

Assessment of mountainous ecosystems in Central Asia by using remote sensing based vegetation cover fractions

D. Klein^{a, *}, S. Asam^a, U. Gessner^b, S. Dech^{a, b}

^a Department of Remote Sensing, Institute of Geography, University of Wuerzburg, Am Hubland, 97074 Wuerzburg, Germany
– (doris.klein, sarah.asam)@uni-wuerzburg.de

^b German Aerospace Center, German Remote Sensing Data Center, 82230 Wessling, Germany – (stefan.dech, ursula.gessner)@dlr.de

Abstract – Mountainous regions are especially vulnerable to climate and land use change. For a monitoring of subtle changes detailed and continuous land cover information is necessary. Therefore vegetation cover fractions are derived for the Naryn river catchment in Kyrgyzstan using a remote sensing multi-scale approach. First a very high resolution QuickBird image is classified into discrete classes using a hybrid approach. This result is aggregated to the spatial resolution of Landsat and used as training data for a random forest and Landsat data to derive sub-pixel fractions of woody vegetation, herbaceous vegetation and bare soil. These results are again spatially aggregated and training pixels are extracted to derive finally vegetation fractions for the whole river catchment using a random forest with moderate resolution MODIS data. The results have a RMSE between 8 and 17 and show that woody cover can be found mainly on northern slopes, while bare soils are predominant on southern slopes.

Keywords: fractional vegetation cover, regression trees, multi-scale, ecosystem monitoring, mountains, Central Asia

1. INTRODUCTION

Mountainous regions are fragile ecosystems and especially vulnerable to climate change and land use pressure (Messerli, 2004). Therefore, a monitoring of gradual changes in land cover and an assessment of their implications is necessary, which implies the availability of detailed land cover information. Remote sensing data are ideal for monitoring because they are objective, repeatable and spatially continuous. Remote sensing based land cover information can be used directly to assess land dynamics and further more as input parameter e.g. for hydrological models. Vegetation plays an important role for the hydrological cycle regarding infiltration, interception and transpiration. At the same time vegetation cover is important for stabilizing the soil especially in mountainous regions to prevent erosion. For modeling usually categorical land cover classification is used to describe the vegetation cover and its properties. This means that land cover classes are already predefined and do not necessarily meet the model requirements. Vegetation cover fractions could be a more suitable data set giving spatially detailed continuous information. At the same time such a dataset is thematically flexible, as values can be grouped and regrouped into classes depending on the purpose of investigation. Thus, for monitoring of vegetation cover as well as a potential improved model input parameter three relevant continuous layers of land cover are derived on two different spatial scales: woody vegetation, herbaceous vegetation and bare soil. In this paper we

present the methodology as well as the accuracy of the results and the spatial distribution of the vegetation cover within the Central Asian mountains.

2. STUDY AREA

The study site comprises the catchment of the Naryn river in Kyrgyzstan, Central Asia, a tributary to Syr Darya and thus to the Aral Sea. The area is dominated by the high mountains of Tien Shan rising up to almost 7500 m in the east (Wilson, 1997). The focus area of the field work and the high resolution classification was situated in the middle of the Naryn river catchment in south-eastern Kyrgyzstan, close to the city of Naryn within Naryn Oblast (see Fig. 1). Precipitation is low in the area, in the city of Naryn (2041 m) less than 300 mm rainfall with high interannual variations is measured. However,, its spatial distribution is highly dependent on topography. Accordingly, different vegetation zones can be found along the elevational gradient and on differently exposed slopes including mountainous xerophilous vegetation, semi-desert, montane and meadow steppe with shrub vegetation, insular coniferous forests, alpine meadows and cryophytes (Agakhanyants, 1981).

3. DATA AND METHODS

The derivation of vegetation cover for a large area is accomplished by a multi-scale methodology (Gessner, 2009, Hansen, 2002) integrating spatially very high, high and moderate resolution images (Fig. 1). In this study we used three different remote sensing data sets: A QuickBird image with the spatial resolution of 0.6 m, a Landsat 5 TM image with the spatial resolution of 30 m (both from August 2009) and the MODIS vegetation indices and spectral data time series for the year 2009 with a resolution of 232 m.

A three step-procedure was applied to achieve vegetation cover fractions on two spatial resolutions. First, the spatially very high resolution QuickBird image with an extend of 80 km² is classified by using categorical land cover classes. This classification is aggregated to the spatial resolution of 30 m and the cover fractions for woody, herbaceous and bare surface are calculated accordingly. By using the resulting cover fractions as training data for a regression tree (Breimann, 2001) the cover fractions can be extrapolated to the area of 180 x 190 km² covered by the Landsat image. On this scale the same aggregation and upscaling procedure as for QuickBird is applied by generating training samples from the Landsat result and calculating cover fractions for the whole river catchment using MODIS time series data for the year 2009.

Before processing, the QuickBird and Landsat images were corrected for topographic effects using the ATCOR method

(Richter, 2009) and orthorectified using the ASTER digital elevation model. To reduce the data the spatial resolution of QuickBird was resampled to 1.2 m. Also the highly correlated blue band was substituted by information on contrast of the infrared band obtained through a Sobelfilter.

The MODIS data pre-processing included an interpolation of noisy and cloudy pixels according to the MODIS quality flags using the software TISEG (Colditz, 2008). Additionally the data had to be mosaicked and subsetting according to the study site.

For the classification of the very high resolution image field data were sampled and geolocated via GPS during a campaign in August 2009 according to the land cover classification scheme LCCS developed by FAO (Gregorio, 2007). This assessment was supplemented by hemispherical photos, which were classified and used for cover fraction derivation (Rich, 1990). These field measurements were then used in a hybrid classification approach as training data. First, by using an object oriented approach, trees (respectively forest stands) as well as shadows and shrubs which can be identified as single objects are classified using support vector machines. Secondly, grasslands were further refined according to their herbaceous cover fraction into 5 classes representing 20 percent steps on the very high resolution scale. Water and settlement, snow and clouds as well as cloud shadows were masked (Asam, 2010).

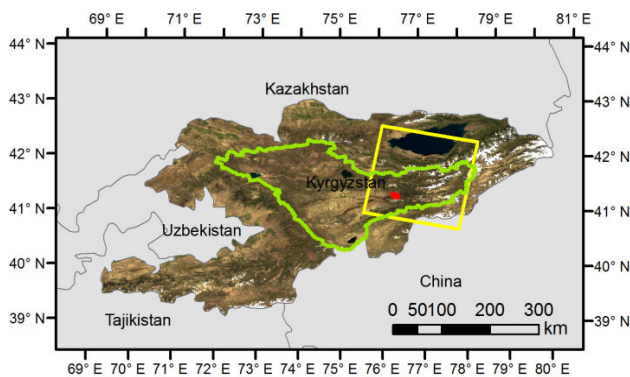


Figure 1. Location of the QuickBird data (small red polygon), Landsat data (yellow rectangle) and the Naryn river catchment (green polygons) within Kyrgyzstan overlaid on MODIS. In the northern part lake Issyk Kul can be seen.

Next, training pixels with a spatial resolution of 30 m were selected by stratified sampling from the aggregated QuickBird classification. The training pixels, together with the spectral features and indices (Normalized Difference Vegetation Index, Soil Adjusted Vegetation Index, Brightness, Greenness, Wetness, and Variance) of the Landsat image were then used to estimate cover fractions with random forest regression trees. For every cover type (woody vegetation, herbaceous vegetation and bare soil) an individual random forest was constructed using the training data and applied to the entire Landsat image.

The results at Landsat scale were again aggregated to the spatial resolution of the MODIS time series. Training and validation samples were extracted and random forest regression trees were applied based on statistical features (median, mean, standard deviation, range) derived from MODIS vegetation indices as well

as the spectral information (blue, red, near infrared and middle infrared wavelength) for the year 2009.

4. RESULTS AND DISCUSSION

4.1 Validation of the results

Woody, herbaceous and bare cover fractions could be estimated on two different spatial scales (QuickBird, Landsat and MODIS) covering, at MODIS resolution, an area as large as 590 by 450 km². A first plausibility analysis of the results was performed by adding the values of the cover fractions of all cover types, which ideally should sum up to 100 for natural areas. Overall, the mean sum for the Landsat extend is 98.5 with a standard deviation of 7.1. The slight underestimation is reasonable given the fact, that artificial areas like settlements and roads are part of the land cover but not included in any of the three calculated cover layers.

However when calculating the sum over the MODIS derived cover fractions for the Naryn catchment the mean value is 113.1 with a standard deviation of 9.1. The overestimation on this scale can be explained by the relative concentration of training pixels in the eastern part of the catchment (Fig. 1). Thus, for a better accuracy new training data from other sites, capturing the variability of land cover within the catchment should be included. This is confirmed when examining the Root Mean Square Error (RMSE) for the two different scales (Tab. A).

Table A. Estimations of root mean square error for the fractional vegetation cover at spatial resolutions of 30 m and 232 m.

	Woody	herbaceous	Bare
Landsat (30 m)	14.7	17.4	7.8
MODIS (232 m)	16.5	13.7	11.6

The validation of the cover fractions is based on the fraction values derived from the higher scale data. The best accuracy is achieved for the Landsat scale of 'bare' cover fractions with a RMSE of 7.8. The lowest RMSE is achieved for Landsat scale for herbaceous cover with an RMSE of 17.4. One reason for the lower accuracies of the herbaceous cover might be the higher phenological variability which might be better captured by using multi-temporal Landsat images (Gessner et al., 2009). This seems to be confirmed by the decrease of the RMSE when deriving herbaceous cover by using the complete annual time series of MODIS data which captures the seasonal variation. The higher RMSE than in other comparable studies using this approach in Kenya and Namibia (Gessner et al, 2009) might also be due to the impact of topographic shading effects.

Few other regional or local remote sensing classifications exist for this area which could be used for comparing the results of this study. One classification in the Naryn area reveals accuracies of 67 % for valleys and 97 % for mountain areas but this classification uses only discrete classes (Dernedde, 2009). In general, the accuracy assessments of the fraction cover presented give a reasonable indication of the quality of the results. However, they are based on validation samples only from the subset close to Naryn and therefore need to be further expanded spatially for a better representation of the whole area.

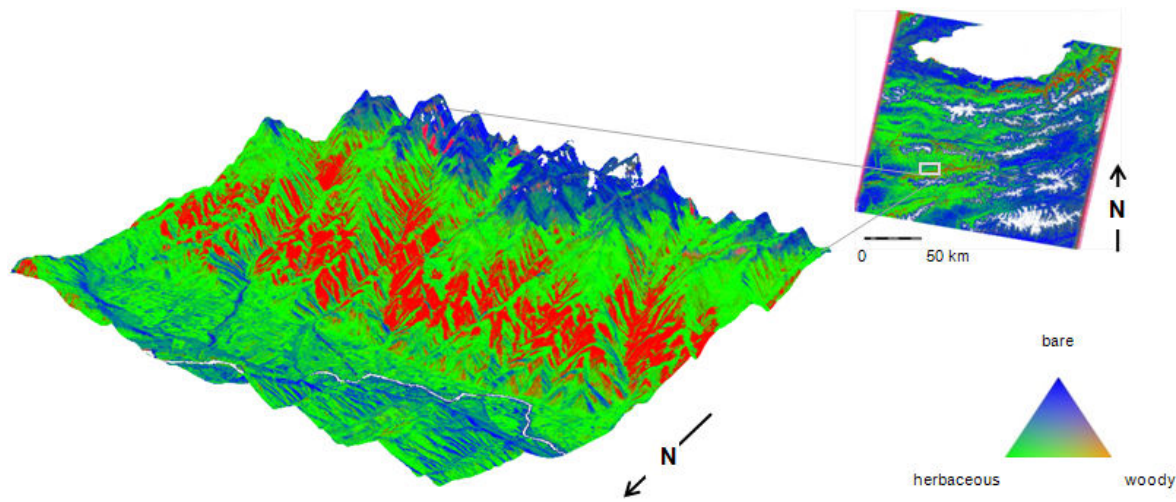


Figure 2. A false color composite displaying woody (red), herbaceous (green) and bare (blue) cover fractions derived from Landsat and a subset close to Naryn overlaid on an ASTER DEM.

4.2 Distribution of vegetation cover

The spatial distribution of the vegetation cover with respect to topography was investigated based on the Landsat results. It generally corresponds well with the field observations. Tree cover can be found mainly on the more humid north facing slopes, which are less exposed to insolation especially at this latitude. The drier southern slopes on the other hand have a higher degree of uncovered soil, as well as areas in altitudes over 4000 m. Herbaceous vegetation cover however is distributed rather homogeneously (Fig. 2).

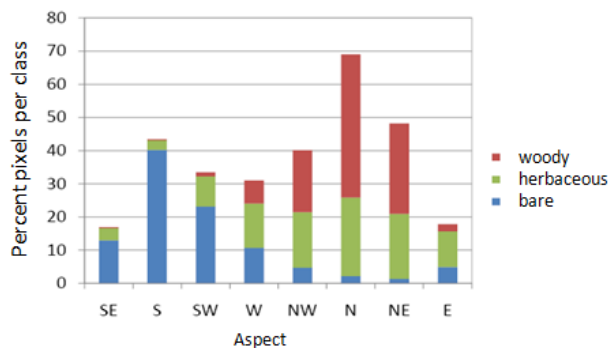


Figure 3: Distribution of homogenous woody, herbaceous and bare cover fractions in relation to aspect

The relation between aspect and the distribution of woody vegetation as well as bare soil cover is confirmed by a quantitative analysis. A statistic of the mean cover of bare soil, woody, and

herbaceous and bare vegetation, stratified according to different aspects (NW-NE, NE-SE, SE-SW and SW-NW) and elevation classes (every 1000 m), show that the overall percentage of bare soils on southern slopes is higher than on northern slopes, especially for the elevation of 2000 m to 4000 m where more than 80% of the slopes can be found. On the southern slopes bare soil cover is 41.2% for the elevation of 2000 m to 3000 m and 46.3 % for 3000 m to 4000 m compared to 25 % and 36 % on the north facing slopes respectively. In contrast woody vegetation covers 18% of the northern slopes at an elevation between 2000 m and 3000 m and 9 % at 3000 m to 4000 m compared to 6.5% and 6.6 % respectively on the southern slopes.

This pattern is even more pronounced if only areas covered by one vegetation type (with values higher than 90%) are considered. For this analysis, random homogeneous pixels were selected by stratified sampling to achieve an equal numbers of pixels per cover (Fig.3).

These results on distribution of vegetation cover can contribute to erosion modeling, indicating that southward facing slopes could be more prone to erosion because of a higher fraction of exposed bare soils than north facing slopes with a higher degree of woody and also herbaceous cover.

5. CONCLUSION AND OUTLOOK

The results of this multi-scale approach proof that it is possible to derive woody, herbaceous and bare cover fractions for large areas also in mountainous regions by using high resolution classifications as training data. These data can be used to analyze the spatial distribution of the vegetation structure in relation to topography. However for improving the accuracy of the estimation additional training areas (from the high and very high resolution images) are needed to capture the area wide land cover variability,

i.e. the spectral variation and seasonal variability of the different vegetation cover within the catchment. Therefore, further analysis will expand the training and validation data to additional high resolution classifications. Furthermore cover fractions for other years shall be calculated in order to analyze changes in land cover over the years.

data”, *Remote Sensing of Environment*, vol. 97, p.p. 322-336, 2005.

REFERENCES

- O.E. Agakhanyants, “Aridye gory SSSR (The Arid Mountains of the USSR)”, Mysl Publishers, Moscow, 270 p., 1981.
- S. Asam, D. Klein, U. Geßner, C. Conrad, C. Beiekuhnlein, S. Dech, „Ableitung des Vegetationsbedeckungsgrades aus multiskaligen Fernerkundungsdaten für hydrologische Modellierung in Zentralasien“, J. Strobl, T. Blaschke, G. Griesebner,(Eds.), “Angewandte Geoinformatik 2010. Beiträge zum 22. AGIT-Symposium Salzburg”, p.p. 284 -289, 2010.
- L. Breiman, J.H. Friedman, R.A. Olshen and C.J. Stone, “Classification and Regression Trees”, Chapman & Hall/CRC, 1984.
- R. Colditz, C. Conrad, T. Wehrmann, M. Schmidt, S. Dech, “TiSeG: A Flexible Software Tool for Time-Series Generation of MODIS Data Utilizing the Quality Assessment Science Data Set”, *IEEE Transactions on Geoscience and Remote Sensing* 46, p.p. 3296-3308, 2008.
- Y. Dervede, “Analyse der Landnutzungsänderung in Naryn, Kirgistan, unter Verwendung von Landsat TM und ETM+ Daten“, Diploma thesis, University Marburg, 2009.
- U. Gessner, D. Klein, C. Conrad, M. Schmidt and S. Dech, “Towards an automated estimation of vegetation cover fractions on multiple scales: Examples of Eastern and Southern Africa”, 33rd International Symposium on Remote Sensing of Environment, Stresa, Italy, 4. – 8. May 2009.
- M. Hansen, R.S. DeFries, J.R.G. Townshend, R. Sohlberg, C. Dimiceli and M. Carroll, “Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data”, *Remote Sensing of Environment*, vol. 83, p.p. 303-319, 2002.
- M. Hansen, R. DeFries, J. Townshend, M. Carroll, C. Dimiceli, and R. Sohlberg, “Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS Vegetation Continuous Fields Algorithm”, *Earth Interactions*, vol. 7, p.p. 1-15, 2003.
- C. Huang and J.R.G. Townshend, “A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover”, *International Journal of Remote Sensing*, vol. 24, p.p. 75-90, 2003.
- A. D. Gregorio and L.J.M. Jansen, “Land cover classification system (LCCS)”, Classification concepts and user manual, FAO, Rome, 2007.
- B. Messerli, D. Viviroli, R. Weingartner, „Mountains of the world: vulnerable water towers for the 21st century”, *Ambio, Spec* No 13, p.p. 29-34, November 2004.
- P. Rich, “Characterizing plant canopies with hemispherical photographs”, *Remote Sensing Reviews*, vol. 5, p.p. 13- 29, 1990.
- R. Richter, “Atmospheric / Topographic Correction for Airborne Imagery”, DLR report DLR-IB 565-02/09, Wessling, Germany, 2009.
- R. Wilson, “Livestock, pastures, and the environment in the Kyrgyz Republic, Central Asia”, *Mountain Research and Development*, vol. 17, p.p. 57-68, 1997.
- M. Xu, P. Watanachaturaporn, P.K. Varshney and M.K. Arora, “Decision tree regression for soft classification of remote sensing