LETTER

doi:10.1038/nature23021

Global patterns of drought recovery

Christopher R. Schwalm[1,2], William R. L. Anderegg[3], Anna M. Michalak[4], Joshua B. Fisher[5], Franco Biondi[6], George Koch[2], Marcy Litvak[7], Kiona Ogle[8], John D. Shaw[9], Adam Wolf[10], Deborah N. Huntzinger[11], Kevin Schaefer[12], Robert Cook[13], Yaxing Wei[13], Yuanyuan Fang[4], Daniel Hayes[14], Maoyi Huang[15], Atul Jain[16] & Hanqin Tian[17]

Drought, a recurring phenomenon with major impacts on both human and natural systems[1–3], is the most widespread climatic extreme that negatively affects the land carbon sink[2,4]. Although twentieth-century trends in drought regimes are ambiguous[5–7], across many regions more frequent and severe droughts are expected in the twenty-first century[3,7–9]. Recovery time—how long an ecosystem requires to revert to its pre-drought functional state—is a critical metric of drought impact. Yet the factors influencing drought recovery and its spatiotemporal patterns at the global scale are largely unknown. Here we analyse three independent datasets of gross primary productivity and show that, across diverse ecosystems, drought recovery times are strongly associated with climate and carbon cycle dynamics, with biodiversity and CO_2 fertilization as secondary factors. Our analysis also provides two key insights into the spatiotemporal patterns of drought recovery time: first, that recovery is longest in the tropics and high northern latitudes (both vulnerable areas of Earth's climate system[10]) and second, that drought impacts[11] (assessed using the area of ecosystems actively recovering and time to recovery) have increased over the twentieth century. If droughts become more frequent, as expected, the time between droughts may become shorter than drought recovery time, leading to permanently damaged ecosystems and widespread degradation of the land carbon sink.

The present understanding of drought recovery has generally focused on precipitation that ends a drought by alleviating water deficit[12,13], as opposed to restoring function in plants[14,15], and recovery is often assumed to be rapid and complete once hydrological drought is ameliorated[16]. Yet the time to recovery of plant function is critical for ecosystem function, because if a new drought arrives before full recovery, an ecosystem may transition to a new state[17]. Recovery time is thus a critical metric for assessing ecosystem resilience, because its duration shapes the odds of crossing a "tipping point"[10]. In the context of drought, tipping point thresholds are associated with widespread tree mortality—and concomitant degradation in land carbon uptake[18]—as well as social disruption of human systems in water-stressed regions[19].

Here we quantify post-drought recovery time of gross primary productivity (GPP) at grid (0.5° spatial resolution) to global scales using three independent datasets: (1) satellite data from the Moderate Resolution Imaging Spectroradiometer (MODIS)[20]; (2) upscaled FLUXNET[21] observations; and (3) an ensemble of observation-driven land surface models from the Multi-scale synthesis and Terrestrial Model Intercomparison Project (MsTMIP)[22]. We focus on GPP because it is the largest carbon flux and the largest carbon input for terrestrial ecosystems, its sensitivity to drought is well documented, and its spatiotemporal patterns can be estimated in several ways[20–23]. Drought events are defined using a multiscalar drought metric, the Standardized Precipitation–Evapotranspiration Index (SPEI)[24], in which more negative values indicate more severe drought relative to average long-term conditions. SPEI can be based on a range of integration times (for example, 24-month SPEI integrates water status over the previous 24 months; see Methods), so we evaluate recovery time as a function of both integration time (1-, 6-, 12-, and 24-month SPEI) and of GPP data. This allows multiple lines of evidence to be comprehensively assessed, while providing an extensive sample of about 4.5 million drought and recovery events.

Using response functions derived with the machine learning algorithm 'Random Forests' (see Methods) we show that recovery time varies from immediate to multiple years across gradients of climate, vegetation, disturbance and drought. These factors explain 76–89% (Extended Data Table 1) of the variance in recovery time. Furthermore, we use standard metrics of variable importance (see Methods) to quantify and rank how each factor influences recovery time (Extended Data Fig. 1). Finally, given the consistency of results across datasets and integration times—in particular between the MsTMIP simulations and the two strongly observation-constrained datasets (see Methods)—we report results that integrate all combinations of GPP product and integration time to facilitate examination of emergent patterns.

Post-drought temperature and precipitation conditions (Fig. 1a, b and c) were the factors most strongly associated (Extended Data Fig. 1) with recovery time. Unsurprisingly, wetter conditions shortened and drier conditions lengthened drought recovery (Fig. 1c). Temperature extremes, both hot and cold, acted to lengthen recovery time. However, increases in drought recovery time were longer for warmer post-drought temperatures (Fig. 1a and b), suggesting that anticipated future temperature trends will further lengthen drought recovery.

Higher GPP (Fig. 1d and e) was also strongly associated (Extended Data Fig. 1) with longer recovery times. Here, GPP is expressed as the departure from the 2000–2010 mean seasonal cycle by grid cell (see Methods). Larger GPP amplitudes led to longer recovery times, especially in eastern North America and central Africa (Extended Data Fig. 2). In contrast, the value of pre-drought GPP functioned as a switch: if pre-drought GPP was already depressed—probably as a result of some other stressor—there was no further effect on recovery time. However, when the pre-drought GPP baseline was positive (indicating above average carbon uptake immediately preceding the drought event) recovery time increased dramatically.

Differences in CO_2 concentrations and biodiversity were also associated with differences in recovery time (Extended Data Fig. 1), but to a lesser extent than post-drought temperature and precipitation conditions or GPP itself. Increasing CO_2 concentration (Fig. 1f) has acted to monotonically shorten recovery times by about 4 months over the twentieth century, that is, drought impacts would be worse in the

[1]Woods Hole Research Center, Falmouth, Massachusetts 02540, USA. [2]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, Arizona 86011, USA. [3]Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA. [4]Department of Global Ecology, Carnegie Institution for Science, Stanford, California 94305, USA. [5]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, USA. [6]DendroLab and Graduate Program of Ecology, Evolution, and Conservation Biology (EECB), University of Nevada-Reno, Reno, Nevada 89557, USA. [7]Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA. [8]Informatics and Computing Program, Northern Arizona University, Flagstaff, Arizona 86011, USA. [9]Rocky Mountain Research Station, US Forest Service, Ogden, Utah 84401, USA. [10]Arable Labs Inc., 40 North Tulane Street, Princeton, New Jersey 08542, USA. [11]School of Earth Sciences and Environmental Sustainability, Northern Arizona University, Flagstaff, Arizona 86011, USA. [12]National Snow and Ice Data Center, Boulder, Colorado 80309, USA. [13]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA. [14]School of Forest Resources, University of Maine, Orono, Maine 04469, USA. [15]Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, Washington 99354, USA. [16]Department of Atmospheric Sciences, University of Illinois, Urbana, Illinois 61801, USA. [17]International Center for Climate and Global Change Research and School of Forestry and Wildlife Sciences, Auburn University, Auburn, Alabama 36849, USA.

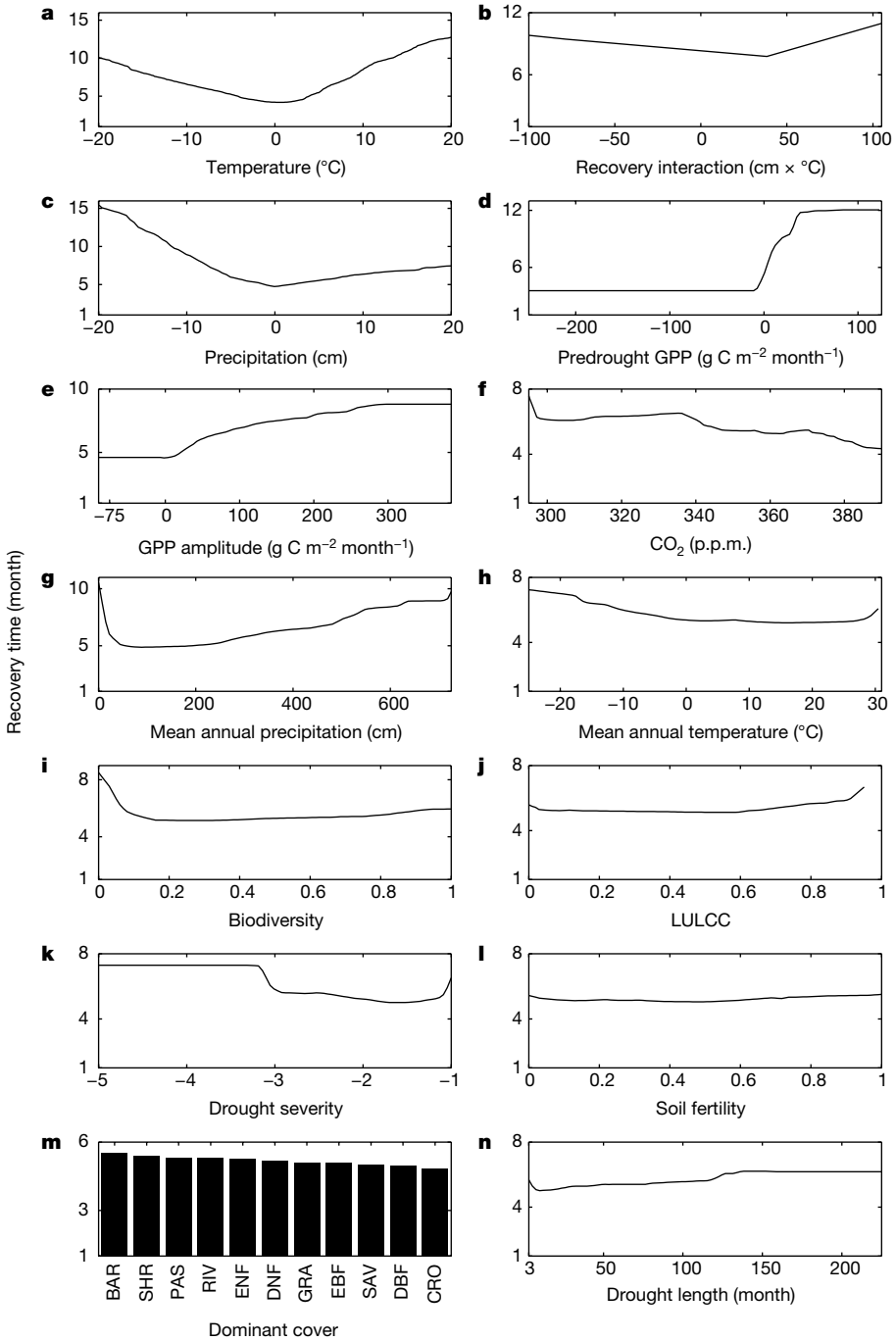


Figure 1 | Response functions for drought recovery time. Variables are ordered by descending importance from **a** to **n**, each showing recovery time across a single variable while holding the others constant. Note differences in the *y*-axis scales. Covariates **a** to **e** are departures from a baseline (see Methods); positive (negative) indicates above (below) average values. **b** shows the interaction between precipitation (in centimetres) and temperature (degrees Celsius). For land-use and land-cover change relative to pre-industrial baseline (LULCC; **j**), unity indicates that land cover is perfectly anti-correlated with its preindustrial baseline whereas zero indicates no change. Types of dominant cover (**m**) are BAR (barren), SHR (shrublands), PAS (pasture), RIV (water/river), ENF (evergreen needleleaf forest), DNF (deciduous needleleaf forest), GRA (grassland/steppe), EBF (evergreen broadleaf forest), SAV (savannah), DBF (deciduous broadleaf forest) and CRO (croplands).

absence of CO_2 fertilization. In contrast, biodiversity—normalized species richness of native species only (see Methods)—exhibited a threshold effect (Fig. 1i). Above a threshold of 0.2—primarily non-boreal forested systems and the tropics (Extended Data Fig. 3)—recovery time became longer as biodiversity increased across space. Although this is unexpected given the positive relationship between biodiversity and productivity[25], experimental evidence supporting greater drought resilience for high-biodiversity forests is lacking[26]. Possible mechanisms leading to longer recovery times in such systems include (1) more intensive use, and thus quicker exhaustion, of plant-available water[27], or (2) a historical lack of drought events and therefore a lack of adaptations to buffer against drought[28].

In contrast to during-drought temperature and precipitation, long-term mean temperature and precipitation were only of secondary importance (Extended Data Fig. 1), that is, deviations from mean climate were more important than the mean state itself. For temperature, extremely cold (for example, northern high latitudes; Extended Data Figs 2 and 3) and, to a lesser extent, extremely hot regions (for example,

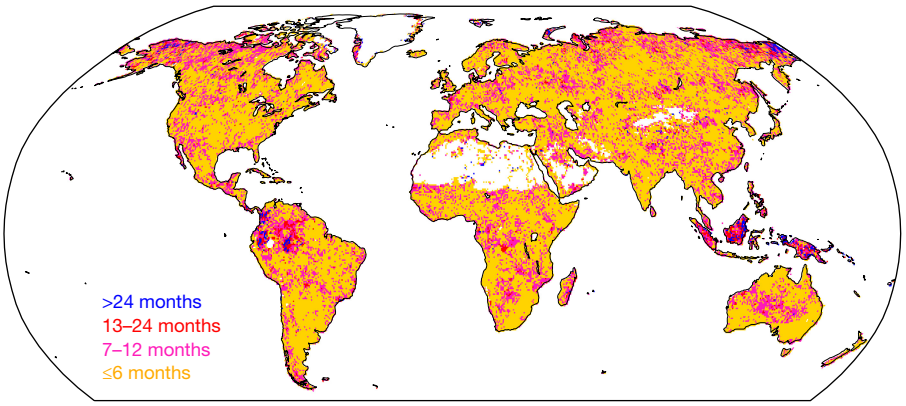


Figure 2 | Spatial pattern of drought recovery time. Recovery time by grid cell across all combinations of GPP and integration time (see Methods). White areas are water, barren, or did not experience any relevant drought events.

the Sahel region; Extended Data Figs 2 and 3) experienced longer recovery times (Fig. 1h). In very dry systems (<50 cm per annum; Extended Data Fig. 3) recovery times, as expected, increased with decreasing precipitation (Fig. 1g). Above this threshold, however, an inverse relationship was observed, with recovery times increasing as precipitation increases. This reflects a general trend where more productive regions (for example, higher-precipitation and higher-biodiversity regions; Extended Data Fig. 3) experienced longer drought recovery times.

More severe and longer droughts were, as expected, associated with longer recovery times (Fig. 1k and n). Drought metrics were, however, of tertiary importance as evidenced by their relatively low importance scores (Extended Data Fig. 1). Moderate droughts ($-1.05 \leq \mathrm{SPEI} \leq -1$) did lead to increased recovery times. Moderate droughts had return times (that is, number of months between successive drought events; Extended Data Fig. 4), that were shorter than those between non-moderate droughts. This suggests stress fatigue, whereby ecosystem recovery is successively degraded by exposure to repeated stressors at a frequency outside the evolutionary history of the affected system[29].

We also examined whether soil, fire or land-use characteristics influence recovery times, and show that these factors, similar to drought characteristics, had relatively little influence (Extended Data Fig. 1). Recovery time varied by up to a month across the global range of soil fertility (Fig. 1l) and fire regime—a composite of longer-term fire characteristics including size, frequency, intensity, season and extent (see Methods). Fires occurring during drought and subsequent recovery acted to increase recovery times by up to two months (Extended Data Fig. 5). Using land-cover/land-use change as the broadest form of disturbance, there was a weak tendency for recovery times to increase as land-use/land-cover diverged from its pre-industrial baseline (Fig. 1j).

Key spatial patterns of drought recovery were also evident (Fig. 2). Across all combinations of GPP and integration time (Extended Data Fig. 1), recovery times were typically longest in the tropics—in Amazonia and Indonesia more so than in Africa—as well as the northern high latitudes, especially the Russian Far East and Alaska. These patterns were linked to spatial gradients in productive potential.

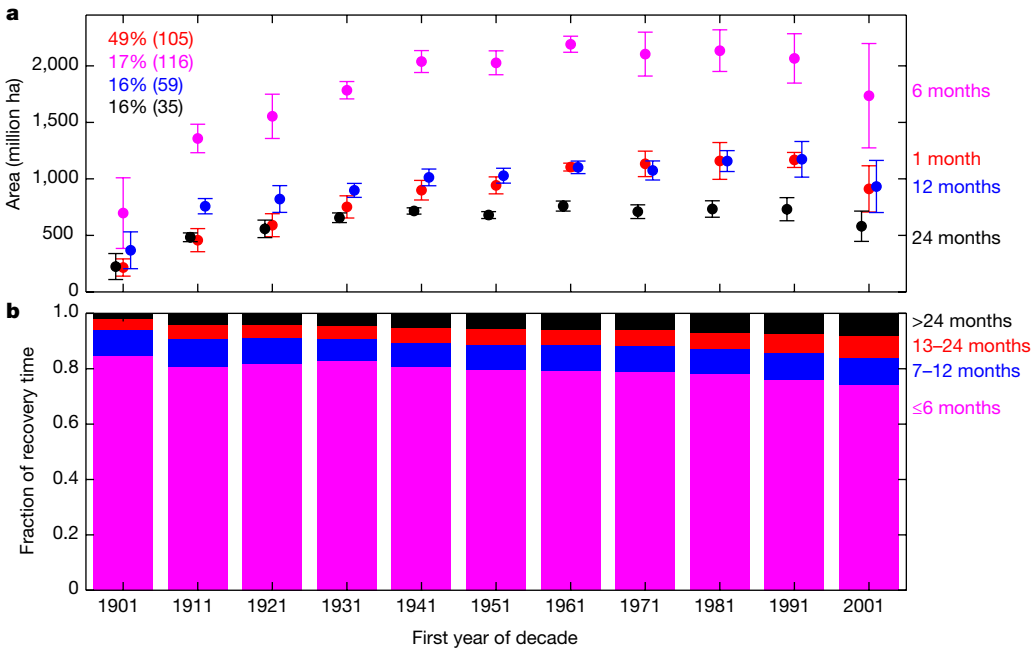


Figure 3 | Decadal changes in drought recovery. a, Global area undergoing a drought recovery for 1-month (red), 6-month (magenta), 12-month (blue), and 24-month SPEI (black). Data are mean ± s.d. ($n = 10$) from annual values. The colour-coded values are linear trend estimates (all $^*P < 0.05$) in per cent relative to first decade and in million hectares per decade (in parentheses). Trends derived using weighted linear regression with weights given by $1/\sigma^2$. **b**, Recovery time distribution. The range in months is colour-coded. Values are across all integration times—for MsTMIP only—and events are temporally sorted by end date.

That is, the northern high latitudes exhibited low biodiversity and low rainfall (Extended Data Fig. 3), whereas the converse held for the tropics. These spatial patterns also reflected dominant cover (Fig. 1m), where barren areas (tundra, desert, polar desert, rock, and ice) showed the longest recovery times.

Finally, we found that drought impacts have increased over the twentieth century. As the only dataset with a centennial-scale time domain, we used MsTMIP simulations to diagnose temporal changes in recovery time. MsTMIP simulations are consistent with the observation-based datasets in terms of drought patterns and explanatory factors during their overlapping time period, and are therefore expected to be representative during the earlier period as well (see Methods). The areal extent of the vegetated biosphere recovering from drought has increased since 1901 (Fig. 3a). Linear trend values range from 16% to 54% or from 35 to 116 million hectares per decade—though uncertainties remain on the absolute magnitudes of area affected in the initial and final decades (see Methods). In addition, multi-seasonal and multi-year recovery times have become more frequent (Fig. 3b). This trend towards longer drought recovery was similarly apparent for extreme recovery times. These increased monotonically from 15 and 20 months (95th and 97th percentiles, respectively) in 1901–1910 to 36 and 58 months in 2001–2010; suggesting that the trend towards more "more extreme extremes"[8] is already under way. Thus, even though the statistical trends in hydrological drought are mixed over the twentieth century[5], we find evidence that the impacts of drought on ecosystems have been increasing over this period.

Given the changes in temperature as well as the increases in drought frequency and severity projected for the twenty-first century, our results suggest that terrestrial ecosystems will take longer to recover after droughts, which could increase the vulnerability of these systems to drought. Many ecological processes, from physiological acclimation of individuals to species turnover in communities, may act to buffer drought impacts in the future, but the effectiveness and timescales over which these processes can act relative to changing drought regimes is unknown. The interplay between expected longer recovery times and more frequent drought events is of particular concern in the Amazon, where longer recovery times are the norm (Fig. 1) and where successive "once-in-a-century" droughts in 2005 and 2010 have prevented full recovery[30]. Our findings indicate that a chronic state of incomplete recovery may become established over the remainder of the twenty-first century, with adverse consequences for the land carbon sink.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 August 2016; accepted 6 June 2017.

1. Haines, A., Kovats, R. S., Campbell-Lendrum, D. & Corvalán, C. Climate change and human health: impacts, vulnerability and public health. *Public Health* **120,** 585–596 (2006).
2. Schwalm, C. R. *et al.* Reduction in carbon uptake during turn of the century drought in western North America. *Nat. Geosci.* **5,** 551–556 (2012).
3. Touma, D., Ashfaq, M., Nayak, M. A., Kao, S. C. & Diffenbaugh, N. S. A multi-model and multi-index evaluation of drought characteristics in the 21st century. *J. Hydrol.* **526,** 196–207 (2015).
4. Reichstein, M. *et al.* Climate extremes and the carbon cycle. *Nature* **500,** 287–295 (2013).
5. Trenberth, K. E. *et al.* Global warming and changes in drought. *Nat. Clim. Chang.* **4,** 17–22 (2014).
6. Sheffield, J., Wood, E. F. & Roderick, M. L. Little change in global drought over the past 60 years. *Nature* **491,** 435–438 (2012).
7. Seneviratne, S. I. *et al.* in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (eds Field, C. B. *et al.*) 109–230 (Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, Cambridge Univ. Press, 2012).
8. Alexander, L. V. *et al.* Climate extremes: progress and future directions. *Int. J. Climatol.* **29,** 317–319 (2009).
9. Cook, B. I., Ault, T. R. & Smerdon, J. E. Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Sci. Adv.* **1,** e1400082 (2015).
10. Lenton, T. M. *et al.* Tipping elements in the Earth's climate system. *Proc. Natl Acad. Sci. USA* **105,** 1786–1793 (2008).
11. Mishra, A. K. & Singh, V. P. A review of drought concepts. *J. Hydrol.* **391,** 202–216 (2010).
12. DeChant, C. M. & Moradkhani, H. Analyzing the sensitivity of drought recovery forecasts to land surface initial conditions. *J. Hydrol.* **526,** 89–100 (2014).
13. Pan, M., Yuan, X. & Wood, E. F. A probabilistic framework for assessing drought recovery. *Geophys. Res. Lett.* **40,** 3637–3642 (2013).
14. Martorell, S., Diaz-Espejo, A., Medrano, H., Ball, M. C. & Choat, B. Rapid hydraulic recovery in *Eucalyptus pauciflora* after drought: linkages between stem hydraulics and leaf gas exchange. *Plant Cell Environ.* **37,** 617–626 (2014).
15. Secchi, F. & Zwieniecki, M. A. Down-regulation of plasma intrinsic protein1 aquaporin in poplar trees is detrimental to recovery from embolism. *Plant Physiol.* **164,** 1789–1799 (2014).
16. Anderegg, W. R. L. *et al.* Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. *Science* **349,** 528–532 (2015).
17. Luo, Y., Keenan, T. F. & Smith, M. Predictability of the terrestrial carbon cycle. *Glob. Change Biol.* **21,** 1737–1751 (2015).
18. Nepstad, D. C., Stickler, C. M., Soares-Filho, B. & Merry, F. Interactions among Amazon land use, forests and climate: prospects for a near-term forest tipping point. *Phil. Trans. R. Soc. Lond. B* **363,** 1737–1746 (2008).
19. Banerjee, O., Bark, R., Connor, J. & Crossman, N. D. An ecosystem services approach to estimating economic losses associated with drought. *Ecol. Econ.* **91,** 19–27 (2013).
20. Zhao, M. & Running, S. W. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* **329,** 940–943 (2010).
21. Jung, M. *et al.* Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res. Biogeosci.* **116,** G00J07 (2011).
22. Huntzinger, D. N. *et al.* The North American Carbon Program Multi-scale synthesis and Terrestrial Model Intercomparison Project—Part 1: Overview and experimental design. *Geosci. Model Dev.* **6,** 2121–2133 (2013).
23. Beer, C. *et al.* Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science* **329,** 834–838 (2010).
24. Beguería, S., Vicente-Serrano, S. M., Reig, F. & Latorre, B. Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *Int. J. Climatol.* **34,** 3001–3023 (2014).
25. Tobner, C. M., Paquette, A., Reich, P. B., Gravel, D. & Messier, C. Advancing biodiversity–ecosystem functioning science using high-density tree-based experiments over functional diversity gradients. *Oecologia* **174,** 609–621 (2014).
26. Grossiord, C. *et al.* Tree diversity does not always improve resistance of forest ecosystems to drought. *Proc. Natl Acad. Sci. USA* **111,** 14812–14815 (2014).
27. Grossiord, C., Granier, A., Gessler, A., Jucker, T. & Bonal, D. Does drought influence the relationship between biodiversity and ecosystem functioning in boreal forests? *Ecosystems* **17,** 394–404 (2014).
28. Engelbrecht, B. M. *et al.* Drought sensitivity shapes species distribution patterns in tropical forests. *Nature* **447,** 80–82 (2007).
29. Hacke, U. G., Stiller, V., Sperry, J. S., Pittermann, J. & McCulloh, K. A. Cavitation fatigue. Embolism and refilling cycles can weaken the cavitation resistance of xylem. *Plant Physiol.* **125,** 779–786 (2001).
30. Saatchi, S. *et al.* Persistent effects of a severe drought on Amazonian forest canopy. *Proc. Natl Acad. Sci. USA* **110,** 565–570 (2013).

Acknowledgements Funding for this research was provided by the National Science Foundation (NSF) grant DEB EF-1340270. C.R.S. was also supported by National Aeronautics and Space Administration (NASA) grants NNX12AK12G, NNX12AP74G, NNX10AG01A and NNX11AO08A. J.B.F. contributed to this paper from the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. Government sponsorship acknowledged. Support was provided to J.B.F. by NASA grants NNN13D504T (CARBON), NNN13D202T (INCA), and NNN13D503T (SUSMAP). Funding for the MsTMIP activity was provided through NASA grant NNX10AG01A. Data management support for preparing, documenting and distributing model driver and output data was performed by the Modeling and Synthesis Thematic Data Center at Oak Ridge National Laboratory (http://nacp.ornl.gov), with funding through NASA grant NNH10AN681. Finalized MsTMIP datasets are archived at the Oak Ridge National Laboratory Distributed Active Archive Center (http://daac.ornl.gov). This is MsTMIP contribution number 10.

Author Contributions C.R.S. and W.R.L.A. designed the analysis. C.R.S. carried out the analysis and wrote the manuscript with contributions from all authors. W.R.L.A., A.M.M. and J.B.F. contributed to the framing of the paper. D.N.H. is the overall lead of the MsTMIP effort; R.C., J.B.F., A.M.M., K.S., C.R.S., Y.F. and Y.W. serve as the MsTMIP core team. D.H., M.H., A.J. and H.T. contributed to MsTMIP results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to C.S. (schwalm.christopher@gmail.com).

Reviewer Information *Nature* thanks M. Migliavacca and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

SPEI. Drought severity is quantified using the multiscalar SPEI metric[23] (http://sac.csic.es/spei/database.html) calculated using monthly precipitation and potential evapotranspiration from the CRU TS3.2 climate dataset[31] (http://www.cru.uea.ac.uk/cru/data/hrg/). First, a climatic water balance is used; the difference between monthly precipitation and potential evapotranspiration is calculated. These values represent either a water surplus or deficit for a given month and are then aggregated over the desired integration timescale: 1-, 6-, 12-, and 24-month periods for this study. That is, a 6-month SPEI value is based on the cumulative water deficit or surplus over the preceding 6 months. As such, water status can change from surplus to deficit, and vice versa, over integration windows longer than one month. Thus, the integration timescale does not measure drought length; rather, it serves as an aggregation scheme to quantify hydrologic state at one point in time. After aggregation, values are then normalized using a three-parameter log-logistic distribution[23]. This standardized value is the SPEI index, with a mean of zero and standard deviation of unity, where negative values indicate drier conditions. The FAO-56 Penman–Monteith estimation of potential evapotranspiration is used throughout. We use −1 as an event threshold throughout to focus on conditions when drought is expected to be the main driving factor of carbon metabolism while also allowing sufficient events to be sampled. This dataset is generated at 0.5° spatial resolution and monthly time step and covers the 1901 to 2010 time period.

Drought events and recovery time. Drought severity is tracked for all vegetated land pixels and is solely a function of the SPEI metric. A drought event begins when SPEI ≤ -1, for at least three consecutive months, and ends when SPEI > -1. Recovery time is tracked starting the first post-drought month. When post-drought GPP returns to its pre-drought level drought recovery has occurred. The pre-drought GPP level value is defined as pre-drought mean monthly GPP across the same number of months as the relevant drought event. Recovery is based on detrended and deseasonalized GPP (that is, the mean monthly seasonal cycle GPP is removed relative to a common base period (2000 to 2010) for each pixel and for all datasets). A 3-month forward window is used to smooth sub-seasonal variation for both SPEI and GPP. Drought events and subsequent recovery are pixel-specific and are calculated for every combination of GPP and SPEI integration time. Of the 4,490,680 events catalogued, 624,330 show an overlap between the GPP pre-drought baseline and the recovery time of the previous drought. Excluding these events does not change any conclusion presented herein. In addition, only drought and recovery events that are fully contained in the 1901 to 2010 time period are analysed. A drought event that, for example, begins before 2010 but ends afterwards is not resolvable. Similarly, an event that starts before 1901 but ends afterwards is also not resolvable. All such events are excluded from our analysis. This leads to a likely underestimation of drought recovery, especially areal extent, in the years immediately after 1901 and before 2010 (Fig. 3), but does not affect the overall trend. Finally, as SPEI is retrospective (that is, the 24-month SPEI is calculated over the previous 24 months), the initial values for longer-term SPEI integration times for MsTMIP are biased low because the data record begins only in 1901.

MODIS. This follows the MOD17 GPP algorithm[20]. GPP is given by the product of maximum light-use efficiency, the fraction of absorbed photosynthetically active radiation, incoming radiation, and two scalar reduction factors that represent limitations on photosynthesis through temperature and vapour pressure deficit. These functions act to depress GPP given low temperatures and high vapour pressure deficit. The maximum light-use efficiency as well as the minimum and maximum threshold values for both scalar reduction factors vary by biome. Gap-filling is used to fill cloud-contaminated satellite data. The MOD17 GPP algorithm is forced using daily minimum temperature, daytime temperature, daily average temperature, daily vapour pressure, and daily total downward short-wave solar radiation—45% is assumed photosynthetically active radiation—from the NCEP/DOE reanalysis II (ref. 32) at a daily time step. This dataset is generated at 1 km spatial resolution (resampled to 0.5° for this study) and composited from a daily to monthly time step for the 2000 to 2010 time period.

Upscaled FLUXNET. GPP is reconstructed using empirical upscaling of site-level eddy covariance FLUXNET data[21]. A 25-member ensemble of individual model trees, using fivefold cross-validation, is trained using gridded explanatory variables and flux-partitioned GPP from about 250 co-located globally distributed FLUXNET sites. The trained algorithm is then applied to all vegetated land pixels. The final GPP for each pixel-month is the median across the full ensemble. Explanatory variables used in training include the remotely sensed fraction of absorbed photosynthetically active radiation, climatic fields, land cover data, and indicator variables for the photosynthetic pathway. This dataset is generated at 0.5° spatial resolution and monthly time step and covers the 1982 to 2008 time period.

MsTMIP. The land surface model ensemble is drawn from MsTMIP—the Multi-scale synthesis and Terrestrial Model Intercomparison Project[22] (http://dx.doi.org/10.3334/ORNLDAAC/1225). MsTMIP uses a standardized simulation protocol—historical forcing data, boundary conditions, and spin-up are uniform across all models[33]—to isolate structural differences. MsTMIP runs are global (0.5° spatial resolution) and monthly from 1901 to 2010 and use a semi-factorial set of simulations where historical time-varying climate, CO_2 concentration (that is, CO_2 fertilization), land cover, and nitrogen deposition are sequentially 'turned on' after steady state is reached. For this study an ensemble mean based on MsTMIP Version 1 models (CLM[34], CLM4VIC[35], DLEM[36], ISAM[37] and TEM6 (ref. 38)) that include all time-varying factors (simulation BG1 (ref. 22)) is used.

Response functions. Partial dependence plots based on the Random Forest algorithm[39] are used to visualize the relationship between explanatory covariates and recovery time, independent of other covariates. Partial dependence gives the marginal effect of a covariate on the response variable, so the y axis is only interpretable within and not across covariates. The Random Forest algorithm here uses 100 binary decision trees, 1 covariate chosen at random from the full set to determine the spitting rule, with a minimal terminal node size of 5 (nodes with less than 5 observations cannot be split). Before fitting a variance stabilizing transformation (3/4-power transform from the Box–Cox family of transformations) is applied to the target variable, recovery time. A Random Forest is then fitted to every combination of GPP and SPEI integration time. Partial dependence is evaluated using equally spaced bins that span the range in each covariate. The explanatory covariates used are: drought severity (mean SPEI value during the drought event), drought length (number of consecutive months, using a forward 3-month moving window, where SPEI ≤ -1), pre-drought GPP (mean GPP immediately before drought), GPP amplitude (difference between largest pre-drought baseline and the smallest GPP flux during the drought event and subsequent recovery), soil fertility (normalized cation exchange capacity from the MsTMIP value-added version of the Harmonized World Soil Database[33]), biodiversity (normalized species richness[40] of native species only where unity is the highest value), CO_2 concentration[33] (to assess the CO_2 fertilization effect), long-term climate normals (1971–2000 means from the CRU-NCEP dataset[33]) of temperature and precipitation, integrated anomalies (relative to 1971–2000 mean seasonal cycle) of during-recovery temperature and precipitation[33] as well as their interaction, pre-industrial versus drought event land-cover/land-use[41] similarity (based on multidimensional scaling with pair-wise correlations across harmonized land-use classes), and dominant land cover[42]. The interaction between drought severity and length was considered initially but as its inclusion does not improve the trained Random Forest we did not use it. We also control for fire effects using fire regime[43] (including unknown/missing as an additional regime) and burned area[44]. For pre-drought GPP we tested three formulations: (1) the pre-drought time horizon mirrors drought length, (2) a fixed pre-drought time horizon of 6 months, and (3) a fixed pre-drought time horizon of 3 months. Our conclusions are unchanged regardless of aggregation scheme and we use the mirror approach in the main text.

Variable importance. With the Random Forest algorithm[39] importance scores for each covariate can be calculated. These scores reflect how important each covariate is in determining the fitted values of drought recovery (Extended Data Fig. 1). Variable importance is calculated for each combination of GPP dataset and integration time as the total decrease in residual sum of squares, averaged over all trees, from splitting on the target covariate[39]. This value is then normalized, that is, all importance values sum to unity for each GPP dataset and integration time combination. We use variable importance as a heuristic to group covariates into tiers. In general, an important covariate has a response function that displays interpretable features such as tipping points, well defined minima/maxima, or a slope. In contrast, variables with flat response functions, that is, no change in recovery time over the range of the target variable, are typically relatively unimportant. In our study several relatively unimportant variables (Fig. 1) show response functions that vary by less than one month, that is, less than the underlying time step of the input data. The judgment of low importance in such cases is intuitive as it is impossible to diagnose and attribute sub-monthly variations with monthly data.

Cross-product consistency. The use of all three data sources to jointly diagnose drought recovery time from 1901 to 2010 is based on their agreement during the period of overlap. The model-based MsTMIP dataset, which is used as an extrapolator across the twentieth century, is consistent with both observation-based datasets, upscaled FLUXNET and MODIS, when examined over a common 2000–2008 time frame. This is supported by high cross-product consistency in (1) spatial patterns of drought recovery time (Extended Data Fig. 6), (2) number of drought events (inset values in Extended Data Fig. 6), and (3) areal extent of drought recovery. Here, mean annual areal extent by decade ($n = 10$ by dataset and decade) overlap at 95% confidence (false discovery rate of 0.05) during the 1981

and 1991 decades for upscaled FLUXNET and MsTMIP and in the 2001 decade for all three datasets.

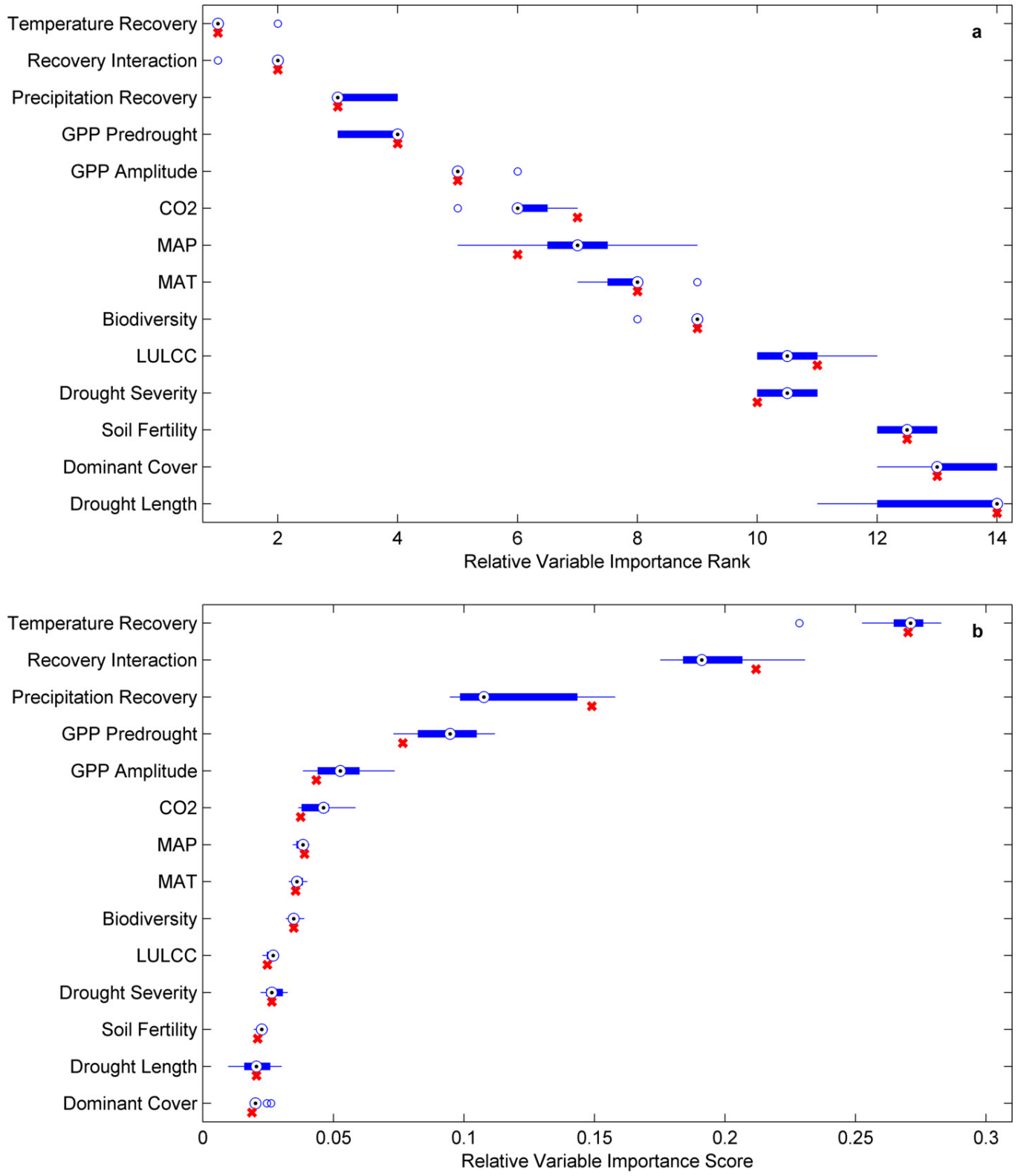
High cross-product similarity scores. The median correlation between response functions across all possible combinations of dataset and integration time during the 2000–2008 common time frame is 0.90 (Extended Data Fig. 7) with over 75% of all similarity scores greater than 0.70 (unity is perfect agreement; Extended Data Fig. 7o shows cumulative distribution function of all similarity scores). Holding integration time constant improves the median correlation to 0.93. This cross-product correspondence is reinforced by the high and highly similar proportion (range 81–89%) of variance explained (in a least-squares sense based on out-of-bag observations only) in the common time period (Extended Data Table 1).

Overall, this indicates that differences in response functions derived using the full time period of each GPP dataset are due to differences in sample size but not underlying mechanisms. That is, MsTMIP simulations sample a longer time period but exhibit the same underlying patterns of drought recovery time as the upscaled FLUXNET and MODIS GPP datasets. We leverage this convergence (see main text) by aggregating all 12 combinations of GPP dataset and SPEI integration time (Figs 1 and 2) with a weighted median that weights each combination based on length of data record and using MsMTIP to extrapolate in time (Fig. 3).

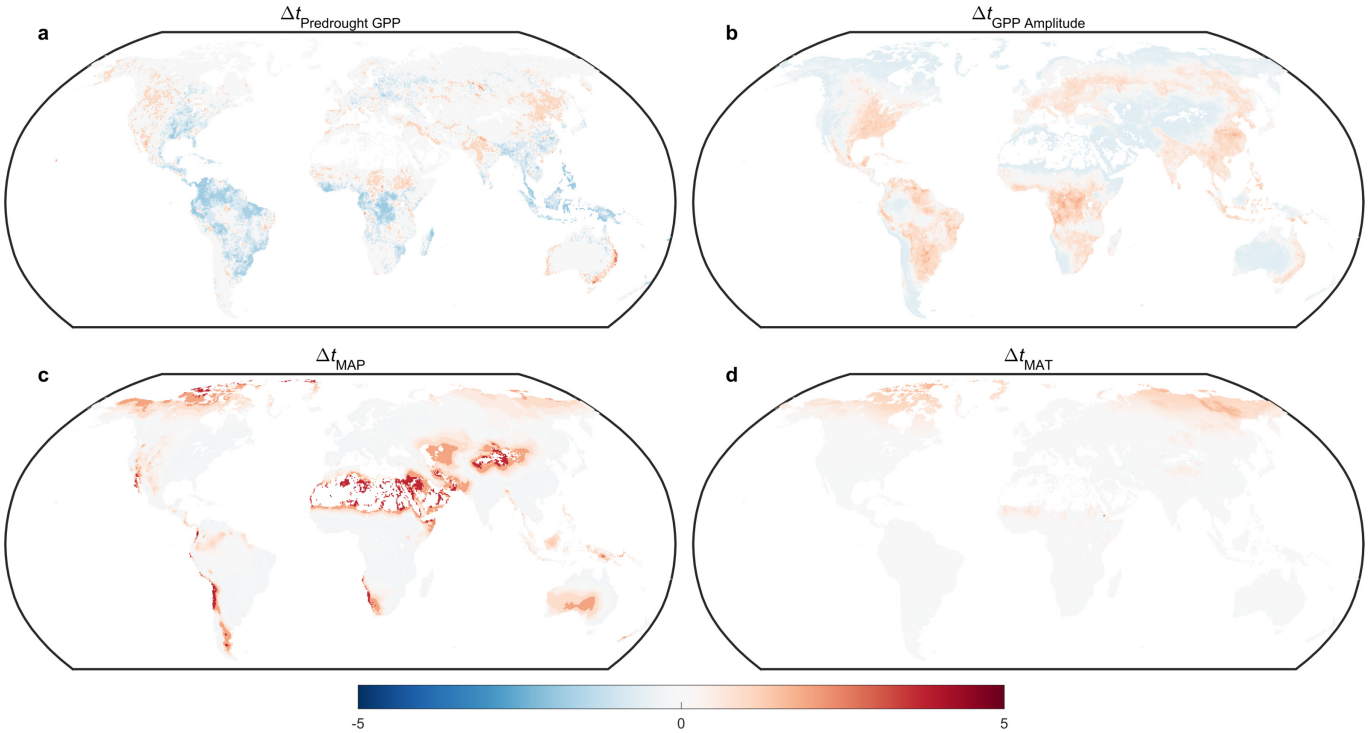
Pre-drought GPP baseline. In the main text the pre-drought value is based on a mirroring approach, that is, the pre-drought time horizon mirrors drought length. To evaluate how changing the baseline period affects our findings we compare this default approach across the common time frame for all three datasets (Extended Data Fig. 8) with (1) a fixed pre-drought time horizon of 3 months (Extended Data Fig. 8a) and (2) a fixed pre-drought time horizon of 6 months (Extended Data Fig. 8b). This analysis reveals high similarity scores (median correlation: 0.97 and 0.96 respectively) between response functions regardless of pre-drought baseline and indicates that our results are not sensitive to the aggregation scheme used. Lastly, the high and highly similar proportion of variance explained is maintained regardless of pre-drought baseline used (Extended Data Table 1).

Data availability. The underlying data used in this study is available in the public domain. The drought metric, the Standardized Precipitation–Evapotranspiration Index (SPEI), is archived at the State Agency for Scientific Research (CSIC) of Spain (http://sac.csic.es/spei/database.html). MODIS GPP is retrievable from the University of Montana's Numerical Terradynamic Simulation Group (http://www.ntsg.umt.edu/project/mod17). Upscaled FLUXNET data are available from M. Jung (mjung@bgc-jena.mpg.de) upon reasonable request. MsTMIP simulation output is archived at the Oak Ridge National Laboratory (https://doi.org/10.3334/ORNLDAAC/1225). Source Data for Figs 1–3 are additionally provided as Excel spreadsheets in the online version of the paper.

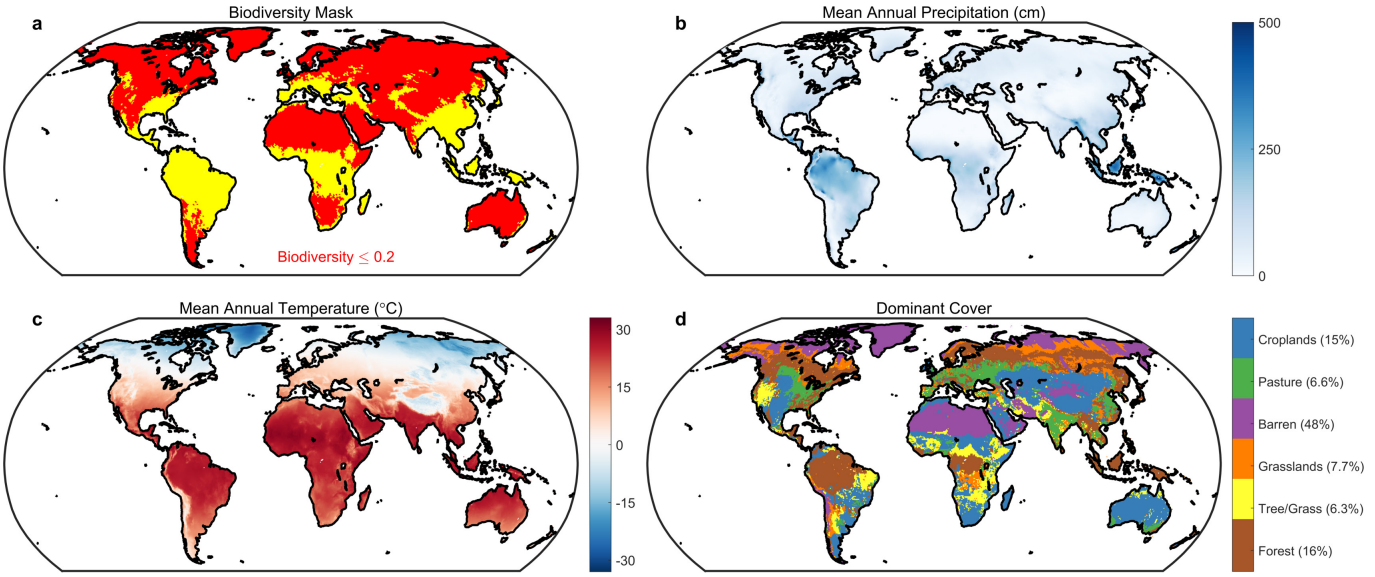
31. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. *Int. J. Climatol.* **34,** 623–642 (2014).
32. Kanamitsu, M. *et al.* NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Am. Meteorol. Soc.* **83,** 1631–1643 (2002).
33. Wei, Y. *et al.* The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project—Part 2: Environmental driver data. *Geosci. Model Dev.* **7,** 2875–2893 (2014).
34. Mao, J., Thornton, P. E., Shi, X., Zhao, M. & Post, W. M. Remote sensing evaluation of CLM4 GPP for the period 2000-09. *J. Clim.* **25,** 5327–5342 (2012).
35. Lei, H. *et al.* Sensitivity of global terrestrial gross primary production to hydrologic states simulated by the Community Land Model using two runoff parameterizations. *J. Adv. Model. Earth Syst.* **6,** 658–679 (2014).
36. Tian, H. Q. *et al.* Century-scale response of ecosystem carbon storage to multifactorial global change in the Southern United States. *Ecosystems* **15,** 674–694 (2012).
37. Jain, A. K. & Yang, X. Modeling the effects of two different land cover change data sets on the carbon stocks of plants and soils in concert with CO_2 and climate change. *Glob. Biogeochem. Cycles* **19,** *GB2015* (2005).
38. Hayes, D. J. *et al.* Is the northern high-latitude land-based CO_2 sink weakening? *Glob. Biogeochem. Cycles* **25,** *GB3018* (2011).
39. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2,** 18–22 (2002).
40. Ellis, E. C., Antill, E. C. & Kreft, H. All is not loss: plant biodiversity in the Anthropocene. *PLoS One* **7,** e30535 (2012).
41. Hurtt, G. C. *et al.* The underpinnings of land-use history: three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands. *Glob. Change Biol.* **12,** 1208–1229 (2006).
42. Meiyappan, P. & Jain, A. K. Three distinct global estimates of historical land-cover change and land-use conversions for over 200 years. *Front. Earth Sci.* **6,** 122–139 (2012).
43. Archibald, S., Lehmann, C. E., Gómez-Dans, J. L. & Bradstock, R. A. Defining pyromes and global syndromes of fire regimes. *Proc. Natl Acad. Sci. USA* **110,** 6442–6447 (2013).
44. Giglio, L., Randerson, J. T. & Werf, G. R. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *J. Geophys. Res. Biogeosci.* **118,** 317–328 (2013).



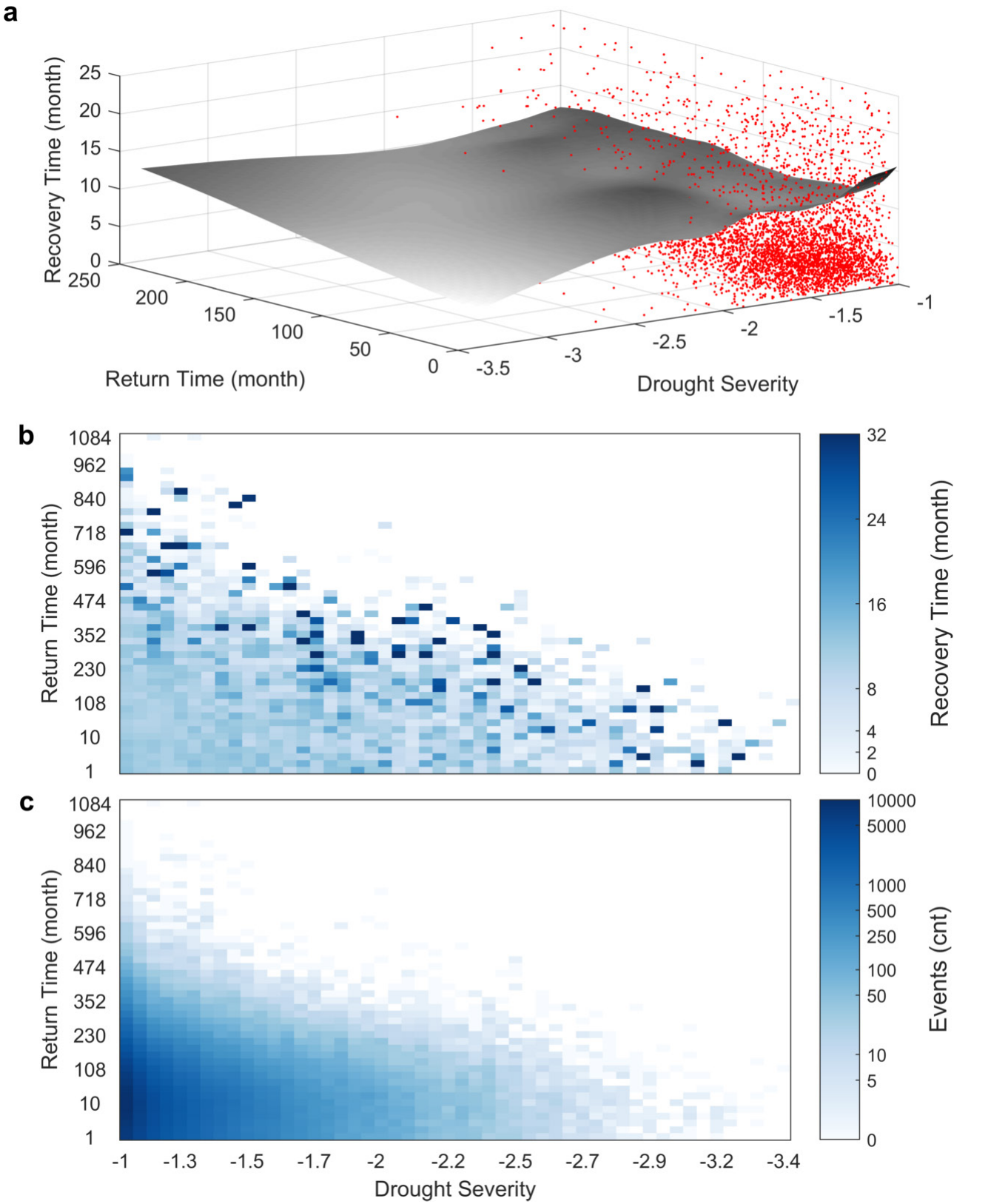
Extended Data Figure 1 | Variable importance scores. a, Covariates by rank of variable importance. **b**, Covariates by value of relative variable importance. Scores allow the skill of a trained Random Forest to be decomposed and mapped to individual covariates. Box plots are based on all possible combinations of GPP dataset and integration time. The box range covers the 25th to 75th percentiles; whiskers extend from about the 0.50th to 99.50th percentiles; outlying values are given by open circles. Red crosses give the composite value from the weighted median with weights given by the length, in years, of each GPP dataset. Both panels are sorted by median value from smallest rank (highest importance) to largest rank (lowest importance). Variable importance is calculated as the total decrease in residual sum of squares, averaged over all trees, from splitting on the target covariate (see Methods). CO2 indicates CO_2 fertilization (p.p.m.), LULCC is land-use and land-cover change relative to the pre-industrial baseline, MAP is mean annual precipitation 1971–2000, and MAT is mean annual temperature 1971–2000.



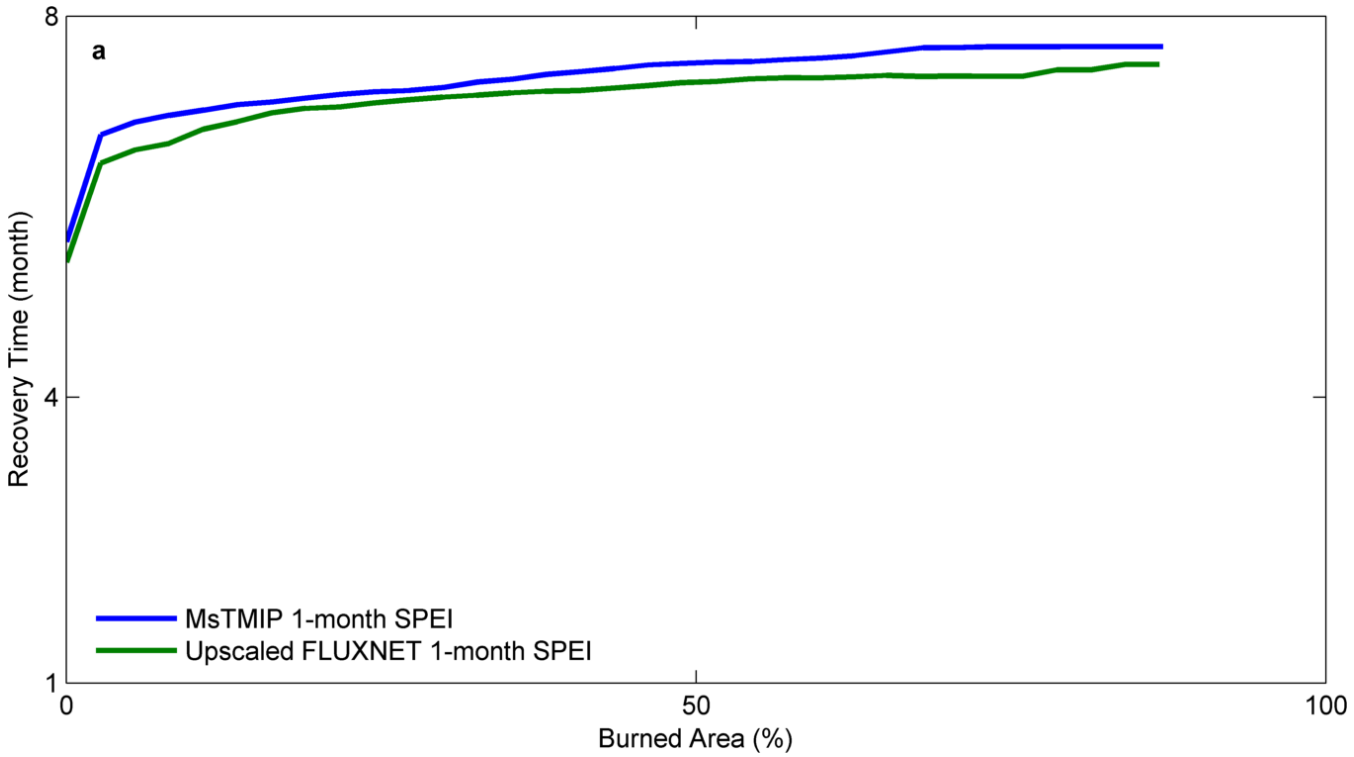
Extended Data Figure 2 | Spatial gradients of recovery time. a, Pre-drought GPP. **b**, GPP amplitude. **c**, MAP. **d**, MAT. Mapped values for **a** and **b** are based on the mean across all events (all GPP datasets, all integration times, and all time steps) by grid cell. For **c** and **d** the covariates from training the Random Forests algorithm (see Methods) are used directly. Recovery time is then projected spatially by matching response function values with each mapped factor by grid cell. To highlight spatial patterns values are expressed as difference in recovery time (Δt) relative to median recovery time by factor; red (blue) values indicate where a given factor is spatially associated with longer (shorter) drought recovery times. White areas are water; grey areas are barren or did not experience any relevant drought events.



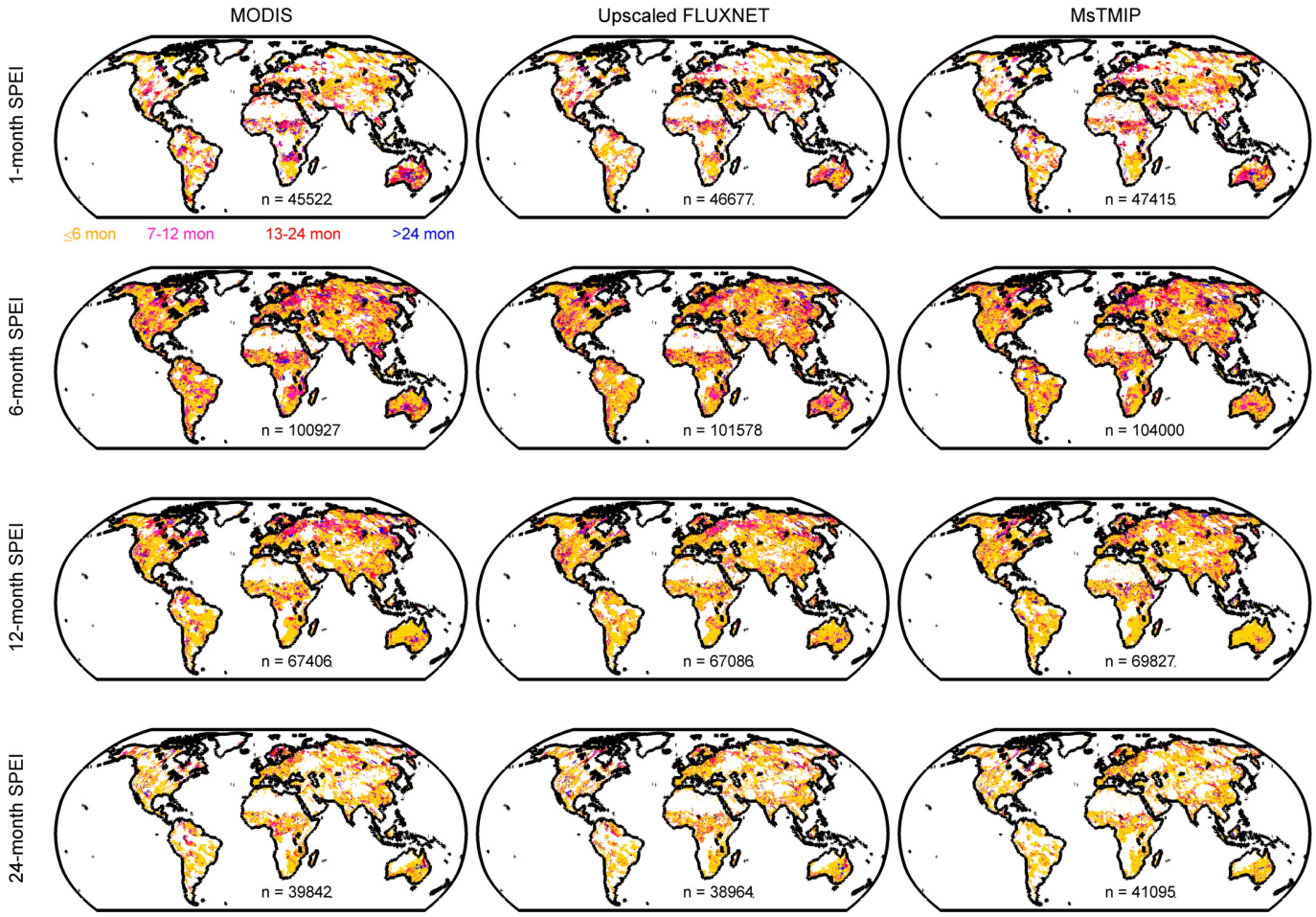
Extended Data Figure 3 | Enviroclimatic spatial gradients associated with recovery time. a, Mask of normalized biodiversity ≤ 0.2; these areas show an increase in recovery time with decreasing biodiversity across space. **b**, MAP. **c**, MAT. **d**, LULCC. Values in parentheses give areal extent as percentage. Aggregated types include Barren (tundra, desert, polar desert, rock, ice), Tree/Grass (shrublands, savannah, riparian/riverine systems) and Forest (all forest types).



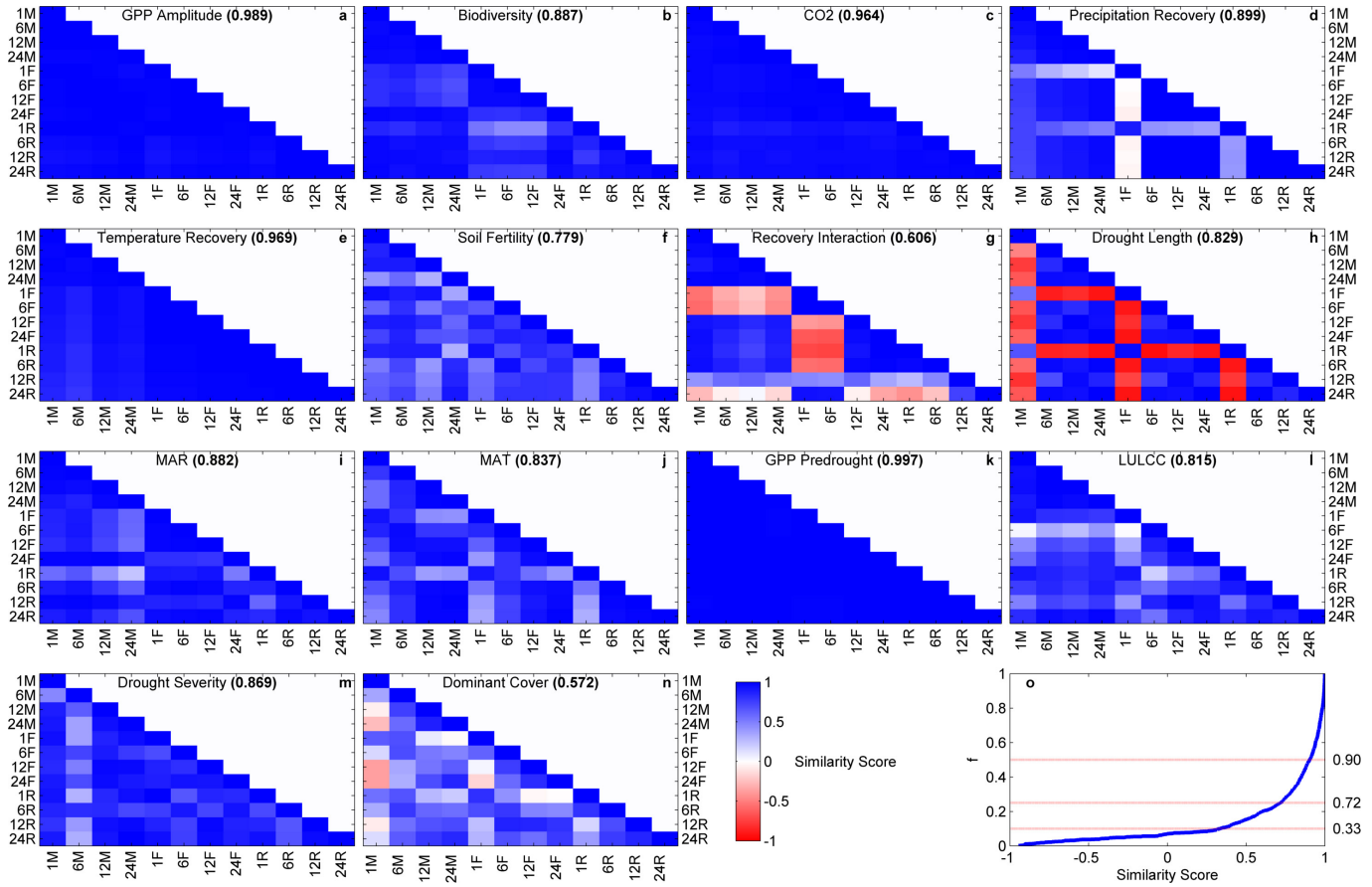
Extended Data Figure 4 | Recovery time for moderate droughts. **a**, Recovery time as a function of drought severity (mean SPEI by drought event) and return interval—calculated as the number of months between successive drought events for a given pixel but excluding the first drought event and droughts where no recovery is observed. This is a representative subsample of 5,000 events from MsTMIP using 1-month SPEI as shown. The grey surface is a smoothed surface to aid visual interpretation. **b**, Mean recovery time by 50 equidistant bins of return time and drought severity. **c**, Number of drought and recovery events for each bin in **b**; note log scale.



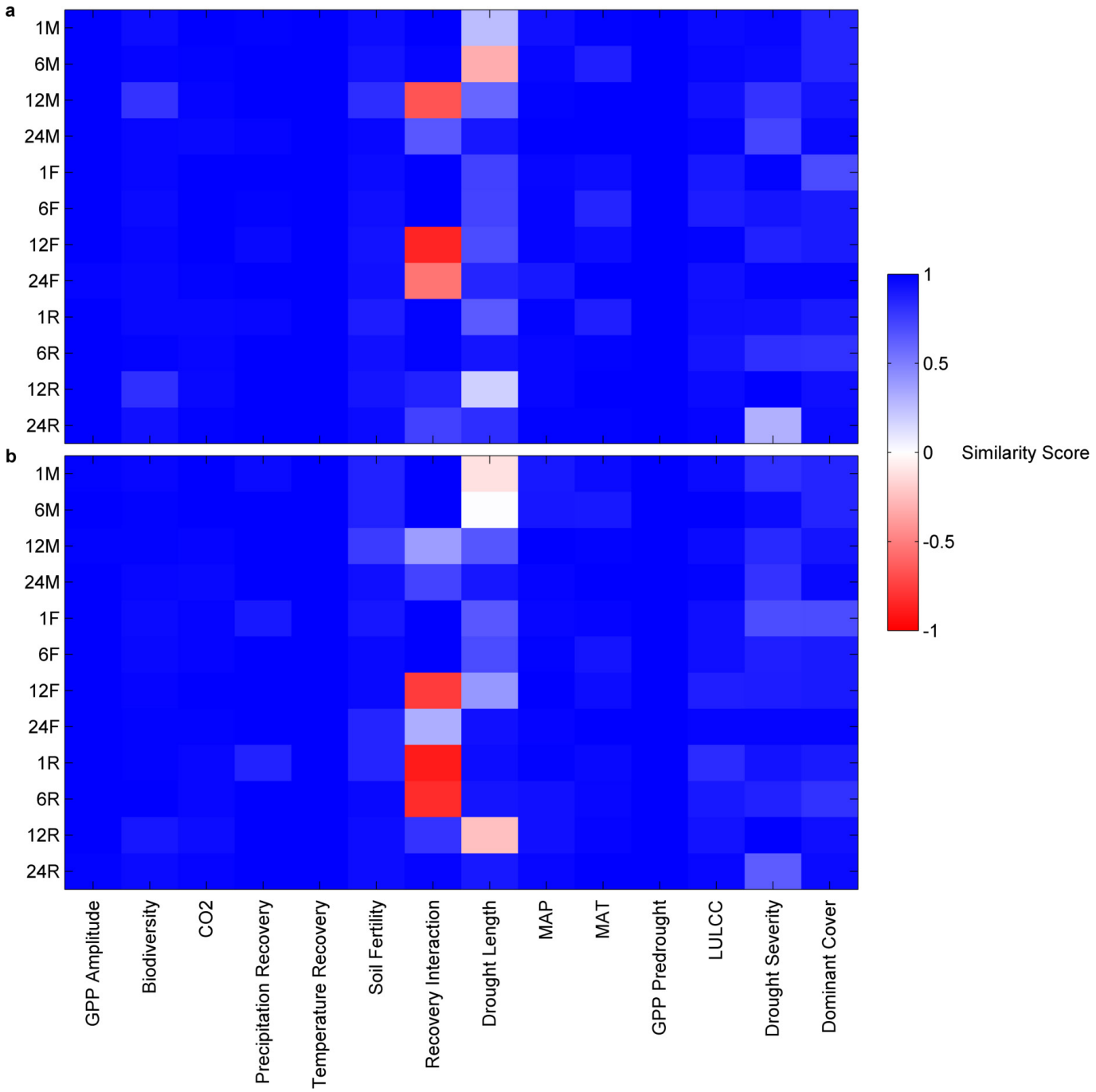
Extended Data Figure 5 | Response functions of recovery time relative to burned area. MsTMIP and upscaled FLUXNET, with 1-month SPEI integration time, for 1997–2010—overlap with GFED data[44] used to calculate burned area—shown. The trained Random Forests algorithm used both fire regime[43] and burned area[44] in addition to the standard set of covariates. Burned area is highly skewed; its 90th percentile is 0.035, such that the data support is concentrated at smaller burned area fractions. Burned area was used in training the Random Forests algorithm, given by its maximum value, for a given pixel, during drought and subsequent recovery. Both fire regime and burned area have low importance scores and are ranked 12 and 13.5, respectively, out of 16 covariates.



Extended Data Figure 6 | Spatial patterns of recovery time. Panels show mean recovery time by grid cell for all combinations of GPP dataset and integration time. Colours denote recovery time as shown. Inset values show number of drought and recovery events, *n*. White areas are water, barren, or did not experience any relevant drought events.



Extended Data Figure 7 | Response function similarity. Panels **a** to **n** show similarity score (correlation) across all possible combinations of GPP dataset and integration time by covariate. Each panel gives the covariate name and median of all off-diagonal scores. Integration time and GPP dataset are abbreviated in the y-axis labels as number–letter couplets (the number indicates the integration time in months, the letter indicates the GPP dataset (M, MsTMIP; F, upscaled FLUXNET; R, remotely sensed MODIS); for example, 12R indicates MODIS using a 12-month SPEI integration time. Panel **o** gives the cumulative distribution function of all ($n = 924$) off-diagonal scores. Thin horizontal red lines show, from bottom to top, the 10th, 25th and 50th percentiles. Over 75% of all similarity scores are greater than 0.70, with an overall median of 0.90; only 7% of all scores are negative.



Extended Data Figure 8 | Response function similarity. Each colourmap pixel shows the similarity score (correlation) between the mirroring and fixed window approaches to calculate the pre-drought baseline across the common time frame for all three datasets (2000 to 2008). **a**, 3-month fixed window (overall similarity median of 0.97). **b**, 6-month fixed window (overall similarity median of 0.96). Integration time and GPP dataset are abbreviated as number–letter couplets as in Extended Data Fig. 7.

Extended Data Table 1 | Skill of the trained Random Forests algorithm

		Skill			
GPP Data Product	**SPEI**	**Full**	**Common**	**B3**	**B6**
MsTMIP	1-month	0.77	0.86	0.86	0.85
	6-month	0.78	0.86	0.86	0.85
	12-month	0.76	0.83	0.84	0.84
	24-month	0.76	0.81	0.82	0.82
Upscaled FLUXNET	1-month	0.80	0.85	0.85	0.85
	6-month	0.78	0.84	0.84	0.84
	12-month	0.77	0.82	0.83	0.82
	24-month	0.78	0.81	0.81	0.81
MODIS	1-month	0.89	0.89	0.88	0.88
	6-month	0.87	0.87	0.87	0.87
	12-month	0.86	0.86	0.85	0.86
	24-month	0.84	0.84	0.83	0.83

Skill is the proportion of variance explained by the trained Random Forest based on out-of-bag samples, that is, those not used in training for a given tree. Skill is given by experiment as follows: Full, full time period for each dataset; Common, 2000–2008 overlap period for all datasets, B3 and B6; as for Common but using a fixed 3-month and 6-month baseline period for pre-drought GPP, respectively.