K-means clustering of Sentinel-1 time series data analysis for paddy fields mapping: Comparative study of Japan and Indonesia

Sentinel-1 時系列分析における K-means 法クラスタリングを用いた水田域図の作成:日本とインドネシアの比較研究

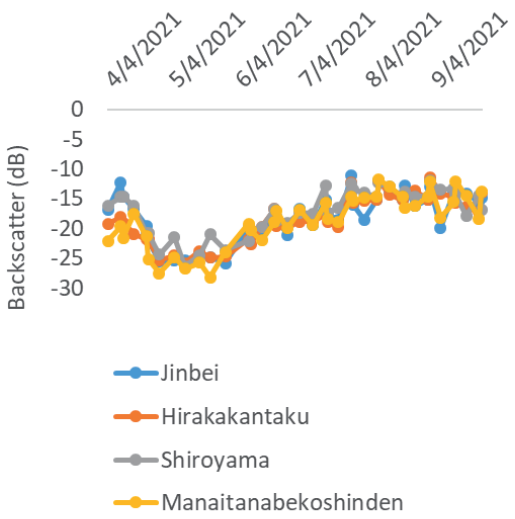
Atiqotun Fitriyah [*], Tasuku Kato[**], and Fumi Okura[***]
〇フィトリヤー アティコトゥン*、加藤 亮**, 大倉芙美***

1. Introduction

Remote sensing application for mapping specific land cover is enormous, and land cover distinction is one of the challenging applications, especially in dense and small-plots agricultural land as in Asian countries. The paddy field mapping mainly relies on single-date satellite imagery. This method might produce errors when the pixels' value coincides with other land covers, such as barren land before cultivation or after harvesting or water body with the condition during the land preparation period. This aspect poses a more significant problem in an ever-dynamic rice cultivation area where the schedule varies even in the same irrigation area. On the other hand, time-series clustering has merit in identifying the dynamic changes in the land cover. Compared to other satellite imagery, Sentinel-1 has superiority due to its open access, high-resolution, and cloud-free SAR images. Sentinel-1 is a powerful tool in studying high cloud coverage areas in recent years. Therefore, we conducted a time series analysis for paddy field mapping using Sentinel-1 in this study.

2 Study area and methods

The comparative study was conducted in two lowland paddy fields area in this study. First is the paddy fields in the Inbanuma Lake area of Chiba Prefecture, Japan. The second is the paddy fields in the upper stream of Indonesia's Bengawan Solo



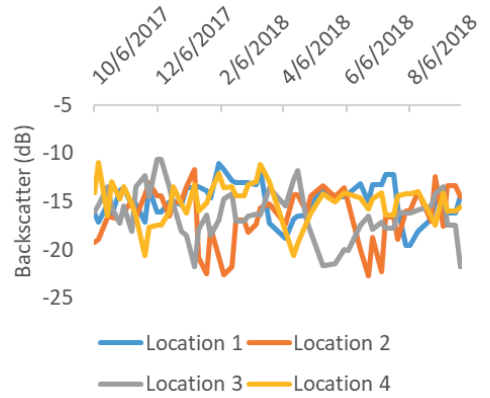


Fig.1 temporal changes of backscattering coefficient σ° for VH polarisation

*東京農工大学農学府農学イノベーションコース International Innovation and Agricultural Science, Tokyo University of Agriculture and Technology **東京農工大学農学研究院 Institute of Agriculture, Tokyo University of Agriculture and Technology ***国際農林水産業研究センター Japan International Research Center for Agricultural Sciences
Keywords:水田、時系列分析, sentinel-1, clustering

river basin. This comparison is intended to see the difference in the mapping performance of paddy fields for the homogenous and heterogenous crop schedules. Unlike paddy cultivation in Japan, paddy can be cultivated multiple times in tropical areas. In actual practice, the cropping schedule could be varied even for farmers in the same irrigation area (the backscatter coefficient change reflected the cropping schedule of rice shown in Fig.1).

In the current study, we used a year-long time series of Sentinel-1 imagery for 2021. Sentinel-1 is C-SAR band imagery produced by the European Space Agency (ESA). We used the dual polarisation Ground Range Detected (GRD) products for both Sentinel-1A and 1B with 10-m spatial resolution, and short revisit time (5-days (1B) and 12-days (1A) temporal resolution for Inbanuma while for Upper solo river area 4 and 8 days temporal resolution (1B). The raw data were accessed through the earthdata.nasa.gov site and then preprocessed using ESA's Sentinel Application Platform (SNAP) software. The preprocessing steps include thermal noise removal, assembling, applying orbit, border noise removal, speckle filter, terrain correction, and conversion of gamma naught from linear backscatter units to logarithmic decibels (dB). In the final step, we performed Euclidean distance k-means clustering to identify the paddy fields area and analyze the dynamics of paddy cropping through the time series data.

3. Ongoing results and discussion

Based on the k-means clustering of the Sentinel-1 backscatter coefficient, the paddy fields in Inbanuma and Upper Solo areas could be detected and distinguished from other land covers (see Fig.2). Paddy fields in the Inbanuma area could be indicated clearly, while in the upper Solo river basin, the river pixels were still mistakenly as paddy fields even in the larger clusters. It may be due to the similarity of the backscatter coefficient value of the river with its neighboring paddy fields. In Inbanuma paddy fields, the backscatter coefficient was very different from that of water bodies so that the paddy area could be differentiated clearly. Though the above issue remains, k-means clustering could classify various time series patterns (as in cropping schedule) of paddy field pixels as one paddy field cluster. Further analysis on the detection of land preparation period and crop schedule is being conducted so that we can complete the mapping of paddy schedule variety over the large area through the satellite imagery.

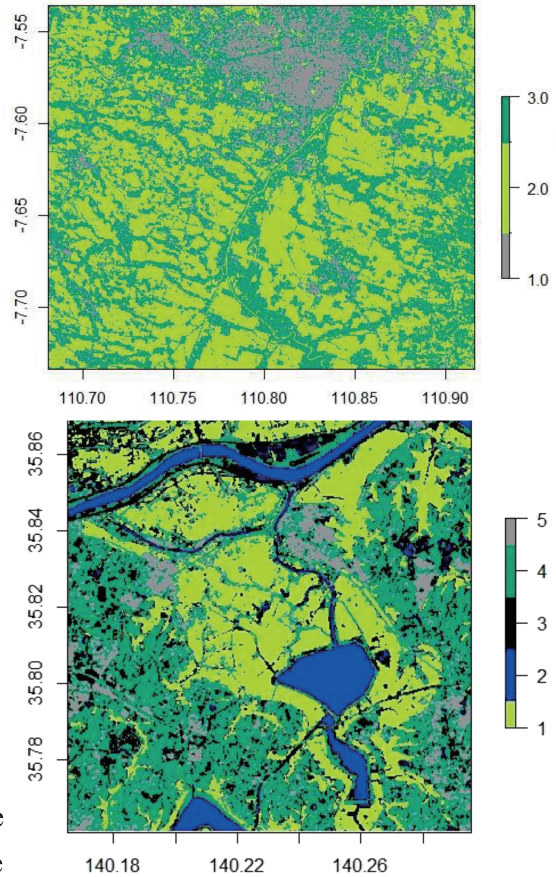


Fig.2 Land classification of Solo river area (up) and Inbanuma area (down) thorugh k-means clustering where the light green area is paddy fields