The 36th International Symposium on Remote Sensing of Environment,
11 – 15 May 2015, Berlin, Germany, ISRSE36-706

# THE USE OF GEOGRAPHICALLY WEIGHTED PCA TO CLASSIFY LAND COVER FROM MULTISPECTRAL IMAGE DATA

P. Harris [a], N. Tsutsumida [b], A.J. Comber * [c]

[a] Rothamsted Research, North Wyke, Okehampton, Devon, EX20 2SB – paul.harris@rothamsted.ac.uk
[b] Graduate School of Global Environmental Studies, Kyoto University, Kyoto, 606-8501, Japan–naru@kais.kyoto-u.ac.jp
[c] Department of Geography, University of Leicester, Leicester, LE1 7RH, UK– ajc36@le.ac.uk

**THEMES: NONE – none of the themes listed fit – this is METHOD development**

**KEY WORDS:** GW models; spatial heterogeneity; classification

**ABSTRACT:**

Image classification in remote sensing typically employs heuristics to determine which image bands are selected for analysis and classification. The heuristics are frequently derived from personal experience, (remote sensing) community expertise and specific knowledge of the domain or study area under consideration. Studies requiring information on biomass and vegetation, for example, typically use combinations of red and infra-red image bands, reflecting the need to quantify the red edge vegetation reflectance near the infrared range in the electromagnetic spectrum. However, it is well known that at different locations within the area under consideration, specific combinations of image bands may identify local land cover classes more reliably. Hitherto, the local variation in image requirements have been difficult to systematically quantify and typically a uniform band combination is applied to the entire image scene. In addition, when the number of image bands are large, dimension reduction techniques, such as Principal Components Analysis (PCA) are required. Image bands are transformed to ordered components, reflecting the amount of variance explained. Scores data for the first few components are then inputted into the classification model.

This paper attempts to address both issues concurrently: (i) the need to account for a certain spatial heterogeneity in the land cover classification and (ii) the need to represent the multivariate structure of the image data by only a few components. In this respect, a Geographically Weighted PCA (GWPCA) was applied to MODIS image data with 7 bands (MOD09A1) for a case study in Jakarta, Indonesia. Here the localised loadings from a GWPCA are fed into a classification algorithm where the scale of the spatial heterogeneity (i.e. the kernel bandwidth parameter of the GWPCA) is determined optimally in the calibration of the GWPCA model itself.

The results of applying GWPCA are compared with a classic remote sensing classification using a predetermined combination bands and a standard validation / accuracy assessment. The results show considerable benefits in accommodating the spatial and multivariate structure of the image data within the (remote sensing) classification model. Geographically Weighted modelling is well established in other areas of spatial science, and some critical considerations are discussed, including the interplay between the number of components to retain, bandwidth selection, and image resolution.

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.