletin of Volcanology, 2010. 72(2): p. 221-232.
[3]. Duarte, R.M., et al., Estimation of surface roughness and stone cover in burnt soils using SAR images. Catena, 2008. 74(3): p. 264-272.
[4]. Saepuloh, A., et al., Identifying surface materials on an active volcano by deriving dielectric permittivity from polarimetric SAR data. IEEE Geoscience and Remote Sensing Letters, 2015. 12(8): p. 1620-1624.
[5]. Van de Griend, A. and M. OWE, On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. International Journal of remote sensing, 1993. 14(6): p. 1119-1131.
[6]. Carlson, T.N. and D.A. Ripley, On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote sensing of Environment, 1997. 62(3): p. 241-252.
[7]. Jiménez-Muñoz, J.C. and J.A. Sobrino, A single-channel algorithm for land-surface temperature retrieval from ASTER data. IEEE Geoscience and Remote Sensing Letters, 2009. 7(1): p. 176-179.
[8]. Stathopoulou, M. and C. Cartalis, Daytime urban heat islands from Landsat ETM+ and Corine land cover data: An application to major cities in Greece. Solar Energy, 2007. 81(3): p. 358-368.
[9]. Long, D.G., R. Collyer, and D.V. Arnold, Dependence of the normalized radar cross section of water waves on Bragg wavelength-wind speed sensitivity. IEEE Transactions on Geoscience and Remote Sensing, 1996. 34(3): p. 656-666.
[10]. Campbell, B.A. and M.K. Shepard, Lava flow surface roughness and depolarized radar scattering. Journal of Geophysical Research: Planets, 1996. 101(E8): p. 18941-18951.