Summarizing the state of the terrestrial biosphere in few dimensions

Guido Kraemer[1,2], Gustau Camps-Valls[2], Markus Reichstein[1,3], and Miguel D. Mahecha[1,3]

[1]Max Planck Institute for Biogeochemistry, 07745 Jena, Germany
[2]Image Processing Lab, Universitat de València, 46980 Paterna (València), Spain
[3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103 Leipzig, Germany

Correspondence: Guido Kraemer (gkraemer@bgc-jena.mpg.de)

Abstract. In times of global change, we must closely monitor the state of the planet in order to understand gradual or abrupt changes early on. In fact, each of the Earth's subsystems—i.e. the biosphere, atmosphere, hydrosphere, and cryosphere—can be analyzed from a multitude of data streams. However, since it is very hard to jointly interpret multiple monitoring data streams in parallel, one often aims for some summarizing indicator. Climate indices, for example, summarize the state of atmospheric circulation in a region. Although such approaches are also used in other fields of science, they are rarely used to describe land surface dynamics. Here, we propose a robust method to create indicators for the terrestrial biosphere using principal component analysis based on a high-dimensional set of relevant global data streams. The concept was tested using 12 explanatory variables representing the biophysical states of ecosystems and land-atmosphere water, energy, and carbon fluxes. We find that two indicators account for 73% of the variance of the state of the biosphere in space and time. While the first indicator summarizes productivity patterns, the second indicator summarizes variables representing water and energy availability. Anomalies in the indicators clearly identify extreme events, such as the Amazon droughts (2005 and 2010) and the Russian heatwave (2010), they also allow us to interpret the impacts of these events. The indicators also reveal changes in the seasonal cycle, e.g. increasing seasonal amplitudes of productivity in agricultural areas and in arctic regions. We assume that this generic approach has great potential for the analysis of land-surface dynamics from observational or model data.

1 Introduction

Today, humanity faces the global impacts of land use and land cover change (Song et al., 2018), global warming (IPCC, 2014), and associated losses of biodiversity (IPBES, 2019), to only mention the most prominent transformations. Over the past decades, new satellite missions (Berger et al., 2012), along with the continuous collection of more ground based measurements (Baldocchi et al., 2001; Baldocchi, 2008), and the generation of model data to anticipate future dynamics in the Earth system (Eyring et al., 2016) have increased our capacity to monitor the Earth's surface enormously. However, there are still large knowledge gaps limiting our capacity to monitor and understand the current changes of the Earth system (Rockström et al., 2009). Regional trends of vegetation greening and browning that have been attributed to fertilization effects on the one hand, and long-term climate change on the other, need to be understood (de Jong et al., 2011; Zhu et al., 2016; Wright et al., 2017). Changes in the seasonal cycles of primary production, e.g. decreased seasonal amplitudes in "cold" ecosystems due to warmer winters (Stine et al., 2009) or increased seasonal amplitude in agricultural areas due to the so called "green revolution", are

expected (Zeng et al., 2014). In general, phenological patterns are changing in the wake of climate change, leading primarily to changes in the onset of spring (Schwartz, 1998; Parmesan, 2006). Additionally, we are confronted with cascading effects induced by today's increasing frequencies and magnitudes of extreme events (Barriopedro et al., 2011; Reichstein et al., 2013) which are yet to be fully understood (Flach et al., 2018; Sippel et al., 2018). The question is, how to uncover and summarize

30   effects of this kind from the wealth of available global data streams? Do we need to develop specific solutions for every observed phenomenon or can we develop a single approach to uncover a wide variety of phenomena.

Extracting the dominant dynamics from high-dimensional observations is a well-known problem in many disciplines. In climate science, for example, it is common to summarize atmospheric states using Empirical Orthogonal Functions (EOF), also known as Principal Component Analysis (PCA; Pearson, 1901). The rationale is that dimensionality reduction retains

35   the main data features, but makes them better accessible to intuitive interpretations. One of the most prominent examples is the description of the El Niño Southern Oscillation (ENSO) dynamics in the multivariate ENSO index (MEI; Wolter and Timlin, 2011), an indicator describing the state of the regional circulation patterns at a certain point in time. The MEI is a very successful index that can be easily interpreted and used in a variety of ways, most basically it provides a measure for the intensity and duration of the different quasi-cyclic ENSO events but it can also be associated with its characteristic impacts:

40   E.g. seasonal warming, changes in seasonal temperatures and overall dryness in the Pacific Northwest of the United States (Abatzoglou et al., 2014), drought related fires in the Brazilian Amazon (Aragão et al., 2018), and yield anomalies (Najafi et al., 2019).

In plant ecology, indicators based on dimensionality reduction methods are used to describe changes to species assemblages along unknown gradients (Legendre and Legendre, 1998; Mahecha et al., 2007a). The emerging gradients can be interpreted

45   using additional environmental constraints, or based on internal plant community dynamics (van der Maaten et al., 2012). It is also common to compress satellite based Earth Observations via dimensionality reduction to get a notion of the underlying dynamics of terrestrial ecosystems. For instance, Ivits et al. (2014) showed that one can understand the impacts of droughts and heatwaves based on a compressed view of the relevant vegetation indices. In general, dimensionality reduction is the method of choice to compress high-dimensional observations in a few (ideally) independent components with little loss of information.

50   Understanding changes in land-atmosphere interactions is a complex problem, as all aforementioned changes may occur and interact: Land cover change may alter biophysical properties of the land surface such as albedo with consequences for the energy balance. Long-term trends in temperature, water availability, or fertilization may impact productivity patterns and biogeochemical processes. In fact, these land surface dynamics have multidimensional implications and require monitoring of biophysical state variables such as leaf area index, albedo, etc., as well as associated land-atmosphere fluxes of carbon, water,

55   and energy.

Here, we aim to summarize these high-dimensional surface dynamics and make them accessible to subsequent interpretations. Specifically, we seek a set of independent, yet comprehensive, state indicators. We want to have a set of very few indicators that represent the most dominant features of the above described temporal ecosystem dynamics. The approach should also give an idea of the general complexity contained in the available data streams. If more than a single indicator is required

60   to describe land surface dynamics accurately, then these indicators shall describe very different aspects. While one indicator

may describe global patterns of change, others could be only relevant in certain regions, for certain types of ecosystems, or for specific types of impacts. The indicators shall have a number of desirable properties: (1) Representing the overall state of parts comprising the system in space and time. (2) Carrying sufficient information to allow for reconstructing the original observations faithfully from these indicators. (3) Being of much lower dimensionality than the number of observed variables.

65 (4) Allowing intuitive interpretations.

2 Methods

2.1 Data

Table 1. Variables used describing the biosphere, for a description of the variables, see appendix F.

Variable	Details	Source
Black Sky Albedo	Directional reflectance	Muller et al. (2011)
Evaporation	$[\mathrm{mm\,day^{-1}}]$	Martens et al. (2017)
Evaporative Stress	Modeled water stress	Martens et al. (2017)
fAPAR	fraction of absorbed photosynthetically active radiation	Disney et al. (2016)
Gross Primary Productivity (GPP)	$[\mathrm{gCm^{-2}day^{-1}}]$	Tramontana et al. (2016)
Latent energy (LE)	$[\mathrm{Wm^{-2}}]$	Tramontana et al. (2016)
Net Ecosystem Exchange (NEE)	$[\mathrm{gCm^{-2}day^{-1}}]$	Tramontana et al. (2016)
Root-Zone Soil Moisture	$[\mathrm{m^3m^{-3}}]$	Martens et al. (2017)
Sensible Heat (H)	$[\mathrm{Wm^{-2}}]$	Tramontana et al. (2016)
Surface Soil Moisture	$[\mathrm{mm^3mm^{-3}}]$	Martens et al. (2017)
Terrestrial Ecosystem Respiration (TER)	$[\mathrm{gCm^{-2}day^{-1}}]$	Tramontana et al. (2016)
White Sky Albedo	Diffuse reflectance	Muller et al. (2011)

Table 1 gives an overview of the data streams used in this analysis (for a more detailed description in appendix F). For an effective joint analysis of more than a single variable, the variables have to be harmonized and brought to a single grid in space

70 and time. The Earth System Data Lab (ESDL; www.earthsystemdatalab.net) curates a comprehensive set of data streams to describe multiple facets of the terrestrial biosphere and associated climate system. The data streams are harmonized as analysis ready data on a common spatiotemporal grid (0.25° in space and 8 days in time), forming a 4d hypercube, which we call a *data cube*. The ESDL not only curates Earth system data, but also comes with a toolbox to analyze this data efficiently.

In this study, each variable was normalized globally to zero mean and unit variance to account for the differences in scales.

75 Because the area of the pixel changes with latitude, the pixels were weighted according to the represented surface area.

2.2 Dimensionality Reduction with PCA

As a method for dimensionality reduction, we used a modified principal component analysis (PCA) to summarize the information contained in the observed variables. PCA transforms the set of d centered and, in this case, standardized variables into a subset of p ($1 \leq p \leq d$) principal components (PCs). Each component is uncorrelated with the other components, while the
80 first PCs explain the largest fraction of variance in the data.

The data streams consist of d observed variables at the same time and location. Each observation is defined in a d-dimensional space, $\mathbf{x}_i \in \mathbb{R}^d$, and we define the dataset by collecting all samples in the matrix $\mathbf{X} = [\mathbf{x}_1 | \cdots | \mathbf{x}_n] \in \mathbb{R}^{d \times n}$. The observations are repeated in space and time and lie on a grid of lat $\times$ lon $\times$ time, which in our case are $n = \#\text{lat} \times \#\text{lon} \times \#\text{time} = 720 \times 1440 \times$ 506, where $\#$ denotes the length of the dimension. Note that the actual number of observations was lower, $n = 106360156$,
85 because we considered land points only and removed missing values.

To derive the PCs, we used an eigendecomposition of the covariance matrix,

$$\mathbf{Q} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \in \mathbb{R}^{d \times d}.$$

The covariance matrix, in this case, is equal to the correlation matrix because we standardized the variables to unit variance. $\mathbf{\Lambda}$ is a diagonal matrix with the the eigenvalues, $\lambda_1, \dots, \lambda_d$, in the diagonal in decreasing order and $\mathbf{V} \in \mathbb{R}^{d \times d}$, the matrix with
90 the corresponding eigenvectors in columns. $\mathbf{V}$ can project the new incoming input data $\mathbf{x}_i$ (centered and standardized) onto the PCs:

$$\mathbf{y}_i = \mathbf{V}^T \mathbf{x}_i \in \mathbb{R}^d, \tag{1}$$

where $\mathbf{y}_i$ is the projection of the observation $\mathbf{x}_i$ onto the d PCs.

Because the observations were centered, the covariance matrix can be calculated by using a simple formula

$$95 \quad \mathbf{Q} = \frac{1}{n-1} \mathbf{X} \mathbf{X}^T = \frac{1}{n-1} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T. \tag{2}$$

Given that the data cube lies on a regular $0.25°$ grid, estimating $\mathbf{Q}$ as above would lead to overestimating the influence of dynamics in high latitudes compared to lower latitudes where each data point represent largers areas. Hence, we used a weighted approach to calculate the covariance matrix

$$\mathbf{Q} = \frac{1}{w} \sum_{i=1}^{n} w_i \mathbf{x}_i \mathbf{x}_i^T, \tag{3}$$

100 where $w_i = \cos(\text{lat}_i)$ and lat_i is the latitude of observation i. $w = \sum_{i=1}^{n} w_i$ is the total weight, and n the total number of observations. Equation (3) has the additional property that it can be computed sequentially on very big data sets, such as our Earth system data cube, by consecutively adding observations.

The actual calculation of the covariance matrix was more complicated, because summing up many floating-point numbers one by one can lead to large inaccuracies due to precision issues of floating-point numbers and instabilities of the naive

105 algorithm (Higham, 1993; the same goes for the implementations of the `sum` function in most software used for numerical computing). Here, the Julia package `WeightedOnlineStats.jl`[1] (implemented by the first author of this paper) is used, which uses numerically stable algorithms for summation, higher precision numbers, and a map-reduce scheme that further minimizes floating point errors.

The canonical measure of the quality of a PCA is the fraction of explained variance, calculated as

$$\frac{\lambda_i}{\sum_{i=1}^{d} \lambda_i}, \tag{4}$$

where λ_i is the i-th eigenvalue of the covariance matrix $\mathbf{Q}$. To get a more complete measure of the accuracy of the PCA, we used the "reconstruction error" in addition to the fraction of explained variance. PCA allows a simple projection of an observation onto the first p PCs and a consecutive reconstruction of the observations from this p-dimensional projection. This is achieved by

$$\mathbf{Y}_p = \mathbf{V}_p^T \mathbf{X} \in \mathbb{R}^{p \times n} \text{ and } \mathbf{X}_p = \mathbf{V}_p \mathbf{Y}_p \in \mathbb{R}^{d \times n}, \tag{5}$$

where $\mathbf{Y}_p$ is the projection on the first p PCs, $\mathbf{V}_p$ the matrix consisting of the eigenvectors belonging to the p largest eigenvalues, and $\mathbf{X}_p$ the observations reconstructed from the first p PCs.

The reconstruction error, $\mathbf{e}_i$, was calculated for every point, $\mathbf{x}_i$ in the space-time-domain based on the reconstructions from the first p principal components:

$$\mathbf{e}_i = \mathbf{V}_p \mathbf{V}_p^T \mathbf{x}_i - \mathbf{x}_i \in \mathbb{R}^d. \tag{6}$$

As this error is explicit in space, time and variable, it allows for disentangling the contribution of each of these domains to the total error. This can be achieved by estimating e.g. the (weighed) mean square error

$$\text{MSE} = \frac{1}{w} \sum_i w_i \mathbf{e}_i^2 \tag{7}$$

where $w_i = \cos(\text{lat}_i)$, lat_i the latitude of $\mathbf{e}_i$, $w = \sum_i w_i$ the total weight. Therefore, this approach can give a better insight into

125 the compositions of the error than a single global error estimate based on the eigenvalues.

2.3 Pixel-wise analyses of time series

When calculating slopes using measured data, ordinary least squares regression is not the optimal choice because outliers can significantly change the estimator. One possible solution is using the Theil-Sen estimator which is robust to up to 29.3% of outliers (Theil, 1950; Sen, 1968). The calculation of the estimator consists simply on computing the median of the slopes

130 spanned by all possible pairs of points

$$\text{slope}_{ij} = \frac{z_i - z_j}{t_i - t_j}, \tag{8}$$

---

[1]DOI: 10.5281/zenodo.3360311, repository: https://github.com/gdkrmr/WeightedOnlineStats.jl/

where z_i is the value of the response variable at time step i and t_i the time at time step i. In our experiments, we computed the slopes separately per pixel and principal component where time is the predictor and the value of the principal component is the response variable.

135   To test the slopes for significance, we used the Mann-Kendall statistics (Mann, 1945; Kendall, 1970) and adjusted the resulting p-values with the Benjamini-Hochberg method to control for the false discovery rate (Benjamini and Hochberg, 1995). Slopes with an adjusted $p < 0.05$ were deemed significant.

For the calculation of the number of breakpoints, the generalized fluctuation test framework (Kuan and Hornik, 1995) was used to test for the presence of breakpoints. The framework uses recursive residuals (Brown et al., 1975), and a breakpoint is

140   identified when the mean of the recursive residuals deviates from zero. We used the implementation in Zeileis et al. (2002). For practical reasons, here we only focus on the biggest breakpoint.

Hysteresis was calculated as the area, A, inside the polygon formed by the mean seasonal cycle of PC1 and PC2

$$A = \frac{1}{2} \sum_{i=1}^{n} x_i(y_{i+1} - y_{i-1}),$$

where $n = 46$, the number of time steps in a year, x_i and y_i the mean seasonal cycle of PC_1 and PC_2 at time step i, respectively.

145   The polygon is circular, i.e. the indices wrap around the edges of the polygon. This formula gives the actual area of the polygon if it is non-self-intersecting and the vertices run counterclockwise. If the vertices run clockwise, the area is negative. If the polygon is shaped as an 8, the clockwise and counterclockwise parts will cancel each other (partially) out, e.g. the green trajectory in fig. 2b. Trajectories that cover a larger range will also tend to have larger areas.

3   Results and Discussion

150   ### 3.1   The PCA embedding

Figure 1a shows the explained fraction of variance (Eq. 4) for the global PCA based on the entire data cube. We see that the first two components explain 73% of the variance from the 12 variables; additional components contribute little $< 10\%$ explained variance each. This results in a "knee" at component 2, which suggests that two indicators are sufficient to capture the major global dynamics of the terrestrial land surface and therefore we focus on these components in the following analyses.

155   Using PCA as a method for dimensionality reduction means that we are assuming linear relations among features. A nonlinear method could possibly be more efficient in reducing the number of variables, but would also have significant disadvantages. In particular: nonlinear methods typically require tuning of specific parameters, objective criteria are often lacking, a proper weighting of observations is difficult, and it is harder to interpret the resulting indicators due to their nonlinear nature (Kraemer et al., 2018). The salient feature of PCA is that an inverse projection is well defined and allows for a deeper inspection of the

160   errors, which is not the case for nonlinear methods due to the pre-imaging problem (Mika et al., 1999; Arenas-Garcia et al., 2013).

The contributions of each variable to the resulting indicators can be understood from the rotation matrix (Eq. 1, fig. 1b). The first PC summarizes variables that are closely related to vegetation primary productivity (GPP, LE, NEE, fAPAR). These
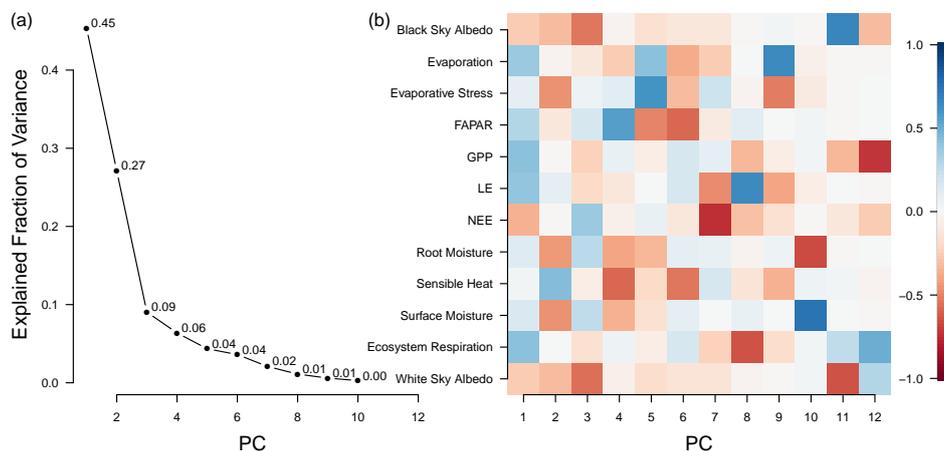
Figure 1. (a) Fraction of explained variance of the PCA by component. Components three and higher do not conrtibute much to total variance. (b) Rotation matrix of the global PCA model, axis one describes primary productivity related variables, axis two describe water availability.

variables are related because they are all directly related to primary productivity. The energy for photosynthesis comes from solar radiation, an indicator for the fraction of light used for photosynthesis is given by fAPAR. Photosynthesis fixes carbon from gaseous CO_2 producing sugars to maintain the metabolism of plants, this total uptake of CO_2 is reflected in GPP. However, the CO_2 uptake is closely related to water consumption. The actual uplift of water within the plant is not only essential to enable photosynthesis, but also drives the transport of nutrients from the roots and is ultimately reflected in transpiration—together with evaporation from soil surfaces one can observe the integrated latent energy needed for the phase transition (LE). However, ecosystems also respire and hence CO_2 is produced by plants in energy consuming processes as well as by the decomposition of dead organic materials via soil microbes and other heterotrophic organisms. This total respiration can be observed as terrestrial ecosystem respiration (TER). The difference between GPP and TER is the net ecosystem exchange (NEE) rate of CO_2 between ecosystems and the atmosphere (Chapin et al., 2006).

On the second axis we observe variables that are related to the surface hydrology of ecosystems. Surface moisture, evaporative stress, root-zone soil moisture, and sensible heat, are all essential indicators for the state of plant available water. While surface moisture is a rather direct measure, evaporative stress is a modeled quantity summarizing the level of plant stress, a value of zero means that there is no water available for transpiration, while a value of one means that transpiration the potential transpiration (Martens et al., 2017). Root-zone soil moisture is the moisture content of the root zone in the soil, the moisture directly available for root uptake. If this quantity is below the wilting point, there is no water available for uptake by the plants. Sensible heat is the exchange of energy by a change of temperature, if there is enough water available, then most of the surface heat will be lost due to evaporation (latent heat), with decreasing water availability more of the surface heat will be lost due to latent heat, making this also an indicator of dryness.

The first two principal components form a triangle (gray background in fig. 2). On one edge of the first principal component we find ecosystems in a high state of primary productivity (high values of GPP, fAPAR, LE, TER, and evaporation), mostly

185   limited by radiation, while on the lower end of the principal component one we find states of low productivity. Ecosystem states of low productivity are further separated by the second principal component: Low productivity can coincide with radiation limitation (the negative extreme of the second principal component) as seen in the lower left corner of the distribution in fig. 2a and b or with water limitation (the positive extreme of the second principal component, the upper left corner in fig. 2a). This pattern reflects the two essential global limitations of GPP in terrestrial ecosystems (Anav et al., 2015).

190   Both axes form the space in which most of the variability of ecosystems takes place. Axis one describes productivity and axis two the limiting factors to productivity. Therefore, we can see that most ecosystems with high values on axis one (a high productivity) are at the approximate center of axis two. When ecosystems are found outside the center of axis two, they have lower values on axis one (lower productivity) because they are limited by water or temperature (see fig. 2b).

Heat transfer from the surface into the atmosphere can happen either by radiative transfer (sensible heat) or evaporation
195   (latent heat). Their ratio is the "Bowen ratio", $B = \frac{\text{LE}}{\text{H}}$, (Bowen, 1926; see also fig. C1), if there is enough moisture, then most of the energy will be dissipated by evaporation, $B < 1$, resulting in a high latent heat flux, but when the surface does not contain much moisture the transfer by latent heat will be low and most of the incoming energy has to be dissipated via sensible heat, $B > 1$. In higher latitudes, there is not much incoming radiation and the temperatures are low, therefore there is not much energy to be dissipated and both heat fluxes are low. A high sensible heat flux is an indicator for water limitation and both low
200   sensible and latent heat flux are indicators for ecosystems that are limited by low temperatures and low amounts of incoming radiation. We can see that the bowen ratio embedds well into the space spanned by the first two PCs.

3.2   Trajectories

The principal components may be used to summarize the movement of a spatiotemporal pixel in variable space, so that they represent the current state of the ecosystem at a certain location in space and time (fig. 2a) or time of year of the mean seasonal
205   cycle of the pixel (fig. 2b).

Because the underlying data are 8-daily resolution, we can observe the seasonal variability and find that the seasonal cycles of very different regions of the world can substantially overlap. We see that very different ecosystems may reach very similar states in the course of the season, even though their seasonal dynamics are very distinct. For instance, mid-latitude areas (blue trajectory in fig. 2) show very similar characteristics to tropical forests during their peak growing season because their
210   patterns of productivity and water availability are similar (see also SI fig. D1). Likewise, many high latitude areas show similar characteristics to mid-latitude areas during winter (low latent and sensible energy release as well as low GPP) and many dry areas such as deserts show similar characteristics to areas with a pronounced dry season, e.g. the Mediterranean.

Ecosystems states shift from limitation to growth during the year (fig. 2b, e.g. Forkel et al., 2015). For example, the orange trajectory in fig. 2, an area close to Moscow, shifts from a temperature limited state in winter to a state of very high productivity
215   during summer. Other ecosystems remain in a single limitation state with only slight shifts, such as the red trajectory in fig. 2. In the corner of maximum productivity of the distribution, we find tropical forests characterized by a very shallow seasonality. We also observe that very different ecosystems can have very similar characteristics during their peak growing season, e.g.
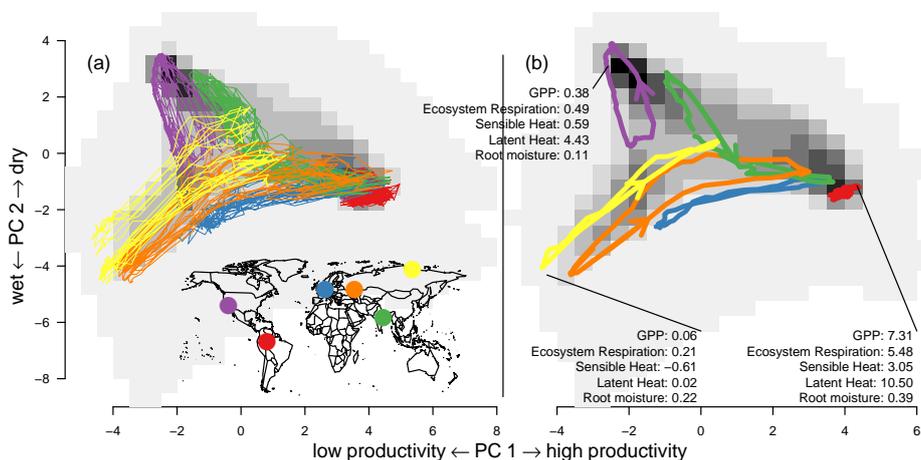
Figure 2. Trajectories of some points (colored lines) and the area weighted density over principal components one and two (the gray background shading shows the density) for (a) the raw trajectories and (b) the mean seasonal cycle. The trajectories were chosen to fill a large area in the space of the first two principal components. Some of the trajectories in (b) have an arrow indicating the direction. The numbers illustrate the value of some variables, for units, see tab. 1. Description of the points: Red: Tropical Rainforest, 67.625°W, 2.625°S; Blue: Maritime climate, 7.375°E, 52.375°N; Green: Monsoon climate, 82.375°E, 22.375°N; Purple: Subtropical, 117.625°W, 34.875°N; Orange: Continental climate, 44.875°E, 52.375°N; Yellow: Arctic climate, 119.875°E, 72.375°N;

green (located in north east India), blue (north west Germany), and orange (located close to Moscow) trajectories have very similar characteristics during peak growing season compared to the red trajectory.

3.2.1 The Mean Seasonal Cycle of Trajectories

As with ordinary variables, we can compute the Mean Seasonal Cycle (MSC) of the principal components summarizing the average state of the ecosystem during the course of the year. Removing year-to-year variability and long-term trends reveals a general characterization of the local ecosystem (cf. fig. 2b).

The global main pattern of the first principal component follows the productivity cycles during summer and winter (3, left column) of the northern hemisphere, with positive values (high productivity, green) during summer and negative values (low productivity, brown) during winter. The tropics show high productivity all year. The global pattern shows the well known green wave (Schwartz, 1994, 1998) because the first dimension integrates over all variables that correlate with plant productivity.

The second principal component (fig. 3, right column) tracks water deficiency: red and light red values indicate water deficiency, light blue values excess water, and dark blue water growth limitation due to cold. Areas which are temperature limited during winter but have a growing season during summer, such as boreal forests, change from dark blue in winter to light blue during the growing season. Areas which have low productivity during a dry season change their coloring from red to light red during the growing season, e.g the north west of Mexico/south west of the United States.
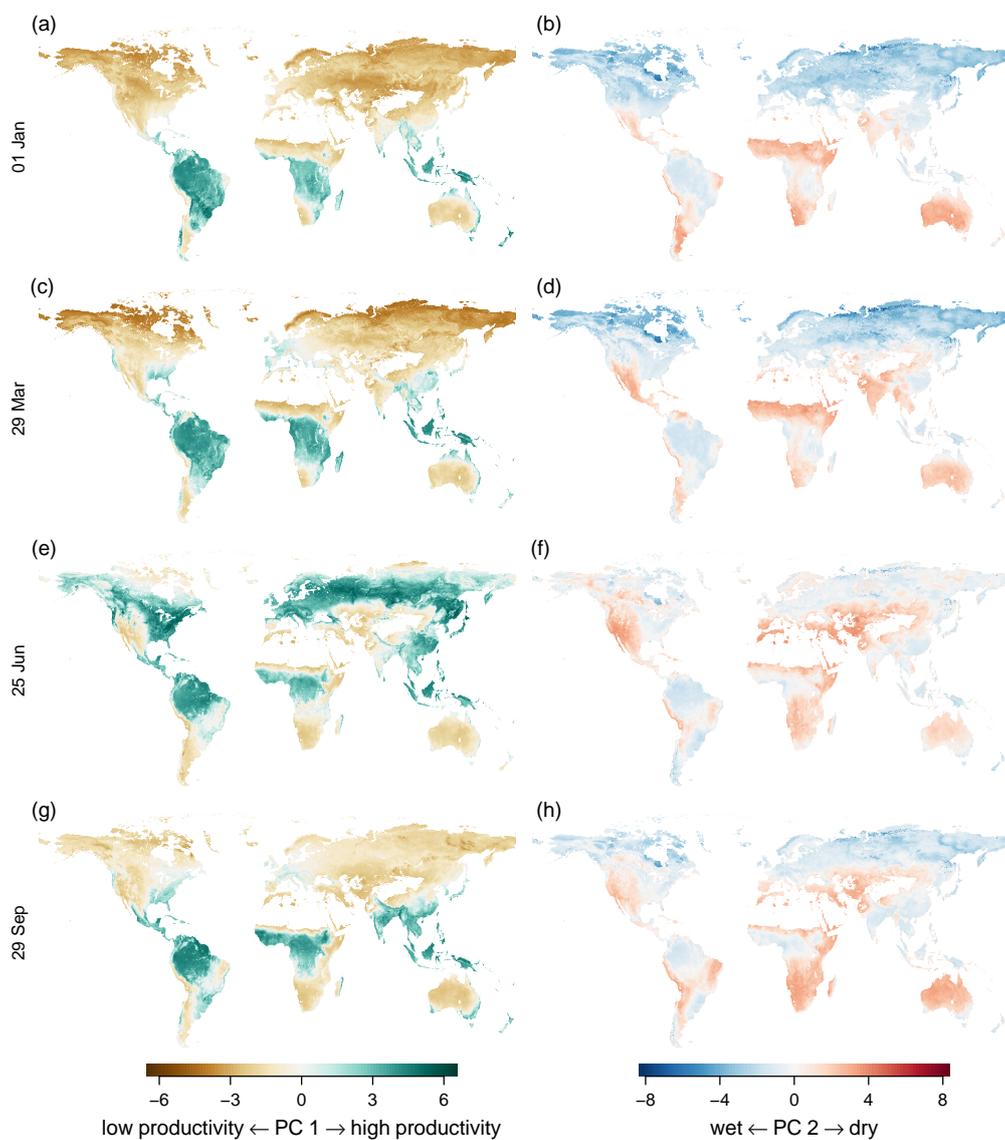
Figure 3. Mean seasonal cycle of the first principal component during the year. Left column: first principal component. Right column: second Principal Component. Rows from top to bottom: equally spaced intervals during the year.

Observing the mean seasonal cycle of the principal components gives us a tool to characterize ecosystems and may also serve as a basis for further analysis, such as a global comparison of ecosystems (Metzger et al., 2013; Mahecha et al., 2017).

3.2.2 Hysteresis
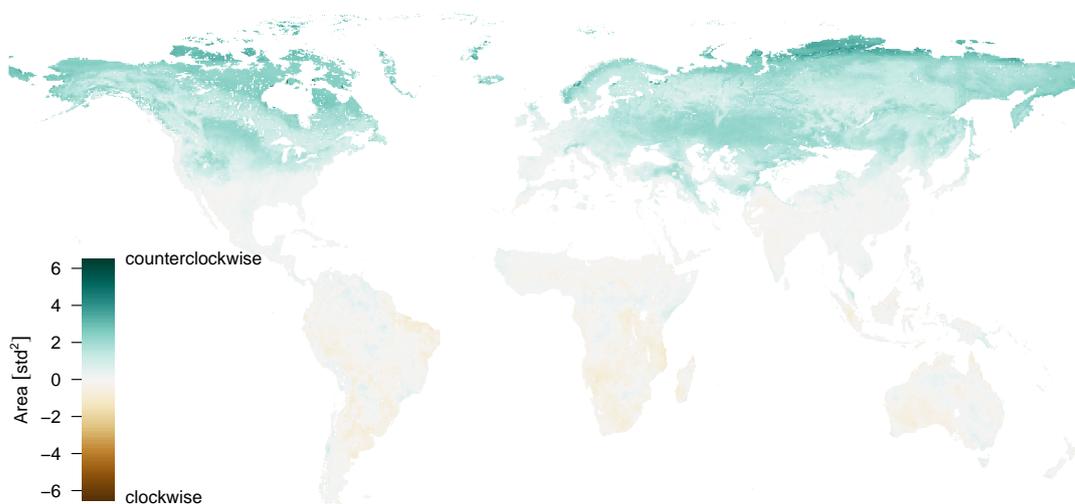
Figure 4. The area inside the mean seasonal cycle of PC1 and PC2, area is positive if direction is counterclockwise and negative if the direction is clockwise. We can observe that most of the trajectories need a pronounced seasonal cycle to show a hysteresis effect. E.g. the green trajectory of 2b does not show up here, because it is shaped like an 8 and therefore the clockwise and counterclockwise parts cancel each other out.

Hysteresis in ecology means that the pathways $A \rightarrow B$ and $B \rightarrow A$ between stable states A and B can be different (Beisner et al., 2003. These alternative paths arise from the ecosystem tracking seasonal changes in the environmental condition, e.g. summer–winter or dry–rainy seasons (fig. 2b)).

Hysteresis is a common occurrence in ecology, e.g. in community ecology it is often cited as the reason why communities may not recover after a disturbance, it is usually attributed to memory and lag effects (Folke et al., 2004; Blonder et al., 2017; Renner et al., 2019). For instance, a hysteresis loop can be found when plotting soil respiration against soil temperature (Tang et al., 2005). The sensitivity of soil respiration to soil temperature changes seasonally due to changing soil moisture and photosynthesis (by supplying carbon to rhyzosphere) producing a seasonally changing hysteresis effect (Gaumont-Guay et al., 2006; Richardson et al., 2006; Zhang et al., 2018). Biological variables also show a hysteresis effect in their relations with atmospheric variables, e.g. Mahecha et al. (2007b) found a hysteresis effect between seasonal NEE, temperature, and a number of other ecosystem and climate related variables.

Looking at some mean seasonal cycles of trajectories, e.g. the orange trajectory (area close to Moscow) in fig. 2b shows that the paths between maximum and minimum productivity can be very different, in contrast to the blue trajectory located in the

north west of Germany which also has a very pronounced yearly cycle but shows no such effect. The trajectories that show a
250   more pronounced hysteresis effect seem to have pronounced growing, dry, and wet seasons and therefore shift their limitations
more strongly during the year, i.e. the moisture reserves deplete during growing season and therefore the return path has higher
values on the second principal component. We can also see that most trajectories that show hysteresis turn counterclockwise
for the same reason (see fig. 4). Usually plant growth starts when there is enough water available (low values on component
2), leading to increasing values on the first component. At the end of the growing season water resources deplete (increasing
255   values on component 2) and productivity decreases (decreasing values on component 1).

3.2.3   Anomalies of the Trajectories

The deviation of the trajectories from their mean seasonal cycle should reveal anomalies and extreme events. These anomalies
have a directional component and can be therefore be interpreted the same way as the original PCs which contain information
of the underlying variables that were affected. In this sense, one can infer the state of the ecosystem during an anomalous state.
260   For instance the well-known Russian heatwave in summer 2010 appears in fig. 5 as a dark brown spot in the southern part of the
affected area, indicating lower productivity and as a thin green line in the northern parts, indicating an increased productivity.
This confirms earlier reports that only the southern agricultural ecosystems were negatively affected by the heatwave, while
the northern predominantly forest ecosystems rather benefited from the heatwave in terms of primary productivity (Flach et al.,
2018).

265   Another example of an extreme event that we find in the PCs is the very wet November rainy season of 2006 in the Horn of
Africa after a very dry rainy season in the previous year. This event was reported to bring heavy rainfall and flooding events
which caused an emergency for the local population but also an increased ecosystem productivity (Nicholson, 2014). The
rainfall event appears as green and blue spots in fig. 5, preceded by the drought events which appear as red and brown spots.

Fig. 5e and 5f also show the strong drought events in the Amazon, particularly the droughts of 2005 and 2010 (Doughty
270   et al., 2015; Feldpausch et al., 2016) appear strongly north and south of the Amazon basin. The central Amazon basin does
not show these strong events, because the observable response of the ecosystem was buffered due to the large water storage
capacity in the central Amazon basin.

3.2.4   Single Trajectories

Observing single trajectories can give insight into past events that happen at a certain place, such as extreme events or permanent
275   changes in ecosystems. The creation of trajectories is an old method used by ecologists, mostly on species assembly data of
local communities, to observe how the composition changes over time (e.g. Legendre et al., 1984; Ardisson et al., 1990). In this
context, we observe how the states of the ecosystems inside the grid-cell shift over time, which comprises a much larger area
than a local community but is probably also less sensitive to very localized impacts than a community level analysis. One of
the main differences of the method applied here to the classical ecological indicators is that the trajectories observed here are
280   embedded into the space spanned by a single global PCA and therefore we can compare a much broader range of ecosystems
directly.
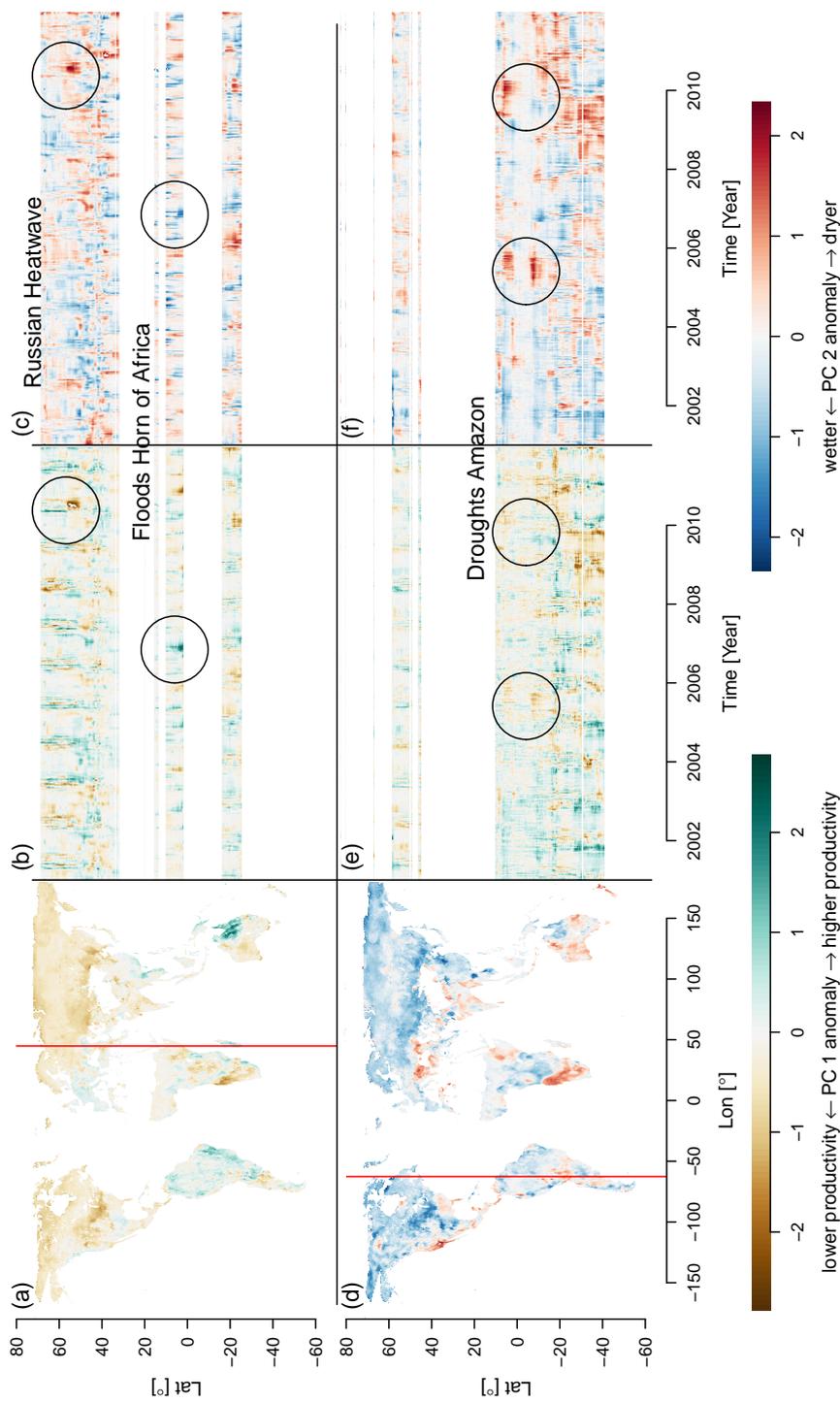
Figure 5. Anomalies of the first two principal components; Brown–green contrast shows the anomalies on PC1, a relative low productivity or greening respectively. Blue–green contrast shows the anomalies on PC2, a relative wetness or dryness respectively. (a) Map showing the PC1 anomalies on the 1/1/2001. (b) and (c) show longitudinal cuts of PC1 and PC2 at the red vertical line in sub-figure (a) respectively. The effects of of the drought on the Horn of Africa (2006) and the Russian heatwave (2010) are highlighted by circles. (d) Map showing the PC2 anomalies on the 1/1/2001. (e) and (f) show longitudinal cuts of PC1 and PC2 at the red vertical line in sub-figure (d) respectively. Strong droughts in the Amazon during 2005 and 2010 can be observed as large red spots on the fringes of the Amazon basin (highlighted by circles).
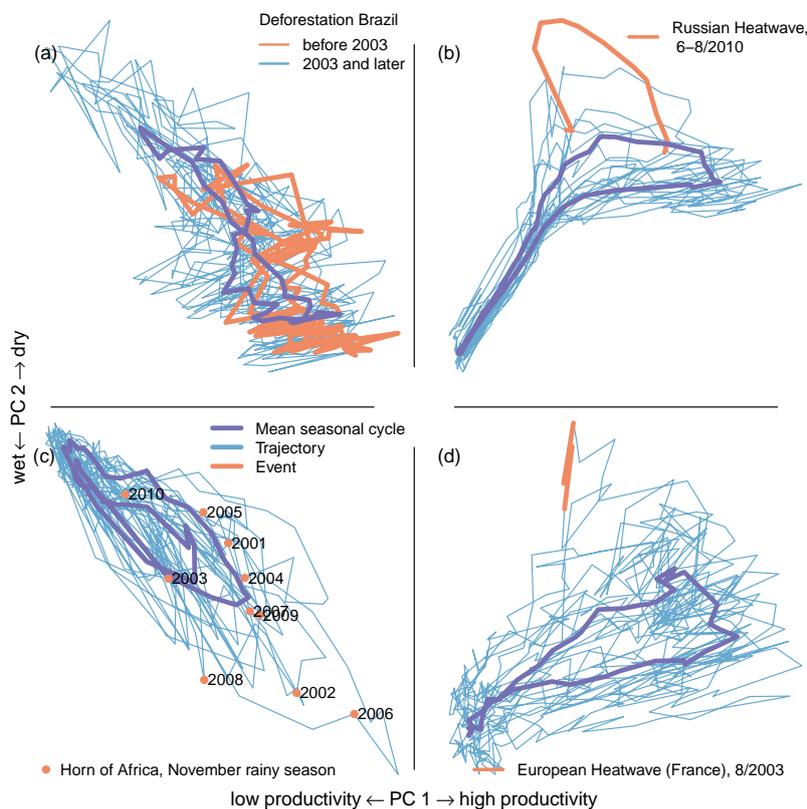
Figure 6. Trajectories of the first two Principal Components for single pixels. (a) Deforestation increases the seasonal amplitude of the first two PCs (Brazilian rainforest, 9.5°S 63.5°W). (b) The heatwave is clearly visible in the trajectory (red, Russian heatwave, summer 2010, 56°N 45.5°E). (c) Rainfall in the short raining season (November/December) influences agricultural yield and can cause flooding (extreme flooding after drought, 11/2006, 3°N 45.5°E). (d) European heatwave in Summer 2003 was one of the strongest on record (France, 47.2°N 3.8°E). The mean seasonal cycle of the trajecotries is shown in purple.

Figure 6a shows an area in the Brazilian Amazon in Rondônia (9.5°S, 63.5°W) which has been affected by large scale land use change and deforestation. It can be seen that the seasonal amplitude increases strongly after the beginning of 2003. Reasons for this increased amplitude could lie in any of the following reasons or a combination of them: Deforestation decreases water 285 storage capability and dries out soils causing larger variability in ecosystem productivity. Therefore, during periods of no rain, large scale deforestation can cause a shift in local scale circulation patterns causing lower local precipitation (Khanna et al., 2017). Crop growth and harvest causes an increased amplitude in the cycle of productivity. An analysis of the trajectory can point to the nature of the change, however finding the exact causes for the change requires a deeper analysis.

Figure 6b shows the deviation of the trajectory during the Russian heatwave (red line) in an area east of Moscow (56°N 290 45.5°E). In the southern grass- and croplands, the heatwave caused the productivity to drop significantly during summer due to a depletion of soil moisture. In the northern forested parts affected, the heatwave caused an increase in ecosystem productivity
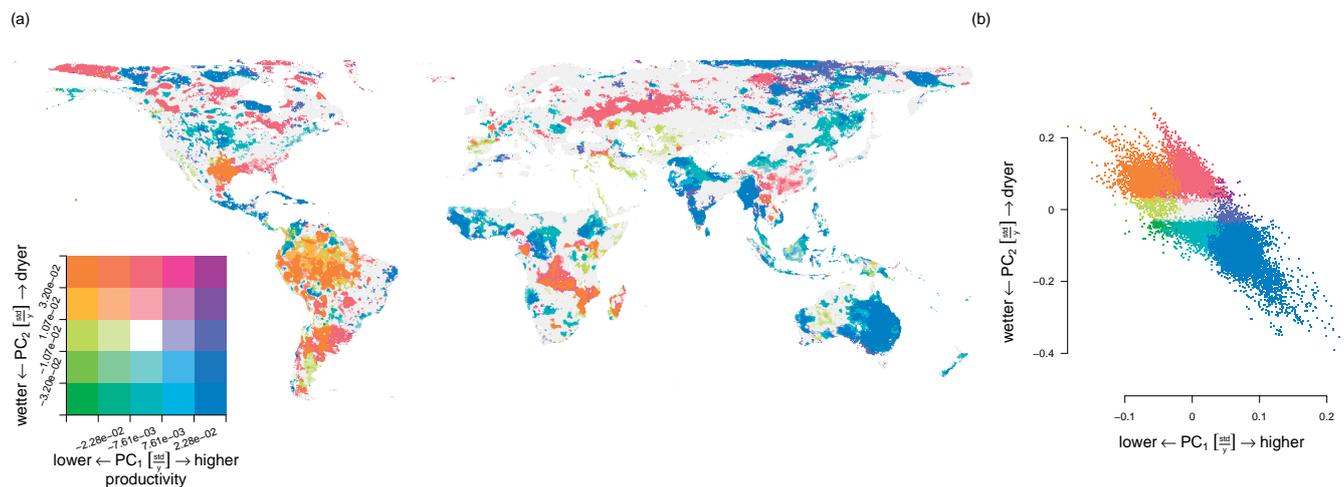
Figure 7. Trends in PC1 and PC2 indicators. Trends were calculated using the Theil-Sen estimator. (a) The spatial distribution of slopes, only significant slopes are shown ($p < 0.05$, Benjamini-Hochberg adjusted). The maximum cutoff for the legend limits was set symmetrically around zero to the maximum absolute value of the 0.1 and 0.9 quantiles. (b) Distribution of spatial points in the space of the first two PCs. The colors correspond to the ones used in (a).

during spring due to higher temperatures combined with sufficient water availability. This shows the compound nature of this extreme event (see fig. 5a and Flach et al. 2018). The analysis of the trajectory points directly towards the different types of extremes and responses that happened in the biosphere during the heatwave.

295   Variability of rainfall during the November rainy season in the Horn of Africa (3°N 45.5°E, fig. 6c) shows the trajectory and points in November of the observed time. The November rain has implications for food security because the second crop season depends on it. In 2006, the rainfall events were unusually strong and caused widespread flooding and disaster but also higher ecosystem productivity (cf. also fig. 5). This was especially devastating because it followed a long drought that caused crop failures. Note also the two rainy seasons in the mean seasonal cycle (purple line if fig. 6c).

300   Figure 6d shows the trajectory during the August 2003 heat wave in Europe (France, 47.2°N 3.8°E). The heatwave was unprecedented and caused large scale environmental, health, and economics losses (Ciais et al., 2005; García-Herrera et al., 2010; Miralles et al., 2014). The 2010 heatwave was stronger than the 2003 heatwave but the strongest parts of the 2010 heatwave were in eastern Europe (cf., fig. 5), while the center of the 2003 heatwave was located in France.

    As we have seen here, observing single trajectories in reduced space can give us important insights into ecosystem states 305 and changes that occurring. While the trajectories can point us towards abnormal events, they can only be the starting points for deeper analysis to understand the details of such state changes.

3.2.5 Trends in Trajectories

The accumulation of CO_2 in the atmosphere should cause an increase in global productivity of plants due to CO_2 fertilization, while large and more frequent droughts and other extremes may counteract this trend. Satellite observations and models have shown that during the last decades the world's ecosystems have greened up during growing seasons. This is explained by CO_2 fertilization, nitrogen deposition, climate change and land cover change (Zhu et al., 2016; Huang et al., 2018; Anav et al., 2015). Tropical forests especially showed strong greening trends during growing season.

To find local trends, we used the Theil-Sen estimator to calculate robust slopes on the trajectories. Figure 7 shows positive and negative trends of the principal components over time. General patterns that can be observed are a positive trend (higher productivity) on the first principal component in the arctic regions and higher temperatures. A large scale dryness trend can be observed across large parts of western Russia. Increasing productivity can also be observed on almost the entire Indian subcontinent and eastern Australia. Negative trends in the first component can also be observed: they are generally smaller and appear in regions around the Amazon and the Congo basin, but also in parts of western Australia. The main difference from previous analyses on the observations presented here is that e.g. Zhu et al. (2016) looked only at trends during the growing season while this analysis uses the entire time series to calculate the slope.

In the Amazon basin, we find a dryness trend accompanied by a decrease in productivity; In the Congo basin, we find a wetness trend and an increasing productivity in the northern parts, while the southern part and woodland south of the Congo basin show a strong dryness trend with decreased productivity. This is different to the findings of Zhou et al. (2014), who found a widespread browning of vegetation in the entire Congo basin for the April-May-June seasons during the period 2000–2012. The finding of Zhou et al. (2014) is not reflected in our data, especially compared to the areas surrounding the Congo basin, we can find only minor browning effects. Inside the basin and our findings are more in line with the global greening (Zhu et al., 2016), which show a browning mostly outside the Congo basin.

Almost the entire Indian subcontinent shows a trend towards higher productivity and an overall wetter climate. The greening trend in India happens mostly over irrigated cropland, however browning trends over natural vegetation have been observed but do not show up in our analysis (Sarmah et al., 2018).

In the Arctic, a general trend towards higher productivity can be observed, vegetation models attribute this general increase in productivity to CO_2 fertilization and climate change. The changes also cause changes to the characteristics of the seasonal cycles (Forkel et al., 2016). Stine et al. (2009) found a decreased seasonal amplitude of surface temperature over norther latitudes due to winter warming.

The seasonal amplitude of atmospheric CO_2 concentrations has been increasing due to climate change causing longer growing seasons and changing vegetation cover in northern ecosystems (Forkel et al., 2016; Graven et al., 2013; Keeling et al., 1996). Therefore we checked for trends in the seasonal amplitude, but because each time series only consists of 11 values (one amplitude per year), after adjusting the p-values for false discovery rate, we could not find a significant slope. However, there were many significant slopes with the unadjusted p-values, see the appendix, fig E1.

4 Conclusions

340

To monitor gradual and abrupt changes in times of global change, we used PCA to construct indicators from a large number of data streams that track ecosystem state in space and time on a global scale. We showed that a large part of the variability of the terrestrial biosphere can be summarized using two indicators. The first emerging indicator represents carbon exchange, while the second indicator shows the availability of water in the ecosystem. The distribution in the space of the first two principal

345 components reflects the general limitations of ecosystem productivity. Ecosystem production can either be limited by water or energy.

The first two indicators can detect many well-known phenomena without analyzing variables separately due to their compound nature. We showed that the indicators are capable of detecting seasonal hysteresis effects in ecosystems, as well as breakpoints, e.g. large scale deforestation. The indicators can also track other changes to the seasonal cycle such as patterns

350 of changes to the seasonal amplitudes and trends in ecosystems. Deviations from the mean seasonal cycle of the trajectories indicate extreme events such as the large scale droughts in the Amazon during 2005 and 2010 and the Russian heat wave of 2010. The events are detected ina a similar fashion as with classical multivariate anomaly detection methods while directly providing information on the underlying variables.

Using compound indicators we gain a high level overview of phenomena in ecosystems and the method therefore provides

355 an interesting tool for analyses where it is required to capture a wide range of phenomena which are not necessarily known a priori. Future research should consider nonlinearities, and work to include different subsystems, such as the atmosphere or the anthroposphere.

Code and data availability. The data are available and can be processed at https://www.earthsystemdatalab.net/index.php/interact/data-lab/, last accessed 28 June 2019. The data can be downloaded from https://www.earthsystemdatalab.net. The code to reproduce this analysis is

360 available at https://github.com/gdkrmr/BioIndicators.jl.

Appendix A: Breakpoints in Trajectories
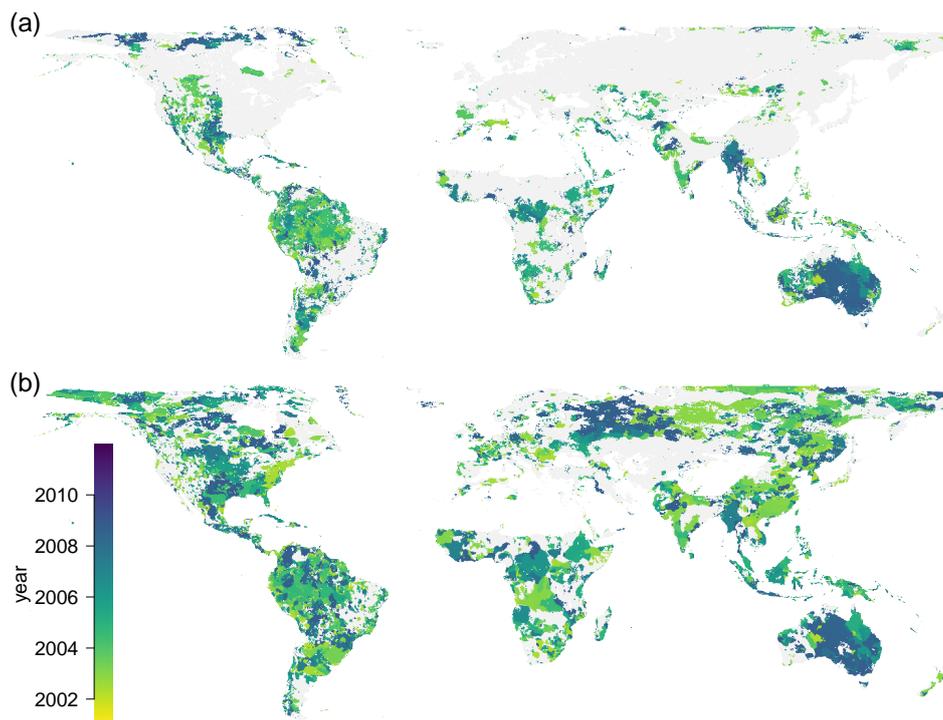
Figure A1. Breakpoint detection, (a) on PC1, (b) on PC2, the color indicates the year of the biggest breakpoint if a significant breakpoint was found, grey if there was no significant breakpoint found.

As the environmental conditions change, due to climate change and human intervention, the local ecosystems may change gradually or abruptly. Detecting these changes is very important for monitoring the impact of climate change and land use change onto the ecosystems. We applied breakpoint detection on the trajectories (fig. A1).

365   Breakpoints on the first component were found in the entire Amazon and the largest breakpoint is dated in the year 2005 during the large drought event. The entire eastern part of Australia shows its largest breakpoint towards the end of the time series because of a La Niña event, which caused lower temperatures and higher rainfall than usual during the years 2010 and 2011.

Appendix B: Reconstruction Error
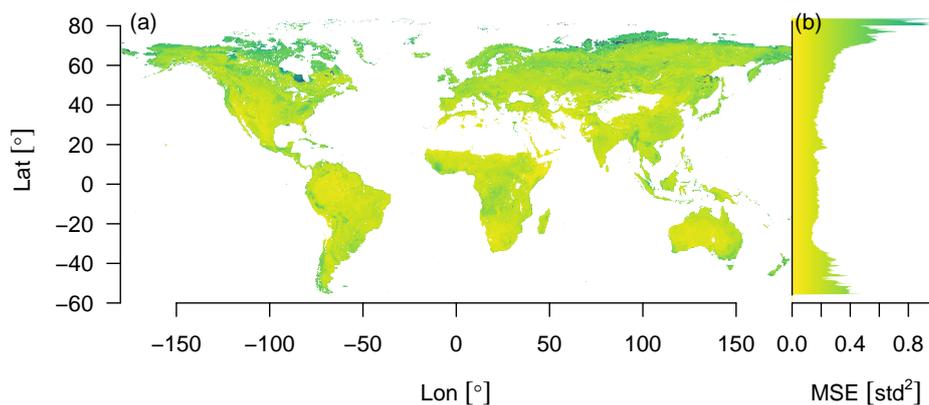
Figure B1. The reconstruction error of the first two pca dimensions aggregated over variables an time by the mean of the square error. The right plot shows the mean reconstruction error aggregated over latitudes.

370   In order to find ecosystems that do no fit well your model of two indicators, we calculated the reconstruction error of the first two PCA axes. Ecosystems that do not fit our model well show a higher reconstruction error, see fig. B1. Higher reconstruction errors appear in extreme latitudes, areas with especially high reconstruction error are at the southern part of the Hudson Bay area. Very limited regions in central and eastern Russia and northern Siberia.

Appendix C: Bowen Ration
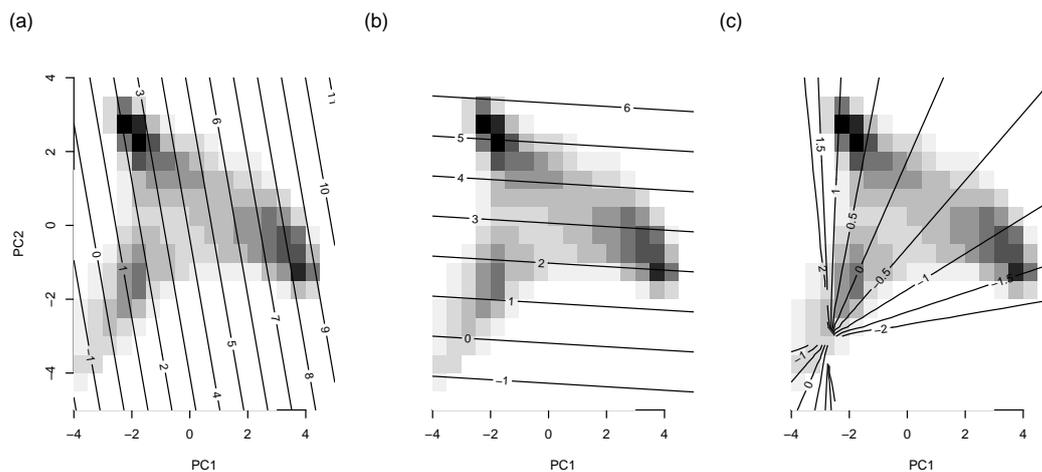
Figure C1. The background shading show the distribution of the mean seasonal cycle of the spatial points (see fig. 2). The contour lines represent the reconstruction of the variables from the first two principal components. The reconstructed variables are (a) Latent Heat, (b) Sensible heat, and (c) $\log_{10}\left(\frac{\text{Latent Heat}}{\text{Sensible Heat}}\right)$, the $\log_{10}$ of the Bowen Ratio.

Appendix D: Mean Seasonal Cycle Extrema
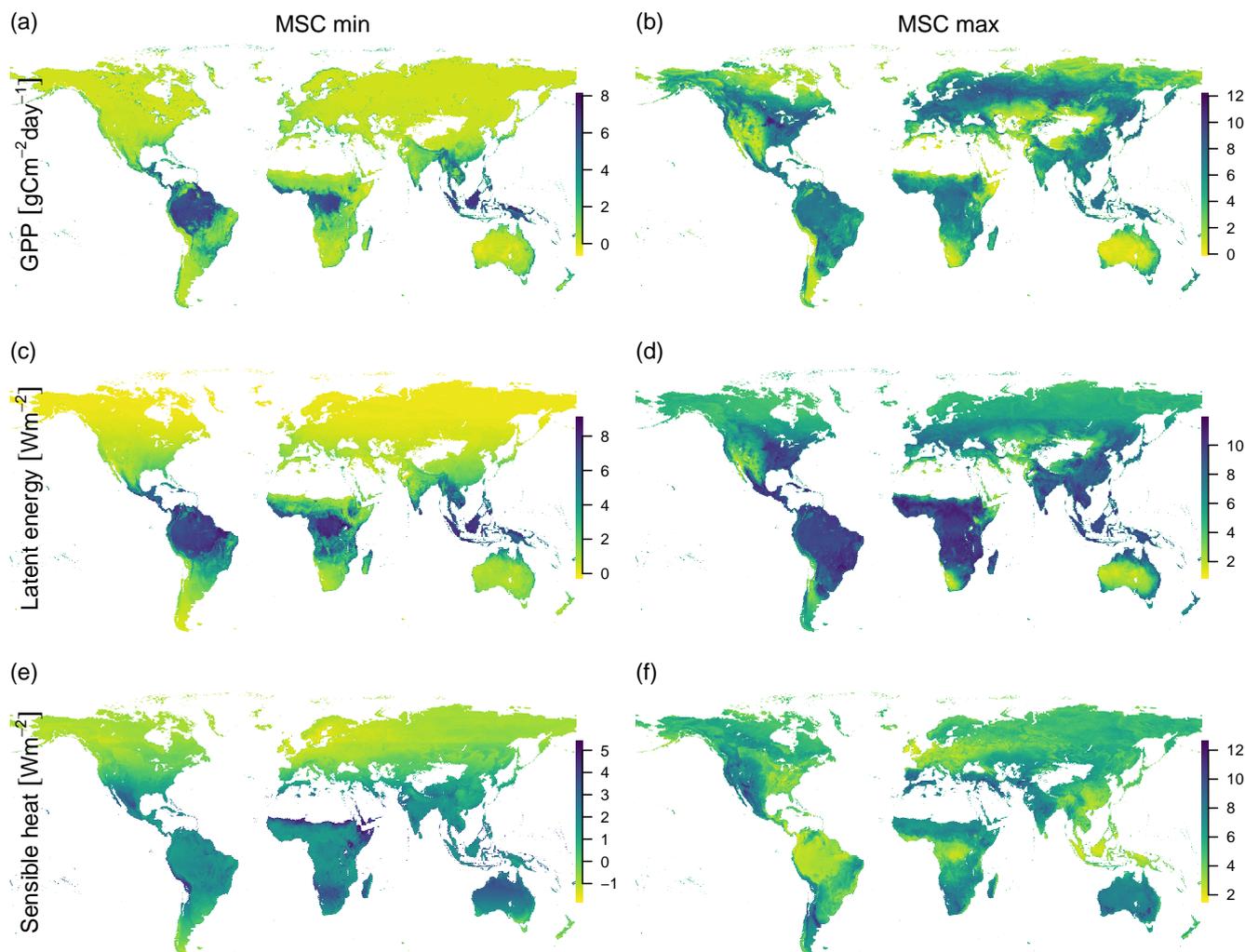
Figure D1. Shows the minimum (left column) and maximum (right column) mean seasonal cycles of GPP (upper row), Latent Heat (middle row), and Sensible heat (lower row). This illustrates the similarity of possibly very different ecosystems in terms of productivity and limitations. During peak growing season, many mid latitude areas have a similar productivity and latent energy release as tropical rainforests (subfigure b and d). The highest maximum seasonal sensible heat loss can be found in dry areas around the world and is lowest in areas with a wet climate such as tropical rainforests and maritime climates (subfigure f).

Appendix E: Changes in the Seasonal Amplitude
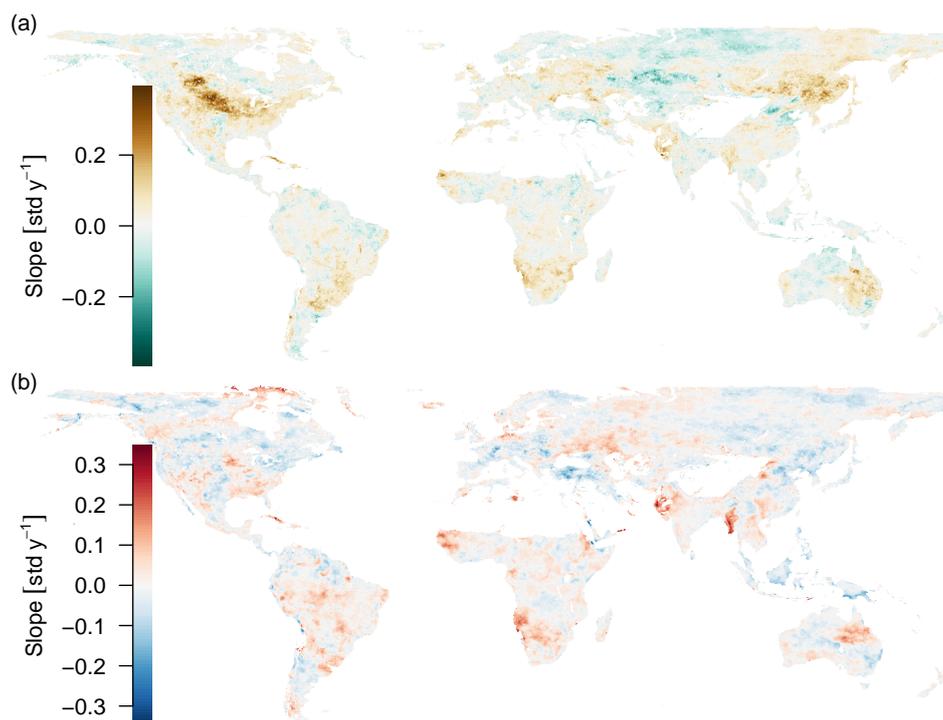
Figure E1. Trends in the amplitude of the yearly cycle, Theil-Sen estimators only significant slopes ($p < 0.05$), *unadjusted*, are shown. Because there is only a single amplitude per year and therefore only 11 data points per time series, the adjusted significances are not significant.

Appendix F: Description of variables

Variables used describing the biosphere can be found in tab. 1, here we provide a more complete description of all variables:

Black Sky Albedo is the reflected fraction of total incoming radiation under direct hermispherical reflectance, i.e. direct
380 illumination (Muller et al., 2011).

White Sky Albedo is the reflected fraction of total incoming radiation under bihemispherical reflectance, i.e. diffuse illumination (Muller et al., 2011). Together with black sky albedo it can be used to estimate the albedo under different illumination conditions.

Evaporation $[\mathrm{mm/day}]$ is the amount of water evaporated per day (Martens et al., 2017), depends on the amount of available
385 water and energy.

Evaporative Stress modeled water stress for plants, zero means that the vegetation has no water available for transpiration and one means that transpiration equals potential transpiration (Martens et al., 2017).

fAPAR the fraction of absorbed photosynthetically active radiation, a proxy for plant productivity (Disney et al., 2016).

Gross Primary Productivity (GPP) $[\mathrm{gCm^{-2}day^{-1}}]$ the total amount of carbon fixed by photosynthesis (Tramontana et al.,
390 2016).

Terrestrial Ecosystem Respiration (TER) $[\mathrm{gCm^{-2}day^{-1}}]$ the total amount of carbon respired by the ecosystem, includes autotrophic and heterotropic respiration (Tramontana et al., 2016).

Net Ecosystem Exchange (NEE) $[\mathrm{gCm^{-2}day^{-1}}]$ The total exchange of carbon of the ecosystem with the atmosphere
$\mathrm{NEE} = \mathrm{GPP} - \mathrm{TER}$ (Tramontana et al., 2016).

395 **Latent energy (LE)** $[\mathrm{Wm^{-2}}]$ the amount of energy lost by the surface due to evaporation (Tramontana et al., 2016).

Sensible Heat (H) $[\mathrm{Wm^{-2}}]$ the amount of energy lost by the surface due to radiation (Tramontana et al., 2016).

Root-Zone Soil Moisture $[\mathrm{m^3m^{-3}}]$ the moisture content of the root zone, estimated by the GLEAM model (Martens et al., 2017).

Surface Soil Moisture $[\mathrm{mm^3mm^{-3}}]$ the soil moisture content at the soil surface (Martens et al., 2017).

400 *Author contributions.* GK and MDM designed the study in collaboration with MR and GCV. GK conducted the analysis and wrote the manuscript with contributions from all co-authors

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study is funded by the Earth system data lab—a project by the European Space Agency. MDM and MR have been supported by the H2020 EU project BACI under grant agreement No 640176. GCV work has been supported by the EU under the
405 ERC consolidator grant SEDAL-647423. We thank Fabian Gans and German Poveda for useful discussions. We thank Jake Nelson for proofreading the manuscript.

References

Abatzoglou, J. T., Rupp, D. E., and Mote, P. W.: Seasonal Climate Variability and Change in the Pacific Northwest of the United States, Journal of Climate, 27, 2125–2142, https://doi.org/10.1175/JCLI-D-13-00218.1, 2014.

410 Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review: GPP Spatiotemporal Patterns, Reviews of Geophysics, 53, 785–818, https://doi.org/10.1002/2015RG000483, http://doi.wiley.com/10.1002/2015RG000483, 2015.

Aragão, L. E. O. C., Anderson, L. O., Fonseca, M. G., Rosan, T. M., Vedovato, L. B., Wagner, F. H., Silva, C. V. J., Silva Junior, C. H. L., Arai, E., Aguiar, A. P., Barlow, J., Berenguer, E., Deeter, M. N., Domingues, L. G., Gatti, L., Gloor, M., Malhi, Y., Marengo, J. A., Miller, J. B., Phillips, O. L., and Saatchi, S.: 21st Century Drought-Related Fires Counteract the Decline of Amazon Deforestation Carbon Emissions, Nature Communications, 9, 536, https://doi.org/10.1038/s41467-017-02771-y, 2018.

Ardisson, P.-L., Bourget, E., and Legendre, P.: Multivariate Approach to Study Species Assemblages at Large Spatiotemporal Scales: The Community Structure of the Epibenthic Fauna of the Estuary and Gulf of St. Lawrence, Canadian Journal of Fisheries and Aquatic 420 Sciences, 47, 1364–1377, https://doi.org/10.1139/f90-156, 1990.

Arenas-Garcia, J., Petersen, K. B., Camps-Valls, G., and Hansen, L. K.: Kernel Multivariate Analysis Framework for Supervised Subspace Learning: A Tutorial on Linear and Kernel Multivariate Methods, IEEE Signal Processing Magazine, 30, 16–29, https://doi.org/10.1109/MSP.2013.2250591, 2013.

Baldocchi, D.: 'Breathing' of the Terrestrial Biosphere: Lessons Learned from a Global Network of Carbon Dioxide Flux Measurement 425 Systems, Australian Journal of Botany, 56, 1, https://doi.org/10.1071/BT07151, 2008.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U, K. T. P., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U, K. 430 T. P., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, Bulletin of the American Meteorological Society, 2001.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and García-Herrera, R.: The Hot Summer of 2010: Redrawing the Temperature Record Map of Europe, Science, 332, 220–224, https://doi.org/10.1126/science.1201224, 2011.

435 Beisner, B., Haydon, D., and Cuddington, K.: Alternative Stable States in Ecology, Frontiers in Ecology and the Environment, 1, 376–382, https://doi.org/10.1890/1540-9295(2003)001[0376:ASSIE]2.0.CO;2, 2003.

Benjamini, Y. and Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society. Series B (Methodological), 57, 289–300, 1995.

Berger, M., Moreno, J., Johannessen, J. A., Levelt, P. F., and Hanssen, R. F.: ESA's Sentinel Missions in Support of Earth System Science, 440 Remote Sensing of Environment, 120, 84–90, https://doi.org/10.1016/j.rse.2011.07.023, 2012.

Blonder, B., Moulton, D. E., Blois, J., Enquist, B. J., Graae, B. J., Macias-Fauria, M., McGill, B., Nogué, S., Ordonez, A., Sandel, B., and Svenning, J.-C.: Predictability in Community Dynamics, Ecology Letters, 20, 293–306, https://doi.org/10.1111/ele.12736, 2017.

Bowen, I. S.: The Ratio of Heat Losses by Conduction and by Evaporation from Any Water Surface, Physical Review, 27, 779–787, https://doi.org/10.1103/PhysRev.27.779, 1926.

445 Brown, R. L., Durbin, J., and Evans, J. M.: Techniques for Testing the Constancy of Regression Relationships over Time, Journal of the Royal Statistical Society. Series B (Methodological), 37, 149–192, 1975.

Chapin, F. S., Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D. D., Clark, D. A., Harmon, M. E., Schimel, D. S., Valentini, R., Wirth, C., Aber, J. D., Cole, J. J., Goulden, M. L., Harden, J. W., Heimann, M., Howarth, R. W., Matson, P. A., McGuire, A. D., Melillo, J. M., Mooney, H. A., Neff, J. C., Houghton, R. A., Pace, M. L., Ryan, M. G., Running, S. W., Sala, O. E.,
450 Schlesinger, W. H., and Schulze, E.-D.: Reconciling Carbon-Cycle Concepts, Terminology, and Methods, Ecosystems, 9, 1041–1050, https://doi.org/10.1007/s10021-005-0105-7, 2006.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., Noblet, N. D., Friend, A. D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D.,
455 Vesala, T., and Valentini, R.: Europe-Wide Reduction in Primary Productivity Caused by the Heat and Drought in 2003, Nature, 437, 529, https://doi.org/10.1038/nature03972, 2005.

de Jong, R., de Bruin, S., de Wit, A., Schaepman, M. E., and Dent, D. L.: Analysis of Monotonic Greening and Browning Trends from Global NDVI Time-Series, Remote Sensing of Environment, 115, 692–702, https://doi.org/10.1016/j.rse.2010.10.011, 2011.

Disney, M., Muller, J.-P., Kharbouche, S., Kaminski, T., Voßbeck, M., Lewis, P., and Pinty, B.: A New Global fAPAR and LAI Dataset
460 Derived from Optimal Albedo Estimates: Comparison with MODIS Products, Remote Sensing, 8, 275, https://doi.org/10.3390/rs8040275, 2016.

Doughty, C. E., Metcalfe, D. B., Girardin, C. a. J., Amézquita, F. F., Cabrera, D. G., Huasco, W. H., Silva-Espejo, J. E., Araujo-Murakami, A., da Costa, M. C., Rocha, W., Feldpausch, T. R., Mendoza, A. L. M., da Costa, A. C. L., Meir, P., Phillips, O. L., and Malhi, Y.: Drought impact on forest carbon dynamics and fluxes in Amazonia, Nature, 519, 78–82, https://doi.org/10.1038/nature14213, https://www.nature.
465 com/articles/nature14213, 2015.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Feldpausch, T. R., Phillips, O. L., Brienen, R. J. W., Gloor, E., Lloyd, J., Lopez-Gonzalez, G., Monteagudo-Mendoza, A., Malhi, Y., Alarcón,
470 A., Dávila, E. A., Alvarez-Loayza, P., Andrade, A., Aragao, L. E. O. C., Arroyo, L., C, G. A. A., Baker, T. R., Baraloto, C., Barroso, J., Bonal, D., Castro, W., Chama, V., Chave, J., Domingues, T. F., Fauset, S., Groot, N., Coronado, E. H., Laurance, S., Laurance, W. F., Lewis, S. L., Licona, J. C., Marimon, B. S., Marimon-Junior, B. H., Bautista, C. M., Neill, D. A., Oliveira, E. A., dos Santos, C. O., Camacho, N. C. P., Pardo-Molina, G., Prieto, A., Quesada, C. A., Ramírez, F., Ramírez-Angulo, H., Réjou-Méchain, M., Rudas, A., Saiz, G., Salomão, R. P., Silva-Espejo, J. E., Silveira, M., ter Steege, H., Stropp, J., Terborgh, J., Thomas-Caesar, R., van der Heijden, G.
475 M. F., Martinez, R. V., Vilanova, E., and Vos, V. A.: Amazon Forest Response to Repeated Droughts, Global Biogeochemical Cycles, 30, 964–982, https://doi.org/10.1002/2015GB005133, 2016.

Flach, M., Sippel, S., Gans, F., Bastos, A., Brenning, A., Reichstein, M., and Mahecha, M. D.: Contrasting biosphere responses to hydrometeorological extremes: revisiting the 2010 western Russian heatwave, Biogeosciences, 15, 6067–6085, https://doi.org/https://doi.org/10.5194/bg-15-6067-2018, https://www.biogeosciences.net/15/6067/2018/, 2018.

480   Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., and Holling, C.: Regime Shifts, Resilience, and Biodiversity in Ecosystem Management, Annual Review of Ecology, Evolution, and Systematics, 35, 557–581, https://doi.org/10.1146/annurev.ecolsys.35.021103.105711, 2004.

Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U., and Carvalhais, N.: Codominant Water Control on Global Interannual Variability and Trends in Land Surface Phenology and Greenness, Global Change Biology, 21, 3414–3435,
485   https://doi.org/10.1111/gcb.12950, 2015.

Forkel, M., Carvalhais, N., Rodenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Zaehle, S., and Reichstein, M.: Enhanced Seasonal CO2 Exchange Caused by Amplified Plant Productivity in Northern Ecosystems, Science, 351, 696–699, https://doi.org/10.1126/science.aac4971, 2016.

García-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J., and Fischer, E. M.: A Review of the European Summer Heat Wave of 2003, Critical
490   Reviews in Environmental Science and Technology, 40, 267–306, https://doi.org/10.1080/10643380802238137, http://www.tandfonline.com/doi/abs/10.1080/10643380802238137, 2010.

Gaumont-Guay, D., Black, T. A., Griffis, T. J., Barr, A. G., Jassal, R. S., and Nesic, Z.: Interpreting the Dependence of Soil Respiration on Soil Temperature and Water Content in a Boreal Aspen Stand, Agricultural and Forest Meteorology, 140, 220–235, https://doi.org/10.1016/j.agrformet.2006.08.003, 2006.

495   Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., Welp, L. R., Sweeney, C., Tans, P. P., Kelley, J. J., Daube, B. C., Kort, E. A., Santoni, G. W., and Bent, J. D.: Enhanced Seasonal Exchange of CO2 by Northern Ecosystems Since 1960, Science, 341, 1085–1089, https://doi.org/10.1126/science.1239207, 2013.

Higham, N. J.: The Accuracy of Floating Point Summation, SIAM J. Scientific Computing, 14, 783–799, https://doi.org/10.1137/0914050, 1993.

500   Huang, K., Xia, J., Wang, Y., Ahlström, A., Chen, J., Cook, R. B., Cui, E., Fang, Y., Fisher, J. B., Huntzinger, D. N., Li, Z., Michalak, A. M., Qiao, Y., Schaefer, K., Schwalm, C., Wang, J., Wei, Y., Xu, X., Yan, L., Bian, C., and Luo, Y.: Enhanced Peak Growth of Global Vegetation and Its Key Mechanisms, Nature Ecology & Evolution, 2, 1897–1905, https://doi.org/10.1038/s41559-018-0714-0, 2018.

IPBES: Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Summary for policymakers, IPBES, 2019.

505   IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Tech. rep., IPCC, Geneva, Swizerland, 2014.

Ivits, E., Horion, S., Fensholt, R., and Cherlet, M.: Drought Footprint on European Ecosystems between 1999 and 2010 Assessed by Remotely Sensed Vegetation Phenology and Productivity, Global Change Biology, 20, 581–593, https://doi.org/10.1111/gcb.12393, 2014.

Keeling, C. D., Chin, J. F. S., and Whorf, T. P.: Increased Activity of Northern Vegetation Inferred from Atmospheric CO 2 Measurements,
510   Nature, 382, 146, https://doi.org/10.1038/382146a0, 1996.

Kendall, M. G.: Rank Correlation Methods., Griffin, London, 1970.

Khanna, J., Medvigy, D., Fueglistaler, S., and Walko, R.: Regional dry-season climate changes due to three decades of Amazonian deforestation, Nature Climate Change, 7, 200–204, https://doi.org/10.1038/nclimate3226, https://www.nature.com/articles/nclimate3226, 2017.

Kraemer, G., Reichstein, M., and Mahecha, M. D.: dimRed and coRanking - Unifying Dimensionality Reduction in R, The R Journal, 10,
515   342–358, https://doi.org/10.32614/RJ-2018-039, 2018.

Kuan, C.-M. and Hornik, K.: The Generalized Fluctuation Test: A Unifying View, Econometric Reviews, 14, 135–161, https://doi.org/10.1080/07474939508800311, 1995.

Legendre, P. and Legendre, L.: Numerical Ecology: Second English Edition, Developments in environmental modelling, 20, 1998.

Legendre, P., Planas, D., and Auclair, M.-J.: Succession des communautés de gastéropodes dans deux milieux différant par leur degré
520   d'eutrophisation, Canadian Journal of Zoology, 62, 2317–2327, https://doi.org/10.1139/z84-339, http://www.nrcresearchpress.com/doi/
       10.1139/z84-339, 1984.

Mahecha, M. D., Martínez, A., Lischeid, G., and Beck, E.: Nonlinear Dimensionality Reduction: Alternative Ordination Approaches
       for Extracting and Visualizing Biodiversity Patterns in Tropical Montane Forest Vegetation Data, Ecological Informatics, 2, 138–149,
       https://doi.org/10.1016/j.ecoinf.2007.05.002, 2007a.

525   Mahecha, M. D., Reichstein, M., Lange, H., Carvalhais, N., Bernhofer, C., Grünwald, T., Papale, D., and Seufert, G.:
       Characterizing Ecosystem-Atmosphere Interactions from Short to Interannual Time Scales, Biogeosciences, 4, 743–758,
       https://doi.org/https://doi.org/10.5194/bg-4-743-2007, 2007b.

Mahecha, M. D., Gans, F., Sippel, S., Donges, J. F., Kaminski, T., Metzger, S., Migliavacca, M., Papale, D., Rammig, A., and
       Zscheischler, J.: Detecting Impacts of Extreme Events with Ecological in Situ Monitoring Networks, Biogeosciences, 14, 4255–4277,
530   https://doi.org/https://doi.org/10.5194/bg-14-4255-2017, 2017.

Mann, H. B.: Nonparametric Tests Against Trend, Econometrica, 13, 245–259, https://doi.org/10.2307/1907187, 1945.

Martens, B., Miralles, D. G., Lievens, H., Schalie, R. v. d., Jeu, R. A. M. d., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest,
       N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903–1925,
       https://doi.org/https://doi.org/10.5194/gmd-10-1903-2017, https://www.geosci-model-dev.net/10/1903/2017/, 2017.

535   Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Sayre, R., Trabucco, A., and Zomer, R.: A High-resolution Bioclimate Map of the
       World: A Unifying Framework for Global Biodiversity Research and Monitoring, Global Ecology and Biogeography, 22, 630–638,
       https://doi.org/10.1111/geb.12022, 2013.

Mika, S., Scholkopf, B., Smola, A., Muller, K., Scholz, M., and Ratsch, G.: Kernel PCA and De-Noising in Feature Spaces, in: ADVANCES
       IN NEURAL INFORMATION PROCESSING SYSTEMS 11, edited by Kearns, MS and Solla, SA and Cohn, DA, vol. 11 of *Advances*
540   *in Neural Information Processing Systems*, pp. {536–542}, 12th Annual Conference on Neural Information Processing Systems (NIPS),
       DENVER, CO, NOV 30-DEC 05, 1998, 1999.

Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C., and Vilà-Guerau de Arellano, J.: Mega-heatwave temperatures due to combined soil
       desiccation and atmospheric heat accumulation, Nature Geoscience, 7, 345–349, https://doi.org/10.1038/ngeo2141, https://www.nature.
       com/articles/ngeo2141, 2014.

545   Muller, J.-P., Lewis, P., Fischer, J., North, P., and Framer, U.: The ESA GlobAlbedo Project for mapping the Earth's land surface albedo for
       15 years from European sensors, Geophysical Research Abstracts, 13, 2, 2011.

Najafi, E., Pal, I., and Khanbilvardi, R.: Climate Drives Variability and Joint Variability of Global Crop Yields, Science of The Total Envi-
       ronment, 662, 361–372, https://doi.org/10.1016/j.scitotenv.2019.01.172, 2019.

Nicholson, S. E.: A detailed look at the recent drought situation in the Greater Horn of Africa, Journal of Arid Environments, 103, 71–79,
550   https://doi.org/10.1016/j.jaridenv.2013.12.003, http://www.sciencedirect.com/science/article/pii/S0140196313002322, 2014.

Parmesan, C.: Ecological and Evolutionary Responses to Recent Climate Change, Annual Review of Ecology, Evolution, and Systematics,
       37, 637–669, https://doi.org/10.1146/annurev.ecolsys.37.091305.110100, 2006.

Pearson, K.: On Lines and Planes of Closest Fit to Systems of Points in Space, Philosophical Magazine, 2, 559–572, 1901.

Reichstein, M., Bahn, M., Ciais, P., Frank, D., Mahecha, M. D., Seneviratne, S. I., Zscheischler, J., Beer, C., Buchmann, N., Frank, D. C.,
555 Papale, D., Rammig, A., Smith, P., Thonicke, K., van der Velde, M., Vicca, S., Walz, A., and Wattenbach, M.: Climate extremes and the
carbon cycle, Nature, 500, 287–295, https://doi.org/10.1038/nature12350, http://www.nature.com/articles/nature12350, 2013.

Renner, M., Brenner, C., Mallick, K., Wizemann, H.-D., Conte, L., Trebs, I., Wei, J., Wulfmeyer, V., Schulz, K., and Kleidon, A.: Using
Phase Lags to Evaluate Model Biases in Simulating the Diurnal Cycle of Evapotranspiration: A Case Study in Luxembourg, Hydrology
and Earth System Sciences, 23, 515–535, https://doi.org/https://doi.org/10.5194/hess-23-515-2019, 2019.

560 Richardson, A. D., Braswell, B. H., Hollinger, D. Y., Burman, P., Davidson, E. A., Evans, R. S., Flanagan, L. B., Munger, J. W., Savage, K.,
Urbanski, S. P., and Wofsy, S. C.: Comparing Simple Respiration Models for Eddy Flux and Dynamic Chamber Data, Agricultural and
Forest Meteorology, 141, 219–234, https://doi.org/10.1016/j.agrformet.2006.10.010, 2006.

Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin, III, F. S., Lambin, E., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J.,
Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sorlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark,
565 M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., and Foley, J.: Planetary
Boundaries: Exploring the Safe Operating Space for Humanity, ECOLOGY AND SOCIETY, 14, 2009.

Sarmah, S., Jia, G., and Zhang, A.: Satellite View of Seasonal Greenness Trends and Controls in South Asia, Environmental Research Letters,
13, 034 026, https://doi.org/10.1088/1748-9326/aaa866, 2018.

Schwartz, M. D.: Monitoring Global Change with Phenology: The Case of the Spring Green Wave, International Journal of Biometeorology,
570 38, 18–22, https://doi.org/10.1007/BF01241799, 1994.

Schwartz, M. D.: Green-Wave Phenology, Nature, 394, 839, https://doi.org/10.1038/29670, 1998.

Sen, P. K.: Estimates of the Regression Coefficient Based on Kendall's Tau, Journal of the American Statistical Association, 63, 1379–1389,
https://doi.org/10.2307/2285891, 1968.

Sippel, S., Reichstein, M., Ma, X., Mahecha, M. D., Lange, H., Flach, M., and Frank, D.: Drought, Heat, and the Carbon Cycle: A Review,
575 Current Climate Change Reports, 4, 266–286, https://doi.org/10.1007/s40641-018-0103-4, 2018.

Song, X.-P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., and Townshend, J. R.: Global Land Change from
1982 to 2016, Nature, 560, 639, https://doi.org/10.1038/s41586-018-0411-9, 2018.

Stine, A. R., Huybers, P., and Fung, I. Y.: Changes in the Phase of the Annual Cycle of Surface Temperature, Nature, 457, 435–440,
https://doi.org/10.1038/nature07675, 2009.

580 Tang, J., Baldocchi, D. D., and Xu, L.: Tree Photosynthesis Modulates Soil Respiration on a Diurnal Time Scale, Global Change Biology,
11, 1298–1304, https://doi.org/10.1111/j.1365-2486.2005.00978.x, 2005.

Theil, H.: A Rank-Invariant Method of Linear and Polynomial Regression Analysis, I, II, III, Proceedings of the Koninklijke Nederlandse
Akademie van Wetenschappen, 53, 386–392, 521–525, 1397–1412, 1950.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G.,
585 Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET
sites with regression algorithms, Biogeosciences, 13, 4291–4313, https://doi.org/https://doi.org/10.5194/bg-13-4291-2016, https://www.
biogeosciences.net/13/4291/2016/, 2016.

van der Maaten, L., Schmidtlein, S., and Mahecha, M. D.: Analyzing Floristic Inventories with Multiple Maps, Ecological Informatics, 9,
1–10, https://doi.org/10.1016/j.ecoinf.2012.01.005, 2012.

590 Wolter, K. and Timlin, M. S.: El Niño/Southern Oscillation Behaviour since 1871 as Diagnosed in an Extended Multivariate ENSO Index
(MEI.Ext), International Journal of Climatology, 31, 1074–1087, https://doi.org/10.1002/joc.2336, 2011.

Wright, J. S., Fu, R., Worden, J. R., Chakraborty, S., Clinton, N. E., Risi, C., Sun, Y., and Yin, L.: Rainforest-Initiated Wet Season Onset over the Southern Amazon, Proceedings of the National Academy of Sciences, 114, 8481–8486, https://doi.org/10.1073/pnas.1621516114, 2017.

595 Zeileis, A., Leisch, F., Hornik, K., and Kleiber, C.: Strucchange: An R Package for Testing for Structural Change in Linear Regression Models, Journal of Statistical Software, 7, 1–38, https://doi.org/10.18637/jss.v007.i02, 2002.

Zeng, N., Zhao, F., Collatz, G. J., Kalnay, E., Salawitch, R. J., West, T. O., and Guanter, L.: Agricultural Green Revolution as a Driver of Increasing Atmospheric CO_2 Seasonal Amplitude, Nature, 515, 394–397, https://doi.org/10.1038/nature13893, 2014.

Zhang, Q., Phillips, R. P., Manzoni, S., Scott, R. L., Oishi, A. C., Finzi, A., Daly, E., Vargas, R., and Novick, K. A.: Changes in Photosynthesis
600 and Soil Moisture Drive the Seasonal Soil Respiration-Temperature Hysteresis Relationship, Agricultural and Forest Meteorology, 259, 184–195, https://doi.org/10.1016/j.agrformet.2018.05.005, 2018.

Zhou, L., Tian, Y., Myneni, R. B., Ciais, P., Saatchi, S., Liu, Y. Y., Piao, S., Chen, H., Vermote, E. F., Song, C., and Hwang, T.: Widespread Decline of Congo Rainforest Greenness in the Past Decade, Nature, 509, 86–90, https://doi.org/10.1038/nature13265, 2014.

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng,
605 L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., Pan, Y., Peng, S., Peñuelas, J., Poulter, B., Pugh, T. A. M., Stocker, B. D., Viovy, N., Wang, X., Wang, Y., Xiao, Z., Yang, H., Zaehle, S., and Zeng, N.: Greening of the Earth and Its Drivers, Nature Climate Change, 6, 791–795, https://doi.org/10.1038/nclimate3004, 2016.