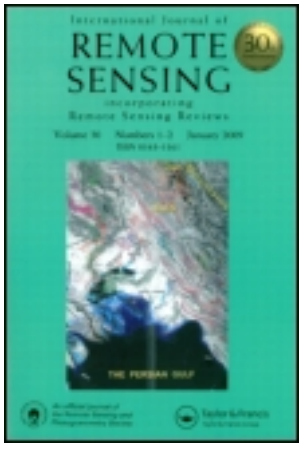

International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

A multiprocess model of adaptable complexity for impervious surface detection

Li Luo [a] & Giorgos Mountrakis [a]

[a] Department of Environmental Resources Engineering, SUNY College of Environmental Science and Forestry, Syracuse, NY, 13210, USA

Available online: 02 Nov 2011

To cite this article: Li Luo & Giorgos Mountrakis (2012): A multiprocess model of adaptable complexity for impervious surface detection, International Journal of Remote Sensing, 33:2, 365-381

To link to this article: http://dx.doi.org/10.1080/01431161.2010.532177

International Journal of Remote Sensing
Vol. 33, No. 2, 20 January 2012, 365–381

Taylor & Francis
Taylor & Francis Group

A multiprocess model of adaptable complexity for impervious surface detection

LI LUO and GIORGOS MOUNTRAKIS*
Department of Environmental Resources Engineering, SUNY College of Environmental Science and Forestry, Syracuse, NY 13210, USA

(*Received 1 June 2010; in final form 15 September 2010*)

Accurately estimating impervious surfaces is of considerable importance towards understanding human impacts on the environment. In this article, a sequential multiprocess model is presented to estimate the distribution of impervious surfaces using Landsat Enhanced Thematic Mapper Plus (ETM+) satellite imagery. Unlike typical single-thread classification processes, the multiprocess model consists of a series of classification steps. By separating the classification task into several subtasks of variable difficulty, multiple classifiers are integrated while adapting to different levels of complexity. The results show statistically significant improvements for the proposed model. An interesting finding is the uneven distribution of simple classifiers in either highly rural or suburban/urban areas, suggesting that the classification difficulty is concentrated in rural areas. The proposed methodology also supports spatially explicit accuracy metrics, facilitating incorporation of the results obtained in interdisciplinary modelling efforts and providing a guide for further algorithmic refinement.

1. Introduction

Imperviousness is a key indicator effectively addressing a host of complex environmental issues and can be explicitly quantified, managed and controlled in land development (Schueler 1994, Arnold and Gibbons 1996). Impervious surface distribution also indicates the extent of urban growth and sprawl. Therefore, accurately estimating impervious surface distribution contributes to a wide range of ecosystem studies, such as hydrology, land use planning and resource management.

Remotely sensed data (from both satellite and aerial platforms) and methods have become increasingly important in impervious surface detection. Remote sensing methods reduce the cost, enhance the speed of ground survey and eliminate the problems of surface access. Several impervious surface estimation studies have been conducted based on a variety of sensors, such as Landsat Thematic Mapper/Enhanced Thematic Mapper Plus (TM/ETM+; Haack *et al.* 2002, Phinn *et al.* 2002, Zha *et al.* 2003), Satellite Pour l'Observation de la Terre (SPOT; Shaban and Dikshit 2001), IKONOS (Cablk and Minor 2003, Lu and Weng 2009) and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER; Zhu and Blumberg 2002, Weng *et al.* 2009).

*Corresponding author. Email: gmountrakis@esf.edu

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2012 Taylor & Francis
http://www.tandf.co.uk/journals
http://dx.doi.org/10.1080/01431161.2010.532177

Various methods have been explored to accurately estimate the distribution of imperviousness and encouraging results have been achieved. These methods include: (1) regression models (Ridd 1995, Bauer *et al.* 2005, Yang 2006), which were used to find the relationship between land cover types in remotely sensed data and spectral characteristics; (2) spectral mixture analysis models (Wu and Murray 2003, Braun and Herold 2004, Powell *et al*. 2007, De Voorde *et al.* 2009), which were used to calculate impervious surface fractions within a pixel and model a mixed spectrum as a combination of spectra for pure land cover types, called endmembers (Roberts *et al.* 1998); (3) machine learning algorithms, which adopted advanced machine learning methods, such as decision tree models (Smith 2000, Herold 2003, Yang *et al.* 2003, Dougherty *et al.* 2004, Crane *et al*. 2005), neural networks (Civco and Hurd 1997, Flanagan and Civco 2001, Iyer and Mohan 2002, Lee and Lathrop 2006) and support vector machines (Mountrakis *et al.* 2011); and (4) contextual classification models, which took into account labelling of neighbours when seeking to determine the most appropriate class for a pixel (Richards and Jia 2006, Luo and Mountrakis 2010, in press, Mountrakis and Luo 2011).

Impervious surfaces, defined as the sum of roads, parking lots, sidewalks, rooftops and other impermeable surfaces of the urban landscape, exhibit heterogeneous spatial and spectral patterns for different types of man-made materials, such as asphalt and concrete (Schueler 1994). Most of the existing impervious surface detection approaches explored only a single classifier, that is only one classification algorithm is applied on the entire dataset. However, different classifiers have their own strengths and limitations and therefore selection of a given method implicitly suggests a tradeoff. For example, the maximum likelihood classifiers produce accurate classification results when the image data are normally distributed. Neural networks may be especially useful when image data are complicated and without a known statistical distribution. Studies have demonstrated that, in general, by combining multiple algorithms together, the classification results are more accurate than a single classification approach (Hansen and Salamon 1990, Krogh and Vedelsby 1995, Breiman 1996). The idea of using hybrid approaches has been applied to impervious surface detection. Liu *et al*. (2004) developed a hybrid classification approach combining decision tree and ARTMAP neural networks and also provided a confidence map via majority voting and other rules. Steel (2000) discussed the land cover mapping methods of combining multiple classification rules as a single rule and using spatial information for classification. Their results suggested that the product rule may increase map accuracy with little additional expense or effort and the spatial classifier is useful for increasing accuracy when patterns exist in the spatial distribution of land cover. Shackelford and Davis (2003) presented a combined fuzzy pixel-based and object-based approach for multispectral IKONOS imagery classification. The object-based classifier adopted the fuzzy pixel-based classification as input and then used shape, spectral and neighbourhood features to determine the final classification of the segmented image. Franke *et al*. (2009) used a hierarchical approach and applied Multiple Endmember Spectral Mixture Analysis (MESMA) to map four levels of complexity of land cover classes. Mountrakis *et al*. (2009) developed an expert-based hybrid multiprocess system to estimate the distribution of impervious surfaces from satellite imagery. Luo and Mountrakis (2010) presented the concept of 'intermediate inputs', where partially classified images act as intermediate layers for additional input production to enhance the classification process.

In the current study, we built up an expert-based hybrid multiprocess classification system to estimate the distribution of impervious surfaces at a binary level from satellite imagery. This multiprocess classification system deals with the impervious surface estimation problem by integrating multiple classifiers within a series of steps that are arranged in a serial sequence. In each step, a single classifier is applied on a certain subset of the entire dataset. The classifiers during subsequent steps are performed on the remaining dataset from early steps. The complexity and resource costs of the classifiers increase as the steps go on. Mathematically simple and computationally efficient classifiers are used during the early steps. As the number of unidentified pixels decreases, the classification difficulty increases. Therefore, complex and computationally intensive classifiers are progressively introduced during the subsequent steps.

The specific application of the hybrid multiprocess classification system was on Landsat ETM+ data acquired in 2001; the aim was to detect the binary impervious surface distribution within the study area of central New York. The adopted hybrid multiprocess classification model consisted of eight classification steps and used two types of classifiers: simple classifiers using only two inputs and complex classifiers with numerous inputs using backpropagation neural network models. The working hypothesis was that the proposed hybrid model would outperform single process classifiers in terms of classification accuracy. To evaluate this hypothesis, the classification results obtained were compared with the results of three single classifiers, namely a single neural network model, a decision tree model and a maximum likelihood model.

2. Methodology

2.1 *Study site and reference data description*

The study area is composed by a Landsat 7 ETM+ scene taken on 28 April 2001 (path 15, row 30) (figure 1). It is located in the central New York region and the area is approximately 173 km × 177 km. The Landsat ETM+ scene contains six bands

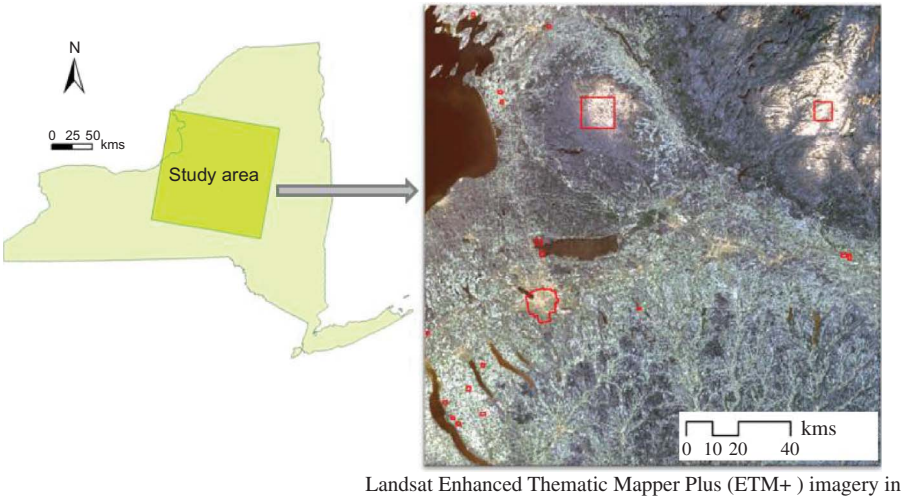

Figure 1. Study area located in the central New York region.

(blue, green, red, near-infrared (IR) and two mid-IR bands) with a spatial resolution of 30 m. Typical land cover classes within this region are tree, grass, bare soil, wetland, water and impervious surfaces; there is also snow on mountain areas located in the north of the scene. Our task was to create a binary impervious map at the 30 m resolution, where any presence of impervious areas within a pixel would translate into an impervious pixel. The underlying idea was to use this method to develop impervious masking products, followed in the future by other subpixel algorithms targeting specifically pixels with an impervious presence.

Nineteen sites representing various land cover types were selected throughout the entire scene to derive the reference data. The first 18 sites used Digital Orthophoto Quarter Quads (DOQQs) acquired in 2001 with a 0.6 m spatial resolution. Binary reference data within these 18 sites were derived based on the on-screen interpretation and digitization of the DOQQ imagery. For site 19, the immediate land cover map with five land cover classes produced by Myeong *et al*. (2001) based on 0.61 m Emerge imagery was overlapped and compared with the Landsat ETM+ scene to derive reference data with a 30 m spatial resolution and two classes of interest: impervious surface areas (ISAs) and non-impervious surface areas (NonISAs). The Landsat pixels overlapping any imperviousness pixels on the high-resolution land cover map were assigned as the impervious class, and all others belonged to the non-impervious class. Although the classification accuracy of Myeong's land cover map was 84.8%, there were sufficient Emerge pixels overlapping each Landsat pixel to ensure that the reference data within site 19 were accurate enough. There is also a 2-year gap between the Emerge imagery and Landsat ETM+ scene. However, site 19 covers the majority of the city of Syracuse, whose urban areas remained mostly steady during the 2 years. The final reference dataset included 101 919 impervious pixels and 280 334 non-impervious pixels from all 19 sites; these were randomly divided into a calibration dataset and a validation dataset with a ratio of 3:7. All algorithmic training took place on the calibration data and the accuracy metrics were reported on the validation dataset.

The six bands of the Landsat ETM+ scene and various band statistics and combinations, normalizations and spatial filtering masks were produced as candidate inputs. These candidate inputs take advantage of spectral and/or spatial characteristics of the remote sensing image, which include established tasseled cap transformations, normalized difference vegetation index (NDVI), normalized differences between other bands, multiple band-to-band advanced correlations and neighbourhood metrics using a standard deviation mask (e.g. 5×5, 7×7), and others. For details of these candidate inputs, refer to Mountrakis *et al.* (2009). The inputs used in each classification step were selected from a random search from all candidate inputs.

2.2 *The hybrid multiprocess model*

The potential benefits of building an integration framework for collaborative operation of multiple algorithms in remote sensing applications have been described previously (Mountrakis 2008, Mountrakis *et al.* 2009). The underlying idea was to develop a method that adjusts algorithmic complexity and transparency to classification difficulty. In other words, use simple classifiers for relatively easy classification tasks and leave complex methods only where they are necessary. The focus of this article was to present a hybrid multiprocess model that is easy to train and understand. The major differences from prior work are the following. First, the structure

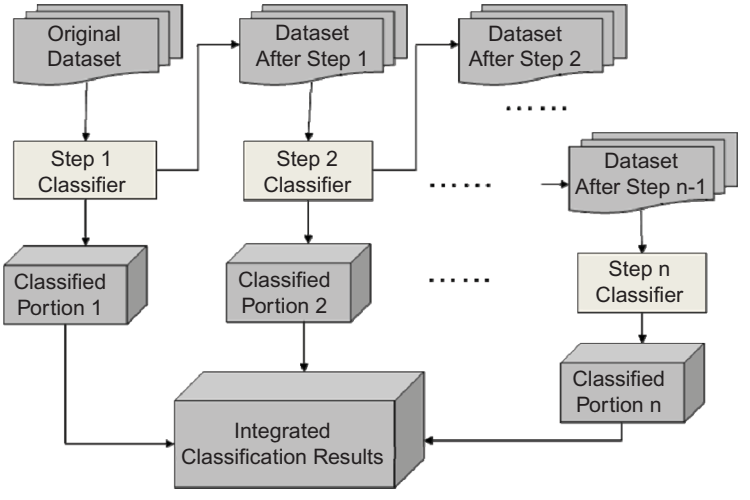

Figure 2. The multiprocess hybrid model framework.

of the proposed hybrid classification model is more explicit as it is composed of a series of sequential steps (figure 2) as opposed to a more complicated tree-type structure. In each step, parts of the dataset are classified using an appropriate classifier towards high classification accuracy while the remaining pixels are forwarded to subsequent classifiers. Second, the proposed hybrid classification model targets a single class in every step (impervious or non-impervious) allowing specific class patterns to prevail and be expressed in a single classifier. The only exception is the last classification step where the remaining pixels are unavoidably classified into both classes. Third, the training process of this proposed method supports a higher automation level, since user involvement is limited in only the identification of the simple classification steps in the early training stages.

In the sequential classification process of the hybrid multiprocess model, easy classification tasks are targeted at the beginning using fairly simple classifiers (e.g. water extraction). As classification progresses, it becomes more difficult to label leftover pixels. Complex classifiers such as neural network classifiers are adopted in the later steps, sacrificing some of the model transparency in exchange for higher classification accuracy. After all the pixels have been identified, the classification results from each step are integrated together to derive the final classification results of the entire dataset. Furthermore, a spatially explicit accuracy map can be generated based on accuracies obtained from each classification step, communicating to potential non-experts algorithmic performance and limitations, an often overlooked process in classification processes.

3. Results and discussion

3.1 *Structure of the hybrid multiprocess model*

The deducted hybrid multiprocess classification model comprised eight steps. During the first four steps, preference was given to simple classifiers operating in low two-dimensional (2D) space. A multivariant analysis was used to select the two inputs with best impervious/non-impervious separation from all the candidate inputs. Then

scatter plots of all the training pixels were drawn in these 2D input spaces. Simple geometrical shapes (e.g. line, rectangle, ellipse) were used to extract pixels belonging to one of the two classes. From each step, all extracted pixels were assigned to either the impervious or the non-impervious class. For example, in step 1, the two inputs adopted were the normalized difference between wetness and band 4 and the normalized difference between bands 4 and 3 (NDVI). After plotting all of the pixels, a rectangle was used to identify non-impervious pixels (figure 3, step 1). The portion of the input space lying within the rectangle range was identified as non-impervious pixels (most of them are water pixels). Points falling outside the rectangle region remained 'unknown' and were forwarded to subsequent steps. The process was repeated creating four sequential steps. The classifications at the input space along with the spatial footprint of each classifier are depicted in figure 3. Table 1 lists the two selected inputs for each simple classifier and the extracted class type during the first four steps.

After the first four steps, where pixels were extracted at a high accuracy, it became progressively more difficult to classify leftover pixels using simple classifiers. To address the high complexity, backpropagation neural networks were the algorithmic choice for the subsequent classification processes. In each step, 2000 candidate neural networks were tested and the neural network with the best overall accuracy in the training dataset was the chosen classifier for that step. In terms of network architecture of the candidate networks, the following assignment process took place: the input layer consisted of 10 randomly selected nodes from the available inputs, the node number in the hidden layers was randomly selected between 8 and 16 for the first layer and between 0 to 6 for the second layer, and the network output layer contained two nodes, each containing a logistic function bounded between 0 and 1. One output layer node represented the impervious class and the other the non-impervious class. We should note that all pixels extracted during the first seven steps were assigned to either the impervious or the non-impervious class but not to both. However, step 8 is the last step of the multiprocess method and therefore all leftover pixels had to be classified; therefore step 8 identified both impervious and non-impervious pixels.

3.2 *Algorithmic complexity spatial assessment*

To contrast the applicability of simple vs. complex algorithms, we simulated the entire Landsat scene to detect ISAs and NonISAs. The spatial footprint of each algorithmic step is illustrated in figure 4; figure 5 shows the portions classified by simple (steps 1–4) and complex classifiers (steps 5–8). Summary statistics show that 60.36% of the entire Landsat scene was identified by using simple classifiers and 39.64% was classified using complex classifiers. When looking into the spatial footprints of the simple and complex classifiers, pixels extracted from simple classifiers possessed unique spectral characteristics from other pixels (such as water pixels), as expected. Therefore, only two inputs were adopted in every simple classifier to achieve high classification accuracy. By contrast, complex classifiers were adopted targeting the identification of pixels that were relatively difficult to separate (such as impervious and soil pixels).

Further statistics on simple vs. complex classifiers were calculated illustrating their distribution within class (table 2). Most ISA pixels were classified by using complex classifiers (about 87.98%) and most NonISA pixels were classified by simple classifiers (about 64.53%). Among the pixels classified by simple classifiers, almost all of them were identified as NonISA pixels while only a small portion of pixels were classified in

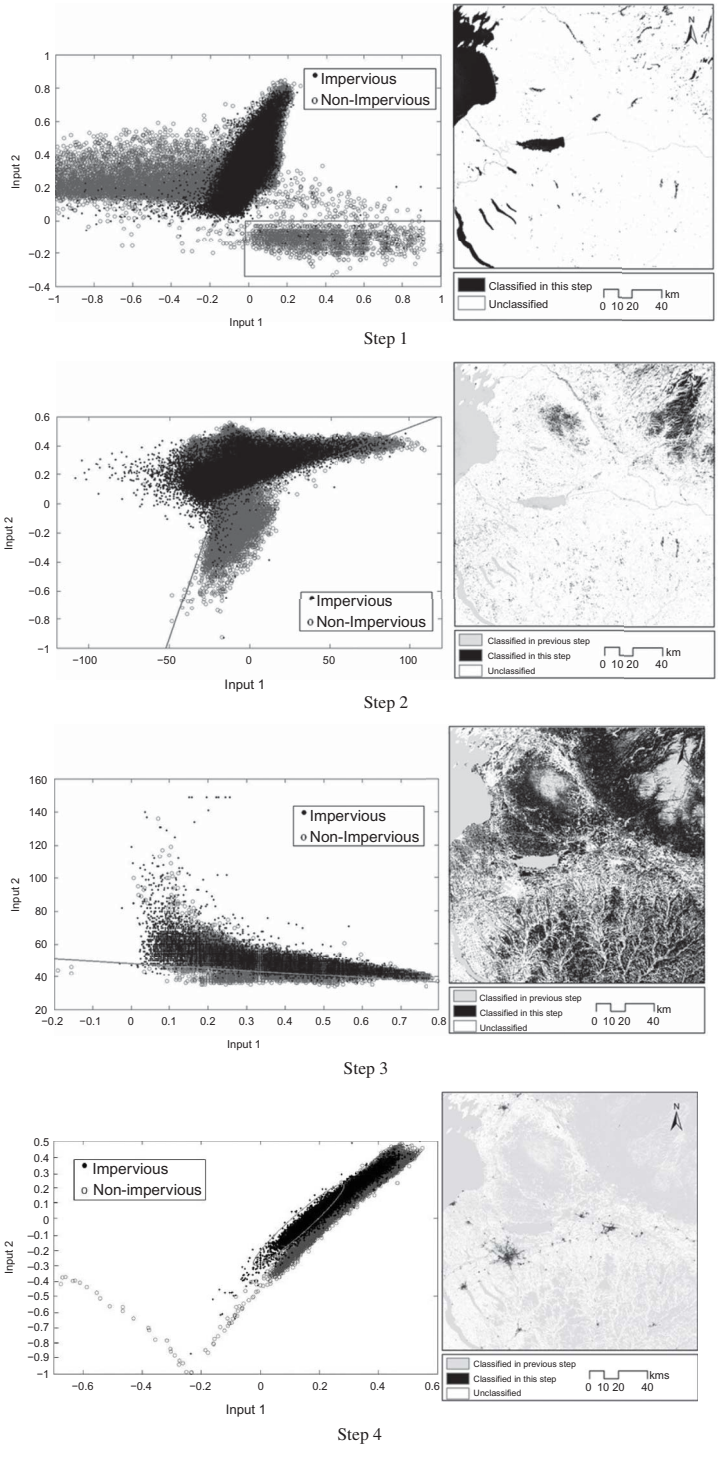

Figure 3. Classification process during the first four steps. The scatter plot (*left*) identifies the selected pixels from each step in the two-dimensional input space and the image (*right*) displays the spatial footprint of the selected pixels.

Table 1. Classifiers, inputs and extracted class types during the first four steps.

Step	Class type	Input
1	NonISA	1: Normalized difference between wetness and band 4 2: Normalized difference between band 4 and band 3 (NDVI)
2	NonISA	1: Greenness 2: Normalized difference between band 5 and band 2
3	NonISA	1: Normalized difference between band 4 and band 3 (NDVI) 2: Band 1
4	ISA	1: Normalized difference between band 5 and band 2 2: A 5×5 filter on normalized difference between brightness and band 5

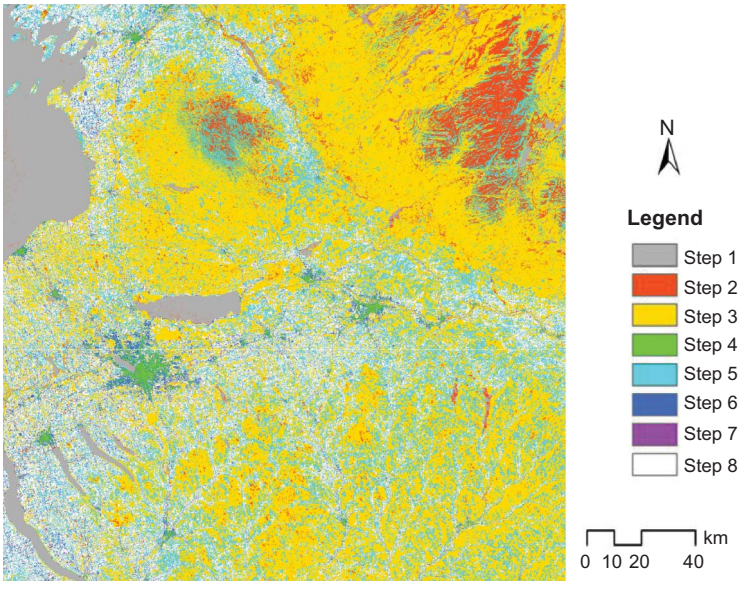

Figure 4. The classification spatial footprint for each algorithmic step.

the ISA class. Therefore, when complex classifiers were involved, a higher probability of impervious pixel presence is expected.

A further investigation was carried at neighbourhood level to assess whether simple classifiers tend to be uniformly distributed when the urbanization level of a given neighbourhood varied. The entire scene was divided equally into 64 × 64 blocks and each block covered an area of approximately 2.75 × 2.80 km^2. ISA percentages and simple classifier percentages were calculated for each block. In figure 6, the grey bars show the histograms of the ISA percentage within each block using a bin size of 10% (see right y axis). Overlaid on the bar graph are aggregated statistics on the simple algorithm percentages; each white box represents the 25% percentile, the median value and the 75% percentile of the simple classifier percentage within each block falling in each ISA percentage range. The mean values within each bin are displayed as crosses.

From the statistics reported in table 2, it was expected that the higher the ISA proportion within a neighbourhood, the more difficult the classification would be and

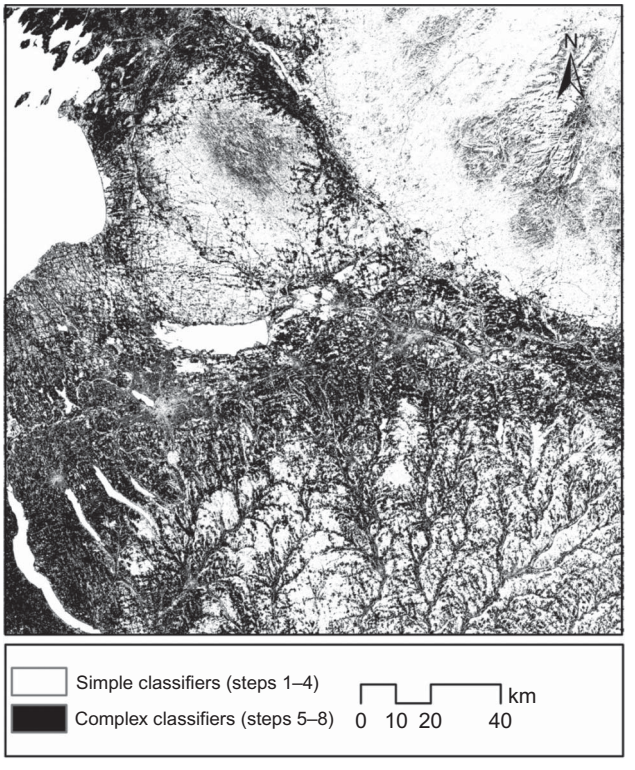

Figure 5. The classification spatial footprint of simple and complex classifiers.

Table 2. Extracted percentage within each land cover class and each classifier.

	Land cover class				Classifier			
	ISA		NonISA		Simple		Complex	
	Simple	Complex	Simple	Complex	ISA	NonISA	ISA	NonISA
Extracted percentage	12.02	87.98	64.53	35.47	1.58	98.42	17.65	82.35

therefore the lower the proportion of simple classifiers. This pattern is evident in figure 6, following the black thick line connecting the mean values of every bin. However, an interesting finding was that the trend of the complexity increase (i.e. smaller simple algorithm proportion) did not continue for higher than 30–40% ISA portions as the portion of simple classifiers started to increase after the 30–40% ISA portions. There seems to be a turning point where, if a neighbourhood becomes highly urbanized, more simple classifiers start contributing. In other words, the most challenging part of the classification is not where the ISA portions reach high values (urban areas), but where low ISA portions exist, in essence in rural and semi-rural areas.

Four representative blocks with an ISA percentage of 10–20, 30–40, 50–60 and 80–90%, respectively, were selected to show the spatial ISA patterns of different simple

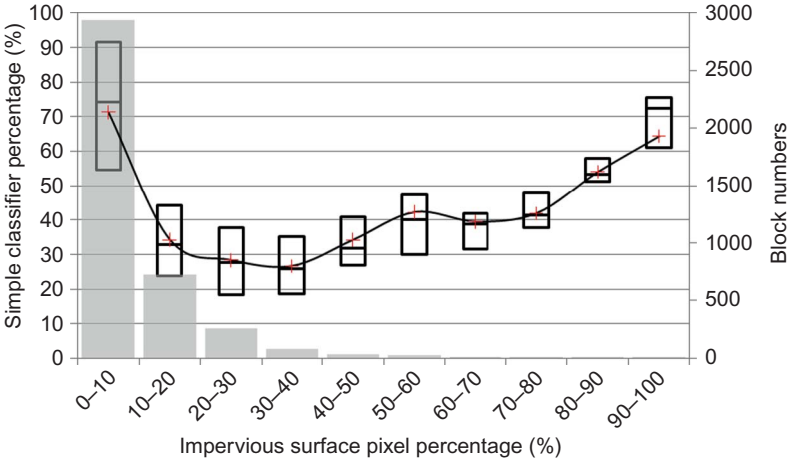

Figure 6. Box plot and histogram distribution of blocks.

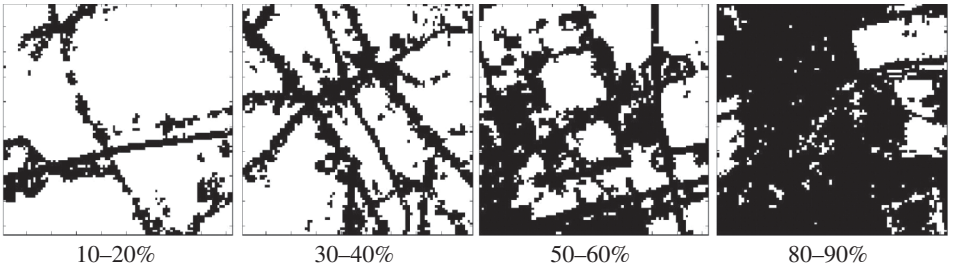

Figure 7. Spatial patterns of representative blocks, percentages show ISA proportions.

classifier percentages (figure 7). Among these blocks, the one that represented the ISA percentage of 30–40% was the most difficult to classify, requiring the lowest simple algorithm proportion. Its spatial pattern illustrated that this block contained complex road networks with a limited amount of buildings. There were only simple road networks appearing in the block with an ISA portion of 10–20% and the other two blocks with ISA portions of 50–60% and 80–90% were located in urban areas and contained a higher number of buildings. The specific land cover patterns and land cover types also affected the classification. In the block with the ISA portion of 30–40%, most pixels belonged to the road and soil classes, which were spectrally similar. Therefore, more complex classifiers were involved to assist in the classification of such blocks. By contrast, blocks with lower or higher ISA portions may have allowed specific texture statistics to resolve spectral ambiguity.

3.3 *Accuracy assessment of the hybrid multiprocess model*

We first assessed the internal algorithmic performance within the hybrid model. Note that this section uses the validation dataset rather than the previous section where statistics were calculated on the entire Landsat scene. Table 3 shows the progressive classification of the validation dataset expressed as the extracted portion and accuracies for the ISA class and the NonISA class. The accuracies of pixels from all of the

Table 3. Classification accuracy of the validation dataset.

	Extracted portion per class (%)			Accuracy (%)				
				Producer's		User's		
Step	ISA	NonISA	Overall	ISA	NonISA	ISA	NonISA	Overall
1 (NonISA)	0.15	1.23	1.01	–	100	–	97.05	97.05
2 (NonISA)	1.98	40.18	32.60	–	100	–	98.79	98.79
3 (NonISA)	2.83	18.40	15.31	–	100	–	96.33	96.33
4 (ISA)	66.73	0.08	13.31	99.49	–	100	–	99.49
5 (NonISA)	1.03	38.47	31.04	–	100	–	99.34	99.34
6 (ISA)	22.66	0.11	4.59	98.05	–	100	–	98.02
7 (ISA)	1.95	0.02	0.40	95.77	–	100	–	95.77
8 (Both)	2.66	1.5	1.73	65.31	93.06	80.53	85.93	84.60
Total	100	100	100	99.81	98.58	94.20	99.95	98.81

steps were calculated directly from the entire validation dataset; in essence we did not use a weighted average of the individual accuracies in each step. Within the validation dataset, the majority of ISAs were identified in step 4 (66.73%) and step 6 (22.66%) and NonISAs were mostly extracted in step 2 (40.18%) and step 5 (38.47%). The overall accuracies in the first seven steps were relatively high, above 95%. As more pixels were distinguished from step 1 to step 7, the classification task became more challenging, which resulted in a relatively low overall accuracy in step 8 (84.60%). This high accuracy variability is desired because it provides a major benefit for our approach, namely the ability to identify problematic areas. If further improvements are pursued, experts could easily target a subset of the problem, in this case pixels belonging to step 8. By doing so, we do not disregard prior classification efforts, rather we build on them to target future areas for improvement.

Figure 8 illustrates the accuracy map of the entire scene based on the calculated accuracies within the validation dataset for each step. The wide range of accuracies obtained reflects the complexity of the problem. In some portions of the scene the classification task was relatively easy (e.g. the Adirondack area in the northeast), while in others some difficulties do exist (e.g. the Finger Lakes area in the southwest). Accuracy maps such as the one in figure 8 are essential to communicate to non-experts successes and challenges of remote sensing products, especially if these products will act as input layers to models from other disciplines (e.g. hydrological modelling).

To investigate the benefits of adopting the integration of multiple classifiers over a single classifier, we also performed the classification by using three single classifiers: the single neural network (NN), decision tree (DT) and maximum likelihood (MLE) classifiers. The exact 31 inputs used in the multiprocess model were introduced as inputs for the single models to constrain any improvements to algorithmic behaviour.

The candidate structure and the training process of the single neural network model was the same with each neural network used in the multiprocess model to allow direct comparisons. It was composed of one input layer, two hidden layers and one output

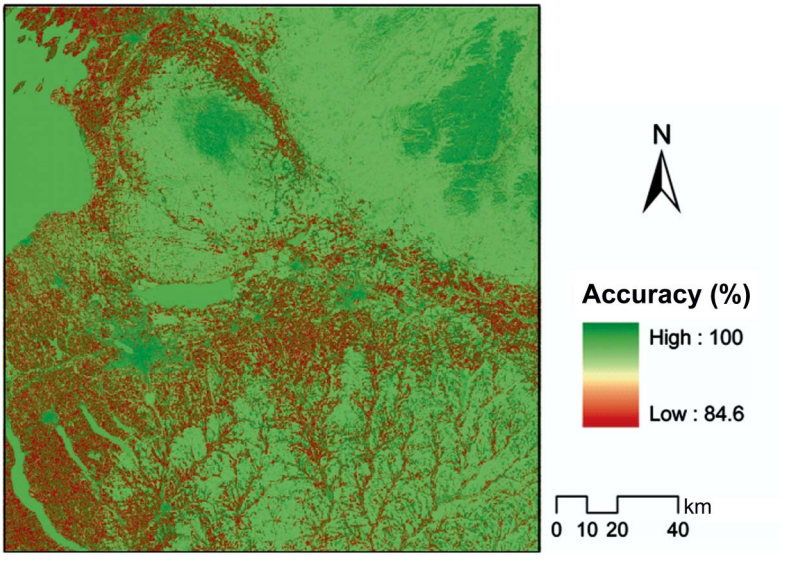

Figure 8. Spatially distributed accuracy extrapolated from the validation dataset.

Table 4. Classification comparison of hybrid multiprocess model with single classifiers.

		Hybrid model	Single neural network	Decision tree	Maximum likelihood
Producer's (%)	ISA	99.81	94.69	94.83	4.70
	NonISA	98.58	99.55	98.92	51.98
User's (%)	ISA	94.20	98.11	95.61	2.37
	NonISA	99.95	98.70	98.72	68.77
Overall (%)		98.81	98.58	98.11	42.59
Kappa (%)		96.19	95.49	94.04	−31.60

layer; 31 inputs were used as nodes in the input layer. The node number in the hidden layers was randomly selected between 8 and 16 for the first layer and between 0 to 6 for the second layer. The output layer contained two nodes representing imperviousness and non-imperviousness. The decision tree model was trained using the calibration dataset and built-in functions within Matlab. The maximum likelihood model was implemented by calculating the distances from each pixel to the mean of the training dataset for each class.

The producer's, user's and overall accuracies as well as kappa statistics of all the single models and the multiprocess model are listed in table 4. The hybrid multiprocess model performed better than the other three single models. Of particular note are the improvements observed in the kappa statistic as classifications by chance are excluded. There the hybrid method outperformed the single neural network by 0.7% and the decision tree by 2.1%, a significant improvement considering the room for improvement. The single neural network model had better accuracies than the decision tree classifier and the maximum likelihood model fitted the validation dataset poorly. The overall accuracy improvements in the hybrid classification model were small in absolute values compared to the single neural network classifier (0.23%) and the decision tree classifier (0.7%). However, in this implementation the two classifiers had already reached high accuracy levels, limiting the available range for improvement. To identify relative improvement terms, the normalized relative accuracy difference between two classification models was used. It was calculated by dividing the improvements by the range potential for improvement of the single classifier (improvement compared to 100% accuracy). The hybrid classification model showed a relative classification accuracy improvement of 16.2% compared to the single neural network classifier and 37.04% compared to the decision tree classifier, and these values were significant. Similar relative improvements were achieved for the kappa statistic: 15.5% when contrasted with the single neural network and 36.1% when compared with the decision tree method.

Furthermore, as stated in §3.2, the most challenging part of the impervious surface detection lies in the low ISA portion areas (i.e. rural and semi-rural regions). Any ISA omission errors are highly pronounced in these areas as they can have a significant impact on the overall ISA estimation. These small changes are of particular interest to further scientific investigations, as it has been shown that small ISA changes at low levels can have a significant impact on ecosystems (e.g. Klein 1979, Forman and Alexander 1998).

To further quantify the benefit of a hybrid multiprocess model over single classifiers, we conducted a Z-test between kappa statistics using different classification models.

Table 5. Z-test between different classifiers, p-values in parentheses.

	Maximum likelihood	Neural network	Decision tree
Neural network	860.76 (<0.0001)	–	–
Decision tree	823.10 (<0.0001)	13.64 (<0.0001)	–
Hybrid model	555.08 (<0.0001)	3.46 (<0.0005)	10.43 (<0.0001)

The null hypothesis is: kappa using a hybrid classification model = kappa using a single classification model. The Z-score between different kappa values is calculated as (Cohen 1960):

$$Z\text{-score} = \frac{(K_1 - K_2)}{\sqrt{(\text{ASE}\,(K_1))^2 + (\text{ASE}\,(K_2))^2}} \tag{1}$$

where K_1 and K_2 represent the kappa values and ASE(K_1) and ASE(K_2) represent the asymptotic standard errors (ASEs) of the two models, respectively. A p-value was calculated using the Z-score expressing the confidence level that these two models can produce identical results. We set the confidence level at 0.05, but note that because of the low p-values our confidence level could have been set lower. The comparison results between pairs of classifiers are shown in table 5, in which the calculated p-values for the three pairs of comparisons are much smaller than the confidence level. Therefore, the hypothesis can be rejected, meaning that the use of a hybrid multiprocess model provided a statistically significant improvement compared to single classifiers.

4. Conclusions

The focus of this article was on developing a hybrid multiprocess classification model for impervious surface detection. The proposed model comprised numerous classifiers, some with simple structures and others with complex algorithmic underpinnings. Experiments suggested that simple classifiers may be used in either highly rural or highly urbanized areas, while the complex classifiers took over in other areas. The hybrid multiprocess classification model was more accurate with regard to overall accuracies and kappa statistics compared to three single-process classifiers, especially when the relative room for improvement was considered. A Z-test also indicated that the improvements were significant. The spatially distributed accuracy map was a natural by-product of the multiple algorithms within the model and supports non-expert usage of the results obtained, but also acts as a guide for further algorithmic refinements by experts. The hybrid multiprocess model presented here was for a binary classification task, but it can be easily extended to multiclass problems. Finding a way to incorporate multiple targeted algorithms remains an issue of investigation.

Acknowledgements
This research was partially supported by the National Science Foundation (award GRS-0648393), the National Aeronautics and Space Administration (awards NNX08AR11G, NNX09AK16G) and the Syracuse Center of Excellence CARTI Program.

References

Arnold, C.L. and Gibbons, C.J., 1996, Impervious surface: the emergence of a key urban environmental indicator. *American Planning Association Journal*, **62**, pp. 243–258.

Bauer, M., Loeffelholz, B. and Wilson, B., 2005, Estimation, mapping and change analysis of impervious surface area by Landsat remote sensing. In *Proceedings of the Pecora 16 Conference*, 23–27 October 2005, Sioux Falls, SD. Available online at: http://land.umn.edu/documents/MinnesotaImpervious_Pecor16.pdf (accessed 5 January 2010).

Braun, M. and Herold, M., 2004, Mapping imperviousness using NDVI and linear spectral unmixing of ASTER data in the Cologne-Bonn region (Germany). In *Proceedings of the SPIE 10th International Symposium on Remote Sensing*, 8–12 September 2003, Barcelona. Available online at: http://www.geogr.uni-jena.de/~c5hema/pub/spie_paper_mhfinal_4.pdf (accessed 20 March 2010).

Breiman, L., 1996, Bagging predictors. *Machine Learning*, **24**, pp. 123–140.

Cablk, M.E. and Minor, T.B., 2003, Detecting and discriminating impervious cover with high-resolution IKONOS data using principal component analysis and morphological operators. *International Journal of Remote Sensing*, **24**, pp. 4627–4645.

Civco, D.L. and Hurd, J.D., 1997, Impervious surface mapping for the state of Connecticut. In *Proceedings of the ASPRS/ACSM 1997 Annual Convention*, 7–10 April 1997, Seattle, WA (Bethesda, MD: American Society for Photogrammetry and Remote Sensing), CD-ROM.

Cohen, J., 1960, A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, pp. 37–46.

Crane, M., Xian, G. and McMahon, C., 2005, Estimation of sub-pixel impervious surface using Landsat and ASTER imagery for assessing urban growth. In *Proceedings of the ISPRS Joint Conference: 5th International Symposium on Remote Sensing and Data Fusion over Urban Areas (URBAN 2005) 5th International Symposium Remote Sensing of Urban Areas (URS 2005)*, 14–16 March 2005, Tempe, AZ. Available online at: http://www.isprs.org/proceedings/XXXVI/8-W27/crane.pdf (accessed 15 February 2010).

De Voorde, V., De Roeck, T. and Canters, F., 2009, A comparison of two spectral mixture modelling approaches for impervious surface mapping in urban areas. *International Journal of Remote Sensing*, **30**, pp. 4785–4806.

Dougherty, M., Dymond, R.L., Goetz, S.J., Jantz, C.A. and Goulet, N., 2004, Evaluation of impervious surface estimates in a rapidly urbanizing watershed. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 1275–1284.

Flanagan, M. and Civco, D.L., 2001, Subpixel impervious surface mapping. In *Proceedings of the American Society for Photogrammetry and Remote Sensing 2001 Annual Conference*, 23–27 April 2001, St Louis, MO (Bethesda, MD: American Society for Photogrammetry and Remote Sensing), CD-ROM.

Forman, R.T.T. and Alexander, L.E., 1998, Roads and their major ecological effects. *Annual Reviews of Ecology and Systematics*, **29**, pp. 207–231.

Franke, J., Roberts, D.A., Halligan, K. and Menz, G., 2009, Hierarchical multiple urban environments. *Remote Sensing of Environment*, **113**, pp. 1712–1723.

Haack, B.N., Solomon, E.K., Bechdol, M.A. and Herold, N.D., 2002, Radar and optical data comparison/integration for urban delineation: a case study. *Photogrammetric Engineering and Remote Sensing*, **68**, pp. 1289–1296.

Hansen, L.K. and Salamon, P., 1990, Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, pp. 993–1001.

Herold, N., 2003, Mapping impervious surfaces and forest canopy using classification and regression tree (CART) analysis. In *Proceedings of ASPRS 2003 Annual Convention*, 5–9 May, Anchorage, AK (Bethesda, MD: American Society for Photogrammetry and Remote Sensing), CD-ROM.

Iyer, S.V. and Mohan, B.K., 2002, Urban landuse monitoring using neural network classification. In *Proceedings of the IEEE Symposium on Geoscience and Remote Sensing*, Vol. 5, pp. 2959–2961.

Klein, R.D., 1979, Urbanization and stream water quality impairment. *Water Resources Bulletin*, **15**, pp. 948–963.

Krogh, A. and Vedelsby, J., 1995, Neural network ensembles, cross validation and active learning. In *Advances in Neural Information Processing Systems*, G. Tesauro, D. Touretzky and T. Leen (Eds.), pp. 231–238 (Cambridge: MIT Press).

Lee, S. and Lathrop, R.G., 2006, Subpixel analysis of Landsat ETM+ using self-organizing map (SOM) neural networks for urban land cover characterization. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 1642–1654.

Liu, W., Gopal, S. and Woodcock, C.E., 2004, Uncertainty and confidence in land cover classification using a hybrid classifier approach. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 963–971.

Lu, D. and Weng, Q., 2009, Extraction of urban impervious surfaces from an IKONOS image. *International Journal of Remote Sensing*, **30**, pp. 1297–1311.

Luo, L. and Mountrakis, G., 2010, Integrating intermediate inputs from partially classified images within a hybrid classification framework: an impervious surface estimation example. *Remote Sensing of Environment*, **114**, pp. 1220–1229.

Luo, L. and Mountrakis, G., in press, Converting local spectral and spatial information from a priori classifiers into contextual knowledge for impervious surface classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66**, pp. 579–587.

Mountrakis, G., 2008, Next generation classifiers: focusing on integration frameworks. *Photogrammetric Engineering and Remote Sensing*, **74**, pp. 1178–1180.

Mountrakis, G., Im, J. and Ogole, C., 2011, Support vector machines in remote sensing: a review. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66**, pp. 247–259.

Mountrakis, G. and Luo, L., 2011, Enhancing and replacing spectral information with intermediate structural inputs: a case study on impervious surface detection. *Remote Sensing of Environment*, **115**, pp. 1162–1170.

Mountrakis, G., Watts, R., Luo, L. and Wang, J., 2009, Developing collaborative classifiers using an expert-based model. *Photogrammetry Engineering and Remote Sensing*, **75**, pp. 831–843.

Myeong, S., Nowak, D.J., Hopkins, P.F., and Brock, R.H., 2001, Urban cover mapping using digital, high-spatial resolution aerial imagery. *Urban Ecosystems*, 5, pp. 243–256.

Phinn, S., Stanford, M., Scarth, P., Murray, A.T. and Shyy, T., 2002, Monitoring the composition and form of urban environments based on the vegetation-impervious surface-soil (VIS) model by sub-pixel analysis techniques. *International Journal of Remote Sensing*, **23**, pp. 4131–4153.

Powell, R.L., Roberts, D.A., Dennison, P.E. and Hess, L.L., 2007, Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment*, **106**, pp. 253–267.

Richards, J.A. and Jia, X., 2006, *Remote Sensing Digital Image Analysis* (Heidelberg, Germany: Springer).

Ridd, M.K., 1995, Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *International Journal of Remote Sensing*, **16**, pp. 2165–2185.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G. and Green, R.O., 1998, Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, **65**, pp. 267–279.

Schueler, T.R., 1994, The importance of imperviousness. *Watershed Protection Techniques*, **1**, pp. 100–111.

Shaban, M.A. and Dikshit, O., 2001, Improvement of classification in urban areas by the use of textural features: the case study of Lucknow city, Uttar Pradesh. *International Journal of Remote Sensing*, **22**, pp. 565–593.

Shackelford, A.K. and Davis, C.H., 2003, A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 1920–1932.

Smith, A.J., 2000, Subpixel estimates of impervious surface cover using Landsat TM imagery. MA scholarly paper, University of Maryland, College Park.

Steele, B.M., 2000, Combining multiple classifiers: an application using spatial and remotely sensed information for land cover type mapping. *Remote Sensing of Environment*, **74**, pp. 545–556.

Weng, Q., Hu, X. and Liu, H., 2009, Estimating impervious surfaces using linear spectral mixture analysis with multitemporal ASTER images. *International Journal of Remote Sensing*, **30**, pp. 4807–4830.

Wu, C. and Murray, A.T., 2003, Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, **84**, pp. 493–505.

Yang, L., Huang, C., Homer, C.G., Wylie, B.K. and Coan, M.J., 2003, An approach for mapping large-area impervious surfaces: synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing*, **29**, pp. 230–240.

Yang, X., 2006, Estimating landscape imperviousness index from satellite imagery. *IEEE Geoscience and Remote Sensing Letters*, **3**, pp. 6–9.

Zha, Y., Gao, J. and Ni, S., 2003, Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, **24**, pp. 583–594.

Zhu, G. and Blumberg, D.G., 2002, Classification using ASTER data and SVM algorithms: the case study of Beer Sheva, Israel. *Remote Sensing of Environment*, **80**, pp. 233–240.