Figure S1: Locations of each of the four land covers with the highest number of detected flash drought events from 2000–2020, as studied in Figure 1. Only pixels with at least one detected flash drought event are shown in colour. While this figure is based on the land cover map from 2020 as an illustration, interannual variations in land cover are accounted for in the flash drought composite analyses.
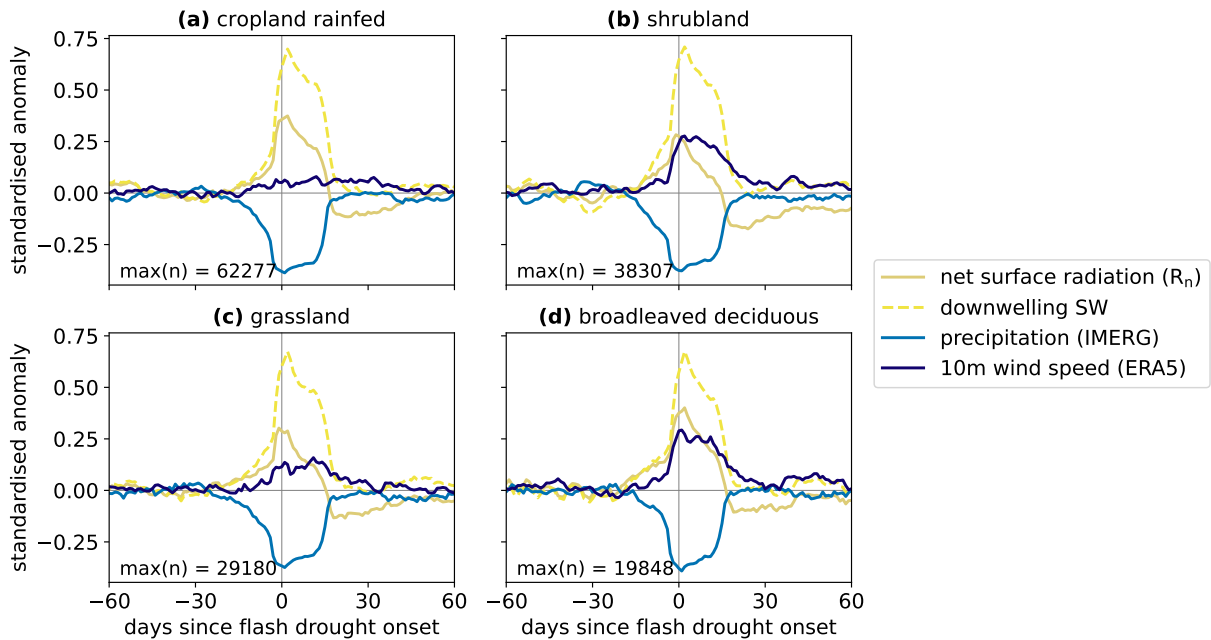
Figure S2: Additional variables to accompany Figure 2. Each variable is composited over all flash droughts during the period 2000–2020 with onset dates during the peak growing season, for the four land cover classes with the highest numbers of events. The R_n composites shown here are identical to those in Figure 2, reproduced to provide context for the downwelling shortwave (SW) radiation anomalies.

Table S1: Summary of datasets composited around flash drought event onsets.

Variable	Data source	Years used	Temporal resolution
Surface soil moisture	ESA CCI Soil Moisture combined active/passive v08.1 (Dorigo et al., 2017; Gruber et al., 2019; Preimesberger et al., 2021)	2000–2020	Running pentad mean of daily data
Land surface temperature	ESA CCI MODIS (Moderate resolution Infra-red Spectroradiometer) on Aqua, level 3 collated (L3C) global product v4.00 (Ghent et al., 2019)	Jul 2002–Dec 2020	Daily (approx 1330 local solar time), clear-sky only
	ESA CCI All-weather MicroWave Land Surface Temperature (MW-LST) global data record v2.33 (Jimenez and Prigent, 2023)	2000–2020	Daily at 1800 local solar time
2m air temperature	ERA5 (Hersbach et al., 2020)	2000–2020	Daily maximum between 1200–1800 local solar time (from hourly data)
Sensible heat flux (ΔT)	Computed as LST-T2m from ESA CCI LST MODIS Aqua v4.00 and ERA5 T2m	Jul 2002–Dec 2020	Daily (approx 1330 local solar time), clear-sky only
	Computed as LST-T2m from ESA CCI MW-LST v2.33 and ERA5 T2m	2000–2020	Daily at 1800 local solar time, all-sky
Precipitation	GPM IMERG v06 (Huffman et al., 2019)	Jun 2000–Dec 2020	Daily
Latent heat flux	GLEAM 4.2a (Miralles et al., 2024)	2000–2018	Daily
Root-zone soil moisture	GLEAM v4.2a (Miralles et al., 2024)	2000–2018	Daily
Standardised Evaporative Stress Ratio	GLEAM v4.2a (Miralles et al., 2024)	2000–2020	Daily
	ERA5 (Hersbach et al., 2020)	2000–2020	Daily
Vapour pressure deficit	ERA5 (Hersbach et al., 2020)	2000–2020	Daily mean between 1200–1800 local solar time (from hourly data)
10m wind speed	ERA5 (Hersbach et al., 2020)	2000–2020	Daily mean between 1200–1800 local solar time (from hourly data)
Net radiation at surface	CERES SYN1deg-Day Edition 4A (Rutan et al., 2015)	Mar 2000–Dec 2020	Daily
Surface downwelling shortwave radiation	CERES SYN1deg-Day Edition 4A (Rutan et al., 2015)	Mar 2000–Dec 2020	Daily
Vegetation Optical Depth	Vegetation Optical Depth Climate Archive (VODCA) v2, CXKu-band (Zotta et al., 2024)	2000–2020	Daily (night-time only)
Solar Induced Fluorescence	Spatially downscaled GOME-2 (Duveiller et al., 2020), using retrieval method of (Köhler et al., 2015)	2007–2018	8-daily
	Spatially downscaled GOME-2 (Duveiller et al., 2020), using retrieval method of (Joiner et al., 2013)	2007–2018	8-daily
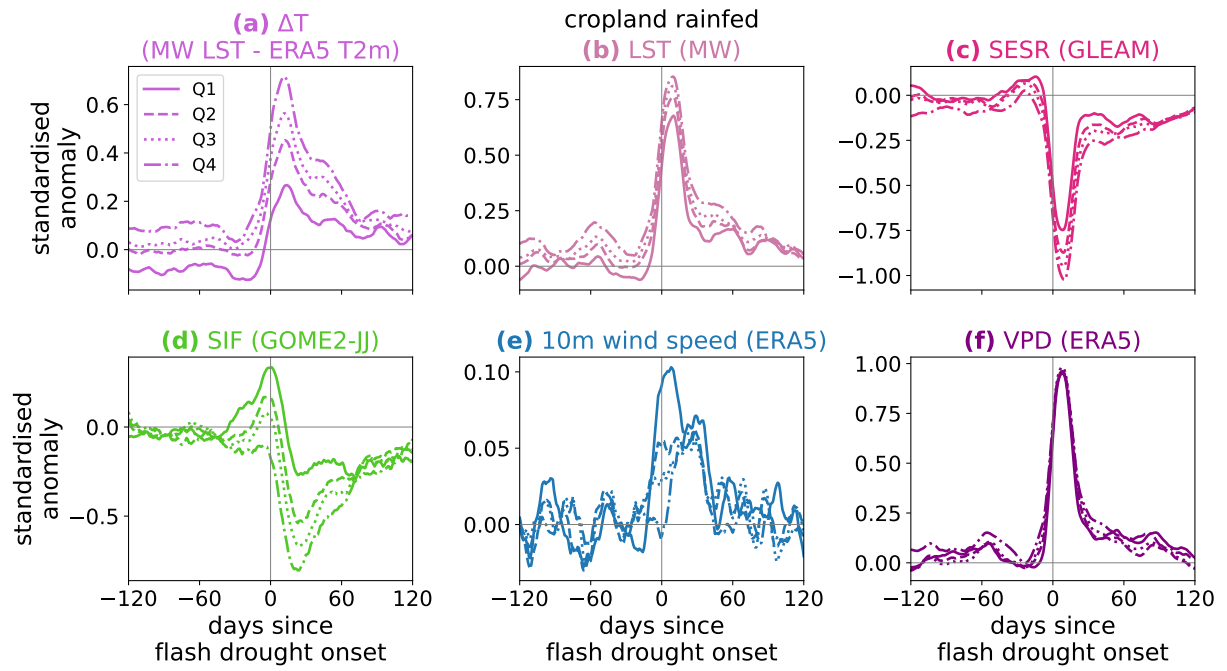
Figure S3: Additional variables to accompany Figure 3. Each panel shows a composite over all events in rainfed cropland during the period 2000–2020. Events are split into quartiles based on the maximum ΔT anomaly (computed with MODIS Aqua LST) 0–20 days after onset. All composites have been smoothed with a 10-day running mean. In (f), VPD is computed from ERA5 2m air temperature and 2m dewpoint temperature.
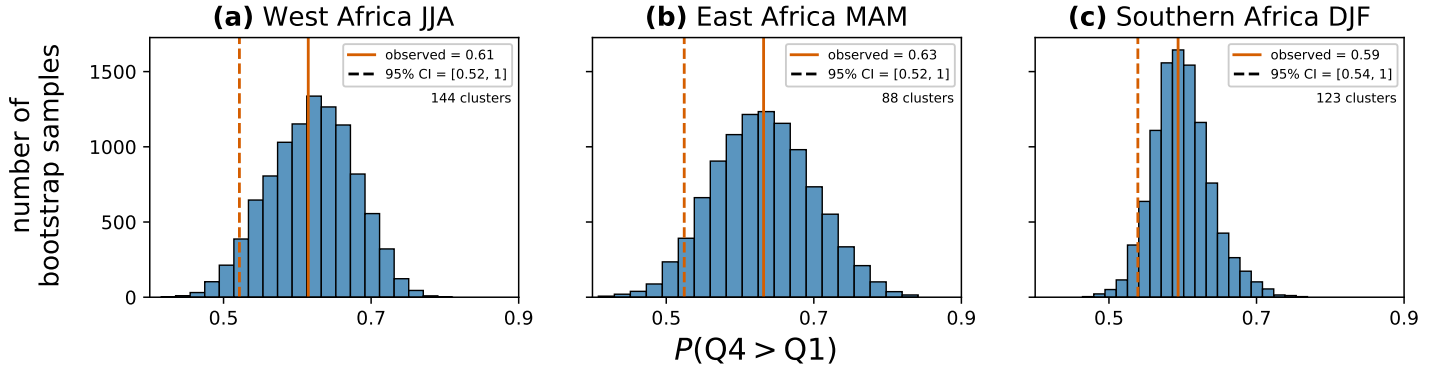
Figure S4: Distribution of $P(\text{Q4} > \text{Q1})$ for 10000 cluster bootstrap samples, to test the relationship between peak sensible heat flux and air temperature for each region shown in Figure 5a. Probabilities above 0.5 indicate that peak temperature anomalies associated with the highest quartile of peak sensible heat fluxes are stochastically larger than those associated with the lowest quartile of peak sensible heat fluxes. The solid red vertical line shows the value of $P(\text{Q4} > \text{Q1})$ for the actual observed pixels (i.e., no resampling). The dashed red line shows the lower end of the one-tailed 95% confidence interval. The number of clusters in the region is indicated below each legend.

Cluster bootstrapping for identifying land surface feedbacks to air temperature

Flash drought events identified at the 0.25° scale are not all independent of one another due to the fact that many occur as part of events with a larger spatial scale. Therefore, when testing whether the distributions of peak air temperature anomalies differ between events with weak and strong sensible heat flux anomalies, we use a cluster bootstrapping method, which accounts for clusters of 0.25° events not being independent (Davison and Hinkley, 1997; Field and Welsh, 2007; Ren et al., 2010). We define clusters by aggregating any 0.25°-scale pixels that are in flash drought conditions on the same day and are within 7 pixels' distance of each other (equivalent to approximately 200 km). Through this aggregation procedure, hundreds of pixels across multiple weeks may be identified as belonging to the same cluster. We assume that events occurring in different clusters are independent.

We use the Mann-Whitney U test to determine whether peak air temperature anomalies are higher for events in quartile 4 (Q4) of peak sensible heat flux anomalies than for those with Q1 sensible heat flux. This is a non-parametric, ranking-based test, which, for two distributions X and Y, computes the probability that a randomly selected observation from X will be larger than a randomly selected observation from Y (Mann and Whitney, 1947). We denote the probability that a temperature drawn from the distribution based on Q4 sensible heat fluxes is higher than one drawn from Q1 by $P(\text{Q4} > \text{Q1})$.

For each region tested, we perform 10000 bootstrap iterations. Each iteration is conducted as follows. There are n clusters of events in the region. From these, we sample n clusters with replacement. Within the cluster, we sample without replacement, i.e. we simply include all 0.25° events; this method has been found to be preferable to performing an additional sampling-with-replacement step within clusters (e.g. Davison and Hinkley, 1997, p. 100–102). All events from the resampled clusters are pooled and $P(\text{Q4} > \text{Q1})$ is calculated for this dataset. In this way, we obtain 10000 values of $P(\text{Q4} > \text{Q1})$, representing the results if different sets of large-scale flash droughts had been observed from within the overall population.

Figure S4 shows the distributions of these $P(\text{Q4} > \text{Q1})$ values for each of the three regions studied. We compute a one-tailed confidence interval, since we wish to test only for the possibility that higher sensible heat fluxes increase peak temperatures. The lower bound of this interval is determined by the value of $P(\text{Q4} > \text{Q1})$ that is exceeded by 95% of the bootstrap samples. For all three regions, the 95% confidence interval lies entirely above 0.5, so we conclude that events in the highest quartile of peak sensible heat flux anomalies have higher peak air temperature anomalies than events in the lowest quartile.

An analogous procedure is performed to test whether lower precursor VOD anomalies are associated with higher peak temperature anomalies, this time computing $P(\text{Q1} > \text{Q4})$. The resulting distributions for each region are shown in Figure S5. For West Africa (in JJA) and East Africa (in MAM), the 95% confidence interval for $P(\text{Q1} > \text{Q4})$ lies entirely above 0.5, so we conclude that events with precursor VOD in the lowest quartile (i.e. lowest vegetation water content) exhibit higher air temperature anomalies during the peak of the flash drought months later, compared to events with precursor VOD in the highest quartile. For Southern Africa (DJF), the 95% confidence interval overlaps 0.5 so we cannot conclude that there is a significant difference in temperature anomaly distributions between the VOD quartiles.
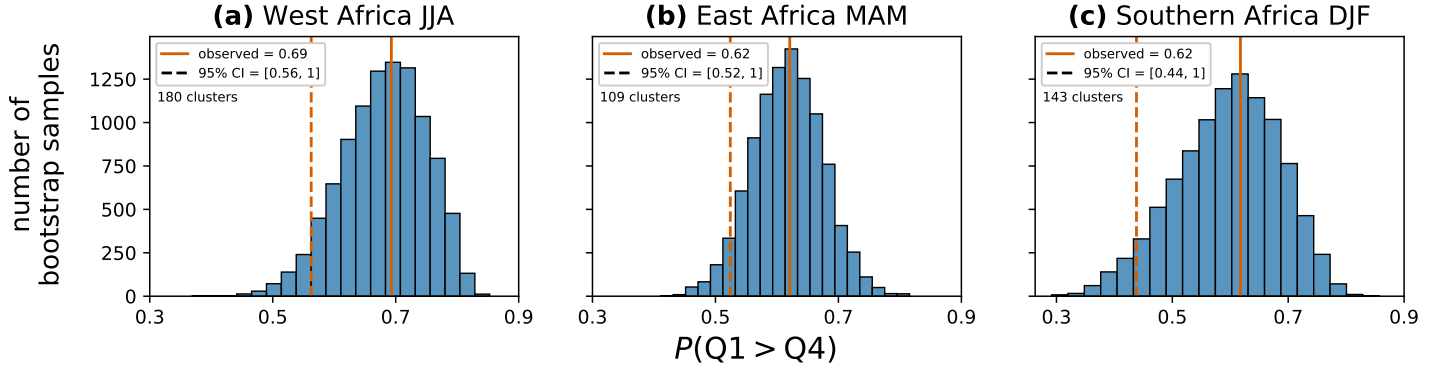
Figure S5: Distribution of $P(\text{Q1} > \text{Q4})$ for 10000 cluster bootstrap samples, to test the relationship between precursor VOD and peak air temperature for each region shown in Figure 5a. Probabilities above 0.5 indicate that peak temperature anomalies associated with the lowest quartile of precursor VOD (30–60 days before onset) are stochastically larger than those associated with the highest quartile of precursor VOD. The solid red vertical line shows the value of $P(\text{Q1} > \text{Q4})$ for the actual observed pixels (i.e., no resampling). The dashed red line shows the lower end of the one-tailed 95% confidence interval. The number of clusters in the region is indicated below each legend.
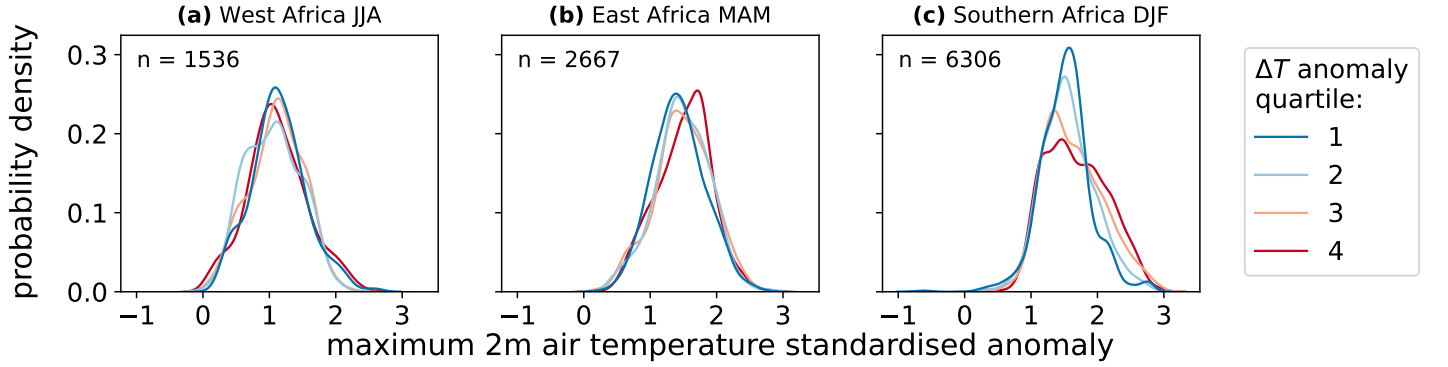
Figure S6: Probability distributions of maximum T2m standardised anomalies during flash droughts, stratified by quartiles of precursor sensible heat flux (30–60 days before drought onset, computed as $\Delta T = \text{MODIS Aqua LST} - \text{ERA5 T2m}$), for each region shown in Figure 5a. Red distributions indicate events with higher precursor sensible heat flux and blue distributions indicate events with lower precursor sensible heat flux. The number of 0.25° flash drought events in each region/season for which both ΔT and T2m data are available is denoted by n.
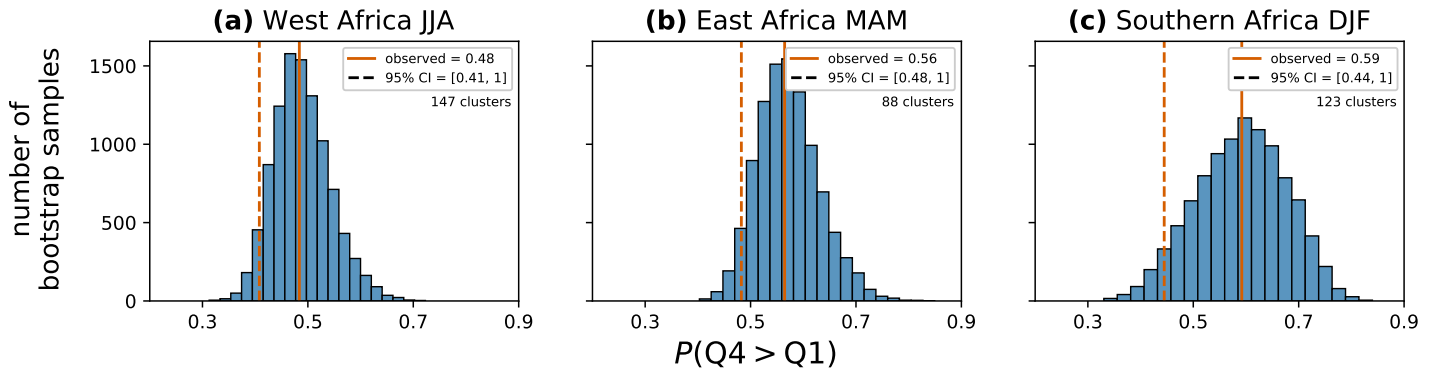
Figure S7: Distribution of $P(\text{Q1} > \text{Q4})$ for 10000 cluster bootstrap samples, to test the relationship between precursor ΔT (computed using MODIS Aqua LST) and peak air temperature for each region shown in Figure 5a. Probabilities above 0.5 indicate that peak temperature anomalies associated with the highest quartile of sensible heat flux (30–60 days before drought onset, computed as $\Delta T = \text{MODIS Aqua LST} - \text{ERA5 T2m}$) are stochastically larger than those associated with the lowest quartile of precursor sensible heat flux. The solid red vertical line shows the value of $P(\text{Q4} > \text{Q1})$ for the actual observed pixels (i.e., no resampling). The dashed red line shows the lower end of the one-tailed 95% confidence interval. The number of clusters in the region is indicated below each legend.

References

Davison, A. C. and Hinkley, D. V.: Bootstrap Methods and their Application, Cambridge University Press, Cambridge, UK, ISBN 9780521573917, https://doi.org/10.1017/CBO9780511802843, 1997.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215, https://doi.org/10.1016/J.RSE.2017.07.001, 2017.

Duveiller, G., Filipponi, F., Walther, S., Köhler, P., Frankenberg, C., Guanter, L., and Cescatti, A.: A spatially downscaled sun-induced fluorescence global product for enhanced monitoring of vegetation productivity, Earth System Science Data, 12, 1101–1116, https://doi.org/10.5194/ESSD-12-1101-2020, 2020.

Field, C. A. and Welsh, A. H.: Bootstrapping clustered data, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69, 369–390, https://doi.org/10.1111/J.1467-9868.2007.00593.X, 2007.

Ghent, D., Veal, K., Trent, T., Dodd, E., Sembhi, H., and Remedios, J.: A New Approach to Defining Uncertainties for MODIS Land Surface Temperature, Remote Sensing, 11, 1021, https://doi.org/10.3390/rs11091021, 2019.

Gruber, A., Scanlon, T., Van Der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology, Earth System Science Data, 11, 717–739, https://doi.org/10.5194/ESSD-11-717-2019, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/QJ.3803, 2020.

Huffman, G., Stocker, E., Bolvin, D., Nelkin, E., and Tan, J.: GPM IMERG final precipitation L3 1 day 0.1 degree x 0.1 degree v06 [data set], https://doi.org/10.5067/GPM/IMERGDF/DAY/06, 2019.

Jimenez, C. and Prigent, C.: ESA Land Surface Temperature Climate Change Initiative (LST_cci): All-weather MicroWave Land Surface Temperature (MW-LST) global data record (1996-2020), v2.33, https://doi.org/10.5285/a7e811fe11d34df5abac6f18c920bbeb, 2023.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F., Yoshida, Y., and Frankenberg, C.: Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, Atmospheric Measurement Techniques, 6, 2803–2823, https://doi.org/10.5194/AMT-6-2803-2013, 2013.

Köhler, P., Guanter, L., and Joiner, J.: 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, 2015.

Mann, H. B. and Whitney, D. R.: On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other, Annals of Mathematical Statistics, 18, 50–60, https://doi.org/10.1214/AOMS/1177730491, 1947.

Miralles, D. G., Bonte, O., Koppa, A., Villanueva, O. B., Tronquo, E., Zhong, F., Beck, H. E., Hulsman, P., Dorigo, W., Verhoest, N. E. C., and Haghdoost, S.: GLEAM4: global land evaporation dataset at 0.1° resolution from 1980 to near present, https://doi.org/10.21203/rs.3.rs-5488631/v1, 2024.

Preimesberger, W., Scanlon, T., Su, C. H., Gruber, A., and Dorigo, W.: Homogenization of Structural Breaks in the Global ESA CCI Soil Moisture Multisatellite Climate Data Record, IEEE Transactions on Geoscience and Remote Sensing, 59, 2845–2862, https://doi.org/10.1109/TGRS.2020.3012896, 2021.

Ren, S., Lai, H., Tong, W., Aminzadeh, M., Hou, X., and Lai, S.: Nonparametric bootstrapping for hierarchical data, Journal of Applied Statistics, 37, 1487–1498, https://doi.org/10.1080/02664760903046102, 2010.

Rutan, D. A., Kato, S., Doelling, D. R., Rose, F. G., Nguyen, L. T., Caldwell, T. E., and Loeb, N. G.: CERES Synoptic Product: Methodology and Validation of Surface Radiant Flux, Journal of Atmospheric and Oceanic Technology, 32, 1121–1143, https://doi.org/10.1175/JTECH-D-14-00165.1, 2015.

Zotta, R.-M., Moesinger, L., van der Schalie, R., Vreugdenhil, M., Preimesberger, W., Frederikse, T., de Jeu, R., and Dorigo, W.: VODCA v2: multi-sensor, multi-frequency vegetation optical depth data for long-term canopy dynamics and biomass monitoring, Earth System Science Data, 16, 4573–4617, https://doi.org/10.5194/ESSD-16-4573-2024, 2024.