Using Remote Sensing derived Spatial Metrics for the Calibration of Land-Use Change Models

Johannes van der Kwast,
Inge Uljee
and Guy Engelen
Unit Environmental modeling
Flemish Institute for Technological Research (VITO)
Mol, Belgium
Email: hans.vanderkwast@vito.be

Tim Van de Voorde
and Frank Canters
Cartography and GIS research Group
Department of Geography
Vrije Universiteit Brussel
Brussels, Belgium

Carlo Lavalle
Joint Research Centre (JRC)
European Commission
Ispra, Italy

***Abstract*— More than ever before, planners and policy makers need tools to anticipate and assess the impact of their decisions on the spatial system that they are to manage. A growing number of high resolution models is currently being developed for this purpose. The calibration of these models remains a major challenge. Typically the required time series of land-use maps based on identical and consistent mapping methodologies, legends and scales are missing. The availability of images from earth observation satellites is much larger. However, conventional remote sensing based land-use classifications result in land cover maps, based on reflective properties of the surface, rather than land-use maps representing the functional classes needed for urban land-use change modeling. Recently, landscape metrics or spatial metrics have been introduced in the field of urban land-use mapping and modeling to characterize the spatial dynamics of such systems. The question raised in the study presented is whether spatial metrics directly applied to remote sensing images can be used to calibrate and validate land-use models of urban systems. The underlying hypothesis is that a methodology can be developed which enables to calculate metrics on both the remote sensing image and the predicted land-use map, which quantify the same distinguishing spatial structures at some level of abstraction. The study demonstrates the potential of spatial metrics to simplify and speed up the calibration procedures in so far that the development of land-use maps could be avoided.**

I. Introduction

An increasing number of dynamic land-use change models is currently being developed in order to provide tools for planners and policy makers who need them to anticipate and assess the impact of their decisions on the spatial system that they are to manage [1]. For a proper calibration of these models typically a time series of land-use maps is needed. Often, however, time series of land-use maps are lacking. Even when time series exist, inconsistencies in mapping methodologies, legends and scales result in measured land-use changes that are caused by mismatches in the mapping procedures rather than an indication of real changes in the land-use patterns of interest.

Since several decades, earth observation satellites provide images of the earth's surface. The long time series of medium resolution sensors like Landsat TM/ETM+ and SPOT HRV offer interesting possibilities of increasing the data to be used for the historic calibration of land-use change models. Compared to land-use maps, the potential temporal availability of medium resolution remote sensing images is relatively high and depends on their ground track repeat cycle which is typically between 1 and 16 days. In addition, the spatial and temporal consistency of a time series of remote sensing images is better than for land-use maps. However, the spatio-temporal consistency and availability heavily depend on atmospheric conditions. Land-use change models typically predict land-use in time steps of 1 year and a spatial resolution of 50-500 meters, making the temporal and spatial resolution of medium resolution remote sensing images sufficient for their calibration.

Remote sensing based land-use classification algorithms can easily derive reflective properties of the earth's surface, resulting in land cover maps. For the purpose of land-use change modeling land-use classes are needed, which are linked to socio-economic activities and, as such, cannot be directly inferred from spectral information only [2]. Previous studies, however, have demonstrated a strong relationship between the spatial structure of the built-up environment and its functional characteristics [3]. The link between land-use and urban form is a key element in visual interpretation of remotely sensed imagery of urban areas. It also led to the development of different (semi-)automated approaches for urban mapping that make use of structural and contextual information present in the image, or derived from ancillary data sources [4].

A rather novel approach to describe urban form and structure is by means of spatial metrics. Spatial metrics describe various properties of the spatial heterogeneity and configuration of land cover in a given area. Originally developed for landscape ecological research, they have recently been shown to have considerable potential for the analysis of urban environments [5]–[8].

978-1-4244-3461-9/09/$25.00

Research associating spatial metrics and changes in these metrics to urban change processes is still at an early stage, yet the potential of linking changes in spatial metrics to specific processes of urban development has already been demonstrated [6], [7].

This research proposes a framework for the historic calibration of land-use change models using remote sensing data. The framework includes three steps: (1) Pixel-based classification of medium resolution remote sensing images; (2) contextual re-classification based on spatial metrics using the Optimized Spatial Reclassification Kernel (OSPARK) algorithm, and (3) Calibration of a land-use change model, based on spatial metrics calculated at different levels of spatial abstraction. The performance of the OSPARK algorithm will be tested for different scenarios of pixel-based classifications that are used as an input to the algorithm.

The proposed framework has been developed for the calibration of the MOLAND (Monitoring Land Use/Cover Dynamics) land-use change model [9] applied to Dublin. The model will be introduced in the next section, followed by an explanation of the concept of the calibration framework, the OSPARK algorithm and the different pixel-based classification scenarios that are tested. The scenario with the best performance will be used for the calibration of the MOLAND model in future studies.

II. Methods

A. The MOLAND Land-Use Change Model

The calibration framework proposed in this study has been developed for the MOLAND (Monitoring Land Use/Cover Dynamics) model [9], applied to Dublin (Ireland). The core of this urban and regional growth model is a constraint cellular automata [10] that models the spatial dynamics of the system. The model input consists of five GIS datasets for the geographical area of interest: (a) actual land use types; (b) accessibility of the area to the transport network; (c) inherent suitability of the area for different land-uses; (d) zoning status (i.e. legal constraints) of the area for different land-uses; (e) socio-economic characteristics (e.g. population, income, production, employment) of the area. The model explores the likely future development of land-use over the next thirty years in time-steps of one year for each grid cell of 4 ha, given alternative planning and policy scenarios and socio-economic trends.

The calibration of any model with the level of complexity of the MOLAND model is not trivial and requires time and effort. The main reason is that in principle every model grid cell represents at least one state variable in the model. The task of the calibration is to ensure that the model behaves in a realistic manner and is able of generating existing spatial patterns. The calibration of the MOLAND model [9] is a heuristic procedure based on trial-and-error. It requires a reference (or 'training' or 'historical') land-use map, from which the actual map (i.e. the most recent available map) is reconstructed. The comparison between the reconstructed and the actual map is performed by means of dedicated goodness-of-fit measures. These consist of a number of statistical indicators such as mean patch area, shape index and proximity index, Simpson's diversity index, Kappa Statistic, the Kappa Histo, the Kappa Location and the Fuzzy Kappa [11]. In addition to the calculation of statistics, the comparison is performed by means of GIS procedures.

The calibration consists of four steps: (1) a first set of approximate values of the parameters describing the neighbor influence function (attraction-repulsion parameters) is fixed. Generally, parameters are coming from previous applications of the model; (2) a second parameter (the stochastic parameter α) is then fixed. α determines largely the scatterness of the land-use patterns and size of clusters; (3) further information such as maps of suitability, accessibility and zoning are then introduced; (4) then the fine tuning of the model starts (repeating of loop 1-2-3 with dedicated statistical analysis). The steps are repeated until the reconstructed map satisfactorily matches with the actual map.

For a successful calibration, the calibration period should be sufficiently long in order to give the underlying processes in the system enough time to manifest themselves in a representative way. The poor availability of high quality and comparable land-use maps often constrains the choice of the calibration period. In the Dublin application of the MOLAND model, used in this study, the required maps are only available for 1990 and 2000, constraining the calibration to this period. It is hypothesized that the proposed remote sensing based calibration framework, explained next, can increase the amount data for the historic calibration, resulting in better predictions of the land-use change model.

Table I gives the land-uses classes in the MOLAND model and their surface area in 1990 and 2000 based on the land-use maps.

B. Concept of the Calibration Framework

Figure 1 shows the concept of using spatial metrics for the calibration of land-use change models that is proposed in this study. The procedure can be applied each time a remote sensing image is available within the model calibration period (Fig. 2). The remote sensing image is first classified using pixel-based techniques. Next, spatial metrics are used to identify the land-use classes. Existing land-use maps are used for calibrating the metrics-based classification. Finally, the remote sensing derived land-use map can be compared with the output of the land-use change model for the same year. Depending on the objective of the land-use change model or the stage in its development (calibration or application) the modeler has the choice to optimize: (1) the overall pattern in model predictions, based on landscape-level metrics; (2) patterns within land-use classes, based on class-level metrics and (3) patterns at the moving window-level or pixel level.

The spatial metrics used in the calibration of land-use change models should be chosen in accordance with the objectives of the model. This study will focus on the evolution of urban sprawl. For urban sprawl the metrics should quantify density characteristics, land-use and activity characteristics, fragmentation and scattering, decentralization characteristics

TABLE I

LAND-USE CLASSES OF THE MOLAND MODEL AND THEIR SURFACE AREA (NR. OF CELLS OF 4 HA) IN THE LAND-USE MAPS OF 1990 AND 2000. RECODE INDICATES THE RECLASSIFICATION IN MERGED CLASSES (SEE TABLE VI

Class	**Area 1990**	**Area 2000**	**Recode**
Arable land	4248	6736	1
Pastures	20082	15952	1
Heterogeneous agricultural areas	0	115	1
Forests	1213	1295	2
Semi-natural areas	2247	2219	2
Wetlands	102	113	2
Abandoned	0	0	3
Residential continuous dense urban fabric	28	29	3
Residential continuous medium dense urban fabric	24	29	3
Residential discontinuous urban fabric	56	1057	3
Residential discontinuous sparse urban fabric	4078	3774	3
Industrial areas fabric	662	900	4
Commercial areas	205	236	4
Public and private services	546	597	4
Port areas	23	25	4
Construction sites	147	193	5
Road and rail networks and associated land	92	135	4
Airport	219	213	6
Mineral extraction sites	118	94	5
Dump sites	68	97	5
Artificial non-agricultural vegetated areas	2140	2451	7
Restricted access areas	10	8	8
Water bodies	5692	5732	8

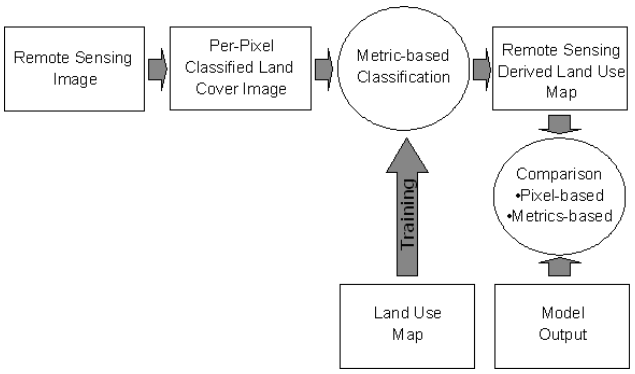


Fig. 1. Concept of using remote sensing derived metrics in the calibration of land-use change models

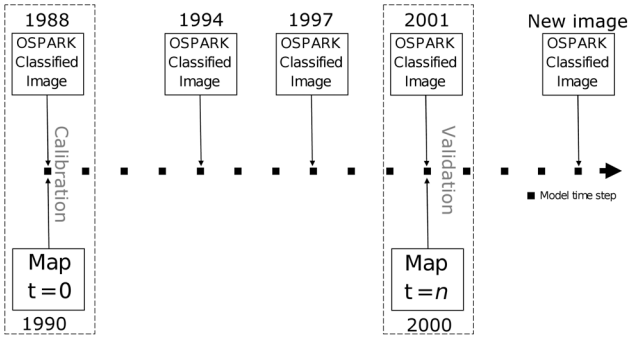


Fig. 2. Application of the Calibration Framework

and accessibility characteristics [6]. In this study some of the spatial metrics suggested by Torrens [6] are used. For the equations of the metrics it is referred to the publication of Torrens [6]. The Shannon's Diversity Index (SHDI) is used to quantify land-use and activity characteristics. The SHDI value reaches zero when there is no diversity at all and increases with increasing diversity. For fragmentation and scattering the Perimeter-Area Fractal dimension (PAFRAC), Contagion (CONTAG) and the Interspersion and Juxtaposition Index (IJI) are used. The PAFRAC metric will result in a value of 1 when the land-use is unfilled and a value of 2 when the urban landscape is compact. CONTAG is used to calculate the spatial configuration and structure of activities, it is zero for landscapes without fragmentation and 100 for regions that have a maximum fragmentation. IJI is similar, but for patches in stead of pixels. The metrics are calculated at landscape-level and moving window-level and the results are compared for the land-use maps and remote sensing images.

The proposed framework is calibrated and validated using the land-use maps of 1990 and 2000 [9] and the Landsat TM5 image of June 13 1988 and the Landsat ETM7 image of

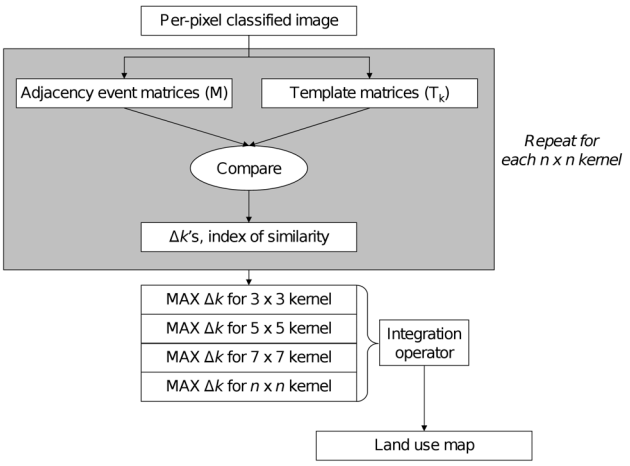


Fig. 3. Flowchart of the OSPARK algorithm

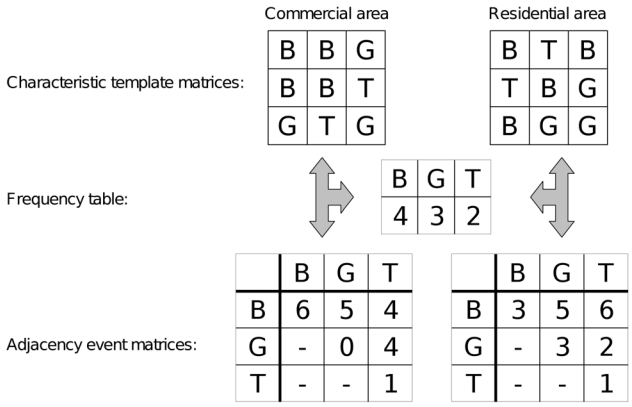


Fig. 4. Simulated 3 by 3 kernels of different urban land-use types with corresponding frequency table and adjacency event matrix. Only the upper triangle is considered because $\mathbf{M}_{ij} = \mathbf{M}_{ji}$. B = building, G = grass and T = trees. (After [13])

May 24 2001 (Fig. 2). In future research the framework will be applied to Landsat TM5 images of 1994, 1997 and other images within the calibration period.

C. The Optimized Spatial Reclassification Kernel (OSPARK)

For the metric-based classification the OSPARK algorithm has been developed. The algorithm facilitates the reclassification of per-pixel classified remote sensing images, using moving window-level spatial metrics. The algorithm is based on the SPAtial Reclassification Kernel (SPARK) [12] algorithm, which has been extended to automatically optimize the kernel size (size of the moving window) to the spatial variation detected around the center pixel of the kernel. The flowchart of the algorithm is shown in Fig. 3.

In this study the adjacency event spatial metric is calculated for each pixel iteratively for kernel sizes with a radius ranging from 1 to 30 pixels. This metric results for each pixel in a matrix ($\mathbf{M}$). Within each $\mathbf{M}$-matrix the frequency and spatial arrangement of the pixel-based classes positioned next to each other as well as diagonally are counted. Each pair of pixels is called an adjacency event. The results of counting the adjacency events are expressed by an adjacency event matrix (Fig. 4).

During each iteration step, i.e. for each kernel size, the adjacency event matrix of each pixel is compared with the template adjacency event matrices ($\mathbf{T}_k$). The $\mathbf{T}_k$ matrices are calculated from template kernels that are representative for the land-use classes to be derived and are in that sense comparable to training areas in per-pixel classifications. Since OSPARK iterates over the kernel sizes, the size of the template kernels should match the kernel radius of each iteration step. The comparison of the $\mathbf{M}$ matrix with the $\mathbf{T}_k$ matrix results in a Δk value, which is the index of similarity. Δk is calculated with [12]:

$$\Delta k = 1 - \sqrt{0.5 \cdot N^{-2} \cdot \sum_{i=1}^{c} \sum_{j=1}^{c} \left(\mathbf{M}_{ij} - \mathbf{T}_{k_{ij}}\right)^2} \qquad (1)$$

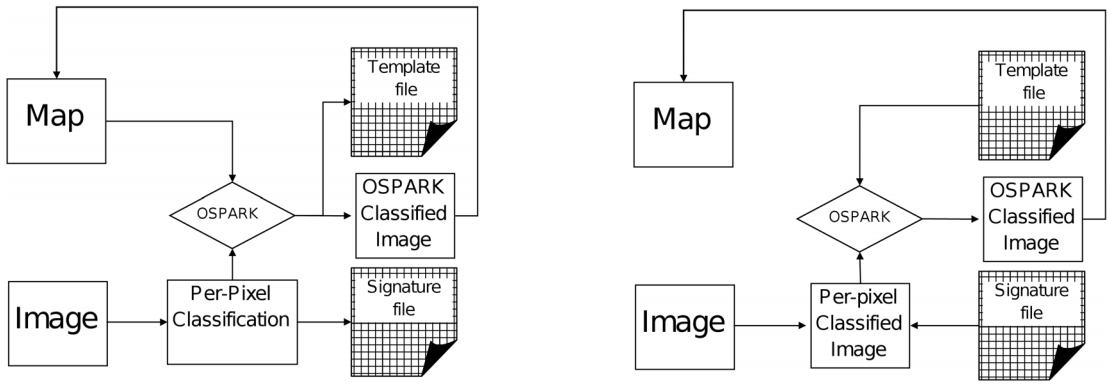
where $\mathbf{M}$ is the adjacency event in an i by j matrix, $\mathbf{T}_k$ is a template matrix for land-use class k, N is the total number of adjacency events in the kernel and c is the number of classes in the per-pixel classified input map. Δk can range from 0 to 1. If Δk equals 0, $\mathbf{M}_{ij}$ is completely different from any $\mathbf{T}_k$, while a value of 1 means that they are identical.

Finally, after all iterations, each pixel in the input image is assigned to the land-use class of the corresponding $\mathbf{T}_k$ matrix with the first local maximum in Δk when increasing the kernel size. The local maximum is chosen, because larger template matrices tend to be less unique for a particular land-use type [8]. Therefore after one or several local maxima Δk will increase with increasing kernel size. The local maximum should be above a user-defined minimum Δk value to prevent classification results with pixels with a low Δk value. The procedure is summarized in Fig. 3.

D. Calibration and validation of OSPARK

In order to apply the OSPARK algorithm to different images, the algorithm needs to be calibrated and validated. It is assumed that the Landsat TM5 image of June 13 1988 corresponds with to land-use map of 1990 and the Landsat ETM7 image of May 24 2001 corresponds to the land-use map of 2000. The calibration procedure is illustrated in Fig. 5.

OSPARK applied to the Landsat image of 1988 is calibrated with the land-use map of 1990, using the following procedure (Fig. 5(a)): (1) Calculate a per-pixel classification; (2) Store the signatures or endmembers of the per-pixel classification; (3) Select center coordinates of template kernels for the desired land-use classes by random sampling locations from the land-use map and store x-coordinate, y-coordinate and class number. 50 samples per land-use class were used in this study; (4) Run OSPARK and store the $\mathbf{T}_k$ matrices for each kernel size; (5) In order to compare the OSPARK classification results (30 m resolution) with the land-use map (200 m resolution), a majority filter with a window size of 4 ha is applied to the OSPARK result; (6) Assess the accuracy of the result by



(a) OSPARK Calibration (b) OSPARK Validation

Fig. 5. Calibration (a) and Validation (b) of the OSPARK algorithm

comparing the OSPARK result at 200 m resolution with the land-use map of 1990. This procedure is applied to different scenarios of pixel-based classifications (see next subsection).

The scenario that gives the highest accuracy with the OSPARK algorithm is applied to the Landsat ETM7 image of 2001 and validated with the land-use map of 2000 (Fig. 5(b)), using the following procedure: (1) Apply the per-pixel classification that gave the best results for the image of 1988 with the stored signatures or endmembers to the image of 2001; (2) Apply OSPARK to this per-pixel classification result with the stored $\mathbf{T}_k$ matrices of 1988; (3) Apply the majority filter in the same way as for 1988; (4) Assess the accuracy of the OSPARK result of 2001 with the land-use map of 2000.

If the validation results are good enough, the per-pixel signatures or endmembers and OSPARK templates are transferable to other images.

E. Pixel-Based Classification Scenarios

Four different pixel-based classification scenarios have been evaluated as an input for the OSPARK algorithm: (1) Iterative Self-Organizing Data Analysis (ISODATA [14]); (2) Kohonen Self-Organizing Maps (SOM [15]); (3) Linear Spectral Mixture Analysis (LSMA [16]) with endmembers representing high albedo substrate, bright vegetation and dark surfaces (SVD [17]) and (4) Linear regression analysis for impervious fraction. This scenario will be referred to as "Impervious".

1) ISODATA: ISODATA [14] is an unsupervised classification algorithm that is commonly used. It is a iterative clustering algorithm: (1) The algorithm initializes by assigning an arbitrary initial cluster vector; (2) each pixel in the image is classified to the nearest cluster; (3) New cluster mean vectors are calculated based on all pixels in a cluster. Step (2) and (3) are repeated until the difference between the iterations are small. The algorithm performs best when clusters are spherical and have the same variance, which is often not the case for remote sensing images, especially in urban areas.

2) SOM: Another unsupervised classification approach is based on Kohonen self-organizing maps (SOM) [15]. A SOM is a type of artificial neural network that was originally developed to visualize topologies and hierarchical structures of multi-dimensional data by transforming the input space into an ordered two dimensional map. The SOM architecture consists of two network layers: an input layer, which is fully connected to a typically two dimensional array of nodes called Kohonen layer or codebook vector map. A weight vector of the same dimension as the input data vectors is associated with each node. The weight vectors are initalised randomly or by evenly sampling the subspace formed by the first two principal components. The SOM is trained by passing an input vector (i.e. a pixels spectral values) to the network, and by choosing a winning node based on the smallest Euclidian distance between the input vector and the weight vectors. Then, the weights of the winning node and its neighbors are adjusted in order to reduce the nodes distance to the input vector. After each image pixel or a representative set of image pixels is passed to the SOM during training, the built model can be applied to any part of the image and even to other images when atmospheric or other calibration constraints are taken into account. Because the trained SOM network assigns each pixel to a particular node in the codebook vector, each such node can be considered to represent a certain information category. This is similar to other unsupervised classification approaches, except that nodes or classes that are closer to each other on the codebook vector are also more spectrally similar. In this research, we applied a SOM with a 3 by 5 Kohonen layer, which divides the image into 15 spectral classes.

3) SVD: Linear spectral mixture analysis (LSMA) is a common approach to sub-pixel classification whereby a pixels observed reflectance is modeled as a linear combination of spectrally pure endmember reflectances. Each endmember contributes proportionally to the overall spectral response according to its relative abundance within the sensors instantaneous field of view (IFOV). To estimate the fractional cover of each endmember within a given pixel, the following equation has to be solved for all image bands simultaneously, using a

least squares approach:

$$R_b = \sum_{i=1}^{n} f_i r_{i,b} + e_b \quad (2)$$

where R_b is the reflectance of the pixel for band b, f_i is the proportion of endmember i within the pixel, $r_{i,b}$ is the reflectance of endmember i for band b, n is the number of endmembers and e_b the error of fit for band b [18]. Inverting this system of mixing equations to retrieve endmember fractions that best fit the observed mixed reflectances implies determining the optimal location of endmembers in feature space. The VIS model is a useful concept for urban analysis because it allows representing any urban area by three physical components: vegetation (V), impervious surfaces (I) and soil (S), in addition to water [19]. However, not all pure vegetation, impervious surfaces or bare soil pixels occupy extreme positions in feature space and can, as such, not be directly used as endmembers for unmixing. Instead, the apexes of the typical triangular shaped feature space correspond to true biophysical endmembers representing high albedo substrate (S), bright vegetation (V) and dark surfaces (D) [17]. Any pixel falling inside the convex hull circumscribing the apexes can be considered as a mixture of these three components, and not of V-I-S. One reason for this is that endmembers are spectrally variable because of brightness differences [16]. In mixture space, pure vegetation pixels are mostly located on the vegetation dark axis, indicating binary mixing between these two endmembers. Darker vegetation types such as trees are located closer to the dark endmember, while brighter vegetation types such as grass or crops are typically found closer to the vegetation endmember. Binary mixing on the grey axis between the dark and substrate endmembers represents different types of urban surfaces, e.g. asphalt versus concrete, while binary mixing on the vegetation high albedo substrate axis is extremely rare [17]. This complicates the direct use of the VIS ternary as an appropriate model for unmixing. Instead, the SVD unmixing model is applied on the Dublin study area. The S and D endmembers for each image are chosen to be identical to the bright and dark pseudo-invariant features that were used for the radiometric calibration. The V endmembers are selected from the extreme pixels on the vegetation axis in a feature space visualization of each image of our time-series by means of high-order principal components. To discretize the continuous output of the LSMA, each of the three components of the SVD unmixing model is first split into 5 subcategories based on equal proportion intervals. Then, the 53 theoretically possible endmember class combinations are assigned a unique identifier (CN) based on:

$$CN = S_c \cdot 5^2 + V_c \cdot 5 + D_c \quad (3)$$

where S_c, V_c and D_c are the class numbers (0-4) of the substrate, vegetation and dark surface endmembers respectively.

4) Impervious: Because LSMA with the SVD endmembers does not provide actual land-cover information, a second subpixel classification approach is adopted. In this approach, a linear regression model is developed to estimate the proportion of vegetation cover within each Landsat pixel. Discounting for a moment the presence of water and exposed soil, the vegetation fraction within a pixel can be considered as the complement of the man-made, sealed surface proportion. The reference data required to build the model are obtained from downsampling an existing land-cover map, which was created from the available Quickbird image. Temporal differences in vegetation cover between the ETM+ images and the land-cover map were filtered out by a temporal filtering technique based on iterative linear regression between NDVI values [20]. Since the regression model only estimates vegetation, and therefore does not explicitly distinguish urban from non-urban surface cover, an urban mask is developed to indicate pixels belonging to urban land cover. Only pixels within this mask are subjected to the regression model. The sealed surface proportions of all pixels outside the mask are considered zero. To create this mask, we apply a non-parametric unsupervised classification approach based on Kohonen self-organizing maps (SOM) [15] and enhance the output map with knowledge-based post-classification rules [21], [22]. The resulting urban mask has an overall accuracy of 93% and a kappa index of agreement of 89%. The RMSE error of the estimated sealed surface fractions is 0.14. To make the estimated sealed surface proportions discrete, they are split into 5 classes using natural breaks [23].

F. Spectral Normalization

Prior to applying the best OSPARK scenario to images of different dates according to the calibration framework, the reflectance in the images need to be normalized. The raw digital numbers of the images in our time-series are converted to exoatmospheric reflectance according to the formulas and calibration parameters presented by The Landsat 7 Users Handbook [24]. While this conversion removes predictable effects caused by differences in solar irradiance and solar angle, it does not take into account the influence of atmospheric condition and sensor drift on the measured radiances. To quantify changes in surface reflectance between the two acquisition dates with spectral mixture analysis, the impact of temporal spectral variability that is not caused by changes in surface reflectance should therefore be minimized. Because no atmospheric data or field measurements of ground reflectance are available, a relative reflectance calibration based on the identification of pseudo-invariant features [25] is applied. By visually comparing the four images, nine sites are selected for which the surface reflectance is not expected to have changed in between the acquisition dates. Three are chosen to represent high albedo surfaces: highly reflective roofs of commercial or industrial buildings in the port area. Three low albedo sites are selected on the Liffy River in central Dublin, and three sites with asphalt are selected on the airport runway. Because at-sensor radiances vary linearly with ground reflectances for visible and short wave infrared wavelengths [26] and because this relationship can be extended to multidate images [27], the pseudoinvariant features of each image can be linearly transformed to give them the same apparent reflectance as

TABLE II
ACCURACY ASSESSMENT OF OSPARK FOR DIFFERENT PER-PIXEL CLASSIFICATION SCENARIOS (ALL CLASSES). K = KAPPA

Scenario	K	Fr. Correct	Fuzzy K	Fuzzy Fr. Correct
ISODATA	0.25	0.33	0.11	0.48
SOM	0.32	0.44	0.21	0.57
SVD	0.40	0.52	0.30	0.66
Impervious	0.24	0.32	0.14	0.33

TABLE III
NUMBER OF CLASSES (ALL CLASSES) FOR WHICH THE STATISTICS SHOW A VALUE LARGER THAN 0.5. USER = USER'S ACCURACY, PRODUCER = PRODUCER'S ACCURACY, K = KAPPA

Scenario	K	User	Producer	Fuzzy K	Total
ISODATA	1	3	5	1	10
SOM	2	4	8	2	16
SVD	2	3	9	2	16
Impervious	4	5	4	2	15

TABLE IV
ACCURACY ASSESSMENT OF OSPARK FOR DIFFERENT PER-PIXEL CLASSIFICATION SCENARIOS (MERGED CLASSES). K = KAPPA

Scenario	K	Fr. Correct	Fuzzy K	Fuzzy Fr. Correct
ISODATA	0.60	0.73	0.44	0.80
SOM	0.57	0.71	0.41	0.78
SVD	0.63	0.75	0.47	0.81
Impervious	0.65	0.81	0.47	0.83

TABLE V
NUMBER OF CLASSES (MERGED CLASSES) FOR WHICH THE STATISTICS SHOW A VALUE LARGER THAN 0.50. USER = USER'S ACCURACY, PRODUCER = PRODUCER'S ACCURACY, K = KAPPA

Scenario	K	User	Producer	Fuzzy K	Total
ISODATA	4	3	5	2	14
SOM	3	4	5	1	13
SVD	4	4	7	3	18
Impervious	3	4	3	2	12

the 1994 reference image [28]. The estimated linear function between the pseudo-invariant sites of the four images can then be applied on the entire 1988, 1997 and 2001 images to reduce temporal spectral variability caused by the combined impact of differences in illumination, sensor drift and atmosphere.

III. RESULTS

A. Per-Pixel Classification Scenarios

Table II shows the contingency matrix statistics of the accuracy assessment when all 23 MOLAND land-use classes are considered. The statistics show that SVD gives the best overall classification result, while the linear regression analysis for impervious fraction gives the worst result. All Kappa's and Fuzzy Kappa's, however, are lower than 0.50 indicating many classification errors. Table III shows for each scenario the number of classes that has a value larger than 0.50 for Kappa, user's accuracy, producer's accuracy [29] and Fuzzy Kappa. Only for SVD these classes were functional urban classes (e.g. residential areas).

B. Merged Classes

When all classes of the MOLAND model are considered, the accuracies of the OSPARK classification for each class remain low for all pixel-based classification scenarios. Therefore the original 23 classes are recoded to 8 classes following Table I. The legend is given in Table VI. Reclassification is also needed, because the area of some classes is too small to be detected by a contextual classifier. Furthermore, inconsistencies between the two land-use maps for the residential classes were observed.

The accuracy assessment has also been done on merged land-use classes. Table IV shows that the overall accuracy results of the merged classes are much better than the results for 23 classes (Table II). All Kappa's are above 0.50 and the (Fuzzy) fraction correct is high for all scenarios. On the other hand, the Fuzzy Kappa's are lower than 0.50, indicating that the overall spatial patterns are not well reproduced by the classification scenarios. The Impervious scenario shows the best overall results.

Table V shows for each scenario the number of classes that has a value larger than 0.50 for Kappa, user's accuracy, producer's accuracy and Fuzzy Kappa, comparable to Table III. The scores are better for the merged classes than for all 23 classes. The best results are obtained for the SVD scenario. The Impervious scenario gives the worst results when the classes are considered individually. Therefore, the SVD method has been chosen to provide the input pixel-based classification for the OSPARK algorithm, deriving the merged classes of Table VI. The next section gives the results of the validation of OSPARK using data for the year 2000.

C. Validation of OSPARK

The SVD pixel-based classification scenario that gave the best results for the year 1990 has been applied to the Landsat ETM7 image of 2001. The output land-use classes resulting from the OSPARK algorithm were merged (Table VI) and compared with the land-use map of the year 2000. The validation results (Table VI) show that for all merged land-use classes the results of 2000 are worse compared to 1990. The most important differences are the decrease in producer's accuracy for "Non-residential urban areas" and "Airports" below a value of 0.50. Although the accuracies of the classification for the year 2000 are reduced, the results for residential areas are acceptable. Residential areas are the most important land-use class to be considered, when studying urban sprawl.

In order to evaluate the land-use structures, the differences

TABLE VI
COMPARISON OF USER'S AND PRODUCER'S ACCURACY OF OSPARK APPLIED TO 1990 AND 2000

Recode	Class	Year 1990		Year 2000	
		User	*Producer*	*User*	*Producer*
1	Agricultural land	0.94	0.80	0.92	0.66
2	(Semi-) natural land	0.52	0.74	0.49	0.67
3	Residential areas	0.71	0.62	0.65	0.59
4	Non-residential urban areas	0.31	0.53	0.23	0.41
5	Construction, mining & dump sites	0.09	0.66	0.05	0.55
6	Airports	0.24	0.58	0.06	0.29
7	Artificial non-agricultural areas	0.28	0.19	0.14	0.09

between the land-use maps and OSPARK classification were checked for the different metrics that are important indicators for urban sprawl, i.e. PAFRAC, CONTAG, SHDI and IJI. The metrics are calculated in a circular moving window of 1600 m radius. This radius corresponds with the definition of the distance decay functions of the cellular automata model in the MOLAND model [9]. Next, the Fuzzy Kappa is calculated for different map pairs. The results show for all metrics that the Fuzzy Kappa is higher between the land-use maps of 1990 and 2000, than between the land-use maps and their corresponding OSPARK classification. Also the Fuzzy Kappa between the OSPARK classifications is high.

IV. DISCUSSION

In this paper a framework for the historic calibration of land-use change models using remote sensing data has been proposed. The framework needs calibration using two land-use maps. Inconsistencies between the land-use maps can affect the calibration accuracy. The results show on the other hand that the land-use maps of 1990 and 2000 are very comparable in terms of urban sprawl, in contrast with their corresponding OSPARK classification. Also the OSPARK classifications of different dates give comparable results for the urban sprawl metrics. This suggests that the amount of noise in the remote sensing classification and the amount of generalization in the land-use maps is comparable.

The current application of the OSPARK algorithm uses the adjacency event metric. Further research should be done to de incorporation of other spatial metrics or sets of spatial metrics. If a set of metrics proves to be generally applicable to urban landscapes to differentiate between different urban land-use classes, the training phase has to be done only once to derive the appropriate Landscape Metrics Signatures (LMS) [30] for each class. When other metrics are used, the Δk similarity index should be replaced by an appropriate goodness-of-fit measure. Another modification to the OSPARK algorithm is the use of a circular kernel in stead of a square kernel, as used in the cellular automata algorithm of the MOLAND model. This could reduce artifacts in the classification.

The OSPARK classification of the Landsat ETM7 image of 2001 gave less good results. This could be caused by errors in the normalization, which can give errors in the transfer of endmembers and templates. Also inconsistencies between the land-use maps can give errors in the validation of the results. A third problem is the discretization of SVD. For different images it can result in a different number of unique combinations. When template kernels contain SVD classes that do not exist in the image to which it is applied, the template is not used. When the image to which OSPARK is applied has some other SVD classes than the image used for calibration of the OSPARK algorithm, kernels containing these pixels can not be classified with the current algorithm.

V. CONCLUSIONS

Spatial metrics derived from remote sensing data can increase the amount of data used for the calibration of land-use change models, making them less dependent on land-use maps. This study proposed a framework for the calibration of land-use change models using remote sensing data. The framework consists of a contextual, metric-based, classification algorithm, OSPARK. OSPARK results have been evaluated for four scenarios of pixel-based classifiers that provided input data for the algorithm. The OSPARK algorithm in combination with the SVD unmixing algorithm gave the best results for the Landsat TM5 image of 1988. The accurate discrimination of all 23 land-use classes of the MOLAND model was not possible. Reclassification into 8 classes, while preserving the differentiation in important classes gave good results.

Application of OSPARK to an SVD classification of a Landsat ETM7 image of 2001 gave worse results. Further research should focus on the improvement of the transferability of OSPARK to other images. Improvements to the OSPARK algorithms have been suggested.

ACKNOWLEDGMENTS

We would like to express our acknowledgments to BELSPO (Belgian Science Policy) for funding the MAMUD research project (SR/00/105). The land-use data used for Dublin are made available by the MOLAND project of the EU - Joint Research Centre in Ispra, Italy.

REFERENCES

[1] C. Agarwal, G. Green, J. Grove, T. P. Evans, and C. Schweik, "A review and assessment of land-use change models: dynamics of space, time, and human choice," US Department of Agriculture, Forest Service, Northeastern Research Station, Newton Square, PA, Gen. Tech. Rep. NE-297, 2002.

[2] P. Gong, D. J. Marceau, and P. J. Howarth, "A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data," *Remote Sens. Environ.*, vol. 40, pp. 137–151, 1992.

[3] M. J. Barnsley and S. L. Barr, "Distinguishing urban land-use categories in fine spatial resolution land-cover data using a graph-based, structural pattern recognition system," *Comput., Environ. and Urban Systems*, vol. 21, no. 3/4, pp. 209–225, 1997.

[4] S. M. de Jong and F. van der Meer, *Remote sensing image analysis, including the spatial domain*, ser. Remote sensing and digital image processing. 5: Kluwer academic publishers, 2004.

[5] D. C. Parker, T. P. Evans, and V. Meretsky, "Measuring emergent properties of agent-based land-cover/land-use models using spatial metrics," in *Seventh annual conference of the international society for computational economics*, Yale University, 2001, p. 24.

[6] P. Torrens, "A toolkit for measuring sprawl," *Appl. Spatial Analysis*, vol. 1, pp. 5–36, 2008.

[7] M. Herold, X. Liu, and K. C. Clarke, "Spatial metrics and image texture for mapping urban land use," *Photogramm. Eng. Rem. S.*, vol. 69, no. 9, pp. 991–1001, 2003.

[8] J. van der Kwast and G. Engelen, "Calibration of land use change models: the potential of spatial metrics," in *28th EARSeL Symposium and Workshops, Remote sensing for a changing Europe*, R. Goossens, Ed. Istanbul: Millpress, Rotterdam, in press.

[9] G. Engelen, R. White, and I. Uljee, "The MURBANDY and MOLAND models for dublin. final report," RIKS, Tech. Rep., 2002.

[10] R. White, G. Engelen, and I. Uljee, "The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics," *Environment and Planning B: Planning and Design*, vol. 24, pp. 323–343, 1997. [Online]. Available: http://www.envplan.com/abstract.cgi?id=b240323

[11] A. Hagen-Zanker, B. Straatman, and I. Uljee, "Further developments of a fuzzy set map comparison approach," *Int. J. Geogr. Inf. Sci.*, vol. 19, no. 7, pp. 769 – 785, 2005.

[12] M. Barnsley and S. Barr, "Inferring urban land use from satellite sensor images using kernel-based analysis and classification," *Photogramm. Eng. Rem. S.*, vol. 62, no. 8, pp. 949–958, 1996.

[13] A. Alimohammadi and M. Shirkavand, "Optimization of the SPARK algorithm for land use mapping from remotely sensed data," in *Geomatics conference 83*. National Cartographic Center, 2004.

[14] G. H. Ball and D. J. Hall, "ISODATA, a novel method of data analysis and pattern classification," Stanford Research Institute, Tech. Rep., Apr. 1965.

[15] T. Kohonen, *Self-Organizing Maps*, 3rd ed. Berlin: Springer, 2001, vol. 30.

[16] C. Wu, "Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery," *Remote Sens. Environ.*, vol. 93, pp. 480–492, 2004.

[17] C. Small, "The landsat ETM+ spectral mixing space," *Remote Sens. Environ.*, vol. 93, pp. 1–17, 2004.

[18] F. van der Meer, "Image classification through spectral unmixing," in *Spatial Statistics for Remote Sensing*, A. Stein, F. van der Meer, and B. Gorte, Eds. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999, pp. 185–193.

[19] M. Ridd, "Exploring a vis (vegetationimpervious surfacesoil) model for urban ecosystem analysis through remote sensing - comparative anatomy for cities," *Int. J. Remote Sens.*, vol. 16, pp. 2165–2185, 1995.

[20] T. van de Voorde, J. Vlaeminck, and F. Canters, "Comparing different approaches for mapping urban vegetation cover from landsat ETM+ data: a case study on brussels," *Sensors*, vol. 8, no. 6, pp. 3880–3902, 2008.

[21] T. van de Voorde, W. de Genst, and F. Canters, "Improving pixel-based VHR land-cover classifications of urban areas with post-classification techniques," *Photogramm. Eng. Rem. S.*, vol. 73, pp. 1017–1027, 2007.

[22] T. van de Voorde, L. Demarchi, and F. Canters, "Multi-temporal spectral unmixing to characterise change in the urban spatial structure of the greater dublin area," in *28th EARSeL Symposium and Workshops, Remote sensing for a changing Europe*, R. Goossens, Ed. Istanbul: Millpress, Rotterdam, in press.

[23] G. Jenks and F. Caspall, "Error on choroplethic maps: definition, measurement, reduction," *Ann. Assoc. Am. Geogr.*, vol. 61, no. 2.

[24] R. Irish, *Landsat 7 science data users handbook*. NASA, 2007.

[25] J. Schott, C. Salvaggio, and W. Volchok, "Radiometric scene normalization using pseudoinvariant features," *Int. J. Remote. Sens.*, vol. 26, pp. 1–16, 1988.

[26] J. Cornel, "Determination of surface reflectance and estimates of atmospheric optical depth and single scattering albedo from landsat thematic mapper data," *Int. J. Remote Sens.*, vol. 11, pp. 783–828, 1990.

[27] V. Caselles and M. Garcia, "An alternative simple approach to estimate atmospheric correction in multitemporal studies," *Int. J. Remote Sens.*, vol. 10, pp. 1127–1134, 1989.

[28] F. Hall, D. Strebel, J. Nickeson, and S. Goetz, "Radiometric rectification: toward a common radiometric response among multidate multisensor images," *Remote Sens. Environ.*, vol. 35, pp. 11–27, 1991.

[29] M. Story and R. G. Congalton, "Accuracy assessment: a user's perspective," *Photogramm. Eng. Rem. S.*, vol. 52, no. 3, pp. 397–399, 1986.

[30] M. Herold and G. Menz, "Landscape metric signatures (LMS) to improve urban land use information derived from remotely sensed data," in *A decade of trans-european remote sensing cooperation. Proceedings of the 20th EARSeL Symposium*, M. F. Buchroithner, Ed. Balkema, Rotterdam, 2001, pp. 251–256.