INTEGRATION OF THE PIXEL AND OBJECT DOMAIN FOR THE MAPPING OF NEW URBAN LANDSCAPES IN THE MEDITERRANEAN WITH A FOCUS ON OUTDOOR WATER CONSUMPTION

Nils Wolf[1], and Angela Hof[2]

1. University of Bochum, Department of Geography, Geomatics and Remote Sensing Group, Bochum, Germany; nils.wolf(at)rub.de
2. University of Bochum, Department of Geography, Landscape Ecology/Biogeography, Bochum, Germany; angela.hof(at)rub.de

ABSTRACT

Land use demands of tourism and residential development drive the spread of leisure infrastructure and new urban landscapes in the Mediterranean. In particular along the coasts, golf courses, irrigated landscaping and swimming pools are becoming characteristic features of already densely populated locations that are among the areas with the greatest water deficits. Against this background, the present paper assesses the potential of different high-resolution imagery in combination with innovative image analysis techniques for an automated, targeted mapping of water-related urban features. The mapping task is conducted with WorldView-2, IKONOS and airborne imagery in three different urban study areas in Spain and Greece. Object-based feature extraction and the Random Forests algorithm are applied to the classification problem of separating turf grass, swimming pools, other vegetation and non-vegetated areas. The classifier performance is evaluated against susceptibility to reduced training set size and high-dimensional features spaces, variations in the training set (stability of results), varying feature subspaces, and the inclusion of uncertain – hence potentially mislabelled – pixels in the model calibration stage. The results indicate that best discrimination can be obtained if complementing the standard spectral feature space by object features from multi-scale segmentation. Furthermore, it is confirmed that Random Forests can handle high-dimensional feature spaces with large amounts of potentially redundant or irrelevant features. Comparing the results across image sources and study areas reveals different quality levels, but indicates that water-related urban landscape elements can be mapped, if the remote sensing imagery meets several requirements such as very high spatial resolution.

INTRODUCTION: BACKGROUND AND MOTIVATION OF THE STUDY

Remote sensing enables measuring, mapping and modelling of urban environments in order to understand complex spatial patterns and dynamics and to support sustainable urban development. (1). However, the spatial and spectral heterogeneity of urban areas poses challenges to many remote sensing applications (2). While several publications reported on the requirements for sensors and image processing for mapping urban environments (3,4), this paper further focuses on urban remote sensing towards a targeted mapping approach, designed for water consumption studies in Mediterranean urban coastal areas. The main objective of the paper is the elaboration and testing of image analysis methods. By gearing image analysis towards research questions from urban studies, the paper moves beyond measuring and monitoring of physical properties of urban areas and looks into image analysis for the extraction of the relevant urban features for a specific set of environmental issues outlined in the following paragraphs.

Geocomputation and remote sensing challenges addressed in the paper

For some time now, technological developments of spaceborne systems have led to new application ranges which enable a much more detailed look into the urban environment. Today's high resolution satellite sensors offer alternatives to airborne products which traditionally have been the source for detailed urban applications (5). In this paper, we conduct the mapping task for different study areas and with different spaceborne (WorldView-2 and IKONOS) and airborne (digital RGB,

DOI: 10.12760/01-2013-1-07

40 cm GSD) remote sensing images in order to report on mapping accuracies obtained in these different settings and to discuss the case specific factors, like seasonality or spatial and spectral resolution.

Urban areas with their vertical structures, non-horizontal surfaces and a large range of surface materials are challenging to remote sensing and require innovative image analysis in order to obtain reliable results. In this paper, we test the performance of the Random Forests (RF) machine learning technique (6) and the value of object-based feature extraction (7) to generate spatial descriptors complementing conventional per-pixel spectral classification. We suggest to access object features on pixel level, because objects as sampling and processing units are of non-uniform size and thereby complicate the acquisition of accurate and unbiased data samples for the calibration and adequate testing of classification. Within this pixel level sampling and processing framework, it is investigated whether per-pixel spectral values and object features can be jointly used in the predictor space in order to increase mapping accuracy. For this purpose, the performance of joint feature sets is compared against feature sets containing exclusively either the object or the spectral per-pixel feature descriptors. Furthermore, it is examined how RF respond to varying amounts of training data and high dimensionality, the latter resulting from the manifold possibilities of extracting object-based features (e.g. spectral mean, variance, minima, textures or geometries and context) on multiple segmentation scales.

Urban leisure landscapes are privately owned to a large extent and therefore hardly accessible for comprehensive in-field data collection. At the same time, appropriate official data are usually not available. This places additional demands on the remote sensing data which preferably should enable the analyst to collect reference data by visually interpreting directly from the image. However, information extraction from remote sensing data is affected by various sources of uncertainty, ranging from noisy and technically limited sensor measurement, degrading processing for signal storage and data transfer and ambiguous feature extraction to inherently vague concepts of classes (8). Therefore, this paper also assesses how uncertainty limits the reference data collection in the different image data and study sites at hand and how uncertainty influences map accuracy estimates.

As an application, this paper presents targeted mapping of high-resolution imagery to derive environmental information for research on water consumption patterns in urban areas. The next section outlines the changing geographies of water consumption in urban-tourist areas in the Mediterranean to outline the requirements for image processing that delivers data which are fit for the purpose. Subsequently, the methodological approach of targeted mapping (9) and its application are illustrated in three different study areas with four types of imagery. Finally, the results are discussed and a brief outlook on follow-up research is given.

Current urbanisation trends and environmental information of interest in new Mediterranean urban landscapes

In order to put the potential contribution of image analysis methods elaborated and tested in this paper into a wider geographical research perspective, it is instructive to outline the issue of changing water consumption trends in the Mediterranean. The Mediterranean is one of the regions that are most affected by global climate change. An increase in environmental conflicts and growing competition for already scarce water resources are expected, with negative implications towards current and future sustainability (10). This development is exacerbated by accelerated expansion of urban and tourist areas, in particular in Mediterranean coastal zones. Land cover change detected from comparison of CORINE Land Cover data for 1990 and 2000 showed that already highly populated coastal strips were hot spots of urban sprawl and fragmentation (11). To the extent that low density urban areas are gaining ground in the Mediterranean, a substantial increase in water consumption for outdoor uses can be expected (12,13).

Water consumption and water-related land use patterns

The interconnection of water balance, water policy and socio-economic aspects under the process of suburbanisation and urban sprawl has been studied at different places around the world (12,13,14,15). Research that has explicitly accounted for urban form and the interrelatedness of land use and water consumption patterns in the Mediterranean has revealed that water consumption in the domestic urban sector is positively correlated with water-intensive urban features. Gardens and swimming pools are the main water-demanding outdoor features driving the recent increase in water consumption levels. Water consumed annually by the garden reaches 30% of the total water consumed in the household and can attain 50% during the summer season (16,17). While Mathieu et al. (18) have shown that object-oriented analysis of IKONOS satellite imagery identifies >90% of gardens from other land covers in an urban area in New Zealand, the present approach moves beyond this objective and targets for a separation of grass-dominant versus shrub/tree dominant garden covers. This is because the general observation is that the higher the proportion of turf grass and the lower the proportion of shrubs and trees, the higher the irrigation and water demand (14,16). Next to gardens, the targeted mapping automatically extracts swimming pools that are considered to be the second most problematic outdoor uses in water consumption studies. At more than 1 million swimming pools in 2008, Spain ranks second (after France) in the number of swimming pools in Europe and 30,000 pools were built annually before the current real estate crisis. To add a swimming pool to the home represents an increase of 155 litres per day in water consumption (19). Operational water consumption assessment studies over a range of urban forms ultimately depend on automated mapping of the relevant features in urban environments. To this end, the present approach is geared towards a targeted mapping workflow applicable to different kinds of image data to map the number and area of swimming pools, the extent, distribution, density and design of urban green spaces (here: gardens) as features of interest to address environmental implications of new urban landscapes. This is a typical example of research context in which applying remote sensing to tackle urban problems is deemed useful and important.

The study areas described in the following section were selected as representative case study areas, where the urbanisation processes described above have transformed the urban model. Moreover, the study areas represent the proliferation of gardens, swimming pools and golf courses as residential resort features.

Summarizing, the main objective of this paper is to develop and test methods for the automated mapping of Mediterranean urban landscape elements with relevance to water consumption studies, using different high resolution remote sensing imagery. For this purpose, the experiments are set up in order to:

- assess object features on pixel level for an integrative approach which combines the pixel and object domains and simplifies the statistically reliable implementation of machine learning
- test the RF classifier and its response to varying training set size, dimensionality, feature sets and noisy/uncertain reference data
- discuss mapping accuracies with regard to sensor- and scene-specific properties of different remote sensing imagery.

MATERIALS AND METHODS

Study areas

Low density urban coastal areas and their land use patterns are compared across the Mediterranean. Nova Santa Ponsa is located in the southwest of Mallorca Island and belongs to Calvià municipality, one of the major Mediterranean tourist resorts with an internationally acclaimed Local Agenda 21 and sustainable development as its overarching policy objective. Nova Santa Ponsa is a sub-municipal low density residential tourism area where 88% of the parcels are used by single residential houses, 60% of them are used as second homes (17). San Roque municipality is located in Cádiz province in the autonomous region of Andalusia. Like Nova Santa Ponsa, San Roque is a high-quality tourist destination and residential area, offering access to nearby golf courses, marinas and beaches. The actual study area is the sub-municipal district of Sotogrande, which is promoted as an unusual Spanish experience due to its ‘American style’ settlement structure with a sprawling suburban feel. Characteristic are large properties with single family dwellings and a spacious, resort-like suburban design. The two Spanish case study areas have experienced accelerated urban development since the early 1990s when the development of a strong private sector housing market in Spain coincided with a reorientation of the tourism industry towards diversification and provision of high-quality tourist services. Chania is one of the four prefectures on the island of Crete, Greece. Located in the northwest of Crete, Chania prefecture attracts one fifth of the tourists and ranks second after the prefecture of Iraklion in tourist numbers in Crete Island. Among the three study areas (cp. Figure 1), Chania represents an urban area which is predominantly characterized by mass tourist land use and its characteristic dominance and density along a rather narrow coastal strip. Nova Santa Ponsa is typical of the low density, non-hotel tourist resorts that have developed next to established mass tourist centres in the vicinity of golf courses. Finally, Sotogrande represents a luxury tourist resort. All three study areas are characterized by Mediterranean climate with dry hot summers and wet and mild winters. The hydrogeological and climatic situation is paired with high water demands by agriculture, the domestic and the tourist sectors. Dependency on groundwater and the increase of demand mainly from agriculture and tourism has exacerbated a water supply situation which is associated with groundwater overdraft, groundwater level decline, and water quality deterioration due to saltwater intrusion as a result of excessive pumping of the aquifers (20,21,22). Agricultural water consumption has been progressively overtaken by tourism and urban uses that have become a sizeable sector of demand. Sustainable water resources management is a major, yet unresolved issue in all three study areas.

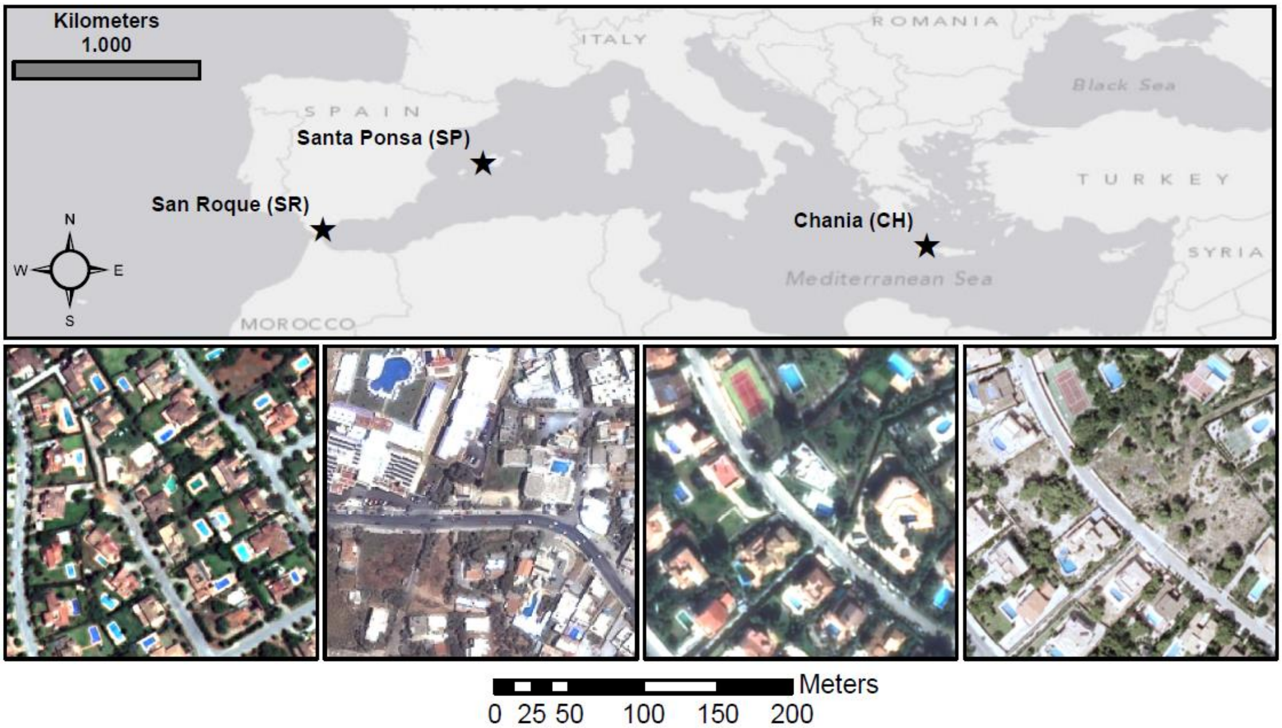

Figure 1: From left to right: pansharpened WorldView-2, San Roque (WV-SR); pansharpened IKONOS, Chania (IK-CH); pansharpened WorldView-2, Santa Ponsa (WV-SP); aerial image, Santa Ponsa (AI-SP).

Imagery

Two WorldView-2 and one IKONOS satellite scenes as well as one digital airborne image are available for the study. Table 1 lists the specifications of the images and Figure 1 shows ~4ha subsets to get a visual impression of the scenes. In case of the three spaceborne datasets which are available in panchromatic and multispectral mode, the Hyperspherical Colour Space method (23) was applied in order to generate pansharpened products. The original panchromatic and multispectral layers were discarded from the analysis. The IKONOS layer stack is only available with three bands, the near-infrared being missing. For each of the images, subsets are defined which restrict the actual study area to equally 125ha. This should guarantee comparable results of analysis steps relying on random data samples. In the following sections, the four analysis-ready datasets are referred to by alias names (first row of Table 1).

Table 1: Remote sensing data.

Dataset alias	"WV-SR"	"IK-CH"	"AI-SP"	"WV-SP"
Preprocessed product (analysis-ready)	8-band pansharpened	3-band pansharpened (VIS)	-	8-band pansharpened
Sensor/platform	*World View 2*	*IKONOS*	Airborne (CCD)	*World View 2*
Study site	*San Roque (Sotogrande)*	*Chania*	*Santa Ponsa*	*Santa Ponsa*
Bands	8MS+1PAN(VIS-NIR)	4MS+1PAN(VIS-NIR)	3 MS (VIS)	8MS+1PAN(VIS-NIR)
GSD (m)	2(MS)/0.5(PAN)	4(MS)/1(PAN)	0.4	2(MS)/0.5(PAN)
Acquisition date	July 16, 2010	July 16, 2006	March 2002	February 16, 2011
Off-nadir (°)	11.90	13.50	n/a	19.61
Sun elevation (°)	69.92	70.04	n/a	36.41
Catalog ID	1030050005F96700	20060716092555000 00011109985	05MA2231-A	103001000942BC00

Reference data collection

A 4-class scheme is applied for targeted mapping of water-relevant urban features by rigidly decomposing a scene into turf grass (vegetation), non-turf grass (vegetation), swimming pool (non-vegetation) and non-swimming pool (non-vegetation). In each scene, 2,500 reference pixels are randomly selected and manually assigned to one of the target classes by visually interpreting directly from the image. It has to be noted that obtained reference sets are characterized by imbalanced class distributions with turf grass and especially pools being the minorities, the latter accounting for only 1-2% of the total.

The 4-class scheme assumes pixels to be discrete which neglects the existence of class boundaries at sub-pixel scales, classes which grade into one another, and the uncertainty when visually interpreting the image. In order to take this into account and to study the influence of uncertain reference information for the training and testing of classification models, each pixel is also annotated by one of three uncertainty categories which are defined as follows:

- UNC0: annotated to pixels which can be assigned exclusively and doubtlessly to one class (usually pixels within homogeneous image regions).
- UNC1: annotated to pixels for which a most appropriate class can be easily identified. However, assignment is not completely free of doubts.
- UNC2: annotated to strongly uncertain pixels. Nevertheless, those cases are assigned to a class as well by relying on the best guess.

Integration of pixel and object domain

Several studies are published, where authors either compare pixel- and object-based image classification (and regression, respectively) or investigate the benefits of their synergetic use (24,25). Usually, pixel and object level outputs are obtained in separate systems and afterwards combined or compared, (e.g., by adding the output of supervised per-pixel classification as categorical layer to a knowledge-based object-oriented system). However, to our knowledge, the integration of both domains into one single system is, if at all, only rarely addressed by researchers. In this paper, we therefore present an integrative approach based on the concept of projecting object-based feature layers of multiple segmentation scales onto the pixel domain. Within a *hierarchical network of image objects* (7), in which the pixel objects constitute the elementary units on the lowest level, each pixel object inherits the object attributes from its corresponding parent objects. Thereby, the spectral feature vector is concatenated with an object feature vector, resulting in a joint feature space for possibly improved classification.

Projecting object-based features to the pixel domain entails several advantages when a representative and unbiased sample of the image data is required, as it is often the case in the context of training, tuning and testing procedures of machine learning applications. One problem concerning objects is that they are of non-uniform size which puts a bias on randomly drawn reference samples (underestimating large objects) (26). This bias would require normalization of the objects' size by corresponding priors (e.g. weights or probabilities). Moreover, using objects as primary reference units assumes the total absence of under-segmentation. Even for fine-scale segmentation levels this is typically not entirely the case. On the contrary, the smallest available image scale, i.e. the pixel domain, guarantees the absence. Projecting object-based feature layers onto the pixel domain in order to extend the original *n*-dimensional matrix of *n*-band imagery supports the statistically reliable application of machine learning.

Features (Predictor Variables)

Image segmentation provides means to represent the image content on multiple scales which is particularly suitable for complex landscapes (27). For this study, we used the Multiresolution Segmentation (28), a region-growing procedure which offers a scale parameter to control the object size by spatial-spectral heterogeneity constraints. For each image, we applied a sequence of systematically incremented scale parameters to generate a five-level segmentation hierarchy. The applied scale sequences are {10, 20, 40, 80, 160} for the WV-SR, IK-CH and WV-SP application cases and {8, 16, 32, 64, 128} for the AI-SP application case. The definition of the scale sequences was guided by a visual check to ensure that the sequence covers the whole span from very fine to very broad image content representation. The other user-defined parameters of the Multiresolution Segmentation were kept as default (shape = 0.1 and compactness = 0.5). Moreover, equal input layer weightings have been maintained, i.e. each layer of the *n* input bands of *n*-band imagery has the same impact on the calculation of the spatial-spectral heterogeneity.

In general, the relationship between segmentation parameters and segmentation result is far from being obvious (29). Therefore, this approach follows the premise that a large set of *uninspected* features is extracted and used for learning a classification model. The set is uninspected, because the features were not pre-evaluated by the human expert concerning their goodness (usefulness) for the classification. Moreover, the features were extracted from systematically incremented scales, which again have not been particularly investigated concerning segmentation quality. The notion is that decisions about the goodness of segmentations and features – two aspects which are strongly interdependent – are preferably handed over to the learning machine (here Random Forests, RF). It is assumed that RF can deal with high-dimensional feature spaces and, furthermore, with redundant variables (multicollinearity) or irrelevant variables which are not correlated with the target variable (i.e. the thematic classes of interest).

Based on the segmentation levels, various object features were generated. Among them are per-object statistics such as the mean, minimum and standard deviation as well as textures (30), geometries (e.g. the object's size) and context information (e.g. intensity difference from neighbour objects). For the empirical experiments, three feature sets are defined in order to assess the use of object-based features in a pixel level processing environment: The first set (PX) contains the *n* standard per-pixel spectral intensity values (of *n*-band imagery), the second set (OB) contains exclusively the features derived from the five segmentation levels. The number of OB features to be extracted depends on the number of input bands, as some object-based features such as the spectral mean value, are calculated per input band. In case of the 8-band data, the total number of OB features amounts to 167 and in case of the 3-band data to 93. The third set (CO) represents the integrative approach, as it combines PX and OB. The three feature sets are compared in terms of classification accuracy.

Random forests

For some time now, RF (6,31) have become increasingly popular in remote sensing research. Several authors (32,33,34,35) applied RF for land cover classification and report good adaptation to high dimensionality and noisy, variant training data. Accuracies are reported to be comparable to the well-established support vector machines, however cheaper to obtain in terms of parameter tuning and computation time.

RF are so-called ensemble learners and predict unknown instances by letting a committee of de-correlated decision trees vote for a particular class. De-correlation is obtained by inducing each tree on bootstrap samples and, furthermore, by selecting features and thresholds at the tree nodes from a random feature subset. The resulting weakness of single trees can be overcome by the ensemble of trees which reduces the generalization error and prevents overfitting (6). The accuracy of RF can be estimated by testing the training instances on those trees for which they have been left out from the bootstrap sampling. According to Breiman (6), the left out cases are referred to as out-of-bag data.

RF offer several possibilities to measure the importance of variables. One is based on the concept of randomly permuting the values of variables at the trees' nodes. The resulting decrease of classification accuracy then provides a measure of variable importance. A formal description of the permutation-based variable importance measure is given by (36). For the current work, the importance scores are used in order to rank image features and to create different feature subspaces, which exclude features beyond a defined rank threshold.

Evaluation criteria

For the estimation of classification accuracies, confusion matrices (37) are calculated based on the out-of-bag data or, as for the evaluation of the effects of uncertain reference information (see Section Uncertainty in Reference Data), with predefined, separate test sets. In order to compare several confusion matrices resulting from different datasets or series of trials, it is beneficial to have one global measure that summarizes the matrix, such as overall accuracy and kappa. However, overall accuracy and kappa are of limited use when assessing highly imbalanced data. Therefore, we assess classification performances by combining per-class producer and user accuracies to one single value (mean producer and user accuracy [MPU]) by simply taking the arithmetic mean.

The classification accuracies are assessed regarding a set of criteria which were suggested by several researchers (38,39,40,41) for the evaluation of machine learning in the field of remote sensing: dimensionality (number of predictor variables), sample size and stability of results. Even though the ensemble approach of RF has stabilization effects (31), instability is an issue, bearing in mind the heterogeneity of urban landscapes and also the extremely unequal class distributions.

RESULTS AND DISCUSSION

Uncertainty in reference data

Figure 2 illustrates that there are considerable degrees of uncertainty associated with all four images. Even for high spatial resolution data such as WorldView-2 or 40 cm GSD airborne imagery, no more than 63 % of the pixels are free of doubt when visually interpreted according to the 4-class scheme. Worst results were obtained for the WV-SP and the IK-CH datasets. WV-SP significantly suffers from unfavourable conditions when the image was taken (off-nadir, sun elevation; cp. to Table 1) and from the hilly terrain of Santa Ponsa. For IK-CH, the high degree of uncertainty is mainly due to the lower spatial resolution, even though Chania represents the most homogeneous urban landscape with rather large structures.

Table 2 shows the classification accuracies obtained with separate train and test subsets of 500 pixels each. The subsets are generated by drawing without replacement from the 2,500-pixel reference set, however, under the condition of a maximally allowed uncertainty level (see Section *Reference data collection*). A UNC0 set, whether it is for training or testing, is a random subsample of only those cases with a UNC0 annotation. A UNC1 set is a random subset of all cases with a UNC0 or UNC1 annotation. Accordingly, a UNC2 set is created by randomly drawing from all 2,500 pixels, irrespective of the annotation. For each constellation, the procedure of subsampling and RF classification is ten times repeated in order to account for variances in the data. Table 2 summarizes the multiple trials by providing the mean and the standard deviations (sd) of recorded MPU results.

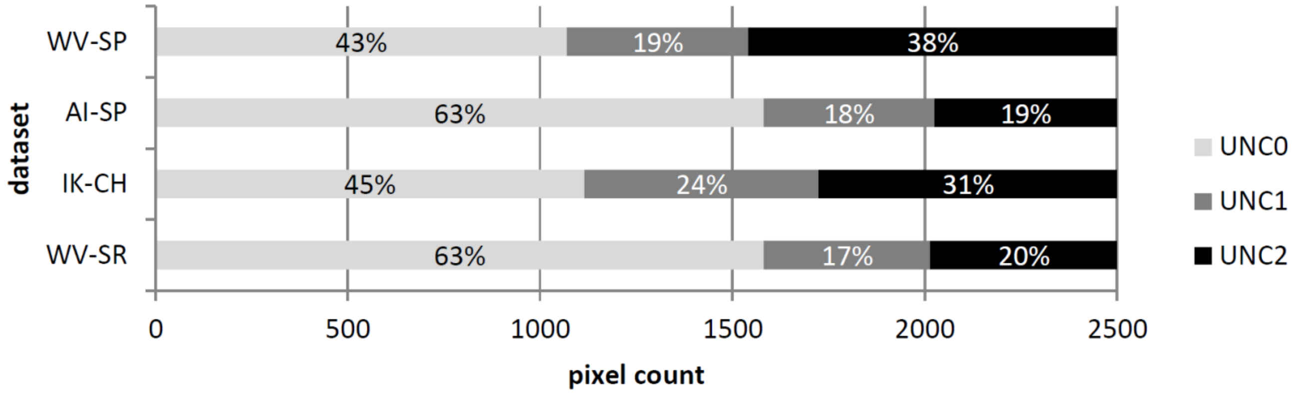

Figure 2: Distribution of uncertain information in data samples (see Table 1 for dataset aliases).

Table 2: Effects of uncertain information in the train and test set on Random Forest performance. MPU are obtained by averaging from ten independent trials; sd denotes standard deviations.

Dataset: WV-SR							*Dataset: AI-SP*						
test data	*UNC0*		*UNC1*		*UNC2*		*test data*	*UNC0*		*UNC1*		*UNC2*	
train data	**MPU**	**sd**	**MPU**	**sd**	**MPU**	**sd**	*train data*	**MPU**	**sd**	**MPU**	**sd**	**MPU**	**sd**
UNC0	0.91	0.02	0.90	0.02	0.90	0.02	*UNC0*	0.81	0.02	0.78	0.04	0.78	0.02
UNC1	0.91	0.03	0.89	0.03	0.89	0.03	*UNC1*	0.82	0.03	0.79	0.03	0.79	0.03
UNC2	0.90	0.02	0.89	0.03	0.89	0.03	*UNC2*	0.80	0.03	0.80	0.03	0.79	0.03
Dataset: IK-CH							*Dataset: WV-SP*						
test data	*UNC0*		*UNC1*		*UNC2*		*test data*	*UNC0*		*UNC1*		*UNC2*	
train data	**MPU**	**sd**	**MPU**	**sd**	**MPU**	**sd**	*train data*	**MPU**	**sd**	**MPU**	**sd**	**MPU**	**sd**
UNC0	0.82	0.04	0.75	0.07	0.72	0.10	*UNC0*	0.89	0.02	0.85	0.02	0.81	0.02
UNC1	0.82	0.03	0.80	0.02	0.76	0.05	*UNC1*	0.88	0.02	0.87	0.03	0.82	0.02
UNC2	0.81	0.03	0.79	0.05	0.77	0.03	*UNC2*	0.84	0.02	0.83	0.02	0.81	0.08

With respect to the four application cases, the accuracy values indicate different responses to the intrusion of vague information. The WV-SR case shows good and stable results across all settings. On the contrary, much stronger effects can be observed for the application cases IK-CH, AI-SP and WV-SP, where the intrusion of uncertainty in the test sets clearly invokes lower values. However, this tendency can be counteracted by using the uncertain information also in the training stage. For example, testing the IK-CH case with a UNC2 sample, the UNC2 train sets lead to stable MPU values around 0.81 (0.03 sd) while UNC0 produced unstable results (0.10 sd) around 0.72.

Generally, it can be noted that training with a limited degree of uncertain information (UNC1) does not deteriorate results if testing against UNC0 and UNC1 sets. However, if one prefers validation with a spatially independent test set (UNC2), the critical, vague cases deliver the crucial decision boundaries of the model, which outweighs the presence of mislabelled cases. The question as to which kind of test data is trusted most can be discussed controversially (42). When estimating the classifier's accuracy, UNC2 test sets provide a spatially independent distribution, however – being highly erroneous themselves – very pessimistic estimates. On the contrary, estimates based on UNC0 test sets are optimistic, because critical image regions are excluded from evaluation. Finally, the choice depends on the application and subsequent use of the product (42). For our work we prefer an exclusion of uncertain cases. Therefore, analysis steps described in the following sections are based on UNC0 data.

Evaluation of feature sets

In this section, we compare the performances of the object and pixel domain and investigate, if combining both can improve the results. For each feature set (PX, OB and CO), the out-of-bag MPU values of 20 independent trials (ensuring a random permutation of train and out-of-bag data for each tree) are recorded. The results are shown in Figure 3, from which it becomes clear that the OB set outperforms the PX set. Moreover, on all four application cases, using the complete set (CO) leads to further improvement, even if marginal. Comparing the performances across the four application cases shows that both 8-band datasets produce better and, in case of the PX sets, also more stable results.

In general, the results indicate that the complementary use of the pixel and object domains, realized within one single system, brings added value.

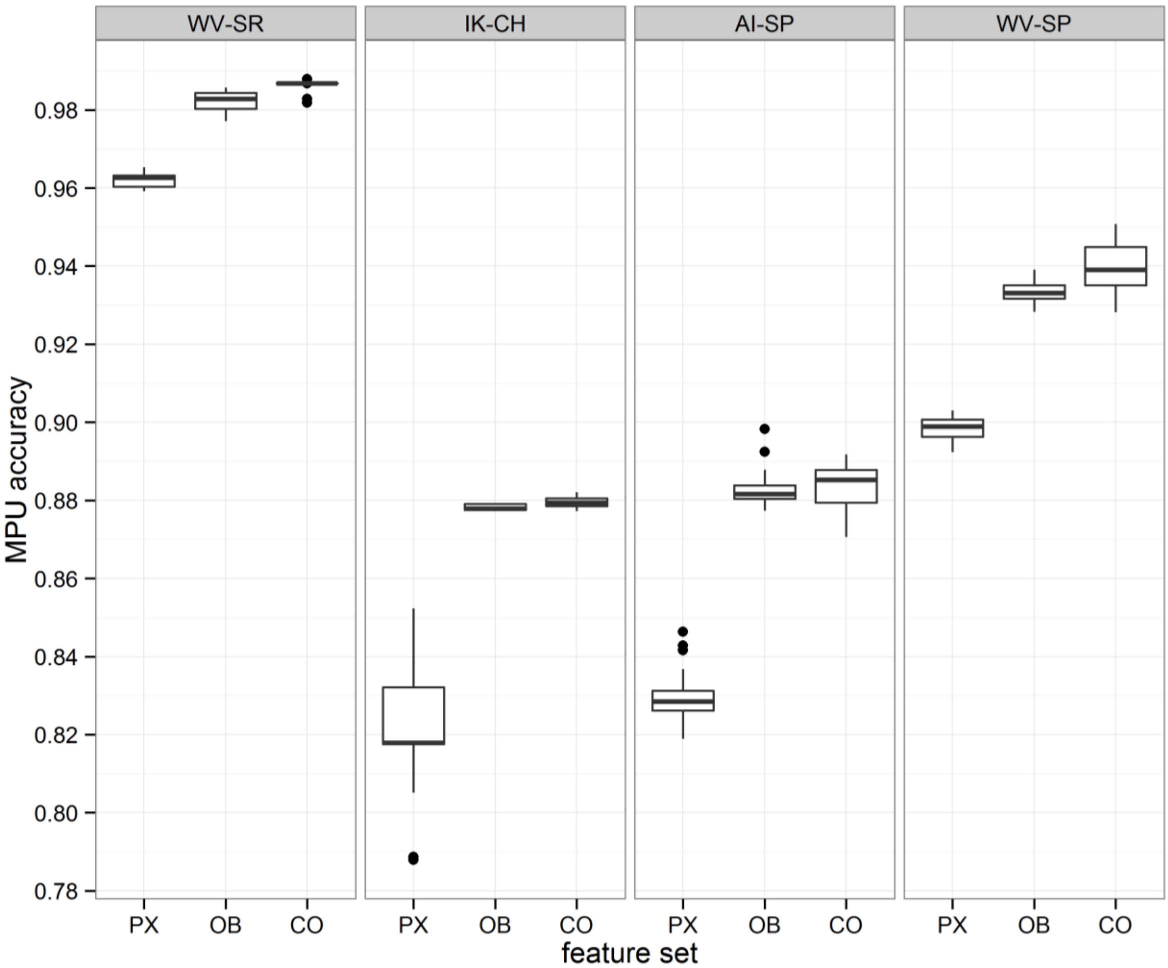

Figure 3: Performance comparison of three feature sets: per-pixel spectral values (PX); object-based features (OB); and the complete set (CO) which combines PX and OB.

Sample size and dimensionality

A high number of features can, on the one hand, increase inter-class separability, but on the other hand degrade accuracies due to the so-called curse of dimensionality – also referred to as Hughes phenomenon (43) – which is related to the number of available training cases (44).

Based on the results of the previous section, we advocate the synergetic use of the pixel and the object domains, but it remains an open question, whether better subspaces exist within CO. In other words, it is to investigate if redundant or irrelevant features need to be discarded in order to obtain an optimized CO feature set for better results. Therefore, the response of the classifier to both training set size and dimensionality is of interest. Within a nested loop, dimensionality and training set size are stepwise increased:

1. RF-based feature ranking (ranks from j = 1 to J); J = total number of features
2. For i = 1 to 100 (logarithmic spaced sequence)
    a. For j = 1 to J (logarithmic spaced sequence)
        i. Draw i % of the reference pixels
        ii. Select j top-ranked features
        iii. Construct RF_{ij} model

Figure 4 shows the interrelationship of classifier performance, training set size and dimensionality. While it is not surprising that higher train set sizes generally increase accuracies, it can be noted that high dimensionality has no significant negative effects, irrespective of the training set size. Considering that added features become weaker (less predictive or more inter-correlated) with increasing dimensionality, declining performance towards the right border of the plots in Figure 4 could be expected. As this is not the case even for small training sets, the Hughes phenomenon does not apply to our experimental settings.

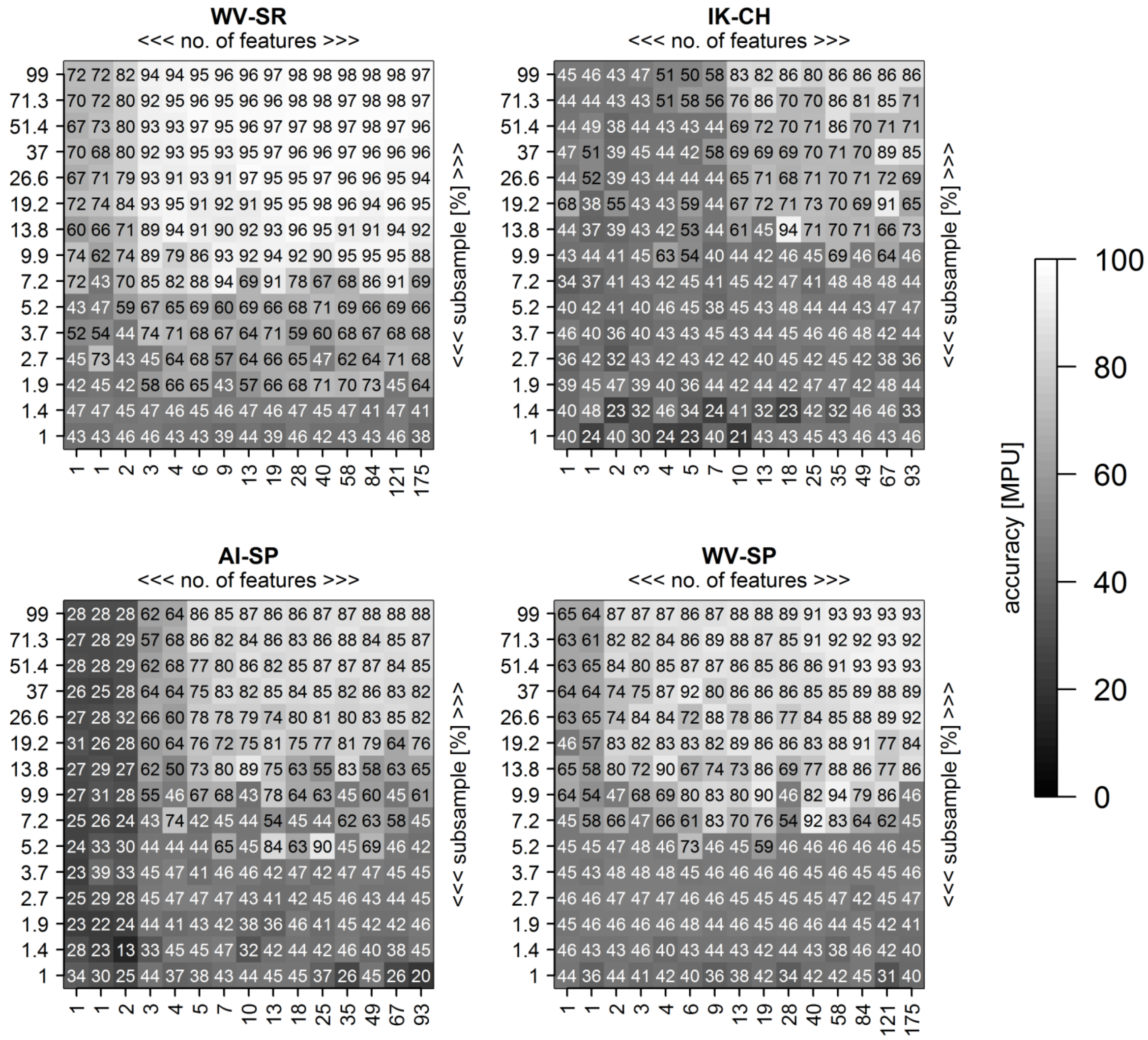

Figure 4: Dependency of RF performance (MPU) on dimensionality (number of features) and training set size.

The requirements for training set size strongly vary across the four application cases. For both 3-band datasets it can be expected that increasing the sample size beyond the amount available for this study could further improve the results, while for both 8-band datasets accuracies seem to saturate with smaller sample proportions (about a quarter for WV-SR and about half for WV-SP).

Table 3 shows the error matrices obtained on a trial using all features (CO) and 100% of the UNC0 sample for training (an example map product for the WV-SR dataset is shown in Figure 5). While the swimming pools are well discriminated on all datasets, the turf grass class shows high confusion rates (i.e., low producer and user accuracies) for the 3-band datasets.

Table 3: Out-of-bag confusion matrices, per-class producer and user accuracies, and MPU values (dark cells), obtained using all features and full UNC0 training set.

WV-SR

cla/ref	turf	non-turf	non-pool	pool	user
turf	144	3	0	0	0.980
non-turf	3	588	1	0	0.993
non-pool	0	1	805	1	0.998
pool	0	0	0	32	1.000
producer	0.980	0.993	0.999	0.970	0.989

IK-CH

cla/ref	turf	non-turf	non-pool	pool	user
turf	12	4	8	0	0.500
non-turf	1	186	8	0	0.954
non-pool	4	15	858	2	0.976
pool	0	0	1	21	0.955
producer	0.706	0.907	0.981	0.913	0.861

AI-SP

cla/ref	turf	non-turf	non-pool	pool	user
turf	38	9	3	0	0.760
non-turf	9	351	24	1	0.912
non-pool	3	32	1073	3	0.966
pool	0	0	3	37	0.925
producer	0.760	0.895	0.973	0.902	0.887

WV-SP

cla/ref	turf	non-turf	non-pool	pool	user
turf	45	6	0	0	0.882
non-turf	3	457	1	0	0.991
non-pool	0	3	525	3	0.989
pool	0	0	6	29	0.829
producer	0.938	0.981	0.987	0.906	0.938

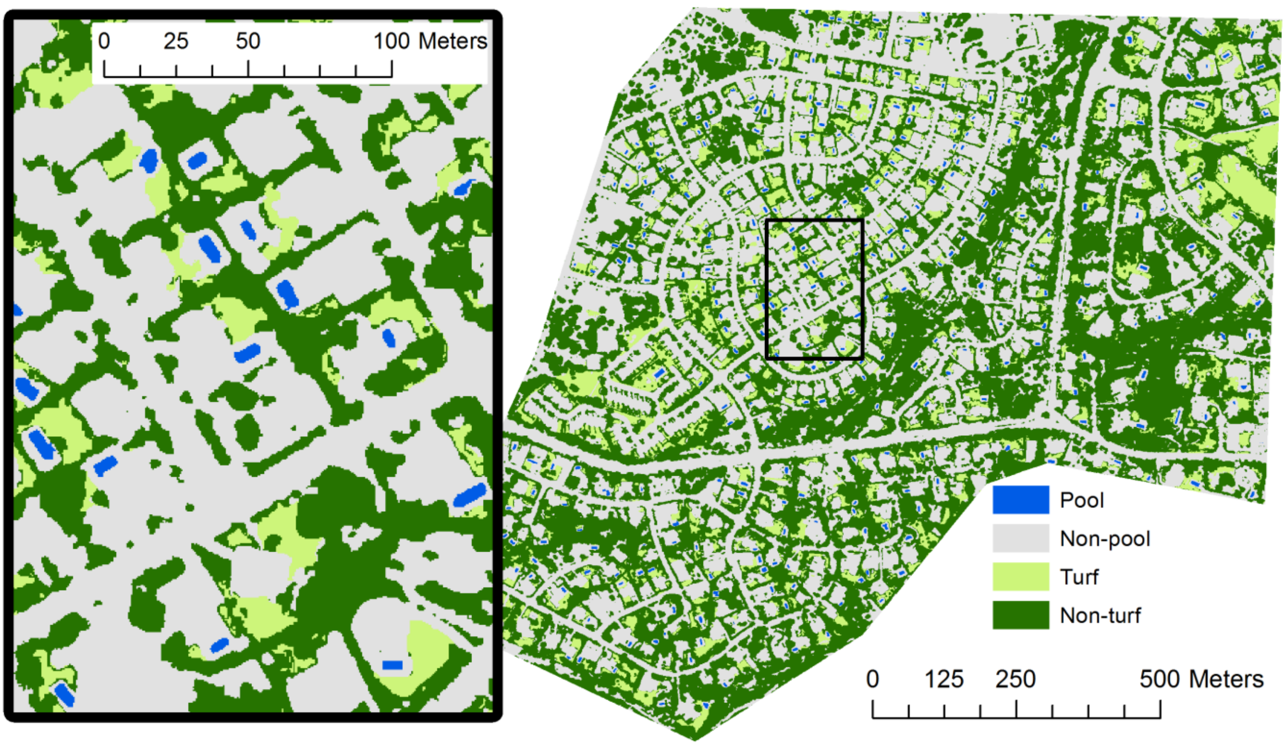

Figure 5: Classification map of the WorldView-2 San Roque dataset (WV-SR) using the complete feature set (CO) and the UNC0 reference data.

SUMMARY AND CONCLUSIONS

This paper has investigated methodological aspects for the automated mapping of Mediterranean urban landscape elements with relevance to water consumption studies. An integrative approach was presented which combines spectral per-pixel features and object features derived from multiple segmentation levels, thereby clearly improving the prediction accuracy of the classification models induced by the RF algorithm. Comparing the results across the four datasets revealed significant differences in terms of map accuracy, training data requirements and the amount of ambiguous or non-interpretable pixels. It is remarkable that significant differences can be observed even for the WV-SR and WV-SP application cases, i.e., datasets from the same sensor (WorldView-2). Thus, apart from spatial and spectral resolution, further aspects need to be considered for the mapping of features like pools and turf grass. Possibly, the worse results of WV-SP can be attributed to the unfavourable acquisition date (February, cp. Table 1) which entails a low sun elevation angle and hence increased shadowing. Furthermore, turf grass of private gardens can be best discriminated from shrub/tree vegetation when maintained and irrigated, as it is predominantly the case in the summer season. The IKONOS image generally produced worst results even though Chania represented the least heterogeneous urban structure among the study areas. Presumably, a combination of relatively low resolution (1 m) and three visible bands is not enough to obtain reliable results. Generally, the results indicate that the combination of high spatial resolution and relatively high spectral resolution, as is the case with WorldView-2, has great potential for accurate and detailed mapping of even small urban structures like private gardens and enables tree/shrub- and turf grass-dominated patches to be discriminated. With favourable image quality it is also possible to reduce the required amount of labelled reference data which are typically difficult or costly to obtain and which often constitute the limiting factor in operational applications of remote sensing-based thematic mapping. Overall, this approach confirms the assumption of Mathieu et al. (18) who mapped private gardens in a low-density residential area, but achieved less thematic granularity in the separation of trees/shrubs and turf grass while requiring considerable efforts for manual intervention at the same time. Hence, they suggested sub-meter resolution imagery, arguing that this can better resolve the landscape elements of interest and, furthermore, add value by textural features.

From an urban resource management point of view, the automated mapping approach presented here derives spatial information that is relevant to a broad range of scientific, planning, and management issues in urban areas. Mapping the number and surface area of swimming pools, the extent, distribution, density and design of urban green spaces and gardens is a first step in assessing and comparing the metabolism of these new urban landscapes with respect to water consumption. Making inventory of water-related outdoor infrastructure is a prerequisite for assessing landscape water demand in follow-up research. For example, a Spanish study inventoried the total household landscape area and the area occupied by turf, trees and shrubs manually from aerial photographs and used a method based on reference evapotranspiration to estimate net landscape irrigation requirements. By analysing water billing records and taking advantage of dual measurements of water supply (separate indoor and irrigation water metering), actual irrigation and net irrigation requirements were compared to assess the level of overirrigation in private household gardens (45). The image analysis approach presented here provides synoptical data on water-related outdoor infrastructure from automated mapping. Ultimately, such geographical research delivers information for decision support and improved water management. While this paper has focused on the Mediterranean region, the methodological approach is transferable to all urban contexts where new landscapes of water consumption evolve.

ACKNOWLEDGEMENTS

The authors would like to thank the Mediterranean Agronomic Institute of Chania (MAICh) and *Geosystems Hellas* for granting access to the IKONOS data used in this study. We are grateful to DigitalGlobe for providing WorlView-2 imagery of Santa Ponsa (catalog ID 103001000942BC00) as part of the 8-Band Research Challenge in 2010 and for permitting further research use of the imagery.

REFERENCES

1 Netzband M, C L Redman & W L Stefanov, 2007. Remote sensing as a tool for urban planning and sustainability. In: Applied Remote Sensing for Urban Planning, Governance and Sustainability, edited by M Netzband, W L Stefanov & C L Redman (Springer), 1-23

2 Herold M & D A Roberts, 2010. The spectral dimension in urban remote sensing. In: Remote Sensing of Urban and Suburban Areas, edited by C Jürgens & T Rashed (Springer), 47-66

3 Herold M, M Gardner & D A Roberts, 2003. Spectral resolution requirements for mapping urban areas. IEEE Transactions on Geoscience and Remote Sensing, 41(9): 1907-1919

4 Sliuzas R, M Kuffer & I Masser 2010. The spatial and temporal nature of urban objects. In: Remote Sensing of Urban and Suburban Areas, edited by C Jürgens & T Rashed (Springer), 67-84

5 Jacobsen K, 2003. Geometric potential of IKONOS and QuickBird images. In: Photogrammetric Weeks 2003, edited by D Fritsch (Wichmann), 101-110

6 Breiman L, 2001. Random Forests. Machine Learning, 45(1): 5-32

7 Benz U, P Hofmann, G Willhauck, I Lingenfelder & M Heynen, 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing, 58: 239-258

8 Atkinson P M & G M Foody, 2002. Uncertainty in remote sensing and GIS: Fundamentals. In: Uncertainty in Remote Sensing and GIS, edited by G M Foody & P M Atkinson (Wiley), 1-18

9 Boyd D S & G M Foody, 2011. An overview of recent remote sensing and GIS based research in ecological informatics. Ecological Informatics, 6 (1, SI): 25-36

10 Iglesias A, L Garrote, F Flores & M Moneo, 2007. Challenges to manage the risk of water scarcity and climate change in the Mediterranean. Water Resources Research, 21: 775-788

11 EEA (European Environment Agency), 2006. Urban Sprawl in Europe - The Ignored Challenge. EEA Report No 10/2006. European Environment Agency, Copenhagen, Denmark, Luxembourg

12 Domene E & D Saurí, 2006. Urbanisation and water consumption: Influencing factors in the Metropolitan Region of Barcelona. Urban Studies, 43(9): 1605-1623

13 March H. & D Saurí, 2010. The suburbanization of water scarcity in the Barcelona metropolitan region: Sociodemographic and urban changes influencing domestic water consumption. The Professional Geographer, 62(1): 32-45

14 Balling Jr R C, P Gober & N Jones, 2008. Sensitivity of residential water consumption to variations in climate: An intraurban analysis of Phoenix, Arizona. Water Resources Research, 44(10): W10401

15 Haase D & H Nuissl, 2007. Does urban sprawl drive changes in the water balance and policy? The case of Leipzig (Germany) 1870-2003. Landscape and Urban Planning, 80(1-2): 1-13

16 Domene E, D Saurí & M Pares, 2005. Urbanization and sustainable resource use: The case of garden watering in the metropolitan region of Barcelona. Urban Geography, 26(6): 520-535

17 Hof A & T Schmitt, 2011. Urban and tourist land use patterns and water consumption: evidence from Mallorca, Balearic islands. Land Use Policy, 28(4): 792-804

18 Mathieu R, C Freeman & J Aryal, 2007. Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery. Landscape and Urban Planning, 81(3): 179-192

19 Vidal M, E Domene & D Saurí, 2011. Changing geographies of water-related consumption: residential swimming pools in suburban Barcelona. Area, 43(1): 67-75

20 Kent M, R Newnham & S Essex, 2002. Tourism and sustainable water supply in Mallorca: a geographical analysis. Applied Geography, 22(4): 351-374

21 Paneque Salgadoa P, S Corral Quintana, A Guimaraes Pereirac, L del Moral Ituarte & B Pedregal Mateos, 2009. Participative multi-criteria analysis for the evaluation of water governance alternatives. A case in the Costa del Sol (Malaga). Ecological Economics, 68(4): 990-1005

22 Voudouris K, Th Mavrommatis, P Daskalaki & G Soulios, 2006. Rainfall variations in Crete island (Greece) and their impacts on water resources. Publicaciones del Instituto Geologico y Minero de España. Serie: Hidrogeologia y aguas subterr´aneas No 18, Karst, Climate Change and Groundwater, Madrid, 453-463

23 Padwick C, M Deskevich, F Pacifici & S Smallwood, 2010. WorldView-2 pan-sharpening. In: ASPRS 2010 Annual Conference (San Diego, California, April 26-30, 2010) (last date accessed: 26.06.2013)

24 Gonçalves L, 2011. A hybrid approach classification of remote sensing images. In: Proceedings of the 31st EARSeL Symposium and 35th General Assembly 2011, edited by L Halounová, 328-335

25 Myint S W, P Gober, A Brazel, S Grossman-Clarke & Q Weng, 2011. Per-pixel versus object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment, 115 (5): 1145-1161

26 Radoux J, P Defourny & P Bogaert, 2008. Comparison of pixel- and object-based sampling strategies for thematic accuracy assessment. Archives of The International Society for Photogrammetry and Remote Sensing, XXXVIII (4/C1)

27 Blaschke T & G J Hay, 2001. Object-oriented image analysis and scale-space: Theory and methods for modeling and evaluating multi-scale landscape structure. Archives of The International Society for Photogrammetry and Remote Sensing, vol. 34, part 4/W5: 22-29

28 Baatz M & Schäpe A, 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. Angewandte Geographische Informationsverarbeitung, XII: 12-23

29 Feitosa R Q, R S Ferreira, C M Almeida, F F Camargo & G A O P Costa, 2010. Similarity metrics for genetic adaptions of segmentation parameters. Archives of The International Society for Photogrammetry and Remote Sensing, XXXVIII-4/C7

30 Haralick R M, K Shanmugam & I Dinstein, 1973. Textural features for image classification. Transactions on Systems, Man and Cybernetics, 3 (6): 610-621

31 Hastie T, R Tibshirani & Friedmann J, 2009. The Elements of Statistical Learning. Data Mining, Inference, and Prediction (2nd edition, Springer, New York) 745 pp.

32 Gislason P O, J A Benediktsson & J R Sveinsson, 2006. Random forests for land cover classification. Pattern Recognition Letters, 27: 294-300

33 Ham J, Y Chen, M M Crawford & J Ghosh, 2005. Investigation of the random forest framework for classification of hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 43: 492-501

34 Pal M, 2005. Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26: 217-222

35 Wolf N, 2011. Feature evaluation for a transferable approach of object-based land cover classification based on Ikonos and QuickBird satellite data. Photogrammetrie, Fernerkundung und Geoinformation, 3: 135-144

36 Strobl C, A-L Boulesteix, T Kneib, T Augustin & A Zeileis, 2008. Conditional variable importance for random forests. BMC Bioinformatics, 9(307)

37 Congalton R G, 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37: 35-46

38 DeFries R S & J Chan, 2000. Multiple criteria for evaluating machine learning algorithms for land-cover classification from satellite data. Remote Sensing of Environment, 74: 503-515

39 Lippitt C D, 2008. Mapping selective logging in mixed deciduous forest: A comparison of machine learning algorithms. Photogrammetric Engineering and Remote Sensing, 74(10): 1201-1212

40 Pal M & P M Mather, 2003. An assessment of the effectiveness of decision tree methods for land-cover classification. Remote Sensing of Environment, 30: 554-565

41 Rogan J, J Franklin, D Stow, J Miller, D A Roberts & C Woodcock, 2008. Mapping land cover modifications over large areas: A comparison of machine learning techniques. Remote Sensing of Environment, 112(5): 2272-2283

42 Plourde L & R G Congalton, 2003. Sampling method and sample placement: How do they affect the accuracy of remotely sensed maps? Photogrammetric Engineering and Remote Sensing, 69: 289-297

43 Hughes G F, 1968. On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory, IT-14(1): 55-63

44 Camps-Valls G & L Bruzzone, 2005. Kernel-based methods for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 43(6): 1352-1362

45 Salvador R, C Bautista-Capetillo, E Playán, 2011. Irrigation performance in private urban landscapes: A study case in Zaragoza (Spain). Landscape and Urban Planning, 100(3): 302-311