Analysis of vegetation and land cover dynamics in north-western Morocco during the last decade using MODIS NDVI time series data

C. Höpfner and D. Scherer

Chair of Climatology, Department of Ecology, Technische Universität Berlin, Rothenburgstraße 12, 12165 Berlin, Germany

Received: 10 February 2011 – Published in Biogeosciences Discuss.: 18 April 2011
Revised: 12 October 2011 – Accepted: 9 November 2011 – Published: 17 November 2011

Abstract. Vegetation phenology as well as the current variability and dynamics of vegetation and land cover, including its climatic and human drivers, are examined in a region in north-western Morocco that is nearly 22,700 km$^2$ big. A gapless time series of Normalized Difference Vegetation Index (NDVI) composite raster data from 29 September 2000 to 29 September 2009 is utilised. The data have a spatial resolution of 250 m and were acquired by the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor.

The presented approach allows to compose and to analyse yearly land cover maps in a widely unknown region with scarce validated ground truth data by deriving phenological parameters. Results show that the high temporal resolution of 16 d is sufficient for (a) determining local land cover better than global land cover classifications of Plant Functional Types (PFT) and Global Land Cover 2000 (GLC2000) and (b) for drawing conclusions on vegetation dynamics and its drivers. Areas of stably classified land cover types (i.e. areas that did not change their land cover type) show climatically driven inter- and intra-annual variability with indicated influence of droughts. The presented approach to determine human-driven influence on vegetation dynamics caused by agriculture results in a more than ten times larger area compared with stably classified areas. Change detection based on yearly land cover maps shows a gain of high-productive vegetation (crop land) of about 259.3 km$^2$. Statistically significant inter-annual trends in vegetation dynamics during the last decade could however not be discovered. A sequence of correlations was respectively carried out to extract the most important periods of rainfall responsible for the production of green biomass and for the extent of land cover types. Results show that mean daily precipitation from 1 October to 15 December has high correlation results (max. $r^2 = 0.85$) on an intra-annual time scale to NDVI percentiles (50 %) of land cover types. Correlation results of mean daily precipitation from 16 September to 15 January and percentage of yearly classified area of each land cover type are medium up to high (max. $r^2 = 0.64$). In all, an offset of nearly 1.5 months is detected between precipitation rates and NDVI values. High-productive vegetation (crop land) is proved to be mainly rain-fed. We conclude that identification, understanding and knowledge about vegetation phenology, and current variability of vegetation and land cover, as well as prediction methods of land cover change, can be improved using multi-year MODIS NDVI time series data. This study enhances the comprehension of current land surface dynamics and variability of vegetation and land cover in north-western Morocco. It especially offers a quick access when estimating the extent of agricultural lands.

1 Introduction

The mapping, quantifying and monitoring of land cover (LC) is essential to understand the current state of landscape. The use of satellite-based remote sensor data has been widely applied to receive cost-effective LC data over large geographic regions (Lunetta et al., 2006) because field surveys are not only an expensive undertaking counting time, labour and costs they also quickly become outdated (Lucas et al., 2007). Especially if validated ground truth knowledge is as scarce as in Morocco, freely available remote sensing data sets are the key to get contiguous spatio-temporal information about LC and vegetation dynamics (Evrendilek and Gulbeyaz, 2008). These dynamics base on factors such as human land use, climate, soil and groundwater conditions, invasion of non-native species, volcanism and tectonics. While focusing on human and climatic drivers as the main factors of land cover change (LCC), a differentiation between human-driven change, climate-driven change, vegetation trends and
inter-annual ecosystem variability is still a challenge for LC research. An understanding of current vegetation dynamics (phenology) and recent inter- and intra-annual variability of LC is required to identify and predict the LCC in the study region (Bradley and Mustard, 2008).

1.1 NDVI, MODIS data and phenology

The Normalized Differenced Vegetation Index (NDVI; Sellers, 1985) as an indicator of vivid green vegetation and as a descriptor of ecosystem functions has proved to be very valuable for assessing ecological responses to environmental changes (Alcaraz-Segura et al., 2009; Pettorelli et al., 2005). NDVI-based approaches are widely used for monitoring vegetation activity using time series data sets (TSD) acquired by sensors aboard satellites with short revisit periods (Goward and Prince, 1995; Beck et al., 2006). The Moderate Resolution Imaging Spectroradiometer (MODIS), as the latest of these sensors, is one of the key instruments of NASA’s Earth Observing System (EOS) and is a major step ahead over previous sensors in terms of its spectral (36 bands), spatial (max. 250 m) and temporal (1–2 d for the entire Earth) resolutions (Xiong and Barnes, 2006).

Plenty of studies use MODIS TSD in combination with ground truth information to map croplands (e.g., Xiao et al., 2005; Chang et al., 2007; Fritz et al., 2008; Wardlow et al., 2007, 2008; Shao et al., 2010), to predict the production of wheat (Ren et al., 2008) or to map irrigated areas (Thenkabail et al., 2005). Characteristic temporal profiles of NDVI (end-members) have been presented for corn, soybean, wheat, alfalfa, paddy rice and fallow ground (Xiao et al., 2005; Wardlow et al., 2007; Chang et al., 2007; Geerken, 2009; Shao et al., 2010). Classification and monitoring studies, apart from crop-related themes, show that the temporal and spatial resolution of MODIS data have the potential to monitor floating vegetation in lakes (Kiage, 2009), to assess regions at risk of desertification (Wang et al., 2007) and to discriminate native and invaded grassland species (Huang, 2009).

In terms of entire natural terrestrial ecosystems Evrendilek and Gulbeyaz (2008) show the potential of MODIS NDVI and Enhanced Vegetation Index (EVI) TSD (16 d composites with 500 m spatial resolution) for the estimation and monitoring of seasonal and inter-annual ecosystem dynamics over Turkey. They explored data according to four LC types, six biogeoclimate zones, four seasons and seven years presenting correlations of MODIS EVI- and NDVI TSD with higher mean values and higher standard deviation of NDVI TSD.

Multiple phenologic parameters related to vegetation activity and derived from NDVI TSD were defined by Reed et al. (1994) to describe seasonal dynamics (phenology) of vegetation. Several studies applied these phenologic metrics, or dealt with additional indicators, to describe vegetation phenology (summary in Bradley and Mustard, 2008; Funk and Budde, 2009). In terms of MODIS NDVI TSD, a general comparison of satellite measured start of season and actual start of season was made by Wardlow et al. (2006). Soudani et al. (2008) evaluated the time of onset of green-up for temperate deciduous broadleaf forests while Archibald and Scholes (2007) examined green-up dates in the African savannah. With regard to rice crops in Italy, Boschetti et al. (2009) produced phenological calendars.

Even though multiple studies already use remote sensing information for an understanding of the phenology of land surface, methods are still in development. Understanding of the phenology is regarded as a key to understand land-surface processes (Reed, 2006), which are mainly based on human and climatic drivers as described.

1.2 Objectives

The major objective of our study is to point out an approach to extract broad knowledge about vegetation and LC dynamics in a widely unknown region. In contrast to the majority of other studies using rastered MODIS NDVI TSD, we present a methodology of analysing vegetation and LC dynamics when ground truth knowledge is scarce. We apply the freely available rastered MODIS NDVI TSD to make this an easy-to-handle approach and to show the amount of information that can be extracted from contiguous spatio-temporal data. The study aims to map, quantify and monitor LC and to contribute to differ between human-driven change, climate-driven change, vegetation trends and inter-annual ecosystem variability. The workflow including the pre-processing is shown in Fig. 1.

In a first step, annual LC maps have been computed by annual LC classifications based on a hierarchically structured decision tree classifier using phenologic metrics. In a second step, areas classified as stable (i.e. areas that do not change their LC) and areas with actual LCC were extracted...
to understand current vegetation and LC dynamics. The extraction of annual endmembers of stably classified areas allows derivation of phenologic metrics of each land cover type (LCT). These phenologic metrics allow to analyse vegetation trends during the last decade and to distinguish vegetation variability. In a last step, climate-driven changes were suspected to be mainly caused by presence or absence of precipitation. Therefore mean NDVI endmembers of stably classified LCT were compared to precipitation data from two measurement stations within the study region. A sequence of correlations was carried out to extract the most important periods of rainfall needed for the production of green biomass and for the extent of LCT. With the focus on areas of high productivity, human-driven impacts caused by agriculture can be shown using a simple approach.

2 Materials and methods

2.1 Study region

The study region (Fig. 2) includes the cities of Casablanca, Rabat and Salé and covers an area from 32°42’N to 34°9’N latitude and from 6°19’W to 8°27’W longitude of 22 677.25 km² (158.75 km x 200.75 km) ashore. The Atlantic Ocean to the north-west and the Middle Atlas to the south-east are natural borders. The mean elevation is about 380 m a.s.l. and the maximum elevation about 1072 m a.s.l. in the Atlas Mountains. The climate near the coast is moderated due to the Canary Current. The mean air temperature in Casablanca is 17.4 °C. Precipitation normally occurs during the winter months with a total mean of yearly 423 mm; ranging between 61 and 707 mm (GHCN DATA 1951-2006, Nouasseur station).

2.2 Data set and pre-processing

The gapless NDVI TSD of MOD13Q1 product (collection 5) covers nine years (Table 1) and has been acquired from the Warehouse Inventory Search Tool (WIST, 2009). In total 207 NDVI data sets of 16 d composites (Huete et al., 1999) were used.

Although the use of composite data already reduces the effect of noise like cloud contamination, shadow, sun angle or aerosol effects (Holben, 1986; Huete et al., 1994), data are still negatively influenced by noise, resulting in generally underestimated NDVI values (Gu, 2006; Hird and McDermid, 2009). Several smoothing techniques like double logistic function-fitting (Beck, 2006; Hird and McDermid, 2009), asymmetric Gaussian function-fitting (Jónsson and Eklundh, 2002; Hird and McDermid, 2009) or Savitzky-Golay filter (Savitzky and Golay, 1964; Chen et al., 2004; Gong et al., 2006; Doraiswamy et al., 2007; Hird and McDermid, 2009; Ren et al., 2008) are used to model a nearly noise-free NDVI time series following the main assumption that vegetation follows a continuous sequence drawn by increase, peak and decrease of NDVI, describing a clear mathematical function. Modelling this mathematical function, smoothing algorithms eliminate depressions of TSD. However, the purpose of these algorithms is not to distinguish between depressions caused by clouds or atmospheric disturbance nor human caused depressions and other natural disturbances (Evrendilek and Gulbeyaz, 2008). In order to keep responses to climate or other drivers it is assumed that noise (mostly technical or physical caused) depresses single NDVI composites while climate conditions, human impacts or other drivers depress a sequence of NDVI composites. Thus, the simple but effective smoothing technique (Eq. 1) introduced by Gu et al. (2006) is applied, which assumes that NDVI is always depressed and never overrated by noise. The algorithm reduces impact of single contaminated data points and keeps the upper envelope. Following Gu et al. (2006) the algorithm was applied three times in order to remove single noise contaminations more efficiently.

\[
\text{NDVI}_{\text{smooth}} = \text{Max}(\text{NDVI}_t, 0.25 \times \text{NDVI}_{t-1} + 0.5 \times \text{NDVI}_t + 0.25 \times \text{NDVI}_{t+1})
\]

The actual time step is represented by \( t \), the previous by \( t - 1 \) and the following by \( t + 1 \).

Data were re-projected into Universal Transverse Mercator (UTM, Zone 29, WGS 84). Incongruous picture elements in the Atlantic Ocean near the coast were eliminated applying a land-water mask. The mask was derived from Shuttle Radar Topography Mission (SRTM) data and includes two additional reservoirs ashore which were masked from SRTM data by catching connected areas of the same altitude of water level.

According to the vegetation break during the summer months and the start of the vegetation period on the first of October, the NDVI mega-file of 207 composites was divided
Table 1. Overview of data used.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Spatial resolution</th>
<th>Corresponding time/time of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Land Cover 2000 (GLC 2000) v.3</td>
<td>1000 m</td>
<td>2000</td>
</tr>
<tr>
<td>MODIS MOD13Q1 collection 5</td>
<td>250 m</td>
<td>29 September 2000–29 September 2009</td>
</tr>
<tr>
<td>Occupation du sol Grand Casablanca</td>
<td>–</td>
<td>2004</td>
</tr>
<tr>
<td>SPOT 5</td>
<td>2.5 m</td>
<td>16 March 2004</td>
</tr>
<tr>
<td>SRTM</td>
<td>~ 90 m</td>
<td>2000</td>
</tr>
</tbody>
</table>

into nine single data files according to the years of 2001 to 2009 with 23 composites each. Thus, one data file effectively covers a time period starting at 29 September/30 September and ending at 28 September/29 September of the following year due to composite length of 16 d. The data set called e.g. “2001” covers the vegetation period starting at the end of September 2000 and finishing at the end of September 2001.

2.3 Land cover classification

LC classifications deliver spatial patterns of LCT. Within the study region, the sole use of NDVI TSD for distinction of LCT apart from very sparsely vegetated and forested areas, is critical because not all grass-, shrub- and croplands have a characteristic productivity. Varying altitudes, varying distances to the coast, varying soils and water availability, as well as multi-crop systems and scarce validated ground truth data hamper a discrimination of these LCT. Analyses of histograms and scatter plots did not result in any possibility to discriminate different LCT specifically. Therefore, LC is discriminated into five defined LCT as: very sparsely vegetated (LCTvsv), sparsely vegetated (LCTsv), forest (LCTfor), high-productive vegetation (LCThpv) and low-productive vegetation (LCTlpv), where high-productive vegetation and low-productive vegetation discriminate LC by productivity of green biomass. This approach allows gathering rain fed, irrigated and multi-crop croplands as well as different grasslands and shrub lands in the defined LCT without discriminating various soils, locations and water supply.

First approaches using classification methods such as the unsupervised ISODATA classifier or the supervised maximum-likelihood classifier lead to unsatisfying results. Traditional methods for LC identification often fail in the Moroccan context due to small parcels of cropland, a high variety of crops and a high heterogeneity of development (V. Simonneaux, 2009, personal communication, 24 November 2009). Thus a hierarchically structured binary decision tree classifier is applied. Five phenologic metrics are deduced (Fig. 3) for each year from the smoothed NDVI TSD to run the decision tree classifier (Fig. 4). In terms of validation, classification accuracy and classification efficiency a decision tree classifier relies heavily on the decision tree chosen. The two most separable types are processed first and the most subtle type pair is discriminated at the bottom of the tree. By so doing, the cumulative error will be minimised (Richards and Jia, 2006).

Following thresholds of NDVI lower than 0.18 (Bradley and Mustard, 2008; Simonneaux, 2007) and lower than 0.2 (Geerken, 2009; Ren et al., 2008), very sparsely vegetated areas were classified by the maximum value of NDVI within the vegetation period (MaxV) lower than 0.2.

The perennial LCT_for with high NDVI values all-the-year is easily to determine in Morocco with freely available Landsat 7 ETM+ scenes. First classification results using training areas showed that it is not possible to delimit LCT_for by single use of thresholds of NDVI values at the beginning of the vegetation period (OnV), nor the NDVI value at the end of the vegetation period (EndV) and the mean NDVI value of the vegetation period (MeanV) because non-forested areas were also classified as LCT_for. Therefore, a combination of the three metrics was applied.

High-productive vegetation growing on agricultural lands has a higher NDVI range (RanV) between MaxV and the minimum of EndV and OnV than wild vegetation according to seed, a dense stand within the vegetation period and a low vegetated surface after crop during the vegetation rest. Benhadj et al. (2008) report NDVI temporal profiles for crop areas in Marrakech (Morocco) showing a RanV greater than 0.4. Other studies report at least a RanV of 0.4 for single crop systems (Chang et al., 2007; Shao et al., 2010).
Fig. 3. Derived phenologic metrics according to Reed et al. (1994): MaxV (maximum NDVI value within the vegetation period), MeanV (mean NDVI value of the vegetation period), OnV (NDVI value at the beginning of the vegetation period), EndV (NDVI value at the end of the vegetation period), RanV (range between maximum value within the vegetation period and minimum of OnV and EndV).

In distinction to LCT\textsubscript{lpv} the remaining vegetation with less productivity is summarised in LCT\textsubscript{lpv} and may also include croplands with moderate or poor yields having less productivity of green biomass than areas with rich yields.

2.4 Validation of land cover classifications

Validation of classification results requires validated in-situ ground truth data, official maps or other validated data. A two-part validation of decision tree classifier results is applied because ground truth knowledge in the study region is limited to the administrative region of Grand Casablanca (boundary see Fig. 11). In a first step, decision tree classification results are visually validated for the entire study region using freely available Landsat 7 ETM + data measured within different vegetation periods. In a second step, final validation using ground truth data is made based upon an official land use/LC map of the administrative region of Grand Casablanca from the year 2004 (Agence Urbaine de Casablanca, 2006). The old medina of Casablanca as very dense city quarter was used as a ground truth polygon for LCT\textsubscript{svs} and less dense city quarters with more vegetation were used as ground truth polygons for LCT\textsubscript{sv} to differ the city region into LCT\textsubscript{svs} and LCT\textsubscript{sv}. Additional ground truth polygons have been acquired directly from a SPOT image such as parks in the city of Casablanca, vegetated areas around the runways of the airport Mohammed V (location see Fig. 6) and the former airport in western Casablanca since the official map does not indicate areas with further vegetation specifications apart from forests and croplands. Global LC classification data as Global Land Cover 2000 (GLC, 2000; Mayaux et al., 2003) and MODIS product MOD12 (Strahler et al., 1999) were additionally validated to further estimate the quality of decision tree classifier results. Out of MOD12 product, LC maps of Plant Functional Types (PFT) are applied for validation since its allocation of LCT fits best to the previously defined LCT in this study. Annual classification maps of PFT contain information from 2001 up to 2004 and GLC2000 contains a LC classification map of 2000.

Usage of confusion matrix and Kappa coefficient is a widely applied method to assess the classification accuracy (Richards and Jia, 2006). In our context it has been used with the identical ground truth data both for the PFT classification of 2004 and decision tree classification result of 2004. A confusion matrix for GLC2000 is not made since ground truth data and classification map concern different years.

2.5 Characterisation of land cover and land cover dynamics

Stably classified areas are extracted to describe vegetation and LC dynamics in the study region from 2001 up to 2009. Derivation of mean endmembers of stably classified areas provides information about intra-annual and inter-annual dynamics of vegetation. NDVI percentiles of stably classified areas, as well as NDVI percentiles of all classified areas, were computed to display inter-annual ecosystem variability. The determination of vegetation trends during the last decade is made by calculating yearly phenologic metrics of endmembers of stably classified areas.

Inter-annual dynamics of LC (LCT A as e.g. LCT\textsubscript{for} changes to LCT B as e.g. LCT\textsubscript{sv}) were identified by comparing all nine annual classification results of 2001 up to 2009.
LCC were assumed to be actual LCC as
2001–2002 classified as LCT A and 2003–2009 classified as LCT B (2 years + 7 years)
2001–2003 classified as LCT A and 2004–2009 classified as LCT B (3 years + 6 years)
2001–2004 classified as LCT A and 2005–2009 classified as LCT B (4 years + 5 years)
2001–2005 classified as LCT A and 2006–2009 classified as LCT B (5 years + 4 years)
2001–2006 classified as LCT A and 2007–2009 classified as LCT B (6 years + 3 years)
2001–2007 classified as LCT A and 2008–2009 classified as LCT B (7 years + 2 years)

by using the requirement that there is exactly one LCC between two LCT within the nine years.

Other patterns of classified results are not regarded as actual LCC to insure that fluctuations caused by method or other reasons like inter-annual precipitation variability do not effect monitoring of LCC.

2.6 Human- and climate-driven impacts

The identification and differentiation of human- and climate-driven impacts through the use of NDVI TSD is difficult due to the diversity of impacts. Apart from areas with actual LCC, human driven impacts on LC are detected by following the approach of the decision tree classifier which assumes that croplands used as productive land surfaces have a higher RanV due to seed and crop. Following this, the mean annual RanV from 2001 up to 2009 (RanV_{mean}) is computed for each pixel of the study region to assess and indicate human-driven impact apart from LC classification for the entire time period.

Climate-driven impacts are suspected to be mainly caused by the presence or absence of precipitation. A cross-correlation is made for mean daily precipitation per composite and endmembers of stably classified areas using a lack of zero up to six composites. Further examinations analyse correlations of mean daily precipitation from different time periods and NDVI percentiles (50 %) of stably classified areas of each LCT in order to identify the time period which has the highest correlation values. For that, varying time periods starting at the first or sixteenth of each month and lasting at least a month or more are applied in order to not overemphasise short term weather events or potential incorrect measurement values. For computation of mean daily precipitation, varying start time (1 July–1 April) and varying end time (+14d up to end of vegetation period) is applied to cover a possible shift of precipitation-vegetation relation. The extraction of results is done firstly by specifying a common time period with high correlation values and secondly by specifying individual time periods with the highest correlation values for each LCT. Analogue, the total annual classified area of each LCT is correlated to mean daily precipitation of different time periods to identify the time period which is the most important for the spatial pattern of LCT.

For all available daily precipitation data of two weather measurement stations (Casablanca station at the former airport in western Casablanca and Nouasseur station at the airport “Mohammed V” is utilised, applying only the days with valid values at both measurement stations. Measurement values from the Casablanca station dating 25 August 2001 (199 mm) and from Nouasseur station dating 2 April 2002 (270 mm) were set as missing values after checking hourly data and weather situation of a wider area because of heavy distrust.

3 Results

3.1 Classification results and validation

Annual classification results of 2002 to 2009 are presented in Fig. 5. The classification result of 2001 is part of Fig. 6.

A visual comparison of annual decision tree classifier results to Landsat 7 ETM+ data shows that analogue to other cities, the city of Casablanca is classified as LCT_{svv} in very dense city quarters. Other parts of the town with lower urban density and higher proportions of vegetation are classified as LCT_{sv}, while areas with detached housing in western Casablanca are classified as LCT_{lpv} due to the higher NDVI signal caused by raised vegetation coverage. Thus, the results of the decision tree classifier image cities smaller than they actually are. Contrarily, GLC2000 and PFT overestimate extensions of cities and underestimate extensions of forests and vegetation apart from cropland. Especially GLC2000 misjudges LC: while urban areas are too big, about 86 % of the land surface is classified as croplands, other vegetated surfaces are classified with less than 10 %. Woody areas in the centre of the study region are too small or not classified as forest at all. A reservoir eastern of Rabat and Salé is not classified as a water body, while a classified water body in the south-east does not even exist. Representatively, Fig. 6 shows a comparison for 2001 with special interest in city region of Casablanca: both, PFT LCT “Urban and built-up” and GLC2000 LCT “Cities” cover the airport and the cities of Casablanca and Mohammed as well as vegetated areas in between both cities, furthermore vegetated areas around Casablanca (mainly south-east of the town) and parts of the Bouskoura forest. In total, PFT classification maps of 2001 up to 2004 reflect LC better than GLC2000 in the study region.

Although a combination of three phenologic metrics was applied to determine LCT_{lpv} and to minimise classification errors of the decision tree classifier, few forested areas were classified as LCT_{lpv} and vice versa. We assume the first case occurs in forested areas with less canopy and higher percentage of shrub- and grasslands in the understorey. In the
second case, classification error in terms of LCT$_{for}$ instead of LCT$_{hpv}$ occurs due to the classification method which discriminates sole monocrop-systems by RanV without considering irrigated multi-crop-systems. Thus, few croplands exceed all thresholds and are misclassified as LCT$_{for}$.

The validation of the LC classification map of 2004, resulting from the decision tree classifier for the administrative region of Grand Casablanca results in an overall accuracy of 80.24 % and a Kappa coefficient of 0.74 (Table 2). The low producer’s accuracy of LCT$_{lpv}$ underlines the described problem of a differentiation between cropland and other vegetation if ground truth data are scarce. While outer parts of the airport runways are classified correctly as LCT$_{hpv}$, some of the inner parts are classified as LCT$_{lpv}$. This phenomenon might result from a raised supply of water due to surface runoff from runways which facilitates enhanced growth of vegetation. However, the user’s accuracy of LCT$_{lpv}$ is about 81.4 %.

An additional validation of PFT classification of 2004 results in an overall accuracy of 35.47 % and a Kappa coefficient of 0.25. This may result by overestimated urban surfaces in the study region, as previously described, but also from determination of ground truth areas within the city region. LCT$_{sv}$ and LCT$_{vsv}$ are not distinguished in the PFT classification, which applies LCT “Urban and built-up” for urban areas. Taking this into account, it was questionable to compare validation results from the decision tree classifier to the validation results of PFT classification. Accordingly a further validation was made with summarised LCT$_{sv}$ and LCT$_{vsv}$ and summarised ground truth areas to test if this is

![Fig. 5. Annual land cover classification results of 2002 to 2009.](image)


**Table 2.** Confusion matrix of decision tree classification result of 2004 using ground truth information from the administrative region of Grand Casablanca. Columns show the ground truth pixels of land cover types and rows show the classified land cover types for these pixels: LCT$_{for}$ (Forest), LCT$_{hpv}$ (High-productive vegetation), LCT$_{lpv}$ (Low-productive vegetation), LCT$_{sv}$ (Sparsely vegetated), LCT$_{vsv}$ (Very sparsely vegetated).

<table>
<thead>
<tr>
<th>land cover type</th>
<th>LCT$_{for}$</th>
<th>LCT$_{hpv}$</th>
<th>LCT$_{lpv}$</th>
<th>LCT$_{sv}$</th>
<th>LCT$_{vsv}$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCT$_{for}$</td>
<td>131</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>131</td>
</tr>
<tr>
<td>LCT$_{hpv}$</td>
<td>11</td>
<td>172</td>
<td>67</td>
<td>0</td>
<td>0</td>
<td>250</td>
</tr>
<tr>
<td>LCT$_{lpv}$</td>
<td>6</td>
<td>0</td>
<td>35</td>
<td>2</td>
<td>0</td>
<td>43</td>
</tr>
<tr>
<td>LCT$_{sv}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>132</td>
<td>4</td>
<td>136</td>
</tr>
<tr>
<td>LCT$_{vsv}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>Total</td>
<td>148</td>
<td>172</td>
<td>102</td>
<td>161</td>
<td>9</td>
<td>592</td>
</tr>
</tbody>
</table>
an artefact of the applied method. The validation of decision tree classification (PFT classification) resulted in an overall accuracy of 85.47% (62.67%) and a Kappa coefficient of 0.80 (0.48).

Following overall higher accuracy and better visual validation results, classification maps of the applied decision tree classifier are considered to be more appropriate to describe local LC in the study region than global LC products of PFT classification and GLC2000.

3.2 Characterisation of land cover and land cover dynamics

While comparing classification results the differences are evident: the years 2001, 2002 and 2007 have larger areas of LCTsv and LCTlpv than the years 2003, 2006, 2008 and 2009 with higher proportions of areas classified as LCThpv. All in all, stably classified areas (Fig. 7) cover 13.9% (~3149 km²) of the study region encompassing other areas among the cities of Casablanca, Rabat, Salé and Khouribga as well as the almost entire forested surfaces. Within the nine years, two different LCT were classified for 64.3% of the study region. Fluctuations occur mainly between LCTlpv and LCThpv especially in the western, southern and south-eastern part of the study region (compare Fig. 5). These areas have altitudes that are mostly higher than 280 m a.s.l. or lower than 140 m a.s.l. Further analyses show that 79.6% (95.7%) of areas, which are stably classified as LCThpv, are in altitudes of 70 up to 280 m a.s.l. (1–450 m a.s.l.) while fluctuations of productivity and resulting fluctuations to LCTlpv mainly occur in higher altitudes. Further transitory zones oscillate between LCTlpv and LCTfor on the borders of forested areas, which result mainly from fix classification thresholds in combination with inter-annual ecosystem variability. Minor fluctuations occur between LCTfor and LCThpv due to the explained classification limits of the applied approach concerning forests and crop systems.

Screening LCC, 1.6% (~374 km²) of the study region have changed LCT actually. Spatial pattern of areas with actual LCC (Fig. 8) show a cluster in the south-western part, one in the eastern part and spots around the cities of Casablanca, Rabat and Salé. Having the advantage of high temporal resolution of NDVI TSD, it is possible to determine the instant when the LC changed. The cluster in the south-west indicates mostly changes from LCTlpv to LCThpv with the highest rates of about 101 km² in 2003 and 63 km² in 2004 indicating human-driven impacts. The eastern cluster indicates succession by a shift from LCTsv to LCTlpv with a maximum rate of nearly 63 km² in 2003. Analyses of dots at the city borders show a gain of LCTsv revealing a loss of LCTlpv and indicating urbanisation as human-driven impact. A matrix with a totalised gain and loss of LCT is presented in Table 3.

Extracted yearly endmembers of stably classified areas characterise LCT intra-annually and show the highest NDVI values between the end of February and the beginning of March. Although LCTlpv and LCThpv are distinguished by thresholds of RanV, the mean endmember of LCTlpv (Fig. 9) shows higher values in off-season than the endmember of LCThpv. This shows that areas classified as LCThpv are less vegetated in off-season since higher NDVI values indicate higher proportion of green vegetation. Therefore our approach of collecting croplands in LCThpv is successful, because agricultural lands have fewer canopies than grass- or shrub lands during the off-season due to harvest.

Results of determination of NDVI percentiles of each year are shown in Fig. 10. When combining the results with percentiles of all classified areas of each LCT, within a single year, the NDVI percentiles of stably classified areas from

![Fig. 7. Stably classified areas of the years 2001–2009.](image)

![Fig. 8. Spatial distribution of areas with actual land cover changes. Colours indicate the new land cover type.](image)
Table 3. Identified actual land cover changes (2001–2009) in total and as percentage of the study region. Columns show the gain of land cover types and rows show the loss of land cover types: LCT_for (Forest), LCT_hp (High-productive vegetation), LCT_lp (Low-productive vegetation), LCT_sv (Sparsely vegetated), LCT_vsv (Very sparsely vegetated).

<table>
<thead>
<tr>
<th>land cover type</th>
<th>LCT_for</th>
<th>LCT_hp</th>
<th>LCT_lp</th>
<th>LCT_sv</th>
<th>LCT_vsv</th>
<th>Total loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCT_for</td>
<td>–</td>
<td>0.8 km² (0.00 %)</td>
<td>13.4 km² (0.06 %)</td>
<td>–</td>
<td>–</td>
<td>14.1 km² (0.06 %)</td>
</tr>
<tr>
<td>LCT_hp</td>
<td>2.8 km² (0.01 %)</td>
<td>–</td>
<td>9.1 km² (0.04 %)</td>
<td>0.1 km² (0.00 %)</td>
<td>–</td>
<td>11.9 km² (0.05 %)</td>
</tr>
<tr>
<td>LCT_lp</td>
<td>15.2 km² (0.07 %)</td>
<td>258.5 km² (1.14 %)</td>
<td>–</td>
<td>6.3 km² (0.03 %)</td>
<td>–</td>
<td>280 km² (1.23 %)</td>
</tr>
<tr>
<td>LCT_sv</td>
<td>–</td>
<td>–</td>
<td>63.6 km² (0.28 %)</td>
<td>–</td>
<td>1.1 km² (0.00 %)</td>
<td>64.7 km² (0.29 %)</td>
</tr>
<tr>
<td>LCT_vsv</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3.1 km² (0.01 %)</td>
<td>–</td>
<td>3.1 km² (0.01 %)</td>
</tr>
<tr>
<td>Total gain</td>
<td>17.9 km² (0.08 %)</td>
<td>259.3 km² (1.14 %)</td>
<td>86.1 km² (0.38 %)</td>
<td>9.4 km² (0.04 %)</td>
<td>1.1 km² (0.00 %)</td>
<td>–</td>
</tr>
</tbody>
</table>

Fig. 9. Mean endmembers of stably classified areas 2001–2009.

Each LCT, show similar patterns of inter-annual ecosystem variability during the last decade. Our early hypothesis, that the use of total classified areas of each class does not lead to reasonable results, could not be shown. It seems possible to estimate inter-annual ecosystem variability in both ways. The shift between both approaches results through the classification method: While stably classified areas encompass the core of areas of each LCT, total classified areas additionally encompass all areas with values of applied phenologic metrics near the thresholds, which leads to class fluctuations as previously described.

The analysis of derived phenologic metrics from annual mean endmembers of stably classified areas (Fig. 12) did not result in any trend during the last decade. Therefore, derived metrics are not presented.

3.3 Drivers of land cover

Areas with actual LCC e.g. from LCT_lp to LCT_hp or from LCT_lp to LCT_sv in areas near cities indicate human-driven impacts. Further on, computing RanV_mean of each pixel for the entire time period of 2001 up to 2009 estimates mean human impact referring to agriculture in the study region (Fig. 11). Computation of RanV_mean is independent of used classification technique. Results show in terms of forested areas and areas with scarce vegetation (e.g. cities) low values of RanV_mean less than 0.3. Thus, areas classified as LCT_hp stick out with values of RanV_mean greater than 0.4 which was the threshold for delimiting LCT_hp. In total 52.54 % (11 915 km²) of the study region show a RanV_mean of at least 0.4. Compared to the stably classified areas of LCT_hp (5.11 %, 1.158 km²) the value is over ten times higher. Hence, RanV_mean respectively seems to be an indicator for a simple and rapid assessment of human driven impact and extent of croplands.

Cross-correlation of daily precipitation values measured at both measurement stations results in the highest $r^2$ of 0.48 with no shift in time between both stations. Thus, the local influence causes approximately 50 % of the daily precipitation and superior climate events contribute to the other half. Further correlations of mean daily precipitation for different time periods rising in monthly steps show the highest values of $r^2$ (0.95) for the time period of October-March including a major part of the rainy season. On a yearly base and using mean daily precipitation, the correlation of both

Fig. 10. NDVI percentiles (5 %, 25 %, 50 %, 75 %, 95 %) for stably classified areas (black) and all classified areas (grey) of land cover types forest/ LCT_for (a), very sparsely vegetated/ LCT_vsv (b), low-productive vegetation/ LCT_lp (c), high-productive vegetation/ LCT_hp (d) and sparsely vegetated/ LCT_sv (e).
measurement stations is lower, with $r^2$ of 0.85 showing that local seasonal precipitation plays an important role.

Comparisons of mean daily precipitation per composite and endmembers of stably classified areas for each LCT (Fig. 12) indicate highest results of $r^2$ using an offset of three composites. This means that the vegetation reacts with a lack of nearly one and a half months to precipitation changes: more precipitation causes increasing NDVI values and weakening precipitation causes decreasing NDVI values, resulting in a vegetation rest during the summer months due to the absence of precipitation. Apart from LCT$_{vsv}$ with lower $r^2$ values and applying an offset of three composites, correlations of mean daily precipitation per composite and endmembers of stably classified areas for each LCT range between $r^2$ of 0.2 and $r^2$ of 0.3 for both stations, where all correlation values of Nouasseur station are higher than correlation values of the Casablanca station. In all, the location of the Casablanca station within the city and near to the sea seems to reflect more local urban effects and sea effects than the Nouasseur station further inland. Thus, the data of Nouasseur station are applied for subsequent examinations.

Highest correlation values of mean daily precipitation and NDVI percentile values (50 %) of stably classified areas of each LCT are identified for the common time period from 1 October to the 15 December. Specifying individual time periods with the highest correlation values for each LCT, just LCT$_{vsv}$ has higher correlation values at other time periods from 16 November to the 15 December. Results are significant at the 1% significance level for all LCT but LCT$_{vsv}$ which is significant at the 5% significance level. Correlations are presented in Fig. 13.

In terms of correlations of mean daily precipitation and the yearly extent of classified areas of each LCT, correlations show the highest $r^2$ values for the common time period from the 16 September to the 15 January. Highest correlation values for individual time periods for each LCT differ from this period except LCT$_{typv}$. Concerning LCT$_{vsv}$ and LCT$_{sv}$ the highest correlation values occur for short time frames before the start of the actual hydrological year. In these time periods practically no precipitation occurs and high correlation values are regarded as artefacts. Therefore, the first time period covering at least a part of the actual hydrological year is considered to be the individual time period with highest correlation values. Correlations as well as time periods are presented in Fig. 14. Results are significant at the 1% significance level (Fig. 14e, h, i, j), at the 5% significance level (Fig. 14a, c, d, f, g) and at the 10% significance level (Fig. 14b).
Correlations of mean daily precipitation (MP) and classified area of each land cover type. Best fit for a common time period from 16 September to 15 January (A–E) and best individual fit for each land cover type (F–J) including second best individual fit for land cover type forest (I). DF = Degrees of freedom.

4 Discussion

Understanding current vegetation dynamics and variability of LC in a widely unknown large region needs remote sensed spatio-temporal data to provide cost-effective information (Lunetta et al., 2006; Evrendilek and Gulbeyaz, 2008). Applied MODIS TSD have a moderate grain of spatial resolution. Thus, a considerable number of mixed pixels, which represent more than one ground LCT, are present (Roosta and Saradjian, 2007). Unmixing strategies and sub-pixel algorithms need knowledge about vegetation, LC and its dynamics. Our study extracts information about vegetation phenology as well as current variability and dynamics of vegetation and LC showing that moderate spatial resolution and high temporal resolution of MODIS TSD are sufficient for this purpose even if ground truth data are scarce as in this case.

Although NDVI is a useful parameter to assess LC in arid regions (Huang and Siegert, 2006), NDVI values may be partly overrated in areas with scarce vegetation due to signal influence from soil under the vegetation (Wang et al., 2008; Edwards et al., 1999). Yearly produced LC maps, which showed to describe LC better than global LC maps, document already higher percentage of areas with sparse vegetation in comparison to global LC maps especially at higher altitudes in the south-eastern study region (e.g. 2001 in Fig. 6). Following the above statement, percentage might be in fact higher.

The aim of this study to yield information about vegetation and LC to differ between human-driven change, climate-driven change, vegetation trends and inter-annual ecosystem variability, even if the ground truth knowledge is scarce, is a challenge to the definition of parameters of the decision tree classifier. It was not possible to use varying thresholds throughout the years since knowledge on inter-annual ecosystem variability in the study region was not known before. The adaptation of classification thresholds to precipitation sums was also impossible because we had no knowledge on the vegetation-precipitation interaction. This way the use of fix classification thresholds ensures to extract the core of each LCT as stably classified areas while the remaining areas vary normally between two LCT as described. By examining the stably classified areas it becomes possible to extract information about the LCT and vegetation (dynamics). There is though, a big difference of LCC every year due to inter-annual fluctuations caused by fix thresholds. In terms of LCC detection we had to find a solution to delimit varying areas from areas with actual LCC. Hence we used the requirements that there is exactly one LCC between two LCT within the years to delimit actual LCC and to receive a sufficient likelihood to call it actual LCC. Apart from the many documented LCC due to inter-annual ecosystem variability documenting e.g. dryer and wetter years it was possible to extract 1.6% (~374 km$^2$) of the study region which have actually changed LCT. In total, change detection of LC for areas with actual LCC has raised value in comparison to commonly change detection techniques because it does not compare two different points in time but a gapless time series of data in 16 d steps. Thus, misleading results due to inter-annual variability (as possible by comparing e.g. Fig. 5: 2007 and 2009) are minimised.

The Influence of human and climatic drivers on vegetation dynamics cannot be shown holistically in a single study. Thus, the approach using $\text{RanV}_{\text{mean}}$ shows a human driven impact, given to extraction of areas with actual LCC indicating urbanisation or ploughing up of grassland. Areas of stably classified LCT show climatically driven inter-
intra-annual variability. Indentations of NDVI values after a lack of precipitation from the middle of December to the middle of February in 2004 indicate the influence of droughts (Fig. 12) and underline the previous assumption that climatic drivers depress a sequence of composites instead of single values, as formulated for the choice of applied smoothing technique.

Climatic drivers affect large regions causing lower or higher biomass production which leads to exceedance or lower deviation of classification thresholds explaining fluctuations between two LCT. For instance the percentage of areas classified as LCT$_{hpv}$ is greater in humid years and lower in arid years. Fluctuations between LCT$_{hpv}$ and LCT$_{lpv}$ (especially in the south-western, southern and south-eastern parts of the study region) do not result from local change in vegetation conditions caused by higher application of fertilisers or implementation of irrigation techniques since they occur in a very widespread manner. As explained, the ground truth knowledge is scarce and we had no information about the extent of irrigation or rain dependency of agricultural used areas within the study region. Nevertheless, correlation values of NDVI percentiles (50 %) and mean daily precipitation of LCT$_{hpv}$ are lower than $r^2$ values of LCT$_{lpv}$ (Fig. 13d, e) but are higher than 0.5 for both LCT. We conclude that both LCT are mainly rain-fed in the study region.

The analysis of correlation values of mean daily precipitation and percentage of yearly classified area of each LCT show a positive correlation of LCT$_{for}$ and LCT$_{hsv}$ as well as a negative correlation for the other LCT. This means the more precipitation occurs, the less percentage of LCT$_{svv}$, LCT$_{sv}$ and LCT$_{lpv}$. In other words, areas classified as LCT$_{sv}$ and LCT$_{lpv}$ increase to some extent through areas fluctuating respectively from LCT$_{svv}$ and LCT$_{sv}$ but at the same time they decrease to the bigger extent through areas fluctuating respectively to LCT$_{for}$ and LCT$_{hpv}$. In terms of fluctuations from LCT$_{hsv}$, we have to differ between areas fluctuating to LCT$_{for}$ which represent the above mentioned transitory zones near the borders of forested areas and fluctuations to LCT$_{hpv}$ due to increased productivity. Thus, percentage of areas classified as LCT$_{hpv}$ range from 16 % in dry years to 70 % in wetter years documenting high fluctuations of productivity. Overall, correlations of mean daily precipitation and percentage of yearly classified area of each LCT are medium up to high. The best fit period of LCT$_{for}$ from 1 February to the 15 May might result from the classification method which uses $EndV$ as one threshold. Thus, the period documents the importance of precipitation falling in later periods to exceed the classification threshold. The second highest correlation occurs for LCT$_{for}$ for the time period from 16 September to the 15 May having $r^2$ of 0.73. This result is significant at the 1 % significance level and underlines the importance of precipitation during the whole vegetation period.

Results of correlations of mean daily precipitation and NDVI percentiles (50 %, presented in Fig. 13) show high and very high $r^2$ values up to 84 %. In all cases the variance in the NDVI not explained by the variance in precipitation is rather small. Unexplained variance could stem from errors in both the NDVI and precipitation time series, or from further predictors missing in the regression model. Although we cannot exclude that other predictors would also be able explaining larger parts of the variance in the NDVI, we can state that precipitation is an appropriate predictor, since there is not only statistical evidence by high correlation but also bio-physical reasons behind the regression model.

We could not detect statistically significant trends in the areas covered by the LCT (using a 10 % significance level) over the studied period. The same holds for the mean NDVI within the stably classified areas. Kennedy et al. (2009) analysed data up to 2007/2008 proclaiming data records of MODIS vegetation to be currently too short to capture long-term changes. Our study could be interpreted such that either there are no significant changes or the time series are still too short as concluded by Kennedy et al. (2009).

5 Conclusions

Our approach shows that moderate spatial resolution and high temporal resolution of MODIS TSD are sufficient for drawing conclusions on LC dynamics and its driving forces in a widely unknown region with sparse validated ground truth knowledge. By deriving phenological parameters from the gapless TSD from the 29 September 2000 to the 29 September 2009 composition of yearly LC maps is possible. Results of applied decision tree classifier, showed to determine LC in the study region better than global LC classifications of PFT and GLC2000. Validation ends in an overall accuracy of 80.24 % and a Kappa coefficient of 0.74 for the administrative region of Grand Casablanca. Drawing conclusions on vegetation dynamics and its drivers is possible. This study contributes to a differentiation between human-driven changes, climate-driven changes and inter-annual ecosystem variability.

Areas of stably classified LCT show climatically driven inter- and intra-annual variability with indicated influence of droughts. The presented approach to determine human-driven influence on vegetation dynamics by RanV$_{mean}$, results in a more than ten times larger area than compared as stably classified areas. It respectively seems to be a good indicator for a simple and rapid assessment of human-driven impacts and extents of annual crop systems. It has still has to be additionally proved for perennial and multi-crop systems.

Change detection based on yearly LC maps shows a gain of LCT$_{hsv}$ (cropland) of about 259.3 km$^2$.

Vegetation in the study region showed to be highly dependent on precipitation, in terms of climate-driven impact, especially croplands, are supposed to be mainly rain-fed. Mean daily precipitation from 1 October to the 15 December and NDVI percentiles (50 %) of LCT show highest correlation
values. Results are significant at the 1% and 5% significance level as shown. Comparable, correlations of mean daily precipitation from 16 September to 15 January and percentage of yearly classified area of each LCT show highest correlation values for a common time period. All in all, an offset of nearly 1.5 months is detected between precipitation rates and NDVI values.

However, the precipitation data originated from two measurement stations delivering the data of a comparatively small area. Therefore, further studies are needed to examine the relation of vegetation and precipitation in the study region in more detail.

The vegetation phenology as well as current variability and dynamics of vegetation and LC, including its climatic and human drivers, could be extracted for a widely unknown region. However, statistically significant trends in vegetation dynamics could not be discovered. We conclude that identification, understanding and knowledge about vegetation phenology, and current variability of vegetation and LC, as well as prediction methods of LCC can be improved using multi-year MODIS NDVI TSD. This study enhances the comprehension of current land surface dynamics and variability of vegetation and LC in north-western Morocco.

Acknowledgements. This study was funded by the German Federal Ministry of Education and Research (BMBF) within the project “Urban Agriculture as an Integrative Factor of Climate-Optimised Urban Development, Casablanca” within the megacity research programme “Research for the Sustainable Development of Megacities of Tomorrow, Focus: Energy- and climate-efficient structures in urban growth centres”. We thank the reviewers for their valuable support, Mr. Vincent Simonneaux for his advice in terms of landscape classification in Morocco, Mr. Marco Otto and Mr. Christopher Brown for revisions and Mr. Jens Rogée for preparing NCDC station data.

Edited by: A. Neftel

References


