Article

Land Cover and Land Use Classification with TWOPAC: towards Automated Processing for Pixel- and Object-Based Image Classification

Juliane Huth [1,*]**, Claudia Kuenzer** [1]**, Thilo Wehrmann** [1]**, Steffen Gebhardt** [2]**, Vo Quoc Tuan** [1] **and Stefan Dech** [1]

[1] German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Oberpfaffenhofen, D-82234 Wessling, Germany; E-Mails: claudia.kuenzer@dlr.de (C.K.); thilo.wehrmann@gmail.com (T.W.); Tuan.VoQuoc@dlr.de (V.Q.T.); stefan.dech@dlr.de (S.D.)

[2] Remote Sensing Unit, Institute of Geography, University of Wuerzburg, Am Hubland, D-97074 Wuerzburg, Germany; E-Mail: gebhardt.steffen@gmail.com

* Author to whom correspondence should be addressed; E-Mail: juliane.huth@dlr.de; Tel: +49-8153-283-281; Fax: +49-8153-281-458.

Abstract: We present a novel and innovative automated processing environment for the derivation of land cover (LC) and land use (LU) information. This processing framework named TWOPAC (TWinned Object and Pixel based Automated classification Chain) enables the standardized, independent, user-friendly, and comparable derivation of LC and LU information, with minimized manual classification labor.

TWOPAC allows classification of multi-spectral and multi-temporal remote sensing imagery from different sensor types. TWOPAC enables not only pixel-based classification, but also allows classification based on object-based characteristics. Classification is based on a Decision Tree approach (DT) for which the well-known C5.0 code has been implemented, which builds decision trees based on the concept of information entropy. TWOPAC enables automatic generation of the decision tree classifier based on a C5.0-retrieved ascii-file, as well as fully automatic validation of the classification output via sample based accuracy assessment. Envisaging the automated generation of standardized land cover products, as well as area-wide classification of large amounts of data in preferably a short processing time, standardized interfaces for process control, Web Processing Services (WPS), as introduced by the Open Geospatial Consortium (OGC), are utilized. TWOPAC's functionality to process geospatial raster or vector data via web resources (server, network) enables

TWOPAC's usability independent of any commercial client or desktop software and allows for large scale data processing on servers. Furthermore, the components of TWOPAC were built-up using open source code components and are implemented as a plug-in for Quantum GIS software for easy handling of the classification process from the user's perspective.

1. Introduction

Land cover describes the physical cover of the Earth's surface including vegetation, non-vegetation and man-made features, while land use is characterized by anthropogenic activities to modify, manage and use certain types of land cover [1–5]. Detailed knowledge of land cover and land use is a key parameter for sustainable planning in the context of agriculture, urban development, and nature conservation. Land cover (LC) and land use (LU) are two examples of value added information, which can be derived from Earth Observation (EO) data.

Land cover information from satellite imagery is a widely-used tool for large area environmental monitoring and analyses of change processes on the Earth's surface [6–9]. Global networks for the assessment of the impact of climate change (e.g., GTOS (Global Terrestrial Observing System), GCOS (Global Climate Observing System), *etc*.) consider land cover as one of 13 essential climate variables beside biomass, leaf area index, fire disturbance, water use, *etc*. [10]. Additionally, land cover products form a major input for further modeling and analysis, e.g., in hydrologic modeling, and trend analysis [11–15], biomass and energy modeling [16–20], population density modeling [21,22], risk and hazard analysis [23–25], or ecosystem assessment [26,27]. Innovative environmental planning and management tools, such as Decision Support Systems, also require high-quality land cover and land use information derived from high and very high resolution data at a regional scale [6]. With the expanding community of scientific and non-scientific users not necessarily experts in remote sensing image analysis techniques (e.g., hydrologists, socio-economists, or political stakeholders and decision makers), the requirements for simplified, automated processing environments yielding the above mentioned LC and LU products are increasing.

During the last decades, new sensors and easily accessible data archives have increased the amount of available remote sensing data. This development necessitates robust, transferable and automated methods and processing services for the derivation of accurate information that can be produced within a short period of time [8,28]. Furthermore, LULC change analyses based on remote sensing data require comparable and stable methods for the initial generation of classification products [29–33].

In general, comprehensive and well-known image processing software packages require a large amount of knowledge about the utilization of the specific software and are mostly limited to single date image classification. Standard software also generally requires high cost software licenses. However, to support classification of large amounts of datasets a comprehensive processing environment is required.

Despite this, very common classification procedures with a large amount of user interaction (e.g., by visual interpretation using local knowledge, image classification followed by manual post-processing and correction in GIS, *etc*.) are still advantageous for solutions that aim for a single derivation of a high accuracy land use map for a small area.

In this paper we introduce the concept and implementation of a methodology to classify image data within an automated processing environment called TWOPAC—*TW*inned *O*bject and *P*ixel based *A*utomated classification *C*hain. The objective is to present a framework to derive LULC products independent from known image classification software packages to allow for large scale data processing. The TWOPAC process chain aims at semi-automated processing of the full classification chain (including sampling, classification, post-processing) to provide a framework for image classification and to reduce user interaction possibilities.

In a nutshell, TWOPAC is a software tool that consists of defined modules for data pre-processing, image classification and validation of the results. The processing steps are conducted in a standardized procedure and implemented in an open source-based software solution. TWOPAC's functionality and processing is built as Open Spatial Consortium (OGC) compliant web services to process geospatial raster or vector data via web resources (server, network).

In the first step TWOPAC is introduced thereby presenting separately the aspects of data pre-processing and preparation, classification, and validation. Secondly, several example applications utilizing the TWOPAC classification chain are presented, including the presentation of classification results, accuracy assessment, and discussion. The paper closes with a conclusion, followed by a comprehensive outlook on further developments.

2. Method

2.1. TWOPAC Overview

The basic classification methodology of TWOPAC is a supervised decision tree classification where the tree is derived using representative training areas of each land cover type of interest [34,35]. Currently, required pre-processing and data input preparation such as image segmentation and feature space generation, are not yet included in the TWOPAC environment. An extensive training and validation data set, which we call "sampling data", is selected from the imagery. The only manual process is the analyst-based assignment of LC and LU classes to these samples. Supervised classification is then performed automatically in the TWOPAC environment with a certain user-defined split percentage of training and validation samples. For the Decision Tree based Classification, (DTC) the well-known C5.0 [36] code has been implemented within TWOPAC, which builds up a set of decision rules based on the concept of entropy. TWOPAC enables not only pixel-based classification, but also allows classification based on object-based characteristics, such as shape parameter and object statistics. Furthermore, classification results are automatically validated using set aside validation data.

The main steps of the TWOPAC approach are presented in detail in the following paragraphs (for an overview, also see Figure 1).

Figure 1. Flowchart of the Twinned Object- and Pixel based Automated classification Chain (TWOPAC).
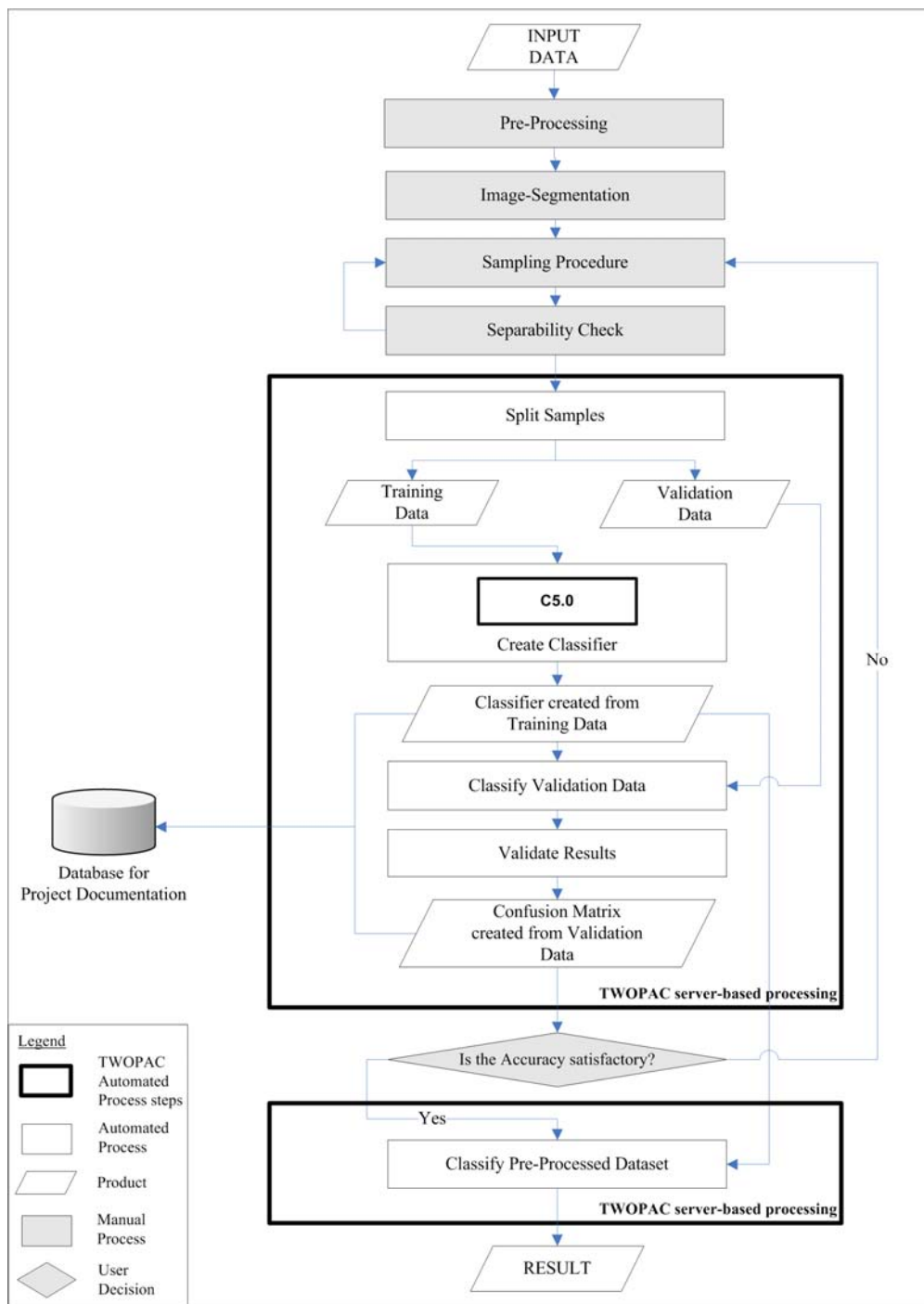
2.2. Data

TWOPAC can analyze multi-spectral and multi-temporal remote sensing data of various spatial resolutions, such as very high resolution data (e.g., Worldview and Quickbird imagery), high resolution data (e.g., Rapid Eye), medium resolution data (e.g., Landsat, SPOT, and Aster), as well as coarse resolution data (e.g., MODIS or AVHRR). Hence, usage of TWOPAC does not depend on spatial resolution or satellite sensor.

For reference data a number of auxiliary information types can be optionally used in order to enhance the identification of samples. These auxiliary datasets could be ground truth data collected during field campaigns, selected vector data, or topographic maps.

2.3. Input Data Generation

The first step in data pre-processing includes geo-referencing and atmospheric correction of the imagery. These work steps generally required by all image classification tools are not yet included in the TWOPAC environment. In our work, geo-referencing has been conducted with ground control points collected in the field as well as in spatially highly accurate TerraSAR-X Stripmap images, which are geo-referenced with high accuracy [37]. For atmospheric correction we utilized the well-known ATCOR-2 atmospheric correction software based on MODTRAN, which allows the derivation of surface reflectance for each pixel [38]. Atmospheric correction results are preferably validated comparing image spectra of pseudo invariant surfaces with spectral reflectance information collected *in-situ*.

Pre-processing in advance of TWOPAC includes the generation of a segmentation layer from an extended image layer stack. Currently, this step is still outside of the TWOPAC framework, realized through the commercial software eCognition Developer. It is envisaged to create these segmentation layers by self-developed or freely accessible segmentation algorithms and to integrate this pre-processing step into the TWOPAC processing environment, in the near future.

Segmentation layers are derived from the geo- and atmospherically corrected images, as the aim is to not classify on pixel basis exclusively, but also to consider the feature information of segments, enabling classification on an object basis as well. ECognition considers the spatial and spectral properties of neighboring pixels of a multi-spectral image stack to group them in homogeneous regions, so called objects or segments [39]. The adjustment of the segmentation layer which is parameterized by scale, color, and shape can be considered as a rather subjective process. However, some measures for objective image segmentation evaluation were proposed by Möller *et al.* and Neubert *et al.* [40,41]. The image objects have to be subsequently checked for their suitability for classification by visual interpretation.

Dependent on the classification approach selected, *i.e.*, pixel- or object-based classification, small methodological differences, related to the segment size and the feature space generation, have to be considered. For the object-based approach segmentation aims for the depiction of real world objects, e.g., channels and river objects are represented by long, thin polygons; agricultural fields are preferably compact objects separated along field borders or irrigation channels, *etc*. (depending on the study region). The feature space which is generated for each segment contains several geometry features such as *length-to-width*, *shape index*, *area*, *compactness*, *etc*. as well as several segment features that describe spectral information (e.g., band statistics such as reflectance *mean values* of all pixels per segment). Features related to spatial information between objects can provide supplementary information for classification [42]. Distance and topological neighboring properties of objects are possible to be used as well, which were not applied for the test cases presented in this study.

For pixel-based classification a segmentation layer which is characterized by strong over-segmentation is generated. The aim is not to represent real world objects (as e.g., rice fields for object-based classification, *etc*.) but to group neighboring pixels together according to their spectral similarity as it

is performed in region growing approaches [43]. Within the sampling procedure (see next subsection) assigning classes to small segments is considered easier than based on single pixels, and is expected to support and accelerate sampling. To ensure the elimination of mixed pixel information from the border of segments, the user applies an inner buffer using a sensor specific buffer size of 1–2 times the pixel size. The pixel-based classification is related to the spectral information of single pixels only. However, additional spectral features/indices are generated to enlarge the feature space. We have typically processed Principal Components (usually PC 1,2,3) as well as Tasselled Cap Components (TCC: Brightness, Greenness, Wetness), and selected indices (e.g., Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI)) as they have been shown to increase classification accuracy [2,5,34,44–49].

To sum up, the data input generation comprises a segment layer for object-based classification, or a segment layer plus a multi-spectral image layer for pixel-based classification. The object-based feature space is described by attributes such as shape and statistical parameters which characterize the spectral values within a segment. The feature space for pixel-based classification is defined by the spectral information of the dataset's multi-spectral bands, and the additionally generated bands (PC, TCC, Indices, *etc*.). The feature space for both, pixel or object-based classification can also be increased with additional information from pre-classification results, information generated from radar imagery, topographical information, *etc*., depending on the availability and suitability of such additional information.

2.4. Sampling Data Generation

In supervised classification, image data are classified based on the analyses of training data, which describe the statistical characteristics (min, max, mean, variability *etc*.) of selected LC and LU types. Unsupervised methods subdivide the image data into a predefined number of spectral classes based on spectral characteristics without requiring analyst-specific training information. The accuracy of unsupervised classification can increase with increasing spectral separability of the classes [34], which however, usually still not lead to sufficient results [35,49]. For this reason, TWOPAC rather follows supervised classification approaches.

Samples for training and validation are selected by the user. Samples are required in ESRI Shapefile format and can be either polygon features, marking the boundaries of a training/validation area, or point features located at the center of the training/validation pixels.

For the sample selection procedure Quantum GIS [50] is used. In general, any GIS software supporting ESRI Shapefiles can be used, but we prefer Quantum GIS since it is a free and easy to handle software with several features supporting such sample data selection. For the sampling procedure the image layer stack is overlaid with the generated segment layer and auxiliary vector layers (if available). In the case of pixel-based classification the segmentation layer is very useful though not mandatory for the sampling process. The segments are selected by the user aiming at an even distribution over the image, and are automatically weighted according to their areal proportion of the classes to avoid over- or under prediction [51]. Selected samples are assigned to a land cover class. The class scheme has to be created according to the intended applications of the classification result. TWOPAC does not depend on specific classification scheme types. The selection of samples requires a

certain amount of knowledge about the satellite image and geographic region. Besides, the reference data need to be of comparable quality to the envisaged features intended to be derived from the imagery.

According to Congalton and Green [52] it is advisable to collect at least 50 samples per class when covering 1 million acres (about 4,050 km² = coverage of one SPOT scene) and less than 12 classes. It is advisable to stick to this rule, but modifications are possible, depending on the data variance within individual classes. Heterogeneous classes for instance, such as the "settlement" or "fruit orchards", require more samples for successful classification than the rather homogeneous class "water".

Figure 2. Sampling procedure—selection and extraction of segments with features.
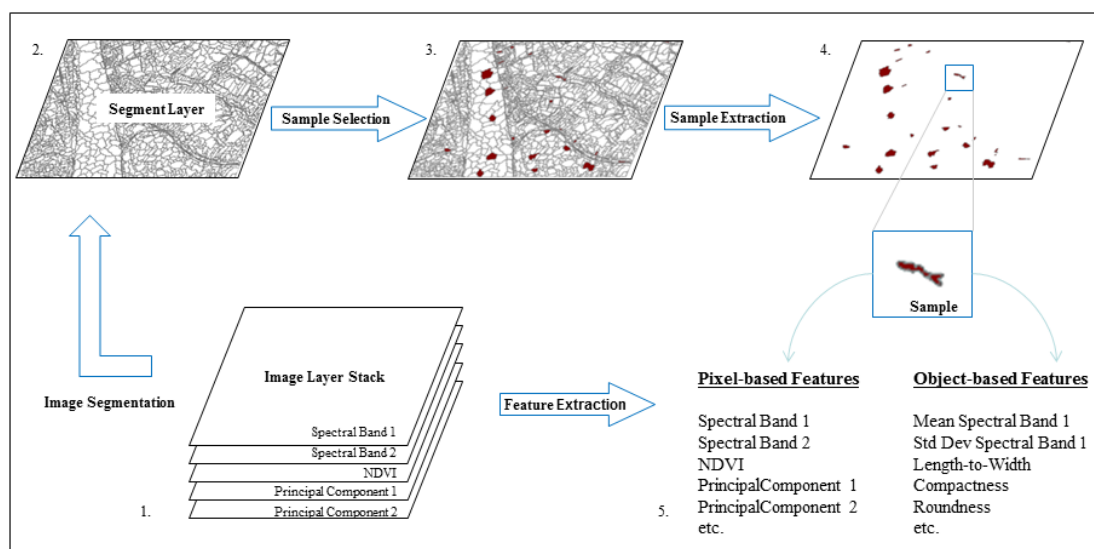
The result of the sampling procedure (see Figure 2) should be an ESRI Shapefile, which contains the sample segments, enriched with their segment features, and class labels as attributes. This dataset is the input for the automated steps of the TWOPAC classification chain.

2.5. TWOPAC Automated Processing Steps

At this point the manual steps outside of the TWOPAC framework are completed and all subsequent processes are conducted in an automated manner. The graphical user interface (GUI) of TWOPAC allows for stepwise, iterative, and interactive implementation of the classification process, comprising the modification of the sample splitting percentage (see Section 2.5.1), the selection between pixel and object-based approaches, as well as the monitoring of intermediate results.

2.5.1. Preparation of the Classification

TWOPAC starts with the division of samples into training and validation data. The percentage of samples used for training and validation can be set by the user (usually 2/3:1/3), but the optimal ratio is not a widely published parameter and is recommended to be tested by the user. The next decision of the user is whether classification will be processed pixel-based or object-based. For pixel-based classification the ID of the originating sample segment is automatically incorporated into separation process to avoid usage of training and validation sample points from the same segment.

The subsequent TWOPAC work step is the creation of the classifier based on the training samples. Following, the term *classifier* is used for the generated rules for separation of training data into predefined classes.

2.5.2. Classification Algorithm

The current implementation of TWOPAC features the C5.0 based decision tree classification [36]. This automatic tree building software conducts data mining and decision tree estimation. The algorithm searches through a given set of training cases to find the best feature for separation into pre-defined classes. The statistical measure *information gain ratio* is used as a splitting criterion to estimate how well a selected feature separates the training cases according to their class membership [36,53,54]. In general, decision tree algorithms are fast, relatively insensitive to noise in the training data, and therefore suitable for classifying large volumes of data [55]. Non-parametric methods, such as decision tree algorithms are not based on the assumption of normality within the training statistics, and hence can handle non-normal, non-homogenous data sets [28,56]. They can also reveal non-linear and hierarchical relationships between training variables. In decision trees, the hierarchical nature of splits, their interactions and their importance for different classes is explicit. This structure can help to understand the classification procedure [56] and allows further interpretation and analysis of the samples and their features.

According to Quinlan [54] the C5.0 algorithm produces classifiers either represented as decision trees or as rule sets. In our work, we prefer rule sets due to the often complex decision trees that are produced, compared to rule sets which are more simple and easier to understand. For example, rules are of the form "if A and B and C,…, then class X" [57,58]. For a complete description of the algorithm behind C5.0 the reader is referred to Quinlan [36,53,54], and for an overview description of decision tree classification as well as further algorithms, to DeFries and Chan [28], Friedl and Brodley [59], and Rokach and Maimon [60].

The classification algorithm C5.0 was chosen to be implemented for two reasons: it has been well documented that it provides very good results for remote sensing applications [9,28,55,56,61,62] and the automatic building of the decision tree eliminates the need for user interaction in contrast to knowledge-based decision tree approaches. Stepwise generation of the decision tree by the user utilizing gathered knowledge from the image about separation boundaries of specific class category definitions is a strong drawback of such knowledge-based decision tree approaches.

The TWOPAC framework is in general not limited to the C5.0 classification method implemented so far. New classification algorithms such as random forest, support vector machines, and maximum likelihood can be integrated as well.

2.5.3. Validation Method

Classification and feature extraction are not complete without quantitative error estimation. Within TWOPAC GUI the result of the automatic accuracy assessment is visualized as a confusion matrix or error matrix, which displays the quality of the classification result compared with the reference information [63]. Moreover, the overall accuracy, the producer's and user's accuracy, and the Kappa values are given [52,64,65]. The validation of the classifier directly after creation from the training

samples avoids unnecessary, time-consuming processing of the complete raster image or segment layer in case the classifier turns out to deliver low accuracies and needs further improvement. This is a noticeable advantage of TWOPAC in comparison to the regular procedures in most of the existing image processing software packages.

Furthermore, direct validation of the classifier facilitates effective and efficient testing of several parameters to find the best classifier. If the result of the accuracy assessment is not satisfying the training data can be modified, enlarged, *etc*., and a new selection of randomly chosen training data can be used for performing a new classification run. Additionally, the implementation of several classification runs under the modification of the training/validation percentage can provide information about the stability of the training data. Moreover, the feature space can be modified very easily with TWOPAC in order to assess the influence of several features on the classification accuracy.

2.5.4. Classification of Pre-Processed Dataset

After the classifier is tested the best variant can be applied to the entire dataset, *i.e.*, to the raster image or to the segment layer. The output is a single band raster image with class categories as pixel values in the case of pixel-based classification, or a shape file with an additional feature holding the respective class labels in the case of object-based classification. The classification result can then immediately be displayed in Quantum GIS.

2.5.5. Information Storage

The last step of the TWOPAC Quantum GIS plug-in is the standardized storage of the most important information of a classification run within a spatial database. The goal is to support consistent documentation and reusability of the classification run and to retain local expert knowledge for potential re-use with other sample data. The database manages information on a classification run in *projects*, *i.e.*, the user creates a project for every classification run to be documented. The project contains information, such as classification type (pixel- or object-based), sensor type, and acquisition date, *etc*. to describe the input data. Within a project every sample is defined by its class membership, spatial location, temporal characteristics, and feature space definition. The classification tree built up from this sample information and the confusion matrix are stored as well.

The future aim of saving training and validation data and its affiliated classifier is to accomplish comprehensive tests to detect stable samples with the most suitable features to describe a specific class allowing for stable classification. This could noticeably reduce time for repeated training data selection, at least for some classes where invariant samples can be stored. Furthermore, this could allow re-use of stable samples and their associated classifiers for spatial and temporal transfer.

2.6. Technical Implementation

The TWOPAC processing environment is installed on a virtual machine on a UNIX server environment. To allow for the automated standardized processing of TWOPAC, several process steps are implemented as Web Processing Services (WPS). The choice of WPS that facilitate standardized

interfaces for process control enables the integration of OGC conform clients, and the services can potentially be published for other users [66].

The introduction of data processing methods following the standards of the OGC allows implementation to be easily understood, reproduced, and integrated into other applications [67]. This specific technology to process geospatial raster or vector data via web resources (server, network) was introduced by the OGC in 2007. So far it was mainly used to perform simple processing steps such as intersection of vector datasets. New to the field of geospatial processing techniques is the usage of WPS for complex processes such as the implementation of image classification algorithms. The WPS interface technology used to wrap the single sub-processes of the classification procedure is available for the main steps of the TWOPAC classification process (as described in Section 2.5). Technical implementation is realized using PyWPS—the Python implementation of the WPS 1.0.0 standard [68]. The single process steps of TWOPAC—preparation of the classification, classification, and validation, are implemented in the Python language as a plug-in for Quantum GIS and the interface is available as a graphical user interface (GUI). The easy-to-handle TWOPAC GUI allows the user to implement the classification processed in a comfortable manner.

3. Results

Several datasets from different sensors and spatial resolutions were processed with the presented TWOPAC framework. Selected classifications are presented to illustrate and directly compare results regarding different aspects.

3.1. Object-Based Classification of Large Mosaic Data Set

We first present the classification of a Rapid Eye data mosaic from 2011. In general, one Rapid Eye image tile has a size of 25 × 25 km whereas the presented 4 × 4 tile mosaic covers an area of 100 × 100 km with spatial resolution of 6.5 m. As can be seen in Figure 3 the final classification map covers most of the central part of the Mekong Delta in southern Vietnam. Paddy rice cultivation is the dominant land use type in the area. Furthermore, the area is characterized by small scale distribution of cover types, such as tree orchards. The Mekong Delta undergoes rapid changes in vegetation coverage, resulting from the strongly varying precipitation due to pronounced dry and wet seasons.

Object-based classification was conducted. The segmented image mosaic contains about 500,000 objects. To characterize 17 spectral classes which could be combined into 14 thematic classes 3,188 segments were chosen as samples. For the delineation of the defined classes 21 features, such as *length to width*, *compactness*, *mean of red band*, *etc.*, were selected on the basis of experience with object-based classification. The classification of the mosaic was conducted with one C5.0 rule set that was built using the sample data combined from all 16 tiles of one acquisition date. The overall accuracy of the classification is 93.7% while Kappa is 0.93. Producer's and user's accuracies of the single classes can be found in Table 1.

Figure 3. Classification result of TWOPAC run with Rapid Eye mosaic (2010-01-27) from the central part of the Mekong Delta—subset of object-based result (**a**); detail on rural area and natural tree area (**b**); detail on urban area (**c**).
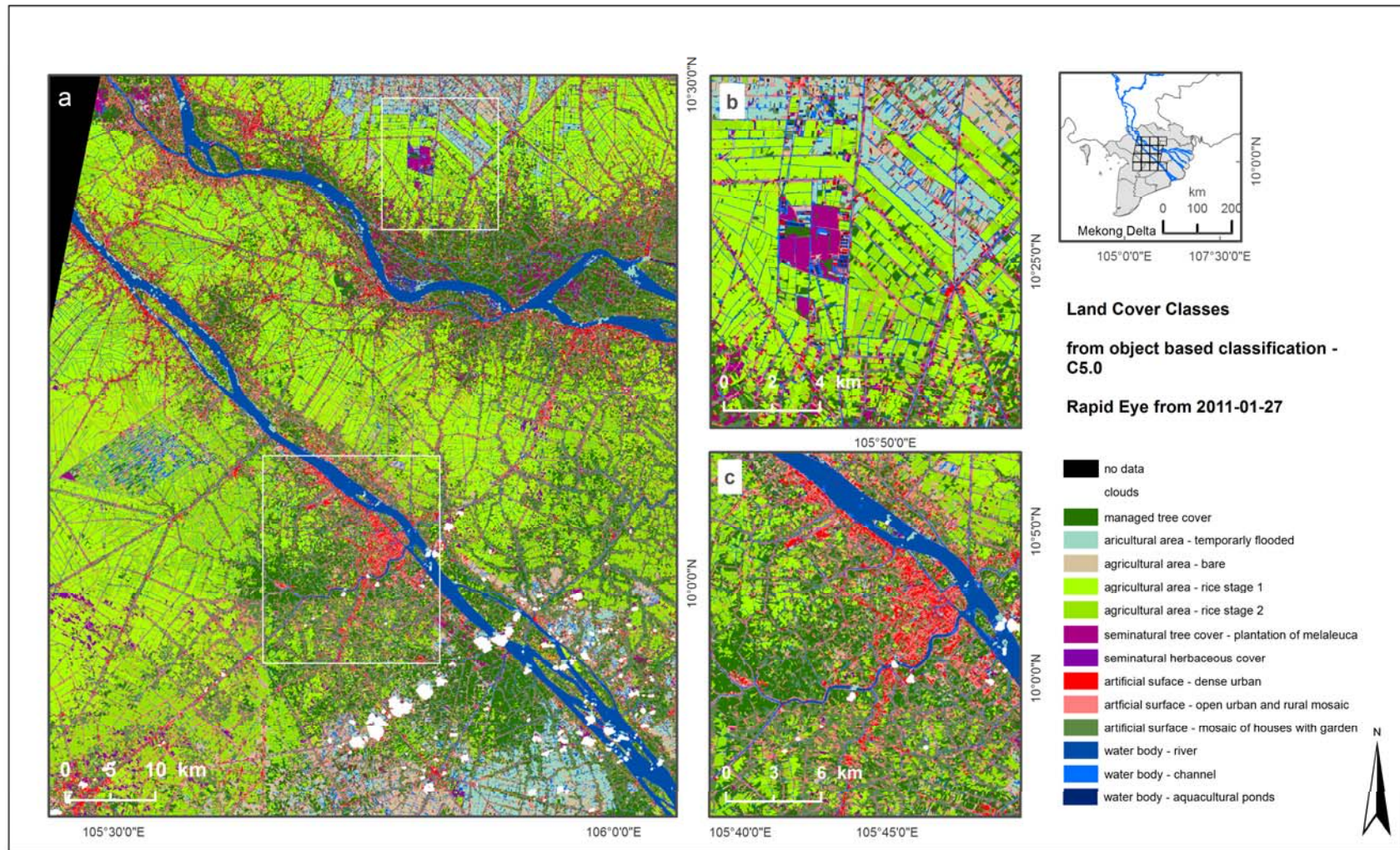
Table 1. Producer's, user's and overall accuracy and kappa coefficient for C5.0 object-based classification result of Rapid Eye-16-tile-mosaic (2010-01-27).

C5.0 Object-Based Class Name	Producer's Accuracy [%]	User's Accuracy [%]
Managed tree cover – fruit tree 1	84.00	75.00
Seminatural tree cover – melaleuca	94.74	90.00
Seminatural herbaceous cover	100.00	100.00
Artificial surface – dense urban	86.67	93.69
Artificial surface – open urban and rural mosaic	94.61	92.68
Artificial surface – open rural with garden	92.16	96.31
Artificial surface – dense urban - large buildings	97.30	95.24
Water body – river	100.00	99.37
Water body – canal	95.87	96.31
Water body – aquacultural ponds	93.94	93.94
Managed tree cover – fruit tree 2	91.80	81.16
Agricultural area – temporarily flooded 1	89.36	95.45
Agricultural area – bare	66.67	81.82
Agricultural area –rice stage 1	100.00	100.00
Agricultural area – rice stage 2	93.55	96.67
Agricultural area – temporarily flooded 1	75.00	78.95
Agricultural area – bare	90.91	75.47
Overall accuracy	93.70	
Kappa	0.93	

3.2. Comparison of Pixel- and Object-Based Classification

The second example is the classification of a SPOT5 single scene from 2008. The dataset has a spatial resolution of 10 m and covers an area of 60 × 60 km located at the west coast of the Mekong Delta. The classification was conducted with both pixel-based and object-based approaches. For the pixel-based classification 7 features including spectral information from the four bands and three spectral indices (Normalized Differenced Vegetation Index (NDVI) [69], Greenness Index (GI) [70], and Modified Triangular Vegetation Index (MTVI) [71]) were generated. For the object-based classification 20 features were chosen, mostly comparable to features chosen for the Rapid Eye scene. The map in Figure 4 shows subsets of the whole scene to allow for the visual comparison of the two approaches. Producer's and user's accuracies of the pixel- and object-based approach are presented in Table 2. For the pixel-based approach checking the class separability led to the combination of sampling data into 9 final classes resulting in an overall classification accuracy of 94.2% and Kappa value of 0.93. For the object-based approach using 12 classes achieved an overall accuracy of 90.8% and a Kappa of 0.90.

Figure 4. Classification result of TWOPAC run with SPOT5 dataset (2008-01-08) from the western coast of the Mekong Delta—subset of pixel-based result ((**a**) detail on rural area and natural tree area in (**b**); detail on urban area (**c**)); subset of object-based result ((**d**) details in (**e**,**f**)).
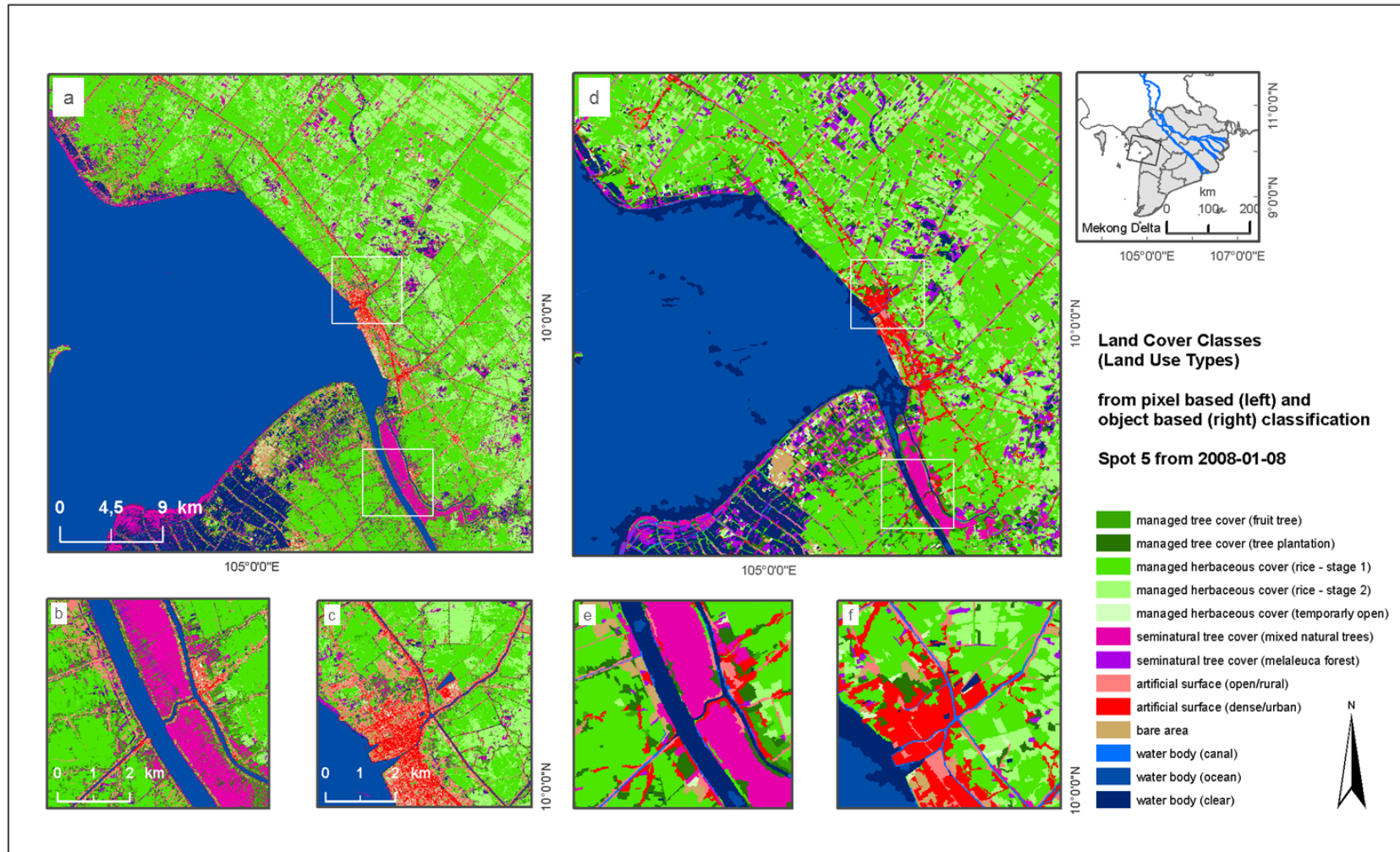
Table 2. Producer's, user's and overall accuracies and kappa coefficients for C5.0 pixel-based and object-based classification results of Spot5 image (2008-01-08).

Class Name	C5.0 Pixel-Based		C5.0 Object-Based	
	Producer's Accuracy [%]	User's Accuracy [%]	Producer's Accuracy [%]	User's Accuracy [%]
Managed tree cover – fruit tree			78.95	78.95
Managed herbaceous cover (rice stage 1)	97.00	96,46	65.38	100.00
Managed herbaceous cover (rice stage 2)	91.23	93,36	94.12	72.73
Managed herbaceous cover (rice stage 3)	86.98	89.39	70.83	58.62
Seminatural tree cover (mixed natural forest)	95.30	97.16	76.19	76.19
Seminatural tree cover (melaleuca)	96.98	93.69	88.24	83.33
Artificial surface – dense urban			87.50	100.00
Artificial surface – open urban and rural mosaic			81.82	75.00
Artificial surface	88.12	85.75		
Bare area			50.00	47.06
Water body – canal	75.60	63.97	80.00	92,31
Water body – ocean	98.55	99.15	96.30	96.30
Water body – aquacultural ponds	90.54	94.13	96.55	96.55
Overall accuracy	94.20		90.80	
Kappa	0.93		0.90	

3.3. Comparison of C5.0 TWOPAC Classification with Maximum Likelihood Classification

Thirdly, we present a classification result of a SPOT4 single scene from 2010 for the Dongting Lake region in Hunan Province of China (Figure 5). The area is characterized by continental, subtropical climatic conditions, and is dominated by agricultural fields for rice and cotton with a small scale distribution similar to the Mekong Delta. Furthermore, managed areas of poplar and reed production are located close to the lake shore as well as wetlands represented by temporarily flooded wetland vegetation and mudflats. Moreover, hilly areas west of Dongting Lake are used for cultivating fruit trees. The classification scheme comprised 12 land use classes and included poplar tree cover, citrus tree cover, reed, and areas used for agricultural production of rice and cotton, *etc*.

For independent comparison of the TWOPAC classification with other approaches commercial image processing software was used to generate a maximum likelihood (ML) classification. The classification was conducted selecting about 120 regions of interest (so-called ROIs). These samples were separated in two even parts as performed for the TWOPAC classifications, by using external GIS software. The maximum likelihood classification also consists of 12 classes.

Accuracy values can be found in Table 3. The overall accuracy of the TWOPAC C5.0 classification result derived from pixel-based approach is 97.8% whereas the ML classification results in 85.9%.

Figure 5. Classification result of TWOPAC run with SPOT 4 dataset (2010-01-02) from the Dongting lake, Hunan province, China: pixel-based result generated with C5.0 ((**a**), details in (**b**,**c**)); for comparison maximum likelihood classification ((**d**), details in (**e**,**f**)).
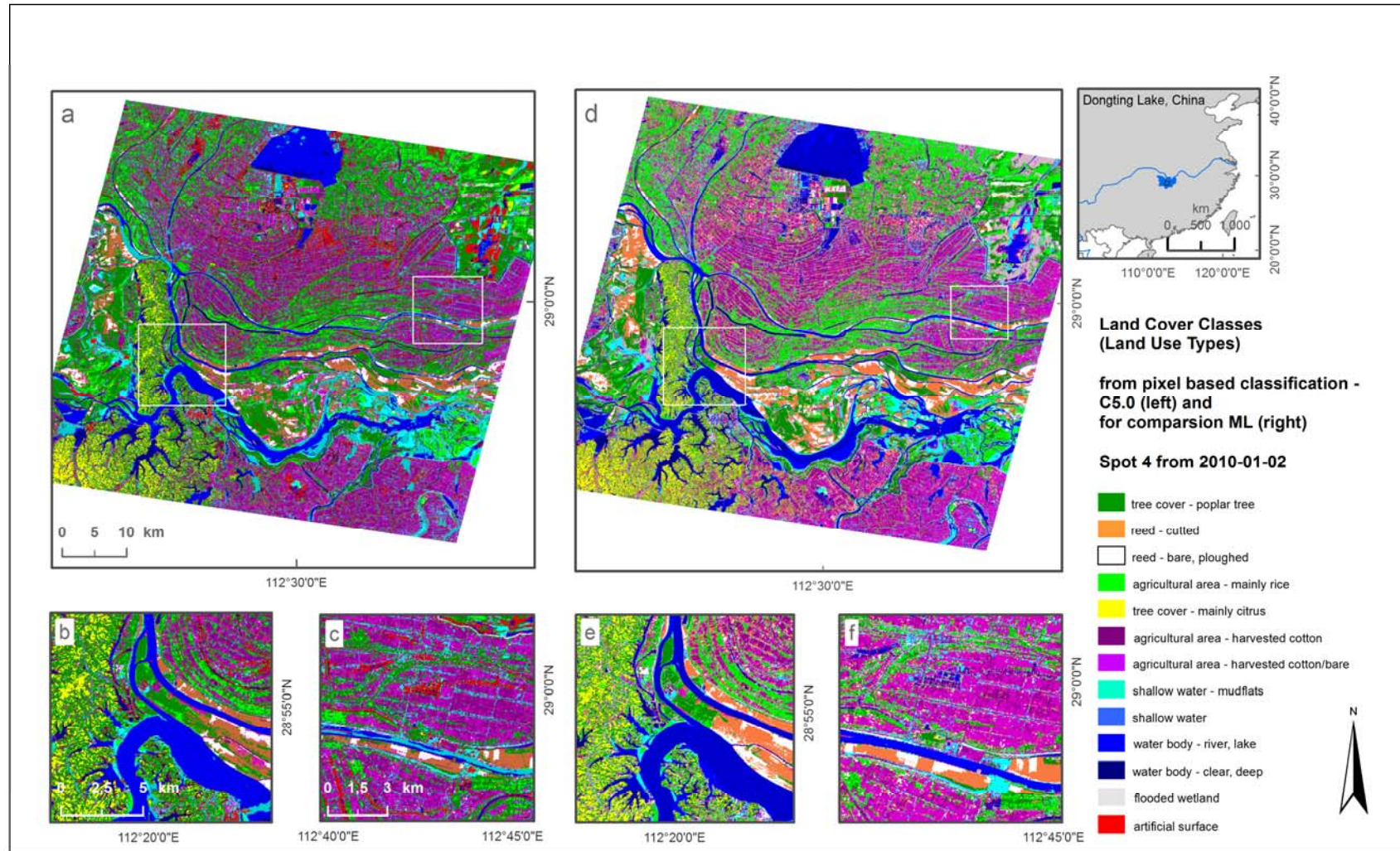
Table 3. Producer's, user's and overall accuracies and kappa coefficients for C5.0 pixel-based TWOPAC and for maximum likelihood classification results of Spot4 image (2010-12-02).

Class Name	C5.0 Pixel-Based		ML Pixel-Based	
	Producer's Accuracy [%]	User's Accuracy [%]	Producer's Accuracy [%]	User's Accuracy [%]
Agricultural area – mainly rice	98.71	97.46	45.45	100.00
Agricultural area – harvested cotton	97.19	98.58	98.04	100.00
Agricultural area – harvested cotton/bare	86.42	93.18	100	99.45
Tree cover – poplar	99.23	98.70	98.66	32.24
Reed – cut	99.59	98.68	100	100.00
Reed – bare, ploughed	98.08	95.63	100	99.82
Tree cover – mainly citrus	96.95	98.45	70.39	96.18
Artificial surface	85.96	95.62	80	99.25
Shallow water – mudflat	99.23	95.19	100	100.00
Shallow water	99.79	100.00	100	100.00
Water body – river, lake	99.7	99.70	100	83.23
Water body – clear, deep	99.7	100.00	73.91	100.00
Overall accuracy	97.80		85.87	
Kappa	0.98		0.84	

3.4. Classification of Large MODIS Time Series Data Set

Finally, we present a classification for Central Asia using MODIS time series data from 2009 with a spatial resolution of 250 m (see Figure 6) processed with TWOPAC. Pre-processing of multi-temporal data which differs substantially from multi-spectral data preparation will not be explained in this study. TWOPAC was used to classify 8 MODIS tiles for Central Asia. Feature space generation resulted in 42 feature parameters, such as maximum and minimum of NDVI in summer, to describe the seasonal phenology of different vegetation types. Sample data were selected with the help of 20 Landsat image datasets which had been segmented as described before. Twelve class categories were selected which exist in Central Asia comprising an area of 7,920,000 km². The single MODIS tiles were classified using the same C5.0 rule set generated from all training samples. The overall accuracy obtained with this TWOPAC classification is 91.2% and the Kappa coefficient is 0.89. Producer's accuracies range from 60.5% for close shrubs to 100% for ice and snow, whereas user's accuracies resulted in values between 41.1% for closed shrubs and 96.1% for bare area with salt. Details on this classification related to pre-processing, confusion matrices, *etc*., can be found in [72].

Figure 6. Classification result of TWOPAC run with multi-temporal MODIS 8 tiles mosaic (2009) for Central Asia: Overview of the classification result for Uzbekistan (**a**); details on specific land cover types—cultivated area and natural vegetation (**b**,**c**).
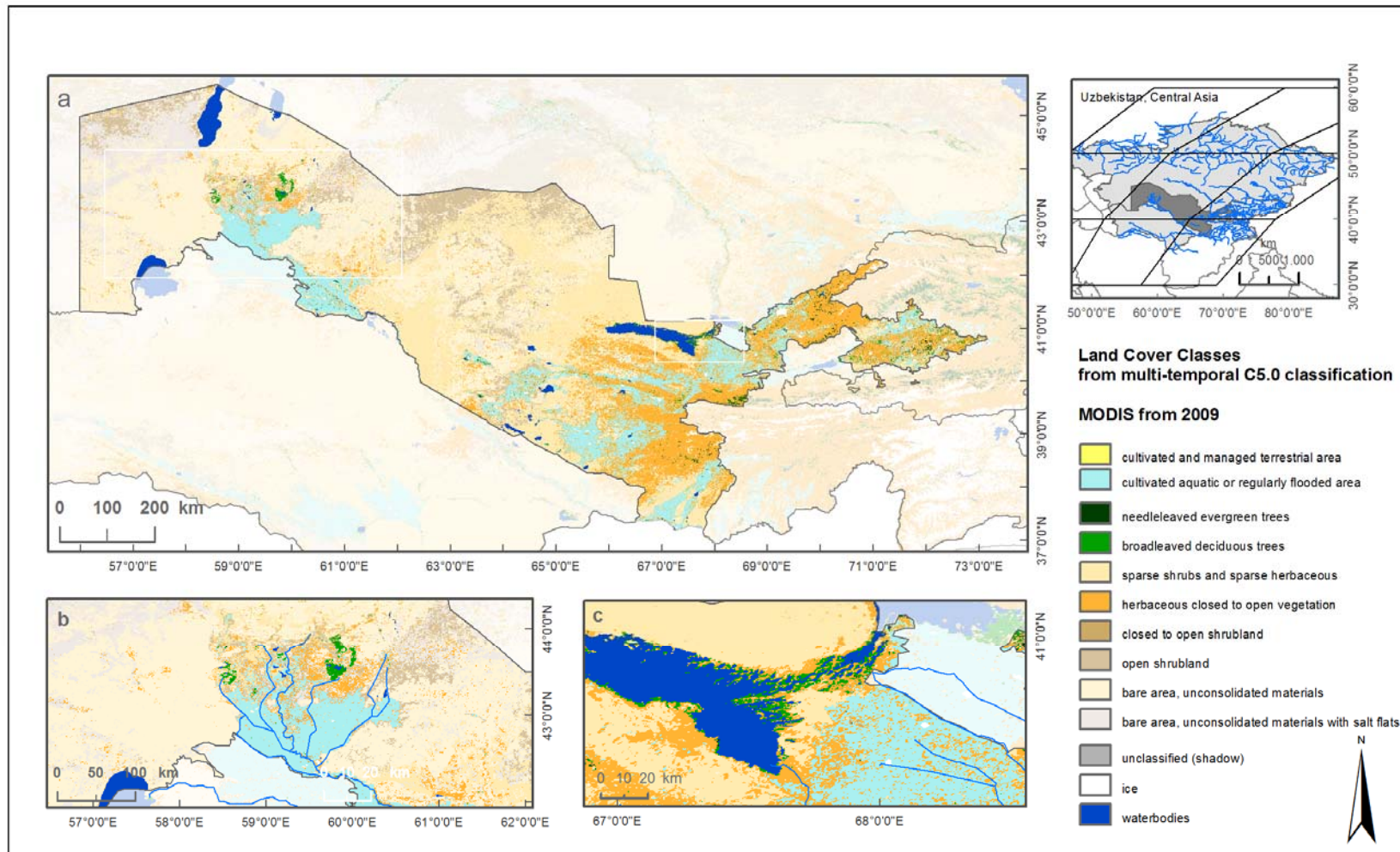
4. Discussion

Four different classification examples were presented to underline the ability of TWOPAC for handling different data set types and sizes, spatial resolutions, and classification approaches (pixel- and object-based). In general, TWOPAC classification results show very high accuracies.

The object-based classification of a large Rapid Eye mosaic results in an overall accuracy of 93.7%. The range of the producer's accuracy is from 66.67% for bare agricultural area to 100% for certain water bodies and vegetated agricultural areas, while the user's accuracies range from 75.0% for certain managed tree cover classes to 100% for certain vegetated agricultural areas and semi-natural herbaceous cover classes (see Table 1). The lower accuracies mark classes which are prone to misclassifications due to high similarities in their spectral class characteristics, e.g., two different tree cover classes, or classes which represent bare surfaces.

The comparison shows that the classifications also result in very good accuracy of 94.2% when using the object-based and of 90.8% when using the pixel-based approach. Although the overall accuracy of the pixel-based classification is unexpectedly higher than for the object-based approach, the results do not differ substantially. The difference could be caused by the significant smaller number of samples for the object-based classification. The comparison of the two TWOPAC classification approaches demonstrate that both object- as well as pixel-based classification of the same dataset can result in high and comparable accuracies as already reported by Duro *et al.* [73]. To underline this, the artificial surface class with a producer's accuracy of 88.1% (pixel-based) is separated into two subclasses using specific object-based characteristics for distinction of dense urban and sparse urban areas. The object-based approach results in a producer's accuracy of 87.5% for dense urban and 81.8% for sparse urban (see Table 2).

For the comparison of a decision-tree-based TWOPAC classification with a maximum likelihood classification based on existing image classification software, a single scene classification was implemented. As can be seen from Table 3, the two classification outputs show significant accuracy differences. Misclassification amongst vegetation classes, as well as amongst water classes are much lower for the non-parametric classification algorithm used in TWOPAC. This could be caused by the insensitivity of decision tree algorithms towards noise in the training data already documented by Friedl *et al.* [55].

The classification of MODIS time series data is presented to underline TWOPAC's capacity to classify large datasets. Creation of one classifier from the complete sample data set was not feasible using common classification systems. The MODIS data set for instance, presented in the section above, required 15 GB storage space after mosaicking and feature creation. Tests had to be interrupted due to the impossibility of reading data larger than 4 GB into a common software system. Hence, the MODIS classification only could have been implemented separately for 4 single MODIS tiles of maximum 4 GB marking the limit of commonly used image classification software. It is anticipated that the overall result would be different since not all envisaged classes occur in every single MODIS tile. Furthermore, the effort for applying the classification process separately to every MODIS tile (including pre-processing, classifier creation, and results harmonization) would be higher than using the TWOPAC approach.

The same applies to the presented object-based Rapid Eye classification. For segmentation, feature creation, and sample selection the mosaic with a dataset size of more than 4 gigabytes had to be split in 4 subsets due to memory problems of the software used. For the large, high resolution dataset mosaic covering 10,000 km² an object-based TWOPAC classification is implemented at once. In addition, outsourcing the processing of large datasets to the TWOPAC server environment is advantageous compared to usage of processing capacity on desktop computers.

5. Conclusion and Outlook

The TWOPAC framework presented in this paper shows great potential for large scale data processing. As proposed in various publications [6,8,9,28], there is a high demand for automation of image classification processes due to, *i.e.*, increasing volumes of accessible data, and the need to process reliable results in a shorter time. As presented in this paper, classification with TWOPAC allows for the fast generation of comparable and standardized results, utilizing server processing capabilities. This is especially advantageous for calculation of memory-intensive data (e.g., mosaics of high resolution data, time series data sets). The TWOPAC classification framework has already been used for several applications and projects in different regions of the world which resulted in overall accuracies of between 85.9% and 97.8%.

TWOPAC features open source-based implementation of the well-known decision tree estimation software C5.0, for automatic building of decision trees from training data. This is considered as a remarkable advantage compared to known commercial image processing software.

Moreover, an expanding community of scientific and non-scientific users, who are increasingly interested in "easily" generated LC and LU products, are given the opportunity to perform image classification in a simplified way. Besides *a priori* local knowledge of the land cover types of the selected geographical region, which is a basic requirement, the user should be aware of necessary pre-processing for data input generation. Additionally, the user needs to have some experience in image interpretation for conducting the sampling procedure. However, the envisaged user of TWOPAC should also have some basic expertise in remote sensing image analysis for evaluation of the classification results, and for understanding the validation tables. In summary, TWOPAC is a very easy-to-use classification system which only requires well-founded basic knowledge of remote sensing image classification.

A further key advantage of the proposed approach is that it is almost fully independent of commercial clients or desktop software for remote sensing image analysis. The utilization of OGC compliant Web Processing Services offers the technical opportunity to integrate the framework into existing processing environments. To fully exploit the possibilities of TWOPAC on an open source basis, the segmentation and spatial feature extraction currently achieved by using eCognition Developer has to be replaced by a freely available segmentation algorithm.

Further enhancements and extensions are focusing on the successive integration of individual pre-processing steps such as geo-rectification and atmospheric correction. Here, we envisage connecting the TWOPAC framework to the already existing automatic pre-processing environment, CATENA (Latin: chain), which has been developed by the Remote Sensing Technology Institute of

the German Aerospace Centre (DLR) [74] and enables the fully automatic geo-rectification and atmospheric correction of optical imagery.

Furthermore, the sampling procedure shall undergo modification, as it is the most time-consuming part of the classification procedure and still includes subjective interpretation and user interaction. Predefinition of classes requires a lot of knowledge of the various classes and their appearance within the satellite image data. Subjective evaluation of class appearance leads to the definition of classes that might not be separable by the classifier. Therefore, supported sampling could include statistical clustering over the feature space and the subsequent selection of samples on the basis of this pre-classification result. Additionally, we envisage improvements in spatial and temporal transferability. Multi-temporal data stacks could be analyzed for pseudo-invariant objects to provide stable sample objects over time. These could easily be found in water bodies, built-up surfaces, areas with evergreen trees, or sandy shore areas. A statistical analysis on the stored sample data regarding sample quality and the re-usability of samples would utilize the comprehensive sampling database approach at the envisaged scale. Sample quality could even be flagged and proposed for reuse.

Finally, the implementation of further sophisticated classification algorithms, such as the random forest classifier [75] or support vector machines [76], would provide the user greater flexibility in their specific classification exercises.

The ultimate goal is then to make TWOPAC available to a larger user community, once the framework is fully based on open source technology.

Acknowledgements

References

1. Brown, D.G.; Duh, J.-D. Spatial simulation for translating from land use to land cover. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 35–60.
2. Campbell, J.B. *Introduction to Remote Sensing*, 4th ed.; Taylor & Francis: London, UK, 2006; p. 626.
3. Cihlar, J.; Jansen, L.J.M. From land cover to land use: A methodology for efficient land use mapping over large areas. *Prof. Geogr.* **2001**, *53*, 275–289.
4. Herold, M. *Assessment of the Status of the Development of the Standards for the Terrestrial Essential Climate Variables. Land. Land Cover*; FAO: Rome, Italy, 2009.
5. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2007.
6. Cihlar, J. Land cover mapping of large areas from satellites: Status and research priorities. *Int. J. Remote Sens.* **2000**, *21*, 1093–1114.
7. Franklin, S.E.; Wulder, M.A. Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. *Progr. Phys. Geogr.* **2002**, *26*, 173–205.

8. Knorn, J.; Rabe, A.; Radeloff, V.C.; Kuemmerle, T.; Kozak, J.; Hostert, P. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sens. Environ.* **2009**, *113*, 957–964.

9. Rogan, J.; Franklin, J.; Stow, D.; Miller, J.; Woodcock, C.; Roberts, D. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sens. Environ.* **2008**, *112*, 2272–2283.

10. Schaaf, C.B.; Barry, R.G.; Brady, M.; Brown, J.; CEOS; Christiansen, H.H.; Cihlar, J.; Clow, G.; Csiszar, I.; Dolman, H.; *et al. Terrestrial Essential Climate Variables: Climate Change Assessment, Mitigation and Adaptation*; FAO: Rome, Italy, 2008.

11. Apel, H.; Thieken, A.H.; Merz, B.; Blöschl, G. Flood risk assessment and associated uncertainties. *Natural Hazards Earth Syst. Sci.* **2004**, *4*, 295–308.

12. Bates, P.D.; Horritt, M.S.; Smith, C.N.; Mason, D. Integrating remote sensing observations of flood hydrology and hydraulic modelling. *Hydrol. Process.* **1997**, *11*, 1777–1795.

13. Delgado, J.M.; Apel, H.; Merz, B. Flood trends and variability in the Mekong river. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 407–418.

14. Matgen, P.; Schumann, G.; Henry, J.B.; Hoffmann, L.; Pfister, L. Integration of SAR-derived river inundation areas, high-precision topographic data and a river flow model toward near real-time flood management. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 247–263.

15. Schumann, G.; Di Baldassarre, G.; Bates, P.D. The utility of spaceborne radar to render flood inundation maps based on multialgorithm ensembles. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2801–2807.

16. Eisfelder, C.; Kuenzer, C.; Dech, S. Derivation of biomass information for semi-arid areas using remote-sensing data. *Int. J. Remote Sens.* **2011**, *33*, 2937–2984.

17. Goetz, S.; Baccini, A.; Laporte, N.T.; Johns, T.; Walker, W.; Kellndorfer, J.; Houghton, R.; Sun, M. Mapping and monitoring carbon stocks with satellite observations: A comparison of methods. *Carbon Balance Manag.* **2009**, *4*, doi:10.1186/1750-0680-4-2.

18. GTOS. *Assessment of the Status of the Development of the Standards for the Terrestrial Essential Climate Variables*; GTOS 67; GTOS: Rome, Italy, 2009.

19. Guenther, K.P.; Borg, E.; Wisskirchen, K.; Schroedter-Homscheidt, M.; Fichtelmann, B. Remote Sensing and Modelling: A Tool to Provide the Spatial Information for Biomass Production Potential. In *Proceedings of 2nd International Baltic Bioenergy Conference IBBC 2006*, Stralsund, Germany, 2–4 November 2006; pp. 33–45.

20. Lu, D. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* **2006**, *27*, 1297–1328.

21. Dobson, J.E.; Bright, E.A.; Coleman, P.R.; Bhaduri, B.L. *LandScan: A Global Population Database for Estimating Populations at Risk*; Taylor & Francis: London, UK, 2003; pp. 267–281.

22. Sutton, P.; Roberts, D.; Elvidge, C.; Baugh, K. Census from Heaven: An estimate of the global human population using night-time satellite imagery. *Int. J. Remote Sens.* **2001**, *22*, 3061–3076.

23. Glade, T.; Birkmann, J. Assessment of different dimensions of vulnerability to natural hazards and climate change. *Natural Hazards Earth Syst. Sci.* **2009**, *9*, 315–1494.

24. Jacquin, A.; Misakova, L.; Gay, M. A hybrid object-based classification approach for mapping urban sprawl in periurban environment. *Landsc. Urban Plan.* **2008**, *84*, 152–165.

25. Wisner, B.; Blaikie, P.; Cannon, T.; Davies, I. *At Risk*; 2nd ed.; Routlegde: London, UK, 2004; p. 471.

26. Kuenzer, C.; Bluemel, A.; Gebhardt, S.; Vo Quoc, T.; Dech, S. Remote sensing of mangrove ecosystems: A review. *Remote Sens.* **2011**, *3*, 878–928.

27. Vo Quoc, T.; Kuenzer, C.; Vo Quang, M.; Oppelt, N. The concept of ecosystem services: A review. *Ecol. Indic.* **2012**, *23*, 431–446.

28. DeFries, R.S.; Chan, J.C.-W. Multiple Criteria for Evaluating Machine Learning Algorithms for Land Cover Classification from Satellite Data. *Remote Sens. Environ.* **2000**, *74*, 503–515.

29. Coppin, P.; Jonckheere, I.; Nackaerts, K.; Muys, B.; Lambin, E. Review ArticleDigital change detection methods in ecosystem monitoring: a review. *Int. J. Remote Sens.* **2004**, *25*, 1565–1596.

30. Dekker, R.J.; Kuenzer, C.; Lacroix, V.; Lehner, M.; Niemeyer, I.; Nussbaum, S.; Schoepfer, E.; Sequeira, V.; Stringa, E. Change Detection Tools. In *Remote Sensing from Space: Supporting International Peace and Security*; Jasani, B., Pesaresi, M., Schneiderbauer, S., Zeug, G., Eds.; Springer: Dordrecht, The Netherlands, 2009; Volume VI, pp. 119–140.

31. Kuenzer, C.; Bartalis, Z.; Schmidt, M.; Wagner, W. Trend analyses of a 15 year global soil moisture time series derived from ERS-1/-2 and METOP scatterometer data—Floods, droughts and long term changes. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, Part B7, 1363–1368.

32. Lu, D.; Mausel, P.; Brondazio, E.; Moran, E. Change detection techniques. *Int. J. Remote Sens.* **2004**, *25*, 2365–2401.

33. Rogan, J.; Chen, D. Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progr. Plan.* **2004**, *61*, 301–325.

34. Lillesand, T.M.; Kiefer, R.W.; Chipman, J.W. *Remote Sensing and Image Interpretation*, 6th ed.; John Wiley & Sons, Inc.: Chichester, UK, 2008; p. 756.

35. Richards, J.A.; Jia, X. *Remote Sensing Digital Image Analysis. An Introduction*, 3rd ed.; Springer: Berlin, Germany, 1999; p. 363.

36. Quinlan, J.R. *C4.5: Programs for Machine Learning*; Morgan Kaufmann: San Mateo, CA, USA, 1993.

37. Roth, A.; Huber, M.; Kosmann, D. Geocoding of TerraSAR-X Data. In *Proceedings of 20th Congress of ISPRS*, Istanbul, Turkey, 12–23 July 2004.

38. Richter, R. A spatially adaptive fast atmospheric correction algorithm. *Int. J. Remote Sens.* **1996**, *17*, 1201–1214.

39. Blaschke, T. A Framework for Change Detection Based on Image Objects. In *Remote Sensing & GIS for Environmental Studies: Applications in Geography*; Erasmi, S., Cyffka, B., Kappas, M., Eds.; Göttinger Geographische Abhandlungen: Göttingen, Germany, 2005; Volume 113, pp. 1–9.

40. Möller, M.; Lymburner, L.; Volk, M. The comparison index: A tool for assessing the accuracy of image segmentation. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 311–321.

41. Neubert, M.; Herold, H. Assessment of Remote Sensing Image Segmentation Quality. In *Proceedings of GEOBIA 2008, Pixels, Objects, Intelligence–GEOgraphic Object Based Image Analysis for the 21st Century*, Calgary, AB, Canada, 5–8 August 2008.

42. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm.* **2010**, *65*, 2–16.

43. Haralick, R.M.; Shapiro, L.G. Image segmentation techniques. *Comput. Vis. Graph. Image Process.* **1985**, *29*, 100–132.

44. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A review of vegetation indices. *Remote Sens. Rev.* **1995**, *13*, 95–120.

45. Crist, E.P.; Cicone, R.C. A physically-based transformation of thematic mapper data—The TM Tasseled Cap. *IEEE Trans. Geosci. Remote Sens.* **1984**, *22*, 256–263

46. Kauth, R.J.; Thomas, G.S. The tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. In *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, Lafayette, IN, USA, 29 June–1 July 1976; pp. 4B-41–44B-51.

47. Schowengerdt, R.A. *Remote Sensing. Models and Methods for Image Processing*, 3 ed.; Elsevier Academic Press: Amsterdam, The Netherlands, 2007; p. 515.

48. Singh, A.; Harrison, A. Standardized principal components. *Int. J. Remote Sens.* **1985**, *6*, 883–896.

49. Tso, B.; Mather, P.M. *Classification Methods for Remotely Sensed Data*, 2nd ed.; Taylor & Francis Ltd: Abingdon, UK, 2001; p. 356.

50. Quantum-GIS. Available online: http://qgis.osgeo.org (accessed on 15 March 2012).

51. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182.

52. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data. Principles and Practices*, 2 ed.; CRC Press: Boca Raton, FL, USA, 2007; p. 183.

53. Quinlan, J.R. Introduction of decision trees. *Machine Learn.* **1986**, *1*, 81–106.

54. Quinlan, J.R. Improved use of continuous attributes in C4.5. *J. Artif. Intell. Res.* **1996**, *4*, 77–90.

55. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409.

56. Hansen, M.; Dubayah, R.; DeFries, R. Classification trees: An alternative to traditional land cover classifiers. *Int. J. Remote Sens.* **1996**, *17*, 1075–1081.

57. Quinlan, J.R., Generating Production Rules from Decision Trees. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*; Morgan Kaufmann Publishers Inc.: Milan, Italy, 1987; Volume 1, pp 304–307.

58. Wu, X.D.; Kumar, V.; Quinlan, J.R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Yu, P.S.; *et al*. Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **2008**, *14*, 1–37.

59. Friedl, M.A.; Brodley, C.E.; Strahler, A.H. Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 969–977.

60. Rokach, L.; Maimon, O.Z. *Data Mining with Decision Trees: Theroy and Applications*; World Scientific Pub. Co. Inc.: Hackensack, NJ, USA, 2008; p. 243.

61. Kamber, M.; Winstone, L.; Gong, W.; Cheng, S.; Han, J.W. Generalization and Decision Tree Induction: Efficient Classification in Data Mining. In *Proceedings of Seventh International Workshop on Research Issues in Data Engineering*, Birmingham, UK, 7–8 April 1997; pp. 111–120.

62. Pal, M.; Mather, P.M. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens. Environ.* **2003**, *86*, 554–565.

63. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46.

64. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psych. Meas.* **1960**, *20*, 37–46.

65. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201.

66. Schut, P. OpenGIS Standards—Web Processing Service (WPS). Available online: http://www.opengeospatial.org/standards/wps (accessed on 15 March 2012).

67. Bielski, C.; Gentilini, S.; Pappalardo, M. Post-disaster image processing for damage analysis using GENESI-DR, WPS and Grid computing. *Remote Sens.* **2011**, *3*, 1234–1250.

68. Cepicky, J.; Becchi geospatial processing via internet on remote servers—PyWPS. *OSGeo J.* **2007**, *1*, 39–42.

69. Rouse, J.W.; Haas, R.H.; Schell, J.A.; and, D.W.D.; Harlan, J.C. *Monitoring the Vernal Advancements and Retrogradation of Natural Vegetation*; NASA/GSFC: Greenbelt, MD, USA, 1974; p. 371.

70. Zarco-Tejada, P.J.; Berjón, A.; López-Lozano, R.; Miller, J.R.; Martín, P.; Cachorro, V.; González, M.R.; de Frutos, A. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sens. Environ.* **2005**, *99*, 271–287.

71. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352.

72. Klein, I.; Gessner, U.; Kuenzer, C. Regional land cover mapping in Central Asia using MODIS time-series. *Appl. Geogr.* **2012**, doi: 10.1016/j.apgeog.2012.06016.

73. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272.

74. Reinartz, P. The CATENA Processing Chain—Multi-Sensor Pre-processing: Orthorectification, Atmospheric Correction, Future Aspects. In *Proceedings of Geoland Forum_6*, Toulouse, France, 24–25 March 2010.

75. Breiman, L. Random Forests. *Machine Learn.* **2001**, *45*, 5–32.

76. Vapnik, V.N.; Chervonenkis, A.Y. *Teoriya Raspoznavaniya Obrazov: Statisticheskie Problemy Obucheniya*; Izdat. Nauka, FizMatLit.: Moscow, Russia, 1974; p. 416.