Spatial heterogeneity of errors in land cover data

Narumasa Tsutsumida
Graduate School of Global Environmental Studies,
Kyoto University

Introduction

Thematic Land cover (LC) maps attempt to describe the Earth's terrestrial surface, encompassing all attributes of the biosphere (International Panel on Climate Change, 2000). LC has been regarded as an important component of the Earth system which physically interacts with climate, topography, human impacts, and their complex interactions. As LC maps are required to cover an area widely from local to global scales, remotely sensed (RS) imagery is often used, that is classified into defined thematic land cover classes by a classification method such as statistical and machine learning models. It is hence important to make an accurate LC classification map for high-quality quantification of the Earth system component. To assess the accuracy of the thematic LC classification map, conventional summary measures of error, such as user's, producer's, and overall accuracies for per-pixel classification, and mean signed deviation (msd), mean absolute error (mae), root mean square error (rmse) and Pearson's correlation coefficient (r) for sub-pixel classification. However, these summary measures of error do not take any spatial information (e.g., spatial heterogeneity) of error into account (Foody, 2005, 2002). A spatially explicit approach for the assessment is helpful to identify spatial characteristics of errors. This study demonstrates one of the spatial measures of error for assessing thematic LC maps. In this paper, a map for forest aboveground biomass (AGB) in the Yucatan peninsula, Mexico, estimated by Rodríguez-Veiga et al. (2016), is assessed.

Methods

A geographically weighted approach is used to generate spatial surfaces of error measures. The GW approach calculates a series of local error measures, using data weighted by their distance to the center of a moving window or kernel to explore spatial heterogeneity (Gollini et al., 2015). The GW versions of msd, mae, rmse, and r are calculated as follows. At any location i in the study area, GW msd: $gw.msd(x_i, y_i)$, GW mae: $gw.mae(x_i, y_i)$, and GW rmse: $gw.rmse(x_i, y_i)$ are defined as:

$$gw.msd(x_i, y_i) = \frac{\Sigma_{j=1}^{n} \omega_{ij}(y_j - x_j)}{\Sigma_{j=1}^{n} \omega_{ij}} \quad (1)$$

$$gw.mae(x_i, y_i) = \frac{\Sigma_{j=1}^{n} \omega_{ij}|y_j - x_j|}{\Sigma_{j=1}^{n} \omega_{ij}} \quad (2)$$

and

$$gw.rmse(x_i, y_i) = \sqrt{\frac{\Sigma_{j=1}^{n} \omega_{ij}(y_j - x_j)^2}{\Sigma_{j=1}^{n} \omega_{ij}}} \quad (3)$$

where x_j and y_j are the reference and predicted values at sample location j, respectively, ω_{ij} weights controlled by a distance-decay kernel function (Equation 8) with respect to location i and j, and n is the total number of sample data points. Observe that this always holds, msd ≤ mae ≤ rmse (Willmott and Matsuura, 2005), and their GW counterparts have the same characteristics.
A GW *r* at any location i, is found using:

$$gw.cor(x_i, y_i) = \frac{c(x_i, y_i)}{s(x_i)s(y_i)} \quad (4)$$

where a GW standard deviation: $s(x_i)$ is

$$s(x_i) = \sqrt{\frac{\Sigma_{j=1}^{n} \omega_{ij}(x_j - m(x_j))^2}{\Sigma_{j=1}^{n} \omega_{ij}}} \quad (5)$$

and a GW mean: $m(x_i)$ is

$$m(x_i) = \frac{\Sigma_{j=1}^{n} \omega_{ij} x_j}{\Sigma_{j=1}^{n} \omega_{ij}} \quad (6)$$

with a GW covariance: $c(x_i, y_i)$

$$c(x_i, y_i) = \frac{\Sigma_{j=1}^{n} \omega_{ij}[(x_j - m(x_j))(y_j - m(y_j))]}{\Sigma_{j=1}^{n} \omega_{ij}} \quad (7)$$

For both case studies, the weights ω_{ij} are found using a bi-square kernel as follows:

$$\omega_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b_i}\right)^2\right)^2 & \text{if } |d_{ij}| < b_i, \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where d_{ij} is the Euclidean distance between locations i and j, and the kernel bandwidth b_i is specified as an adaptive distance, which includes a fixed number of data points for the local error calculation. Its size was arbitrarily defined as 10% of nearby data to location *i*. Observe that the chosen diagnostics complement each other: measures of msd, mae, and rmse and their GW counterparts, all summarize the error in some manner, while *r* and GW *r* measure specifically the slope of the linear relationship between the predicted and reference values. Furthermore, *r* and GW *r* are scale-invariant, meaning that they cannot capture a consistent and uniform over- or under-prediction bias.

Data

An AGB spatial dataset for the Campeche, Yucatan, and Quintana Roo administrative regions in the Yucatan peninsula, Mexico, developed by Rodríguez-Veiga et al. (2016), is used (Figure 1.a). This data has 250 m spatial resolution made by MaxEnt with MODIS, Advanced Land Observing Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) dual-polarization backscatter coefficient images, and the shuttle radar topography mission (SRTM) digital elevation model (DEM), trained with the INFys *in-situ* dataset. Dry forest, moist forest, and mangrove forest are found in the North-Western region, the central region, and the coastal zone of the Yucatan peninsula. The reference data for this case study was provided by the INFyS *in-situ* observation data which record measures of AGB (Mg ha^{-1}) at four nested 0.04 ha subplots within 1 ha field plots (Rodríguez-Veiga et al., 2016). Data from a total of 286 (1 ha) field plots were used as reference measures of AGB for the period 2004-2007 (Figure 1.b). It is noted that the spatial resolution of assessed AGB datasets and reference sample is different, which is a limitation of data availability, similar to the study of Rodríguez-Veiga et al. (2016).

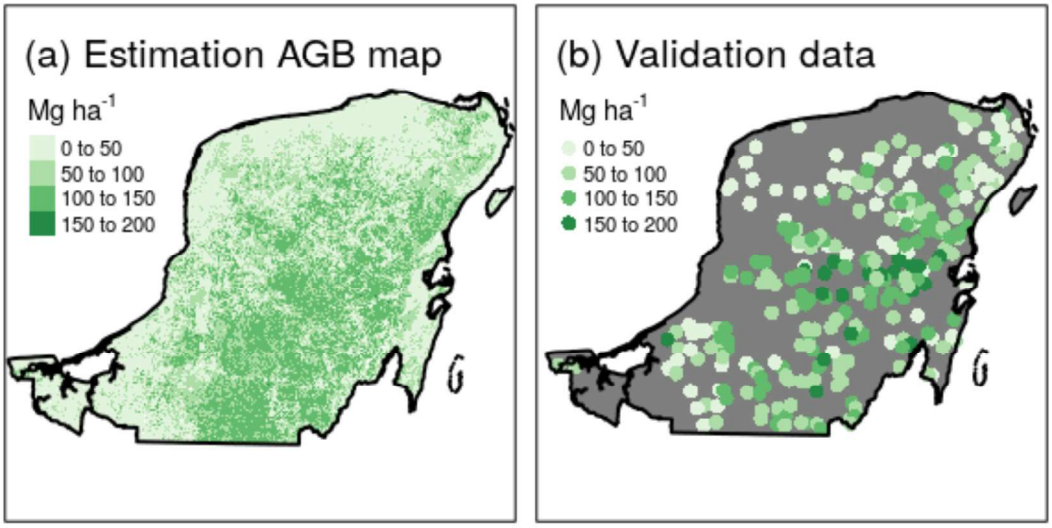

Figure 1. Data used in this study: (a) Forest aboveground biomass map and (b) Validation sample data points.

Results

Conventional global diagnostics for the four AGB datasets assess the errors of the msd (-2.05), the mae (31.52), the rmse (39.42), and the *r* (0.50). However, these only describe overall error characteristics for the map. Figure 2 maps the three GW error diagnostics and GW *r*. In the AGB map, there is a cluster of positive values of GW msd in the dry forests of the North-West, which is coupled with relatively small GW mae and GW rmse values and positive GW *r* values. Forests in this area are often utilized for slash-and-burn agriculture, and the re-growth of trees can influence the remote sensing signals, resulting in potentially large prediction errors, but where it appears, not so large to adversely influence GW mae, and GW rmse, and GW r. Conversely, there are clusters of negative values of GW msd in the moist forests of central-Eastern areas. These areas are coupled with clusters of relatively large GW mae and GW rmse values and a cluster of negative GW *r* values. Thus, this dataset

clearly performs worse in central-Eastern areas, as all four GW diagnostics indicate so. In this central-Eastern area, the forest is matured with large AGBs, so the saturation of spectral data from satellite sensors may be a cause of the inaccurate predictions.

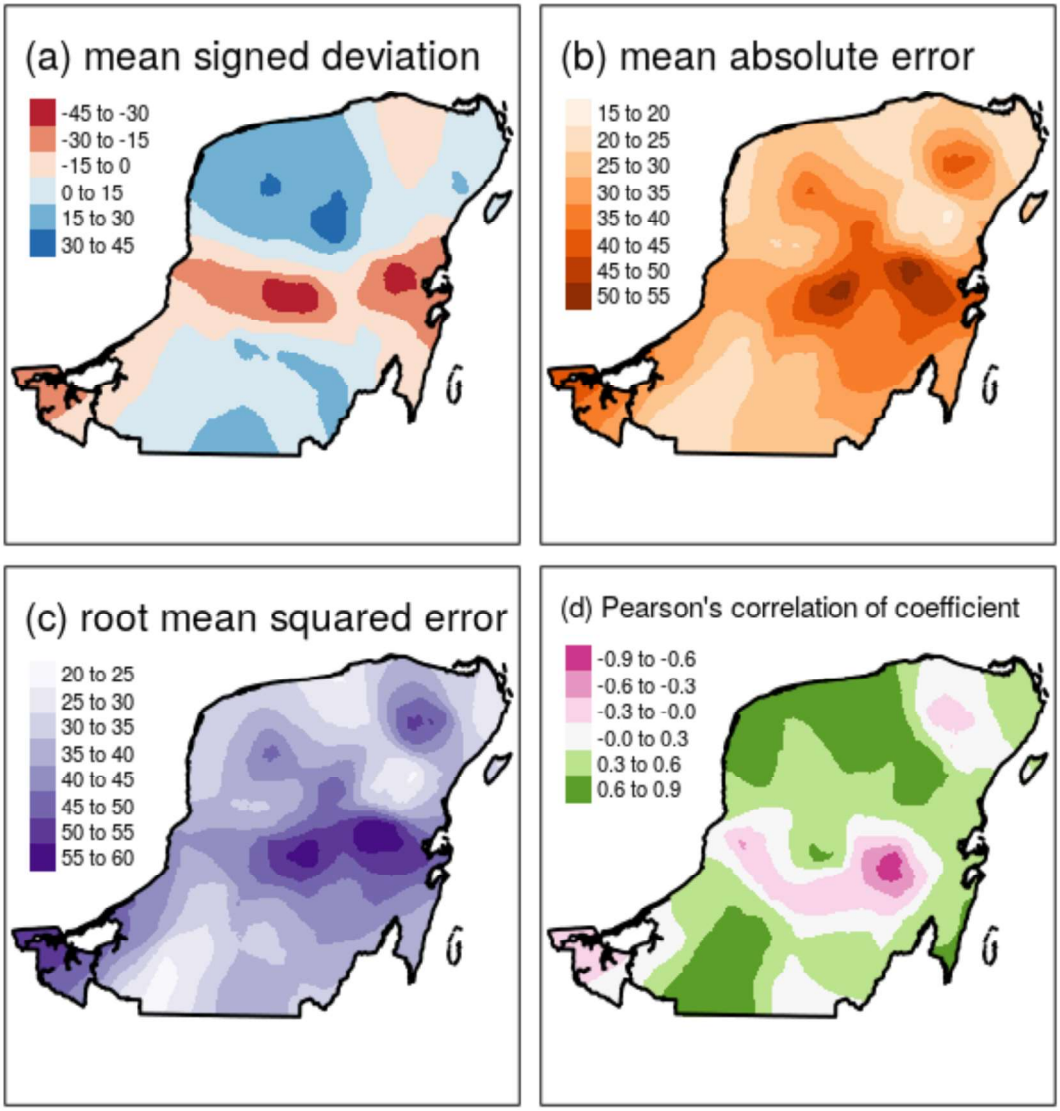

Figure 2. Geographically weighted error measures of (a) mean signed deviation, (b) mean absolute error, (c) root mean squared error, and (d) Pearson's correlation of coefficient for a forest aboveground biomass map.

Conclusion

Conventional diagnostics of error, such as msd, mae, and rmse provide global, 'on-average' measures. These summary measures of error do not capture any spatial information of the error. Ignoring spatial structures in error may result in a false interpretation and misuse of the data that the errors stem from. This work develops and applies localized diagnostics of the error to investigate spatial heterogeneity of each type of these diagnostics.

Notes

This is a summary of a published paper of Tsutsumida et al. (2019) Investigating spatial error structures in continuous raster data, International Journal of Applied Earth Observation and Geoinformation, 74, 259-268.

Acknowledgments

This work was supported by the Joint Support Center for Data Science Research (ROIS-DS-JOINT) under Grant 004RP2019.

Reference

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

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

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, 85–101. https://doi.org/10.18637/jss.v063.i17

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

Rodríguez-Veiga, P., Saatchi, S., Tansey, K., Balzter, H., 2016. Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. Remote Sensing of Environment 183, 265–281. https://doi.org/10.1016/j.rse.2016.06.004

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, 259–268. https://doi.org/10.1016/j.jag.2018.09.020

Willmott, C., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research 30, 79–82. https://doi.org/10.3354/cr030079

Graduate School of Global Environmental Studies,
Kyoto University
Kyoto 606-8501
JAPAN
E-mail address: tsutsumida.narumasa.3c@kyoto-u.ac.jp