Supplement of Earth Syst. Sci. Data, 16, 1333–1351, 2024
https://doi.org/10.5194/essd-16-1333-2024-supplement
© Author(s) 2024. CC BY 4.0 License.

Supplement of

A global estimate of monthly vegetation and soil fractions from spatiotemporally adaptive spectral mixture analysis during 2001–2022

Qiangqiang Sun et al.

Correspondence to: Hong Li (lih5176@126.com) and Danfeng Sun (sundf@cau.edu.cn)

The copyright of individual parts of the supplement might differ from the article licence.

Supplementary Methods

Seasonal Mann-Kendall test

The seasonal Mann-Kendall test is commonly applied to identify trends for seasonal environmental data of interest that is available as time series for which the time intervals between adjacent observations arc less than one year (i.e., daily, weekly, and monthly sequences) (Hirsch et al. 1982). Letting the sequence *X* consists of a complete seasonal record of *n* year that includes *m* seasons per year, the *X* can be expressed by

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

The null hypothesis, H_0, is that the *n* observations come from each of *m* seasons with independent realizations are identically distributed. While the alternative hypothesis (H_A) of a two-sided test is that data presents a monotonic trend. The Seasonal Mann-Kendall test statistic for the *g* th season is

$$S_g = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_{jg} - x_{ig}), g = 1,2,\ldots,m$$

where

$$sgn(\theta) = \begin{cases} 1 \ if \ \theta > 0 \\ 0 \ if \ \theta = 0 \\ -1 \ if \ \theta < 0 \end{cases}$$

Sg is asymptotically normally distributed, thus the mean of *Sg* is $E[S_g] = 0$, and the variance is

$$Var[S_g] = \left\{ n(n-1)(2n+5) - \sum_{j=1}^{p} t_j(t_j - 1)(2t_j + 5) \right\} / 18$$

where *n* is the number of years of each season, *p* is the number of tied groups for data x_{ig}, *i*=1,2,…*n*, in season *g*, and t_j is the number of data points in the *j*th tied group. The seasonal Mann-Kendall test statistic is

$$S = \sum_{g=1}^{m} S_g$$

which is also asymptotically normally distributed where $E[S] = 0$, thus the variance of S is

$$Var[S] = \sum_{g=1}^{m} Var[S_g]$$

And the statistic S is approximately normal distributed provided that the following Z-transformation is employed,

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{Var[S]}}\ if\ S > 0 \\ 0\ if\ S = 0 \\ \dfrac{S+1}{\sqrt{Var[S]}}\ if\ S < 0 \end{cases}$$

For a given α-significance level, the original null hypothesis (H_0) is unacceptable if |Z| ≥ $Z_{1-\alpha/2}$. This implies a significantly upward or downward trend in the series.

Theil-Sen estimator is a method of robust linear regression by selecting the median value of the slope of all lines passing through the paired points. It is also known as Sen's slope estimation (Sen and Kumar, 1968). Here, we detect slope of fractions according to seasonal Sen's method. For sequence *X* consisting of a complete seasonal record of *n* year that includes *m* seasons per year, a set of linear slopes is calculated as,

$$d_{g,j,k} = \frac{x_{g,j} - x_{g,k}}{j-k}$$

For each $x_{i,j}, x_{i,k}\ pair\ i = 1, 2, \ldots, m,\ 1 \leq k < j \leq n$, where *n* is length of *g*th season. and seasonal Sen's slope is then calculated as the median from all slopes.

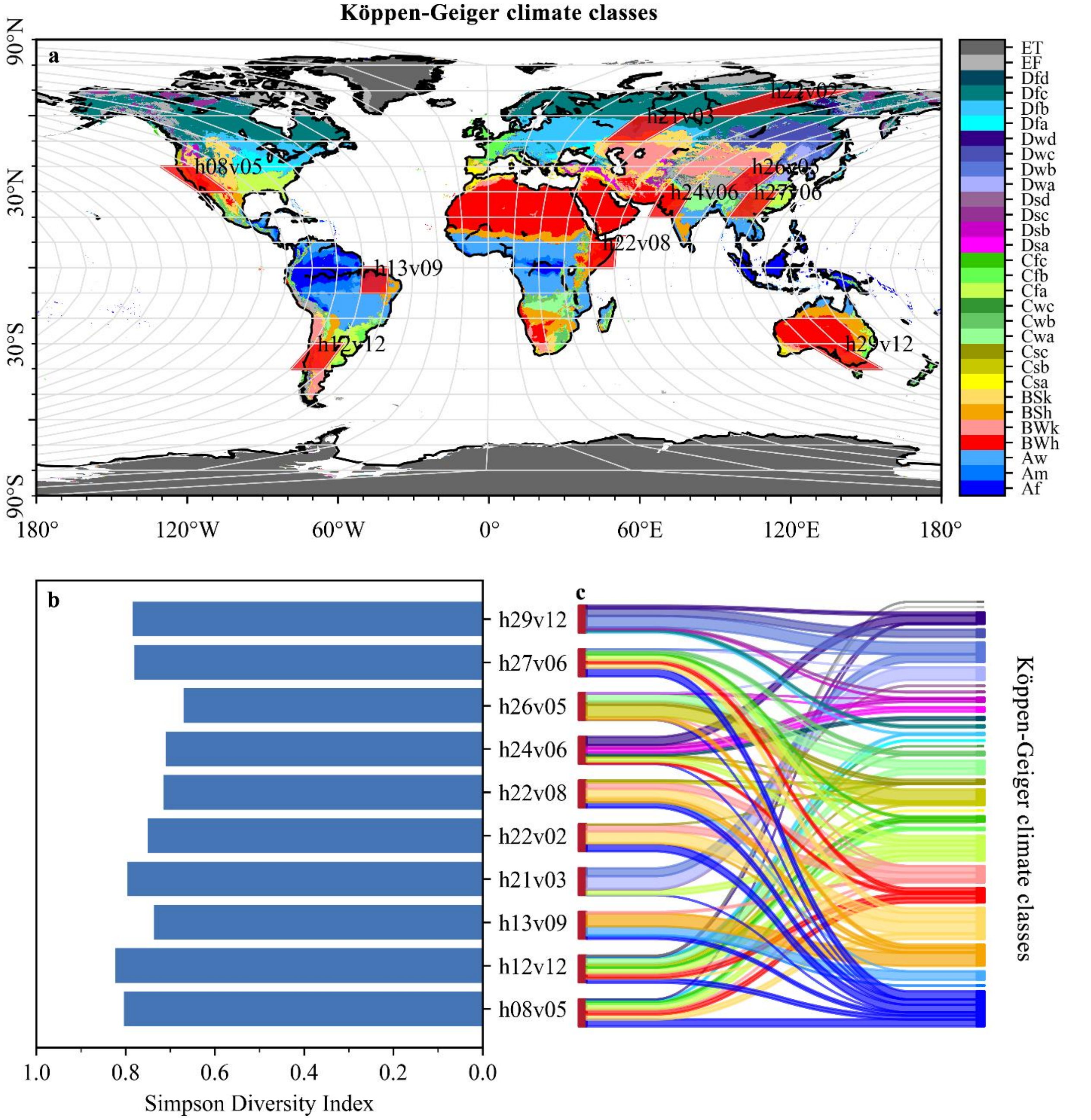


Figure S1: The typical MODIS Grids for endmembers selection. **a**, 10 selected MODIS grids (i.e., h08v05, h12v12, h13v09, h16v01, h21v03, h22v02, h22v08, h24v06, h26v05, h27v06, h29v12), distributed in 6 continents, were colored with red. **b**, The Simpson's Diversity Index (D) greater than 0.6 for selected MODIS grids. **c**, the correspondence between selected MODIS grids and Köppen-Geiger climate classification, indicating each Köppen-Geiger climate classification was represented by selected MODIS grids.

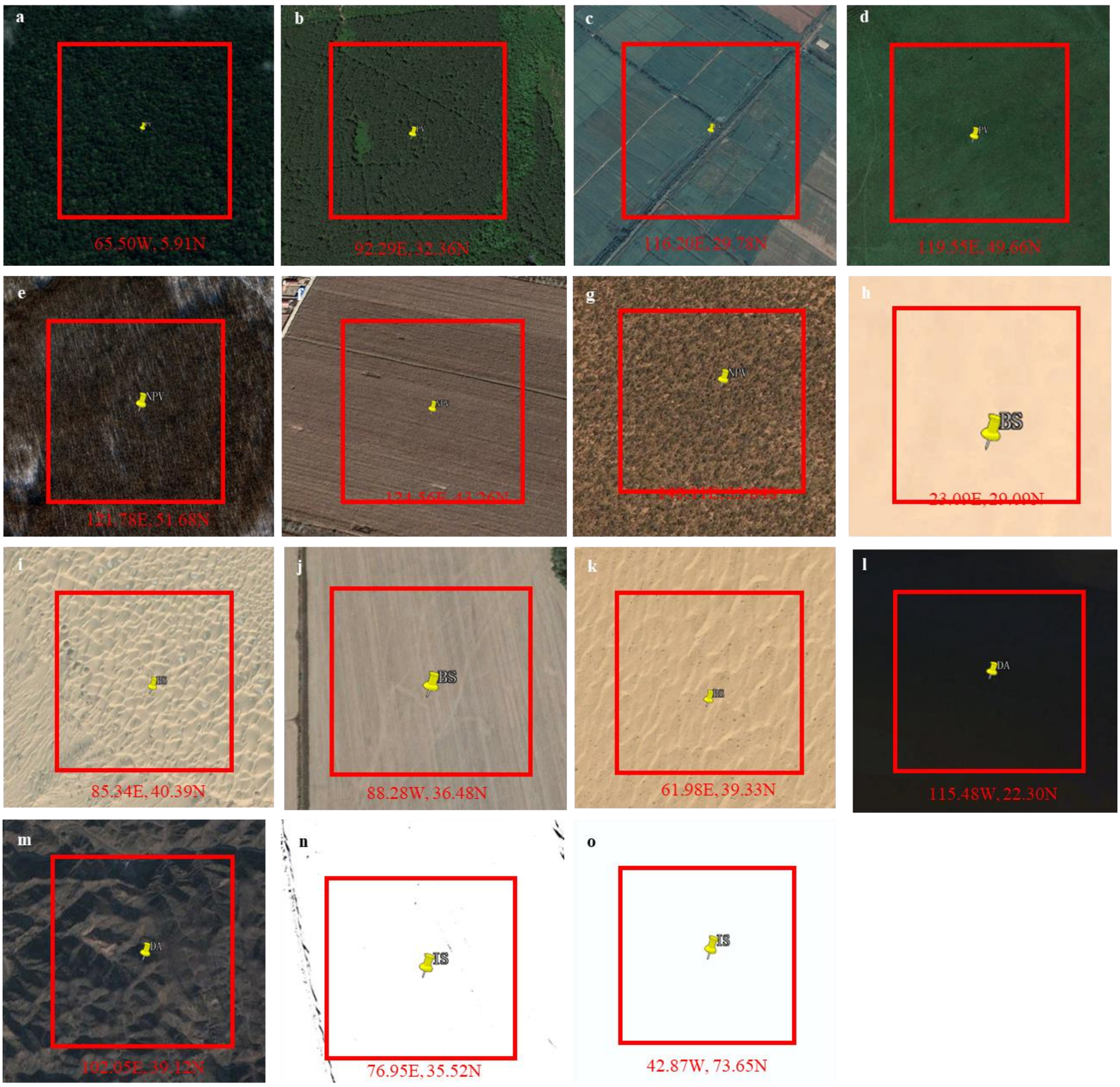


Figure S2: Typical images representing selected pure pixel of each endmember. a-o are tropical rainforest, temperate forest, cropland, grasslands, temperate deciduous forest in winter, crop residues, shrubs in dryland, moving sands, sand dunes, bare ground, moving sands. waters, bare rock, polar glaciers and alpine glaciers.

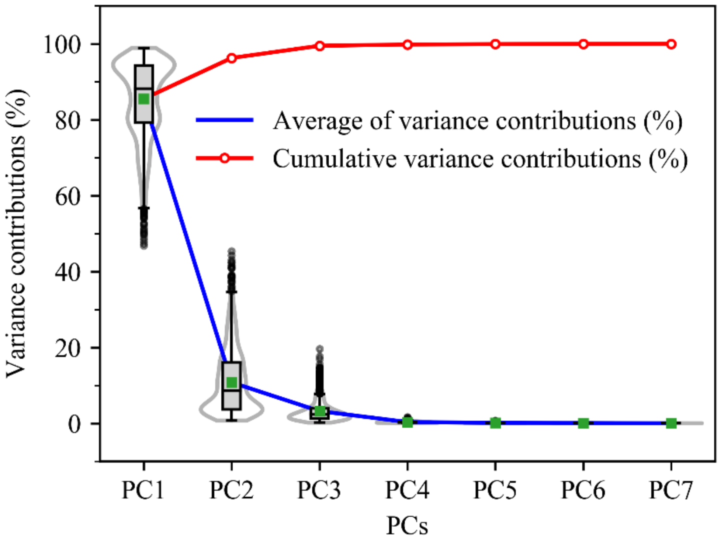


Figure S3: Variance contributions (%) of principal component transformation for each monthly MODIS reflectance image across 10 grids (n=2280). The violin plot and box plot revealed value distribution of each PC. The blue and red lines represented average of variance contributions and cumulative variance contributions of each PC, respectively. The cumulative contribution of the top three PCs has exceeded 99%.

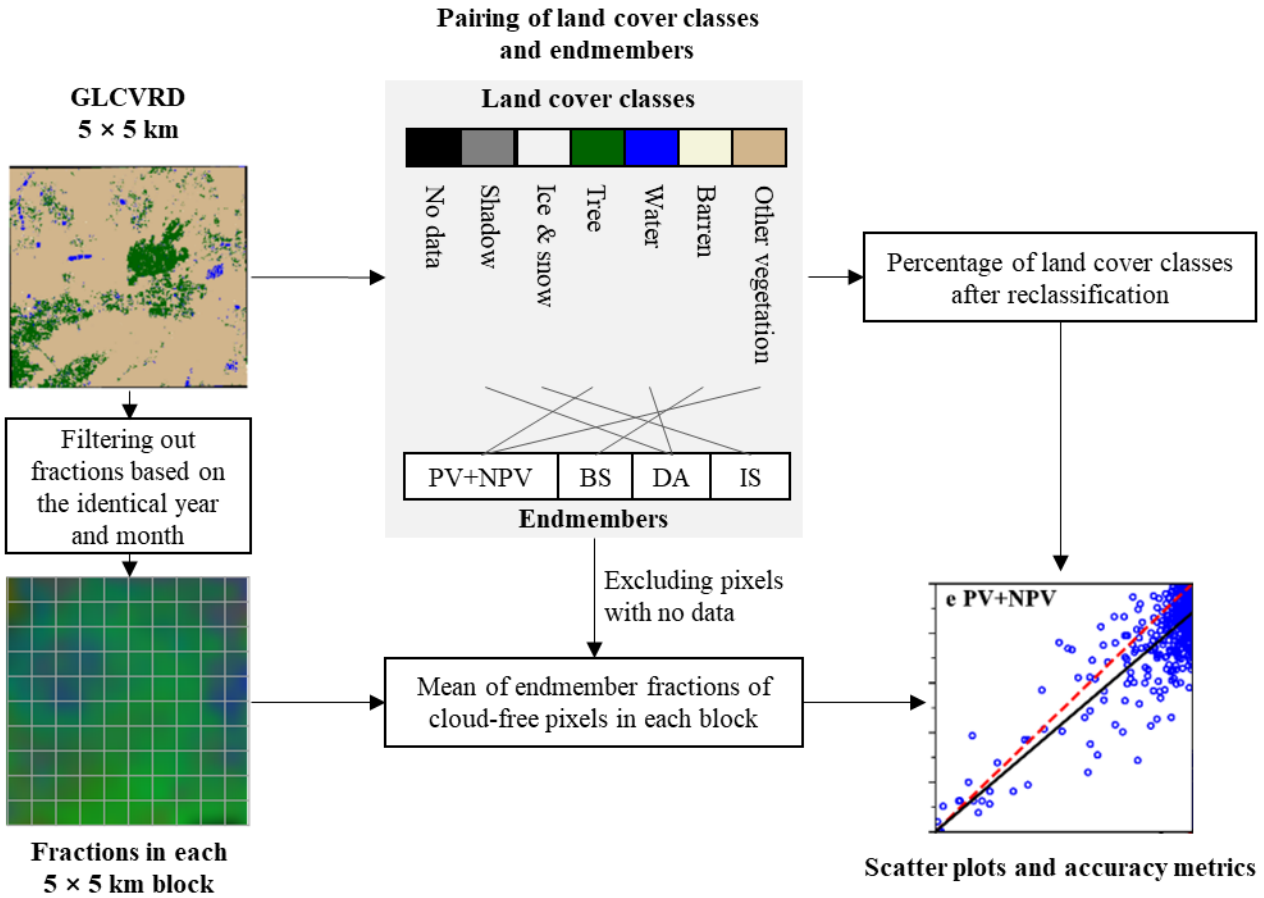


Figure S4: Procedural description for estimating fractional vegetation-soil compared to GLCVRD.

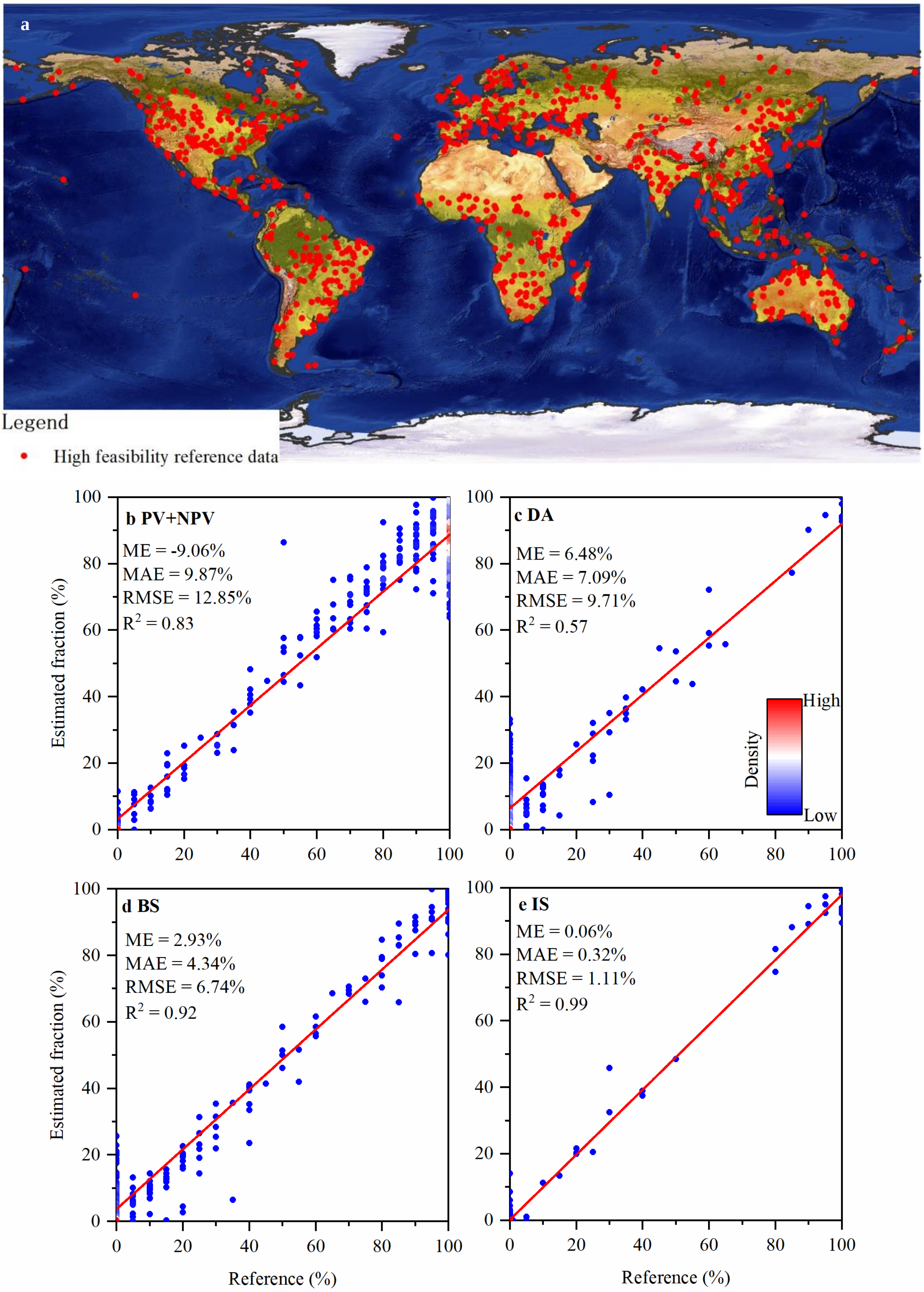
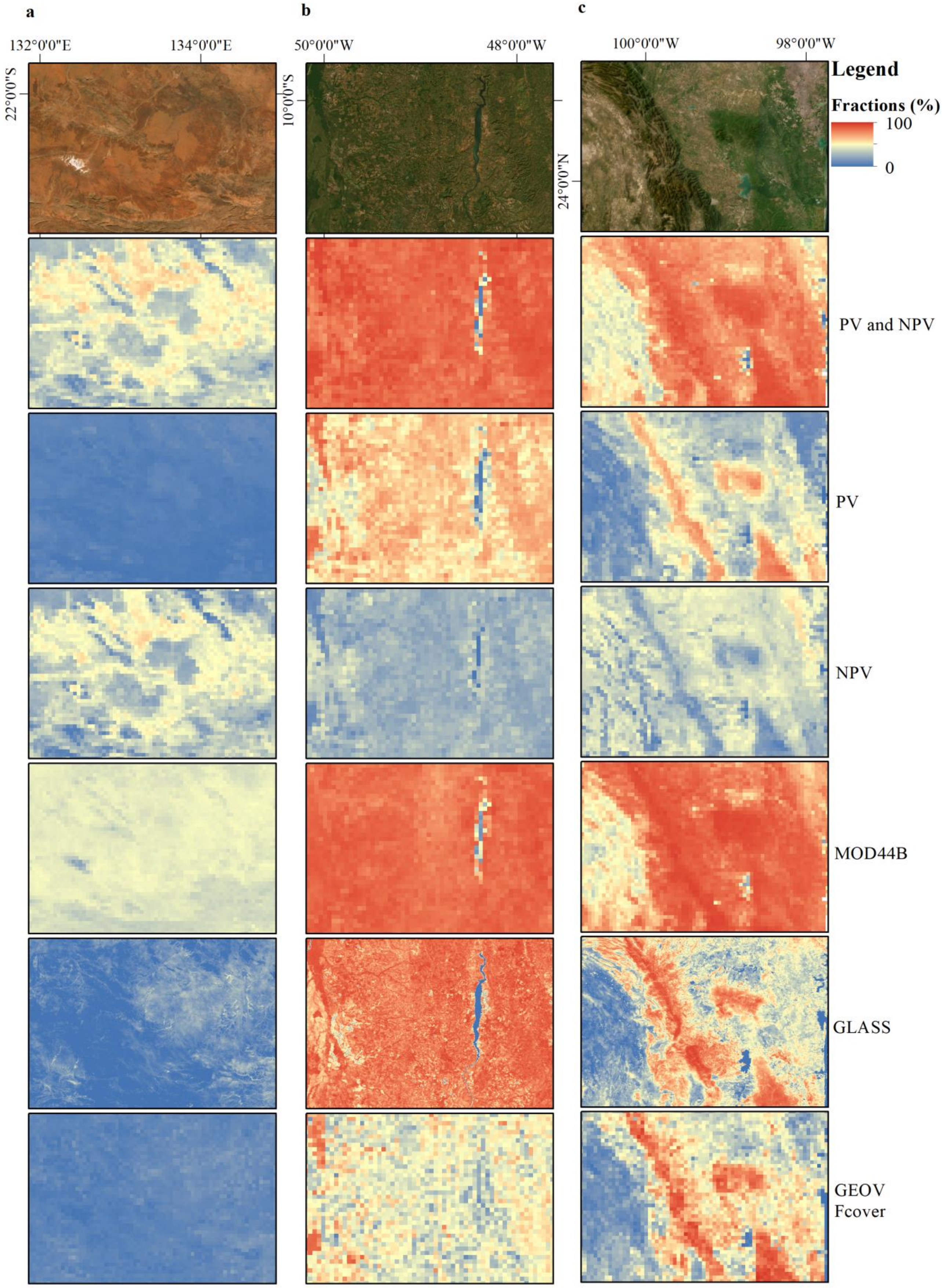


Fig. S5 Evaluation of global fractional endmember estimates based on land cover reference data. a, the location of high-feasibility land cover reference data. **b-d,** Scatter plots of PV+NPV, BS, DA, IS fractions against land cover reference data



97

98 **Fig. S6 The detailed graphs for comparing different datasets. a, b,** and **c** represent comparisons of vegetation

abundance products in different scenarios, specifically, regions with low vegetation cover in arid areas, high vegetation cover in tropical rainforests, and transitional zones from low to high values. The compared products include our produced PV and NPV, MOD44B, GLASS, and GEOV.

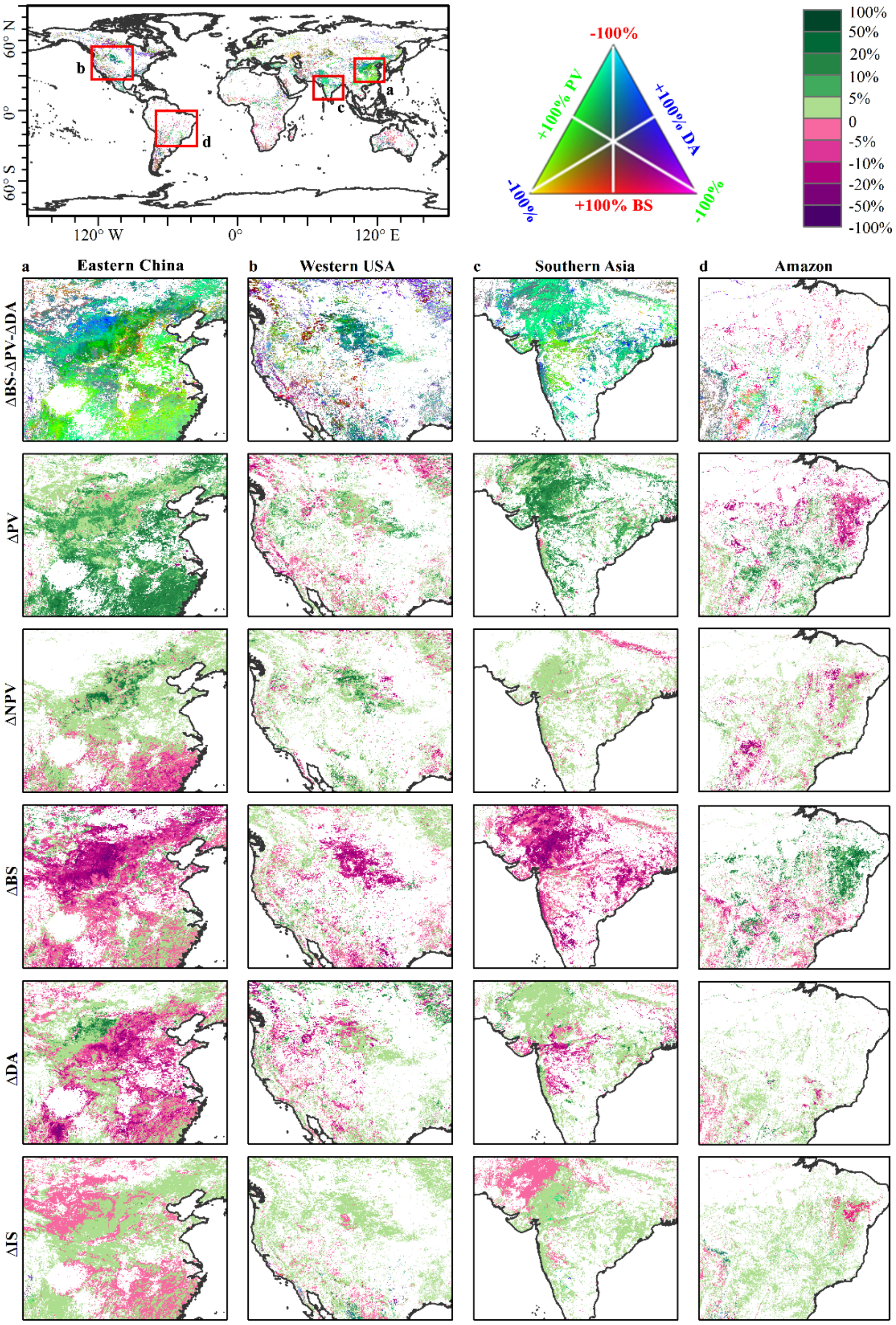


Fig. S7 Regional detailed subsets of changes for endmember fractions. From top to bottom, the first row represented composited images with ΔBS, ΔPV, and ΔDA. the second to sixth row displayed the change magnitude (%) in each pixel for estimated endmembers, i.e., ΔPV, ΔNVP, ΔBS, ΔDA, and ΔIS. Pixels showing a statistically significant trend ($n = 228$, Seasonal Mann–Kendall test, $P < 0.05$) for either endmember are depicted on the change map. a-d, Eastern China, Western USA, Southern Asia, and Amazon.

Table S1 Selected typical MODIS grids for endmember selection and corresponding timing for each endmember.

Girds	Zones	**IS**	**NPV**	**PV**	**BS**	**DA**
h08v05	West Coast of North America	2016-12	2005-02	2010-07	2005-02	2019-10
		2009-12	2005-03	2012-07	2005-03	2007-10
		2017-01	2010-03	2013-07	2010-03	2003-10
h12v12	West Coast of South America	2003-06	2007-09	2007-01	2007-09	2003-04
		2016-05	2011-09	2016-02	2011-09	2009-04
		2005-08	2008-09	2014-02	2008-09	2013-04
h13v09	Amazon Basin	-	2019-10	2016-01	2019-10	
		-	2013-10	2018-02	2013-10	
		-	2016-10	2004-03	2016-10	
h21h03	Western Europe -Central Asia	2018-02	2007-10	2014-09	2017-10	
		2011-02	2016-10	2013-09	2016-10	
		2005-03	2016-10	2004-09	2009-10	
h22h02	Russian Far East	2003-01		2019-06	2019-06	2006-03
		2009-12		2018-06	2018-06	2011-03
		2010-12		2019-07	2019-07	2012-03
h22h08	East Coast of North Africa	-		-	2005-07	2013-04
		-		-	2006-10	2018-05
		-		-	2016-06	2019-10
h24v06	South Asia	-		2018-07	2018-07	
		-		2004-07	2004-07	
		-		2008-07	2008-07	
h26v05	Northwest of Qinghai Tibet Plateau	2008-01		2013-07	2007-03	2007-09
		2019-02		2006-07	2007-11	2017-09
		2008-02		2010-07	2003-11	2008-10
h27v06	Southwest China	-		2013-09	2007-11	
		-		2006-08	2019-10	
		-		2017-07	2004-10	
h29v12	South Australia	-	2007-01	2019-08	2017-05	
		-	2003-01	2016-09	2011-05	
		-	2006-12	2016-08	2012-05	

Table S2 692 combination models. These models include two-endmember model, three-endmember model and four-endmember model.

Models	combinations
two-endmember model (88)	PV+BS (16)
	PV+DA (8)
	PV+IS (8)
	BS+DA (8)
	BS+IS (8)
	DA+IS (4)
	PV+NPV (12)
	BS+NPV(12)
	DA+NPV(6)
	IS+NPV(6)
three-endmember model (252)	PV+BS+DA (32)
	PV+BS+IS (32)
	PV+DA+IS (16)
	BS+DA+IS(16)
	PV+BS+NPV (48)
	PV+DA+NPV (24)
	PV+IS+NPV (24)
	BS+DA+NPV (24)
	BS+IS+NPV(24)
	DA+IS+NPV (12)
four-endmember model (352)	PV+BS+DA+IS (64)
	PV+BS+DA+NPV (96)
	PV+BS+IS+NPV (96)
	PV+DA+IS+NPV(48)
	BS+DA+IS+NPV (48)

Table S3 Evaluation of estimated five vegetation and soil components against GLCVRD reference dataset.

	ME	MAE	RMSE	R^2
PV+NPV	-0.100	0.118	0.149	0.592
BS	0.047	0.075	0.109	0.710
DA	0.047	0.050	0.065	0.156
IS	0.008	0.008	0.020	0.792

Table S4 Global and regional fractional endmembers dynamics. The initial area, gain area, loss area and net change area are calculated for globe and five climate zones, i.e., tropical, arid, temperate, cold, and polar.

Zones	Endmembers	Initial area (km^2)	Loss (km^2)	Gain (km^2)	Net change area (km^2)
Globe	PV	49861610.43	-637802.76	1573227.20	935424.44
	NPV	15128296.27	-564426.41	345275.53	-219150.87
	BS	46994793.90	-1381062.00	866955.99	-514106.02
	DA	32708319.15	-887176.16	660346.40	-226829.76
	IS	22247774.92	-114679.83	139342.04	24662.21
Tropical	PV	16018642.05	-236558.32	231221.99	-5336.33
	NPV	2217754.40	-109780.44	24544.40	-85236.04
	BS	3830525.28	-197155.41	319038.83	121883.42
	DA	2001545.61	-61499.28	29922.32	-31576.96
	IS	133822.65	-3499.33	3765.24	265.91
Temperate	PV	9393375.78	-78191.98	413284.59	335092.61
	NPV	2417754.40	-154026.78	52210.12	-101816.66
	BS	3530525.28	-316382.59	102279.47	-214103.11
	DA	3001545.61	-79412.55	62801.35	-16611.20
	IS	133822.65	-9419.53	6857.89	-2561.64
Arid	PV	7302896.73	-166255.40	431469.49	265214.10
	NPV	2753166.01	-158937.90	149897.07	-9040.83
	BS	34208215.02	-577137.23	402055.35	-175081.89
	DA	6055362.59	-301342.53	226780.57	-74561.96
	IS	3234817.92	-79628.12	73098.70	-6529.42
Cold	PV	15661149.01	-153482.53	488284.80	334802.27
	NPV	6556417.74	-140497.42	114025.71	-26471.71
	BS	3772331.10	-278401.74	38748.72	-239653.02
	DA	17529802.86	-438961.80	334040.19	-104921.61
	IS	9580056.20	-17785.65	54029.72	36244.07
Polar	PV	1485546.86	-3314.53	8966.33	5651.80
	NPV	1183203.72	-1183.88	4598.24	3414.36
	BS	1653197.22	-11985.03	4833.61	-7151.42
	DA	4120062.48	-5960.00	6801.97	841.96
	IS	9165255.50	-4347.20	1590.50	-2756.70