# Postprint: Estimation of Chlorophyll Content in Typical Oasis Vegetation in Arid Regions Based on Sentinel-2 Data

**Authors:** Gu Feng, Ding Jianli, Ge Xiangyu, Gao Shibao, Wang Jingzhe, Ding Jianli

**Date:** 2019-09-11T00:00:00+00:00

## Abstract

Taking the Weigan River-Kuqa River Oasis (Weiku Oasis) as the study area, this study employs the Random Forest regression algorithm, which possesses distinct advantages in machine learning, to estimate and validate the relative chlorophyll content (soil and plant analyzer development, SPAD) of leaves from four typical vegetation types (cotton, reed, poplar, and jujube) within the oasis. Initially, 23 broad-band spectral indices sensitive to chlorophyll were extracted from Sentinel-2 imagery, which is rich in spectral information at the "red edge" region, as well as from its derived first-order and second-order derivative images. Three soil parameters influencing plant growth (soil water content, soil organic matter, and soil electrical conductivity) were incorporated as feature variables affecting leaf SPAD. Based on these feature variables, three distinct SPAD estimation models were established for each vegetation type, thereby enabling monitoring of vegetation chlorophyll within the oasis. The results demonstrate:

Vegetation indices extracted from first-order derivative images play a more important role in SPAD estimation models compared to those from original spectral images, ranking at the top in the importance ordering of the Random Forest algorithm; The SPAD estimation models for all four vegetation types achieved satisfactory results, particularly remarkable for reed leaves, with the coefficient of determination ($R^2$) reaching 0.926; Comparative analysis of model predictive capability under the three schemes reveals that Scheme 3 (including soil parameters) exhibits excellent predictive ability ($2.143 <$ relative percent deviation (RPD) $< 2.692$), with the predictive capability ranking as: Scheme 3 > Scheme 1 > Scheme 2, and soil properties show strong nonlinear correlation with model prediction results. Sentinel-2 data possess ideal potential for estimating oasis vegetation chlorophyll content, providing an efficient, low-cost, and potentially high-precision approach for chlorophyll content estimation, which can offer references for achieving more effective protection and

management of agriculture and ecosystems in arid region oases.

## Full Text

# Estimation of Chlorophyll Content of Typical Oasis Vegetation in Arid Area Based on Sentinel-2 Data

**GU Feng[1,2,3], DING Jian-li[1,2,3], GE Xiang-yu[1,2,3], GAO Shi-bao , WANG Jing-zhe[1,2,3]**

[1]College of Resources and Environmental Sciences, Xinjiang University, Urumqi 830046, Xinjiang, China
[2]Key Laboratory of Wisdom City and Environmental Modeling under the Department of Education of Xinjiang Uygur Autonomous Region, Urumqi 830046, Xinjiang, China
[3]Key Laboratory of Oasis Ecology under the Ministry of Education, Xinjiang University, Urumqi 830046, Xinjiang, China
School of Earth Sciences, Zhejiang University, Hangzhou 310027, Zhejiang, China

## Abstract

The Ogan-Kuqa River Delta Oasis, a typical oasis in the arid zone of China, was selected as the study area. The Random Forest method, which has comparative advantages in machine learning, was chosen to model and estimate the relative chlorophyll contents (SPAD values) of leaves from four representative vegetation types (cotton, reed, poplar, and jujube). Twenty-three broadband spectral vegetation indices sensitive to chlorophyll content were obtained based on the reflectance of original Sentinel-2 imagery with rich spectral information in the "red edge" bands. These vegetation indices were extracted again from the first-derivative Sentinel-2 image and second-derivative Sentinel-2 image. Three soil parameters related to vegetation growth—soil moisture content (SMC), soil organic matter (SOM), and electrical conductivity (EC)—were incorporated as characteristic variables affecting SPAD values. According to these variables, three modeling schemes were developed to monitor the SPAD values of vegetation leaves in the oasis.

The results showed that: (1) Vegetation indices obtained from the first-derivative image played a more important role than the original vegetation indices in the SPAD estimation model. (2) The SPAD-RF regression model based on Sentinel-2 satellite data could effectively monitor the SPAD values of leaves across the four vegetation types, with the estimation model for reed leaves achieving an $R^2$ of 0.926. (3) By analyzing and comparing model prediction capability across the three schemes, Scheme 3 (including soil parameters) demonstrated excellent performance (2.143 < relative percentage deviation

(RPD) < 2.692), and the prediction capability ranked as Scheme 3 > Scheme 1 > Scheme 2. There was a significant nonlinear correlation between soil properties and model prediction results. Overall, Sentinel-2 data shows great potential for predicting chlorophyll content of oasis vegetation, providing an efficient, low-cost, and potentially high-precision solution for SPAD estimation.

**Keywords:** oasis; Sentinel-2 data; SPAD; chlorophyll; vegetation index; random forest; Xinjiang

## References

[11] Shou LN, Jia LL, Chen XP, et al. Using high-resolution satellite image to evaluate nitrogen status of winter wheat in the North China plain[J]. Journal of Plant Nutrition, 2007, 30(10): 1669-1680.

[12] Song XY, Huang WJ, Wang JH, et al. Preliminary application of ASTER images in winter wheat quality monitoring[J]. Transactions of the CSAE, 2006, 22(9): 148-153.

[13] Cheng ZQ, Zhang JS, Meng P, et al. Hyperspectral estimation of chlorophyll content of typical oasis vegetation in arid area[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(8): 264-271.

[22] Green M. A relative error bound for balanced stochastic truncation[J]. IEEE Transactions on Automatic Control, 1988, 33(10): 961-965.

[23] Filella I, Penuelas J. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status[J]. International Journal of Remote Sensing, 1994, 15(7): 1459-1470.

[24] Shang J, Liu J, Ma B, et al. Mapping spatial variability of crop growth conditions using RapidEye data in Northern Ontario, Canada[J]. Remote Sensing of Environment, 2015, 168: 113-125.

[25] Gitelson AA, Gritz Y, Merzlyak MN. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves[J]. Journal of Plant Physiology, 2003, 160(3): 271-282.

[26] Jiang Z, Huete AR, Didan K, et al. Development of a two-band enhanced vegetation index without a blue band[J]. Remote Sensing of Environment, 2008, 112(10): 3833-3845.

[31] Chen Q, Huang JF, Wang RC, et al. Comparative assessment of two methods for estimation of soil organic carbon content by Vis-NIR spectra in Xinjiang Ebinur Lake Wetland[J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(18): 162-168.

[33] Shang J, Liu J, Ma B, et al. Mapping spatial variability of crop growth conditions using RapidEye data in Northern Ontario, Canada[J]. Remote Sensing of Environment, 2015, 168: 113-125.

[34] Cheng ZQ, Zhang JS, Meng P, et al. Hyperspectral estimation of chlorophyll content of typical oasis vegetation in arid area[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(8): 264-271.

*Note: Figure translations are in progress. See original paper for figures.*