

# MAPPING LEAF AREA INDEX OVER SEMI-DESERT AND STEPPE BIOMES IN KAZAKHSTAN USING SATELLITE IMAGERY AND GROUND MEASUREMENTS

*Pavel Propastin and Martin Kappas*

Department of Geography, Georg-August University Göttingen, 37077 Göttingen, Germany;  
{ppropas / mkappas@uni-goettingen.de}

## ABSTRACT

Maps of peak seasonal leaf area index (*LAI*) were produced using the Normalised Difference Vegetation Index (*NDVI*) from SPOT VEGETATION (VEG) satellite at 1 km resolution over a large region in the semi-arid zone of Kazakhstan. Ground measurements of *LAI* were acquired using indirect and direct techniques across a 150·150 km<sup>2</sup> large region. A Landsat Enhanced Thematic Mapper (ETM+) scene at 30 m spatial resolution was used to locate ground sites and to facilitate spatial scaling to 1 km pixels. A high-resolution *LAI* map retrieved from the Landsat ETM+ data was aggregated to 1 km resolution and afterwards used as reference data. The methods tested for transfer function between ETM+ *LAI* and SPOT-VEG were ordinary least squares (OLS) regression, non-linear regression, and reduced major axis (RMA) regression. In this paper, final maps of peak season *LAI* at a 1 km resolution are presented after an assessment of their accuracy using the aggregated ETM+ *LAI* scene. The most appropriate results were attained by RMA. Advantages and shortages of the used regression approaches were analysed and discussed. Errors were mostly caused by uncertainties in co-registration of Landsat ETM+ and SPOT-VEG images as it was demonstrated by a pixel degradation experiment. The methodology presented in this paper can serve as a basis for generation of medium- and coarse-resolution *LAI* satellite products for wide areas of Central Asia and Kazakhstan. The study exposed a general transferability of the developed model for *LAI* estimations at coarser scales. The 1000 m SPOT-VEG model has proved to be fully suitable for utilising with the SPOT-VEG data with resolution of 2 km.

## INTRODUCTION

Most of the models estimating net primary productivity and photosynthesis use as a key input variable the biophysical characteristic of the earth's vegetated surfaces: the leaf area index (*LAI*) (1,2). The leaf area index is a dimensionless quantity, which is defined as one-side green leaf surface area in a vegetation canopy per unit ground surface. The site-specific or stand-specific value of the leaf area index can be estimated with high accuracy using direct and indirect measurement methods including destructive sampling and optical observation techniques (3,4,5). Estimations of *LAI* over large areas were prompted by the need of *LAI* information for biophysical modelling at the scales from sub-regional to global. A number of remote sensing techniques have been developed enabling estimation of *LAI* over large areas. Multi-spectral and multi-temporal information acquired from different types of earth satellites has been proved to be the most effective means for estimates of *LAI* at all scales.

Generally, most of the methodologies of remotely sensed estimations use a conversion of radiometric data into *LAI*. The procedure of the conversion involves a calibration of satellite radiometric data versus ground measurements of *LAI* using a certain transfer function. There are three categories of the transfer function but the most used ones are physical models and empirical methods. Physical models are based on the radiative transfer equation and have been used relatively rarely at local and regional scales. They have gained importance for estimations of the global products of *LAI* with a spatial resolution of 1 km, which are being operationally produced from the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) and have been released to the public recently. Retrieval of the *LAI* variables from MODIS data is based on the algorithms that

contain biome-specific constants such as leaf angle distribution, canopy heterogeneity, and soil and wood optical properties. The theoretical basis of the algorithm is given in (6). A Look-Up-Table (LUT) method is used to inverse the three-dimensional radiative transfer problem (7). Nevertheless, the application of this product at the regional scale is difficult; it requires extensive validation of the product to the regional specifications. The validation work is going on in collaboration with investigators from around the world. In recent literature (8,9,10), there are more and more reports about activities of MODIS *LAI* validation in different earth regions and various biomes. These activities are an integral part of algorithm refinement and product improvement.

Empirical models are considered to be relatively easy to implement and can provide accurate results by application at the scale from local to regional. They have already demonstrated their effectiveness in numerous studies throughout the world. Empirical methods imply the use of any type of regression, artificial neural network or geostatistics to transfer radiometric data into *LAI*. The relationship between various multi-spectral vegetation indices (VI) and in situ measurements is used to establish a model (11,12,13).

Kazakhstan is the world ninth large country (2.7 million km<sup>2</sup>) that turned independent after the collapse of the Soviet Union in 1991. More than 80% of its territory are drylands covered by grassland and scrubland vegetation. Since these areas represent a huge carbon sink and play an important role in the global carbon dioxide cycle, their research has got more and more into the focus of science, policy and economy over the past years (14,15,16,17). Because of the enormous area of these territories and the low level of infrastructure development, a ground survey approach can gain only limited knowledge. Therefore, remote sensing methods are of great importance for investigating and developing the natural resources of this region (18,19).

In this paper, we report about derivation and validation of a data set of *LAI* at 1 km spatial resolution over a large semi-arid region in Kazakhstan. Major issues facing the retrieval of a moderate-resolution *LAI* product such as (a) spatial scaling from ground plot's size to 1 km pixel, (b) choice of a regression model, and (c) accuracy assessment for moderate-resolution *LAI* images are addressed in the underlying paper. Three empirical models between ETM+ and SPOT-VEG *NDVI* were investigated. The associated advantages and drawbacks are discussed and illustrated. The focus of this study is on peak season *LAI* maps. Seasonal variations of *LAI* will be considered in our subsequent studies.

## STUDY AREA

The study region is located between 48°37'N - 49°22'N and 72°25'E - 74°07'E in the middle part of Kazakhstan and occupies an area of approximately 10·10<sup>5</sup> km<sup>2</sup>. Geographically it belongs to the south-western margin of the Kazakh Low Hills. The morphology, vegetation and climate characteristics of the study area are typical for the semi-desert and dry steppe zones in Kazakhstan. In terms of surface structure the study area is divided into two large regions: a plateau of rolling upland in the south, in the west, and in the north with average elevations between 300-600 metres; hills in the central and north-eastern parts with elevations of 600-800 metres. The region is well drained and intersected by a number of small rivers and streams, which are filled with water during 1-2 months at the beginning of the warm period and are dry during the rest of the hydrological year. The soils covering most of the territory are light chernozems and kastanozems with a humus content of 3-6%. In the southern part of the study region soil types of the desert zone such as serozems (grey desert soils) cover wide areas (20).

The climate is semi-arid to arid with an average annual precipitation of 250-300 mm and potential evaporation of 1000-1200 mm. The most part of precipitation falls during the warm period from March to October with two peaks at the end of May-June and the end of July. The temperature amplitude is relatively high: average January temperature is below -12°C while average July temperature is above 26°C (21).

In the vegetation cover xerophyllous bunch grasses in the northern part of the study area and dwarf shrubs in the south of the area are dominant. The grassland is dominated by genera *Festuca* and *Stipa*. Few euryxerophilous forbs occur; the co-dominants are dwarf shrubs of the genus *Artemisia* and sometimes of other genera, particularly *Anabasis* and *Salsola*. Species diversity is about 12-15 species in a square metre. The height of the canopy decreases from 40-50 in the north to 15-20 in the south, while the vegetation cover decreases from 40-60% to 10-20%, and even less. The vegetation growth in the study area is strongly dependent on precipitation dynamics. Grasses and shrubs in the vegetation cover grow during the whole vegetative period, but the vegetation growth is most rapid during May and early June (the period of greatest precipitation) in the southern part, and during June in the northern part of the study area. During droughty months in summer (July and early August) their rate of development is hindered. This period of semi-dormancy occurs throughout the study region (21,22). The area is used for agricultural production: about 10% of the total area is occupied by crops and fodder grass (*Agropyron cristatum* and *Convolvulvis arvensis*) fields.

## DATA

### Satellite data

Two types of satellite data were combined for the derivation of a medium resolution *LAI* map at the regional scale: Landsat ETM and SPOT-VEG data.

This study used a Landsat ETM+ image (path 154/row 26) acquired on June 19, 2004; it had level 1 G processing, a 30 m cell size, and was projected in UTM coordinates (WGS 84 datum). Visual analysis of the image revealed that there was an absence of haze or cloud in the image, and the region is known to have very little aerosol or atmospheric water vapour at this time of the year. This was also supported by analysing meteorological data recorded at climate stations located in the study region. The meteorological data indicated low values for air moisture and very low cloud cover of the sky during the whole of June 2004. The ETM+ row digital numbers were transformed to reflectance values using ENVI 4.3 preprocessing function.

A data set of Normalised Difference Vegetation Index with a spatial resolution of 1000 m was obtained from the VEGETATION instrument (VEG) onboard the Satellite Pour Observation d'Terra (SPOT). The multi-temporal dataset covers a time period of seven sequential years from April 2002 to October 2007 and was composed of maximum 10 day values. Although the use of maximum values significantly reduces the noise due to atmospheric effects, particularly, the amount of clouds in *NDVI* data (23), the used SPOT-VGT data over the study region comprised many areas whose surface reflectance values were contaminated by clouds. The remaining clouds from the 10 day data time series were removed by using a filtering algorithm based on a weighted least-squares regression approach described by (24). This algorithm calculates the 10 day value of each individual pixel as the mean of weighted values of the neighbouring 10 day periods. The closer a 10 day period is to the regression 10 day value, the more weight it receives. The number of the neighbours taken for the calculation can be different. In our calculation, we used a filter with five 10 day periods.

From the pre-processed data set we used three 10 day maximum *NDVI* images covering periods in June 2004 to calculate a monthly maximum *NDVI*. This monthly maximum *NDVI* image for June 2004 was then used in the subsequent analysis.

### *In situ* measurements of *LAI*

Two basic categories, direct and indirect methods, can be discerned for estimating *LAI* in field (4). As direct methods are known to be the most accurate but extremely time-consuming and impractical for a larger number of samples, they are employed relatively seldom (3). A comparison of direct *LAI* with indirect *LAI* is considered to be very useful, particularly when indirect *LAI* is measured in biomes or landscapes where it has not been estimated by this technique before. This was the rea-

son why, for the underlying study, both direct and indirect techniques were employed: we wanted to prove the consistency of indirect *LAI* estimations with the direct measurements.

The location of plots for both direct and indirect measurements was chosen to seize the spatial variability of vegetation types across the study area. The sampling design was based on a stratified sampling method along a 150 km transect through the study region. The stratification was based on the estimated vegetation type and estimated production level of the herbaceous cover. Both direct and indirect measurements were carried out during the second half of June at the peak of the growing season.

The direct method, which was applied, involved harvesting the biomass by destructive sampling of vegetation in 14 plots in the study area and converting it to the leaf area index through multiplication with the specific leaf area (*SLA*). For each of the 14 sites, the aboveground biomass (*AGB*) in (g C)/m<sup>2</sup> was estimated. The dry matter content was measured by drying samples in an oven. The proportion of carbon in biomass was assumed to be 0.47. The value of leaf area index for each of the 14 test sites was calculated as:

$$LAI = AGB \cdot SLA \quad (1)$$

with *AGB* in (g C)/m<sup>2</sup> and *SLA* in m<sup>2</sup>/(g C). The specific leaf area (*SLA*) is an ecophysiological biome dependent constant and has been widely used in models for net primary production. For the underlying study, values of *SLA* for grassland, shrubland and cropland were taken from (25).

Gap fraction analysis based on hemispherical photography was used in this study as indirect technique of ground-based *LAI* measurements. The vertical gap fraction corresponds directly to the complement to unity of the cover fraction and makes it possible to estimate the *LAI* of canopies with an assumption about the spatial organisation of the vegetation elements (26,27,28). We performed ground-based measurements of gap fraction using a WinScanopy Image Acquisition instrument developed by REGENT INSTRUMENTS (<http://www.regentinstruments.com>). Twenty test plots across the whole study region had been selected as sites of indirect *LAI* measurements. The plot size was chosen to correspond to an area observed by nine 30 m Landsat ETM+ pixels. Each plot had a size of 90·90 m<sup>2</sup> with multiple sub-plots over which measurements were averaged to produce a unique value per plot. The measurements were made in 30 m transect spacing within each plot. This spacing was selected on the basis of test trials conducted in reference plots and analysis of fine-scale vegetation patterns at Landsat ETM+ imagery. Plot locations and coordinates were determined using global positioning system (GPS), and hemispherical photos were taken at each of the sub-plots.

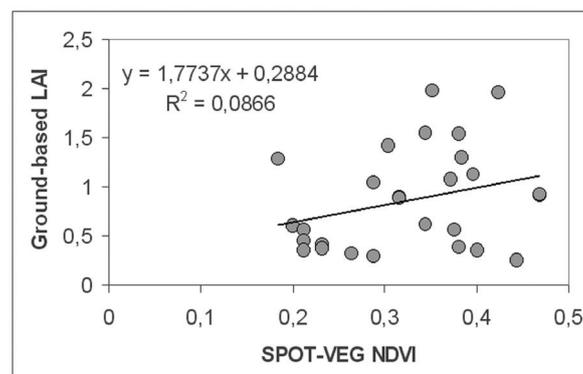


Figure 1: Scatterplot of ground-based *LAI* versus SPOT-VEG NDVI. By direct comparison, the relationship between these both variables is very weak and statistically insignificant. That supports high heterogeneity of the surface reflectance at the local scale (the level of plot size, 90·90 m<sup>2</sup>). This heterogeneity is not fully represented at the scale of 1000 m spatial resolution.

The *Can Eye* software (INRA, France, [http://www.avignon.inra.fr/can\\_eye/](http://www.avignon.inra.fr/can_eye/)) was used for the processing of hemispherical photographs taken at the test plots during the field campaign. *Can Eye* uses a look-up-table approach which is composed of gap fraction measurements in different view

zenith angles and the corresponding *LAI* and average leaf inclination parameters. The look-up-table is derived using the Poisson model and consists of 50 000 simulation cases. *LAI* and average leaf inclination are computed using a cost function between the measured gap fraction distribution and the simulated cases.

The study by (29) compared the direct and indirect *LAI* values in the data set used in the present study. They found a high consistency of the *LAI* values from different sources. With regard to the present study, it means that a combined use of direct and indirect *LAI* for remote sensing-based models will not contaminate the later modelling results.

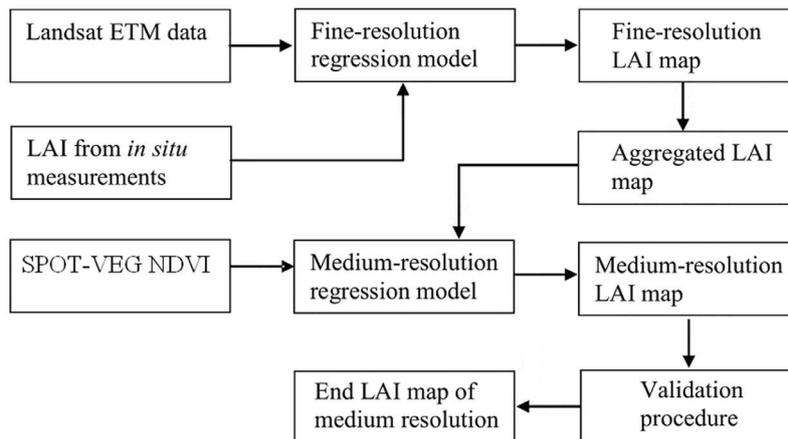


Figure 2: Overview of *LAI* estimation at the regional scale.

## METHODS

The whole *LAI* algorithm with all work sequences used in the present study is presented in Figure 2. These work steps were the following:

- Collection of field data on *LAI* at sample plots located across the study region
- Development of a fine-resolution regression model between the *in situ* measurements and Landsat ETM+ data
- Construction of the fine-resolution *LAI* map from the Landsat ETM+ data
- Aggregating the fine-resolution map to the 1 km spatial resolution of the SPOT-VEG data
- Development of the coarse-resolution regression model using SPOT-VGT data and the aggregated Landsat ETM+ *LAI* map
- Constructing the coarse-resolution *LAI* map
- Validating the coarse-resolution *LAI* map using the aggregated Landsat ETM+ *LAI* map as a reference data set
- Producing the final coarse-resolution *LAI* map over the study region and extrapolating the results to the neighbouring regions.

The individual partitions of the algorithm are briefly described in the following sections.

### Scaling up field-measured *LAI* to fine-resolution model

The algorithm used in the present study is based on empirical models between *LAI* and *NDVI* values in the SPOT-VGT data (Figure 1). To link ground-based *LAI* values with SPOT-VGT *NDVI* values a statistical regression was employed between these both variables. However, a direct comparison between ground measurements and medium- or coarse-resolution data is not recommended because of scale-mismatch and surface feature heterogeneity at the medium/coarse resolution (8,9). The probability of spatial variability in the surface feature is larger for coarse resolution

pixels compared to fine-resolution ground measurements. The simple correlation coefficient between the two variables is expected to be small where there is much variation within the larger scale. Thus, Figure 1 represents the inconsistency between the ground-based *LAI* and SPOT-VEG *NDVI* by a direct comparison. While the size of the ground plots for *in situ* measurements is too small to integrate it directly with medium or coarse resolution satellite data, it was necessary to introduce an intermediate scale between the ground plots and the scale of SPOT-VGT data. A commonly used approach includes two steps:

- 1) generation/calibration of a map/model of the variable of interest from ground measurements and satellite data with a spatial resolution close to that of the ground measurements, and
- 2) aggregation/application of this map/model to a coarser scale or to satellite data with a coarser spatial resolution.

Such an intermediate step has revealed its suitability in numerous recent studies on remote sensing-based estimation and validation of different biophysical variables including MODIS *LAI* (30,31,32). The derived fine-resolution *LAI* map served afterwards as “ground-truth information” for use with SPOT-VEG *NDVI* data. The regression model between this “ground-truth information” and the SPOT-VEG *NDVI* data was a basis for *LAI* estimation from the SPOT-VEG products. In these models the dependent variables were the ETM+ estimates of *LAI* computed for the area covered by the Landsat ETM+ image, and the explanatory variables were the *NDVI* values of SPOT-VEG pixels.

### **Selection of explanatory variables for remote sensing based *LAI* estimations**

A number of vegetation indices that are universally applicable to fine- and medium-resolution satellite data were tested in the analysis. Nonetheless, in this section we describe only two *VIs*, which were the basis for the production of final results.

The normalised difference vegetation index (*NDVI*) was calculated using the red (*R*) and the near infrared (*NIR*) reflectance. *NDVI* is defined by the following equation (33):

$$NDVI = \frac{NIR - R}{NIR + R} \quad (2)$$

Three bands of Landsat ETM+ (*R*, *NIR*, and *MIR*) were used to form the middle infrared corrected normalised difference vegetation index (*NDVIC*). The *NDVIC* is defined as follows (21,22):

$$NDVIC = NDVI \cdot \left( 1 - \frac{MIR - MIR_{min}}{MIR_{max} - MIR_{min}} \right) \quad (3)$$

where *NIR*, *R*, *MIR* are the reflectance in the near infrared, red, and middle infrared bands, respectively. *MIR<sub>min</sub>* and *MIR<sub>max</sub>* are the minimum and maximum *MIR* reflectances found in the associated image and defined as the 1% minimum and maximum cut of points in the histograms of *MIR* reflectance.

The *NDVIC* technique was developed to incorporate middle-infrared information while at the same time accounting for background effects in the observed reflectance (34,35). Another advantage of *NDVIC* against *NDVI* is that the difference between cover types is very much reduced so that the accuracy for *LAI* retrieval for areas with different cover types can be improved and a single *LAI* algorithm can be developed without the use of a co-registered land cover map.

### **Regression models**

Several regression types, i.e., the common least squares OLS regression and the non-linear regression reduced major axis (RMA) regression, were tested in order to find the best prediction model. Validation results were used to refine and improve the end model. Ordinary least square (OLS) regression methods are the most commonly used technique for prediction of different biophysical variables from remote sensing data. The OLS regression is a simple empirical approach designed to predict the dependent variable(s) using a linear relationship between them and the

dependent variable(s). However, previous studies have shown that the relationship between forest structure parameters and remotely sensed data can often be non-linear in nature. Therefore, the use of conventional least squares linear regression analysis is not always suitable. In these cases, the use of non-linear models can be advantageous. In the underlying study we used a logarithmic model with the following equation:

$$y = \alpha + b \cdot \ln x + \varepsilon \quad (4)$$

where  $\alpha$  and  $b$  are regression coefficients to be estimated, and  $\varepsilon$  is the error term.

Reduced major axis (RMA) is an orthogonal regression method that minimises the sum of the squared orthogonal distances from measurement points to the model function, accounting for the errors in both dependent and explanatory variables. The form of the regression model for RMA is identical to that of the OLS regression. However, the regression coefficients are estimated using a different approach than in the OLS regression. The slope  $b$  is defined as the ratio of sample standard deviation for  $y$  over the sample standard deviation for  $x$ . The intercept  $\alpha$  is defined as the sample mean of  $y$  minus the derived value of  $b$  multiplied with the mean of  $x$ . The mathematical equations for RMA parameter estimates are the following:

$$\alpha = \frac{\sigma_y}{\sigma_x} \quad (5)$$

$$b = \bar{y} - \frac{\sigma_y}{\sigma_x} \bar{x} \quad (6)$$

where  $\bar{y}$  and  $\bar{x}$  are the means of  $y$  and  $x$ , while  $\sigma_y$  and  $\sigma_x$  are their standard deviations, respectively.

The RMA technique exhibited promising results in studies on relationships between biophysical variables and remote sensing data reported in recent literature. Studies by (13) and (36) determined that the inflation and reduction of the variance in predictions, normally observed by OLS regression, is not observed in RMA regression. The common OLS regression technique is associated with an attenuation of the variance of its predictions. Values above the mean of the dependent variable tend to be underpredicted and values below the mean tend to be overpredicted (13). That occurs because the coefficients for the regression equation are calculated by minimising the sums of squares error in the dependent variable (37).

### Up-scaling the 30 m resolution model to 1 km pixel resolution

It is accepted that most of the environmental processes are scale-dependent (38,39). It means that a particular scale of measurements or modelling will influence the observed magnitude of spatial variations. The problem of scaling data and models is adequately treated in geostatistics. The geostatistical technique of block kriging has effectively been used to estimate over scales or domains larger than that of the original observations (39,40). In the present study, block kriging was used to re-scale data from the fine-resolution (30 m Landsat ETM+) to the coarser resolution (1 km SPOT-VGT). Since most readers will be familiar with the kriging and block kriging procedures described thoroughly in several books and papers (see, for instance, 40,41), we give here only equations for the block kriging-based up-scaling technique. According to (41), an estimate for the larger spatial domain is calculated from a weighted average of known values at the finer spatial domain:

$$\hat{Z}(B) = \sum_{i=1}^N \lambda_i z_a(x_i) \quad (7)$$

where  $\hat{Z}(B)$  is the value of the property of interest ( $LAI$  in the present study) at the spatial domain  $B$  (at 1 km resolution in the present study);  $N$  is the amount of fine-resolution pixels within a spatial unity of the coarser resolution domain (Landsat ETM+ pixels contained within a 1 km SPOT-VGT

pixel in the present study);  $\lambda_i$  are weights; and  $z_a$  is the value of the property of interest (*LAI*) over a pixel  $x_i$  at the spatial domain  $a$  (30 m resolution of Landsat ETM+ data in the present study). The weights  $\lambda_i$  are chosen to minimise the estimation variance given by:

$$\hat{\sigma}^2(B) = 2 \sum_{i=1}^N \lambda_i \bar{\gamma}(x_i, B) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(x_i, x_j) - \bar{\gamma}(B, B) \quad (8)$$

where  $\gamma(x_i, x_j)$  is the semivariance between points  $x_i$  and  $x_j$ ;  $\bar{\gamma}(x_i, B)$  is the average semivariance between the  $i^{\text{th}}$  point and the larger cell  $B$ ; and  $\bar{\gamma}(B, B)$  is the average semivariance within  $B$  (41). Solving the above equations to obtain the weights  $\lambda_i$  is the basis of kriging where  $\gamma(x_i, x_j)$ ,  $\bar{\gamma}(x_i, B)$  and  $\bar{\gamma}(B, B)$  can be obtained from the variogram model representing the sample data.

Block kriging has been recognised as one of the most optimal techniques for data re-scaling because it accounts for the scale of spatial dependence in the property of interest as represented with the variogram (39).

The geostatistics-based re-scaling was done prior to building the model between the SPOT-VGT *NDVI* and the re-scaled Landsat ETM+ *LAI*. Due to the adequate dealing with the scaling issue the SPOT-VGT *LAI* model was more accurate.

### Evaluation of models

For assessment of the fit and accuracy of the model, we used common accuracy and precision statistics such as the sum of squared residuals (*RSS*), the mean of squared residuals (*MSS*), and tests of statistical significance (*t*- and *p*-values). We also employed the *F*-test to prove the null hypothesis implying no relationship between the dependent and independent variables in a model. The null hypothesis was rejected if the calculated *F*-value was greater than the critical *F*-value. The value of  $R^2$  was also used as a guide for goodness-of-fit of the derived models.

The above statistics characterise the expected value of the regression estimate. However, in some cases, the precision of the regression estimate is just as important as the expected value of this estimate. The issue of bias-precision trade-offs - an important issue that occurs in many types of statistical modelling - was addressed in the present study using the approach based on an idea of (42) who proposed a statistic for selecting an unbiased model from a list of candidates. The Mallows' statistic is calculated as:

$$c_p = \frac{RSS}{\hat{\sigma}^2} + 2p - n \quad (9)$$

where *RSS* is the sum of squared residuals for a given model; *p* is the number of parameters in the model; *n* is the number of calibration points and  $\hat{\sigma}^2$  is the mean squared error estimate from the model. If the model is providing an unbiased estimate then the observed value of *RSS* will be  $c_p$ . On the other hand, if the estimation provided by the model is biased, then the observed *RSS* will be higher than  $c_p$ . In this way, the statistic can be used to make a diagnosis of the bias-precision relationship and determine the model that still provides an unbiased (or slightly-biased) model.

## RESULTS AND DISCUSSION

### Retrieval of the fine-resolution *LAI* map

Ordinary least squares regression models were developed using the in situ measurements of *LAI* from the calibration data set and Landsat ETM+ data. The model with the best accuracy was derived using *NDVI* as independent variable:  $R^2 = 0.68$ ,  $RMSE = 0.22$ . Table 1 shows the fit statistics of this model. These model coefficients were then implemented to the ETM+ scene to obtain a fine-resolution *LAI* map with the pixel size of 30 m · 30 m. The data from the 12 validation plots were used to validate the constructed model. The constructed model was applied to the validation

data set and statistics on the quality of the model fit and precision were calculated. The derived statistics are presented in Table 2. The table also includes similar statistics for the calibration data set. The results show that the model correlated strongly to the *LAI* values in the validation data set. The model explains about 55% of the variance in the ground-based *LAI* in the validation data set and yields the mean squared error (MS) of 0.1655.

The robustness of the model was also assessed with scatterplots of the observed versus predicted values and residuals versus predicted values to check for unbiasedness and homoscedasticity of the model (Figure 3). The values predicted by the model (Figure 3a) are shown together with the 95% prediction confidence interval. The model seems to explain reasonably well the variation of the dependent variable in the validation data set. The residuals of the model against the predicted values are shown in Figure 3a. Statistical tests of regressing the model squared residuals on the corresponding predictors revealed no significant relationships between these variables. The tests of unbiasedness and homoscedasticity had negative results. That means that the errors from the fine-resolution *LAI* map were random. Therefore, the total error in an aggregated 1000·1000 m<sup>2</sup> Landsat ETM+ pixel (1111.111 ETM+ pixels) would be 0.00021. In the further analysis, we consequently assumed that ETM+ *LAI* values - when degraded to the resolution of SPOT-VEG data - had negligible errors and can serve as reference information for the SPOT-VEG *LAI* model. Further in the study, these fine-resolution maps were used as “ground truth” in retrieval and validating *LAI* models at the medium-resolution scale.

Table 1: Fit statistics of the regression model between Landsat ETM+ NDVIc and ground-based *LAI* from the calibration data set.

Parameter	Coefficient	Stand. Error	t-value	p-value	Lower 95%	Upper 95%
A	0.4543	0.0651	6.9695	0.0000	0.3134	0.5952
B	3.1332	0.5942	5.2727	0.0001	1.8494	4.4169

Table 2: Goodness-of-fit and precision statistics derived from the *LAI* model in the calibration and validation data sets.

Data sets	Goodness-of-fit		Residuals accuracy		F-value
	R <sup>2</sup> adjusted	Stand. error	RSS	MS	
Calibration	0.6568	0.2188	0.6226	0.0478	27.8020
Validation	0.5463	0.3068	1.6552	0.1655	8.0614

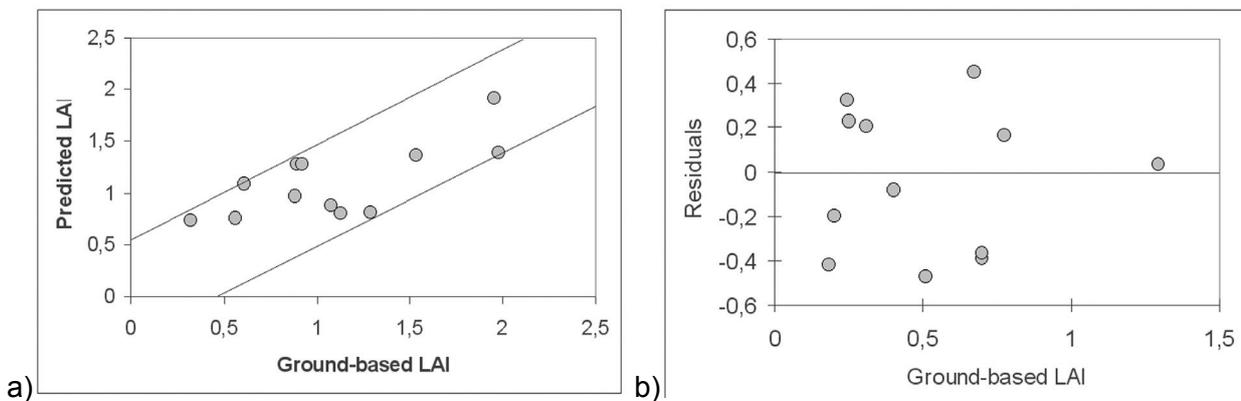


Figure 3: (a) Plot of predicted *LAI* versus ground-based *LAI* in the validation data set. The lines represent the 95% prediction confidence interval. (b) Plot of residuals versus ground-based *LAI* values in the validation data set.

### Statistical models between ETM+ LAI and SPOT-VEG NDVI

The fine-resolution LAI map derived from the Landsat ETM+ imagery was scaled up to the spatial resolution of SPOT-VGT data using the block kriging technique. The dependent variable was the ETM+ estimates of LAI degraded to the spatial resolution of 1000 m, and explanatory variables were the NDVI values of the corresponding SPOT-VEG pixels. Linear, non-linear and reduced major axis regression models were tested. 7800 pixels were selected randomly across the area covered by the Landsat ETM+ image to be used in parametrisations of the model. The parameterisations of the models were based on the overlay of the aggregated fine-resolution LAI map and SPOT-VEG data for these pixels.

The relationship between Landsat ETM+ LAI and SPOT-VEG NDVI in the calibration data set is strong and statistically significant ( $R^2 = 0.58$ ,  $p < 0.001$ ). The medium-resolution NDVI generally corresponds well to the Landsat ETM+ LAI (Figure 4). Most of the data points are located within the 95% prediction confidence interval. However, visual inspection of the scatter plot suggested that the relationship between these both variables seems to be slightly non-linear. The scatterplot reveals the decreasing sensitivity of NDVI to the variability in LAI with values above 0.8. It means that similar values of NDVI may represent a wide range of LAI values. The use of conventional least squares linear regression analysis to estimate LAI from the normalised difference vegetation index could not be fully suitable because of a lower explanatory power of this regression type in the case of its application to a non-linear relationship. As it has been reported by (36), the reduced major axis regression can be advantageous for applications to slightly non-linear relationships.

We employed three types of regression techniques to relate the ETM+ LAI and SPOT-VEG NDVI: OLS regression, RMA, and non-linear regression. We also employed RMA regression with  $\ln(NDVI)$  as independent variable. Table 3 shows the fit statistics of these regression models. All the examined models demonstrated a very high statistical significance of their coefficients ( $p$ -value  $< 0.0000$ ). Standard errors of the regression coefficients are characterised by very low values. The goodness-of-fit for slope coefficient  $b$  was similar for all the regression models, the  $t$ -value ranged only slightly from 49.54 to 50.71. On the contrary, the  $t$ -value of intercept varied very significantly between the regression models. Both RMA models are characterised by much more robust estimations of the intercept coefficient.  $T$ -values of the RMA intercept estimations were 11.23 and 12.68 versus 5.87 and 6.22 for their OLS counterparts. Values of the determination coefficient ( $R^2$ ) were practically similar for the four regression models, about 0.58. Note that the 95% confidence intervals of the individual parameter estimates are wider for the RMA models than for the corresponding OLS models. This indicates a better robustness of the RMA models in comparison to the OLS models.

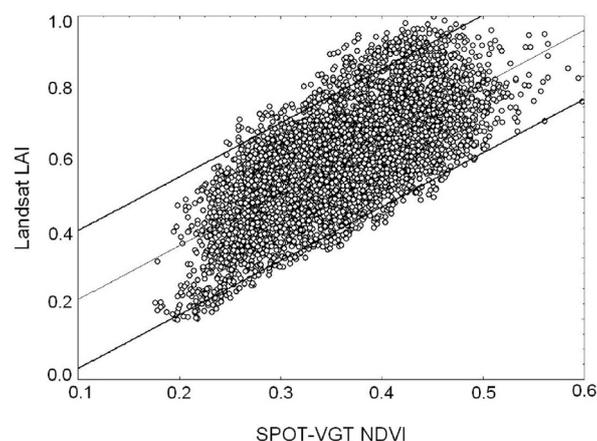


Figure 4: Relationship between SPOT-VEG NDVI and the LAI derived from Landsat ETM+. Thick lines represent 95% prediction confidence limits.

Table 3: Fit statistics of regression models between NDVI from SPOT-VEG and ETM+ LAI.

Regression models and equations	Coefficient	Stand. Error	t-value	p-value	Lower 95%	Upper 95%
<b>Non-linear</b> $LAI=a+b \cdot \ln NDVI$	0.8866 0.3115	0.0038 0.0035	-6.2219 50.5084	0.0000 0.0000	0.8791 0.3045	0.8941 0.3185
<b>OLS</b> $LAI=a+b \cdot NDVI$	0.2431 0.8886	0.0037 0.0104	-5.3212 49.7293	0.0000 0.0000	0.2359 0.8683	0.2504 0.9090
<b>RMA</b> $LAI=a+b \cdot NDVI$	0.1302 1.1254	0.0029 0.0107	12.2897 49.5433	0.0000 0.0000	0.1194 1.0146	0.1528 1.1571
<b>RMA (lnNDVI)</b> $LAI=a+b \cdot \ln NDVI$	1.0205 0.4377	0.0291 0.0073	11.6381 50.7156	0.0000 0.0000	0.8671 0.3964	1.2937 0.4761

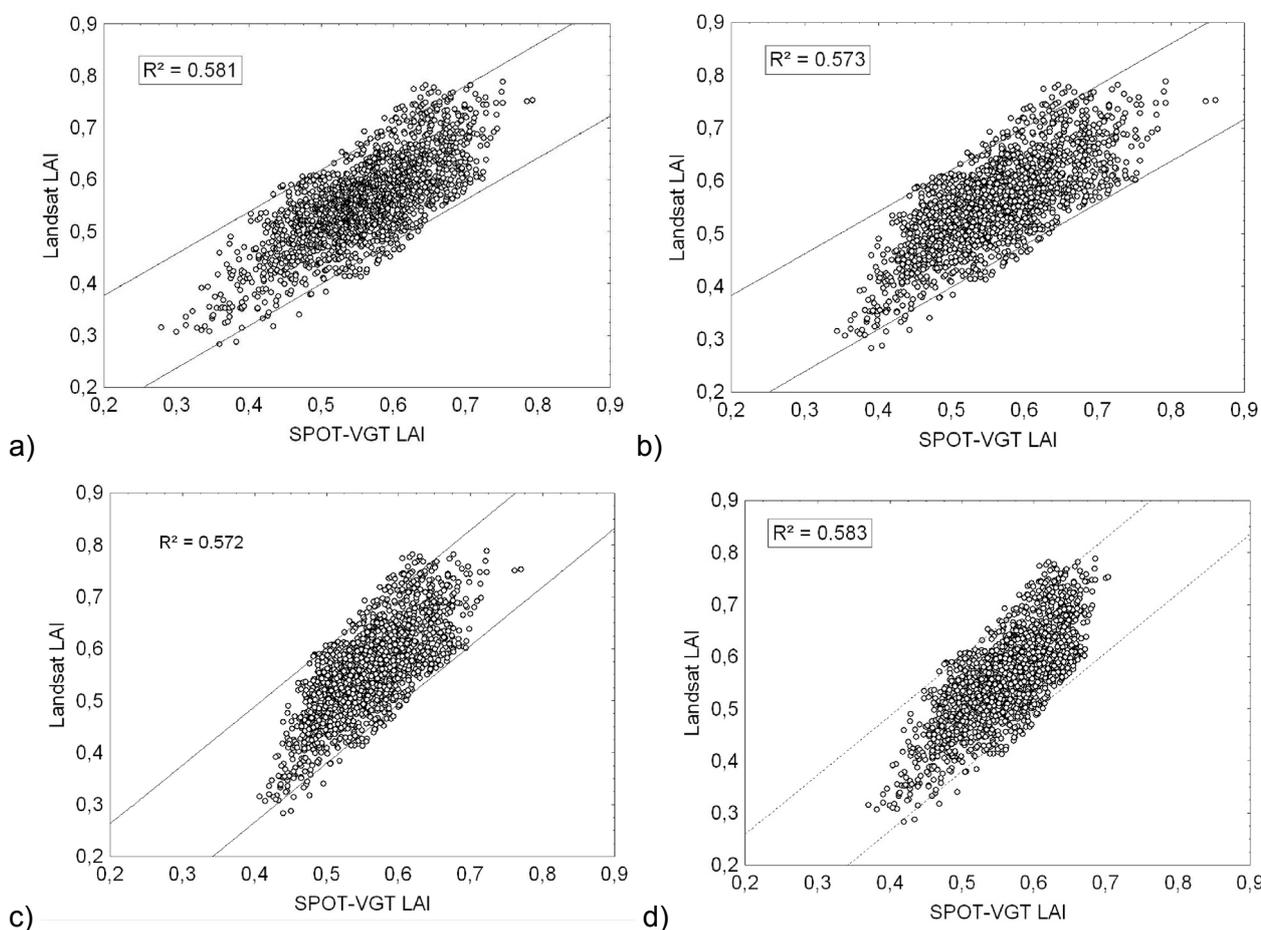


Figure 5: Scatter plots of SPOT-VEG LAI estimates against ETM+ LAI estimates together with 95% confidence intervals for (a) RMA (lnNDVI), (b) RMA, (c) linear OLS model, and (c) non-linear model.

**Comparison of SPOT-VEG estimates with the ETM+ estimates**

The LAI values in SPOT-VEG images were calculated at 1000 m resolution from  $NDVI_c$  values of individual pixels using the regression models described above. To investigate the accuracy of LAI values in individual pixels, a comparison was made between LAI values in medium-resolution images and those in the degraded ETM+ image. For this purpose, we used a validation data set. To build the validation data set, we randomly selected 1800 pixels throughout the area covered by the Landsat ETM+ image. The values predicted by the models are shown together with the 95% confidence intervals of the individual predictions in Figure 5. The models seem to explain reasonably

well the variation of the dependent variable. Even though the scatter of data points in the graphs is significant, most of the regression points are located within the 95% prediction confidence intervals. The models explain about 58% of the corresponding total variation of the Landsat ETM+ *LAI*. Scatter plots in Figure 5 demonstrate a decrease of the 95% prediction confidence intervals when changing from RMA to OLS approach. Another issue is the much smaller range of the *LAI* values predicted by the OLS models. The *LAI* from the Landsat ETM+ model ranges between 0.3 and > 0.8, whereas the *LAI* predicted by the linear and non-linear OLS models ranges between about 0.4 and about 0.7. To some extent, the *LAI* value scope predicted by the RMA approach is wider than in both OLS counterparts and it is closer to that of the Landsat ETM+ prediction.

Using equation (8), the  $c_p$  statistic was calculated for all the *LAI* models in the validation data set. Figure 6 demonstrates the calculated values of  $c_p$  plotted together with the observed values of *RSS*. The vertical lines in the plot show the upper and lower 95% confidence limits of the distribution of  $c_p$  with the black dot showing the expected value. The squares represent the observed value of *RSS*. The optimal combination of bias and precision is achieved when the square exactly overlays the black dot. This case would indicate no bias in the model calibration. When the dot is located within the 95% confidence interval, the model provides a *LAI* prediction with a negligible bias. When the square falls outside of the 95% confidence interval, then the model provides a highly biased estimation of *LAI*. On the basis of this plot, the optimal combination of bias and precision is provided by the RMA(lnNDVI) model. The linear OLS and RMA models provide *LAI* estimations with a negligible bias (within 95% confidence intervals). On the contrary, the non-linear OLS model is characterised by a highly-biased prediction of *LAI*: the observed *RSS* value is located outside of the 95% confidence limits of the  $c_p$  statistic.

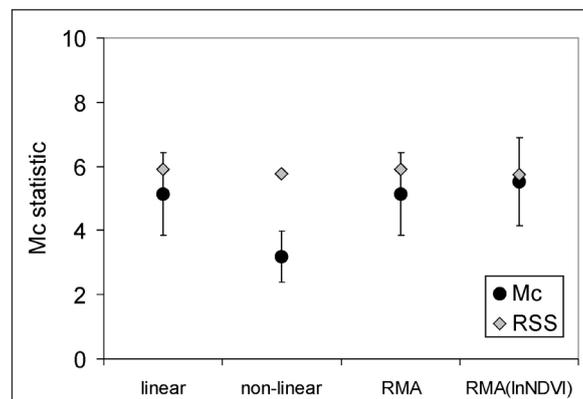


Figure 6: Bias-precision trade-offs of the SPOT-VGT *LAI* models.

Figure 7 shows the distribution of the *LAI* modelled by the used SPOT-VEG algorithms as well as that of the up-scaled ETM+ *LAI*. The comparison of the mean values given in the histograms suggests that both the linear and non-linear models underestimated the ETM+ *LAI* by 0.05 and 0.03, respectively, whereas the RMA model represents the mean value of ETM+ *LAI* precisely. The histograms of the linear and non-linear models are characterised by a much lower range of *LAI* in comparison to the ETM+ histogram. In both models the standard deviation shows a value of 0.15 versus 0.21 in the ETM+ model. The occurrence of pixels located in the right middle range (0.5 – 0.65) of the *LAI* span is too high. This happens at the expense of other *LAI* classes: the pixels located at the lowest (< 0.5) and highest end (> 0.7) of the *LAI* span are not predicted by either of these models.

The distribution of *LAI* from the RMA model shows the closest association with the ETM+ *LAI* distribution. Figure 7c demonstrates that the frequency of pixels in most of the classes is similar to that in the ETM+ *LAI* histogram. These results presume a better suitability of the RMA technique to deal with the inflation and reduction of the variance in *LAI* predictions. The attenuation of the variance of *LAI* predictions presented in the OLS and non-linear models is not observed in the RMA regression. The range of *LAI* values is entirely retained in the RMA prediction. These results sup-

port the conclusion by (36) who found that the RMA regression method is the most practical option for regional *LAI* modelling with coarse resolution inputs.

Generally, the results of the pixel-by-pixel comparison suggested a good agreement between the *LAI* predictions from the ETM+ and the SPOT-VEG data at the scale of 1000 m. Even though the value of  $R^2$  was not very high, the prediction accuracy of the SPOT-VEG models should be considered as sufficient. The values of *RMSE* and bias were relatively low. The distribution of the resulted *LAI*, particularly in the *RMA* model, is very close to the distribution of *LAI* derived from the ETM+ data.

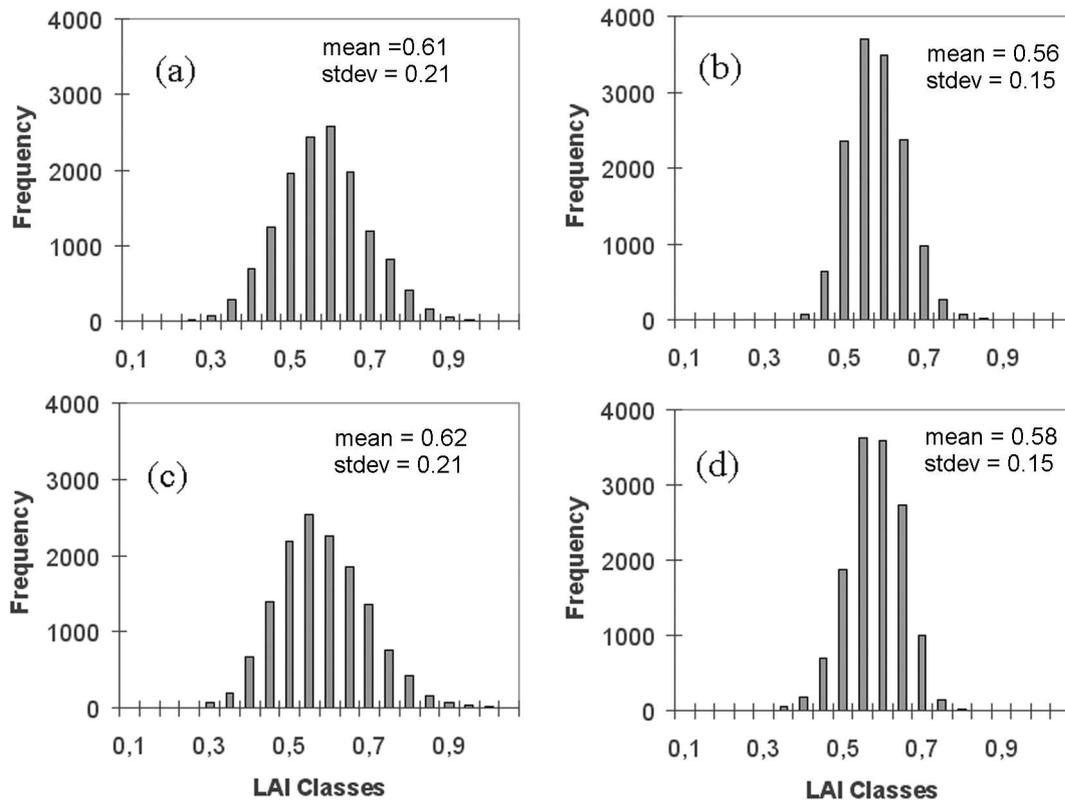


Figure 7: Histograms of the *LAI* derived by the following models: (a) the up-scaled ETM+, (b) SPOT-VEG using linear, (c) RMA, and (d) non-linear algorithms. The y-axis presents the frequency of *LAI* values in each value class.

#### Influence of scale on the modelling results

(37) gives the following reasons for the scattering of data points when plotting the *LAI* obtained from a coarse-resolution satellite image against the *LAI* from an up-scaled fine-resolution image:

- errors in co-registration of two images
- information distortion through aggregation of a fine-resolution ETM+ image to a coarse resolution
- the effect of surface heterogeneity arising at a coarse-resolution scale
- atmospheric noises in data sets.

With respect to our study region, the atmospheric effects in the SPOT-VEG dataset should be considered as insignificant because the region of the study is characterised by a low content of aerosol or atmospheric water vapour at this time of the year. (37) proposed a pixel degradation experiment (i.e. resampling the *LAI* images to be compared to much coarser resolutions) to separate the reasons relating to errors in co-registration of two images from other reasons. We thought that a pixel degradation experiment could be helpful for identification of the degree of factors *a*, *b*, and *c* on the modelling re-

sults. Another issue of a pixel degradation experiment could be an examination of the transferability of the MODIS LAI algorithm retrieved using the data with the spatial resolution of 1000 m to other spatial scales. Both issues were addressed in the study by performing the pixel degradation experiment. The SPOT-VEG NDVI image was up-scaled to spatial resolution of 2 km using a spatial average sampling method. The RMA algorithm was applied to the resampled image. The ETM+ LAI was also up-scaled to 1 km resolution using the same sampling method and compared with the corresponding SPOT-VEG LAI derivatives. Figure 8 shows the scatter plots of this comparison.

The results suggested a general applicability of the regression model retrieved at the scale of 1000 m to the scale of 2000 m. The values of the coarse resolution SPOT-VEG LAI correspond very well to the values of ETM+ LAI. This correspondence is better than that for the scale of 1000 m in Figure 4. Visual inspection of the scatter plots reveals a significant reduction of the scatter of data points. This reduction is also reflected in the value of  $R^2$ : 0.57 and 0.66 for the 1 km and 2 km models, respectively. The reduction of a scatter of data points after degradation to a coarser resolution suggests that many of the errors in the original scatter were caused by errors in co-registration of images (37). The improved comparison gives us assurance in the accuracy of LAI modelling at the origin scale of 1000 m. However, the results of the pixel degradation experiment suggest that greater attention should be given to image preprocessing. Errors introduced in the co-registration are a significant problem in LAI modelling and validation.

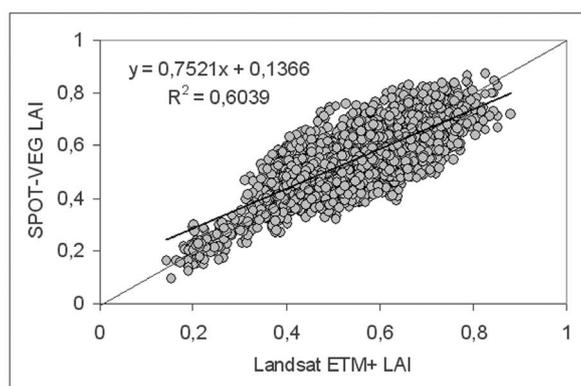


Figure 8: Comparison of LAI derived from the SPOT-VEG RMA algorithm with that from the ETM+ re-sampled to the resolution of 2000 m.

### Final medium-resolution LAI maps

The results of the validation indicate that it is feasible to retrieve medium-resolution LAI maps from SPOT-VEG data using an algorithm based on the overlay of the SPOT-VEG data and the up-scaled LAI map derived from Landsat ETM+ imagery. As first step towards mapping LAI from the medium-resolution SPOT-VEG data at a regional scale, a map of the peak-season LAI for the semi-desert and grassland zones of Kazakhshtan with a spatial resolution of 1000-m was produced from the maximum values of NDVI derived from 10-day SPOT-VEG composites for June 2004 (Figure 9b). The LAI map derived from SPOT-VEG data corresponds well to the up-scaled ETM+ LAI map (Figure 9a).

The spatial distribution of the peak-season LAI demonstrates a strong north-south gradient changing from more than 1.4 to near 0.0. The LAI spatial pattern clearly reflects the distribution of vegetation zones. Blue colour ( $LAI < 0.3$ ) is associated with desert zone. Green and yellow colours ( $0.3 < LAI < 0.6-0.7$ ) correspond to semi-desert zone. Light red colour ( $0.7 < LAI < 1.1-1.2$ ) represents short grassland which dominates dry steppe zone, whereas gloomy red ( $1.2 < LAI$ ) corresponds to long grassland that covers the zone of mixed steppe.

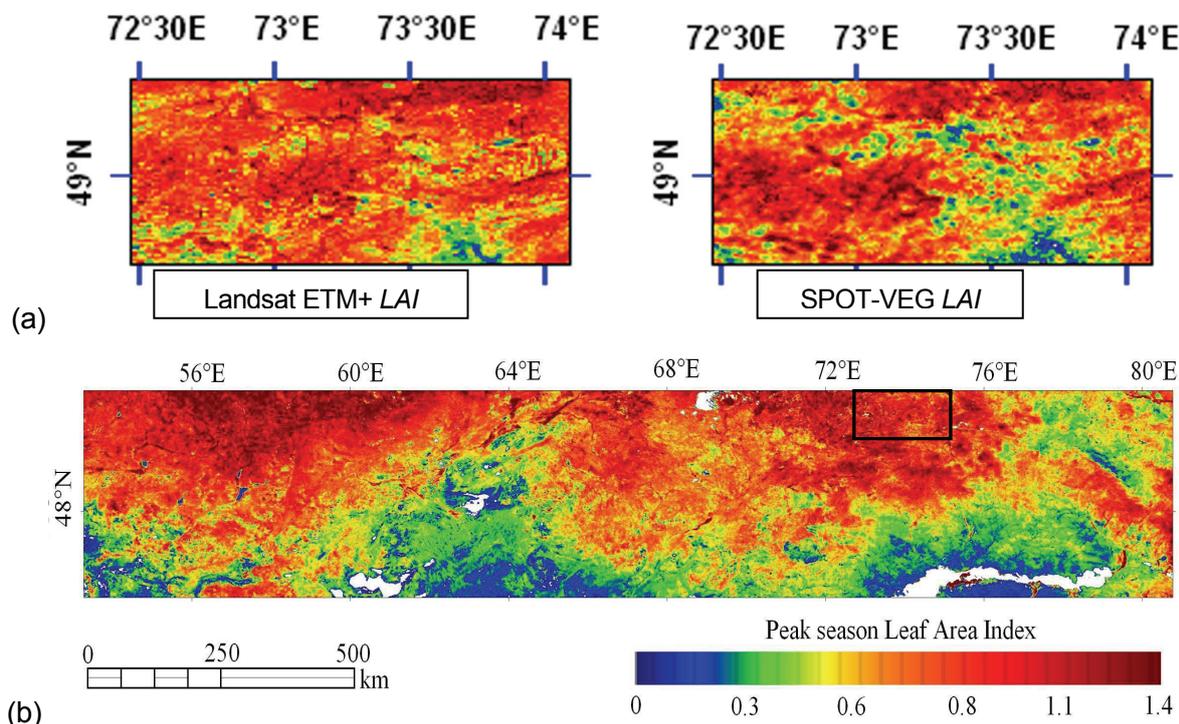


Figure 9: (a) Maps of the peak-season Leaf Area Index at a resolution of 1000-m over the study area retrieved from Landsat ETM+ and SPOT-VEG. (b) Map of the peak-season Leaf Area Index over the semi-desert and dry steppe zones of Kazakhstan. Black square represents the area covered by the maps in panel (a).

### CONCLUSIONS

This study presented an algorithm for remote estimation of LAI over a large semi-arid region in Kazakhstan using ground data collected from field sites, a Landsat ETM+ image, and SPOT-VEG data. The algorithm was based on the use of multi-spectral vegetation indices as variables explaining spatial variability of the ground based LAI. The modelling algorithm presented in this study differs from common empirical models in other related studies, in this study a fine-resolution LAI map retrieved from Landsat ETM+ imagery was used as an intermediate step to relate the ground-based LAI and 1000-m resolution data corresponding SPOT-VEG pixels. Attempting to predict ground-based LAI directly from SPOT-VGT imagery would result in a pure relationship. The Landsat ETM+ image having been aggregated to the SPOT-VEG resolution smoothes the spatial heterogeneity observed at the scale of 30-m pixel and makes it possible to establish statistically significant relationships between LAI and SPOT-VEG data. We conclude that the use of an up-scaled fine-resolution LAI map as ground-truth information is beneficial in the case of high inconsistency between the size of *in situ* plots and the pixel size of satellite data.

Several types of the function between the aggregated ETM+ LAI and SPOT-VGT NDVI were tested and analysed. All three methods, OLS, non-linear and reduced major axis regression, revealed good suitability and could establish a statistically significant relationship providing acceptable values of RMSE and  $R^2$ . However, both the OLS and non-linear models did not preserve the basic statistics (i.e., mean and standard deviation) of the dependent variable in their predictions, they considerably decreased the span of LAI as it was observed in histograms. RMA seems to retain the basic statistics as well as the range of LAI values in the SPOT-VGT prediction. As a result, RMA appears more promising for scaling up from Landsat ETM+ to SPOT-VGT spatial resolution.

The demonstrated approach can be used as an effective tool to produce *LAI* estimations at a region-wide scale. The estimates of *LAI* were realistic both when averaged for large areas and at the pixel scale. Furthermore, models calibrated using 1000 m resolution data may be used with data of coarser spatial resolution to extrapolate *LAI* predictions to larger areas. The results indicate that the models developed for estimating *LAI* over the area covered by the fine-resolution *LAI* map based on SPOT-VEG data with a resolution of 1000 m can also be applied to data with spatial resolution of 2000 m. Furthermore, up-scaling the model to a coarser resolution improves the model accuracy at the pixel scale: the pixel aggregating by a spatial average sampling method eliminates surface heterogeneity and significantly diminishes the errors induced by uncertainties in co-registration of the source images.

## ACKNOWLEDGEMENTS

We are grateful to all the colleagues from the Space Research Institute in Almaty, Kazakhstan, and personally Mrs. Dr. Nadya R. Muratova, for their help in organising and carrying out field work as well as for providing the data set of aboveground biomass used in the estimation of direct *LAI*.

## REFERENCES

- 1 Chen J M, J Liu, J Cihlar & M L Goulden, 1999. Daily canopy photosynthesis model through temporal and spatial scaling from remote sensing applications. *Ecological Modelling*, 124: 99-119
- 2 Monteith J L & M Unsworth, 1990. Principles of Environmental Physics (Arnold, London) 291 pp.
- 3 Gower S T, C J Kucharik & J M Norman, 1999. Direct and indirect estimation of leaf area index, FAPAR, and net primary production of terrestrial ecosystems. *Remote Sensing of Environment*, 70: 29-51
- 4 Jonckheere I, S Fleck, K Nackaerts, B Muysa, P Coppin, M Weiss & F Baret, 2004. Review of methods for in situ leaf area index determination. Part I. Theories, sensors and hemispherical photography. *Agricultural and Forest Meteorology*, 121: 19-35
- 5 Breda N J J, 2003. Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. *Journal of Experimental Botany*, 54: 2403-2417
- 6 Myneni R B, R R Nemani & S W Running, 1997. Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Trans. Geosci. Remote Sens.*, 35: 1380-1393
- 7 Muneni R B, S Hoffman, Y Knyazikhin, J Glassy, Y Tian, Y Wang, X Song, Y Zhang, G R Smith, A Lotsch, M Friedl, J T Morisette, P Votava, R R Nemani, & S W Running, 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, 83: 214-231
- 8 Weiss M, L Beaufort, F Baret, D Allard, N Bruquier & O Marloie, 2000. Mapping leaf area index measurements at different scales for the validation of large swath satellite sensors: first results of the VALERI project. In: 8<sup>th</sup> International Symposium in Physical Measurements and Remote Sensing, edited by J Leroy (Aussois, France) 125-130
- 9 Yang W, B Tan, D Huang, M Rautiainen, N V Shabanov, Y Wang, J L Privette, K F Huemmrich, R Fensholt, I Sandholt, M Weiss, D E Ahl, S T Gower, R R Nemani, Y Knyazikhin & R Muneni, 2006. MODIS leaf area index products: from validation to algorithm improvement. *IEEE Transactions on Geoscience and Remote Sensing*, 44: 1885-1898
- 10 Fensholt R, I Sandholt & M S Rasmussen, 2004. Validation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment*, 91: 490-507

- 11 Fernandes R & S G Leblanc, 2005. Parametric (modified least squares) and non-parametric (Theil-Sen) linear regressions for predicting biophysical parameters in the presence of measurements errors. *Remote Sensing of Environment*, 95: 303-316
- 12 McAllister D M & C Valeo, 2007. A robust method for the remote estimation of LAI in montane and boreal forests. *International Journal of Remote Sensing*, 28: 1891-1905
- 13 Cohen W B, T K Maersperger, S T Gower & D P Turner, 2003. An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment*, 84: 561-571
- 14 Robinson S, E L Milner-Gulland & I Alimaev, 2002. Rangeland degradation in Kazakhstan during the Soviet-era: re-examining the evidence. *Journal of Arid Environments*, 53: 419-439
- 15 Henebry G M, 2009. Carbon in idle croplands. *Nature*, 457: 1089-1090
- 16 Lioubimtseva E, R Cole, J M Adams & G Kapustin, 2005. Impacts of climate and land cover changes in arid lands of Central Asia. *Journal of Arid Environments*, 62: 285-308
- 17 Lal R, 2004. Carbon sequestration in soils of Central Asia. *Land Degradation & Development*, 15: 563-572
- 18 DeBeurs K M & G M Henebry, 2004. Land surface phenology, climatic variation, and institutional change: analysing agricultural land cover change in Kazakhstan. *Remote Sensing of Environment*, 89: 497-509
- 19 Propastin P A, M Kappas & M Muratova, 2008. Inter-annual changes in vegetation activities and their relationship to temperature and precipitation in Central Asia from 1982 to 2003. *Journal of Environmental Informatics*, 12(2): 75-87
- 20 Lavrenko E M & Z V Karamysheva, 1993. Steppes of the former Soviet Union and Mongolia. In: Coupland, R. T. (Editor) *Natural Grasslands* (Elsevier, London) pp. 3-55
- 21 Chupachin B M, 1968. Physical geography of Kazakhstan (Mektep, Alma-Ata) pp. 80-87 (in Russian)
- 22 Propastin P, M Kappas, S Erasmi & N R Muratova, 2007. Remote sensing based study on intra-annual dynamics of vegetation and climate in drylands of Kazakhstan. *Basic and Applied Drylands Research*, 2: 138-154
- 23 Holben B N, 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal Remote Sensing*, 7:1417-1434
- 24 Chen J, P Jönsson, M Tamura, Z Gu, B Matsushita & Eklundh, 2004. A simple method for reconstructing a high quality NDVI time series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment*, 91: 332-344
- 25 White M A, P E Thornton, S W Running R R & Nemani, 2000. Parametrization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: net primary production controls. *Earth Interactions*, 4: 1-85
- 26 Chen J M & J Cihlar, 1995. Plant canopy gap-size analysis theory for improving optical measurements of leaf area index. *Applied Optics*, 34: 6211-6222
- 27 Norman J M & G S Campbell, 1989. Canopy structure. In: R W Pearcy, J Ehleringer, H A Mooney & P W Rundel (Editors) *Plant Physiological Ecology: Field Methods and Instrumentation* (Chapman and Hall, New York) pp. 301-325
- 28 Weiss M, F Baret, G J Smith, I Jonckheere & P Coppin. 2004. Review of methods for in situ leaf area index determination. Part II. Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology*, 121: 37-53

- 29 Propastin P & M Kappas, 2009. Mapping Leaf Area Index in a semi-arid environment of Kazakhstan using fine-resolution satellite data and in situ measurements. GIScience & Remote Sensing, 46(2): 212-231
- 30 Tomppo E, M Nilsson, M Rosengren, P Aalto & P Kennedy, 2002. Simultaneous use of Landsat-TM and IRS- 1C WiFS data in estimating large areas tree stem volume and aboveground biomass. *Remote Sensing of Environment*, 82: 156-171
- 31 Muukkonen P & J Heiskanen, 2007. Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: a possibility to verify carbon inventories. *Remote Sensing of Environment*, 107: 617-624
- 32 Tomppo E, M Nilsson, M Rosengren, P Aalto & PKennedy, 2002. Simultaneous use of Landsat-TM and IRS- 1C WiFS data in estimating large areas tree stem volume and aboveground biomass. *Remote Sensing of Environment*, 82: 156-171
- 33 Tucker C J, C L Vanpra, M J Sharman & G Van Ittersum, 1985. Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980–1984. *Remote Sensing of Environment*, 17: 233-249
- 34 Eklundh L, L Harrie & A Kuusk, 2001. Investigating relationships between Landsat ETM+ sensor data and LAI in a boreal conifer forest. *Remote Sensing of Environment*, 78; 239-251
- 35 Brown L J, J M Chen, S G Leblanc & J Cihlar, 2000. Shortwave infrared correction to the simple ratio: an image and model analysis. *Remote Sensing of Environment*, 77: 16-25
- 36 Barterretche M, A Hudak, W B Cohen, T K Maier-spenger, S T Gower & J Dungan, 2005. Comparison of regression and geostatistical methods for mapping Leaf Area Index with Landsat ETM+ data over a boreal forest. *Remote Sensing of Environment*, 96: 49-61
- 37 Chen M, G Pavlic, L Brown, J Cuhlar, S G Leblanc, H P White, R J Hall, D R Peddle, D J King, J A Trofimow, E Swift, J Van der Sanden & P K E Pellikka, 2002. Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sensing of Environment*, 80: 165-184
- 38 Dark S & D Bram, 2007. The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography*, 31(5): 471-479
- 39 Atkinson P & N Tate, 2000. Spatial scale problems and geostatistical solutions: a review. *Professional Geographer*, 52(4): 607-623
- 40 Chiles J P & P Delfiner, 1999. *Geostatistics: Modeling spatial uncertainty* (John Wiley & Sons, New York)
- 41 Oliver M A, 1996. Geostatistics, rare disease and the environment. In: M Fischer, H J Scholten & D Unwin (Editors). *GISDATA 4: Spatial Analytical Perspectives on GIS* (Taylor & Francis, London) pp. 67-85
- 42 Mallows C L, 1973. Some comments on Cp. *Technometrics*, 15: 661-667