NOAA technical memorandum OAR ARL-285

High Resolution Global Land Surface Datasets using Satellite Measurements for Application to Earth System Models

Wei-Ting Hung, Patrick C. Campbell, Barry Baker

September 2024

National Oceanic and Atmospheric Administration
Oceanic and Atmospheric Research
Air Resources Laboratory
Atmospheric Sciences and Modeling Division
College Park, Maryland

U.S. Department of Commerce
Secretary of Commerce Gina M. Raimondo

Under Secretary of Commerce for Oceans and Atmosphere
Richard W. Spinrand, Ph.D.

Assistant Administrator Oceanic and Atmospheric Research
Steven Thur, Ph.D.

NOAA technical memorandum OAR ARL-285
https://doi.org/10.25923/d06p-2333

High Resolution Global Land Surface Datasets using Satellite Measurements for Application to Earth System Models

By Wei-Ting Hung[1,2,3], Patrick C. Campbell[1,2,3], Barry Baker[1]

[1]NOAA Air Resources Laboratory, National Oceanic and Atmospheric Administration, College Park, Maryland

[2]Cooperative Institute for Satellite and Earth System Studies, University of Maryland, College Park, Maryland

[3]Center for Spatial Information Science and Systems, George Mason University, Fairfax, Virginia

September 2024

National Oceanic and Atmospheric Administration
Oceanic and Atmospheric Research
Air Resources Laboratory
Atmospheric Sciences and Modeling Division
College Park, Maryland

Contents

List of Tables

List of Figures

Abstract

A global high resolution land surface dataset (GLSD) with universal latitude/longitude grid coordinate is developed at NOAA Air Resources Laboratory (ARL) for the applications to land, weather, and atmospheric composition models. The GLSDs include key land surface properties based on various operational satellite measurements (e.g. MODIS, VIIRS and GEDI), including land surface type, vegetation clumping index, leaf area index, vegetative canopy height, and green vegetation fraction. Critical processing steps are taken to extend the surface properties for a full global coverage, as well as gap-filling within data points that are not available. All GLSDs are gridded to universal global latitude/longitude gridded coordinates (-90° – 90° in latitude and 0° – 360° in longitude) with a consistent spatial resolution of 0.01° (~ 1 km). The GLSDs are now available at the National Centers for Environmental Information (NCEI; downloaded from https://doi.org/10.25921/qzm2-zg29) for one year period representative of 2020. Given the high spatiotemporal resolution from satellite products, the novel GLSDs can provide more realistic, updated surface information to the earth modeling systems rather than using prescribed surface type dependent tables and parameterizations, which are commonly used across the NOAA Unified Forecast System (UFS) model components (e.g., land, weather, and atmospheric composition models). The major uncertainties in the GLSDs lie in the discrepancies of surface type classification in different satellite products, and the simple climatological approach used for gap filling. To deal with these issues, additional fractional surface type dataset and improved gap filling processes using machine learning techniques are proposed for future developments.

1. Introduction

Currently, many land surface models (e.g. the Noah land surface model; Ek et al., 2003) used in regional and global Earth System Models (ESMs) rely on prescribed surface type dependent tables and parameterizations to constrain vital physical parameters controlling geophysical fluxes due to the description of the land surface. Since the land surface is an essential component in the Earth system, prescribed land surface data may not explicitly represent the realistic land surface properties in space and time, and could lead to potential uncertainties in ESMs such as the land, weather and air composition model components (Fisher et al., 2020; Johannsen et al., 2019; Nogueira et al., 2020).

To reduce such uncertainty, Earth observation systems such as satellites can provide measurements of the land surface characteristics with improved spatial and temporal resolutions. For instance, the Moderate Resolution Imaging Spectroradiometer (MODIS; https://modis.gsfc.nasa.gov/about/) sensors onboard the Aqua and Terra satellites and the Visible Infrared Imaging Radiometer Suite (VIIRS; https://www.nesdis.noaa.gov/our-satellites/currently-flying/joint-polar-satellite-system/visible-infrared-imaging-radiometer-suite-viirs) sensor onboard the Suomi National Polar-orbiting Partnership (S-NPP) satellite have been widely used in Earth science research. The multiple spectrum channels with a wide spectral range from near infrared to ultraviolet allow these sensors to provide various products of Earth land surface properties and vegetative canopy characteristics (Justice et al., 2002, 2013). In addition, the Global Ecosystem Dynamics Investigation (GEDI; https://gedi.umd.edu/), which is a lidar instrument onboard the International Space Station (ISS), provides the two- and three-dimensional structure of vegetative canopies such as forests, including canopy height, biomass density, and the distribution of stems and leaves (Dubayah et al., 2020). Previous studies have demonstrated that applying satellite land surface products to land models would improve the simulations of soil moisture and surface fluxes (Li et al., 2024), even leading to significant improvements in weather predictions such as surface air temperature (Nogueira et al., 2020; Ruiz-Vásquez et al., 2023). Use of seasonal vegetative canopy information from satellite products could better represent the Earth surface more realistically and improve the atmospheric-surface exchanging processes within the boundary layer (van der Graaf, et al., 2020). However, most satellite products have diverse sources and data formats, as well as limitations in their spatiotemporal coverage, which introduces difficulties for researchers and operational centers to integrate them into ESMs.

Hence, NOAA Air Resources Laboratory (ARL) Atmospheric Sciences Modeling Division (ASMD) has developed a global land surface dataset (GLSD) with universal grid coordinates for ESM

applications based on a suite of operational satellite products. Five major land surface properties are included: surface type (ST), vegetation clumping index (CLU), vegetation leaf area index (LAI), vegetative canopy height (CH), and green vegetation fraction (GVF). The temporal resolution of each dataset may be different based on the characteristics of land surface property and the original temporal frequency of satellite products. The ST and CH datasets are on an annual basis since they are relatively consistent in terms of seasonal variability, while the CLU, LAI and GVF datasets are monthly for their seasonality.

2. Global Land Surface Datasets

2.1 Data Processes

Table 1 summarizes the GLSDs included in this dataset, along with data periods, and the original satellite products and spatiotemporal resolution. Because of the lack of sunlight in high latitudes during winter months, the passive MODIS and VIIRS satellite retrievals (e.g., for CLU, LAI, and GVF products) often have no data beyond 60°N and many gaps in mid to high latitudes across the expansive land areas in the North Hemisphere. These datasets are globally extended for high latitude regions to obtain full global coverage. The gap-filling approaches vary by the GLSDs based on the characteristics of different land surface properties. Details on the gap filling are described in the following sections.

Table 1: List of the GLSDs included in this dataset, along with the data period, frequency, original data source, and spatial and temporal resolutions.

Dataset	Data period/Frequency	Original Spatial/Temporal resolution	Data source
Surface type (ST)	2020/Annual	1 km/Annual	VIIRS
Clumping index (CLU)	2001 – 2017/Monthly	500 m/Monthly	MODIS
Leaf area index (LAI)	2020/Monthly	500 m/8-day	VIIRS
Canopy height (CH)	2020/Annual	10 m/Annual	GEDI
Green vegetation fraction (GVF)	2020/Monthly	4 km/Weekly	VIIRS

After global extension and gap filling for certain GLSDs, all GLSDs are gridded to universal global latitude/longitude gridded coordinates (-90° - 90° in latitude and 0° - 360° in longitude) with a spatial resolution of 0.01°. The re-gridding is done by the griddata function from the Python Scipy package (https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.griddata.html), except for ST whose re-gridding method is described in Section 2.2. Example visualization graphics are generated using Python Matplotlib package (Hunter, 2007; https://matplotlib.org/), and the shapefile of the global coastlines overlaid on the graphics is downloaded from Natural Earth Data (https://www.naturalearthdata.com/).

2.2 Surface Type

The annual ST is based on the VIIRS surface type product (Huang et al., 2023; https://www.star.nesdis.noaa.gov/jpss/st.php), which includes 17 ST categories globally based on the International Geosphere–Biosphere Program (IGBP) classification (Belward and

Loveland, 1997). Figure 1 visualizes the annual global ST of 2020 with 17 categories. Please see the cited reference for detailed description of each ST category.

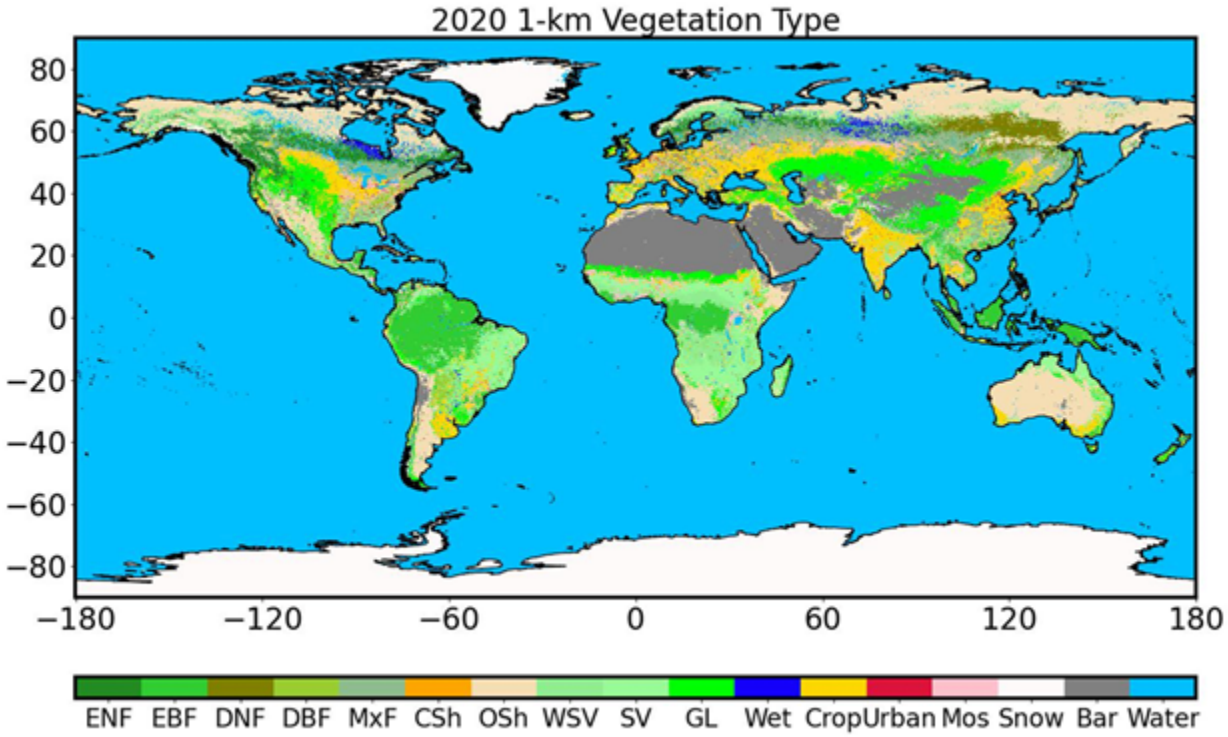

Figure 1. Annual global surface type of 2020. The 17 surface type categories are 1) Evergreen Needleleaf Forests (ENF), 2) Evergreen Broadleaf Forests (EBF), 3) Deciduous Needleleaf Forests (DNF), 4) Deciduous Broadleaf Forests (DEF), 5) Mixed Forests (MxF), 6) Closed Shrublands (CSh), 7) Open Shrublands (OSh), 8) Woody Savannas (WSV), 9) Savannas (SV), 10) Grasslands (GL), 11) Permanent Wetlands (Wet), 12) Croplands (Crop), 13) Urban and Built-up Lands (Urban), 14) Cropland/Natural Vegetation Mosaics (Mos), 15) Snow and Ice (Snow), 16) Barren (Bar), and 17) Water Bodies (Water).

This product is the heritage of the MODIS global land cover product (Friedl et al., 2010; Friedl and Sulla-Menashe, 2015) by following the same retrieval algorithms. Two ST products show a good agreement in general while inconsistencies were found at high latitude regions (Moon et al., 2019; Zhang et al., 2018). The original 1 km coordinate is re-gridded to the global 0.01° x 0.01° coordinate by following a dominant ST approach. The ST value at a grid cell in the new coordinate is determined as the ST value with the most data points from the original coordinate within that grid area. Therefore, the ST dataset provides the dominant terrestrial ST, which describes the surface characteristics of the majority of the land surface within the area of each grid cell.

2.3 Vegetation Clumping Index

The monthly global CLU is based on the MODIS Bidirectional Reflectance Distribution Function (BRDF) product (Wei et al., 2019) downloaded from the GriddingMachine (Wang et al., 2022; https://github.com/CliMA/GriddingMachine.jl). Due to measurement limitation, many missing data are found over mid to high latitude regions in the North Hemisphere during winter. The gaps are filled for high latitude regions by converting summertime values (CLU_s) to wintertime values (CLU_w) based on the ratio of monthly averages in 2003 – 2017 using MODIS analyses from Fang et al. (2021):

$$CLU_w = CLU_s \times R_{w/s} \quad \text{Eq. (1)}$$

where $R_{w/s}$ is the ratio of the CLU average of selected winter month to the CLU average of selected summer month. The monthly averages are ST dependent, which is based on the ST dataset described in Section 2.2. The monthly average in July is used as the summer standard (i.e. CLU_s), and the missing gaps in the North Hemisphere from late fall to early spring (January, February, March, April, May, October, November and December) are filled. Only regions with significant vegetation covered (e.g., ST equals 1 – 12 reported in Fang et al. (2021)) are considered in this process. Figure 2 shows the monthly CLU average in (a) July 2020, (b) January 2020 before gap filling, and (c) January 2020 after gap filling.

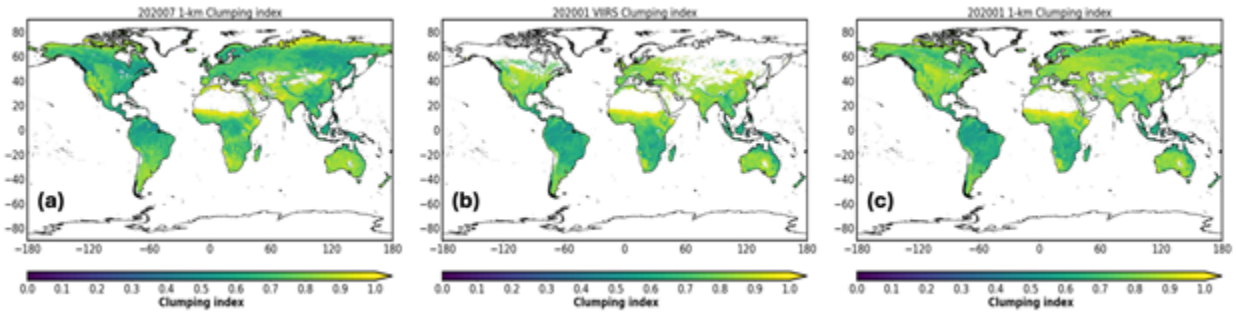

Figure 2. Monthly averages of global vegetation clumping index in (a) July, (b) January before gap filling and re-gridding, and (c) January 2020. The datasets shown in (a) and (c) have been re-gridded to 0.01° coordinate.

Although the cropland/natural vegetation mosaics (Mos; ST = 14, see Figure 1) is not included in this process, the result should not be significantly affected since Mos only accounts for around 1.5% in the North Hemisphere. In addition, Fang et al. (2021) used MODIS land cover type product (Friedl and Sulla-Menashe, 2015; https://lpdaac.usgs.gov/products/mcd12c1v006/) for

ST categorization, while the ST dataset here is based on VIIRS product. Based on the investigations done by previous studies (e.g., Moon et al., 2019) and the same retrieval algorithms used for both products, the ST classifications are assumed to be consistent in two products.

2.4 Vegetation Leaf Area Index

The monthly global LAI is from the VIIRS LAI product (Myneni and Knyazikhin, 2018; https://doi.org/10.5067/VIIRS/VNP15A2H.001). Similar to CLU (Section 2.3), the wintertime high latitude gaps are filled by converting summertime values (LAI_s) to wintertime values (LAI_w) based on the ratio of monthly averages in 2003 – 2017 using MODIS analyses from Fang et al. (2021):

$$LAI_w = LAI_s \times R_{w/s} \quad \text{Eq. (2)}$$

where $R_{w/s}$ is the ratio of the LAI average of selected winter month to the LAI average of selected summer standard month (i.e. LAI_s). Similar to the CLU process, the monthly averages are ST dependent (Section 2.2), and the missing gaps in the North Hemisphere from late fall to early spring (i.e. January, February, March, April, May, October, November and December) are filled for STs equal to 1 – 12 only by assuming a relatively minor contribution from Mos (ST = 14). Figure 3 shows the monthly LAI average in (a) July 2020, (b) January 2020 before gap filling, and (c) January 2020 after gap filling. Note that the Fang et al. LAI climatological monthly averages were also based on MODIS STs, and the assumption of consistent ST classifications from MODIS and VIIRS ST products are made when mapping to VIIRS STs.

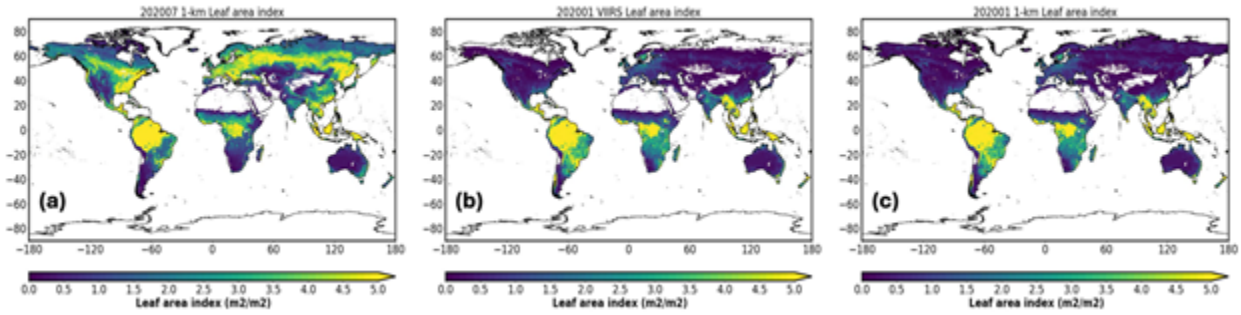

Figure 3. Monthly averages of global vegetation leaf area index in (a) July, (b) January before gap filling and re-gridding, and (c) January 2020. The datasets shown in (a) and (c) have been re-gridded to 0.01° coordinate.

2.5 Vegetative Canopy Height

The annual CH dataset is based on the global GEDI product developed by Lang et al. (2023) and Figure 4 visualizes the annual global CH of 2020. Limited by the space orbit of ISS, the original GEDI measurements only provide a spatial coverage between 52°S and 52°N. In Lang et al. (2023), GEDI CH dataset is extended to a comprehensive global map by fusing it with the satellite images from Sentinel-2 (https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2) using deep machine learning models. Since this product is already globally extended, gap filling process is not required here.

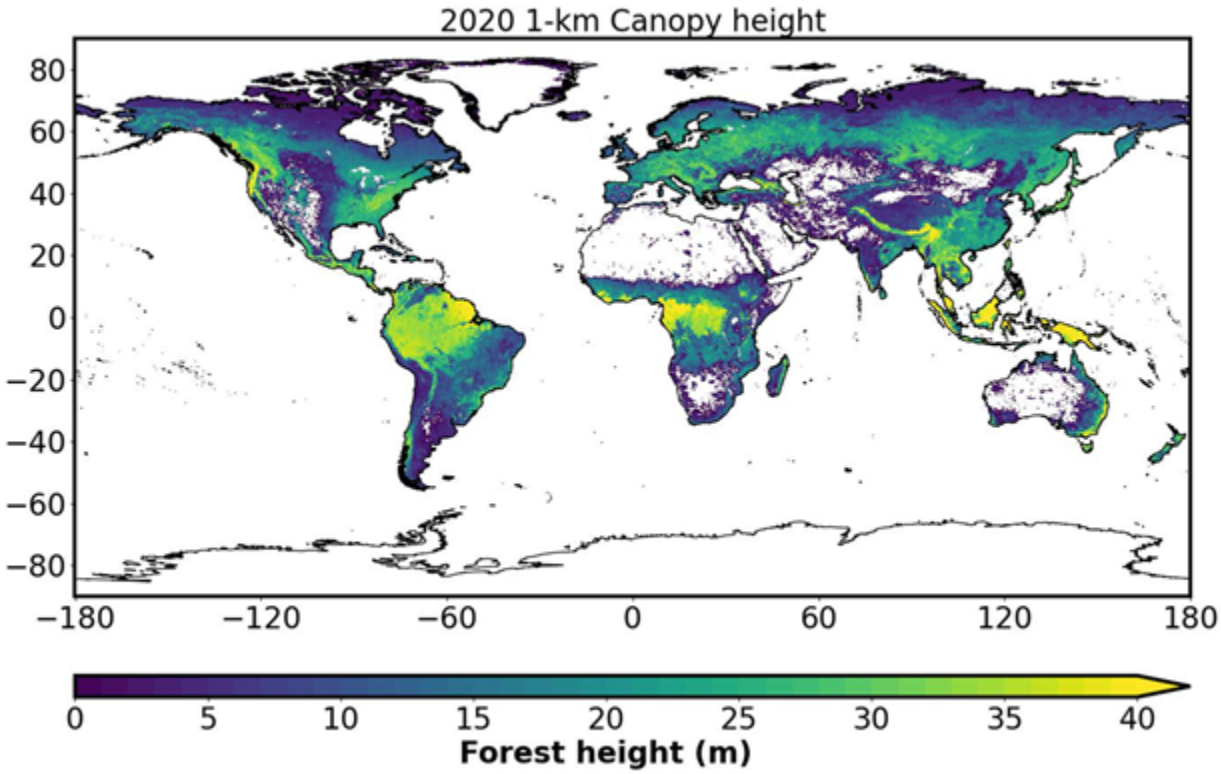

Figure 4. Annual global vegetative canopy height of 2020.

2.6 Green Vegetation Fraction

The annual global GVF is based on the VIIRS green vegetation fraction product (Jiang et al., 2021; https://www.star.nesdis.noaa.gov/jpss/gvf.php), which provides the fractional green vegetation cover on a weekly basis. Similar to CLU and LAI, many missing data are found over high latitude regions in the North Hemisphere due to the lack of sunlight during winter.

The gaps are filled by estimated fractional vegetation cover (FVC) based on vegetation indices (VI) such as CLU and LAI (Gao et al., 2020; Kuusk et al., 2018; Nilson, 1971). LAI is technically defined as the amount of green leaf area per unit ground surface area, and is generally used as an indicator of greenness in canopy models (Fang et al., 2019). Hence, a VI-based estimated FVC can provide a good proxy of the GVF for certain STs (e.g., broadleaf forests) during some seasons (e.g., summer), but this relationship can have more uncertainty when the non-green components of the vegetation are relatively large for other STs (e.g., needleleaf forests) and seasons (e.g., winter). To fill the gaps, a global VI-based estimation of FVC map is first generated following the formulations from Gao et al. (2020) and Kuusk et al. (2018):

$$FVC = 1 - P(0) \quad \text{Eq. (3)}$$

$P(0)$ is the canopy gap fraction at a zenith angle equal to zero degree, and can be calculated as:

$$P(0) = e^{-CLU \times G \times LAI} \quad \text{Eq. (4)}$$

where G is the Ross–Nilson geometry function (the G-function; Ross, 1981) that describes the effect of a unit of leaf area on radiation attenuation. $G = 0.5$ is used here by assuming a zenith angle of zero degree based on Figure 1 in Stenberg (2006). The global gap-filled CLU and LAI maps discussed in Section 2.3 and 2.4, respectively, are used here.

Moreover, to reduce the inconsistency between VIIRS GVF and calculated VIIRS VI-based FVC, the latter is corrected based on the linear regression between two datasets:

$$GVF = a(ST) \times FVC + b(ST) \quad \text{Eq. (5)}$$

The ST dependent regression coefficients $a(ST)$ and $b(ST)$ are calculated based on the data at lower latitude areas (e.g., < 60N) where VIIRS GVF is available. The ST classification is based on the VIIRS-based ST dataset described in Section 2.2. Similar to the gap-filling process for CLU and LAI, the missing gaps in the North Hemisphere from late fall to early spring (i.e. January,

February, March, April, May, October, November and December) are filled for STs equal to 1 – 12 only by assuming a relatively minor contribution from Mos (ST = 14).

Figures 5 and 6 show the scatter density plots and linear regressions of monthly VIIRS GVF and VI-based FVC of January for all ST and each ST category, respectively, while Figure 7 demonstrates the monthly global GVF datasets of 2020 before and after gap filling. The coefficient of determination (R^2) for all ST is 0.79. The deciduous broadleaf forest (DBF, ST = 4) shows the highest R^2 of 0.92, followed by the grassland (GL, ST = 10) with a R^2 value of 0.88. The deciduous needleleaf forest (DNF, ST = 3) shows the lowest R^2 of 0.02, probably associated with the contributions of non-green components of vegetation (e.g., stems) during wintertime.

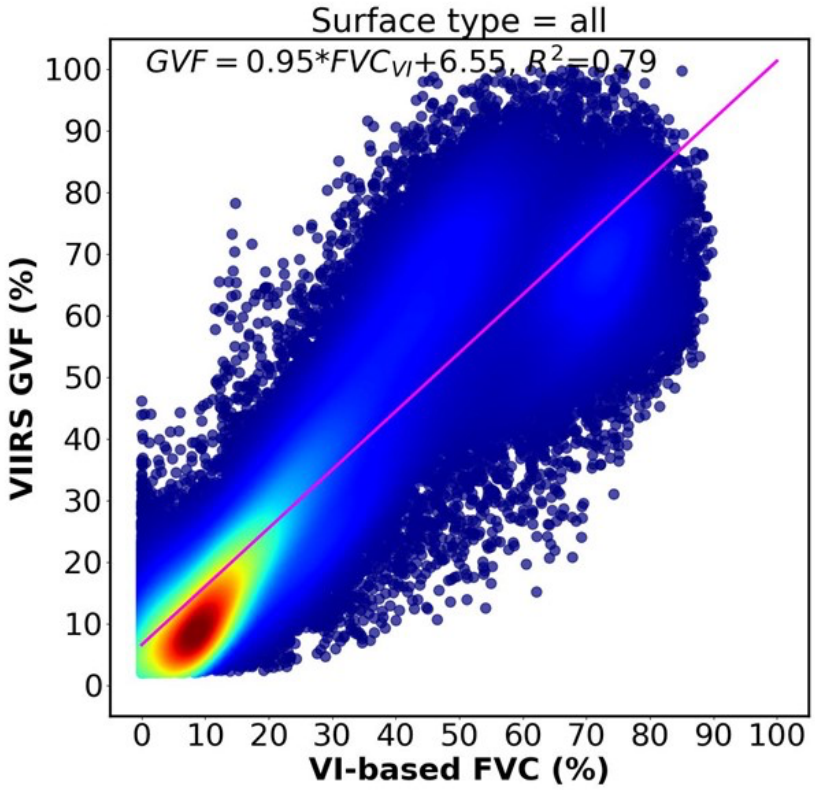

Figure 5. Scatter density plot of VI-based FVC estimation and VIIRS GVF for all ST. The shedding color shows the data density, and the magenta line represents the linear regression of two datasets.

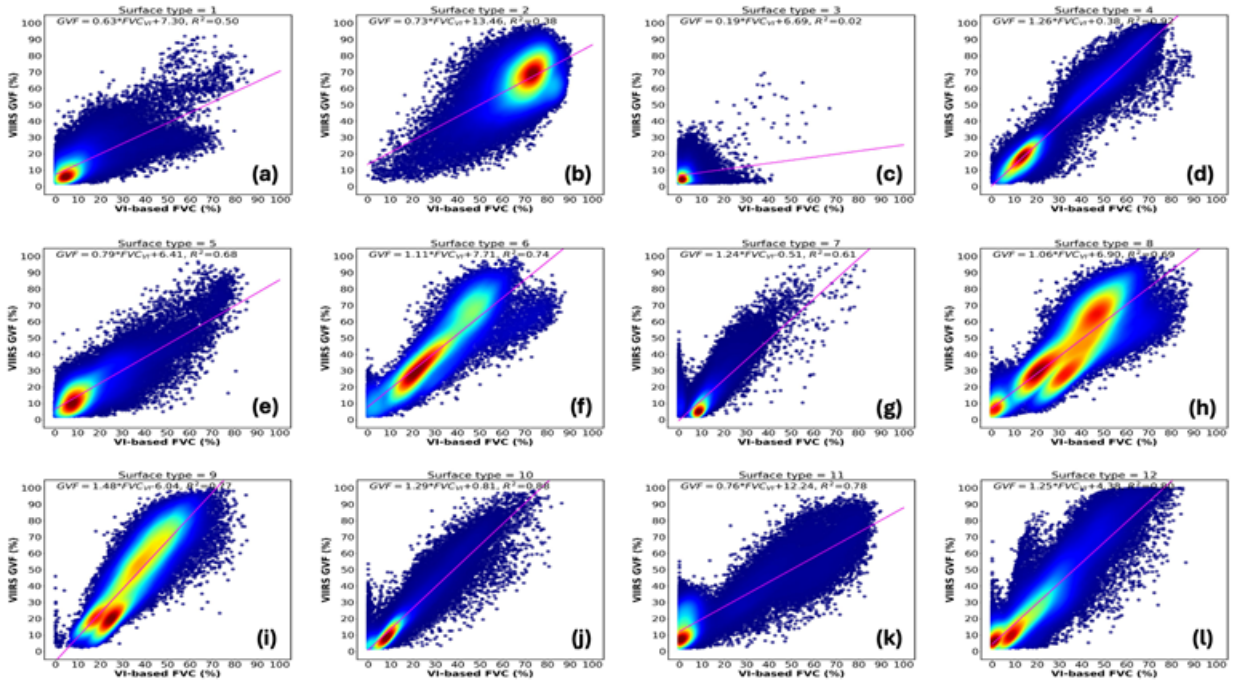

Figure 6. Scatter density plots of VI-based FVC estimation and VIIRS GVF for (a) ENF, (b) EBF, (c) DNF, (d) EDF, (e) MxF, (f) CSh, (g) Osh, (h) WSV, (i) SV, (j) GL, (k) Wet, (l) Crop. The shedding color shows the data density, and the magenta line represents the linear regression of two datasets.

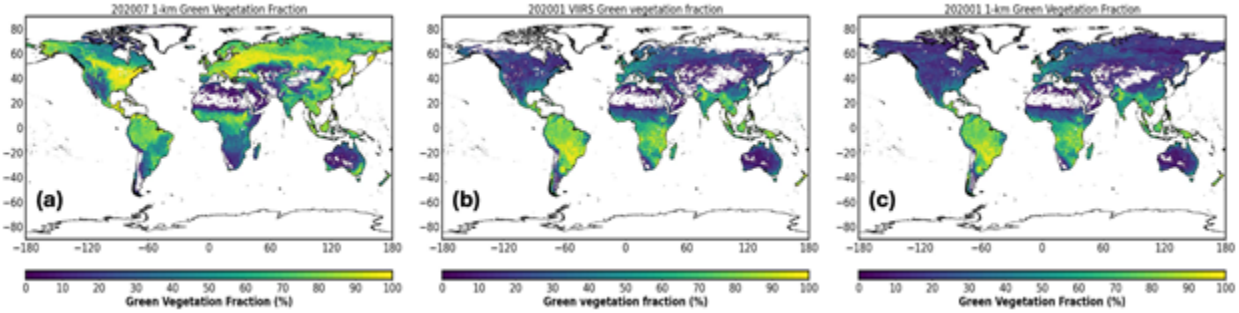

Figure 7. Monthly averages of global green vegetation fraction in (a) July, (b) January before gap filling and re-gridding, and (c) January 2020. The datasets shown in (a) and (c) have been re-gridded to 0.01° coordinate.

3. Discussion

The major uncertainty of the GLSDs comes from the discrepancies of ST classification among different satellite products. Although most MODIS and VIIRS products follow the IGBP classification with standard 17 ST categories and the same retrieval algorithms, the differences in the spectrum channels, instrument calibration processes, and spatial resolution could lead to inconsistencies in the ST classifications. Such inconsistencies in the products of two satellites may affect the global extension and gap filling process, which are ST based, for the CLU, LAI and GVF datasets. Replacing MODIS-based products (e.g., CLU and the climatological values used for gap filling) with VIIRS-based products would improve the reliability of the GLSDs and make them a more unified dataset by removing instrumental uncertainties from and using consistent satellite measurements in the data processing for a single GLSD.

In addition, the ST dataset provides the dominant ST of each grid cell and may miss the fractional contributions from other surrounding vegetation types This may have significant implications for averaging the CH dataset across the area of each grid cell, especially in regions of high spatial heterogeneity. For example, most of the regions over the southeast US are identified as forests (i.e. ST = 1 – 5) and about 30% is mixed forest (ST = 5). However, around 5% of these forest regions show CH values less than 10 m, probably affected by the relatively short, surrounding grasslands and croplands after averaging.

Another uncertainty lies in the simple gap filling approach based on climatological monthly averages, which inherently miss higher temporal variability. Furthermore, the summer standard values used for global extension are also satellite based and could be affected by similar instrument limitations (e.g., the amount and quality of samples).

Finally, the linear regression used to correct VI-based FVC to VIIRS GVF is fairly oversimplified, leading to potential uncertainties. To deal with this issue, machine learning algorithms, such as random forest (Breiman, 2001; Segal, 2004) and neural network (Uhrig, 1995), can be used to improve the gap-filling process. Specifically, a more explicit machine learning based regression model of VI-based FVC and VIIRS GVF can be generated and applied to GVF correction for advanced accuracy.

4. Summary

A global land surface dataset (GLSD) for five major land surface properties including surface type (ST), vegetation clumping index (CLU), vegetation leaf area index (LAI), vegetative canopy height (CH), and green vegetation fractions (GVF) is developed at NOAA ARL. The GLSDs share a universal global latitude/longitude gridded coordinate with a spatial resolution of 0.01° (~ 1 km). The GLSDs are based on a suite of operational satellite products from MODIS, VIIRS and GEDI. Some satellite products (e.g., CLU, LAI, and GVF) are globally extended and gap-filled for mid to high latitude regions in the North Hemisphere. After gap filling, all GLSDs are re-gridded to 0.01° x 0.01° coordinate with a coverage of -90° - 90° in latitude and 0° - 360° in longitude. The major uncertainties lie in the discrepancies of ST classification in different satellite products, and the simple linear regressions used for gap filling and GVF estimation correction.

Currently, GLSDs are available for one year period representative of 2020. The gridded NetCDF data format is beneficial for applications to Earth System Models (ESMs) such as land, weather, and air composition models. The LAI and CH datasets are being applied and tested in the next-generation UFS Air Quality Model (AQM) at NOAA, and while not shown here, can lead to improvement in the spatial representation of explicit vegetative canopy effects on surface air quality (e.g., ozone concentrations) and meteorological (e.g., temperature and wind speed) predictions over the canopy regions in the US. In the future, an additional fractional ST dataset, which introduces the fraction of different STs within each grid cell, and VIIRS-based datasets will be developed and be used to improve the gap filling processes in a more unified fashion. Machine learning algorithms such as neural network and random forest, which are popular machine learning tools for simulating the complex correlations between datasets, can be used to improve the gap filling process and the correction process of VI-based FVC estimation. GLSDs for multiple years will also be available in the future.

5. Data Availability

The GLSDs is now available for one year period representative of 2020 at the National Centers for Environmental Information (NCEI) and can be downloaded from https://doi.org/10.25921/qzm2-zg29.

6. References

Belward, A. S., and Loveland, T. S., 1997. The International Geosphere Biosphere Programme Data and Information System global land cover data set (DISCover). Acta Astronautica, 41, 4-10, doi: 10.1016/S0094-5765(98)00050-2.

Breiman, L., 2001. Random forests. Mach. Learn., 45, 5–32, doi: 10.1023/A:1010933404324.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing, 1: 100002, doi: 10.1016/j.srs.2020.100002.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., et al. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. Journal of Geophysical Research, 108, 8851, doi: 10.1029/2002JD003296, D22.

Fang, H., Baret, F., Plummer, S., Schaepman-Strub, G., 2019. An overview of global leaf area index (LAI): Methods, products, validation, and applications. Reviews of Geophysics. 57, 739–799, doi: 10.1029/2018RG000608

Fang, H., Li, S., Zhang, Y., Wei, S., Wang, Y., 2021. New insights of global vegetation structural properties through an analysis of canopy clumping index, fractional vegetation cover, and leaf area index. Science of Remote Sensing 4, 100027, doi: 10.1016/j.srs.2021.100027.

Fisher, R. A., and Koven, C. D., 2020. Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. Journal of Advances in Modeling Earth Systems, 12, e2018MS001453, doi: 10.1029/2018MS001453.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114, 168-182, doi: 10.1016/j.rse.2009.08.016.

Friedl, M. and Sulla-Menashe, D., 2015. MCD12C1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2023-03-29 from https://doi.org/10.5067/MODIS/MCD12C1.006.

Gao L, Wang X, Johnson BA, Tian Q, Wang Y, Verrelst J, Mu X, Gu X. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A

review. ISPRS J Photogramm Remote Sens. 2020 Jan;159:364-377. doi: 10.1016/j.isprsjprs.2019.11.018. PMID: 36082112; PMCID: PMC7613353.

Hunter, J. D., 2007. Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95.

Lang, N., Jetz, W., Schindler, K., Wegner, J. D., 2023. A high-resolution canopy height model of the Earth. Nat Ecol Evol 7, 1778–1789, doi: 10.1038/s41559-023-02206-6.

Li, L., Bisht, G., Hao, D., Leung, L. R., 2024. Global 1 km land surface parameters for kilometer-scale Earth system modeling, Earth Syst. Sci. Data, 16, 2007–2032, doi: 10.5194/essd-16-2007-2024.

Jiang, Z., Yu, Y., Csiszar, I., Chen, M., Zhao, F., He, Y., 2021. GVF Algorithm Theoretical Basis Document Version 4.0. [Available online: https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_NVPS_GVF_v4.0.pdf]

Johannsen, F., Ermida, S., Martins, J.P.A., Trigo, I.F., Nogueira, M., Dutra, E., 2019. Cold Bias of ERA5 Summertime Daily Maximum Land Surface Temperature over Iberian Peninsula. Remote Sens., 11, 2570, doi: 10.3390/rs11212570.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P., Morisette, J. T., 2002. An overview of MODIS Land data processing and product status. Remote Sens. Environ., 83, 1-2, pp 3-15, doi: 10.1016/S0034-4257(02)00084-6.

Justice, C. O., Román, M. O., Csiszar, I., Vermote, E. F., Wolfe, R. E., Hook, S. J., Friedl, M., Wang, Z., Schaaf, C. B., Miura, T., Tschudi, M., Riggs, G., Hall, D. K., Lyapustin, A. I., Devadiga, S., Davidson, C., Masuoka, E. J., 2013. Land and cryosphere products from Suomi NPP VIIRS: Overview and status. J. Geophys. Res. Atmos., 118, 9753–9765, doi: 10.1002/jgrd.50771.

Kuusk, A, Pisek J, Lang, M, Märdla, S., 2018. Estimation of Gap Fraction and Foliage Clumping in Forest Canopies. Remote Sensing, 10(7):1153, doi: 10.3390/rs10071153.

Moon, M., Zhang, X., Henebry, G. M., Liu, L., Gray, J. M., Melaas, E. K., Friedl, M. A., 2019. Long-term continuity in land surface phenology measurements: A comparative assessment of the MODIS land cover dynamics and VIIRS land surface phenology products. Remote Sensing of Environment, 226, 1, p74-92, doi: 10.1016/j.rse.2019.03.034.

Myneni, R. and Knyazikhin, Y., 2018. VIIRS/NPP Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V001 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2023-03-29 from https://doi.org/10.5067/VIIRS/VNP15A2H.001.

Nilson, T., 1971. A theoretical analysis of the frequency of gaps in plant stands. Agricultural Meteorology, 8, 25-38, doi: 10.1016/0002-1571(71)90092-6.

Nogueira, M., Albergel, C., Boussetta, S., Johannsen, F., Trigo, I. F., Ermida, S. L., Martins, J. P. A., Dutra, E., 2020. Role of vegetation in representing land surface temperature in the CHTESSEL (CY45R1) and SURFEX-ISBA (v8.1) land surface models: a case study over Iberia, Geosci. Model Dev., 13, 3975–3993, doi: 10.5194/gmd-13-3975-2020.

Ross, J., 1981. The radiation regime and architecture of plant stands. Dr. W. Junk Publishers, The Hague, 391 p, doi: 10.1007/978-94-009-8647-3.

Ruiz-Vásquez, M., O, S., Arduini, G., Boussetta, S., Brenning, A., Bastos, A., Koirala, S., Balsamo, G., Reichstein, M., Orth, R., 2023. Impact of updating vegetation information on land surface model performance. Journal of Geophysical Research: Atmospheres, 128, e2023JD039076, doi: 10.1029/2023JD039076.

Segal, M. R., 2004. Machine Learning Benchmarks and Random Forest Regression. UCSF: Center for Bioinformatics and Molecular Biostatistics. Retrieved from https://escholarship.org/uc/item/35x3v9t4.

Stenberg, P., 2006. A note on the G-function for needle leaf canopies. Agricultural and Forest Meteorology, 136(1-2), 76–79, doi:10.1016/j.agrformet.2006.01.009 .

Townshend, J., Hansen, M., Carroll, M., DiMiceli, C., Sohlberg, R., Huang, C. User Guide for the MODIS Vegetation Continuous Fields product Collection 61, version 1. [Available online: https://lpdaac.usgs.gov/documents/1494/MOD44B_User_Guide_V61.pdf]

Uhrig, R. E., 1995. Introduction to artificial neural networks. Proceedings of IECON '95 - 21st Annual Conference on IEEE Industrial Electronics, Orlando, FL, USA, 1995, pp. 33-37 vol.1, doi: 10.1109/IECON.1995.483329.

van der Graaf, S. C., Kranenburg, R., Segers, A. J., Schaap, M., and Erisman, J. W., 2020. Satellite-derived leaf area index and roughness length information for surface–atmosphere exchange modeling: a case study for reactive nitrogen deposition in north-western Europe using LOTOS-EUROS v2.0 , Geosci. Model Dev., 13, 2451–2474, doi: 10.5194/gmd-13-2451-2020.

Wang, Y., Köhler, P., Braghiere, R. K., Longo, M. , Doughty, R., Bloom, A. A., Frankenberg, C., 2022. GriddingMachine, a database and software for Earth system modeling at global and regional scales. Scientific Data 9, 258, doi: 10.1038/s41597-022-01346-x.

Wei, S., Fang, H., Schaaf, C. B., He, L., Chen, J. M., 2019. Global 500 m clumping index product derived from MODIS BRDF data (2001–2017). Remote Sens Environ, 232, 111296, doi: 10.1016/j.rse.2019.111296.

Zhang, R., Huang, C., Zhan, X., Jin, H., Song, X.-P., 2018. Development of S-NPP VIIRS global surface type classification map using support vector machines. International Journal of Digital Earth, 11:2, 212-232, doi: 10.1080