Inter-Comparison and Evaluation of the Global LAI Product (LAI3g) and the Regional LAI Product (GGRS-LAI) over the Area of Kazakhstan

Martin Kappas 1,*, Pavel Propastin 1, Jan Degener 1 and Tsolmon Renchin 2

1 GIS & Remote Sensing Section, Institute of Geography, Cartography, Georg-August University Goettingen, Goldschmidt Street 5, D-37077 Goettingen, Germany; E-Mails: ppropas@gwdg.de (P.P.); jdegene@uni-goettingen.de (J.D.)
2 NUM-ITC-UNESCO Space Science and Remote Sensing Laboratory, National University of Mongolia, 17026 Ulaanbaatar, Mongolia; E-Mail: tsolmonren@gmail.com

* Author to whom correspondence should be addressed; E-Mail: mkappas@gwdg.de; Tel.: +49-551-398-071; Fax: +49-551-398-020.

Academic Editors: Clement Atzberger and Prasad S. Thenkabail

Received: 23 September 2014 / Accepted: 10 March 2015 / Published: 27 March 2015

Abstract: Long-term global datasets of the Leaf Area Index (LAI) are important for monitoring global vegetation dynamics and are an important input for Earth system models (ESM). The comparison of long-term datasets is based on two recently available datasets both derived from AVHRR (Advanced Very High Resolution Radiometer) time series. The LAI3g dataset is developed from the new improved third generation Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) from AVHRR sensors and best-quality MODIS LAI data. The second long-term LAI dataset is based on the 8-km spatial resolution GIMMS-AVHRR data (Goettingen GIS & Remote Sensing, GGRS dataset). The GGRS-LAI product uses a satellite-based LAI. This algorithm uses a three-dimensional physical radiative transfer model, which establishes the relationship between LAI, vegetation fractional cover and given patterns of surface reflectance, view-illumination conditions and optical properties of vegetation. The model incorporates a number of site-/region-specific parameters, including the vegetation architecture variables, such as leaf angle distribution, clumping index and light extinction coefficient. For the application of the model to Kazakhstan, the vegetation architecture variables were computed at the local (pixel) level based on extensive field surveys of the biophysical properties of vegetation in representative grassland areas of Kazakhstan. As a main result of our study,
we could summarize that the differences between both products are most pronounced at the start and the end of the growing season. During the spring and autumn months, the LAI difference maps showed a considerable difference of LAI GGRS and LAI3g. LAI3g is characterized by a considerably earlier start and a later finish to the growing season than LAI GGRS. Moreover, LAI3g showed LAI > 0 during the winter months when any green vegetation is absent in all land covers of Kazakhstan. A direct cause for this could be a too high base level of the LAI3g during the leafless phase.

**Keywords:** long-term LAI products; LAI3g; Kazakhstan-wide LAI product (GGRS-LAI data); remote sensing of vegetation

1. Introduction

A major need stated by the NASA Earth science research strategy and European Space Agency (ESA) in environmental research is to develop long-term, consistent and calibrated data and products that are valid across multiple missions and satellite sensors. These datasets are called ESDR (Earth System Data Records) and are defined as unified and coherent sets of observations of a given parameter of the Earth system (e.g., LAI, FAPAR), which are optimized to meet specific requirements in addressing science questions (e.g., ecosystem modeling). Estimating, validating and conveying systematic errors in long-term Earth system science data records are a major effort to reduce data record uncertainty. ESDRs, like LAI or the fraction of photosynthetically-active radiation (FPAR), data records have been generated globally from different sensors (AVHRR, MODIS, MISR, Polarization and directionality of the Earth’s Reflectance (POLDER), SPOT-VGT, etc.) at different spatial resolutions (250 m to one degree) and temporal frequencies (four-day, eight-day and monthly; for more detailed information, see the Making Earth System Data Records for Use in Research Environments (MEaSUREs) 2006–2010 Program website: http://measures-projects.gsfc.nasa.gov/).

The Eurasian temperate grassland, representing the world’s largest grassland biome, with its east-west extension of more than 6000 km and a north-south extension of about 1490 km, has been beyond the scope of extensive validation efforts. The existing global LAI products have not been checked in any region or biome of Central Asia yet. In our current knowledge, there are no published studies on the validation of any global LAI products in Central Asia. Within this scope, an ongoing research work on validation of the existing global LAI products in a grassland region in Central Kazakhstan is of great importance. A short summary of the available global LAI products before the delivery of the new LAI3g data record [1] and the suitability of the available Global LAI products for applications in Central Asia was presented by [2]. Towards the successful implementation of a global LAI product, various consistent time series of LAI exist with 1 km × 1 km spatial resolution, but independent assessment of product quality is a critical step towards the successful usage of these global LAI products. The satellite-modelled LAI products evidently simplify reality and require extensive validation from the greatest possible number of sites to represent all biomes included in the model.

LAI is defined as the one-sided green leaf area per unit ground area in broadleaf canopies and as half the total needle surface area in coniferous canopies [3,4]. FPAR is the fraction of photosynthetically-active
radiation (400–700 nm) that is absorbed by vegetation. Both are key biophysical variables and have an
effect on energy exchange, mass exchange (e.g., water and CO₂) and momentum between the Earth
surface and atmosphere [5].

LAI is a major variable for modelling land surface processes, as it presents the phenological
development of vegetation. Furthermore, ecosystem modeling is widely needed to ascertain terrestrial
responses to changes in climate. All land surface models, such as the simple biosphere (SiB, [6]) model,
community land model (CLM, [7]), biome-biogeochemical cycles model (Biome-BGC, [8]), terrestrial
ecosystem model (TEM, [5]), and models in hydrology, crop production, such as
Goettingen-BioSTAR [9,10] and others, require a set of land-surface variables that include land cover,
LAI, FPAR and albedo. Land surface parameters, like LAI and FPAR, can only currently be derived by
remote sensing at a regional or global scale.

Tian et al. [11] worked on an inter-comparison of seasonal and spatial variations of global LAI
estimations from MODIS and the common land model (CLM). This inter-comparison showed mostly
consistency between the two datasets during snow-free periods. LAI and FPAR datasets are key input
data for carbon cycle models, such as Biome-BGC and CASA (Carnegie Ames Stanford Approach), for
which long-term changes in LAI indicate significant shifts in vegetation biomass and carbon sequestration.
On the other hand, climate controls the variability in net primary production and heterotrophic
respiration year in and year out. Many other studies refer to the sensitive role of LAI and FPAR in climate
simulations and carbon cycle modeling. Therefore, the demand for good quality time series of global
LAI and FPAR datasets is still high.

This paper is organized as follows. Section 2 describes the production details of the LAI3g and
Kazakhstan-wide (GGRS) datasets. Section 3 is focused on methods of the inter-comparison process for
LAI3g and GGRS-LAI datasets. Section 4 summarizes the results of the comparison between LAI3g
and LAI GGRS and gives some concluding remarks.

2. Evolutionary History of Global LAI3g and Kazakhstan-Wide Datasets (GGRS-LAI)

2.1. LAI3g Global Product

For the generation of the LAI3g dataset, a combination of the Collection 5 Terra MODIS LAI product,
Collection 5 MODIS land cover product MCD12C1, with international Geosphere Biosphere Programme
(IGBP) classes, and the GIMMS NDVI3g data was used. The MODIS-derived datasets were resampled
using the nearest neighbor method to match the spatial resolution of the NDVI3g data (0.0833°), and
after that, the three datasets were composited into a uniform spatial grid and temporal frequency. The
LAI retrieving algorithm employed a feed-forward neural network (FFNN) consisting of four neurons
(input parameters) in the input layer, 11 neurons in the hidden layer and one neuron in the output layer
(LAI3g) to simulate LAI values from GIMMS NDVI3g using MODIS LAI as the base. The four input
parameters of the LAI algorithm were land cover class, pixel-center latitude, pixel-center longitude and
NDVI3g. FFNN models were generated for each month and were trained through the back-propagation
method, whereas ten networks with different initial values were trained for each of the models. The best
fitting network was selected as the final FFNN model for generating the LAI3g dataset. The
 corresponding LAI3g datasets have the following attributes: 15-day temporal frequency, 1/12 degree
spatial resolution and a temporal span of July, 1981, to December, 2011. The entire dataset is described by [1].

2.2. LAI GGRS–Kazakhstan-Wide Product

Based on a radiative transfer model using an inverse method for solving the Beer–Lambert law and LAI algorithm previously derived for applications in Indonesia [12], the study by [13] produced an enhanced algorithm for the retrieval of the Kazakhstan-wide LAI dataset. The algorithm makes use of the GIMMS NDVI dataset obtained from the NOAA AVHRR sensor. This enhanced model was based on a relationship between LAI, the fraction of vegetation cover and given patterns of surface reflectance, view-illumination conditions and optical properties of vegetation. The model used a number of input parameters related to optical and structural characteristics of canopy, such as the extinction coefficient (k), the foliage projection coefficient (G), leaf inclination angle (Θ) and the clumping index (Ω), which were derived from hemispherical photographs in situ and direct measurements of vegetation architecture obtained during field surveys. These parameters were obtained separately for each of the main vegetation types of Kazakhstan and incorporated into the algorithm as biome-specific constants. View-illumination conditions were simulated by a view angle geometry model incorporating the solar zenith angle and the sensor viewing angle. For a detailed description of the models, see the related publications by the authors [12–14]. The LAI dataset was generated from GIMMS NDVI data and ground data on canopy structure measured at 66 sampling sites across the study area. The entire dataset is described by [13] and can be downloaded under Department of Cartography, GIS and Remote Sensing. Available online at: http://www.uni-goettingen.de/de/download-avhrr-lai-data/366739.html.

<table>
<thead>
<tr>
<th>LAI Product/Producer</th>
<th>Spatial Resolution</th>
<th>Spatial Coverage</th>
<th>Temporal Coverage</th>
<th>Input Data</th>
<th>LAI Algorithm</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAI3g (Boston University, USA)</td>
<td>0.0833°</td>
<td>Global</td>
<td>1981–2011</td>
<td>NOAA AVHRR GIMMS3g NDVI MODIS BNU LAI MODIS Land Cover Collection 5 product (2007)</td>
<td>Neural network</td>
<td>Zhu et al. (2013)</td>
</tr>
<tr>
<td>LAI GGRS (Göttingen University, Germany)</td>
<td>0.0833°</td>
<td>Kazakhstan</td>
<td>1982–2010</td>
<td>NOAA AVHRR GIMMS NDVI NOAA AVHRR Land Cover Product (1981–1994) GTOPO30 digital terrain model Solar geometry Vegetation architecture parameters (clumping, leaf inclination, etc.)</td>
<td>Radiative transfer model</td>
<td>Propastin and Kappas (2012)</td>
</tr>
</tbody>
</table>

Field Data

In situ estimations of LAI and vegetation structure parameters were carried out at 66 sampling plots established within the main vegetation types, which cover the majority (99.2%) of Kazakhstan’s land
area: grassland, cropland, shrubland, woodland, mixed forest, deciduous broadleaf forest and deciduous needle leaf forest. Field campaigns were undertaken in June 2008, August 2009, and June 2010. In all cases, the plot size for the in situ measurements was 90 × 90 m, corresponding to an area observed by 3 × 3 Landsat ETM+ pixels [15]. Within each plot, measurements were made along transects with 20-m spacing between data points. In total, 14 measurements were completed within each of the sampling plots, which were then averaged to mean values over corresponding plots.

The in situ estimations were performed using indirect non-contact (hemispherical photography) and direct contact (harvesting) methods. For hemispherical photography, a WinScanopy Image Acquisition instrument developed by Regent Instruments (http://www.regentinstruments.com) was used. For processing of hemispherical photographs and retrieval of vegetation structure variables, we used the Can Eye software (INRA, France, http://www.avignon.inra.fr/can_eye/). The gap fraction was estimated at 5° zenith angle intervals and used for further calculations. LAI and other canopy structure variables (leaf inclination angle, foliage projection function) were calculated using routine procedures included in the Can Eye software following the methods described by [3,4]. The clumping index was calculated using the logarithmic gap averaging technique given by [16].

During the field campaign of 2010, the estimations by hemispherical photography were supported by Li-Cor LAI 2000 plant canopy analyzer (Lincoln, NE, Li-Cor 1992) measurements of LAI. The Li-Cor LAI 2000 measurements were performed in the same plots as the hemispherical photography. LAI was then automatically calculated from transmittance for directional radiation corresponding to Poisson’s distribution, as described by [17]. The Li-Cor-derived LAI values were compared with corresponding LAI values retrieved from hemispherical photography. Both non-contact optical methods provided almost equivalent results (discrepancies were <5%).

The LAI obtained by the indirect non-contact methods were validated versus LAI from the direct contact destructive technique. For this, in 14 sampling plots, LAI was obtained from harvested vegetation and compared with the LAI values computed from the hemispherical photography (Table 2). The results of this comparison revealed good correspondence between the indirect optical and direct contact measured LAI [18].

**Table 2.** In situ measurements of Leaf Area Index (LAI) in Kazakhstan used in this study. For the location of the test sites, see Figure 1.

<table>
<thead>
<tr>
<th>Test Site and Number of Sampling Plots</th>
<th>Date</th>
<th>Vegetation Type</th>
<th>Measurement Method</th>
<th>LAI Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shetsky 1, 14 plots</td>
<td>June 2004</td>
<td>Grassland, Cropland</td>
<td>Direct contact destructive</td>
<td>0.25–1.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.31–0.85</td>
</tr>
<tr>
<td>Shetsky 2, 20 plots</td>
<td>June 2008</td>
<td>Grassland, Cropland, Shrubland</td>
<td>Indirect non-contact optical</td>
<td>0.30–1.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.21–0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.84–2.06</td>
</tr>
<tr>
<td>Almaty, 16 plots</td>
<td>June 2008</td>
<td>Mixed forest, Deciduous broadleaf</td>
<td>Indirect non-contact optical</td>
<td>3.30–5.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>forest, Deciduous needle leaf</td>
<td></td>
<td>3.10–5.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>forest, Woodland</td>
<td></td>
<td>4.35–7.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.42–4.10</td>
</tr>
</tbody>
</table>
Figure 1. Classification of current vegetation for the Republic of Kazakhstan based on the 8-km Advanced Very High Resolution Radiometer (AVHRR) Landcover 1981–1994 global maps. Test sites for in situ measurements are presented by squares and names.

3. Methods of the Inter-Comparison Process for LAI3g and GGRS-LAI

The methods of the analyses of the spatial and temporal differences between satellite-derived datasets are well described by [19]. With respect to the recommendations of [19], in this paper, discrepancies between LAI estimates are evaluated and quantified by the following statistics:

- coefficient of determination ($R^2$) of scatter plots of monthly LAI estimates for 2008 were analyzed over the whole area of Kazakhstan;
- slope and offset of the regression line;
- the value of root mean square error (RMSE);
- histograms of the frequency distribution;
- empirical cumulative distribution functions (ECDF); differences between two ECDFs $F(x)$ and $G(x)$ are quantified by the Kolmogorov–Smirnov statistics $D_{KS} = \max(\abs{F(x) - G(x)})$;
- Agreement coefficient $AC = 1 - SSD/SPOD$, where SSD is the sum of squared differences and SPOD is the sum of potential differences:

$$SSD = \sum_{i=1}^{n} (X_i - Y_i)^2,$$

$$SPOD = \sum_{i=1}^{n} (|\bar{X} - \bar{Y}| + |X_i - \bar{X}|)(|\bar{X} - \bar{Y}| + |Y_i - \bar{Y}|)$$

Validation of LAI products requires a comparison of the LAI dataset with ground measurements of LAI. Since a direct comparison of small-scale ground measurements with coarse-scale LAI products cannot provide any suitable results, the ground truth data have to be scaled up to the coarse resolution of LAI products. The state-of-the-art scaling up approach uses a transfer function and ancillary information, such as fine-resolution satellite data, to scale the ground measurements up to the coarse resolution of the products [20,21]. The database for the validation of the LAI GGRS product over Kazakhstan consists of the ground data of LAI from the main land cover types (grassland, cropland, shrubland, woodland, mixed forest, deciduous broadleaf forest and deciduous needle leaf forest) obtained in three regions of Kazakhstan. These data were collected by an indirect method (hemispherical photography) during a
number of field campaigns in 2008–2010 at 66 sampling plots of a size 90 × 90 m. A detailed description of the validation dataset is given by [13]. The suitability of hemispherical photography for the estimation of LAI in the study area was assessed by comparing the obtained LAI values with the LAI values from the direct destructive method [18].

For validation of the LAI3g versus the LAI GGRS products in this study, we used high-resolution maps of LAI using 30-m resolution Landsat imagery over the validation sites in Kazakhstan (see Table 3), which were produced in our previous studies [18, 22]. The maps were produced using empirical models based on the relationships between the ground measurements of LAI and spectral reflectance in the Landsat imagery. These high resolution LAI maps were degraded to the spatial resolution of LAI3g and LAI GGRS products (0.0833°). The validation was achieved by the pixel-by-pixel comparisons of the LAI products and the spatially-degraded Landsat-based LAI maps.

Table 3. Basic characteristics of the test sites that are used for the validation of the LAI3g versus the LAI GGRS product.

<table>
<thead>
<tr>
<th>Validation Site</th>
<th>Location</th>
<th>Number of Sampling Plots</th>
<th>Vegetation Type</th>
<th>Measurement Method</th>
<th>LAI Range</th>
<th>Date of Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shetsky</td>
<td>Central Kazakhstan</td>
<td>34</td>
<td>Grassland, cropland, shrubland</td>
<td>Direct contact destructive, indirect non-contact optical (hemispherical photography)</td>
<td>0.25–1.55</td>
<td>June 2008</td>
</tr>
<tr>
<td>Almaty</td>
<td>South Kazakhstan</td>
<td>16</td>
<td>Mixed, broadleaf and deciduous forests</td>
<td>Indirect non-contact optical (hemispherical photography)</td>
<td>2.40–7.20</td>
<td>August 2010</td>
</tr>
</tbody>
</table>

4. Results

4.1. Direct Validation of LAI3g as Compared to LAI GGRS

The validation of LAI3g and LAI GGRS with ground-based LAI datasets was carried out by performing regression analysis between LAI from both products (dependent variables) and LAI derived from Landsat (independent variables). The detailed results obtained from this validation are summarized in Table 4 and Figures 2 and 3. Figure 2 shows scatter plots for the pixel-by-pixel comparisons of the LAI products with the spatially-degraded Landsat-derived LAI map from the Shetsky validation site in Central Kazakhstan dominated by grassland, while Figure 3 presents the scatter plots from the Almaty validation site in southern Kazakhstan, with forest cover being dominant.

The results indicate a generally significant correspondence between both LAI products and the Landsat-derived LAI maps in both validation sites. However, the analyses detect that the LAI3g product has much weaker relations with the Landsat LAI in both validation sites, indicating poorer replication of the ground-based LAI by LAI3g dataset, as compared to the LAI GGRS product. Comparisons of LAI3g with the Landsat-derived LAI are characterized by lower values of $R^2$ and higher values of RMSE in both validation sites. The difference between the LAI3g and LAI GGRS validation results is much more pronounced for the Almaty validation site. Thus, LAI3g reproduces only 25% of the variance...
contained in the Landsat-derived LAI, whereas LAI GGRS reproduces 69% of the variance. The RMSE value of LAI3g is much higher than for LAI GGRS (1.55 and 0.70). In terms of relative RMSE (RMSE% = 98%), the validation results of LAI3g in the Almaty test site detect that the user consistency requirements are not fulfilled (RMSE > 0.5 LAI). For other comparisons, differences between the LAI products and the Landsat-derived LAI maps are lower than 0.5 LAI in terms of RMSE, fulfilling user consistency requirements for LAI estimates [23].

**Figure 2.** Validation of LAI products over grassland in the Shetsky test region for June 2008. Pixel-by-pixel \( (n = 720) \) comparison of the aggregated fine-resolution Landsat-derived LAI map versus the LAI GGRS product (a); and versus the LAI3g product (b).

**Figure 3.** Validation of LAI products over forests in the Almaty test region for August 2010. Pixel-by-pixel \( (n = 846) \) comparison of the aggregated fine-resolution Landsat-derived LAI map versus the LAI GGRS product (a); and versus the LAI3g product (b).

The analyses detect statistically-significant positive values of offset for both LAI3g and LAI GGRS comparisons in both validation sites, indicating a higher estimation of low LAI values in the coarse-resolution LAI products as compared to the Landsat-derived LAI for both sites. The magnitude of difference is quite high for LAI3g in both validation sites (offset value of 0.67 versus 0.33 and 0.91 versus 0.41 for LAI3g and LAI GGRS, respectively). In all analyses, slope values are <1.0, indicating a significantly lower estimation of the high values of Landsat-derived LAI by both LAI3g and LAI GGRS retrievals. As compared to LAI3g, the magnitude of the difference is much lower by LAI
GGRS for both validation sites (slope values of 0.33 versus 0.53 and 0.38 versus 0.63 for LAI3g and LAI GGRS, at Sites 1 and 2 respectively).

The better predictions of LAI by LAI GGRS than by LAI3g retrievals are also reflected in the comparisons of LAI histograms from the validation sites (Figure 4). The histograms reveal that the LAI distribution of the LAI GGRS datasets is generally consistent with that of the Landsat-derived LAI datasets in both validation sites, whereas LAI3g distributions are characterized by significant discrepancies with the independent datasets. In the Shetsky test site, the LAI3g histogram shows a sharp-peaked form with a much higher frequency in the value classes 0.5 < LAI < 0.75, 0.75 < LAI < 1.0 and 1.0 < LAI < 1.25 in comparison to the Landsat-derived LAI histograms, while the LAI classes at the left and right side ranges are occupied only slightly. This is also reflected in statistics of the datasets in the Shetsky test area. LAI3g has a mean value of 0.91 with a standard deviation of 0.2 versus a mean value of 0.76 and a standard deviation of 0.31 for LAI GGRS. The statistics of LAI GGRS are much closer to those of LAI ETM+ (mean value of 0.66, standard deviation of 0.43) than the statistics of LAI3g. In the Almaty test site, the LAI3g histogram is characterized by very low frequency of pixels in the lowest LAI class (0 < LAI < 0.25) and excessively high frequency in the third and fourth LAI classes (0.5 < LAI < 0.75 and 0.75 < LAI < 1.0), which is contrary to the Landsat based histogram. The LAI GGRS histogram demonstrates much more similarity with the LAI ETM+ histogram in all value classes (Figure 4b). The LAI value range of LAI3g estimates (0 ≤ LAI ≤ 4.7) is significantly lower than that of the ETM+ LAI reference dataset (0 ≤ LAI ≤ 6.5), highlighting the relatively poor capability of LAI3g to replicate the LAI ETM+ values.

The data distribution is also examined comparing the ECDFs from both validation sites (Figure 4c,d). Generally, for both test sites, the ECDF of LAI GGRS shows much better correspondence with the ECDF of the Landsat-derived LAI than LAI3g. This is reflected in values of the Kolmogorov–Smirnov statistic (DKS), which are considerably lower for the LAI GGRS product. LAI3g demonstrates larger differences in ECDF shapes to Landsat-derived LAI; they are particularly noticeable at the Shetsky test site (Figure 4c). However, the two-sample Kolmogorov–Smirnov test at the 5% significance level accepted the null hypothesis that ECDFs have the same continuous distribution, revealing a general agreement of both LAI products with the Landsat-derived LAI.

In order to support the analyses of the regressions and histograms, we also conducted F-tests of variance equality. Results of the F-tests are shown in Table 5. The results revealed a statistically-significant correspondence between the compared datasets: for all cases, the obtained F-value was lower than the critical F-value, detecting that the null hypothesis (Ho = datasets have similar variance) should be accepted.

In order to support the analyses of the regressions and histograms, we also conducted F-tests of variance equality. Results of the F-tests are shown in Table 5 and are valid at the 95% confidence level. The results revealed a statistically-significant correspondence between the compared datasets, except LAI3g for the Almaty test site. Here, the obtained F-value was higher than the critical F-value, detecting that the null hypothesis (Ho = datasets have similar variance) should be rejected.
**Figure 4.** Frequency distributions (a,b) and empirical cumulative distribution functions (c,d) of Landsat-based LAI, LAI GGRS and LAI3g products at the Shetsky and Almaty validation sites.

**Table 4.** Validation results of LAI3g and LAI GGRS datasets *versus* the ground-based LAI up-scaled using Landsat imagery.

<table>
<thead>
<tr>
<th>Validation Site/Dominate Land Cover</th>
<th>Month/Year</th>
<th>LAI Product</th>
<th>Offset</th>
<th>Slope</th>
<th>$R^2$</th>
<th>RMSE/RMSE%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shetsky/Grassland</td>
<td>June/2008</td>
<td>LAI3g</td>
<td>0.67</td>
<td>0.33</td>
<td>0.44</td>
<td>0.37/49%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LAI GGRS</td>
<td>0.33</td>
<td>0.53</td>
<td>0.61</td>
<td>0.24/31%</td>
</tr>
<tr>
<td>Almaty/Forest</td>
<td>August/2010</td>
<td>LAI3g</td>
<td>0.91</td>
<td>0.38</td>
<td>0.25</td>
<td>1.55/98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LAI GGRS</td>
<td>0.41</td>
<td>0.63</td>
<td>0.69</td>
<td>0.71/43%</td>
</tr>
</tbody>
</table>

**Table 5.** Test statistics and results of the $F$-test conducted for a comparison of the LAI3g and the LAI GGRS product *versus* the ground-based up-scaled Landsat LAI. The null hypothesis of variance equality is accepted when $F$-value < critical $F$-value.

<table>
<thead>
<tr>
<th>Validation Site</th>
<th>LAI product</th>
<th>Mean LAI</th>
<th>Variance</th>
<th>$F$-value</th>
<th>Critical $F$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shetsky</td>
<td>Landsat LAI</td>
<td>0.74</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAI3g</td>
<td>0.92</td>
<td>0.07</td>
<td>0.23</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>LAI GGRS</td>
<td>0.75</td>
<td>0.12</td>
<td>0.18</td>
<td>0.88</td>
</tr>
<tr>
<td>Almaty</td>
<td>Landsat LAI</td>
<td>1.71</td>
<td>1.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAI3g</td>
<td>1.52</td>
<td>1.01</td>
<td>0.90</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>LAI GGRS</td>
<td>1.51</td>
<td>1.17</td>
<td>0.68</td>
<td>0.88</td>
</tr>
</tbody>
</table>
4.2. Comparison of LAI3g and LAI GGRS

The following comparison between the LAI3g and the LAI-GGRS is grouped in terms of spatial and temporal components. For spatial comparison, we used the following three procedures. First, the spatial frequency is compared followed by the spatial analysis based on per pixel comparison being computed. Second, pixel-by-pixel comparisons of both LAI products for the year 2008 were carried out using regression analysis and corresponding scatter plots. Root mean square error (RMSE), the coefficient of determination ($R^2$), the slope of the regression line and intercept were used as guides of statistical association between the LAI products. To assess spatial agreement between the products, AC was calculated for each of the months. The analysis of the temporal evolution of AC aimed to highlight specific periods of spatial disagreement. In addition to these analyses, $F$-tests were conducted for each month from 2008 to test the equality of variances between monthly LAI products. The $F$-tests included all pixels within the territory of Kazakhstan. The null hypothesis that implies the statistically-significant equality of the datasets was rejected when an $F$-value was higher than a critical $F$-value.

In addition to these analyses, $F$-tests were conducted for each month from 2008 to test the equality of variances between monthly LAI products. The $F$-tests included all pixels within the territory of Kazakhstan. The null hypothesis that implies the statistically-significant equality of the datasets was rejected when an $F$-value was higher than a critical $F$-value. Third, differences between LAI3g and LAI GGRS datasets for the individual months over the year 2008 were produced, and LAI difference maps were analyzed. For temporal comparison, we produced seasonal trajectories of LAI3g and LAI GGRS for different land cover types during the period 2005–2008.

4.2.1. Spatial Frequency

In general, the analyzed LAI histograms revealed similar trends between the LAI datasets. Nonetheless, there also are some important differences between the datasets, which are reflected in Figure 5, which shows the LAI histograms for LAI3g and LAI GGRS products in four specific months that cover the main vegetation growing period (April to October) over the area of Kazakhstan. Frequency is given as the total number of pixels over the area of Kazakhstan.

For both LAI products, the frequency of the lowest LAI values is very high at the start and the end of the growing season (Figure 5a,d). It decreases during summer months (Figure 5c) and has the minimum at the peak of the growing season (Figure 5b). These results are consistent with other LAI products [24,25] and reflect the seasonality of vegetation growth in temperate climate zones. The largest histogram differences between the two LAI products occur at the beginning of the growing season during April (Figure 5a), when LAI3g shows a much lower frequency of LAI retrievals within the lowest LAI class ($0 < \text{LAI} < 0.25$) than does LAI GGRS. The second lowest LAI class ($0.25 < \text{LAI} < 0.5$) of the LAI3g histogram is characterized by a much higher frequency as compared to LAI GGRS. Most of the LAI3g pixels belong to this LAI class, whereas LAI GGRS has greatest number of its pixels within the lowest LAI class. The LAI3g product continually, and significantly, exceeds LAI GGRS, also at the end of the growing season in October (Figure 5d). LAI3g shows a much lower frequency of retrievals with LAI < 0.25 and a much higher frequency of retrievals with LAI > 0.25 than does LAI GGRS. This means that the LAI3g product
tends to produce generally higher LAI values at the beginning and at the end of the growing season than does LAI GGRS.

Figure 5. Comparison of the LAI3g/GGRS-LAI value distribution over Kazakhstan: distribution histograms for (a) April, (b) June, (c) August, (d) October (e) and (f) empirical cumulative distribution functions (ECDFs). Number of samples: 48,000. In (e), the months from the growing season (April–October) and in (f) the months from the leafless phase (November–March) are pooled together. Number of samples: 48,000.

The June histograms represent the maximum LAI values over the Kazakhstan territory (Figure 5b); therefore, among all months, they are characterized by the broadest range of values. Both LAI3g and LAI GGRS June histograms show a log-normal distribution. However, there are some differences between the June histograms: the LAI3g histogram shows a sharper kurtosis, while the LAI GGRS histogram shows a continuous and smooth frequency distribution. Most LAI3g pixels are concentrated
in the lowest four LAI classes (from 0.25 to 1.5), whereas the other LAI classes (from 2.0 to >6.0) are occupied more rarely. As compared to LAI3g, the LAI GGRS shows lower frequency in the lower LAI classes and higher frequency in all other classes (Figure 5b). The greatest differences existing between the two LAI histograms is observed in the peak LAI class (0.25 < LAI < 0.5): here, the pixel frequency of the LAI GGRS product is almost two-times lower than that of the LAI3g product. The differences between the June histograms are also reflected in the calculated statistics: the histogram kurtosis of LAI3g has a value of 8.7 versus 5.2 for the LAI GGRS histogram; the histogram shift of LAI3g is 2.5 versus 2.0 for the LAI GGRS histogram. These results show that, in contradiction to a sharp-peaked histogram of LAI3g, LAI GGRS is characterized by a flattened distribution of pixels within the LAI classes. Other statistics also confirmed some significant differences between the June LAI datasets. Thus, the June LAI3g is characterized by a mean value of 0.78 ($p < 0.01$) and a standard deviation of 0.69, while the June LAI GGRS has a mean of 1.03 ($p < 0.01$) and a standard deviation of 0.94. The August histograms also reveal some differences between both LAI products (Figure 5c); however, these differences are not as significant as for the April and October histograms. All in all, the histogram analyses detected a higher similarity between both LAI products during the summer months and a lower similarity between LAI3g and LAI GGRS at the beginning and at the end of the growing season. The overall data distribution is examined comparing the ECDFs, where the months from the growing season and from the leafless phase (November–March) are pulled together (Figure 5e,f). For the growing season months, LAI3g values are lower than the corresponding LAI GGRS for the first left-side class (0 < LAI < 0.25) and somewhat higher for the middle-range classes (0.75 < LAI < 3.5). For the leafless period, differences between LAI3g and LAI GGRS are enormous. The LAI GGRS ECDF contains the greatest number of the pixels (over 80%) in the first class (LAI = 0) and 14% of all pixels in the second class (0 < LAI < 0.1), whereas the LAI3g ECDF shows that the greatest number of pixels is contained in the second and third classes (0 < LAI < 0.1, 0.1 < LAI < 0.25, respectively), and the first class contains only 23%. This means that the LAI3g product detects green vegetation in the most part of the territory of Kazakhstan, whereas the LAI GGRS product does not. Quantitatively, differences in ECDFs are described by the Kolmogorov–Smirnov statistic (DKS) presented in Figure 5. For the growing season ECDFs, the two-sample Kolmogorov–Smirnov test at the 5% significance level accepted the null hypothesis that ECDFs are from the same continuous distribution. For the leafless period, the null hypothesis was rejected, detecting statistically-significant differences between the LAI3g and LAI GGRS distributions.

4.2.2. Spatial Agreement between the LAI3g and GGRS-LAI

Pixel-by-pixel comparisons of both LAI products for the year 2008 supported the results of histogram analyses (Section 4.2.1). Spatial agreement between both LAI products is much lower in April (Figure 6a) and October (Figure 6d) than during the summer months (Figure 6b,c). As compared to the LAI GGRS retrievals, LAI3g tends to a higher frequency of estimates of low LAI values in April and October. The estimation of higher LAI values is particularly pronounced in October, where the regression line slope (0.38) is very different from the 1:1 line. As for April, the regression line slope (0.74) is somewhat closer to the 1:1 line. However, the April scatter plot displays more considerable scattering of data points
around the regression line, detecting worse correspondence between both LAI products in this month ($R^2 = 0.26$ for April versus $R^2 = 0.42$ for October).

Contrary to the April and October comparisons, the June and August scatter plots (Figure 6b,c) demonstrate a better agreement between LAI retrievals of the products. The regression line slope is very close to a value of 1.0 for both June and August (1.05 and 0.93, respectively), indicating good correspondence between both LAI products in the whole range of LAI retrievals. Nonetheless, some differences were detected between the products as detected in the histogram analyses (Section 4.2.1). These differences are highlighted by the scattering of data points.

![Figure 6](image)

**Figure 6.** Comparison of LAI3g and LAI GGRS over the whole area of Kazakhstan for specific months in 2008: (a) April; (b) June; (c) August; and (d) October. Number of samples: 48,000.

Both June and August comparisons revealed much better correspondence between the LAI datasets with RMSE values of 0.42 (43%) and 0.35 (48%) than the April and October retrievals. For these months, differences between LAI3g and LAI GGRS are lower than 0.5 LAI in terms of RMSE, fulfilling user consistency requirements for LAI estimates from multiple sensors [22]. RMSE values of the April (0.22% or 66% in terms of relative RMSE) and October (0.21% or 92%) analyses are higher than 0.5LAI, detecting that the user consistency requirements are not fulfilled in these months.

Figure 7 reports the temporal profiles of the agreement coefficient (AC) and the RMSE. Both statistics demonstrate similar dynamics with a pronounced seasonality. Periods of minimum agreement between LAI3g and LAI GGRS are found during wintertime, while spring/autumn months outside the growing
season are characterized by a low agreement. In fact, during these periods, the two datasets provide quite different LAI dynamics. In particular, during wintertime LAI GGRS minima, LAI3g presents increasing LAI values. This overestimation by the LAI3g product can also be seen in the temporal profiles of LAI from different land cover classes (see section 4.2.4). A possible explanation for this effect may be that, during wintertime and months with a temperature <0 °C, LAI GGRS applies a threshold of LAI = 0 for all non-woody land covers, suggesting no existence of green vegetation during the leafless phase. Obviously, LAI3g does not apply any filter or threshold that could adequately deal with imaginary vegetation during the leafless phase.

![Figure 7. Temporal profiles of the agreement coefficient (AC) and RMSE/mean LAI between LAI3g and LAI GGRS for 2008 over Kazakhstan.](image)

The above results are supported by the results of the F-tests (Table 6). The growing season months (April–September, except October) are characterized by statistically-significant similarity of LAI variance in both products. In contradiction to these, the LAI products of the months outside the growing season (October–March) reveal statistically-significant differences. Further analyses below (see Section 4.2.3 and Figure 7) provide additional evidence for the similarity/difference of the LAI products.

4.2.3. Spatial Differences between the LAI Products

Spatial differences between LAI3g and LAI GGRS products at the pixel scale for specific months in 2008 are shown in Figure 8. Thus, in Figure 8a, we see a very regular distribution of green-colored pixels (0.056–0.13) over the whole territory of Kazakhstan with the exception of some pixels presented as the yellow class (0.14–0.23) in the forested areas in the southern and eastern parts of the country. A similar picture is apparent for the other winter months. In general, this is because LAI3g tends to show values of LAI > 0 for grassland and semi-desert biomes in winter months, while LAI GGRS does not (compare also with Figure 9).
Table 6. Test statistics and results of the $F$-test carried out for a comparison of the LAI3g versus the LAI GGRS product. The number of samples used for the comparison is 48,000. The null hypothesis is accepted for marked months ($F$-value < critical $F$-value).

<table>
<thead>
<tr>
<th>Month</th>
<th>LAI Product</th>
<th>Mean LAI</th>
<th>Variance</th>
<th>$F$-value</th>
<th>Critical $F$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>LAI3g/LAI GGRS</td>
<td>0.085/0.005</td>
<td>0.0049/0.0005</td>
<td>9.35</td>
<td>1.01</td>
</tr>
<tr>
<td>February</td>
<td>LAI3g/LAI GGRS</td>
<td>0.088/0.0038</td>
<td>0.0061/0.0005</td>
<td>12.45</td>
<td>1.01</td>
</tr>
<tr>
<td>March</td>
<td>LAI3g/LAI GGRS</td>
<td>0.327/0.034</td>
<td>0.0694/0.0077</td>
<td>8.99</td>
<td>1.01</td>
</tr>
<tr>
<td>April</td>
<td>LAI3g/LAI GGRS</td>
<td>0.343/0.225</td>
<td>0.0387/0.0610</td>
<td>0.63</td>
<td>0.98</td>
</tr>
<tr>
<td>May</td>
<td>LAI3g/LAI GGRS</td>
<td>0.644/0.695</td>
<td>0.344/0.659</td>
<td>0.522</td>
<td>0.98</td>
</tr>
<tr>
<td>June</td>
<td>LAI3g/LAI GGRS</td>
<td>0.788/1.032</td>
<td>0.476/0.884</td>
<td>0.539</td>
<td>0.98</td>
</tr>
<tr>
<td>July</td>
<td>LAI3g/LAI GGRS</td>
<td>1.076/1.069</td>
<td>1.632/1.856</td>
<td>0.879</td>
<td>0.98</td>
</tr>
<tr>
<td>August</td>
<td>LAI3g/LAI GGRS</td>
<td>0.435/0.377</td>
<td>0.225/0.297</td>
<td>0.762</td>
<td>0.98</td>
</tr>
<tr>
<td>September</td>
<td>LAI3g/LAI GGRS</td>
<td>0.421/0.321</td>
<td>0.213/0.187</td>
<td>1.00</td>
<td>1.01</td>
</tr>
<tr>
<td>October</td>
<td>LAI3g/LAI GGRS</td>
<td>0.217/0.074</td>
<td>0.031/0.011</td>
<td>3.00</td>
<td>1.01</td>
</tr>
<tr>
<td>November</td>
<td>LAI3g/LAI GGRS</td>
<td>0.181/0.049</td>
<td>0.022/0.011</td>
<td>2.177</td>
<td>1.01</td>
</tr>
<tr>
<td>December</td>
<td>LAI3g/LAI GGRS</td>
<td>0.136/0.014</td>
<td>0.0086/0.0021</td>
<td>4.11</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Figure 8. Cont.
Figure 8. Differences (LAI3g minus LAI GGRS) between the compared LAI products for specific months in 2008: (a) January; (b) April; (c) June; (d) August; and (e) October.

LAI3g is characterized by sensitivity to an earlier and more rapid start of the growing season in grassland/shrubland than LAI GGRS. This is reflected in the considerably greater values of LAI3g for these biomes in spring months (see, as an example, April in Figure 8b), whereas LAI of both products for forest areas is characterized by only very small differences. In April, LAI3g of grasslands estimated values of LAI >> 0.5, while LAI GGRS shows significantly lower values (0.2 < LAI < 0.5) (compare also Figure 9). This leads to a considerable broad-scale overestimation of LAI GGRS by the LAI3g product.
LAI values of forest areas in the southern (mountainous) and northern parts of Kazakhstan seem to have only little differences between both LAI products.

The summer months (June and August in Figure 8c,d) are characterized by diverse spatial patterns in the differences between LAI3g and LAI GGRS. Thus, large areas of mixed grasslands in northern Kazakhstan demonstrate in June somewhat lower values of LAI3g than LAI GGRS (yellow color), whereas other areas of grassland are characterized by a small difference between LAI GGRS and LAI3g (reddish color). Nonetheless, some areas of LAI3g show considerably higher values of LAI than LAI GGRS (red color). These areas are mostly located in cropland in the southern and middle parts of Kazakhstan.

Autumn, similarly to the spring, is characterized by significantly higher values of LAI3g in comparison to LAI GGRS (Figure 8e). Higher values of LAI estimates presented by LAI3g range from 0.03 to 0.16 (green color) in the short grassland (central part of Kazakhstan) to 0.17–0.36 (yellow color) in the mixed and long grassland in the northern Kazakhstan. The reason for the generally higher values of LAI3g for autumn is the considerably slower decrease of LAI in the last phase of the growing season. The LAI difference image in Figure 8c supports the scatter plot in Figure 6d, demonstrating the overall overestimation of LAI GGRS by LAI3g.

4.2.4. Temporal Agreement: Seasonal Trajectories

Temporal variations of LAI3g and LAI GGRS over the test sites with the main land cover types of Kazakhstan are compared in Figure 9. The LAI data estimated for the representative test sites over Forest 1, Forest 2 and Reed Grass sites show consistent dynamics between LAI3g and LAI GGRS, but there are some differences in their temporal development. Thus, LAI3g in Forest 1 decreases in the leaf withering period (July–October) somewhat later than does LAI GGRS. For the Forest 2 and Reed Grass sites, the amplitude of LAI3g is much lower than that of LAI GGRS, due to lower values at the peak of the growing season and significantly higher values (LAI = 0.11–0.20) during the leafless phase (November–March).

The test sites in short grassland and mixed grassland (Figure 9d,e) have much less dynamic consistency between the datasets. Both the seasonal magnitude and the timing of the time series are characterized by distinct differences. For the short grass test site, the LAI3g product shows much higher values during the whole vegetation growth period. The typical LAI values for short grassland in the area around the test site are 0.25–0.40 from field observations [19]. The LAI3g retrievals for short grassland are considerably higher. For the mixed grass test site, LAI values at the peak of the growing season are to some extent consistent between both datasets. However, LAI values during the senescence and autumn withering periods are different. During the leafless period, the LAI3g retrievals at the short and mixed grass test sites show LAI values significantly greater than zero (0.1–0.2), while LAI GGRS does not.

Within the study area, the principal mode of vegetation variability is generally associated with intra-annual seasonality of green leaf area. The seasonality in LAI of temperate climate zones is caused by a seasonal increase/decrease in solar radiation and air temperature, which serves as a proximate signal for the start/stop of leaf production. During the period November–March, when the temperature in the study area is below 0 °C, no vegetation activity is possible. The graphs in Figure 9 demonstrate that the LAI3g product tends to ignore the annual dynamics of air temperature over Kazakhstan. The LAI3g product has values above zero for months outside the growing season. A recent publication [26] has also noticed...
stronger differences between the various datasets at the start/end of the growing season. Further investigations, also using time series of \textit{in situ} LAI measurements in the leafless period, are needed to better understand the causes of the observed disagreement and to draw conclusions.

Figure 9. Cont.
Figure 9. Seasonal trajectories of LAI3g and LAI GGRS for test sites in forest (a,b); reed grass (e); short grass (d); and mixed grass (e) during the period 2005–2008.

5. Conclusions

Two coarse-scale LAI products, the global LAI3g and the Kazakhstan national-level LAI GGRS, were compared and validated over the territory of Kazakhstan. The comparison was carried out by means of examining the differences in spatial patterns and spatial frequencies, as well as comparing seasonal LAI trajectories of both products. A direct comparison was enabled using pixel-by-pixel regression analysis and scatter plots of LAI3g versus the LAI GGRS product. After that, fine-resolution ETM+ LAI maps over two test regions in Kazakhstan, optimized by field measurements, were used for the validation of the LAI3g and LAI GGRS retrievals.

Generally, the global LAI3g product explains 25% of the variability in the reference LAI estimates at the forest-dominated validation site in Almaty and 44% at the grass-dominated validation site in Shetsky district of Kazakhstan. The variations explained by the global LAI3g product are much lower than by the regional LAI GGRS product (69% and 68% in the forest and grassland validation sites, respectively), detecting a relatively poor performance of LAI3g in the main land cover types of Kazakhstan. Unfortunately, at the Almaty validation site (forest-dominated land cover), the LAI3g does not fulfil the general user consistency requirements demonstrating a value of RMSE > 0.5LAI. At the Shetsky validation site (grassland-dominated), the user consistency is fulfilled by the LAI3g product (RMSE < 0.5LAI). Nonetheless, LAI estimates of LAI3g have significantly higher values when estimating the lower range of LAI and significantly lower values when estimating the higher range in comparison to the GGRS product. This leads to higher estimates of the mean LAI value in the grassland validation site (0.91 by the LAI3g versus 0.66 by the reference LAI dataset). As compared to the LAI3g product, estimates of LAI GGRS are much closer to the LAI reference datasets, both in the grass-dominated and the forest-dominated validation sites. These facts are revealed in the statistics of the validation results, scatter plots and frequency histograms.

The results have demonstrated that differences between the compared LAI products (LAI3g and LAI GGRS) are most pronounced at the beginning and at the end of the growing season. During the spring and the autumn months, the LAI difference maps showed a considerable difference of LAI GGRS by LAI3g over the most part of the territory of Kazakhstan. Analyses of LAI difference maps and seasonal trajectories revealed that these spring/autumn LAI3g deviations are greater in the semi-desert and steppe grassland land covers, while forest areas demonstrate a much better correspondence between both LAI
products. Generally, LAI3g is characterized by a considerably earlier start and a later finish to the growing season than LAI GGRS. Moreover, LAI3g showed LAI > 0 during the winter months when any green vegetation is absent in all land covers of Kazakhstan. A direct cause for this could be a too high base level of the LAI3g LAI during the leafless phase. Probably, the base LAI level by the beginning and the end of the annual profile is also too high.

From all analyzed months, a reasonable agreement was found between the LAI3g and the LAI GGRS products for the summer months, June and August. This is reflected in the results of the examining of distribution histograms, regressions between both products and analysis of LAI difference maps.

Acknowledgments

We thank Zaichun Zhu and the Remote Sensing team of Boston University for delivering the LAI3g dataset to us.

Author Contributions

These authors contributed equally to this work.

Conflicts of Interest

The authors declare no conflict of interest.

References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).