MAPPING SPATIAL ACCURACY OF FOREST TYPE CLASSIFICATION IN JAXA's HIGH-RESOLUTION LAND USE AND LAND COVER MAP

N. Tsutsumida [1,*], S. Nagai[2], P. Rodríguez-Veiga [1,3,4], J. Katagi[5], K. Nasahara[6], T. Tadono[7]

[1,*] Graduate School of Global Environmental Studies, Kyoto University, Kyoto, Japan - naru@kais.kyoto-u.ac.jp
[2] Department of Environmental Geochemical Cycle Research, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa, Japan - nagais@jamstec.go.jp
[3] Centre for Landscape and Climate Research, University of Leicester, Leicester, UK - prv4@leicester.ac.uk
[4] NERC National Centre for Earth Observation (NCEO), Leicester, UK
[5] Graduate School of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan - kemisknkn37@gmail.com
[6] Faculty of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan - nasahara.kenlo.gw@u.tsukuba.ac.jp
[7] Earth Observation Research Center, Japan Aerospace Exploration Agency (JAXA), Ibaraki, Japan - tadono.takeo@jaxa.jp

Commission III, WG III/2

KEY WORDS: LULC, Geographically weighted, Accuracy assessment, Forest type classification

ABSTRACT:

Accuracy assessment of forest type maps is essential to evaluate the classification of forest ecosystems quantitatively. However, map users do not understand in which regions those forest types are well classified from conventional static accuracy measures. Hence, the objective of this study is to unveil spatial heterogeneities of accuracies of forest type classification in a map. Four forest types (deciduous broadleaf forest (DBF), deciduous needleleaf forest (DNF), evergreen broadleaf forest (EBF), and evergreen needleleaf forest (ENF)) found in the JAXA's land use / cover map of Japan were assessed by a volunteered Site-based dataset for Assessment of Changing LAnd cover by JAXA (SACLAJ). A geographically weighted (GW) correspondence matrix was applied to them to calculate the degree of overall agreements of forest type classes (forest overall accuracy), and the degree of accuracy for each forest class (forest user's and producer's accuracies) in a spatially varying way. This study compared spatial surfaces of these measures with static ones of them. The results show that the forest overall accuracy of the forest map tends to be relatively more accurate in the central Japan, while less in the Kansai and Chubu regions and the northern edge of Hokkaido. Static forest user's accuracy measures for DBF, DNF, and ENF are better than forest producer's accuracy ones, while the GW approach tells us such characteristics vary spatially and some areas have opposite trends. This kind of spatial accuracy assessment provides a more informative description of the accuracy than the simple use of conventional accuracy measures.

1. INTRODUCTION

Land cover (LC) maps describe the Earth's terrestrial surface, encompassing all attributes of the biosphere (International Panel on Climate Change, 2000). LC is a major part of the Earth system which physically interacts with climate, topography, human impacts, and their complex interactions. Remotely sensed (RS) imagery, is often used to produce thematic LC maps aiming to understand wide ranges of terrestrial environments from the properties of the land cover itself such as the forest / non-forest classification (Hansen et al., 2013; Shimada et al., 2014), the urban extent (Schneider et al., 2009), the cropland distribution (Xiong et al., 2017) to the application of biodiversity estimation (Tuanmu and Jetz, 2014), carbon stock estimation (Rodríguez-Veiga et al., 2017), and potential ecosystem services assessment (Andrew et al., 2014). It is hence important to make an accurate LC classification map for high-quality quantification of these properties.

The classification of plant functional types (PFTs) is one of the main and most challenging tasks toward accurate LC mapping. PFT mapping can be considered as a land cover classification in which plants are grouped with regard to their physiology and physiognomy (Moncrieff et al., 2016). In Japan, the forest type classification (i.e. deciduous broadleaf forest (DBF), deciduous needleleaf forest (DNF), evergreen needleleaf forest (ENF), and

evergreen needleleaf forest (ENF)) is required for multiple purposes, as the tree cover rate is reported as approximately 68.5% of the entire of the territory (The World Bank, 2018). However, the accurate classification in the four forest types from RS data has not yet been achieved. Due to the similarity of spectral characteristics captured by satellite sensors, even the separability between Evergreen and Broadleaf forest types is difficult (Sharma et al., 2016).

Some LC products are available to classify PFT-based forest types in Japan. MODerate resolution Imaging Spectroradiometer (MODIS) Land Cover Type Products (MCD12Q1) (Friedl et al., 2010) has a 500 m spatial resolution with the world standard international geosphere-biosphere programme (IGBP) LC class definition and the 12 PFT classification scheme. Climate Change Initiative (CCI) LC data set developed by the European Space Agency (ESA) describes LC by 22 thematic classes at 300 m spatial resolution at a global scale (Bontemps et al., 2013). The Ministry of Environment Japan investigates detailed PFTs in Japan by the combination of field survey and RS image analyses and publishes a 1 km resolution grid map (Biodiversity Center of Japan, 2018). These well-known products are coarse and may stem the mixed pixel problem: pixels will typically contain a mix of LC classes and the propensity for this is greater with medium and coarse spatial resolution data, as every pixel represents the dominant category and neglect the presence of others, resulting in the failure of accurate representations of LC due to the coarse pixel size (Fisher and Pathirana, 1990). Recently, the Japan

---

* Corresponding author

Aerospace Exploration Agency (JAXA) has developed and published a High-Resolution Land Use and Land Cover (HRLULC) map (version 18.03) with 10 thematic classes at the spatial resolution of 30 m for the whole Japan territory (JAXA, 2018) and is expected to overcome such issue.

It is important for users to know how accurate the LC classification map is. Incorrect conclusions or decisions may be derived from the map if mis-classification in the map is neglected (Daly, 2006). The accuracy assessment provides a guide to the quality of the data and their reliability (Foody, 2002). The accuracy of the classification map is usually reported using a reference (ground-truthing) sample. A confusion matrix, which is a cross-tabulation of classification and reference sample categories, is built to report measures of overall, user's, and producer's accuracy (Congalton, 1991, Tsutsumida et al, 2016). Although these accuracy measures are utilized in order to report the degree of correspondence of the classification against a reference sample, they are averaged summaries of the correspondence, meaning that spatial variation of accuracy is not taken into account. Spatial heterogeneity of the accuracy may exist especially when the classification model does not consider spatial configuration of data. The knowledge of such spatial variations of accuracy is required to inspect the map not only for understandings of the spatial distribution of forest types accurately, but also for minimizing error propagation when we analyze this map for further applications. Previous studies have demonstrated a means to represent spatial heterogeneity of accuracy for categorical raster data (Comber et al., 2012, 2017; Comber, 2013; Congalton, 1988; Foody, 2005). They focused on spatial dependency and / or non-stationarity of accuracy measures which are often found in RS-based LC classification map. Specifically, Comber et al. (2017) proposed a geographically weighted (GW) correspondence matrix which is able to build a local correspondence matrix with a spatially weighted pairwise of classification and reference sample data. The GW uses a distance-decayed moving window or kernel over space to calculate any types of models and statistical measures (Gollini et al., 2015; Lu et al., 2014).

Hence, the objective of this study is to unveil spatial heterogeneities of accuracies for four forest type classes in a LC classification map. We use the JAXA's HRLULC map of Japan for the case study. In order to validate estimated classes in the map against the actual land cover situation, we used a ground truth reference data set, called site-based dataset for assessment of changing land cover by JAXA (SACLAJ). A GW correspondence matrix was applied to these data sets to analyze the degree of overall agreements of forest type classes and the degree of accuracy for each forest class in a spatially varying way. This allows to map spatial heterogeneities of such degrees over space, hidden in conventional accuracy measures. Results provide us insightful information for the usefulness of the JAXA's HRLULC map with special attention to local accuracies of forest types against the actual land covers.

2. MATERIALS

2.1 JAXA's High-Resolution Land Use and Land Cover map (HRLULC map)

The JAXA'a HRLULC map version 18.03 at the 30 m spatial resolution is used shown in Figure 1 (JAXA, 2018). This data describes LC in the entire territory of Japan during the period 2014-2016 with 10 thematic classes: water; urban; rice paddy; crop; grass; DBF; DNF; EBF; ENF; and bareland. This product uses multiple geospatial data sets for inputs such as Landsat 8,

digital elevation model provided by Geospatial Information Authority of Japan, ALOS-2/PALSAR-2, Sumoi National Polar-orbiting Partnership (NPP), road geospatial data in OpenStreetMap, and ground truth sampling data in SACLAJ (details below). A kernel density estimation in a Bayesian inference was applied to obtain the class posterior probability and produce the JAXA'a HRLULC map (Hashimoto et al., 2014).
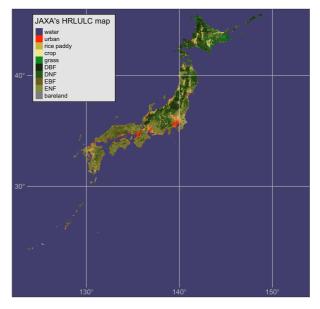

Figure 1. JAXA's land use land cover map.

2.2 SACLAJ

3,014 reference sample points in SACLAJ are used for validation in this study. Such sample points are randomly distributed across the entire territory of Japan and are captured during the period 2014-2016, which is the same as the target period of the JAXA HRLULC map. Approximately 300 data points are allocated in each class category of which definition is the same as the one in the JAXA's HRLULC map. Scientists and engineers who visited field sites took geo-referenced pictures on the ground and interpreted them into one of the land cover categories (Nagai et al., accepted; Tadono et al., 2014). Such labels are also confirmed by very fine spatial resolution satellite images in geospatial visualization tools such as Google Earth. Since 2013, more than 50,000 data points with information on date, LC category, location, and size of the LC, have been recorded across Japan and all over the world.

3. METHODS

3.1 Correspondence matrix

For the purpose of the accuracy assessment of a classification map, it is common to build a correspondence matrix between the classified and reference values and to calculate some accuracy measures. At lth point in total L sample points considered, a pair of the classified class (c_l) and the reference class (r_l) is obtained. A correspondence matrix (M) is built so that rows represents classified data, and columns represent reference data, and the number of correspondence pairs between c_l and r_l corresponding to the class i in the sample is stored at the ith row and ith column in the M. Similarly, the number of pairs between c_l corresponding to the class i and r_l corresponding to the class j in the sample is stored at the ith rows and the jth columns in the

M. Thus, the element m_{ij} of the M at the ith row and the jth column is:

$$m_{ij} = \Sigma_{l=1}^{L} a_{ij(l)},$$
$$a_{ij(l)} = \begin{cases} 1 & \text{if } c_l = i \text{ and } r_l = j, \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $a_{ij(l)}$ is a binary response whether the pair of the classification class c_l and the reference data r_l at the lth of L sample points are the class i and the class j, respectively, or not. It is noted that a coordination (u_l, v_l) at the lth sample point is not considered in the m_{ij}. In this way, M represents a summary of the degree of agreement and disagreement among classes.

Generally, accuracy measures are calculated from the matrix M to assess classification maps and to tell data users how the degree the correspondence between classified and reference class at the sample points is accurate. The overall accuracy describes the proportion of the number of pairs having the same class in the classified and reference data for all classes in total. User's and producer's accuracies are also often reported to represent how the degree the classification is accurate at a class level. User's accuracy describes the proportion of the number of classified class at sample points that actually corresponds to the reference class, whereas producer's accuracy is the proportion of the number of reference class at sample points that correctly corresponds to the classified class.

In this study, we would like to know how much correctly forest types are classified into four forest classes. We extend the concept of overall, user's and producer's accuracy measures to focus only on forest related classes. We define forest overall, forest user's and forest producer's accuracies as:

$$FO = \frac{\Sigma_{i=f_1}^{f_n} m_{ii}}{\Sigma_{i=f_1}^{f_n} \Sigma_{j=f_1}^{f_n} m_{ij}} \quad (2)$$

$$FU_i = \frac{m_{ii}}{\Sigma_{f_1}^{f_n} m_{i+}} \quad (3)$$

$$FP_i = \frac{m_{ii}}{\Sigma_{f_1}^{f_n} m_{+i}} \quad (4)$$

where FO is forest overall accuracy, and FU_i and FP_i are forest user's and producer's accuracies of the class $i \in f_1, \ldots, f_n$, respectively. Here the f_n is the nth forest class and in this study, $n = 4$. $\Sigma_{f_1}^{f_n} m_{i+}$ denotes the marginal sum of m for the classification class i in n forest classes, and $\Sigma_{f_1}^{f_n} m_{+i}$ denotes the marginal sum of m for the reference class i in n forest classes.

These measures are averaged values with different aspects of map accuracy. Although the result is represented by a 'value' which describes how much degrees the overall agreement rate between the classified and the reference class, the classified class corresponds to the one in the reference class, or the reference class corresponds to the one in the classified class, these measures do not consider any spatial configurations in the sample points. There have been argued that they only function when errors are assumed to be spatial homogeneous, while the error in the classification map often represents spatial heterogeneity (Comber et al., 2017; Comber, 2013; Foody, 2005; Tsutsumida and Comber, 2015; Tsutsumida et al., 2019). It is important to tackle with this issue.

3.2 Geographically weighted correspondence matrix

A GW correspondence matrix has been introduced by Brunsdon et al. (2016) and Comber et al. (2017). It builds locally weighted correspondence matrix at each location by a moving-window distance-decayed kernel. Again, we consider the equation (1) but deal with the spatial structure of the sample points by applying GW framework. Here at the location k on a grid with the spatial resolution of r covered in the study area and ω_k is a bisquared kernel defined as:

$$\omega_k = \begin{cases} \left(1 - \left(\frac{d_k}{b}\right)^2\right)^2 & \text{if } |d_k| < b, \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where b is the bandwidth size, which is arbitrary determined. Larger b tends to describe more spatial homogeneity (stationary) closer to the global (conventional) measures of accuracy. The total number of grids K is determined by the spatial resolution r in the study area. The weighted element m_{ij} of the local matrix M_k at the location k is now replaced by the $\rho_{ij(k)}$:

$$\rho_{ij(k)} = \Sigma_{k=1}^{K} \omega_k a_{ij(k)} \quad (6)$$

Sample points found within the distance b from the location k are only used for the calculation of the GW correspondence matrix. As this way, at location k, the GW forest overall ($GWFO_{(k)}$), and GW forest user's, and GW forest producer's accuracy measures for the class i ($GWFU_{i(k)}$ and $GWFP_{i(k)}$, respectively) are:

$$GWFO_{(k)} = \frac{\Sigma_1^{f_n} \rho_{ii}}{\Sigma_{j=f_1}^{f_n} \Sigma_{i=f_1}^{f_n} \rho_{ij}} \quad (7)$$

$$GWFU_{i(k)} = \frac{\rho_{ii}}{\Sigma_{f_1}^{f_n} \rho_{i+}} \quad (8)$$

$$GWFP_{i(k)} = \frac{\rho_{ii}}{\Sigma_{f_1}^{f_n} \rho_{+i}} \quad (9)$$

where the class $i, j \in f_1, \ldots, f_n$. $\Sigma_{f_1}^{f_n} \rho_{i+}$ denotes the marginal sum of ρ for the classification class i in n forest classes, and $\Sigma_{f_1}^{f_n} \rho_{+i}$ denotes the marginal sum of ρ of the reference class i in n forest classes. In this study, b and r are defined as 200 km and 5 km, respectively, to show regional variations of accuracy measures well with relatively less calculation time.

4. RESULTS

In order to confirm the whole-map accuracy at first, the conventional correspondence matrix with JAXA's HRLULC map and SACLAJ reference sample was built (Table 1). The conventional overall accuracy is 82.2% for 10 classes. This table suggests relatively better classifications for water (user's accuracy: 95.3%, producer's accuracy: 97.7%), while some mis-classifications are also found such as between urban and bareland, and between crop and grass. When focusing on the classification between aggregated forest and non-forest classes, 97.4% of the overall agreement rate is achieved, suggesting the high performance of the forest/non-forest distinction. From this result, our interests of accuracy in this map move forwards to the performance of the four forest types classification.

Classified	water	urban	rice paddy	crop	grass	DBF	DNF	EBF	ENF	bareland	user's accuracy (%)
						Reference					
water	304	1	0	1	1	0	0	0	0	12	95.3
urban	1	254	5	3	2	0	0	1	0	30	85.8
rice paddy	1	3	287	34	6	0	0	0	0	5	85.4
crop	0	7	5	220	14	0	0	1	0	15	84.0
grass	0	1	1	28	248	9	3	2	1	25	78.0
DBF	0	0	0	3	14	238	26	16	6	9	76.0
DNF	0	0	0	2	0	29	252	7	4	3	84.6
EBF	1	0	0	2	8	13	4	218	24	4	79.6
ENF	4	0	0	1	1	11	15	56	266	8	73.5
bareland	0	34	0	6	6	0	0	0	0	190	80.5
producer's accuracy (%)	97.7	84.7	95.7	73.3	82.7	79.3	84.0	72.4	88.4	63.1	

Table1. Correspondence matrix of JAXA's land use and land cover map and SACLAJ reference sample. The overall accuracy is 82.2%. Grey color represents a subset matrix related to forest-related classes.

Table 2 represents a summary of FU and FP. All values in FU for four forest types except the ENF class are larger than those in FP, suggesting that these forest classes in the classification map tend to be identified at the reference points correctly well, compared to that in the reference against the classification map. The ENF class shows the opposite pattern. The FO is 82.2%.

	DBF	DNF	EBF	ENF
forest user's accuracy (%)	83.2	86.3	84.2	76.4
forest producer's accuracy (%)	81.8	84.8	73.4	88.7

Table2. Forest user's and producer's accuracy for deciduous broadleaf forest (DBF), deciduous needleleaf forest (DNF), evergreen broadleaf forest (EBF), and evergreen needleleaf forest (ENF).

As expected, such measures inform us the degree of correspondence of classes in the whole map, however it lacks spatial information on accuracy. Thus, GW correspondence matrix was applied to show spatial heterogeneities in accuracy measures. Figure 2 shows a result map of GWFO. It clearly shows spatial heterogeneity of the FO hidden in the conventional overall accuracy measure. The GWFO varies around the FO: 82.2% and on the whole, but it tends to be more accurate in central Japan while less in the Kansai (western mainland) and the northern edge of the main island and Hokkaido (north island).
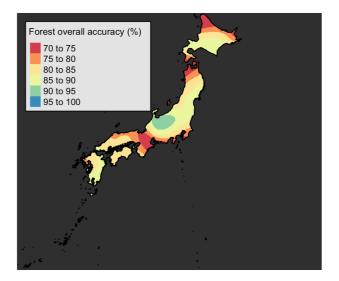

Figure 2. Geographically weighted forest overall accuracy.

Maps of both GWFU and GWFP shown in Figure 3 and 4 demonstrate clear spatial variations of accuracy. The highest variation can be found in the EBF for both of the FU and the FP (the standard deviations are 25.3 and 30.5, respectively). Surfaces of accuracy in some areas in the Japanese territory lacks due to insufficient number of classification and reference sample points, implying that there is no distribution of the class found in such areas. Each map indicates potential spatial clusters of high / low accuracy. For example for the DNF, both FU and FP represent a cluster with high accuracy (over 90%) in the central mainland. It is reasonable to find highly accurate clusters in such regions as the DNF is distributed in cooler regions such as Hokkaido and high-elevated mountainous areas in Japan (mostly found in the central Japan) (Oshida et al., 2009). In the meantime, the DNF are relatively less found in other areas, meaning the

classification tends to fail (less accurate). Relatively lower FU and FP for the EBF are shown especially in the northern mainland and Hokkaido. This occurs because the actual distribution of the EBF can be found in warmer regions below around 37° north latitude (Oshida et al., 2009) and thus the EBF found in the northern Japan tends to be mis-classified. While Table 2 describes FU for DBF, DNF, and ENF are higher than FP, Figure 3 and 4 tell us such characteristics depend on regions and some areas have the opposite trend. For example, FP is larger than FU in the central mainland and the northern edges of Hokkaido for the DBF, in the southern Hokkaido and the north mainland for the DNF, and in the western mainland for the EBF. In contrast, FP is larger than FU for the ENF according to Table 2, however, some areas with larger FU than FP can be found in the middle of Hokkaido and the north-western mainland. These spatial characteristics of accuracy are hidden in the global accuracy diagnostics. Such maps of spatial heterogeneities of accuracy enable data users to pay special attention to the regional difference of the quality of the JAXA's HRLULC map and to areas which have opposite trends from global results.
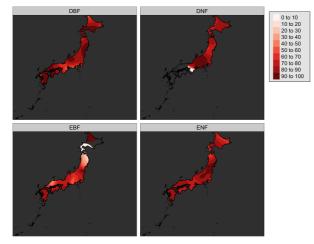

Figure 3. Geographically weighted forest user's accuracy for DBF, DNF, EBF, and ENF.
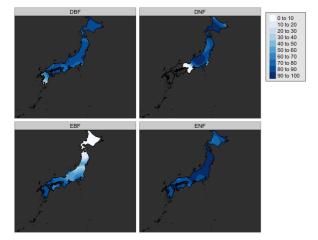

Figure 4. Geographically weighted forest producer's accuracy for DBF, DNF, EBF, and ENF.

5. DISCUSSIONS

Accurate forest type classification is challenging. To understand the quality of a LC classification map, data users have to inspect the accuracy of the map in which regions are well described or not. It has been difficult to infer in which areas these forest types (a major part of PFT) are mis-classified (confused) although forest / non-forest classification has been clearly available. This is a limitation of understandings of the actual spatial distribution of the PFT. To support this, this study highlighted the spatial heterogeneities of accuracy for the forest type classification in the JAXA's HRLULC map using a GW framework as a case study. This study generated a map of GWFO which tells spatial accuracy of the four forest types classification in the JAXA's HRLULC map based on the SACLAJ reference sample. Maps of GWFU and GWFP also describe local accuracies but focus on individual classes. All forest types have spatially-varying surfaces of accuracy, indicating the spatial heterogeneity of accuracy. Such descriptions are informative for data users who are interested in accurate maps of forest types to confirm the quality of the classification over space.

There are some technical discussions which help build bridges to further potential studies. Firstly, the bandwidth size is arbitrary determined (in this study, 200 km). This was explored preliminarily and determined so that the kernel enables to cover sufficient reference sample points over space. For a conventional GW regression (GWR) and other types of GW regression models such as generalized GWR, some optimization approaches for the bandwidth size have been developed. One of the popular methods is a leave-one-out cross validation (LOOCV) (Gollini et al., 2015; Lu et al., 2014). The LOOCV approach can also be applicable to GW summary statistics (Brunsdon et al., 2002; Fotheringham et al., 2002). However, there have not yet proposed such optimization approaches for the GW correspondence matrix. Even if the bandwidth can be optimized for the GW correspondence matrix, it does not produce optimized GW accuracy measures, because the LOOCV calculation requires a statistic which represents a nature of the matrix, and it is not optimal for accuracy measures. Furthermore, the LOOCV can be applicable to each of GW accuracy measures, however optimized bandwidth would not be consistent through measures, resulting in the failure of their comparisons. The topic on bandwidth optimization is highly linked with the question which spatial scale level is appropriate to investigate local accuracy in the map. Hence new developments for the bandwidth optimization are expected. Similar discussion can be found in Tsutsumida et al. (2019). Secondly, the quality of the labelling in reference samples is not considered in this study. Ultimately, any results of accuracy assessments are relative to the reference sample. The SACLAJ reference data set is well organized collected from volunteers who visited field sites and / or from very fine spatial resolution satellite imageries. However, no reports on the quality are found. Previous studies have argued the impact of the use of imperfect reference sample to accuracy assessments (Foody, 2011; Foody et al., 2016; Zhao et al., 2017) and such findings would be helpful. Furthermore, potential studies can focus on the uncertainty between a classification map and imperfect reference sample when these reliabilities are available, by applying GW fuzzy differences (Comber et al., 2012) and / or GW mean absolute error (Tsutsumida et al., 2019). Lastly, our approaches evaluated local accuracy but did not contribute to the improvement of the classification. Future studies will focus on the usefulness of such information on spatial accuracy to improve the classification. A spatial backpropagation approach based on the GW error measures would be worth to be investigated toward minimizing the spatial heterogeneity of accuracy.

6. CONCLUSION

The forest type classification, a major part of the PFT, is essential to estimate aboveground biomass, carbon cycle, biodiversity, and many other environmental studies. Accurate forest type classification maps are required for these applications, however such map has not yet been achieved. This study demonstrated how to estimate the spatial accuracy of classification in four forest types in Japan to inform users the reliability of the classification map. Three types of accuracy measures tell different aspects of local accuracy: the degree of overviews of the correspondence of the classification to reference sample; the degree of how much reference sample points are correctly classified in the predicted map; and the degree of how much classified sample are found in the reference sample. Particularly, this study applied geographically weighted approaches to the forest type classification in the JAXA's HRLULC map and highlighted spatial heterogeneities of these three types of accuracy measures. The spatial accuracy provides a more informative description of accuracy than the simple use of some conventional accuracy measures.

ACKNOWLEDGEMENTS

REFERENCES

Andrew, M.E., Wulder, M.A., Nelson, T.A., 2014. Potential contributions of remote sensing to ecosystem service assessments. Progress in Physical Geography 38(3), pp. 328–353 https://doi.org/10.1177/0309133314528942.

Biodiversity Center of Japan, 2018. Natural Environment Survey http://www.biodic.go.jp/ne_research_e.html#id01 (15 December 2018).

Bontemps, S., Defourny, P., Radoux, J., Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., Kirches, G., Zülkhe, M., Kalogirou, V., Seifert, F., Arino, O., 2013. Consistent global land cover maps for climate modelling communities: current achievements of the ESA' land cover CCI, in: Proceedings of the Esa Living Planet Symposium.

Brunsdon, C., Charlton, M., Harris, P., 2016. Geographically Weighted Contingency Tables https://rpubs.com/IIASA/gwxtabs (15 December 2018).

Brunsdon, C., Fotheringham, A.S., Charlton, M., 2002. Geographically weighted summary statistics — a framework for localised exploratory data analysis. Computers, Environment and Urban Systems 26, pp. 501–524.

Comber, A., Brunsdon, C., Charlton, M., Harris, P., 2017. Geographically weighted correspondence matrices for local error reporting and change analyses: mapping the spatial distribution of errors and change. Remote Sensing Letters 8(3), pp. 234–243 https://doi.org/10.1080/2150704X.2016.1258126

Comber, A.J., 2013. Geographically weighted methods for estimating local surfaces of overall, user and producer accuracies. Remote Sensing Letters 4(4), pp. 373–380 https://doi.org/10.1080/2150704X.2012.736694

Comber, A.J., Fisher, P., Brunsdon, C., Khmag, A., 2012. Spatial analysis of remote sensing image classification accuracy. Remote Sensing of Environment 127, pp. 237–246 https://doi.org/10.1016/j.rse.2012.09.005.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37, pp. 35–46 https://doi.org/10.1016/0034-4257(91)90048-B.

Congalton, R.G., 1988. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric engineering and remote sensing 54(5), 587–592.

Daly, C., 2006. Guidelines for assessing the suitability of spatial climate data sets. International Journal of Climatology 26(6), pp. 707–721 https://doi.org/10.1002/joc.1322.

Fisher, P., Pathirana, S., 1990. The evaluation of fuzzy membership of land cover classes in the suburban zone. Remote Sensing of Environment 34(2), pp. 121–132 https://doi.org/10.1016/0034-4257(90)90103-S.

Foody, G.M., 2011. Impacts of imperfect reference data on the apparent accuracy of species presence-absence models and their predictions. Global Ecology and Biogeography 20, pp. 498–508 https://doi.org/10.1111/j.1466-8238.2010.00605.x.

Foody, G.M., 2005. Local characterization of thematic classification accuracy through spatially constrained confusion matrices. International Journal of Remote Sensing 26(6), pp. 1217–1228 https://doi.org/10.1080/01431160512331326521.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment 80(1), pp. 185–201 https://doi.org/10.1016/S0034-4257(01)00295-4.

Foody, G., Pal, M., Rocchini, D., Garzon-Lopez, C., Bastin, L., 2016. The Sensitivity of Mapping Methods to Reference Data Quality: Training Supervised Image Classifications with Imperfect Reference Data. ISPRS International Journal of Geo-Information 5(11), 199. https://doi.org/10.3390/ijgi5110199.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Chichester.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment 114(1), pp. 168–182 https://doi.org/10.1016/j.rse.2009.08.016.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., Harris, P., 2015. GWmodel : An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. Journal of Statistical Software 63(17), pp. 85–101 https://doi.org/10.18637/jss.v063.i17.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S. a, Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science 342(6160), pp. 850–853 https://doi.org/10.1126/science.1244693.

Hashimoto, S., Tadono, T., Onosato, M., Hori, M., 2014. A New Method to Derive Precise Land-use and Land-cover Maps Using Multi-temporal Optical Data. Journal of The Remote Sensing of Japan 34(2), pp. 102–112 (in Japanese).

International Panel on Climate Change, 2000. Land Use, Land-Use Change, and Forestry. Cambridge University Press, Cambridge. https://doi.org/10.2277/0521800838.

JAXA, 2018. Homepage of High-Resolution Land Use and Land Cover Map Products https://www.eorc.jaxa.jp/ALOS/en/lulc/lulc_index.htm (15 December 2018).

Lu, B., Harris, P., Charlton, M., Brunsdon, C., 2014. The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models. Geo-spatial Information Science 17(2), pp. 85–101 https://doi.org/10.1080/10095020.2014.917453.

Moncrieff, G.R., Bond, W.J., Higgins, S.I., 2016. Revising the biome concept for understanding and predicting global change impacts. Journal of Biogeography 43(5), pp. 863–873 https://doi.org/10.1111/jbi.12701.

Nagai, S., Nasahara, K., Akitsu, T., Saitoh, T., Muraoka, H. Importance of the collection of abundant ground-truth data for accurate detection of spatial and temporal variability of vegetation by satellite remote sensing, in: Biogeochemical Cycles: Ecological Drivers and Environmental Impact. AGU Books, accepted.

Oshida, T., Masuda, R., Ikeda, K., 2009. Phylogeography of the Japanese giant flying squirrel, Petaurista leucogenys (Rodentia: Sciuridae): implication of glacial refugia in an arboreal small mammal in the Japanese Islands. Biological Journal of the Linnean Society 98(1), pp. 47–60 https://doi.org/10.1111/j.1095-8312.2009.01276.x.

Rodríguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K., Balzter, H., 2017. Quantifying Forest Biomass Carbon Stocks From Space. Current Forestry Reports 3(1), pp. 1–18 https://doi.org/10.1007/s40725-017-0052-5.

Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environmental Research Letters 4(4), 044003 https://doi.org/10.1088/1748-9326/4/4/044003.

Sharma, R.C., Tateishi, R., Hara, K., Iizuka, K., 2016. Production of the Japan 30-m land cover map of 2013-2015 using a random forests-based feature optimization approach. Remote Sensing 8(5) https://doi.org/10.3390/rs8050429.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data (2007-2010). Remote Sensing of Environment 155, pp. 13–31 https://doi.org/10.1016/j.rse.2014.04.014.

Tadono, T., Motohka, T., Nasahara, K.N., Kobayashi, K., Watanabe, T., Doutsu, M., Dan, R., 2014. Introduction of "SACLA-J" Site-based dataset for Assessment of Changing LAnd cover by JAXA https://www.eorc.jaxa.jp/ALOS//kyoto/dec2014_kc21/pdf/3-07_KC21_SACLAJ_Tadono-JAXA.pdf (15 December 2018).

The World Bank, 2018. Forest area (% of land area) | Data https://data.worldbank.org/indicator/AG.LND.FRST.ZS (15 December 2018).

Tsutsumida, N., Comber, A.J., 2015. Measures of spatio-temporal accuracy for time series land cover data. International Journal of Applied Earth Observation and Geoinformation 41, pp. 46–55 https://doi.org/10.1016/j.jag.2015.04.018.

Tsutsumida, N., Comber, A.J., Saizen, I., 2016. Challenges in spatio-temporal land cover classification and its accuracy assessment, in: Katsumi, T., Hashimoto, S. (Eds.), Towards Future Earth: Challenges and Progress of Global Environmental Studies. Kaisei Publishing, pp. 61–79.

Tsutsumida, N., Rodríguez-Veiga, P., Harris, P., Balzter, H., Comber, A., 2019. Investigating spatial error structures in continuous raster data. International Journal of Applied Earth Observation and Geoinformation 74, pp. 259–268 https://doi.org/10.1016/j.jag.2018.09.020.

Tuanmu, M.-N., Jetz, W., 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Global Ecology and Biogeography 23(9), pp. 1031–1045 https://doi.org/10.1111/geb.12182.

Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., Thau, D., 2017. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS Journal of Photogrammetry and Remote Sensing 126, pp. 225–244 https://doi.org/10.1016/j.isprsjprs.2017.01.019.

Zhao, Y., Feng, D., Yu, L., See, L., Fritz, S., Perger, C., Gong, P., 2017. Assessing and Improving the Reliability of Volunteered Land Cover Reference Data. Remote Sensing 9(10), 1034. https://doi.org/10.3390/rs9101034.