Monitoring vegetation dynamics in semi-arid rangelands of South Africa by fusion of high temporal resolution MODIS data with high spatial resolution RapidEye data

ANDREAS TEWES[1], FRANK THONFELD[2], MICHAEL SCHMIDT[3], ROELOF J. OOMEN[4], XIAOLIN ZHU[5], GUNTER MENZ[6], JÜRGEN SCHELLBERG[7]

Zeitreihen, die aus räumlich und zeitlich hoch aufgelösten Fernerkundungsdaten gewonnen werden, erfassen Vegetationsdynamiken auf lokaler Skala. Allerdings bleibt die Erstellung, durch unterschiedliche Faktoren bedingt, schwierig. In dieser Studie wurde der ESTARFM-Algorithmus (Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model) auf multi-spektrale Bilder von zwei semi-ariden Weidegebieten in Südafrika angewendet. MODIS 8-Tages Komposite mit 250 m räumliche Auflösung und RapidEye-Bilder mit 5 m wurden miteinander verschnitten, um 38 synthetische Bilder mit RapidEye-Auflösung im Zeitraum Juni 2011 bis Juli 2012 zu erzeugen. Wir evaluierten die Ergebnisse, indem wir die synthetischen Bilder mit tatsächlichen zeitnahen Aufnahmen verglichen. Dies erfolge Band für Band sowie für das jeweils gesamte Bild. Die Ergebnisse zeigen, dass die ESTARFM-Vorhersagen eine hohe Genauigkeit aufweisen, mit einem Determinationskoeffizienten für das rote Band von $0.80 < R^2 < 0.92$, für das nahe Infrarot von $0.83 < R^2 < 0.93$, einem durchschnittlichen relativen Fehler zwischen 6% und 12% für das rote Band und 4% bis 9% für das nahe Infrarot. Heterogene Vegetation auf unter-MODIS Auflösung wird adäquat erfasst: ein Vergleich der NDVI-Zeitreihen aus MODIS, RapidEye und ESTARFM-Daten zeigt, dass die charakteristischen phänologischen Dynamiken der unterschiedlichen Vegetationstypen gut wiedergegeben werden. Weiterhin können wir zeigen, dass die ESTARFM-Zeitreihe Mehrinformationen bietet. Wir folgern, dass der ESTARFM-Algorithmus es erlaubt, synthetische Bilder mit kombinierter hoher zeitlicher und räumlicher Auflösung zu erstellen. Dies bietet Informationen über Vegetationsdynamiken, die für ökologische oder landwirtschaftliche Studien benötigt werden.

1) Andreas Tewes, Institut für Nutzpflanzenwissenschaften und Ressourcenschutz (INRES), Universität Bonn, Katzenburgweg 5, 53115 Bonn; Email: atewes@uni-bonn.de
2) Frank Thonfeld, Zentrum für Fernerkundung der Landoberfläche, Universität Bonn, Walter-Flex-Str. 3, 53113 Bonn; Email: frank.thonfeld@uni-bonn.de
3) Dr. Michael Schmidt, Remote Sensing Research Group, Geographisches Institut, Universität Bonn, Meckenheimer Allee 166, 53115 Bonn; Email: michael.schmidt@uni-bonn.de
4) Roelof J. Oomen, Institut für Nutzpflanzenwissenschaften und Ressourcenschutz (INRES), Universität Bonn, Katzenburgweg 5, 53115 Bonn; Email: roelof.oomen@uni-bonn.de
5) Xiaolin Zhu, Department of Geography, The Ohio State University, 1036 Derby Hall, 154 North Oval Mall, Columbus, OH 43210, USA; Email: zhu.381@osu.edu
6) Prof. Dr. G. Menz, Remote Sensing Research Group, Geographisches Institut, Universität Bonn, Meckenheimer Allee 166, 53115 Bonn; Email: g.menz@geographie.uni-bonn.de
7) PD Dr. Jürgen Schellberg, Institut für Nutzpflanzenwissenschaften und Ressourcenschutz (INRES), Universität Bonn, Katzenburgweg 5, 53115 Bonn; Email: j.schellberg@uni-bonn.de

1 Introduction

Time series derived from remote sensing data exhibiting both high temporal and high spatial resolution capture surface changes and vegetation dynamics and the possibility of relating both to measurements collected in the field. However, the use of temporally dense data at high spatial resolution still is difficult with optical imagery due to low repetition rates or interruption by clouds, cloud shadows, haze or smoke from fires (WATTS ET AL., 2011; SCHMIDT ET AL., 2012). This limits the use for detection of rapid surface changes (GAO ET AL., 2006). Medium to low spatial resolution instruments such as MODIS or AVHRR cover the earth with a frequency that is suitable for monitoring land cover dynamics at large scales. Spatial resolution however is not satisfactory for tracking changes at ecologically relevant resolutions, especially in heterogeneous landscapes. Fusion of remotely sensed data from different sensors with different spatial and temporal characteristics is an efficient solution to enhance the capability of remote sensing for the monitoring of land surface dynamics at varying scales (ZHU ET AL., 2010; KIM AND HOGUE, 2012). Result is a combined time series that exhibits both a high temporal and high spatial resolution (SCHMIDT ET AL., 2012).

Our motivation for this study was to find a method that overcomes the low temporal resolution of time series and yet yields high spatial resolution that enables to capture heterogeneous vegetation dynamics. To achieve this we used ESTARFM to blend high spatial resolution RapidEye data and high temporal resolution MODIS data. ZHU ET AL. (2010) developed the ESTARFM (Enhanced Spatial and Temporal Reflectance Fusion Model) algorithm to combine the spatial resolution of Landsat imagery with the temporal resolution of coarse-resolution sensors such as MODIS. We are not aware of any study on the fusion of RapidEye with a coarse-resolution sensor. The research questions underlying this study are:

- Is the ESTARFM algorithm applicable for generating time series using the combination of RapidEye and MODIS?
- Is a time series combining real RapidEye with ESTARFM-computed synthetic images appropriate for detecting highly dynamic vegetation changes at different small scale bush density classes in semi-arid rangelands in South Africa?

2 Material and Methods

2.1 The Study Area

The study area is situated in the Northern Cape Province of South Africa and lies within the southern outreaches of the Kalahari. The landscape is a savannah biome; vegetation cover comprises of a woody layer of mainly seasonally deciduous trees and shrubs, furthermore a ground layer of annual and perennial grasses and some forbs (VAN ROOYEN ET AL., 2001; PALMER AND AINSLIE, 2005). The study area lies within the southern African summer rainfall zone (approximately October - early April), with dry winters. Rainfall in the Kalahari is highly variable and uncertain. Mean annual precipitation is around 220 - 440 mm (MUCINA AND RUTHERFORD, 2006). Low precipitation has favoured the use of land for pasture farming. Rainfall, fire and grazing are the three key driving forces for vegetation dynamics in the Kalahari ecosystem, as in many other semi-arid rangelands (TRODD AND DOUGILL, 1998; VAN ROOYEN ET

AL., 2001). We formed two subsets for ESTARFM input to reduce data load and shorten image calculations, instead of using the largest overlap of all images (hereafter referred to as Kathu Bushveld and Kuruman Thornveld).

2.2 Data

2.2.1 The RapidEye Satellite System

RapidEye is a commercial remote sensing mission, consisting of five identical satellites placed in a single sun-synchronous orbit. Each satellite carries a five band multi-spectral optical imager that captures radiation in the blue, green, red, red edge and near-infrared spectral range. At nadir, ground sampling distance is 6.5 m and 5 m after orthorectification and resampling, respectively (TYC ET AL., 2005; RAPIDEYE AG, 2012). The RapidEye images were orthorectified and atmospherically corrected using the automated processing chain CATENA developed and maintained by DLR (KRAUß ET AL., 2013). To georeference and orthorectify the RapidEye data, accurate orthophotos were used as reference data.

2.2.2 The MODIS Instrument

To minimize the spatial resolution and acquisition time differences between the MODIS and RapidEye data, the MOD09Q1 product was selected. This product provides surface spectral reflectance as it would be measured at ground level in the absence of atmospheric scattering or absorption. It comprises the MODIS bands 1 and 2 (red and NIR) on the Terra satellite at a 250 m resolution in an 8-day composite product, where each pixel contains the best possible observation during an 8-day period.

2.3 The ESTARFM Algorithm

The ESTARFM algorithm requires two pairs of fine-resolution and coarse-resolution images as input, with each pair (t_m and t_n) captured at the same date. For the desired prediction date t_p, one coarse-resolution image is required. ESTARFM yields a synthetic image at the prediction date t_p with the same spatial resolution as the fine-resolution input images (ZHU ET AL., 2010).

During the ESTARFM computation, four major steps take place: (1) Two fine-resolution images are used to search for pixels similar to the central pixel in a moving search window, (2) the spectral and spatial distance between each similar pixel and the predicted pixel are used to calculate weights of each similar pixel w_i, (3) a linear regression of the coarse-resolution values in the two observed pairs (t_m and t_n) against the fine-resolution values of the similar pixel is used to determine the conversion coefficient v_i, which is then used to convert the change found from the coarse-resolution images to the fine resolution images, and (4) calculation of fine-resolution reflectance from coarse-resolution image at prediction date t_p. The algorithm is explained in detail in ZHU ET AL. (2010).

2.4 ESTARM Implementation

Synthetic images at RapidEye resolution were computed for both study subsets using ten RapidEye scenes (captured between 2011-06-28 and 2012-07-18) as input, combined with the corresponding MODIS images acquired at the closest possible date (i.e. forming ten image pairs). For image pair formation, two MODIS MOD09Q1 products have not been available in 2012, and so the daily surface reflectance product MOD09GQ was used instead. Here, the same acquisition date as of the corresponding RapidEye image was chosen. Coarse-resolution information at prediction dates (t_p) was provided by 38 MODIS images within the time period 2011-07-04 until 2012-07-03. For prediction, three missing MOD09Q1 datasets in 2012 were substituted by MOD09GQ products. The second image pair (i.e. at date t_n) continuously served as first image pair in the next computation step (i.e. at date t_m). Two temporally closest image pairs were always provided as bracketing images for ESTARFM prediction. Areas affected by clouds were assigned no-data values, ignored by ESTARFM in the subsequent calculation and excluded from further study. By checking the quality band of every MODIS product, we found that all MODIS images were provided at best quality. To reduce algorithm computing time, sensor bands not suitable for the purpose of this study were excluded from the datasets. Accordingly, only red and NIR bands were included for NDVI calculation (for MODIS input either from MOD09Q1 or MOD09GQ).

2.5 Accuracy Assessment of ESTARFM Images

The validation of the accuracy of the synthetic images was undertaken using a set of independent images. Nine RapidEye scenes, each captured at a date bracketed by two consecutive RapidEye images used for ESTARFM processing, were used to compare pixel values in the predicted image with the corresponding pixels in the reference RapidEye image, on a band by band basis (following WALKER ET AL. 2012). Based on a scheme proposed by WALD ET AL. (1997) and THOMAS AND WALD (2007), we employed the (1) bias and its value relative to the mean value of the observed image, (2) the standard deviation of the difference image in relative value (hereafter refered to as to per-pixel level of error) and (3) the coefficient of determination (R^2) as quantitative criteria.

2.6 Bush density estimation

Grass and bush canopies show differences in their reflectance, caused by different life-forms and phenology (DOUGILL AND TRODD, 1999). This can result in mixed signals at coarser resolutions, making it difficult to compare values from different sensors at different resolutions. Also, there may be situations where two contrasting changes compensate each other (GAO ET AL., 2006), but may be visible in fine-resolution imagery. We used an unsupervised k-means cluster algorithm to classify a composited NDVI time series, derived from 15 (Kathu Bushveld) and 16 (Kuruman Thornveld) RapidEye images used within our study, into the classes “grass cover”, “bush cover” and “asphalt, bare soil, non-dynamic surface”. We then calculated the percentage area classified as “bush cover” falling into each MODIS pixel at 250 m resolution.

2.7 Monitoring vegetation dynamics

To investigate the usability of ESTARFM for vegetation monitoring purposes, parallel time series of the Normalized Difference Vegetation Index (NDVI) were computed based solely on RapidEye, MODIS and ESTARFM images for the different bush density classes. Despite its appearance to be a poor indicator of vegetation biomass for low ground cover, a number of studies confirm the usefulness of NDVI for monitoring vegetation dynamics in arid and semi-arid environments (e.g. SCHMIDT AND KARNIELI (2000); WEISS ET AL. (2004)).

3 Results

The ESTARFM algorithm yielded 76 synthetic images (38 images for each subset) at RapidEye resolution, covering the time period June 2011 until July 2012. Scatter plots (not displayed here) show that the relationship between the observed and predicted pixel values is closely adherent to the 1:1 line in all cases in both the red and NIR band, proving that reflectance is accurately predicted by ESTARFM for each image. The slopes of the regression lines range from 0.8975 to 1.118 for the red band and from 0.9128 to 1.053 for the NIR band. This indicates only little differences between the observed and predicted images. ESTARFM both under- and overestimated mean reflectance in the red and NIR band, commonly between 0% and 16%. Generally, better prediction results were found for the NIR bands than for the red bands. Per-pixel levels of error vary between 0.06 and 0.012 (standard deviation of 6% to 12%) for the red band and between 0.04 and 0.08 for the NIR band, also indicating a better performance of ESTARFM in the longer waveband.
Correlation between observed and predicted images is generally high. For all dates, the NIR band yielded better results than the red band (Red: $0.80 < R^2 < 0.92$, NIR: $0.83 < R^2 < 0.93$). Prediction accuracy was lowest for the scene predicted for 2012-01-17 while precision improves for the antecedent and precedent scenes, with the best results found for the scenes predicted for 2011-09-24 and 2012-06-30.
Since the ESTARFM predicted red and NIR bands were found to behave well over time, the data was used to calculate NDVI time series suitable for comparison to NDVI time series derived from RapidEye and MODIS images. No temporal filtering was applied. Figure 1 shows MODIS, ESTARFM and RapidEye NDVI time series derived from averaging all pixels at RapidEye resolution found within MODIS pixel size extents that contain less than 5% (and thus grass cover by the majority; first row),between > 50% and ≤ 65% (second row) and more than 95% of bush cover (third row).

Fig. 1: Time series of mean NDVI values calculated for all MODIS pixel size extents classified as containing ≤5% (first row), >50% - ≤65% (second row) and >95% bush cover (last row). The left column shows results from the Kathu Bushveld subset, the right column from the Kuruman Thornveld subset.

4 Discussion

We found strong correlations between the observed and predicted RapidEye reflectance values, for both the red and NIR bands (R^2 values range from 0.80 to 0.92 for the red band and from 0.83 to 0.93 for the NIR band). Results suggest that precision is best during phases of low vegetation dynamics, and deteriorates during phases of strong vegetation growth (December, January). Notably, predictions generally achieved better results for the NIR band than for the red band. This might be due to greater influence of atmospheric effects at shorter wavelengths. Although we took great care in finding and masking cloud pixels in the RapidEye images, remaining deviating pixel values might have affected ESTARFM predictions and results of the accuracy assessment, caused by variations in aerosol loadings and water vapour variations not taken into account by the atmospheric correction procedure.

In almost all cases, the relative mean bias between the observed and predicted reflectance values was unequal to 0, with sets of both positive and negative numbers. We interpret this as a signal of noise likely due to atmospheric and BRDF effects.

The band by band comparison of observed and predicted images for different bush density classes shows strong relationships, suggesting a good performance of ESTARFM at sub-MODIS scale. The comparison of the NDVI time series derived from MODIS, ESTARFM and RapidEye data shows that the ESTARFM-predicted temporal sequence reproduces the characteristic phenological dynamics of different vegetation communities well. The good agreement of RapidEye and ESTARFM-derived curves for heterogeneous vegetation cover (bush cover between 50% and 65%) indicates good prediction behaviour at MODIS sub-pixel scale. The comparison furthermore indicated that the fusion approach provides additional information that would not have been captured by either MODIS or RapidEye series alone.

The MODIS MOD09GQ product is generated by selecting pixels matching different criteria from a number of images collected over a period of 8 days. It is thus probable that the reflectance values for adjacent pixels were collected at different illumination and viewing angles and may be subject to high differences. This may result in non-uniform conversion coefficients that would not have been observed if the entire image had been generated from observations captured on the same day.

Due to the huge differences in spatial resolution of MODIS and RapidEye data, the images could not be co-registered. This might have introduced an error of unknown extent, in case good sub-pixel accuracy was not achieved by using the georeferences provided with the datasets.

5 Conclusions

As related to our research questions, we conclude:

- The ESTARFM algorithm can be used for generating a time series for monitoring vegetation dynamics at RapidEye spatial and MODIS temporal resolution.
- The ESTARFM derived NDVI time series reproduces the characteristic phenological dynamics of different vegetation types well and inherits more information than either MODIS or RapidEye alone.

- Prediction accuracy of ESTARFM is good for the red and NIR bands during phases of little vegetation dynamics, but deteriorates during times of quick vegetation growth. The NIR band generally yields better prediction results than the red band.
- ESTARFM shows strong prediction performance at sub-MODIS scale in heterogeneous vegetated areas, with best results during phases of low vegetation dynamics.

6 References

DOUGILL, A., TRODD, N., 1999. Monitoring and modelling open savannas using multisource information: analyses of Kalahari studies. Glob. Ecol. Biogeogr. 8, 211 – 221.

GAO, F., MASEK, J., SCHWALLER, M., HALL, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. Geosci. Remote Sens. IEEE Trans. 44, 2207 –2218.

KIM, J., HOGUE, T.S., 2012. Evaluation and sensitivity testing of a coupled Landsat-MODIS downscaling method for land surface temperature and vegetation indices in semi-arid regions. J. Appl. Remote Sens. 6, 063569–1–063569–17.

KRAUß, T., D’ ANGELO, P., SCHNEIDER, M., GSTAIGER, V., 2013. The Fully Automatic Optical Processing System CATENA at DLR. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XL-1/W1, 177 – 183.

MUCINA, L., RUTHERFORD, M.C. (Eds.), 2006. The Vegetation of South Africa, Lesotho and Swaziland. South African National Biodiversity Institute, Pretoria.

PALMER, A.R., AINSLIE, A.M., 2005. Grasslands of South Africa, in: Suttie, J.M., Reynolds, S.G., Batello, C. (Eds.), Grasslands of the World. Food and Agriculture Organization of the United Nations, Rome, pp. 77–120.

RAPIDEYE AG, 2012. Satellite Imagery Product Specifications Version 4.1.

SCHMIDT, H., KARNIELI, A., 2000. Remote sensing of the seasonal variability of vegetation in a semi-arid environment. J. Arid Environ. 45, 43 – 59.

SCHMIDT, M., UDELHOVEN, T., GILL, T., RÖDER, A., 2012. Long term data fusion for a dense time series analysis with MODIS and Landsat imagery in an Australian Savanna. J. Appl. Remote Sens. 6, 063512–1–063512–18.

THOMAS, C., WALD, L., 2007. Comparing distances for quality assessment of fused images, in: Bochenek, E. (Ed.), New Developments and Challenges in Remote Sensing. Millpress, Varsovie, Pologne, pp. 101–111.

TRODD, N.M., DOUGILL, A.J., 1998. Monitoring vegetation dynamics in semi-arid African rangelands: Use and limitations of Earth observation data to characterize vegetation structure. Appl. Geogr. 18, 315 – 330.

TYC, G., TULIP, J., SCHULTEN, D., KRISCHKE, M., OXFORT, M., 2005. The RapidEye mission design. Acta Astronaut. 56, 213 – 219.

VAN ROOYEN, N., BEZUIDENHOUT, H., DE KOCK, E., 2001. Flowering plants of the Kalahari dunes. Ekotrust.

WALD, L., RANCHIN, T., MANGOLINI, M., 1997. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. Photogramm. Eng. Remote Sens. 63, 691–699.

WALKER, J.J., BEURS, K.M. DE, WYNNE, R.H., GAO, F., 2012. Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. Remote Sens. Environ. 117, 381 – 393.

WATTS, J.D., POWELL, S.L., LAWRENCE, R.L., HILKER, T., 2011. Improved classification of conservation tillage adoption using high temporal and synthetic satellite imagery. Remote Sens. Environ. 115, 66 – 75.

WEISS, J.L., GUTZLER, D.S., COONROD, J.E.A., DAHM, C.N., 2004. Long-term vegetation monitoring with NDVI in a diverse semi-arid setting, central New Mexico, USA. J. Arid Environ. 58, 249 – 272.

ZHU, X., CHEN, J., GAO, F., CHEN, X., MASEK, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sens. Environ. 114, 2610 – 2623.