Supplement of Biogeosciences, 19, 2557–2581, 2022
https://doi.org/10.5194/bg-19-2557-2022-supplement

Biogeosciences
Open Access

Supplement of

Assimilation of passive microwave vegetation optical depth in LDAS-Monde: a case study over the continental USA

Anthony Mucia et al.

Correspondence to: Jean-Christophe Calvet (jean-christophe.calvet@meteo.fr)

This supplement contains additional information regarding the land use data of ECOCLIMAP, and its representation over CONUS, as well as a demonstration of how the probability distribution function transformation compares to a histogram of correlation scores.

1 ECOCLIMAP-SG dominant land cover over CONUS

Figure S1 displays the classes of patches used in the ECOCLIMAP-SG land use databases, including nine plant types (evergreen broadleaf trees, needle leaf trees, deciduous broadleaf trees, herbaceous, tropical herbaceous, wetlands, C3 crops, C4 crops, and C4 irrigated crops) and the non-vegetation surfaces of rocks, bare soil, and permanent snow and ice. This map can also be used as a reference of dominant regional vegetation types and their relation to the results over the domain.

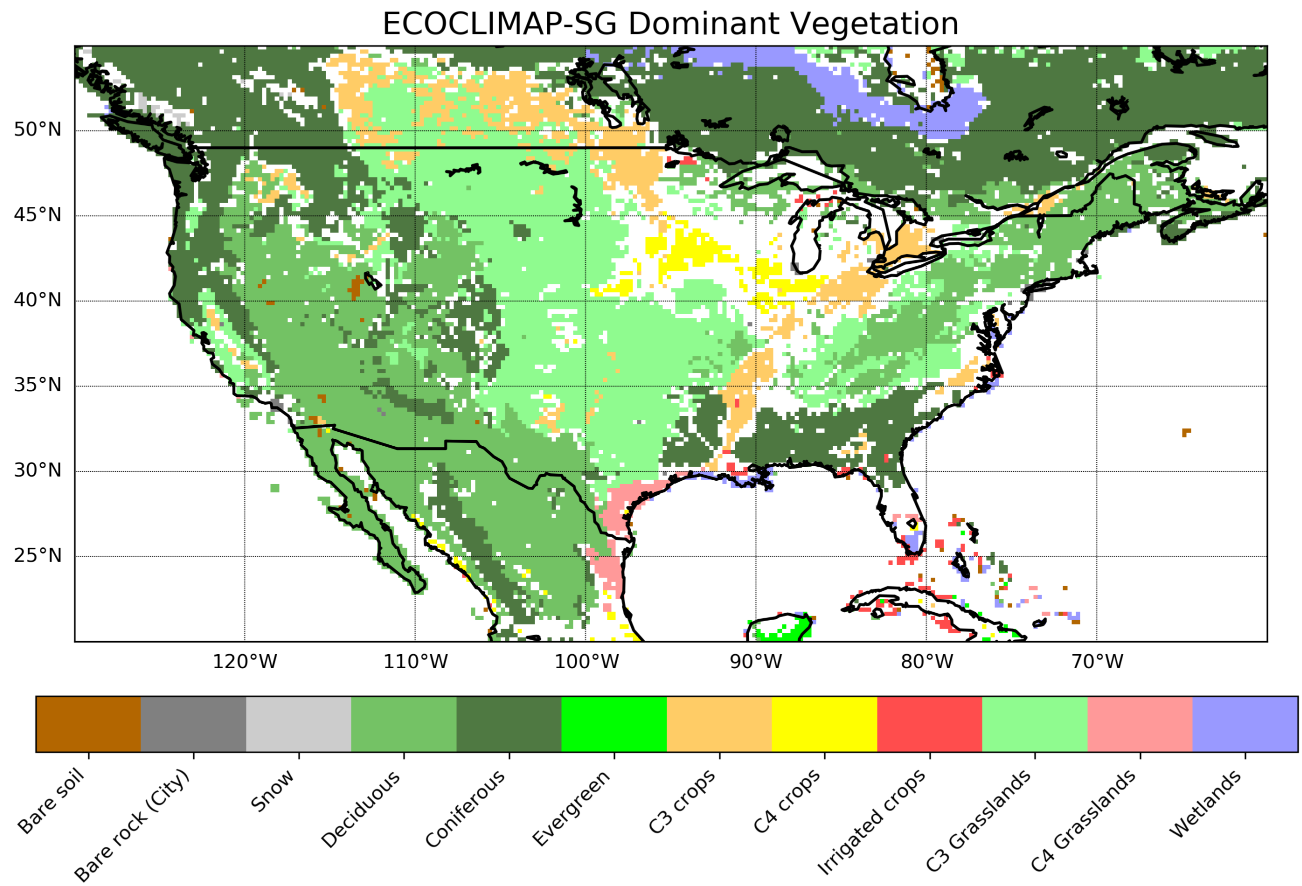

Figure S1. Map of dominant vegetation types (>50% single vegetation) from ECOCLIMAP-SG. White patches over land indicate no single vegetation type covers more than 50% of the tile.

2 VOD re-scaling code segment

The following Python pseudocode segment was used to re-scale VOD values based on LAI observations. This code demonstrates that the re-scaling is done via twelve linear equations (only one month of re-scaling is shown in the code), one for each month of the year, and accounting for the preceding and following month's climatology.

The primary variables are "vodx" and "lai_cgls" which are DataFrames of VODX observations from VODCA and LAI observations from CGLS respectively, over the period of 2003 to 2018 and over the CONUS domain as described in the Methdology. The resulting variable, "rescaled", also a DataFrame, is the linearly re-scaled VODX.

```
for i in range((vodx.shape[1])):
    season=((previous_month) | (considered_month) | (next_month))
    if ((len(vodx[i].ix[season][~np.isnan(vodx[i].ix[season])]) > 0)
        try:
            b = (lai_cgls[i].ix[season][~np.isnan(lai_cgls[i].ix[season])]).std()/\
                    (vodx[i].ix[season][~np.isnan(vodx[i].ix[season])]).std()
            a = (lai_cgls[i].ix[season][~np.isnan(lai_cgls[i].ix[season])]).mean() - \
                    b*(vodx[i].ix[season][~np.isnan(vodx[i].ix[season])]).mean()
            rescaled[i].ix[considered_month] = b*(vodx[i].ix[considered_month]) + a
        except ValueError:
            print('ValueError')
```

3 Verification of Probability Density Function versus Histogram Relationship

For several analyses, probability distribution functions (PDFs) are estimated from the distribution of correlation scores of individual gridcells. These PDFs are derived using a Gaussian kernel density estimation, with "Scott's Rule" calculating the appropriate smoothing bandwidth. These PDFs give a far smoother and readable estimation of correlations when compared to simple histograms. These PDFs are used in order to better visualize and compare the distribution of scores from several experiments at once, specifically to highlight smaller differences between experiments, which may not be easily seen in histograms. Figure S2 A) provides an example of a typical histogram along with the B) accompanying PDF for the distribution of LAI correlations for different experiments over CONUS, demonstrating that they are in fact representing the same distribution.

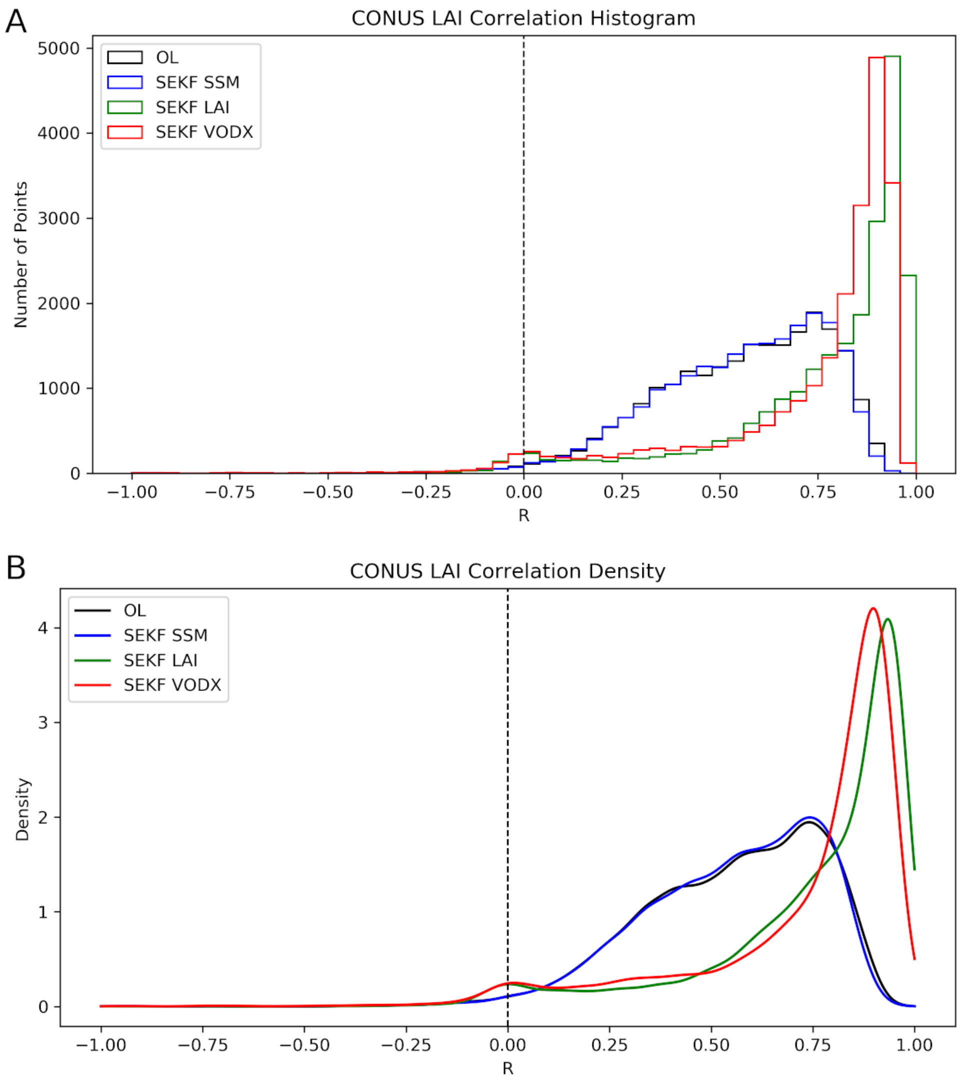

Figure S2. Examples of A) histogram distribution of scores and B) probability distribution function of scores for LAI correlations for OL (black), SEKF SSM (blue), SEKF LAI (green), and SEKF VODX (red) experiments over CONUS. This figure demonstrates that the histogram and PDF represent the same relationship.