From Remote Sensing to Web Mapping - integrated geoinformation for land use monitoring in Central Asia

Dimo Dimov[1*], Lucia Morper-Busch[1], Sarah Schönbrodt-Stitt[1], Christopher Conrad[1]

[1] Department of Remote Sensing, Institute of Geography and Geology, University of Würzburg
Oswald-Külpe-Weg 86, 97074 Würzburg, Germany

* Contact: dimo.Dimov@uni-wuerzburg.de

Abstract

Irrigation agriculture is of a major socio-economic importance for the Central Asian countries Afghanistan, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan. Main aspects to be considered refer to remote agricultural areas, the lack of harmonized agricultural statistics and the necessity of intensification of agriculture combined with water scarcity as well as unsustainable water use management. Monitoring of agricultural land and water use also aims to develop a scientifically sound database for this region and can thus support a sustainable and effective irrigation management. The integrated use of Remote Sensing and web GIS is a very valuable technique for crop monitoring. Thus, within the frame of the Regional Research Network "Central Asian Water" project (CAWa), a web-based mapping platform, the "Water Use Efficiency Monitor in Central Asia" (WUEMoCA), is developed. In this study, we focus on the retrospective land use classification of the Aral Sea Basin (ASB) for the years 2000-2016. We utilize MODIS data and present a machine learning framework for the unsupervised extraction of NDVI crop signatures which are fed into a supervised classifier based on Gradient Boosting, Random Forest and Support Vector Machines. The overall accuracy score is 76%. The classification results are further ingested into the web GIS backend system based on PostGIS spatial database and GeoServer technology. Resulting land use classification is then visualized on a client-based web mapping platform.

Study area and data

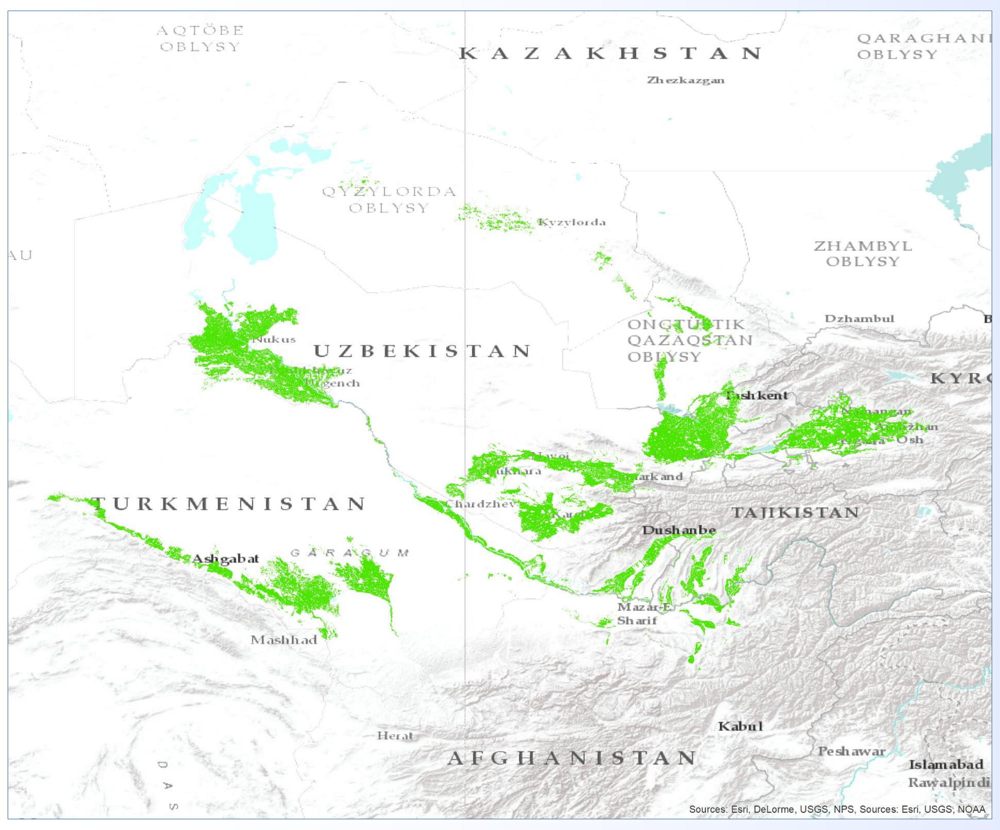

Figure 1:The irrigation zones in the ASB.

The study area covers the irrigation zones in the Aral Sea Basin (ASB) within the Central Asian countries. MOD09Q1 surface reflectance data from the MODIS Terra platform was obtained and annual NDVI time series stacks with 30 8-day intervals (from March to October) were created for the observation period from 2000 to 2016.

Methodology

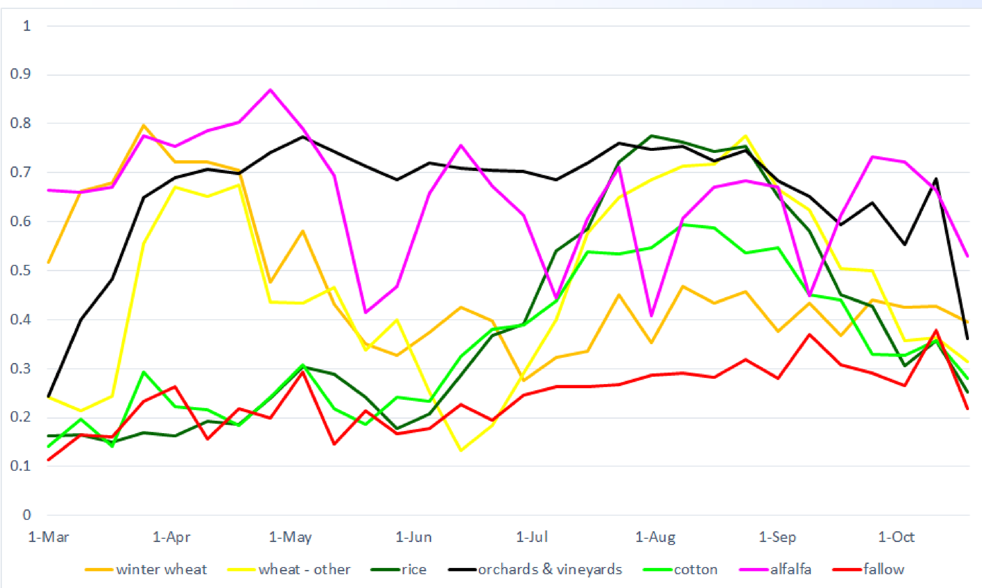

Figure 2: NDVI time series signatures of the key land use classes.

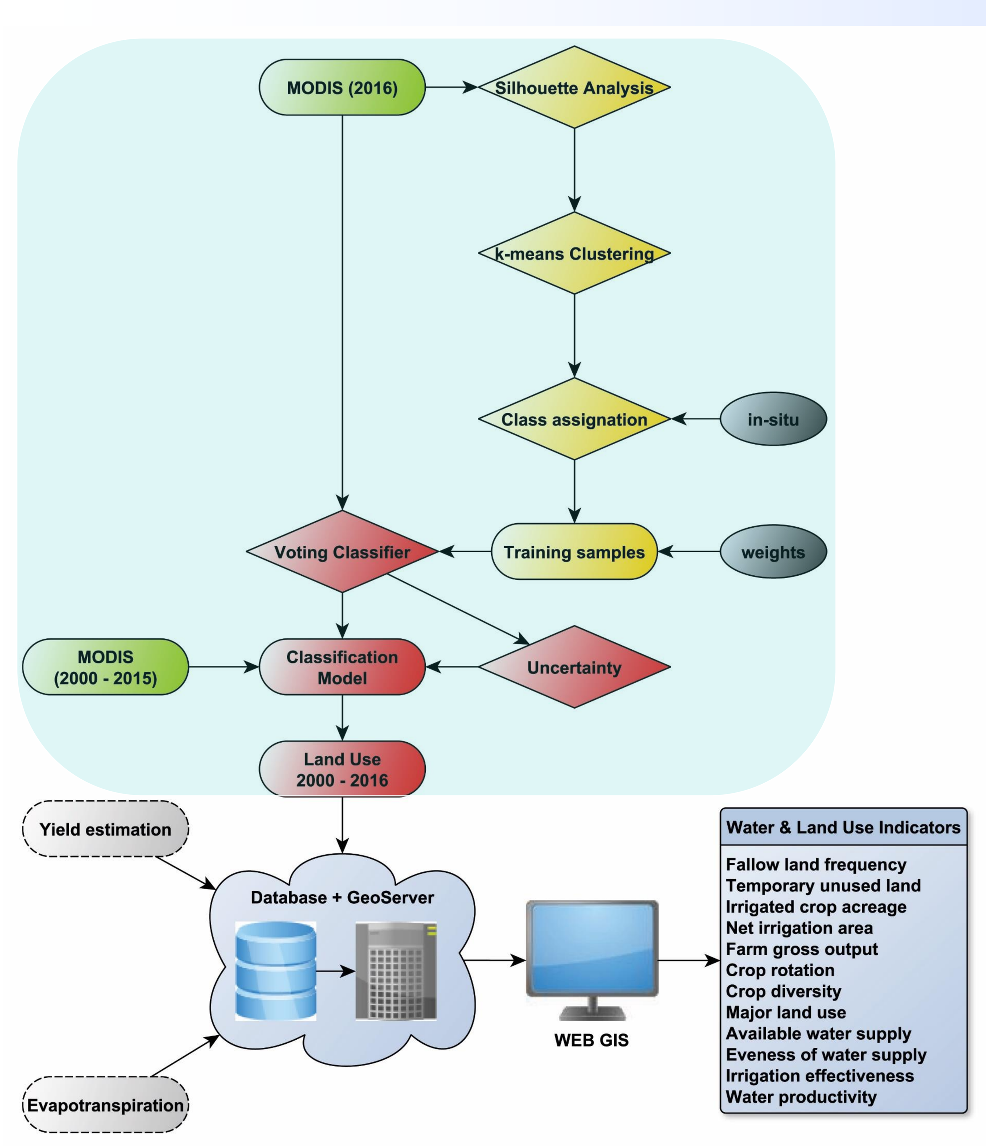

Figure 3: Overall framework of the WUEMoCA monitoring tool. The land use classification process is highlighted in light blue.

The land use classification framework is based on the extraction of training samples from unsupervised *k-means* clustering. The optimum number of *k* is determined through a Silhouette analysis which measures the mean intra-cluster distance. A subset of the clusters which correspond to the target land use classes is selected based on in-situ data and the temporal NDVI signatures. They are fed as training samples into a Voting classifier based on *Gradient Boosting*, *Random Forest* and *linear SVM,* each one of them handling bias and variance of the dataset differently.
In soft voting, the class labels are predicted based on the probability *P* for each classifier of the ensemble. Pixels with *P* lower than 40% are classified as "other". This pre-trained model from from a single year of 2016 is applied on all MODIS data stacks for the entire observation period.

Results

The classification reveals land use maps for the whole observation period. All results are ingested into the PostGIS database and spatially aggregated on administrative boundaries such as Water Use Association (WUA), district and province. This enables the generation of an adequate land use indicator system for an improved land use management. Relevant indicators will provide information on crop diversity, crop rotation frequency, fallow land frequency, irrigated crop acreage and major land use.

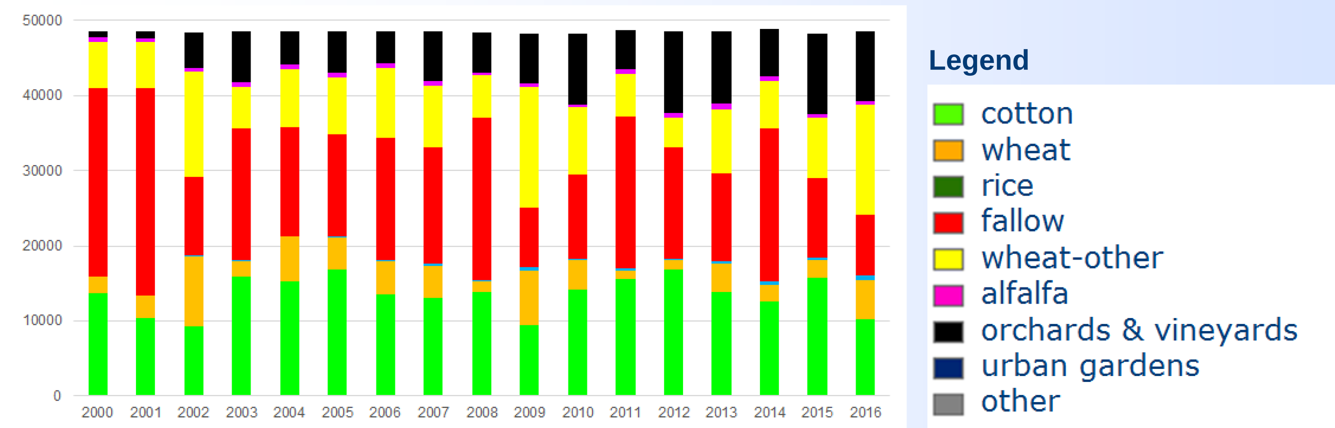

Figure 4: Land use aggregation on province level (Samarkand Province). Y-axis indicates acreage in hectares.

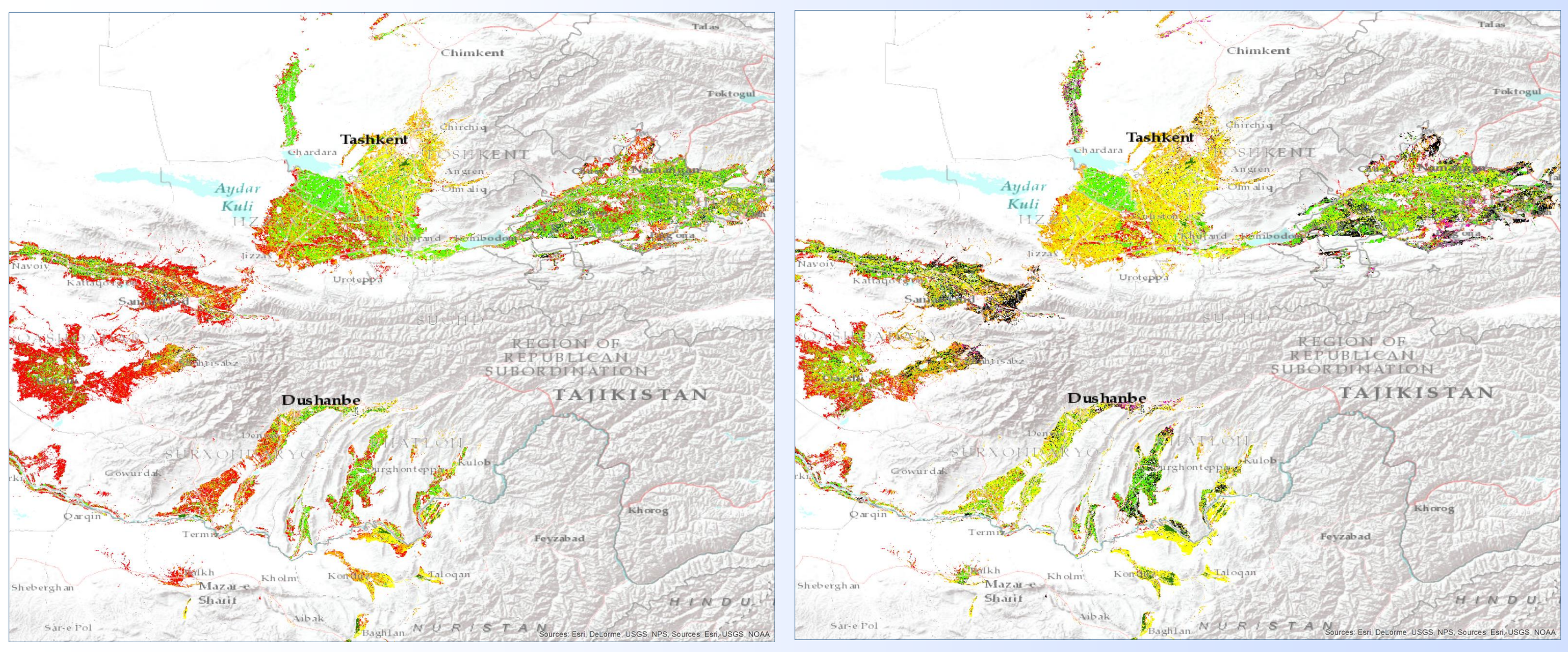

Figure 5: Resulting land use maps in the Eastern ASB for the years 2001 (left) and 2016 (right)

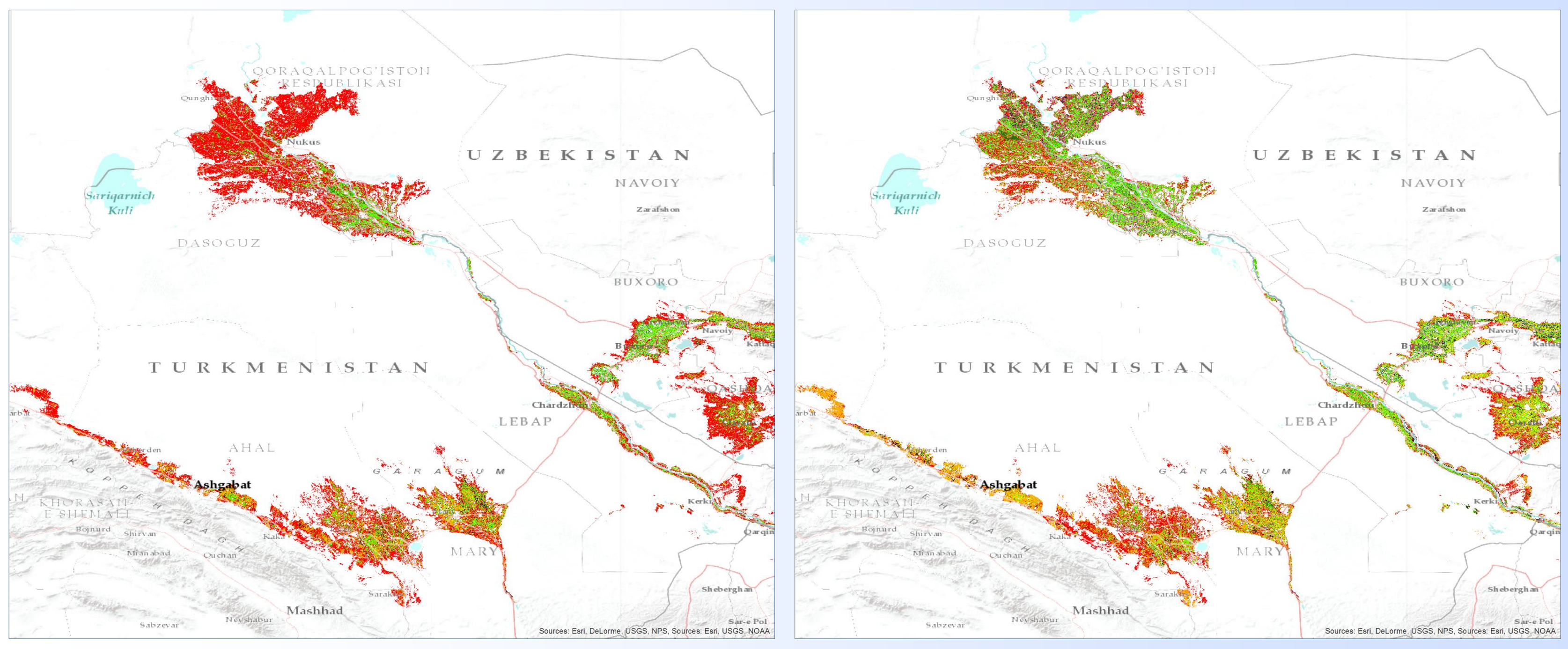

Figure 6: Resulting land use maps in the Western ASB for the years 2001 (left) and 2016 (right)

Challenges and Advantages of Copernicus program

Challenges

- Transferability of models for land use classification, yield estimation and calculation of actual evapotranspiration from MODIS to Sentinel-3 data
- Lower spatial resolution of 300m GSD might increase spatial uncertainty
- Transition from MODIS to Sentinel-3 to ensure data and project continuity
- Large swaths and high temporal resolution requires Big Data handling

Advantages

- Spectral similarity of NIR bands 841-876nm (MODIS) vs. 845-885nm (Sentinel-3) supports model transferability
- 21 spectral bands in 300m resolution (Sentinel-3) vs. 2 spectral bands in 250m resolution (MODIS) may improve spectral and temporal separation of large time series
- Faster provision of near-real time data through Copernicus data hub

References

[1] J. Friedman, "Stochastic gradient boosting: nonlinear methods and data mining", Computational Statistics and Data Analysis, vol. 38, pp. 367-378, 2002.

[2] B. Tso and P.M. Mather, "Classification methods for remotely sensed data", CRC Press, 2009.

[3] M. M.-T. Chang, "Intelligent Choice of the Number of Clusters in K-Means Clustering: An Experimental Study with different Cluster Spreads", Journal of Classification, vol. 27, no. 1, pp. 3-40, 2010.

Acknowledgements:
The authors would like to thank the Scientific Information Center of the Interstate Water Commission of Central Asia (SIC ICWC), the Khorezm Rural Advisory Support Service (KRASS) and the German Federal Foreign Office for project funding (grant nr. AA7090002).