Global High-Resolution Estimates of SIF from Fused SCIAMACHY and GOME-2, V2

Get Data

Documentation Revision Date: 2024-01-25

Dataset Version: 2

Summary

This dataset provides global solar-induced chlorophyll fluorescence (SIF) estimates at a 0.05-degree resolution (approximately 5 km at the equator) for each month from January 2003 through December 2017. SIF data (740 nm) was retrieved from the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) and Global Ozone Monitoring Experiment 2 (GOME-2) instruments onboard the MetOp-A satellite. The data were downscaled to 0.05 degrees using the Random Forest algorithm and predictors from Moderate Resolution Imaging Spectroradiometer (MODIS) and Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) reanalysis, and then harmonized with the cumulative distribution function (CDF) matching technique. The uncertainty of the harmonized SIF estimates was also quantified and provided. Validation of the harmonized product showed that it retained high spatial and temporal consistency with the original SCIAMACHY and GOME-2 SIF retrievals and had good correlations with independent airborne and ground-based SIF measurements. The dataset can inform on the synergy between satellite SIF and photosynthesis and research on drought, yield estimation, and land degradation evaluation. The data are provided in netCDF format.

This dataset is related to the dataset Wen et al. (2021) but uses different SIF retrievals and received minor changes to processing procedures. The temporal coverage of this dataset also differs from Wen et al. (2021).

There are 180 data files in netCDF (*.nc) format included in this dataset.

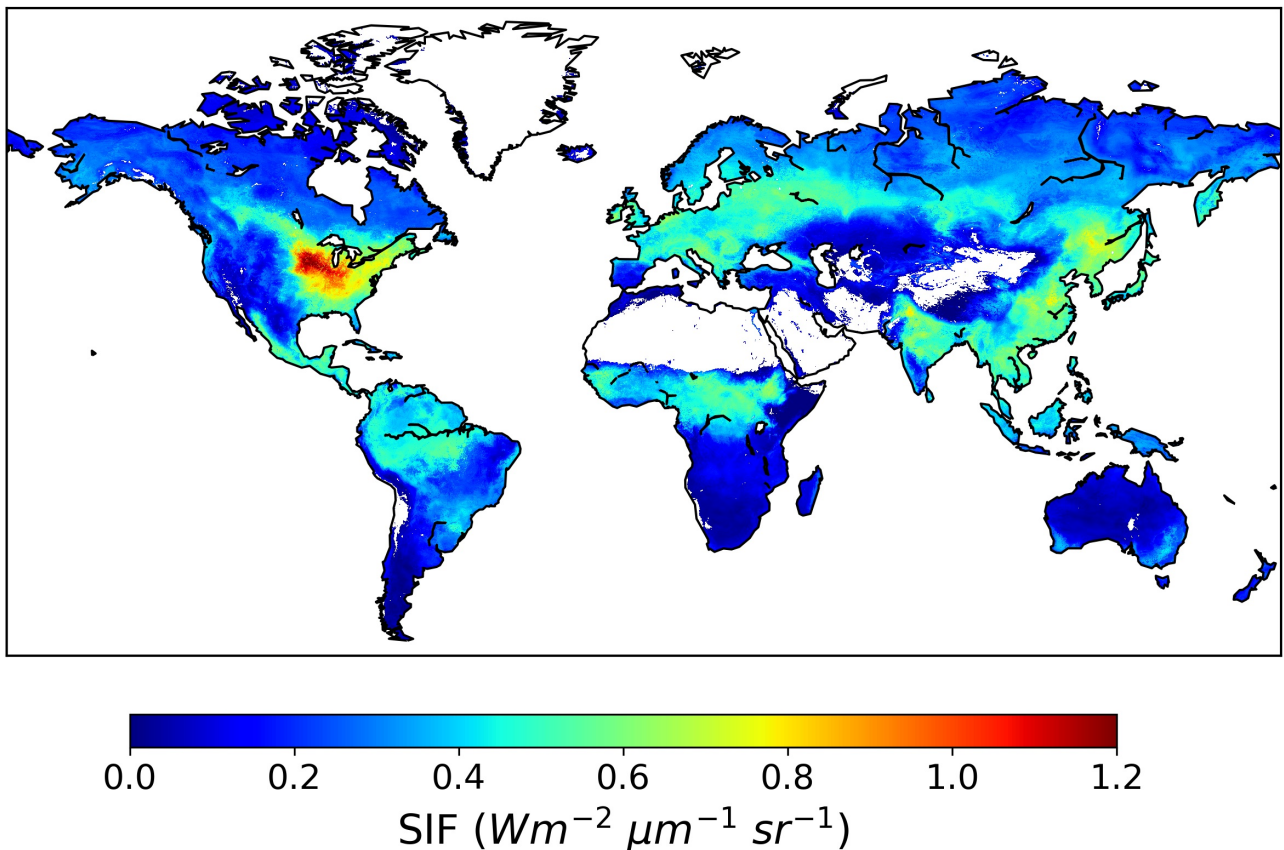## Daily-averaged SIF

SIF ($Wm^{-2} \mu m^{-1} sr^{-1}$)

Figure 1. Daily-corrected SIF at 740 nm for August 2015. The highest SIF values during this time period were observed in agricultural sectors in the Northern Hemisphere.

Citation

Table of Contents

1. Dataset Overview

This dataset provides global solar-induced chlorophyll fluorescence (SIF) estimates at a 0.05-degree resolution (approximately 5 km at the equator) for each month from January 2003 through December 2017. SIF data (740 nm) was retrieved from the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) and Global Ozone Monitoring Experiment 2 (GOME-2) instruments onboard the MetOp-A satellite. The data were downscaled to 0.05 degrees using the Random Forest algorithm and predictors from Moderate Resolution Imaging Spectroradiometer (MODIS) and Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) reanalysis, and then harmonized with the cumulative distribution function (CDF) matching technique. The uncertainty of the harmonized SIF estimates was also quantified and provided. Validation of the harmonized product showed that it retained high spatial and temporal consistency with the original SCIAMACHY and GOME-2 SIF retrievals and had good correlations with independent airborne and ground-based SIF measurements. The dataset can inform on the synergy between satellite SIF and photosynthesis and research on drought, yield estimation, and land degradation evaluation.

This dataset is related to the dataset Wen et al. (2021) but uses different SIF retrievals and received minor changes to processing procedures. The temporal coverage of this dataset also differs from Wen et al. (2021).

Project: Solar Induced Fluorescence Earth Science Data Record

This project is developing a global, observation-based Earth System Data Record (ESDR) for quantifying global vegetation solar induced fluorescence (SIF) and photosynthesis gross primary productivity (GPP) from 1996-2020. It was funded under the 2017 Making Earth System Data Records for Use in Research Environments (MEaSUREs) call (17-MEASURES-0032).

Related Publication

Wen, J., P. Köhler, G. Duveiller, N. C. Parazoo, T. S. Magney, G. Hooker, L. Yu, C. Y. Chang, and Y. Sun. 2020. A framework for harmonizing multiple satellite instruments to generate a long-term global high spatial-resolution solar-induced chlorophyll fluorescence (SIF). Remote Sensing of Environment 239:111644. https://doi.org/10.1016/j.rse.2020.111644

Related Datasets

Wen, J., P. Koehler, G. Duveiller, N.C. Parazoo, T. Magney, G. Hooker, L. Yu, C.Y. Chang, and Y. Sun. 2021. Global High-Resolution Estimates of SIF from Fused SCIAMACHY and GOME-2, 2002-2018. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1864

- This is an earlier version of the fused SCIAMACHY and GOME-2 SIF dataset. It used SIF retrievals from the German Research Center for Geosciences (GFZ) (processed by Dr. Philipp Köhler), which could present spurious trends owing to instrument degradation.

Joiner, J., Y. Yoshida, P. Koehler, C. Frankenberg, and N.C. Parazoo. 2021. L2 Solar-Induced Fluorescence (SIF) from SCIAMACHY, 2003-2012. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1871

Joiner, J., Y. Yoshida, P. Koehler, C. Frankenberg, and N.C. Parazoo. 2023. L2 Daily Solar-Induced Fluorescence (SIF) from MetOp-A GOME-2, 2007-2018. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2083

Acknowledgments

2. Data Characteristics

Spatial Coverage: Global land

Spatial Resolution: 0.05 degree

Temporal Coverage: 2003-01-01 to 2017-12-31

Temporal Resolution: Monthly

Study Area: Latitude and longitude are given in decimal degrees.

Site	Northernmost Latitude	Southernmost Latitude	Easternmost Longitude	Westernmost Longitude
Global Land	90	-90	180	-180

Data File Information

There are 180 data files in netCDF (*.nc) format included in this dataset. The files are named sif005_ *YYYYMM*.nc (e.g., sif005_200301.nc), where *YYYYMM* represents the year (YYYY) and month (MM) of the data.

Table 1. File names and descriptions.

File Names	Description
sif005_*YYYYMM*.nc	Each file provides daily corrected SIF at 740nm and its associated uncertainty.

Data File Details

Missing values are represented by -999. Each file contains 3600 rows and 7200 columns indexed by latitude and longitude dimensions, respectively. The Coordinate Reference System is "WGS84 - World Geodetic System 1984" (EPSG:4326). The spatial resolution is 0.05 degree.

Table 2. Variable names and descriptions.

Variable Name	Units	Description
crs	-	Coordinate Reference System information
EVI_Quality	-	MODIS Enhanced Vegetation Index (EVI) quality; detailed flag information can be found in MOD13C2.006 User Manual Table 5, https://lpdaac.usgs.gov/documents/103/MOD13_User_Guide_V6.pdf
lat	degrees north	Latitude at center of cell
lon	degrees east	Longitude at center of cell
SIF_740_daily_corr	mW m-2 nm-1 sr-1	Daily-corrected SIF at 740nm at a 0.05-degree resolution
SIF_740_daily_corr_SD	mW m-2 nm-1 sr-1	Uncertainty of daily-corrected SIF at 740nm

3. Application and Derivation

This long SIF record can greatly enhance the synergy between satellite SIF and photosynthesis. This product has also been used in applications such as drought monitoring, yield estimation, land degradation evaluation, etc.

4. Quality Assessment

The infinitesimal jackknife method was used to estimate standard errors of random forest predictions. The estimation was accomplished by evaluating the covariance between certain predictions from each individual tree and the number of times each training sample appeared in a bootstrap sample for that tree (Wager et al., 2014; Ghosal and Hooker, 2020). Smaller covariance indicates that the prediction is less affected by whether a certain training sample is in the bootstrapped sample or not; therefore, it had less variance and thus less uncertainty. The R libraries "randomForest" and "randomForestCI" were used in this study.

With MODIS EVI as one important input of the model, a quality flag layer was added and it is recommended that users check this flag before subsequent analysis. Detailed information on the quality flag layer can be found in the MOD13C2.006 User Manual Table 5 (https://lpdaac.usgs.gov/documents/103/MOD13_User_Guide_V6.pdf). Users should extract the value of the desired bit(s) to match the information in the table (see https://modis-images.gsfc.nasa.gov/tools_bit_interpretation.html).

5. Data Acquisition, Materials, and Methods

This dataset was produced based on the framework described in Wen et al. (2020) by fusing SIF retrievals from SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) and Global Ozone Monitoring Experiment 2 (GOME-2) onboard MetOp-A developed at German Research Center for Geosciences. First, the original Level 3 SIF retrievals of SCIAMACHY and GOME-2 (Köhler et al., 2015) were downscaled from their native resolutions to 0.05° using machine learning (ML) algorithms imposed by regionalization constraints to account for the varying relationships between predictors and SIF in space and time. Then the cumulative distribution function matching technique was applied to correct the offset between SIF retrieved from GOME2_005 and SIF retrieved from SCIAMACHY as inherited from the original instrumental discrepancies to generate a harmonized SIF time series. See Wen et al. (2020) for detailed methods.

This dataset is related to the dataset Wen et al. (2021). The data provided by Wen et al. (2021) used SIF retrievals from the German Research Center for Geosciences (GFZ) (processed by Dr. Philipp Köhler), which could present spurious trends owing to instrument degradation. In this version, the same procedures were used but with SIF retrievals from NASA (processed by Dr. Joanna Joiner), which largely corrected the spurious trends. Some minor changes were also made to this version: 1) before machine learning, a spatial smoothing procedure was applied to the SIF retrievals (i.e., by calculating the moving average over nearby 3 x 3 pixels) to reduce noise; 2) in the procedure of CDF matching for values that fall into <5 or >95 percentiles, coefficients from linear regression of all data pairs were used to avoid unreasonable scaling results instead of coefficients from linear regression of data pairs in the specific percentile ranges.

6. Data Access

These data are available through the Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC).

Global High-Resolution Estimates of SIF from Fused SCIAMACHY and GOME-2, V2

Contact for Data Center Access Information:

- E-mail: uso@daac.ornl.gov
- Telephone: +1 (865) 241-3952

7. References

Ghosal, I., and G. Hooker. 2020. Boosting Random Forests to Reduce Bias; One-Step Boosted Forest and Its Variance Estimate. Journal of Computational and Graphical Statistics 30:493–502. https://doi.org/10.1080/10618600.2020.1820345

Joiner, J., Y. Yoshida, P. Koehler, C. Frankenberg, and N.C. Parazoo. 2021. L2 Solar-Induced Fluorescence (SIF) from SCIAMACHY, 2003-2012. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1871

Joiner, J., Y. Yoshida, P. Koehler, C. Frankenberg, and N.C. Parazoo. 2023. L2 Daily Solar-Induced Fluorescence (SIF) from MetOp-A GOME-2, 2007-2018. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2083

Köhler, P., L. Guanter, and J. Joiner. 2015. A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data. Atmospheric Measurement Techniques 8:2589–2608. https://doi.org/10.5194/amt-8-2589-2015

Wager, S., T. Hastie, and B. Efron. 2014. Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. The Journal of Machine Learning Research 15:1625-1651. http://jmlr.org/papers/v15/wager14a.html

Wen, J., P. Köhler, G. Duveiller, N. C. Parazoo, T. S. Magney, G. Hooker, L. Yu, C. Y. Chang, and Y. Sun. 2020. A framework for harmonizing multiple satellite instruments to generate a long-term global high spatial-resolution solar-induced chlorophyll fluorescence (SIF). Remote Sensing of Environment 239:111644. https://doi.org/10.1016/j.rse.2020.111644

Wen, J., P. Koehler, G. Duveiller, N.C. Parazoo, T. Magney, G. Hooker, L. Yu, C.Y. Chang, and Y. Sun. 2021. Global High-Resolution Estimates of SIF from Fused SCIAMACHY and GOME-2, 2002-2018. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1864

8. Dataset Revisions

Version	Release Date	Revision Notes
2.0	2024-01-25	Updated version of the fused SCIAMACHY and GOME-2 SIF dataset. In this version, the same procedures were used but with SIF retrievals from NASA (processed by Dr. Joanna Joiner), which largely corrected the spurious trends. Some minor changes were also made to this version: 1) before machine learning, a spatial smoothing procedure was applied to the SIF retrievals (i.e., by calculating the moving average over nearby 3 x 3 pixels) to reduce noise; 2) in the procedure of CDF matching for values that fall into <5 or >95 percentiles, coefficients from linear regression of all data pairs were used to avoid unreasonable scaling results instead of coefficients from linear regression of data pairs in the specific percentile ranges.
1.0	2021-06-16	Initial release of the fused SCIAMACHY and GOME-2 SIF dataset. It used SIF retrievals from the German Research Center for Geosciences (GFZ) (processed by Dr. Philipp Köhler), which could present spurious trends owing to instrument degradation.

OAK RIDGE
National Laboratory

Privacy Policy | Feedback | Help

🏠 **Home**

About Us
Mission
Data Use and Citation Policy
User Working Group
Partners

Get Data
Science Themes
NASA Projects
All Datasets

Submit Data
Submit Data Form
Data Scope and Acceptance
Data Authorship Policy
Data Publication Timeline
Detailed Submission Guidelines

Tools
TESViS
THREDDS
SDAT
Daymet
Airborne Data Visualizer
Soil Moisture Visualizer
Land - Water Checker

Resources
Learning
Data Management
News
Earthdata Forum ⤴

💬**Contact Us**