# 1. Supplementary Material

## 1.1. Classification Methodologies

Currently two main image analysis techniques exist for urban mapping: spectral indices and classification algorithms. Spectral indices such as the Normalised Difference Built-up Index (NDBI) have been used to delineate urban areas from non-urban (Zha, Gao, and Ni 2003). NDBI identifies built up regions using a ratio of the shortwave-infrared (SWIR) and near infrared (NIR) wavebands and assumes that built up areas have higher SWIR reflectance (Xu 2008). In comparison to ground truth observations, NDBI-derived classification from Landsat Thematic Mapper (TM) over Nanjing, China was found to result in an overall accuracy of 92.6% (Zha, Gao, and Ni 2003). However, due to the heterogeneous nature of urban environments, the identification of built up areas by thresholding a spectral index is not always reliable (Xu 2008).

The Impervious Build-up Index (IBI) attempts to mitigate for this by using a combination of a number of thematic indices namely: NDBI, the Soil Adjusted Vegetation Index (SAVI) and the Modified Normalised Difference Water Index (MNDWI) (Xu 2008). The index amplifies the identification of built up land through the inclusion of ancillary information on the presence of bare surfaces (SAVI) and water bodies (MNDWI) resulting in positive values for pixels identified as being urban (Xu 2008). Nevertheless, urban areas often remain an inseperable mix of impervious and bare earth surfaces which require additional post-processing to delineate (Stathakis, Perakis, and Savin 2012; Sun et al. 2015; Zha, Gao, and Ni 2003).

In the second instance, classification algorithms are defined as parametric (e.g. Maximum Likelihood (ML)) or non-parametric (e.g. decision trees, DT), depending on whether training samples can be represented by a Gaussian probability density function (Donnay and Unwin 2001; Jensen 2005). Maximum Likelihood accounts for the variance-covariance within class distributions and has been implemented for monitoring land cover change and to derive sub-pixel proportions (Shalaby and Tateishi 2007; Atkinson, Cutler, and Lewis 1997). However, due to the parametric assumption of multivariate normal data, the ML classifier can often fail to represent land cover that might be multimodal (Otukei and Blaschke 2010; Melgani and Bruzzone 2004; Mountrakis, Im, and Ogole 2011). An example of this issue is illustrated in semi-arid locations, such as grasslands, which are sensitive to precipitation timing and volume that can result in differing multimodal spectral-temporal profiles (Friedl et al. 2002). A decision tree methodology was utilised to generate the USA National Land Cover Database 2001 (NLCD 2001) resulting in a non-parametric approach able to handle continuous and nominal data, interpretable classification rules and swift application (Homer et al. 2004). Nevertheless, DTs can be negatively affected by pruning methods, for example pessimistic error pruning (PEP) introduces a continuity correction value, within error estimation on no theoretical basis, resulting in under or over pruning (Otukei and Blaschke 2010; Mahesh Pal and Mather 2003; Esposito et al. 1997).

More recently, Machine Learning Algorithms (MLA) or ‘expert systems’ (e.g. Support Vector Machine, SVM) have been implemented for image classification (Jensen 2005; Okujeni et al. 2014) using an automated inductive approach for identification of patterns in data (Cracknell and Reading 2014). SVM is a non-parametric binary statistical learning methodology that separates a dataset into example classes (training data) based on a decision boundary, or hyperplane, with an aim to minimise misclassification. The optimal maximum margin separating hyperplane divides the data into a predefined number of classes, with points on the margins termed ‘support vectors’ (Foody and Mathur 2004; Foody and Mathur 2006). The underlying benefit of SVM pertains to structural risk minimisation, whereby SVMs are able to minimise error on unseen data without prior assumptions on the distribution (Mountrakis, Im, and Ogole 2011; Vapnik and Chervonenkis 1971). SVMs are linear binary classifiers which, when deriving more than two classes, require implementation of an additional process, either a one-against-all or one-against-one analysis. One-against-all solves for the multiple optimisation problem, which separates one class from the remaining classes. Comparatively one-against-one combines multiple classifiers and performs pair-wise comparisons using a ‘voting’ process to assign a pixel to a land cover class, based on the class assigned the most votes (Chih-Wei, Chih-Chung, and Chih-Jen 2008; Mountrakis, Im, and Ogole 2011; M. Pal and Mather 2005). Within SVMs implementation of soft margin and kernel methods aid separability through the introduction of additional variables that ignore hyperplane outliers and transform data into high dimensional feature spaces (Euclidean or Hilbert) utilising non-linear functions to identify linear solutions respectively (Braun, Weidner, and Hinz 2012; Mountrakis, Im, and Ogole 2011; Melgani and Bruzzone 2004; Cortes and Vapnik 1995).

SVMs have been extensively used for classification purposes, due to their ability to ignore inherent image errors and to avoid overfitting (Mountrakis, Im, and Ogole 2011; Foody and Mathur 2006). SVMs have obtained broad applicability for land cover classification using data from a multitude of sensors such as HyMAP (Camps-Valls et al. 2004), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (Zhu and Blumberg 2002) and Landsat (Knorn et al. 2009) producing classification results with accuracies between 85% and 95%.

# References

Atkinson, P., M. Cutler, and H.G. Lewis. 1997. “Mapping Sub-Pixel Proportional Land Cover with AVHRR Imagery.” *International Journal of Remote Sensing* 18 (4): 917–935. doi:10.1080/014311697218836.

Braun, Andreas Ch, Uwe Weidner, and Stefan Hinz. 2012. “Classification in High-Dimensional Feature Spaces-Assessment Using SVM, IVM and RVM with Focus on Simulated EnMAP Data.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5 (2): 436–443. doi:10.1109/JSTARS.2012.2190266.

Camps-Valls, Gustavo, Luis Gómez-Chova, Javier Calpe-Maravilla, José David Martín-Guerrero, Emilio Soria-Olivas, Luis Alonso-Chordá, and José Moreno. 2004. “Robust Support Vector Method for Hyperspectral Data Classification and Knowledge Discovery.” *IEEE Transactions on Geoscience and Remote Sensing* 42 (7): 1530–1542. doi:10.1109/TGRS.2004.827262.

Chih-Wei, Hsu, Chang Chih-Chung, and Lin Chih-Jen. 2008. “A Practical Guide to Support Vector Classification.” *BJU International* 101 (1): 1396–1400. doi:10.1177/02632760022050997.

Cortes, C, and V Vapnik. 1995. “Support-Vector Networks.” *Machine Learning* 20 (3): 273–297. doi:10.1007/BF00994018.

Cracknell, Matthew J., and Anya M. Reading. 2014. “Geological Mapping Using Remote Sensing Data: A Comparison of Five Machine Learning Algorithms, Their Response to Variations in the Spatial Distribution of Training Data and the Use of Explicit Spatial Information.” *Computers & Geosciences* 63: 22–33. doi:10.1016/j.cageo.2013.10.008.

Donnay, Jean-Paul, and David Unwin. 2001. *Modelling Geographical Distributions in Urban Areas*. *Remote Sensing and Urban Analysis*. London, UK: Taylor & Francis.

Esposito, F., D. Malerba, G. Semeraro, and J. Kay. 1997. “A Comparative Analysis of Methods for Pruning Decision Trees.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (5): 476–491. doi:10.1109/34.589207.

Foody, Giles M., and Ajay Mathur. 2004. “A Relative Evaluation of Multiclass Image Classification by Support Vector Machines.” *IEEE Transactions on Geoscience and Remote Sensing* 42 (6): 1335–1343. doi:10.1109/TGRS.2004.827257.

Foody, Giles M., and Ajay Mathur. 2006. “The Use of Small Training Sets Containing Mixed Pixels for Accurate Hard Image Classification: Training on Mixed Spectral Responses for Classification by a SVM.” *Remote Sensing of Environment* 103 (2): 179–189. doi:10.1016/j.rse.2006.04.001.

Friedl, M.A, D.K McIver, J.C.F Hodges, X.Y Zhang, D Muchoney, A.H Strahler, C.E Woodcock, et al. 2002. “Global Land Cover Mapping from MODIS: Algorithms and Early Results.” *Remote Sensing of Environment* 83 (1–2): 287–302. doi:10.1016/S0034-4257(02)00078-0.

Homer, Collin, Chengquan Huang, Limin Yang, Bruce Wylie, and Michael Coan. 2004. “Development of a 2001 National Land-Cover Database for the United States.” *Photogrammetric Engineering & Remote Sensing* 70 (7): 829–840. doi:10.14358/PERS.70.7.829.

Jensen, JR. 2005. *Introductory Digital Image Processing 3rd Edition*. New Jersey, USA: Prentice Hall.

Knorn, Jan, Andreas Rabe, Volker C. Radeloff, Tobias Kuemmerle, Jacek Kozak, and Patrick Hostert. 2009. “Land Cover Mapping of Large Areas Using Chain Classification of Neighboring Landsat Satellite Images.” *Remote Sensing of Environment* 113 (5): 957–964. doi:10.1016/j.rse.2009.01.010.

Melgani, F., and L. Bruzzone. 2004. “Classification of Hyperspectral Remote Sensing Images with Support Vector Machines.” *IEEE Transactions on Geoscience and Remote Sensing* 42 (8): 1778–1790. doi:10.1109/TGRS.2004.831865.

Mountrakis, Giorgos, Jungho Im, and Caesar Ogole. 2011. “Support Vector Machines in Remote Sensing: A Review.” *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (3): 247–259. doi:10.1016/j.isprsjprs.2010.11.001.

Okujeni, Akpona, Sebastian van der Linden, Benjamin Jakimow, Andreas Rabe, Jochem Verrelst, and Patrick Hostert. 2014. “A Comparison of Advanced Regression Algorithms for Quantifying Urban Land Cover.” *Remote Sensing* 6: 6324–6346. doi:10.3390/rs6076324.

Otukei, J. R., and T. Blaschke. 2010. “Land Cover Change Assessment Using Decision Trees, Support Vector Machines and Maximum Likelihood Classification Algorithms.” *International Journal of Applied Earth Observation and Geoinformation* 12 (SUPPL. 1): 27–31. doi:10.1016/j.jag.2009.11.002.

Pal, M., and P. M. Mather. 2005. “Support Vector Machines for Classification in Remote Sensing.” *International Journal of Remote Sensing* 26 (5): 1007–1011. doi:10.1080/01431160512331314083.

Pal, Mahesh, and Paul M. Mather. 2003. “An Assessment of the Effectiveness of Decision Tree Methods for Land Cover Classification.” *Remote Sensing of Environment* 86 (4): 554–565. doi:10.1016/S0034-4257(03)00132-9.

Shalaby, Adel, and Ryutaro Tateishi. 2007. “Remote Sensing and GIS for Mapping and Monitoring Land Cover and Land-Use Changes in the Northwestern Coastal Zone of Egypt.” *Applied Geography* 27 (1): 28–41. doi:10.1016/j.apgeog.2006.09.004.

Stathakis, D, K Perakis, and I Savin. 2012. “Efficient Segmentation of Urban Areas by the VIBI.” *International Journal of Remote Sensing* 33 (20): 6361–6377. doi:10.1080/01431161.2012.687842.

Sun, Genyun, Xiaolin Chen, Xiuping Jia, Yanjuan Yao, and Zhenjie Wang. 2015. “Combinational Build-Up Index (CBI) for Effective Impervious Surface Mapping in Urban Areas.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1–12. doi:10.1109/JSTARS.2015.2478914.

Vapnik, V, and A Chervonenkis. 1971. “On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities” XVI (2): 264–280. doi:10.1007/978-3-319-21852-6\_3.

Xu, H. 2008. “A New Index for Delineating Built‐up Land Features in Satellite Imagery.” *International Journal of Remote Sensing* 29 (14): 4269–4276. doi:10.1080/01431160802039957.

Zha, Y., J. Gao, and S. Ni. 2003. “Use of Normalized Difference Built-up Index in Automatically Mapping Urban Areas from TM Imagery.” *International Journal of Remote Sensing* 24 (3): 583–594. doi:10.1080/01431160304987.

Zhu, Guobin, and Dan G. Blumberg. 2002. “Classification Using ASTER Data and SVM Algorithms;” *Remote Sensing of Environment* 80 (2): 233–240. doi:10.1016/S0034-4257(01)00305-4.