Int J Appl Earth Obs Geoinformation 74 (2019) 259–268

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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

Investigating spatial error structures in continuous raster data

Narumasa Tsutsumida[a,*], Pedro Rodríguez-Veiga[b,c], Paul Harris[d], Heiko Balzter[b,c], Alexis Comber[e]

[a] Graduate School of Global Environmental Studies, Kyoto University, Kyoto, 606-8501, Japan
[b] Centre for Landscape and Climate Research, University of Leicester, Leicester, LE1 7RH, UK
[c] NERC National Centre for Earth Observation (NCEO), University of Leicester, Leicester, LE1 7RH, UK
[d] Sustainable Agriculture Sciences, Rothamsted Research, North Wyke, Okehampton, EX20 2SB, UK
[e] School of Geography, University of Leeds, Leeds LS2 9JT, UK

ARTICLE INFO

Keywords:
Error distribution
Spatial accuracy
Local error diagnostics
Spatial heterogeneity

ABSTRACT

The objective of this study is to investigate spatial structures of error in the assessment of continuous raster data. The use of conventional diagnostics of error often overlooks the possible spatial variation in error because such diagnostics report only average error or deviation between predicted and reference values. In this respect, this work uses a moving window (kernel) approach to generate geographically weighted (GW) versions of the mean signed deviation, the mean absolute error and the root mean squared error and to quantify their spatial variations. Such approach computes local error diagnostics from data weighted by its distance to the centre of a moving kernel and allows to map spatial surfaces of each type of error. In addition, a GW correlation analysis between predicted and reference values provides an alternative view of local error. These diagnostics are applied to two earth observation case studies. The results reveal important spatial structures of error and unusual clusters of error can be identified through Monte Carlo permutation tests. The first case study demonstrates the use of GW diagnostics to fractional impervious surface area datasets generated by four different models for the Jakarta metropolitan area, Indonesia. The GW diagnostics reveal where the models perform differently and similarly, and found areas of under-prediction in the urban core, with larger errors in peri-urban areas. The second case study uses the GW diagnostics to four remotely sensed aboveground biomass datasets for the Yucatan Peninsula, Mexico. The mapping of GW diagnostics provides a means to compare the accuracy of these four continuous raster datasets locally. The discussion considers the relative nature of diagnostics of error, determining moving window size and issues around the interpretation of different error diagnostic measures. Investigating spatial structures of error hidden in conventional diagnostics of error provides informative descriptions of error in continuous raster data.

1. Introduction

All spatial data are subject to error. Remotely sensed (RS) imagery routinely contains sensor-related errors, atmospheric effects, and geometric errors. Environmental datasets that describe landscape features and properties from RS products (e.g. forest aboveground biomass, species distribution, and climate change scenarios) inherently contain prediction errors. Errors can manifest themselves as systematic deviations and/or noise which require careful assessment in order to avoid mis-interpretations of the data, to support reliable conclusions and to make informed decisions (Daly, 2006; Foody, 2002). Error assessments provide a guide to data quality and reliability (Foody, 2002) and can provide earth observation (EO) scientists with an understanding of the sources of error both in RS imagery and products (Liu et al., 2007; Stehman and Czaplewski, 1998). However, conventional summary measures of error do not take any spatial information (e.g. spatial heterogeneity) of error into account (Foody, 2005, 2002). Spatially explicit approach for the assessment is hence important.

In EO studies, spatial extensions of conventional diagnostics of error or accuracy have been demonstrated for categorical raster data, such as land cover classification data (Comber et al., 2017, 2012; Comber, 2013; Congalton, 1988; Foody, 2005). These approaches spatially extend the usual method of estimating and reporting accuracy through a confusion matrix, which is the cross-tabulation of predicted and reference classes to generate measures of user's and producer's accuracy that correspond to commission and omission errors, respectively, along

* Corresponding author at: Yoshidahonmachi, Sakyo, Kyoto, 606-8501, Japan.
E-mail address: naru@kais.kyoto-u.ac.jp (N. Tsutsumida).

https://doi.org/10.1016/j.jag.2018.09.020
Received 10 July 2018; Received in revised form 28 September 2018; Accepted 28 September 2018

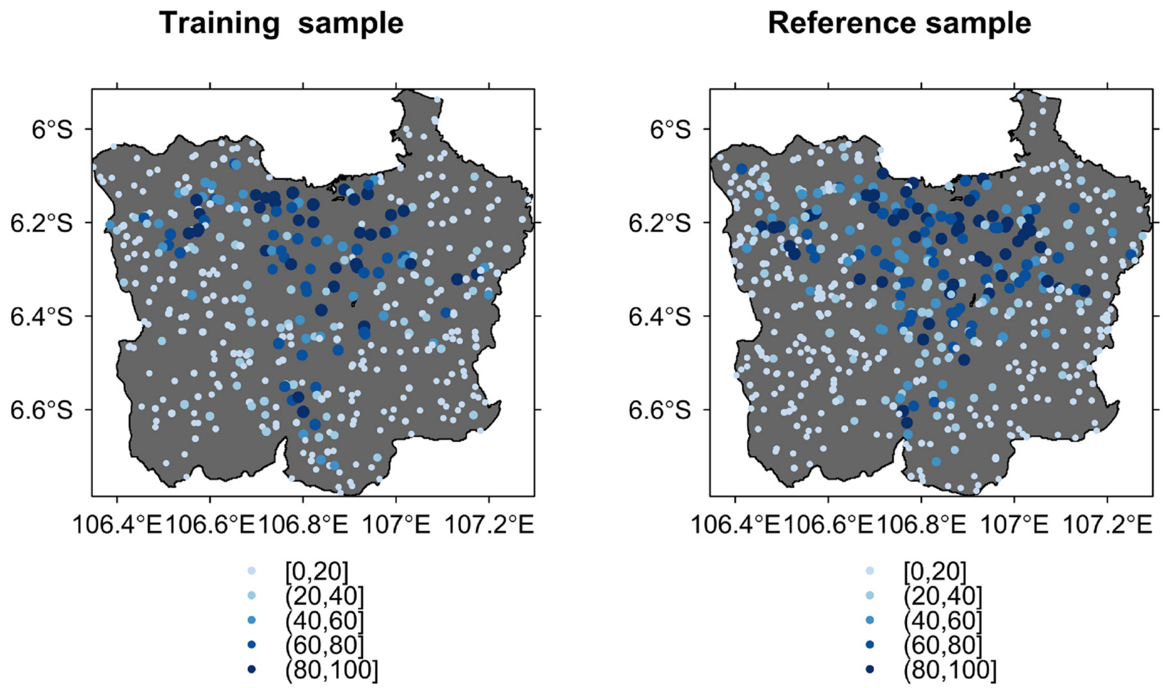

Fig. 1. The spatial distribution of the training (left) and reference (right) sample of fractional impervious surface area (%) in the Jakarta metropolitan area, Indonesia.

with an overall accuracy (Congalton, 1991; Stehman and Czaplewski, 1998). Specifically, Comber (2013) demonstrated the use of a geographically weighted (GW) approach to generate spatial surfaces of these measures. The GW approach calculates a series of local diagnostics of accuracy, using data weighted by their distance to the centre of a moving window or kernel to explore spatial heterogeneity (Gollini et al., 2015). This has been used to compare global land cover datasets (Comber et al., 2013), to assess the consistency of such classification over time (Tsutsumida and Comber, 2015), and to construct hybrid global land cover datasets from multiple inputs (See et al., 2015). Comber et al. (2017) proposed GW confusion matrices for further generic applications. The GW framework itself (Fotheringham et al., 2002; Gollini et al., 2015; Lu et al., 2014) has been widely adopted across many scientific disciplines (e.g. Geography, Ecology, Health), where GW regression (Brunsdon et al., 1996) is the most popular GW model.

The developments of spatially explicit approaches for error assessment in continuous raster data in the EO domain have been limited. Comber et al. (2012) proposed a fuzzy GW difference analysis which estimates absolute deviations between the predicted and reference fuzzy membership, essentially applying a fuzzy generalization of the categorical accuracy measures. Khatami et al. (2017) proposed a spatial interpolation approach for soft classification maps in which a linear kernel function was applied to interpolate spatial deviations between predicted and reference proportions, with a focus on weight of spectral or class proportion as a soft classification measure. Willmott and Matsuura (2006) described maps of cross-validation error. Continuous raster data are commonly assessed using mean signed deviation (msd), mean absolute error (mae), root mean square error (rmse) and Pearson's correlation coefficient (r). Accurate predictions are reflected by msd, mae and rmse to be zero, coupled with r to be one. Although these conventional diagnostics are useful in reporting error, each of them provides an overall, global or 'whole map' measure only. In this respect, Harris and Juggins (2011) demonstrated GW r for assessing UK freshwater acidification prediction accuracy. Harris et al. (2013) demonstrated GW mae for UK freshwater acidification and London house price prediction accuracy, as separate case studies. Monteys et al. (2015) demonstrated GW r for assessing water depth prediction accuracy in Irish coastal waters. These studies either directly extend GW summary statistics (e.g. GW averages, GW variances) as first proposed by Brunsdon et al. (2002), or directly use GW r (Fotheringham et al., 2002), but in a model accuracy context. Further advances of GW summary statistics can be found in Harris and Brunsdon (2010) and Harris et al. (2014). However, the previous studies have only reported spatial error briefly as part of a suite of diagnostics. That is, spatial extensions of conventional diagnostics of error for continuous raster data have not been described in a comprehensive way, specifically in an EO context. Here we demonstrate the linked use of all four diagnostics, msd, mae, rmse and r, through their GW msd, GW mae, GW rmse and GW r counterparts and advance them through the application of Monte Carlo permutation tests to identify unusual clusters of error applied to two EO case studies. The first case study evaluates datasets of the fractional impervious surface area (%ISA) with the aim of investigating spatial structures of error in multiple predictions by four different models. The second case study evaluate four different forest aboveground biomass (AGB) datasets in order to compare spatial structures of error in multiple independent datasets.

2. Case study data

2.1. Study 1

In order to explore how spatial structures of error can differ according to different models, four independent predictions of %ISA in the Jakarta Metropolitan Area (JMA), Indonesia, for 2012 were produced. The %ISA was inferred from the enhanced vegetation index (EVI) stored in moderate resolution imaging spectroradiometer (MODIS) MOD13Q1 product, which are 16-days composite RS imagery with a 231 m spatial resolution. Annual minimum, mean, maximum, and standard deviation of EVI were calculated on a pixel by pixel basis from the 24 images in 2012. These data were classified and assessed using training and reference (validation) samples collected at 984 randomly selected grid squares of the same size and at the same locations as the MODIS MOD13Q1 product. The %ISA was visually interpreted from fine resolution images in available Google Earth from the same year (Comber et al., 2016; Tsutsumida et al., 2016; Tsutsumida