Supplementary materials for the article "Utility of LAI for Phenology Monitoring of Russian Forests" by Shabanov et al

The objective of materials below is to provide an overview of the accuracy assessment of the IKI land satellite products and highlight their refinements over Russia by comparison to the baseline global products and ground measurements.

A. Intercomparison and validation of the NASA and IKI MODIS LAI products

The IKI LAI product was generated using an improved Collection 6 NASA MODIS LAI/FPAR algorithm for daily retrievals. The daily algorithm was adjusted for better performance at high northern latitudes by extending RT simulations to cover low Solar Zenith Angles, most significant adjustments of simulations were implemented for DNF to better match ground observations. Ancillary input landcover was replaced with IKI landcover product. Daily LAI retrievals were composited and temporally interpolated with IKI algorithms. Changes are discussed in the main text.

As IKI LAI retrievals are based on the original NASA technology, we will rely on accuracy assessment of the baseline NASA product (MCD15). Below we briefly review evaluation/validation of this product, next we intercompare IKI LAI and MCD15 data, following by reports on several recent efforts on evaluation of IKI LAI with independent satellite LAI products and ground measurements over sparse forests.

Validation of the NASA MODIS LAI product has been performed over the past 20+ years. Garrigues et al (2008) reported on the intercomparision with other global satellite LAI da sets (GLOBCARBON and CYCLOPES) and climatology (ECOCLIMAP). RMSE wrt satellite products was 0.82-1.25, R^2 was 0.5-0.74 depending on forest vegetation class. State of the art validation was implemented within the the BigFoot project over a set of plots in the North America (Cohen et al., 2006). Technology included statistical LAI point measurements on the ground and generation of high-resolution 5x5 km reference LAI maps at 30m resolution to allow scaling of point measurements to coarse resolution of the NASA MODIS LAI product. Discrepancy over ENF and DBF forests plots in North America was within 20%. Kobayashi et al (2010) performed large scale intercomparison of their layered LAI product with Collection 4 MOD15 product over DNF of eastern Siberia. They found systematic overestimation of seasonal maximum (RMSE 1.24).

Evaluation of the latest version of the NASA MODIS LAI product (Collection 6) is reported in Yan et al., (2016). Multi-level comparison with global satellite LAI products (GLASS, CYCLOPES and GEO1), validation over the BELMANIP network plots and correlation analysis with the meteorological data (temperature and precipitation) was performed. Summary plots on accuracy assessment wrt references are shown in Figures S1 and S2. RMSE wrt satellite product is 0.80, wrt ground measurements is 0.66.

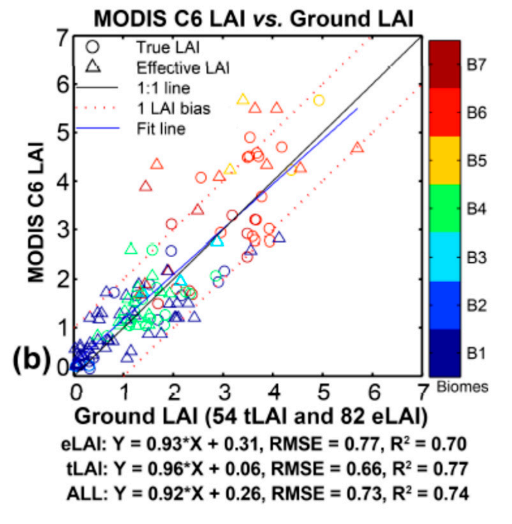


Figure S1. Comparisons between ground measured LAI with MODIS C6 retrievals. Fifty four true LAI, 82 effective LAI and true LAI are used here. The 3 km × 3 km sites dominated by different biome types are depicted by different colors. Circles (triangles) represent ground LAI measurements corrected (not corrected) for clumping. Color coded are biomes: B6- deciduous broadleaf forests, B7- evergreen needleleaf forests, are relevant to this study. Materials are taken from Yan et al. (2016).

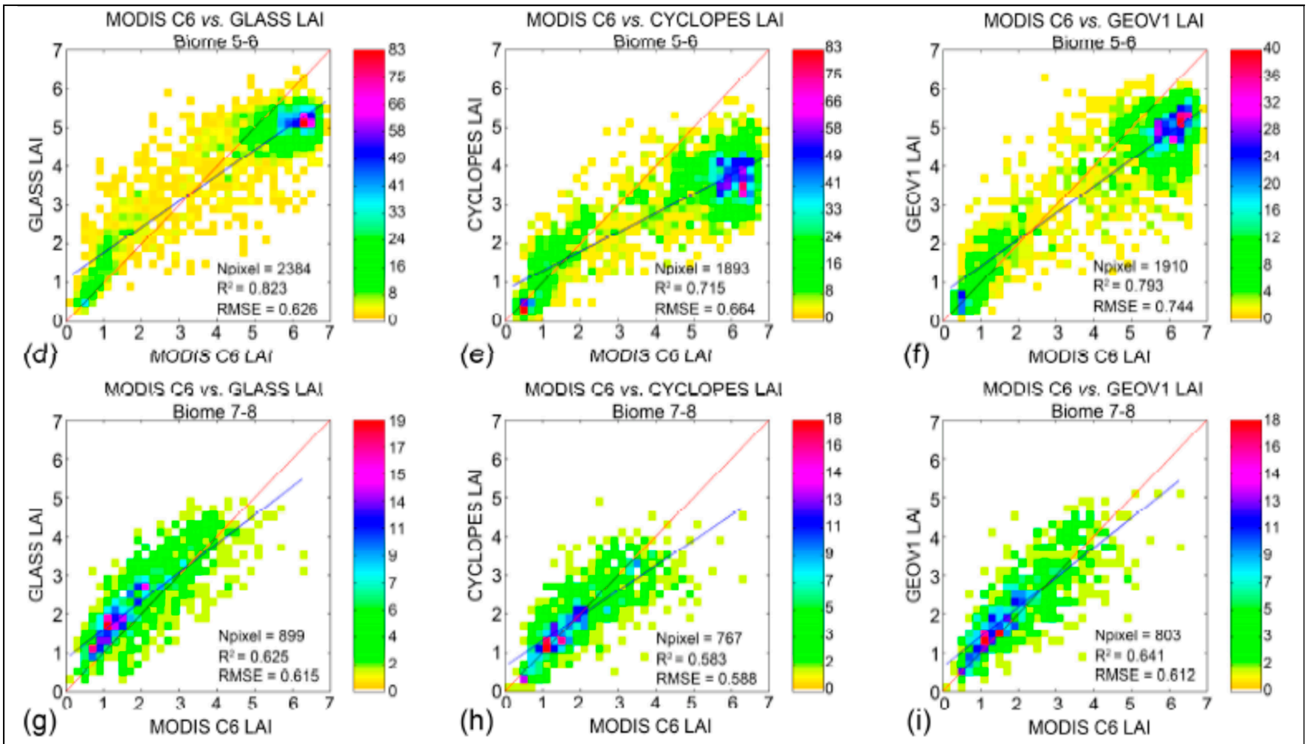
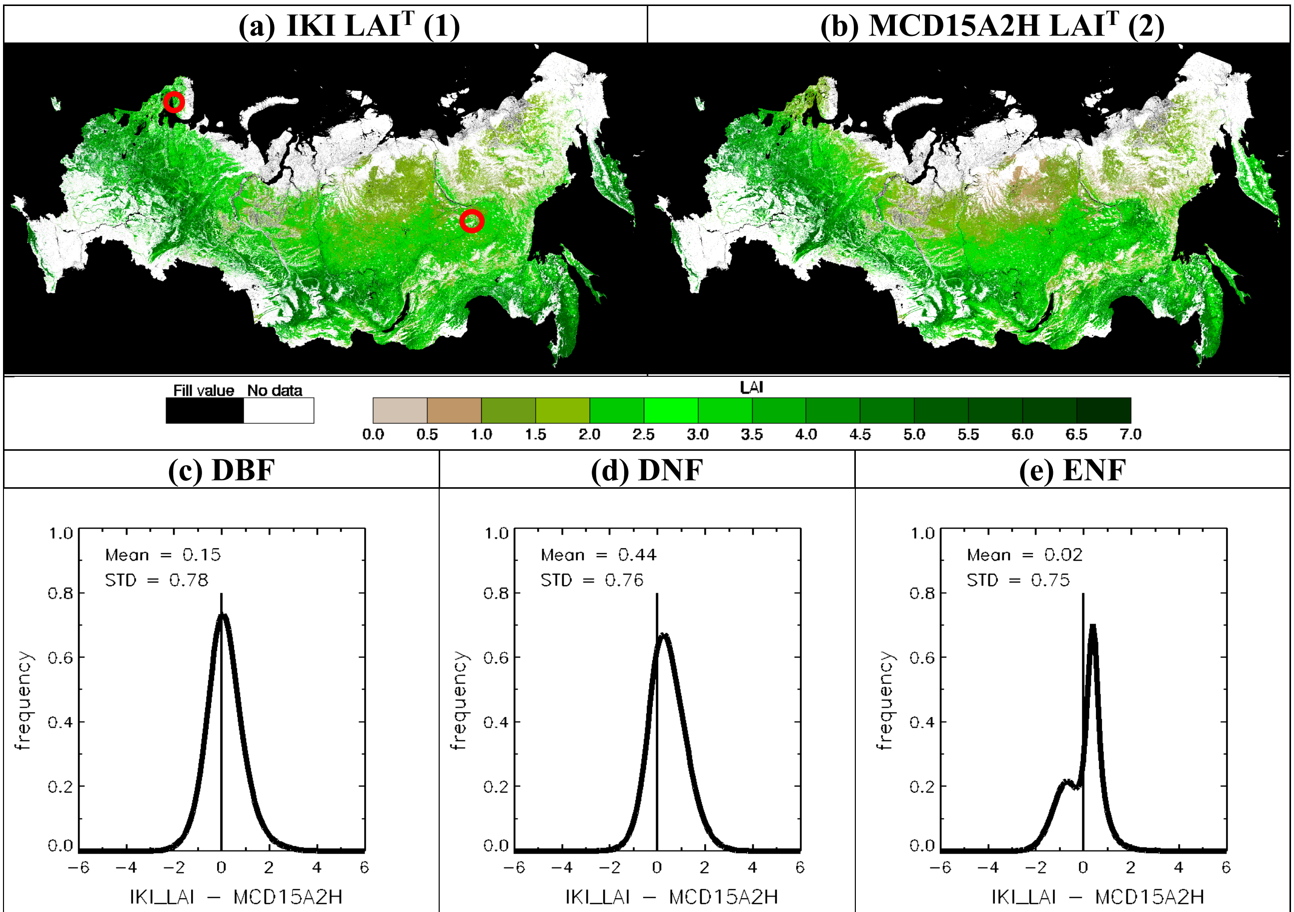


Figure S2. Density scatter plots of monthly MODIS LAI and three other LAI products (left: GLASS; middle: CYCLOPES; right: GEOV1) over BELMANIP sites during the time period from 2001–2005. The plots show a correlation between MODIS and other products for broadleaf forests ((d–f) Biomes 5 and 6) and needle leaf forests ((g–i) Biomes 7 and 8). The red lines and blue lines are the 1:1 lines and regression lines derived from the scatter plots, respectively. Materials are taken from Yan et al., (2016).

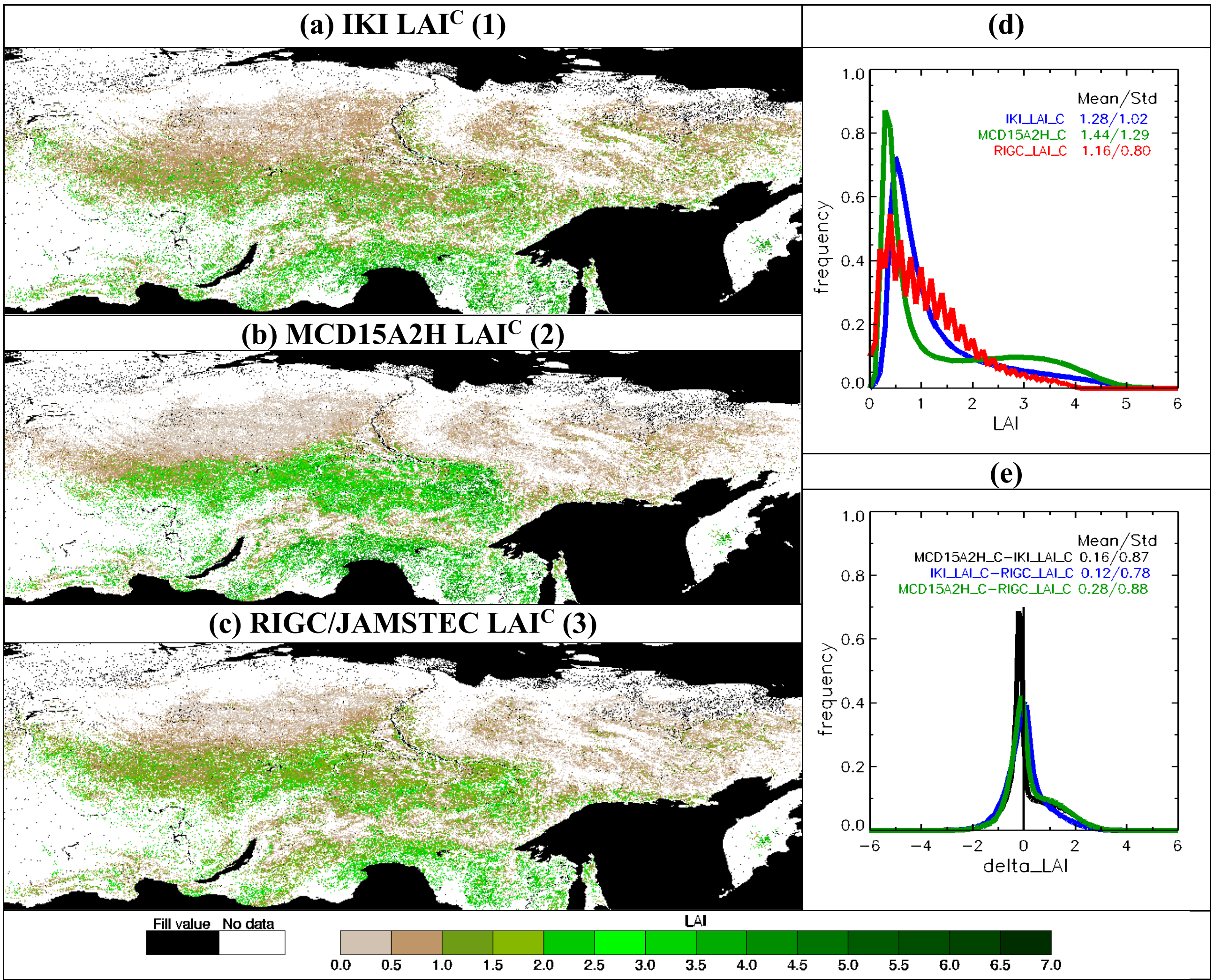
Below we report on analysis of consistency between IKI LAI product and the latest version of the NASA MODIS LAI product (MCD15A2H v. 6.1) over whole extent of the Russian forests. The MCD15A2H is 8-day composited 500-m global LAI product for 2010. We intercompared total forest LAI. Both products were averaged over the growing season (JJA) and gridded to common projection, Albers at 230m. Forest mask was applied, based on the IKI forest species product. Statistics of comparison were calculated both for all forest pixels and separately for the three forest classes (DNF, ENF, DBF). Results are shown in Figure S3. Spatial distribution of products is fairly similar, however most changes are noticeable over Siberia: while the MCD15A2H exhibits significant latitudinal gradients, latitudinal changes in the IKI LAI are much lower. Histograms of LAI differences between products indicate that the IKI LAI is slightly higher than MCD15A2H LAI. Most significant discrepancy is observed for DNF, due to adjustment of RT simulations in the IKI algorithm. Overall consistency statement: absolute difference, mean/std=0.17/0.76, while relative difference 9.2%/28.21%. Given accuracy of the MCD15A2H product of <20% and discrepancy between the IKI and NASA products of 9.2% over forests of Russia, and also fair correlation (R^2/RMSE=0.64/0.81) we can state that products are fairly consistent and large scale accuracy level of forest total LAI is similar.



	Regression equation	R^2	RMSE	pix #
1:2, DBF	Y=1.09768 + 0.70068*X	0.44737	0.79943	34,762,320
1:2, DNF	Y=0.49117 + 0.72720*X	0.50388	0.88665	49,489,271
1:2, ENF	Y=0.04783 + 0.97109*X	0.56599	0.76912	77,867,880
1:2, ALL	Y=0.30062 + 0.84273*X	0.64029	0.81311	162,119,471

Absolut Mean/STD (1-2,ALL) = 0.17/0.76
Relative Mean/STD (1-2,ALL) = 9.2%/28.21%

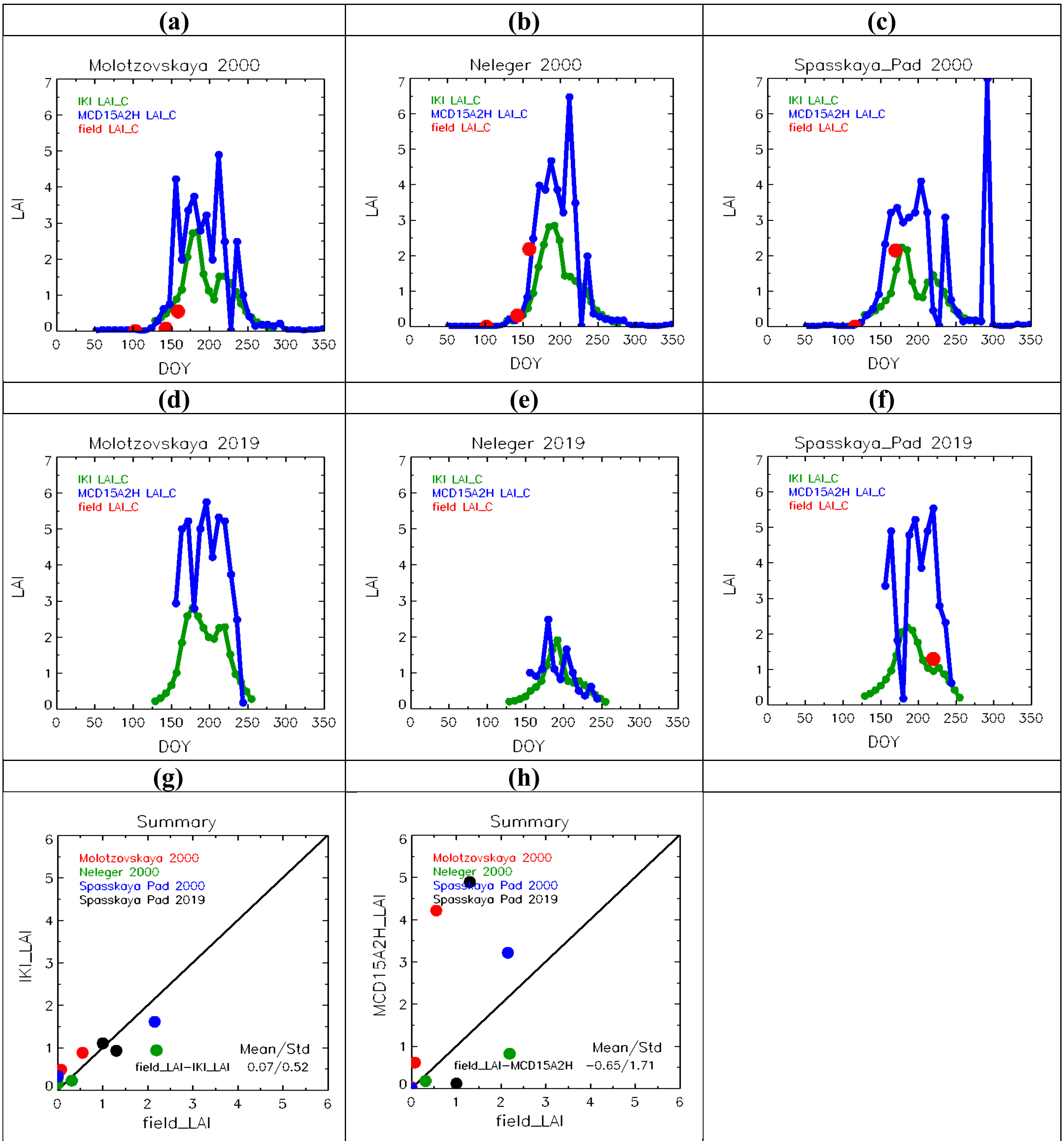
Figure S3. Intercomparison of the IKI LAI^T (set 1) and MCD15A2H LAI^T(set 2) over Russian forests. Data were averaged over the growing season (June-July-August, JJA) for 2010. Spatial distribution is shown in Panels (a)-(b). Red circles denote two sites with sparse forest, where recent validation was performed using ground measurements (cf. Figures S5 and S6). Data are in Albers projection at 230m. Fill values (white) correspond to non-forested areas. Most significant differences are in Siberia, where UMD TCC often underestimates tree crown coverage (cf. Fig. A for examples). Histograms and statistics for three classes (DBF, DNF and ENF) are shown in panels (c)-(e).



	Regression equation	R^2	RMSE	pix #
1:3	Y=0.50368 + 0.51466*X	0.4232	0.7941	6,786,922
2:3	Y=0.49887 + 0.46154*X	0.5475	0.9224	6,786,922
1:2	Y=0.23851 + 0.93697*X	0.5457	0.8857	6,786,922

Figure S4. Intercomparison of the crown LAI (LAI^C) from IKI LAI (set 1) and MCD15A2H LAI (set 2) and RIGS/JAMSTEC LAI (set 3) products over forests of eastern Siberia. LAI data were averaged over the growing season (JJA) for 2010. Data were screened to retain pixels occupied by larch. Spatial distribution of IKI, MCD15A2H and RIGS/JAMSTEC LAI^C products are shown in panels (a)-(c). Maps are shown in Geographic projection at 1/112° resolution. Comparison of histograms of three data sets is presented in panel d, while histograms of difference between IKI and MCD15A2H and the reference are shown in panel (e).

In view that the former study (Kobayashi et al., 2010) has reported an overestimation of the seasonal LAI variations in the NASA MODIS LAI over DNF in the eastern Siberia we performed intercomparison of the MCD15A2H and IKI LAI products with the reference RIGC LAI from the above study. For consistency of products we converted total LAI from MCD15A2H to crown LAI, using our algorithm. Results of intercomparison are shown in Figure S4. The MCD15A2H product features much stronger latitudinal gradients compared to RIGC and IKI LAI data sets (Figures S4a-c). This is also seen in the histograms- histogram of MCD15A2H product indicates underestimation for low values and overestimation at the high end compared to those of the other two products. The IKI LAI product show better consistency with RIGC LAI.



Source	Regression equation	R^2	RMSE	pix #
1:2	Y= 0.33726 + 0.46664 *X	0.73097177	0.501683	10
1:3	Y= 0.62951280+ 1.0237254*X	0.21326333	1.74399	10

Figure S5. Intercomparison and validation of the IKI LAI^C (set #1) and MCD15A2H LAI^C (set #2) data at the local scale of three forest sites near Yakutsk city in eastern Siberia- Molotzovskaya (62.25°N, 130.9112583°E), Neleger (62.31416667°N, 129.4975°E) and Spasskaya Pad (62.255°N, 129.6188°E) (set #3)The dominant forest species for all sites is larch. Seasonal LAI profiles are shown for 2000 and 2019.

Still, there exist discrepancy (intercompare R^2, RMSE), arising due to differences in inputs and processing algorithms. Further assessment of the products accuracy was performed at local scale of three plots in vicinity of city of Yakutsk in eastern Siberia- Molotsovskaya, Neleger and Spasskaya Pad. Seasonal profiles of crown LAI from MCD15A2H and IKI LAI were verified with the seasonal ground measurements for 2000 and 2019 (data described in the main article). According to Figure S5 MCD15A2H substantially overestimates seasonal LAI variations, while IKI LAI is fairly close to ground observations. Also, due to temporal interpolation, the IKI LAI product suppresses noise of intercomposite LAI variations, significantly affecting MCD15A2H product. The improvements further can be seen by comparing summary plots for two products: Mean/Std (IKI_LAI-field data) = 0.07/0.52 and R^2/RMSE = 0.73/0.50, while Mean/Std (MCD15A2H -field data) =-0.65/1.71 and R^2/RMSE = 0.2/1.30

The last case study of products intercomparison and accuracy assessment at the local scale was performed at the plot of sparse forest located off the southern foothill of Khibiny Mountains at the Kola Peninsula (Figure S6). Forest cover changes from spruce-birch at north-west (NW) to birch-spruce at south-east (SE) (Figure Sb), resulting in corresponding LAI gradient in this direction. Crown fraction measurements were performed with UAV technology. Crown fraction was converted to crown LAI according to our algorithm. Our algorithm was also used to separate total LAI from MCD15A2H into crown and understory portions. More details on data and processing is shown in Sec 2.3 of the main article. Figure S6c-g show spatial distribution of the IKI LAI^T, MCD15A2H LAI^T, IKI LAI^C, MCD15A2H LAI^C and UAV LAI^C, respectively. While maps don't exhibit an exact match, but all of them feature LAI gradients from NW to SE. Also there is a substantial contrast between the total and crown LAI. Comparison of histograms of total (crown) LAI histograms from MCD15A2H and IKI LAI product are shown in Figure S6h and Figure S6i, respectively. Histograms of differences are shown in Figure S6j. Overall, distribution of total and crown LAI match between products and with reference. IKI LAI shows slight improvements, as MCD15A2H product slightly underestimates ground reference.

Overall, country-wide analysis indicate that IKI MODIS LAI product is generally consistent with the baseline NASA MODIS LAI (MCD15A2H), thus overall accuracy statement of the baseline product is applicable to the IKI version. Still the latter indicate regional improvements, especially over larch forests (DNF) of eastern Siberia and also forests of high northern latitudes.

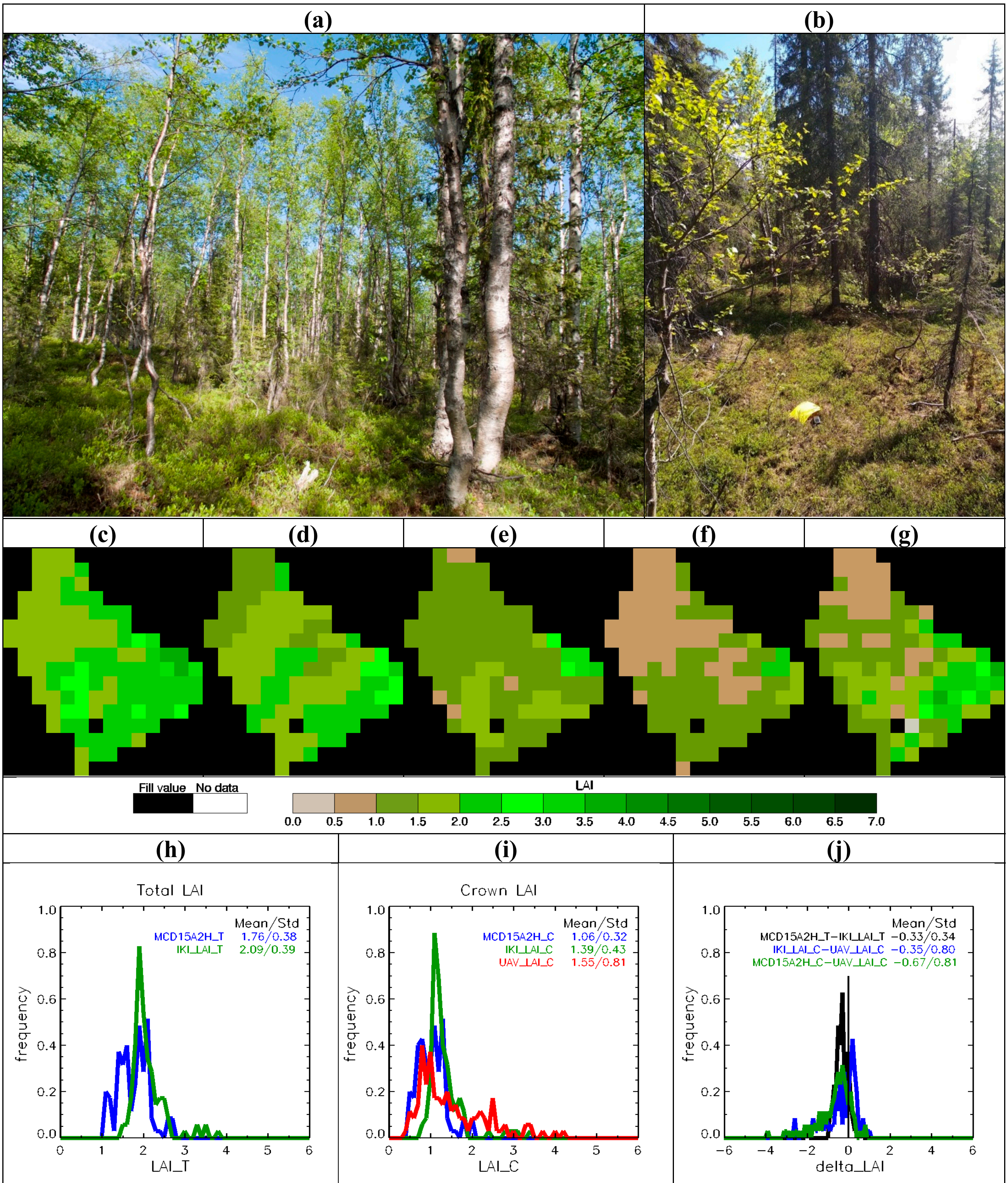


Figure S6. Validation of tree crown LAI at local scale of forest site at Kola Peninsula using UAV-based LAI estimates. Photos demonstrate vegetation variability at the site: dominance of broadleaved forests (at the southern part, panel (a)) and needleleaf forests (at the northern part, panel (b)). Panels (c)-(g) show maps of IKI LAI^T, MCD15A2H LAI^T, IKI LAI^C, MCD15A2H LAI^C and UAV LAI^C, respectively. All data were resampled to UTM projection (WGS-1984, zone 36N at 230m resolution). This site features sparse overstory and substantially developed understory. MCD15A2H and IKI LAI products were validated using UAV-based LAI estimates. Difference between products and reference UAV LAI data is quantified in Table B.

B. Validation of the IKI forest species product

Validation of the IKI forest species product is in progress. For the accuracy assessment we utilize ground measurements, high resolution satellite data and also refer to regional experts. An initial accuracy assessment was performed in terms of fractional areas occupied by species, defined as portion of total forested area of Russia, occupied by particular species (Figure S7). Comparison of estimates by IKI forest species map for 2009 and ground surveys reports from the State Forest Accounting (Gosudarstvenni Lesnoi Reestr) for 1 January 2009 results in R^2=0.99. We also intercompared fractional areas occupied by species in individual federal districts, resulting in the range of R^2 from 0.81 (for aspen) up to 0.98 (for larch).

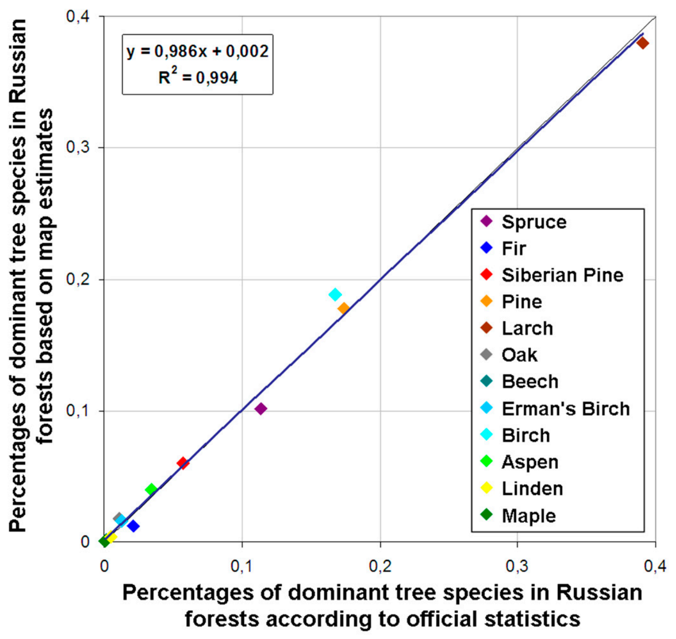


Figure S7. Comparison of fractional areas of dominant tree species estimates by IKI forest species product and ground surveys reports from State Forest Accounting (Gosudarstvenni Lesnoi Reestr) for 1 January 2009.

Recently, validation of the IKI forest species map was implemented using data from country-wide surveys of the Russian Forest Management framework (*Lesoustroistvo*). Those surveys provide information on multiple characteristics of forest cover, including dominant tree species at the forest stand level. Using the stratified random sampling strategy, set of 177 50x50 km raster reference maps (plots) at 230-m resolution was formed, to allow pixel-by-pixel comparison with the IKI forest species map. Spatial distribution of reference plots over territory of Russia and an example of forest species raster derived from the forest stands surveys vector data at one of the plot locations is shown in Figure S8.

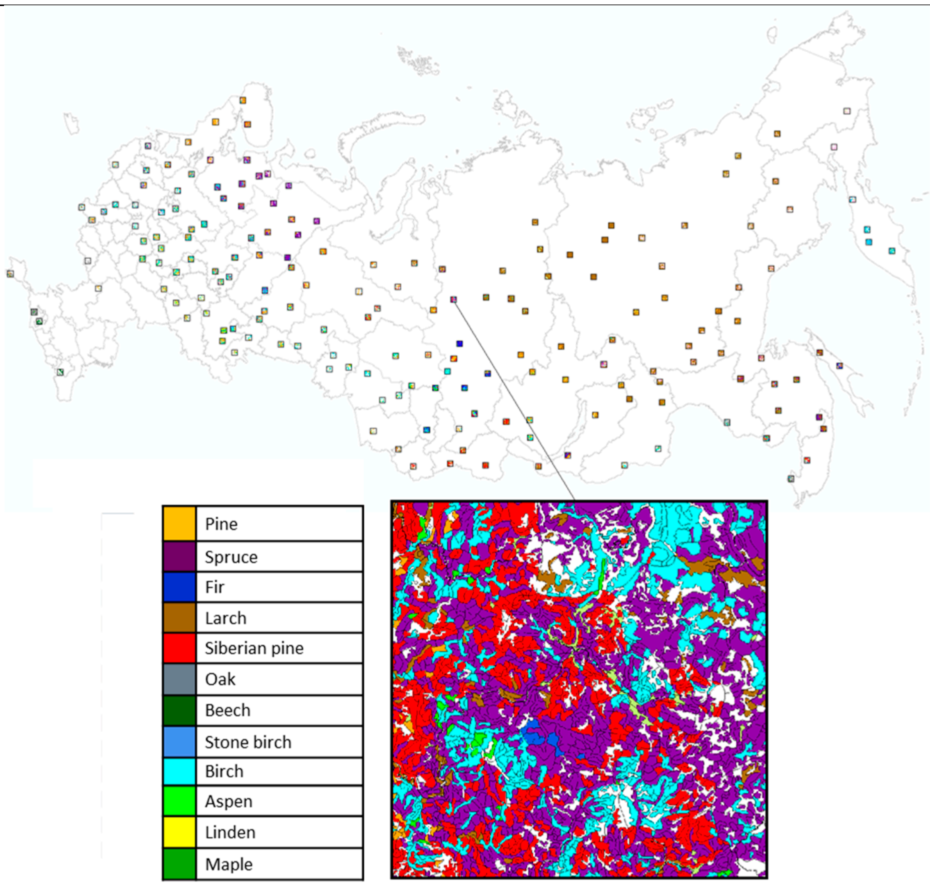


Figure S8. Validation of IKI forest species product using data from country-wide surveys of the Russian Forest Management framework (*Lesoustroistvo*). Survey data were used to generate a set of 177 50x50 km raster reference maps (plot) at resolution of 230 m. Spatial distribution of plots over territory of Russia and an example of forest species raster derived from the forest stands surveys vector data at one of the plot locations is shown.

Data at each plot were pre-processed to minimize data noise. Namely, individual plots contain information on forest species over various years. To avoid bias, pixels with forest cover disturbances over the period 2001-2019 were eliminated from the analysis. To do so, we reference the Global Forest Change product (Hansen et al., 2013), which is ETM data based product, mapping forest loss at spatial resolution of 30-m. Also, to eliminate the impact of the local data heterogeneity on validation of relatively coarse spatial resolution of MODIS data, the reference was also screened with 3x3 moving window to screen-out mixed pixels. If more than single class populates the window, central pixel is set to fill value. In this study the IKI forest species map for 2019 was validated. This map was also screened to retain only those pixels, where results of classification remained unchanged in the current and adjacent years. After applying all filters specified above, total of 1,458,721 of pixels pairs (map-references) has been retained.

Table S1. Confusion matrix between IKI forest species and reference forest stands. Overall accuracy is 83,7%.

		Reference forest stands												Type II error
IKI forest species		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Pine	**1**	**150455**	7100	785	17228	12647	119	0	51	7104	475	1	0	23,2%
Spruce	**2**	5782	**133586**	7395	10876	2859	0	0	47	2316	206	2	0	18,1%
Fir	**3**	74	5879	**21956**	174	2211	0	0	0	2142	956	0	0	34,2%
Larch	**4**	18470	9992	237	**696916**	6969	37	0	383	10243	179	3	0	6,3%
Siberian Pine	**5**	3256	1848	2670	11894	**23429**	38	0	307	899	177	72	0	47,5%
Oak	**6**	82	17	0	12	389	**33631**	478	66	1034	85	173	7	6,5%
Beech	**7**	4	0	113	0	0	2457	**4208**	0	4	11	0	4	38,1%
Stone Birch	**8**	0	104	12	100	0	0	0	**27150**	848	0	0	0	3,8%
Birch	**9**	8266	13307	2354	13684	712	249	188	11118	**118107**	5349	389	27	32,0%
Aspen	**10**	431	515	1350	793	172	57	0	40	12966	**9111**	583	11	65,0%
Linden	**11**	10	21	1	0	1278	767	0	49	122	132	**1860**	36	56,5%
Maple	**12**	0	4	0	0	0	16	0	0	63	34	2164	**951**	70,6%
Type I error		19,5%	22,5%	40,5%	7,3%	53,8%	10,0%	13,7%	30,8%	24,2%	45,5%	64,6%	8,2%	**83,7%**

Confusion matrix shown in the Table S1 quantifies accuracy of the IKI forest species product. Overall map accuracy is 83,7%. In terms of individual species best accuracy is achieved for larch (Type I and II errors are 6.3% and 7.3% respectively). Among needleleaf forests good accuracy is observed for pine and spruce (Type I and II errors 18.1-23.2%). Lower accuracy is observed for fir and Siberian pine (Type I and II errors 34.2-53.8%), mostly due to expected confusion with other dark needleleaf species and pine, respectively. Among broadleaf species best accuracy was achieved for oak (Type I and II errors are 6.3% -7.3%). Other broadleaved species, except linden, have low Type I errors (8.2%-13.7%), but poses higher Type II errors (38.1-70.6%). Linden has the lowest accuracy (Type I and II errors 56.5-64.6%), which is partially due to its low population. Among small-leaved species moderate accuracy is archived for birch (Type I and II errors are 24.2% -32.0%) and stone birch (Type I and II errors are 30.8% -3.8%). Aspen expectedly demonstrate low accuracy due to confusion with birch (Type I and II errors are 45.5% -65.0%).

References

Bartalev S., Egorov V., Zharko V., Loupian E., Plotnikov D., Khvostikov S., Shabanov N. (2016) Land cover mapping over Russia using Earth observation data. Moscow. Russian Academy of Sciences' Space Research Institute, 2016. - 208 p

Cohen, W. B., Maiersperger, T. K. , Turner, D. P., Ritts, W. D. , Pflugmacher, D. , Kennedy, R. E., Kirschbaum, A. , Running, S. W., Costa, M. and Gower S. T. (2006), MODIS land cover and LAI Collection 4 product quality across nine sites in the Western Hemisphere, IEEE Trans. Geosci. Remote Sens., 44, 1843–1857

Garrigues, S., Lacaze, R., Baret, F., Morisette, J.T., Weiss, M., Nickeson, J., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R., Yang, W. (2008). Validation and Intercomparison of Global Leaf Area Index Products Derived from Remote Sensing Data. Journal of Geophysical Research, 113, G02028, doi:10.1029/2007JG000635

Kobayashi, H., Delbart, N., Suzuki, R., and Kushida, K. (2010). A satellite based method for monitoring seasonality in the overstory leaf area index of Siberian larch forest. J. Geophys. Res.-Biogeosci. 115, G01002, doi:10.1029/2009jg000939.

Yan, K., et al., (2016). Evaluation of MODIS LAI/FPAR Product Collection 6. Part 2: Validation and Intercomparison. Remote Sensing, 8, 460, doi:10.3390/rs8060460