Article

Large Chinese land carbon sink estimated from atmospheric carbon dioxide data

Jing Wang[1,2,3,4], Liang Feng[3,4], Paul I. Palmer[3,4 ✉], Yi Liu[1,2 ✉], Shuangxi Fang[5,6 ✉], Hartmut Bösch[7], Christopher W. O'Dell[8], Xiaoping Tang[9], Dongxu Yang[1,2], Lixin Liu[6] & ChaoZong Xia[9]

Limiting the rise in global mean temperatures relies on reducing carbon dioxide (CO_2) emissions and on the removal of CO_2 by land carbon sinks. China is currently the single largest emitter of CO_2, responsible for approximately 27 per cent (2.67 petagrams of carbon per year) of global fossil fuel emissions in 2017[1]. Understanding of Chinese land biosphere fluxes has been hampered by sparse data coverage[2–4], which has resulted in a wide range of a posteriori estimates of flux. Here we present recently available data on the atmospheric mole fraction of CO_2, measured from six sites across China during 2009 to 2016. Using these data, we estimate a mean Chinese land biosphere sink of −1.11 ± 0.38 petagrams of carbon per year during 2010 to 2016, equivalent to about 45 per cent of our estimate of annual Chinese anthropogenic emissions over that period. Our estimate reflects a previously underestimated land carbon sink over southwest China (Yunnan, Guizhou and Guangxi provinces) throughout the year, and over northeast China (especially Heilongjiang and Jilin provinces) during summer months. These provinces have established a pattern of rapid afforestation of progressively larger regions[5,6], with provincial forest areas increasing by between 0.04 million and 0.44 million hectares per year over the past 10 to 15 years. These large-scale changes reflect the expansion of fast-growing plantation forests that contribute to timber exports and the domestic production of paper[7]. Space-borne observations of vegetation greenness show a large increase with time over this study period, supporting the timing and increase in the land carbon sink over these afforestation regions.

The global mass and growth of atmospheric CO_2 can be determined by surface-based measurements (www.esrl.noaa.gov/gmd/ccgg/trends/). The largest anthropogenic emissions of CO_2 include fossil fuel combustion, cement production and human-driven changes to land use[1]. Reconciliation of bottom-up emission estimates of these anthropogenic processes, accounting for widely assumed uncertainties, with atmospheric CO_2 measurements reveals a surprisingly invariant airborne fraction of 44 ± 14%[8]. The remainder is absorbed by the ocean and land biosphere[9]. Sparse measurements of partial pressure of CO_2 in the surface ocean, p_{CO_2}, taking advantage of large-scale spatial–temporal correlations, allow estimation of ocean CO_2 fluxes[10], although regional campaigns highlight large flux variations in successive years[11]. The residual of these mass balance terms yields the magnitude of the land biosphere carbon balance[9]. Current understanding of the land biosphere, encapsulated in numerical models, requires additional uptake to be reconciled with the data-driven residual term, even accounting for uncertainties[1]. Therein lies the crux of the challenge faced by the science and policy communities: effective mitigation of fossil fuel CO_2 emissions within a large-scale, dynamic natural carbon cycle that we do not quantitatively understand.

China overtook the United States in 2006 as the single largest emitter of CO_2, mainly due to the fossil fuel combustion and cement production sectors[12]. These emission estimates have large uncertainties[13] that translate into larger uncertainties for residual terrestrial biosphere flux estimates[2,3] (see Methods). Sparse in situ observations over China result in a large model spread of a posteriori terrestrial fluxes estimates[2,3,14], with values highly sensitive to available data[4].

Constraints on CO_2 flux estimates

In this study, we use newly available in situ CO_2 mole fraction data (collectively describing weekly flask and hourly continuous measurements) over China (see Methods section 'Data' and Supplementary Information; Extended Data Fig. 1) from 2009 to 2016 to show how these data can revise the provincial and national carbon balance of China. We corroborate our unexpected a posteriori flux estimates through independent satellite remote-sensing measurements of vegetation greenness, soil water availability, a posteriori fluxes inferred from satellite column observations of CO_2, and consecutive Chinese forest censuses.

We use several sources of data to investigate the Chinese carbon balance (see Methods and Supplementary Information): in situ measurements of CO_2 mole fraction, satellite column observations of CO_2, leaf area index (LAI), normalized difference and enhanced vegetation

[1]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]National Centre for Earth Observation, University of Edinburgh, Edinburgh, UK. [4]School of GeoSciences, University of Edinburgh, Edinburgh, UK. [5]College of Environment, Zhejiang University of Technology, Hangzhou, China. [6]Meteorological Observation Centre, China Meteorological Administration, Beijing, China. [7]National Centre for Earth Observation, University of Leicester, Leicester, UK. [8]Colorado State University, Fort Collins, CO, USA. [9]Academy of Forest Inventory and Planning, State Forestry Administration, Beijing, China. ✉e-mail: paul.palmer@ed.ac.uk; liuyi@mail.iap.ac.cn; fangsx@cma.gov.cn
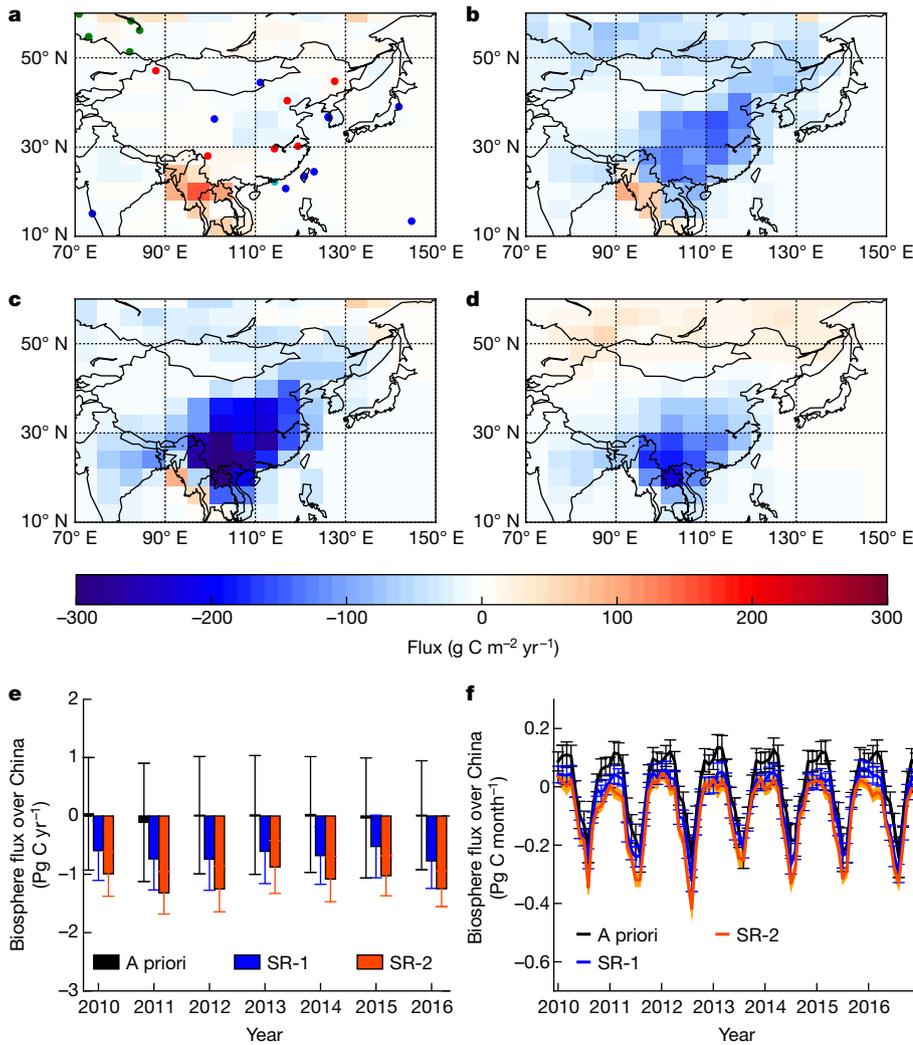
Fig. 1 | Chinese terrestrial biosphere CO_2 fluxes. **a**–**c**, Annual mean spatial distributions of CO_2 fluxes (g C m^{-2} yr^{-1}) for 2010–2016 (inclusive) corresponding to the a priori inventory (**a**) and to a posteriori CO_2 fluxes corresponding to SR-1 (**b**) and SR-2 (**c**). **d**, Difference between SR-2 and SR-1 (g C m^{-2} yr^{-1}). In **a**, the blue, cyan, green and red dots denote NOAA ObsPack stations, the Hok Tsui (HKG) station from World Data Centre for Greenhouse Gases (WDCGG), Japan–Russia Siberia Tall Tower Inland Observation Network and the China Meteorological Administration (CMA) regional background stations, respectively. **e**, Annual CO_2 fluxes from the a priori inventory, and a posteriori fluxes inferred from (SR-1) NOAA ObsPack CO_2 mole fraction measurements, and (SR-2) NOAA + Siberian + CMA + HKG CO_2 mole fraction measurements from 2010 to 2016, inclusive. **f**, Corresponding monthly fluxes. In **e** and **f**, vertical bars and the orange envelope denote a priori and a posteriori uncertainties.

indices (NDVI and EVI, respectively), liquid water equivalent thickness derived from gravity anomalies, and above-ground biomass inferred from microwave measurements. The atmospheric data are interpreted using the GEOS-Chem atmospheric transport model coupled with a Bayesian inversion tool to infer CO_2 fluxes from the atmospheric data (see Methods).

CO_2 fluxes inferred from atmospheric data

First, we use an atmospheric inversion to estimate the magnitude and distribution of land biosphere CO_2 fluxes over China (see Methods). Inversion SR-1, which uses only the ObsPack GLOBALVIEWplus data (Fig. 1a) provided by the Earth System Research Laboratory (ESRL) of the National Oceanic and Atmospheric Administration (NOAA), shows a CO_2 sink over most of China (Fig. 1b), mainly focused over central, northeast and southern China. These data are consistent with a mean annual net uptake estimate for the terrestrial biosphere of −0.66 ± 0.52 petagrams of carbon per year (Pg C yr^{-1}; 2010–2016), varying from −0.52 ± 0.54 Pg C yr^{-1} in 2015 to −0.77 ± 0.47 Pg C yr^{-1} in 2016 (Fig. 1e, Table 1). We find that the annual Chinese carbon sink is mostly due to uptake during the growing season, particularly during June to August (Fig. 1f).

Including the Siberian tower data (see Data and Supplementary Table 1) broadly increases CO_2 uptake over China and India. The additional Chinese mainland sites (Extended Data Fig. 1), together with the Siberian and Hong Kong data, in inversion SR-2 (Fig. 1c) have two major impacts on our understanding of the Chinese carbon budget. First, these data bring into sharper focus the distribution of Chinese CO_2

fluxes (Fig. 1c), increasing the net uptake over the southwest (mainly Yunnan, Guizhou and Guangxi provinces), and increasing net emissions over the northeast (mainly Heilongjiang, Jilin and Liaoning provinces) relative to SR-1 (Fig. 1d). Second, the resulting uncertainties of the Chinese CO_2 budget are also much smaller than when using only the NOAA ObsPack data (Fig. 1f, Supplementary Fig. 6).

Table 1 reports annual CO_2 fluxes for China. Using inversion SR-2, we find that the mean annual Chinese land biosphere sink is −1.11 ± 0.38 Pg C yr^{-1} over 2010–2016. The magnitude of this sink term is 68% larger than the a posteriori values inferred from SR-1 for the same period (−0.66 ± 0.52 Pg C yr^{-1}), with uncertainties reduced by about 27% (Fig. 1f, Table 1; Supplementary Fig. 6). This revised sink term is substantially larger (that is, more uptake) than many recent estimates[3,4,14–18] that cover various periods between 1980 and 2013 and inferred using different methods and data, including the inversion of atmospheric data (Table 1).

Our a posteriori estimate of CO_2 uptake over southwest China (18–30° N, 95–110° E) accounts for about 32% of the uptake over the Chinese mainland. We also find evidence of substantial year-to-year variability of carbon uptake across southwest China (Fig. 2a, b), with the largest values during boreal summer months. As we discuss below, this variation is at least partly due to the young age structure of this forest ecosystem[5,19,20] and variations in meteorology (Supplementary Fig. 13). The relatively smaller uptake over northeast China (compared with SR-1) is due to the elevated sink during the growing season (Fig. 2c, d, Extended Data Fig. 2) that is compensated by emissions during the rest of the year.

Table 1 | Carbon dioxide fluxes over China and East Asia

This study (Pg C yr⁻¹)			Previous studies (Pg C yr⁻¹)			
Year	SR-1	SR-2	Method	Period	Carbon balance	Reference
2010	−0.60 ± 0.51	−0.99 ± 0.38	Inventory-satellite-based or process-based model estimation	1961–2005	−0.179[a]	16
2011	−0.74 ± 0.53	−1.32 ± 0.36		1980-1999	−0.177 ± 0.073[b]	15
2012	−0.74 ± 0.53	−1.25 ± 0.39		1990-2009	−0.224 ± 0.141[b,c]	14
2013	−0.61 ± 0.55	−0.87 ± 0.45		2001–2010	−0.966[b]	18
2014	−0.68 ± 0.49	−1.08 ± 0.39		2006–2009	−0.33[b]	4
2015	−0.52 ± 0.54	−1.03 ± 0.34	Atmospheric inversion	1996–2005	−0.35 ± 0.33[a]	15
2016	−0.77 ± 0.47	−1.25 ± 0.31		1990–2009	−0.270 ± 0.507[b,c]	14
				2001–2010	−0.33[a]	17
				2008–2012	−0.46 (−1.18 to −0.01)[a,c,d]	3
2010–2016 mean	−0.66 ± 0.52	−1.11 ± 0.38		2006–2009	−0.45[a]	4

In this study, inversion SR-1 denotes NOAA ObsPack in situ data (including one site over China); inversion SR-2 denotes in situ data from NOAA ObsPack, Siberia, Hong Kong and from the Chinese mainland. Uncertainties represent one standard deviation from the mean value.
[a]Net CO_2 flux of terrestrial biosphere.
[b]Net CO_2 flux of terrestrial biosphere, including the contribution from the oxidation of reduced carbon.
[c]East Asia, including China, Japan, North and South Korea, and Mongolia.
[d]Ensemble model median and range.

Understanding the CO_2 flux distributions

In the absence of independent downwind data that are sensitive to our a posteriori changes in Chinese CO_2 fluxes, we also infer independent CO_2 flux estimates from satellite column observations of CO_2 from the Japanese Greenhouse gases Observing Satellite (GOSAT)[21] and the NASA Orbiting Carbon Observatory (OCO-2) (see Methods). Using these satellite data, we find that a posteriori CO_2 fluxes during the growing season (May to September) show larger uptake over central, northeast and southern parts of China than SR-1, consistent with the SR-2 inversion (Extended Data Fig. 3). Differences between these flux estimates can be partly explained by seasonal changes in clear-sky views (Extended Data Fig. 4) that are determined by cloud cover and aerosol loading. We find major data gaps over China in winter, where the a priori information plays a larger part in the a posteriori solution (Extended Data Fig. 5).

Using reports from quinquennial Chinese forest inventories[22], we find that, since the 1990s, northeast and southwest China have experienced a period of rapid growth in forest area (Extended Data Fig. 6, Extended Data Table 1), reflecting national policies[23]. These findings are supported by remote-sensing observations of land cover (Extended Data Fig. 7), vegetation indices (Fig. 3a, Extended Data Fig. 8, Supplementary Figs. 11, 12) and above-ground biomass (Fig. 3b) that show small but significant upward trends. A consequence of the rapid increase of afforestation during the past 30 years is that Chinese forests contain a large fraction of young and mid-aged trees (Extended Data Fig. 9), which are associated with a high rate of carbon sequestration[6,20,24,25].

The southwest region, in particular the Guangxi autonomous region, is populated by eucalyptus, which is a fast-growing and high-yielding tree species with high potential biomass carbon sequestration[26]. These eucalyptus plantations are irrigated and fertilized, factors that would need to be considered in the overall environmental value of this carbon
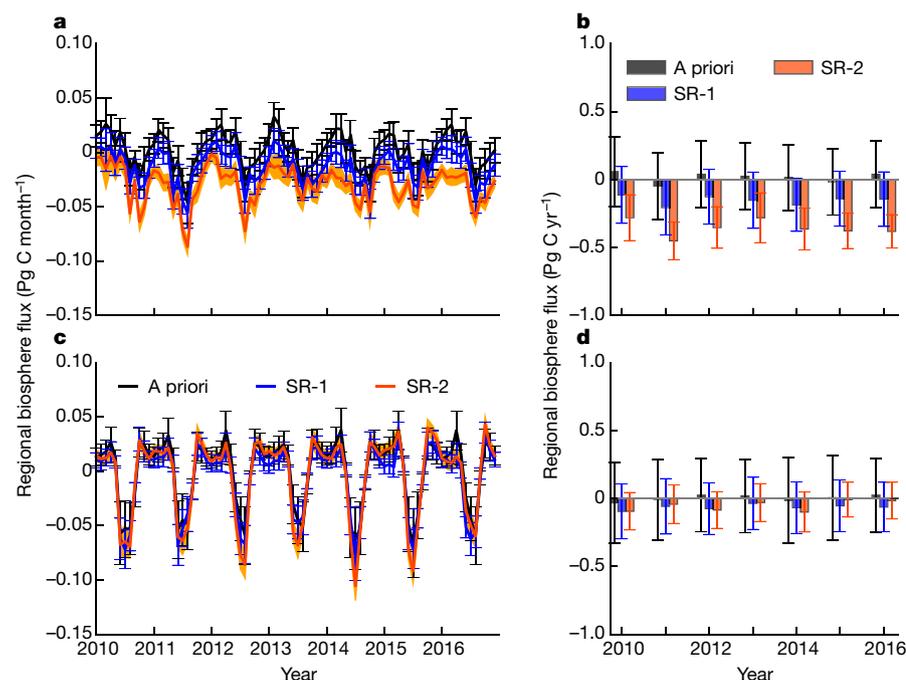

Fig. 2 | Terrestrial biosphere CO_2 fluxes. a, b, Over northeast China (within 38–54° N, 120–135° E); c, d, over southwest China (within 18–30° N, 95–110° E). Monthly CO_2 fluxes (Pg C month⁻¹) from the a priori inventory and a posteriori fluxes inferred from SR-1 and SR-2 are shown from 2010 to 2016 over northeast China (a) and southwest China (c); and corresponding annual fluxes (Pg C yr⁻¹) over northeast China (b) and southwest China (d). Vertical bars and the orange envelope (a and c) denote a priori and a posteriori uncertainties.
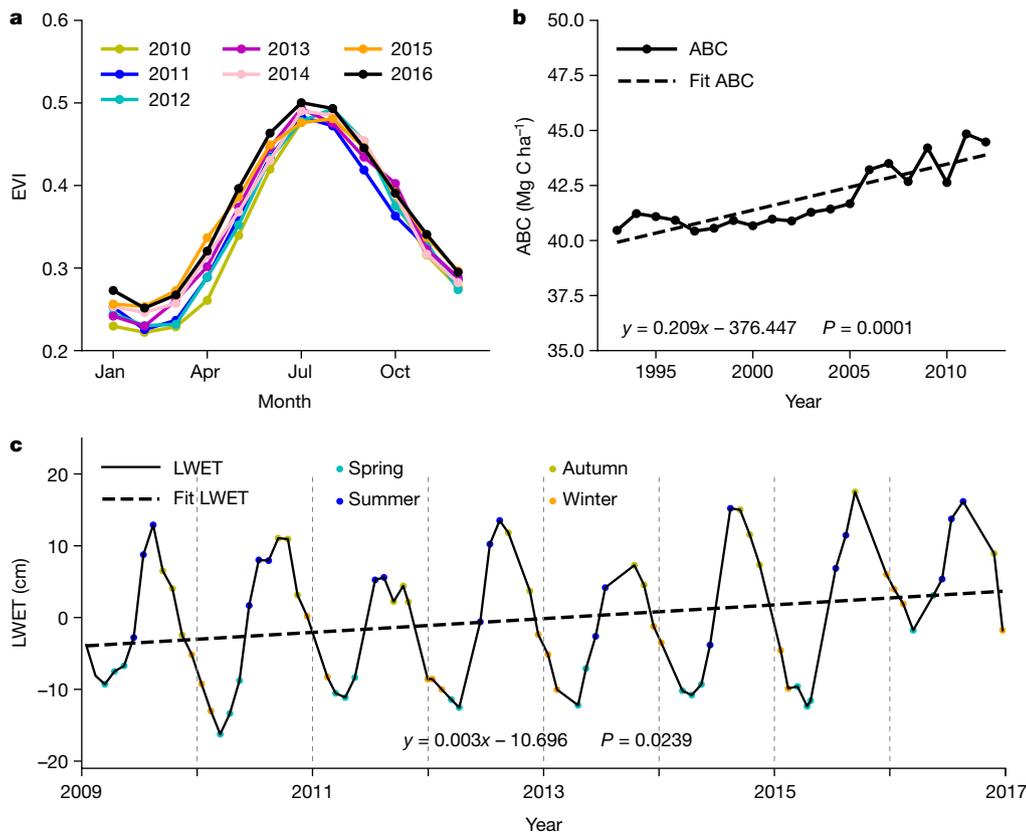
Fig. 3 | Satellite observations relating to vegetation growth. **a**, Leaf phenology: successive years (2010–2016) of superimposed monthly EVI from MODIS. **b**, Above-ground biomass harmonized from several sensors (1993–2012). **c**, Liquid water equivalent thickness (LWET) from the Gravity Recovery and Climate Experiment (GRACE) (2009–2016) over southwest China (18–30° N, 95–110° E). Data record lengths vary according to the sensor lifetime. The linear best-fit lines (dashed lines) and the associated *P* values are shown in **b** and **c**. More detailed data descriptions can be found in Supplementary Information.

sink. Correlative satellite data of leaf phenology and groundwater (Fig. 3, Supplementary Information) clearly show upward trends that together are consistent with a progressively larger biological carbon sink. Because it is a managed ecosystem, the southwest region appears to be insensitive to wider changes in climate (Supplementary Figs. 13, 14), particularly during the 2009–2010 drought associated with reduced precipitation[27].

We have shown that using atmospheric CO_2 measurements collected at a small number of additional sites changes our understanding of the distribution of CO_2 fluxes over China. This reflects the sparse and uneven distribution of ground-based measurements. Our use of satellite observations to verify a posteriori fluxes inferred from the in situ data will enable extensive improvement in CO_2 monitoring.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2849-9.

1. Le Quéré, C. et al. Global carbon budget 2018. *Earth Syst. Sci. Data* **10**, 2141–2194 (2018).
2. Peylin, P. et al. Global atmospheric carbon budget: results from an ensemble of atmospheric CO₂ inversions. *Biogeosciences* **10**, 6699–6720 (2013).
3. Thompson, R. L. et al. Top-down assessment of the Asian carbon budget since the mid 1990s. *Nat. Commun.* **7**, 10724 (2016).
4. Jiang, F. et al. A comprehensive estimate of recent carbon sinks in China using both top-down and bottom-up approaches. *Sci. Rep.* **6**, 22130 (2016).
5. Pan, Y. et al. A large and persistent carbon sink in the world's forests. *Science* **333**, 988–993 (2011).
6. Zhang, Y., Yao, Y., Wang, X., Liu, Y. & Piao, S. Mapping spatial distribution of forest age in China. *Earth Space Sci.* **4**, 108–116 (2017).
7. Williams, R. A. Mitigating biodiversity concerns in eucalyptus plantations located in South China. *J. Biosci. Med.* **3**, 57099 (2015).
8. Barlow, J. M., Palmer, P. I., Bruhwiler, L. M. & Tans, P. Analysis of CO₂ mole fraction data: first evidence of large-scale changes in CO₂ uptake at high northern latitudes. *Atmos. Chem. Phys.* **15**, 13739–13758 (2015).
9. Schimel, D. S. et al. Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature* **414**, 169–172 (2001).
10. Takahashi, T. et al. Climatological mean and decadal change in surface ocean *p*CO₂, and net sea–air CO₂ flux over the global oceans. *Deep Sea Res. II* **56**, 554–577 (2009).
11. Rödenbeck, C. et al. Data-based estimates of the ocean carbon sink variability — first results of the Surface Ocean *p*CO₂ Mapping intercomparison (SOCOM). *Biogeosciences* **12**, 7251–7278 (2015).
12. Gregg, J. S., Andres, R. J. & Marland, G. China: emissions pattern of the world leader in CO₂ emissions from fossil fuel consumption and cement production. *Geophys. Res. Lett.* **35**, L08806 (2008).
13. Liu, Z. et al. Reduced carbon emission estimates from fossil fuel combustion and cement production in China. *Nature* **524**, 335–338 (2015).
14. Piao, S. L. et al. The carbon budget of terrestrial ecosystems in East Asia over the last two decades. *Biogeosciences* **9**, 3571–3586 (2012).
15. Piao, S. et al. The carbon balance of terrestrial ecosystems in China. *Nature* **458**, 1009–1013 (2009).
16. Tian, H. et al. Net exchanges of CO₂, CH₄, and N₂O between China's terrestrial ecosystems and the atmosphere and their contributions to global climate warming. *J. Geophys. Res. Atmos.* **116**, G02011 (2011).
17. Zhang, H. et al. Net terrestrial CO₂ exchange over China during 2001–2010 estimated with an ensemble data assimilation system for atmospheric CO₂. *J. Geophys. Res. Atmos.* **119**, 3500–3515 (2014).
18. Wang, Q. et al. Primary estimation of Chinese terrestrial carbon sequestration during 2001–2010. *Sci. Bull.* **60**, 577–590 (2015).
19. Fang, J. Y., Chen, A. P., Peng, C. H., Zhao, S. Q. & Ci, L. Changes in forest biomass carbon storage in China between 1949 and 1998. *Science* **292**, 2320–2322 (2001).
20. Pan, Y., Luo, T., Birdsey, R., Hom, J. & Melillo, J. New estimates of carbon storage and sequestration in China's forests: effects of age-class and method on inventory-based carbon estimation. *Clim. Change* **67**, 211–236 (2004).
21. Yokota, T. et al. Global concentrations of CO₂ and CH₄ retrieved from GOSAT: first preliminary results. *Sci. Online Lett. Atmos.* **5**, 160–163 (2009).
22. Zeng, W., Tomppo, E., Healey, S. P. & Gadow, K. V. The national forest inventory in China: history — results — international context. *For. Ecosyst.* **2**, 23 (2015).
23. Liu, J., Li, S., Ouyang, Z., Tam, C. & Chen, X. Ecological and socioeconomic effects of China's policies for ecosystem services. *Proc. Natl Acad. Sci. USA* **105**, 9477–9482 (2008).
24. Wang, S. et al. Relationships between net primary productivity and stand age for several forest types and their influence on China's carbon balance. *J. Environ. Manage.* **92**, 1651–1662 (2011).
25. Yu, Y. et al. Influence of site index on the relationship between forest net primary productivity and stand age. *PLoS One* **12**, e0177084 (2017).
26. Du, H. et al. Carbon storage in a eucalyptus plantation chronosequence in Southern China. *Forests* **6**, 1763–1778 (2015).
27. Xu, K., Yang, D., Xu, X. & Lei, H. Copula based drought frequency analysis considering the spatio-temporal variability in Southwest China. *J. Hydrol.* **527**, 630–640 (2015).

Article

Methods

Data

In situ CO_2 mole fraction observations. We use (weekly) discrete flask air samples from 105 sites and (hourly) continuous observations from 52 sites that are part of the global atmospheric surface CO_2 observations network. These are currently described by the Observation Package (ObsPack) obspack_co2_1_GLOBALVIEWplus_v4.1_2018-10-29 data product[28] from NOAA's ESRL. We also use an additional Chinese data from WDCGG at Hok Tsui (HKG) provided by Hong Kong Observatory.

We use CO_2 mole fraction data from nine towers collected by the Japan–Russia Siberia Tall Tower Inland Observation Network (JR-STATION; Supplementary Table 1, Extended Data Fig. 1) over Siberia from 2009–2016[29,30]. We also use weekly flask measurements and hourly continuous CO_2 mole fraction measurements during 2009–2016 of six new regional background sites over China (Longfengshan, Shangdianzi, Linan, Shangri-La, Jinsha and Akedala) (Supplementary Table 2, Extended Data Fig. 1) are operated by the China Meteorological Administration[31,32]. Further details about these CO_2 observations are in Supplementary Information.

GOSAT observations. We use column CO_2 data collected by GOSAT, launched in January 2009 in a Sun-synchronous orbit with a local overpass time over China around 13:00[33]. We use total column CO_2 (XCO_2) full-physics retrievals from GOSAT: v7.3 data product from the Atmospheric CO_2 Observations from Space (ACOS) from the Jet Propulsion Laboratory[34].

OCO-2 observations. We use column CO_2 data collected by NASA's OCO-2, launched in July 2014 in a sun-synchronous orbit with a local crossing time over China around 13:30[35]. We use bias-corrected OCO-2 V8r data[36] in this study. We use only nadir observations over land with a warning level <12. After considering observation error correlations, we have thinned the data with a frequency of 0.1 Hz, corresponding to an approximate distance of 75 km between successive measurements.

NDVI and EVI represent composite properties of leaf area, chlorophyll and canopy structure. We use MOD13C2 (MODIS/Terra Vegetation Indices Monthly L3 Global 0.05° CMG V006)[37] to get NDVI AND EVI information. The data are only retained with pixel reliability values masked as good data ('0') or marginal data ('1').

GRACE provides information about changes in the water column[38–40]. Rooting depths of Chinese terrestrial ecosystems are likely to be sufficiently deep that we cannot establish a direct and immediate relationship between vegetation and changes in precipitation. Changes in gravity due to changes in water column depth provide a much stronger relationship with vegetation access to water. We use the surface mass change data based on the RL05 spherical harmonics from the Center for Space Research at University of Texas, Austin (CSR), Jet Propulsion Laboratory (JPL) and Geoforschungs Zentrum Potsdam (GFZ). The three different processing groups chose different parameters and solution strategies when deriving month-to-month gravity field variations from GRACE observations. We use the ensemble mean of the three data fields and multiply the data by the scaling grid provided.

Above-ground biomass carbon. We use above-ground biomass carbon (ABC) estimates (1993–2012) based on harmonized vegetative optical depth data for 1993 onwards derived from a series of passive microwave satellite sensors[41]. These sensors include the Special Sensor Microwave Imager (SSM/I), Advanced Microwave Scanning Radiometer for Earth Observation System (AMSR-E), FengYun-3B Microwave Radiometer Imager (MWRI) and Windsat.

Models

To describe the relationship between surface fluxes of CO_2 and atmospheric CO_2, we use the GEOS-Chem global 3-D chemistry transport model (v9.02). For our experiments we run GEOS-Chem with a horizontal resolution of 4° (latitude) × 5° (longitude), driven by the GEOS-5 (GEOS-FP from 2013) meteorological analyses from the Global Modelling and Assimilation Office Global Circulation Model based at NASA Goddard Space Flight Centre.

To describe atmospheric CO_2, we use well established emission inventories as our a priori flux estimates: (1) weekly biomass burning emissions[42]; (2) monthly anthropogenic emissions[43–45] (we use 88% of ODIAC anthropogenic[43,44] inventory in this study, details in Supplementary Information); (3) monthly climatological ocean fluxes[10]; and (4) three-hourly terrestrial biosphere fluxes[46].

We use an ensemble Kalman Filter (EnKF) framework[47–49] to infer CO_2 fluxes from the in situ or space-based measurements of atmospheric CO_2 concentrations. Surface flux $f^a(x, t)$ (in kg m^{-2} s^{-1}) at location x and time t is given by[47,49]:

$$f^a(x, t) = f^0(x, t) + \sum_i^N c_i BF_i(x, t),$$

where $f^0(x, t)$ describes the a priori estimates (in kg m^{-2} s^{-1}). The basis function set $BF_i(x, t)$ is defined as the pulse-like (monthly) CO_2 fluxes over pre-defined geographical regions (in kg m^{-2} s^{-1}). For each individual basis function, we estimate the unitless scaling coefficient c_i by optimally fitting the model to observation data. The coefficients for the basis function set represent the state vector $\mathbf{c}$.

We use a total of 792 monthly basis functions split between 317 oceanic regions and 475 land regions. These regions are based on established 22 regions in TransCom-3[50]. To infer Chinese CO_2 fluxes on a spatial resolution of 4° (latitude) × 5° (longitude) we further divide the Eurasia temperate region into 78 sub-regions.

We solve for the a posteriori state vector using the Kalman gain matrix K:

$$\mathbf{c}_a = \mathbf{c}_0 + K\left(\mathbf{y}_{obs} - H(\mathbf{c}_0)\right),$$

where $\mathbf{c}_a$ and $\mathbf{c}_0$ represent the a priori and a posteriori estimates, respectively; $\mathbf{y}_{obs}$ denotes the atmospheric CO_2 observations; and H describes the observation operator that relates surface fluxes to the observation data. In our case, H represents the GEOS-Chem model sampled at measurement locations, and in the case of satellite data also taking into account instrument sensitivity to CO_2.

We assume a 50% uncertainty for monthly land terrestrial fluxes, and 40% for monthly ocean fluxes[49]. We assume land (ocean) a priori fluxes are correlated with a correlation length of 500 (800) km.

We determine the terrestrial biosphere flux by subtracting the fossil fuel and cement production emission estimate (FF). This is a common approach[2–4,51], based on an assumption about our knowledge of FF flux. Consequently, our biosphere fluxes depend to a certain degree on what we assume for anthropogenic emissions (Supplementary Table 3).

In addition to our control run (SR-2), we also report Chinese land biosphere fluxes that correspond to the ODIAC anthropogenic inventory[43,44]. We find that our total net a posteriori emissions of CO_2 over China (1.35 ± 0.38 Pg C yr^{-1}) are not significantly different to our control inversion (1.37 ± 0.38 Pg C yr^{-1}), suggesting that our a posteriori fluxes are robust. However, our a posteriori land biosphere fluxes represent a larger sink of CO_2 (-1.47 ± 0.38 Pg C yr^{-1}) compared to our control flux estimates (-1.11 ± 0.38 Pg C yr^{-1}). Our larger a posteriori sink estimate is expected since we are subtracting a larger FF flux from the net emission estimate. This calculation also provides a crude measure of the land biosphere flux uncertainty due to uncertainties in FF fluxes.

In the EnKF framework, the Kalman gain matrix K is approximated by[46]:

$$K \approx \Delta C \Delta Y^T (\Delta Y \Delta Y^T + R)^{-1},$$

where R is the observation error covariance matrix. ΔC represents the ensemble of perturbations to the state vector that are used to construct

the a priori error covariance, $P \approx \Delta C \Delta C^T$ where T denotes the matrix transpose[47]. To compare with observation, ΔC is projected onto observation space using the observation operator $\Delta Y = H(\Delta C)$. We assume no observation error correlations but include an additional 1.0 parts per million uncertainty to the reported observation errors to account for model transport errors.

The methodology we use to estimate annual uncertainty estimates from the ensemble Kalman filter with a lagged window is described in detail in Supplementary Information.

Data availability

GRACE data are available from http://grace.jpl.nasa.gov. ABC data are available from https://www.wenfo.org/wald/global-biomass. The NDVI, EVI and LAI data were retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center, USGS/Earth Resources Observation and Science Center (https://lpdaac.usgs.gov/data_access/data_pool). CO_2 mole fraction data from the Chinese sites used in this study are available at https://doi.org/10.17632/w3bwmr6rfg.1 on http://data.mendeley.com. Requests for further information about these data should be directed to S.F. Source data are provided with this paper.

Code availability

We acknowledge the Python Software Foundation: Python Language Reference, version 3.7.7; available at http://www.python.org. We also acknowledge Matplotlib (v3.1.3, 10.5281/zenodo.3984190)[52]. The community-led GEOS-Chem model of atmospheric chemistry and transport is maintained centrally by Harvard University (http://acmg.seas.harvard.edu/geos/) and is available on request. The ensemble Kalman filter code is publicly available as PyOSSE (https://www.nceo.ac.uk/data-tools/atmospheric-tools/).

28. Cooperative Global Atmospheric Data Integration Project. Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957–2017; obspack_co2_1_GLOBALVIEWplus_v4.1_2018_10_29 (NOAA Earth System Research Laboratory, Global Monitoring Division, 2018); https://doi.org/10.25925/20181026

29. Sasakawa, M. et al. Continuous measurements of methane from a tower network over Siberia. *Tellus B* **62**, 403–416 (2010).

30. Saeki, T. et al. Carbon flux estimation for Siberia by inverse modeling constrained by aircraft and tower CO_2 measurements. *J. Geophys. Res. Atmos.* **118**, 1100–1122 (2013).

31. Fang, S. X. et al. In situ measurement of atmospheric CO_2 at the four WMO/GAW stations in China. *Atmos. Chem. Phys.* **14**, 2541–2554 (2014).

32. Fang, S. X., Tans, P. P., Steinbacher, M., Zhou, L. X. & Luan, T. Comparison of the regional CO_2 mole fraction filtering approaches at a WMO/GAW regional station in China. *Atmos. Meas. Tech.* **8**, 5301–5313 (2015).

33. Kuze, A., Suto, H., Nakajima, M. & Hamazaki, T. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.* **48**, 6716–6733 (2009).

34. OCO-2 Science Team. *ACOS GOSAT/TANSO-FTS ACOS Level 2 Bias-Corrected XCO2 and Other Select Fields from the Full-Physics Retrieval Aggregated as Daily Files V7.3* (GES DISC, 2016); https://disc.gsfc.nasa.gov/datacollection/ACOS_L2_Lite_FP_7.3.html (accessed October 2017).

35. Eldering, A. et al. The Orbiting Carbon Observatory-2 early science investigations of regional carbon dioxide fluxes. *Science* **358**, eaam5745 (2017).

36. OCO-2 Science Team. *OCO-2 Level 2 Bias-Corrected XCO2 and Other Select Fields from the Full-Physics Retrieval Aggregated as daily files. Retrospective Processing V8r* (GES DISC, 2017); https://doi.org/10.5067/W8QGIYNKS3JC

37. Didan, K. *MOD13C2 MODIS/Terra Vegetation Indices Monthly L3 Global 0.05 Deg CMG V006* (NASA EOSDIS Land Processes DAAC, 2015); https://doi.org/10.5067/modis/mod13c2.006

38. Swenson, S. & Wahr, J. Post-processing removal of correlated errors in GRACE data. *Geophys. Res. Lett.* **33**, L08402 (2006).

39. Swenson, S. C. *GRACE Monthly Land Water Mass Grids NETCDF Release 5.0 V5.0* (PODAAC, 2012); https://doi.org/10.5067/TELND-NC005 (accessed 27 September 2017).

40. Landerer, F. W. & Swenson, S. C. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resour. Res.* **48**, W04531 (2012).

41. Liu, Y. Y. et al. Recent reversal in loss of global terrestrial biomass. *Nat. Clim. Change* **5**, 470–474 (2015).

42. van der Werf, G. R. et al. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys.* **10**, 11707–11735 (2010).

43. Oda, T. & Maksyutov, S. A very-high-resolution (1 km × 1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmos. Chem. Phys.* **11**, 543–556 (2011).

44. Oda, T. & Maksyutov, S. *ODIAC Fossil Fuel CO_2 Emissions Dataset (ODIAC2016)* (Center for Global Environmental Research, National Institute for Environmental Studies, 2015); https://doi.org/10.17595/20170411.001 (accessed 1 June 2017).

45. Shan, Y. et al. China CO_2 emission accounts 1997–2015. *Sci. Data* **5**, 170201 (2018).

46. Olsen, S. C. & Randerson, J. T. Differences between surface and column atmospheric CO_2 and implications for carbon cycle research. *J. Geophys. Res. Atmos.* **109**, D02301 (2004).

47. Feng, L., Palmer, P. I., Bösch, H. & Dance, S. Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman filter. *Atmos. Chem. Phys.* **9**, 2619–2633 (2009).

48. Feng, L. et al. Evaluating a 3-D transport model of atmospheric CO_2 using ground-based, aircraft, and space-borne data. *Atmos. Chem. Phys.* **11**, 2789–2803 (2011).

49. Feng, L. et al. Estimates of European uptake of CO_2 inferred from GOSAT XCO2 retrievals: sensitivity to measurement bias inside and outside Europe. *Atmos. Chem. Phys.* **16**, 1289–1302 (2016).

50. Gurney, K. R. et al. Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models. *Nature* **415**, 626–630 (2002).

51. Saeki, T. & Patra, P. K. Implications of overestimated anthropogenic CO_2 emissions on East Asian and global land CO_2 flux inversion. *Geosci. Lett.* **4**, 9 (2017).

52. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

**Extended Data Fig. 1 | In situ atmospheric CO_2 measurements across East Asia. a**, Map showing the location of atmospheric CO_2 measurement sites over East Asia used in our numerical experiments. Coloured dots represent individual measurement sites referred to by letter codes (Supplementary Tables 1, 2). Blue dots represent NOAA ObsPack stations, the cyan dot represents the HKG station from WDCGG, brown dots represent Siberia stations and orange dots represent CMA regional background stations. **b–g**, CO_2 mole fraction observations (in parts per million, ppm) at the six sites across China from 2009 to 2017 used in this study. Discrete (weekly) flask air samples, denoted by blue dots, and continuous (hourly) observations, denoted by black dots, are collected and analysed by the China Meteorological Administration. CRDS, cavity ring-down spectrometer.
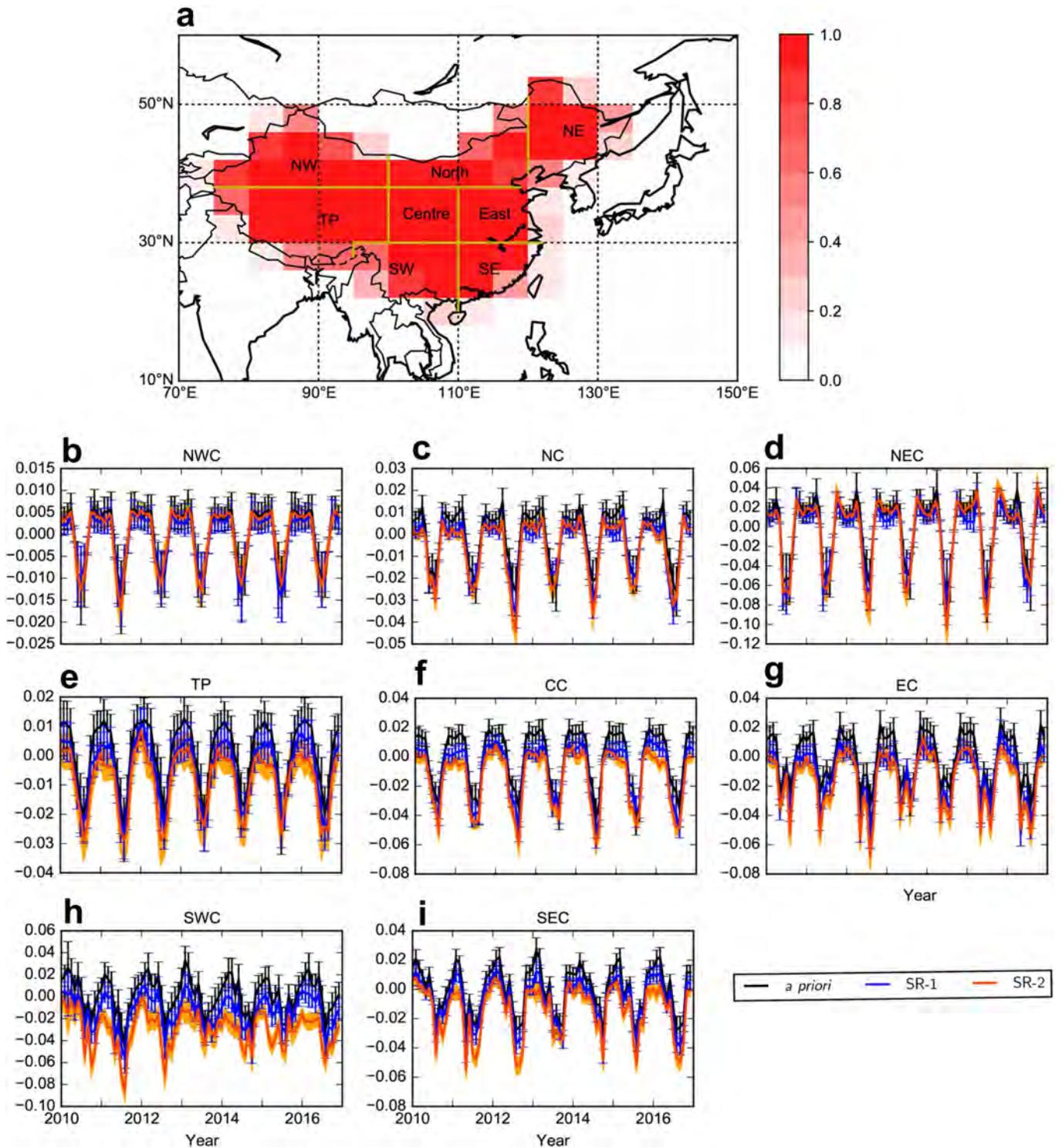
Extended Data Fig. 2 | Monthly subcontinental Chinese CO_2 flux estimates.
a, Map showing the subcontinental geographical regions over China where we report a posteriori CO_2 fluxes. Colours denote the proportion of each region that falls within mainland China. **b–i,** Monthly regional a priori and a posteriori biosphere CO_2 fluxes (Pg C month^{-1}) over China during 2010–2016. A posteriori fluxes are inferred from data used in the SR-1 and SR-2 inversions (see main text). Vertical bars and the orange envelope denote a priori and a posteriori uncertainties. NWC, northwest China; NC, northern China; NE, northeast China; TP, Tibetan Plateau; CC, central China; EC, eastern China; SW, southwest China; SE, southeast China, as shown in **a.**
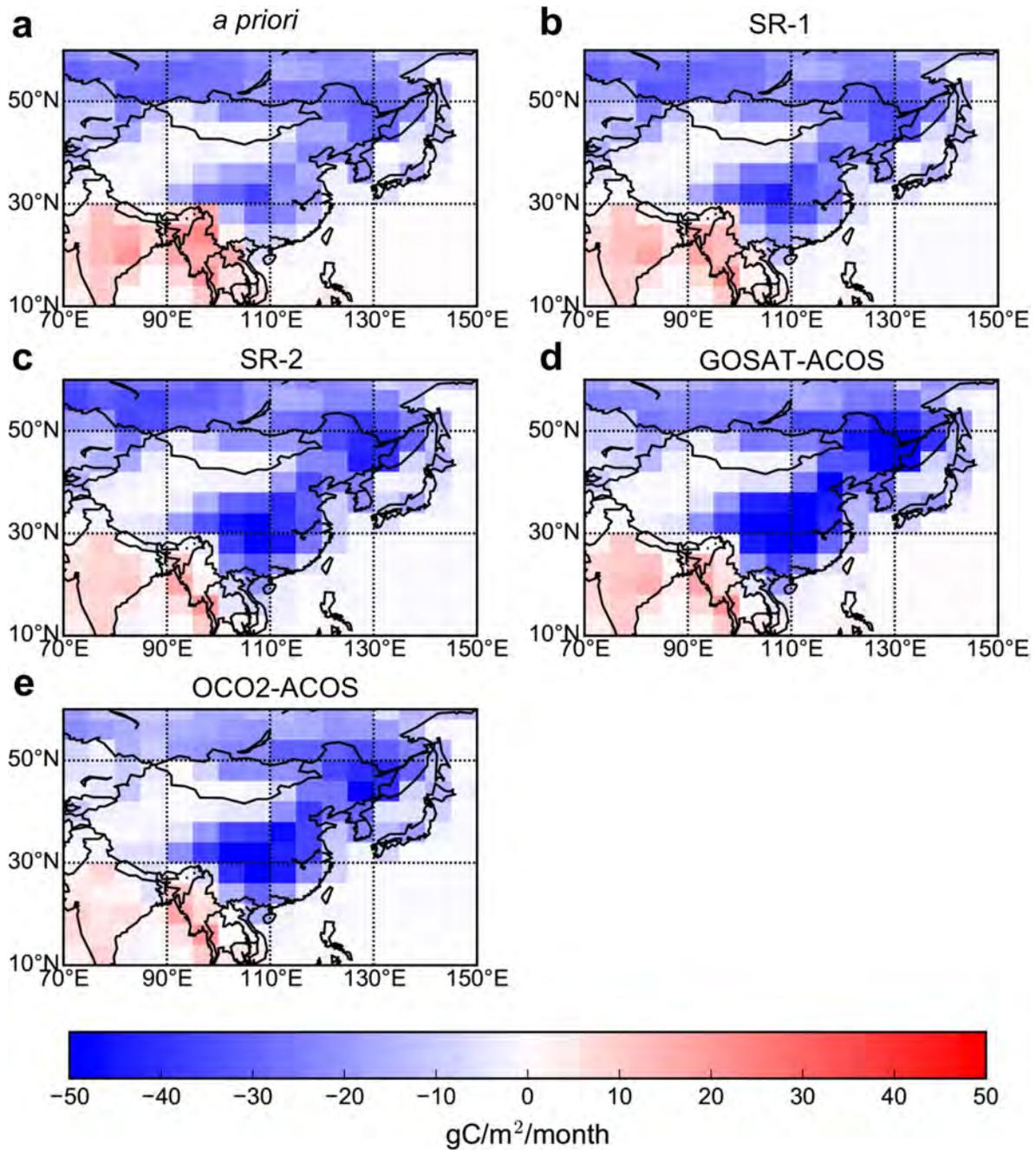
Extended Data Fig. 3 | Mean spatial distribution of a priori and a posteriori land biosphere CO_2 fluxes from May to September, inferred from in situ and satellite observations of CO_2. a, Our a priori fluxes. **b**, **c**, The a posteriori fluxes corresponding to inversions SR-1 and SR-2 (see main text) that use in situ data. **d**, **e**, A posteriori fluxes inferred from column observations of CO_2 from GOSAT and from NASA's OCO-2, respectively. Flux estimates reported represent a temporal mean from 2010 to 2015, except for **e**, which is only for 2015.
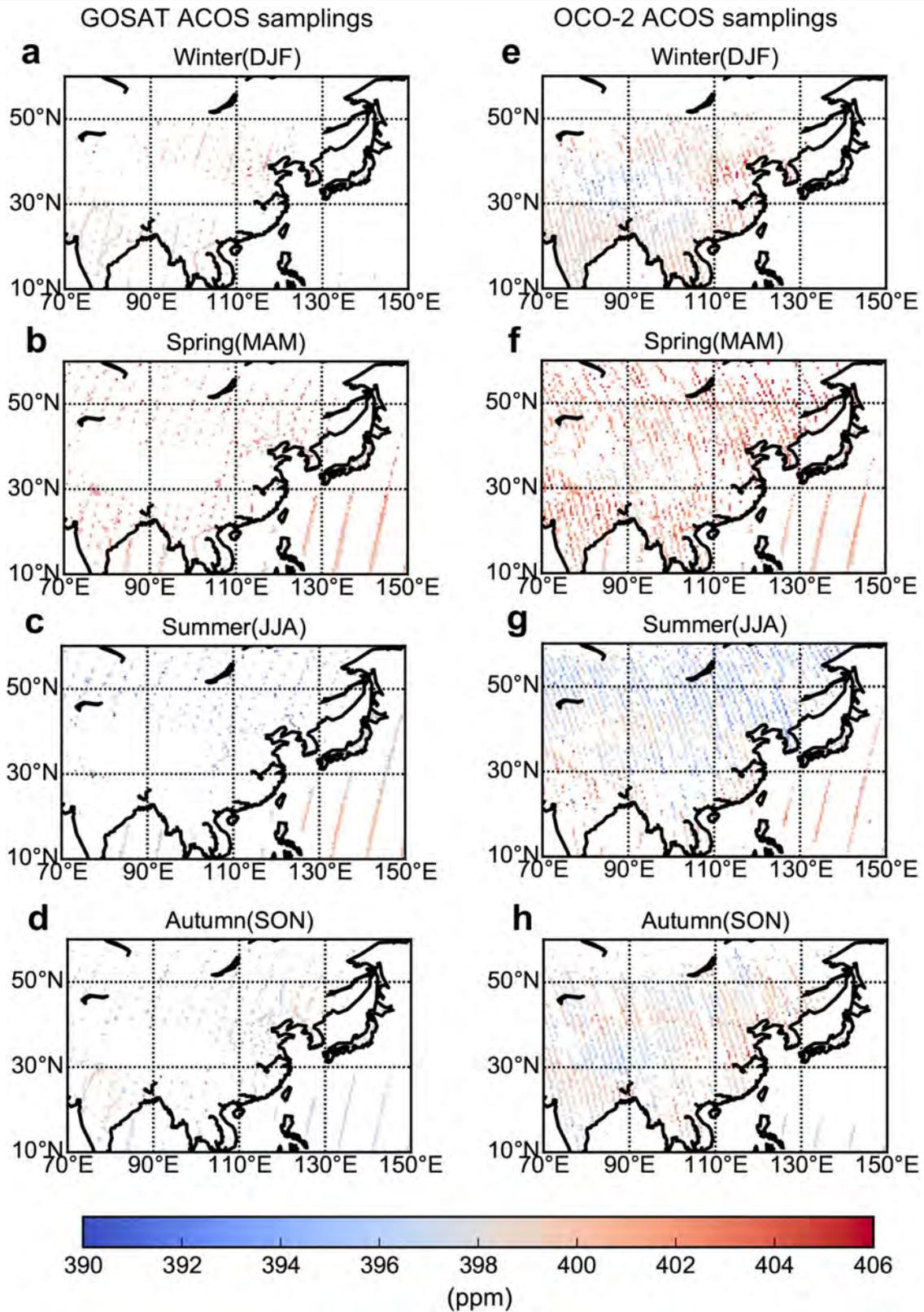
**Extended Data Fig. 4 | Seasonal distribution and magnitude of satellite retrievals of column CO₂ from December 2014 to November 2015. a–d,** Data from GOSAT (v7.3 ACOS). **e–h,** Data from OCO-2 (v8r). See Methods section 'Data' for further information.
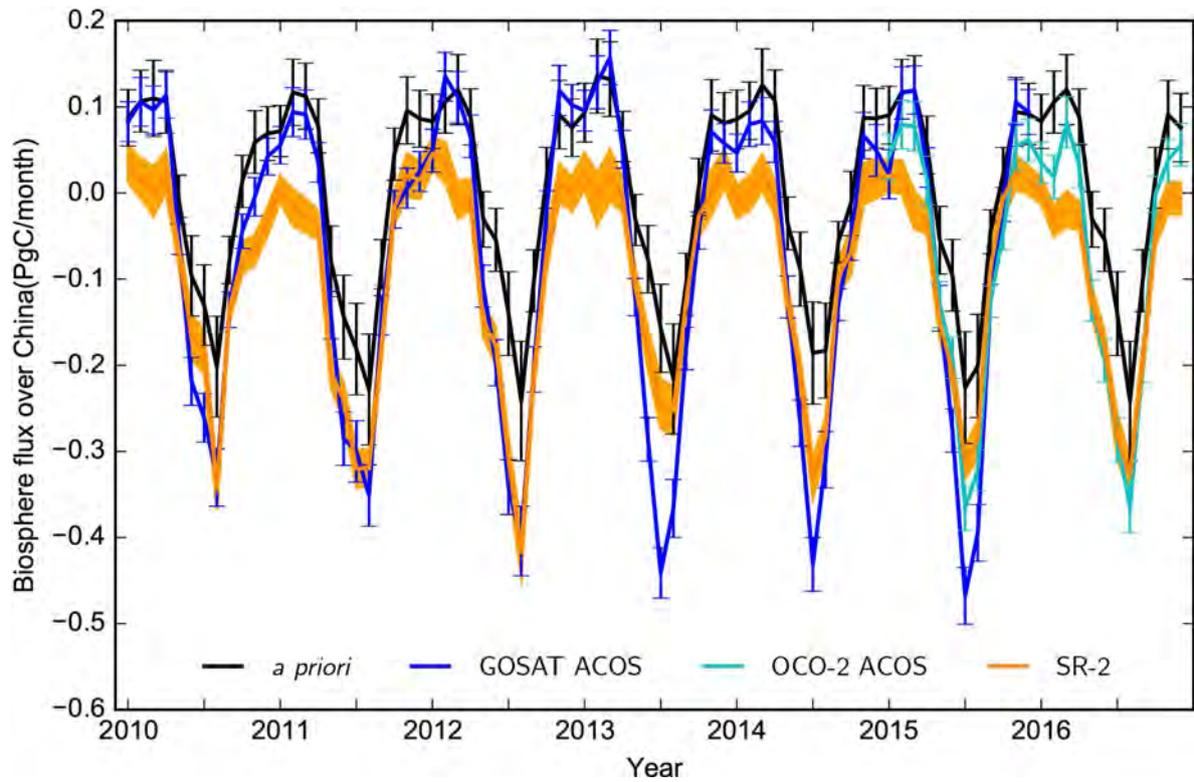
Extended Data Fig. 5 | A priori and a posteriori monthly biosphere CO$_2$ fluxes over China inferred from the GOSAT CO$_2$ column data and in situ data from 2010 to 2016. We also report a posteriori fluxes inferred from OCO-2 CO$_2$ column from September 2014 to December 2016. Vertical bars and shaded envelopes denote a posteriori uncertainties. Higher annual fluxes inferred from GOSAT and OCO-2 are due mainly to higher a posteriori fluxes during winter months when data coverage is sparse (Extended Data Fig. 4) and the fluxes are more influenced by a priori values.
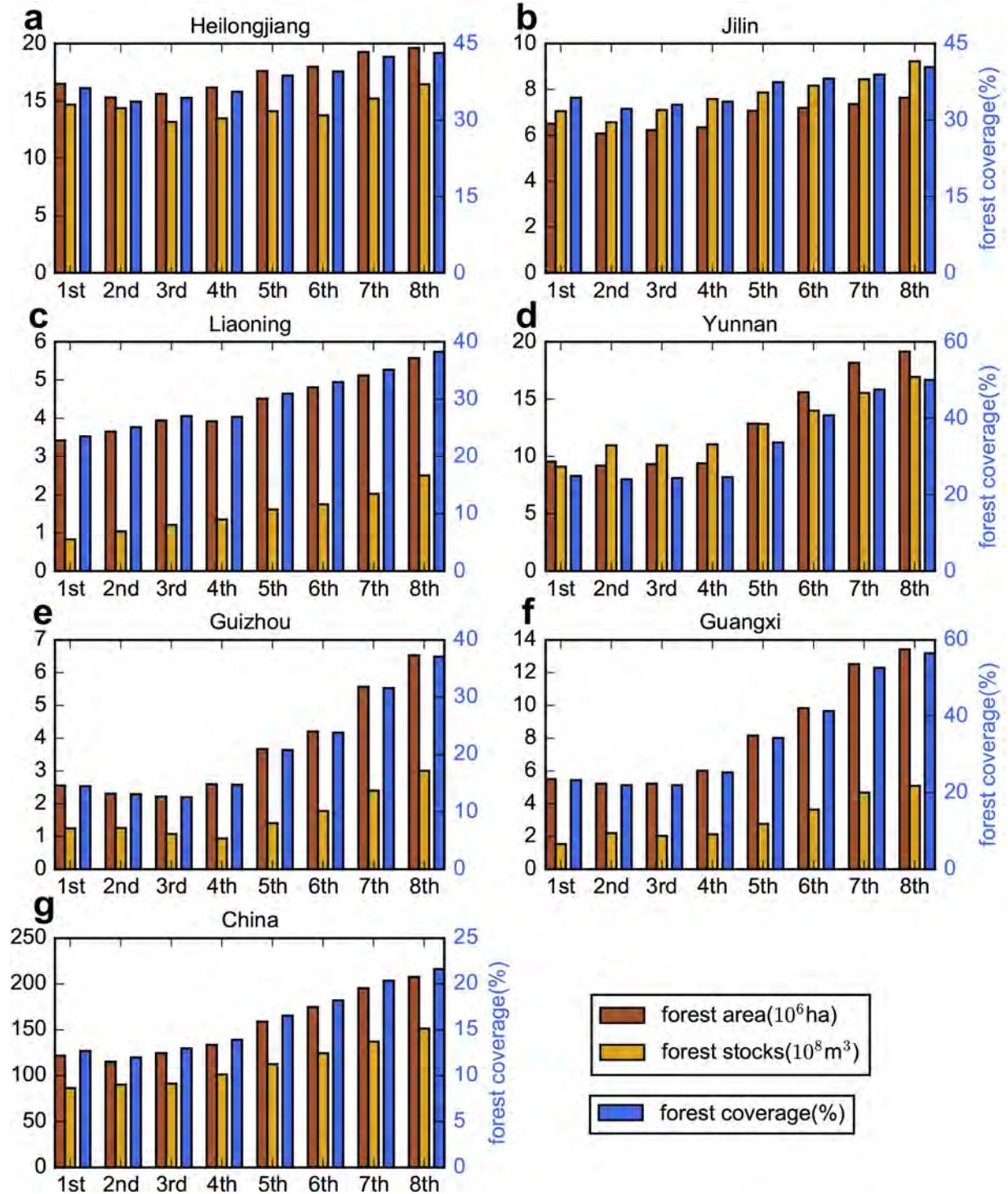
Extended Data Fig. 6 | Changes in forest area, forest stocks and forest coverage over six key forested Chinese provinces and over the whole of China, 1973–2013. a–f, Values for Heilongjiang, Jilin, Liaoning, Yunnan, Guizhou and Guangxi provinces. **g,** Values for China. The *x*-axis labels refer to the National Forest Inventory of China's State Forestry Administration: 1st (1973–1976), 2nd (1977–1981), 3rd (1984–1988), 4th (1989–1993), 5th (1994–1998), 6th (1999–2003), 7th (2004–2008) and 8th (2009–2013).
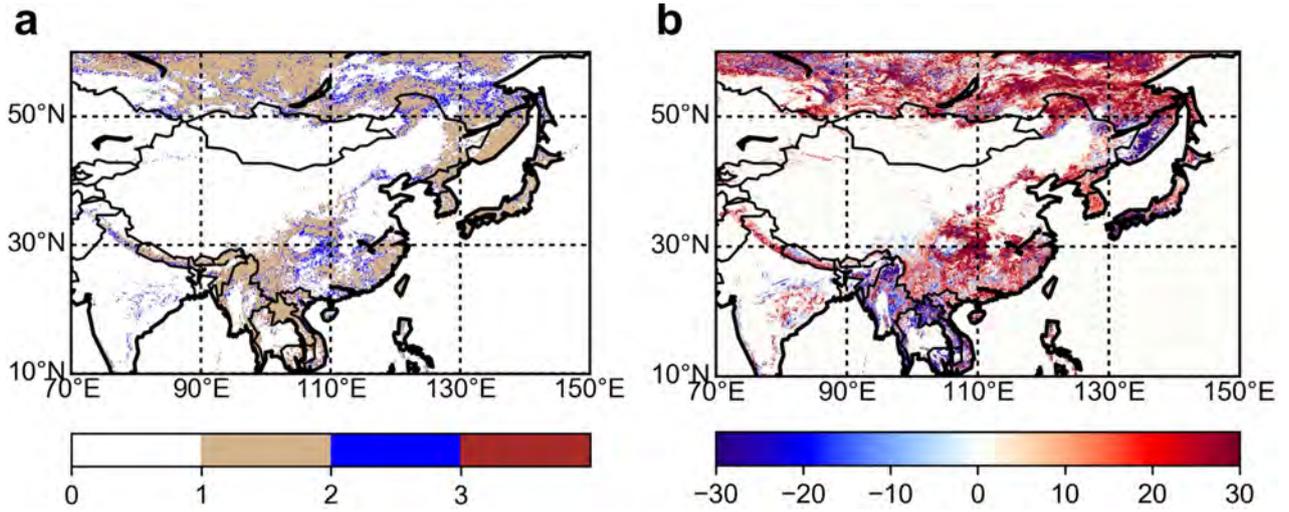
Extended Data Fig. 7 | Forest cover change over China during the period 2002–2012. a, How land cover changed from 2002 to 2012: 0 (1) denotes non-forest (forest) in both 2002 and 2012, 2 denotes conversion from non-forest to forest between 2002 and 2012, and 3 denotes conversion from forest to non-forest between 2002 and 2012. **b**, The forest percentage change per grid box from 2002 to 2012. Data are presented on a 0.05° × 0.05° spatial grid.
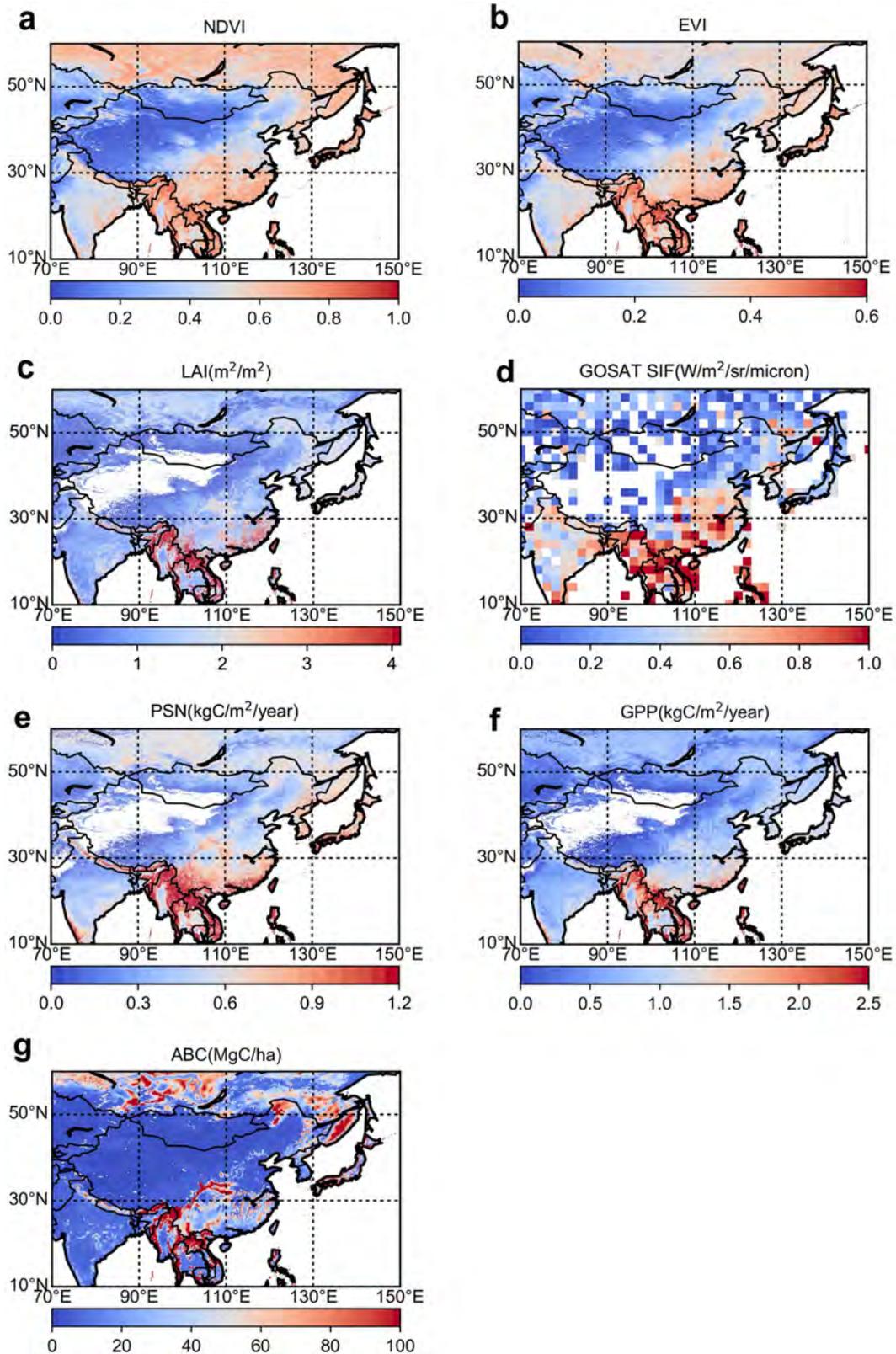
Extended Data Fig. 8 | Multi-year mean of satellite observations of vegetation indices. a, NDVI. **b**, EVI. **c**, LAI. **d**, Solar induced fluorescence (SIF). **e**, Net photosynthesis (PSN). **f**, Gross primary production (GPP). **g**, Above-ground biomass carbon (ABC). See Supplementary Information for details. The mean is over 2010–2012, inclusive, for ABC, and over 2010–2016 for the other data.

a Heilongjiang
5.4%
11.7%
18.5%
41.4%
23.0%

b Jilin
5.2%
18.0%
25.5%
26.6%
24.7%

c Liaoning
2.2% 10.1%
14.4%
26.5%
46.9%

d Yunnan
7.6%
12.6%
16.6%
28.3%
34.9%

e Guizhou
0.8% 4.8%
11.3%
33.2%
50.0%

f Guangxi
1.3% 6.2%
7.8%
39.4%
45.3%

g China
6.3%
13.1%
15.5%
32.0%
33.1%

h Heilongjiang
9.7% 7.7%
16.7%
24.8%
41.1%

i Jilin
8.0% 9.5%
27.6%
31.5%
23.5%

j Liaoning
3.7%
20.2%
22.3%
31.4%
22.4%

k Yunnan
16.4% 20.0%
20.8%
19.2%
23.5%

l Guizhou
1.8%
11.9%
19.9%
42.9%
23.5%

m Guangxi
1.8%
11.8%
15.5%
49.6%
21.3%

n China
11.0%
16.5%
24.1%
20.5%
27.8%

Legend:
- Young forest
- Mid-age forest
- Near-mature forest
- Mature forest
- Overmature forest

Extended Data Fig. 9 | Forest area and stock of five forest-stand age groups over six key forested provinces and over the whole of China. a–f, Forest area for Heilongjiang, Jilin, Liaoning, Yunnan, Guizhou and Guangxi provinces. **g**, Forest area for China. **h–m**, Forest stock for Heilongjiang, Jilin, Liaoning, Yunnan, Guizhou and Guangxi provinces. **n**, Forest stock for China. Forest stands are divided into five age groups: young, mid-aged, near-mature, mature and overmature. Data are taken from the 8th NFI.

Extended Data Table 1 | Summary statistics calculated from the 6th to 8th (and 9th where available) National Forest Inventory of China's State Forestry Administration for China and six Chinese Provinces

Period	Forest land (10^6 ha)	Forest stock (10^6 m³)	Plantation forest area (10^6 ha)	Plantation forest stock (10^6 m³)	Fraction of planted forest area (%)	Fraction of planted forest stock (%)	Forest coverage (%)
China							
6th (1999-2003)	171.1	12455.9	53.3	1504.5	31.1	12.1	18.2
7th (2004-2008)	183.5	13720.8	61.7	1960.5	34.0	14.7	20.4
8th (2009-2013)	193.3	15137.3	69.3	2483.3	36.3	16.8	21.6
Guangxi							
6th (2000)	9.8	364.8	4.5	132.1	45.8	36.2	41.4
7th (2005)	10.3	468.8	5.2	171.3	49.9	36.5	52.7
8th (2010)	11.2	509.4	6.4	222.7	56.8	43.7	56.5
9th (2015)	14.3	677.5	7.3	345.2	51.3	50.9	60.2
Yunnan							
6th (2002)	15.0	1399.3	2.5	52.0	16.8	3.7	40.8
7th (2007)	16.5	1553.8	3.3	72.6	19.8	4.7	47.5
8th (2012)	17.5	1693.1	4.1	110.2	23.7	6.5	50.0
9th (2016)	21.2	1894.9	5.3	262.1	24.8	13.8	59.3
Guizhou							
6th (2000)	4.2	178.0	1.8	54.5	43.7	30.6	23.8
7th (2005)	4.6	240.1	2.0	87.2	43.5	36.3	31.6
8th (2010)	5.4	300.8	2.4	115.6	44.2	38.4	37.1
9th (2015)	7.7	391.8	3.2	165.9	40.9	42.3	43.8
Heilongjiang							
6th (2000)	17.98	1375.02	1.73	94.50	9.60	6.87	39.54
7th (2005)	19.27	1521.05	2.36	135.20	12.23	8.89	42.39
8th (2010)	19.62	1644.87	2.47	164.24	12.56	9.98	43.16
Jilin							
6th (1999)	7.19	816.455	1.48	90.63	20.60	11.10	38.13
7th (2004)	7.36	844.12	1.49	95.95	20.24	11.37	38.93
8th (2009)	7.63	922.57	1.61	103.97	21.04	11.27	40.38
Liaoning							
6th (2000)	4.64	174.77	2.68	63.10	57.66	36.11	32.97
7th (2005)	4.84	202.27	2.83	72.99	58.53	36.09	35.13
8th (2010)	5.17	250.46	3.07	94.88	59.37	37.88	38.24

The period defined for each China-wide census reflects the range of years of the individual provincial censuses.