PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Research

Cite this article: Burton C, Rifai S, Malhi Y. 2018 Inter-comparison and assessment of gridded climate products over tropical forests during the 2015/2016 El Niño. *Phil. Trans. R. Soc. B* **373**: 20170406.
http://dx.doi.org/10.1098/rstb.2017.0406

Accepted: 25 July 2018

One contribution of 22 to a discussion meeting issue 'The impact of the 2015/2016 El Niño on the terrestrial tropical carbon cycle: patterns, mechanisms and implications'.

Subject Areas:
environmental science

Keywords:
reanalysis, validation, El Niño southern oscillation, tropical forests

Author for correspondence:
C. Burton
e-mail: c.a.burton123@gmail.com

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4211753.

THE ROYAL SOCIETY PUBLISHING

Inter-comparison and assessment of gridded climate products over tropical forests during the 2015/2016 El Niño

C. Burton, S. Rifai and Y. Malhi

Environmental Change Institute, School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK

CB, 0000-0003-3048-8484; SR, 0000-0003-3400-8601; YM, 0000-0002-3503-4783

To understand the impacts of extreme climate events, it is first necessary to understand the spatio-temporal characteristics of the event. Gridded climate products are frequently used to describe climate patterns but have been shown to perform poorly over data-sparse regions such as tropical forests. Often, they are uncritically employed in a wide range of studies linking tropical forest processes to large-scale climate variability. Here, we conduct an inter-comparison and assessment of near-surface air temperature fields supplied by four state-of-the-art reanalysis products, along with precipitation estimates supplied by four merged satellite-gauge rainfall products. Firstly, spatio-temporal patterns of temperature and precipitation anomalies during the 2015–2016 El Niño are shown for each product to characterize the impact of the El Niño on the tropical forest biomes of Equatorial Africa, Southeast Asia and South America. Using meteorological station data, a two-stage assessment is then conducted to determine which products most reliably model tropical climates during the 2015–2016 El Niño, and which perform best over the longer-term satellite observation period (1980–2016). Results suggest that eastern Amazonia, parts of the Congo Basin and mainland Southeast Asia all experienced significant monthly mean temperature anomalies during the El Niño, while northeastern Amazonia, eastern Borneo and southern New Guinea experienced significant precipitation deficits. Our results suggest ERA-Interim and MERRA2 are the most reliable air temperature datasets, while TRMM 3B42 V7 and CHIRPS v2.0 are the best-performing rainfall datasets.

This article is part of a discussion meeting issue 'The impact of the 2015/2016 El Niño on the terrestrial tropical carbon cycle: patterns, mechanisms and implications'.

1. Introduction

In recent decades, a multitude of global and quasi-global gridded climate products has been produced by the world's major meteorological organizations (e.g. European Centre for Medium-range Weather Forecasts, ECMWF; National Oceanic and Atmospheric Administration, NOAA; National Aeronautics and Space Administration, NASA) and has found wide use in the scientific research community. These climate products, which include atmospheric retrospective analysis (reanalysis) models and satellite-based remote sensing products, combine millions of irregularly distributed satellite and on-the-ground observations from multiple observing networks, and, through data assimilation algorithms (and climate models in the case of reanalysis) produce physically coherent, spatio-temporally complete gridded climate datasets [1,2]. For many gridded climate products, tropical forest areas present the greatest challenge as they are regions where on-the-ground observations are sparse and land-atmosphere feedbacks are complex and fine-scale. Despite ingesting similar observations into their assimilation algorithms, significant differences in the spatio-temporal patterns of climate trends have been identified between different climate

datasets, and where assessment has been undertaken, many of the products show poor correlations with on-the-ground observations [3–5]. This is particularly the case in data sparse but functionally important regions such as the tropical forest biome. Regardless of these known concerns, these climate products are often uncritically employed in investigations linking tropical forest processes with major climate modes of variability such as the El Niño Southern Oscillation (ENSO) [6–9]. Thus, there is an imperative to discern which climate products are most reliable over tropical forests to ensure findings are robust against the uncertainties introduced by the limitations of these products.

Here, we present an inter-comparison and assessment of the near-surface air temperature fields supplied by four state-of-the-art reanalysis products, and the precipitation estimates supplied by four merged satellite-gauge rainfall products. Our aim is to determine which of the most widely used climate products reliably model air temperature and rainfall over tropical forests at the continental scale, both during the most recent major El Niño event, and over the entire satellite observation period from 1980 to 2016. First, the spatio-temporal patterns of temperature and precipitation anomalies during the 2015–2016 El Niño are shown for each product, both to characterize where and when the most recent El Niño impacted upon the tropical biome, and to highlight the discrepancies and similarities between products. We then conduct a two-stage assessment using station data compiled from across the three tropical continents. The first stage of assessment uses station data from the 2015–2016 El Niño period to determine which products most accurately modelled the latest El Niño. In the second assessment stage, station data collected over the period 1980–2016 are used to determine which products performed most reliably over this extended period.

2. Data and methods

(a) Temperature and precipitation datasets

For the near-surface (2-m) temperature inter-comparison and assessment, four state-of the-art reanalysis datasets were used: ERA-Interim, produced by ECMWF [10]; Climate Forecast System Reanalysis (CFSR), produced by National Centers for Environmental Prediction (NCEP) [11,12]; Modern-Era Retrospective Analysis 2 (MERRA2), produced by NASA [13]; and the Japanese 55-year Reanalysis (JRA55) produced by the Japanese Meteorological Agency (JMA) [14,15]. Specifications of the four models are included in the electronic supplementary material, table S1. Reanalysis models use a forecast model with data assimilation schemes for both conventional and satellite weather observations to produce an estimate of the past state of the atmosphere [16]. Forecast models are used to propagate information on the atmosphere forward in time from previous analyses, and an assimilation scheme uses the vast array of observations to blend the input observations with the results of the forecast model. The output of the assimilation scheme is then used as the initialization data for the next forecast [1]. Reanalysis products are subject to a combination of errors and uncertainties that arise due to gaps and errors in observations, shortcomings in the model physics, shortcomings in the assimilation algorithms, computational limitations resulting in coarse spatial resolutions relative to meteorological phenomena and the non-stationarity of observing systems that introduce time-dependent biases [2,5].

Precipitation estimates from reanalysis products are known to possess significant biases over the tropics [4,5,17] and are normally not used operationally. Satellite-based rainfall estimates are generally considered to produce better estimates of historical rainfall. We examined them here in place of reanalysis precipitation fields. Satellite-based rainfall estimates are based on observations from either thermal infrared (TIR) bands or passive microwave (PMW) sensors. Although satellite-based rainfall estimates provide high-resolution spatio-temporal coverage of regions that otherwise lack reliable rain gauges, they suffer from several limitations that stem from their reliance on remotely sensed spectral data rather than observations of rainfall itself [18–20]. To overcome these limitations, several new rainfall datasets merge TIR and PMW observations with *in situ* gauge measurements to produce a best-estimate rainfall product, and these products have been shown to produce more accurate results than the satellite-only versions [3,20–22]. Four merged satellite-gauge products are evaluated here: the Tropical Rainfall Measuring Mission (TRMM) 3B42 V7 Daily accumulated rainfall product [23]; the Climate Hazards Infrared Precipitation with Stations (CHIRPS) V2.0 dataset produced by the U.S. Geological Survey and the Climate Hazards Group (CHG) at the University of California, Irvine [24]; the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Climate Data Record (PERSIANN-CDR) produced by the Centre for Hydrometeorology and Remote Sensing at the University of California [25]; and the Climate Prediction Centres (CPC) Merged Analysis of Precipitation (CMAP) dataset produced by NOAA/NCEP [26]. Details of these products are also included in the electronic supplementary material, table S1.

(b) Pantropical temperature and precipitation anomalies

The 2015–2016 El Niño was probably the most significant climate event in the tropics since the 1997–1998 El Niño, and possibly since much earlier. Mean and maximum temperature and precipitation anomalies during the 2015–2016 peak El Niño period are quantified and mapped here to provide an account of the spatio-temporal impact of the latest El Niño on the tropical biosphere. The peak El Niño period is defined throughout this paper as occurring from January 2015 to May 2016, which corresponds to the peak sea surface temperatures (SST) anomalies in the Niño 3.4 region [27].

Monthly 2 m air temperature anomalies were calculated against a 1980–2016 baseline on the native resolution of each reanalysis product. The mean and maximum temperature anomalies during the peak 2015–2016 El Niño period were then calculated on a per-pixel basis. Each product was then bilinearly interpolated to a common 0.5° grid to facilitate comparison between products. The commonly gridded temperature anomaly maps were masked to include only lowland evergreen tropical forest using a modified version of the Hansen Global Forest Change (GFC) v1.4 product [28], remapped via a nearest neighbour method to match the resolution of the climate datasets. Regions of tropical forest in the GFC occurring 1000 m.a.s.l. were removed, and the mask was clipped to 21° north and south of the equator. Limiting the extent of tropical forest to include only the lowlands regions partly eliminates the complication of assessing

relatively coarse spatial resolution climate products in regions of complex topography where large biases can occur due to the steep temperature gradients of mountainous terrain. The extent of the tropical forest mask used in the analysis is presented in electronic supplementary material, figure S1.

Monthly total precipitation datasets underwent a similar set of pre-processing operations, except the datasets were remapped to a common 0.25° grid (excluding CMAP whose 2.5° resolution was deemed too coarse to resample down to 0.25°) and the reference baseline for calculating anomalies was set at 1998–2016 to account for the shorter time-span of the TRMM dataset.

Time-series of monthly mean temperature anomalies and monthly total precipitation anomalies for each tropical continent were produced by calculating a simple average of all cells within each continent at each time-step. The continentally averaged time-series presented in the results extend from January 2014 to December 2016 to cover the full timeline of positive SST anomalies.

(c) Performance evaluation using station observations

To go beyond an inter-comparison between climate datasets and begin to evaluate which products reliably model both the most recent El Niño event and the longer climate record in the tropics, a point-to-pixel station assessment was conducted. Several sources of error are associated with this method of assessment. Station air temperature data are point measurements and therefore not necessarily representative of the grid boxes in a model output where values for a grid represent the grid average. Additionally, station temperature measurements are subject to measurement errors, microclimatology effects and inhomogeneities from changes to measuring procedures [16,29]. Rain gauge observations are subject to the same errors as temperature observations but contain additional sources of error from the systematic under-sampling of rainfall by gauges owing to wind-field deformation above the gauge opening, evaporation and wetting on the internal walls [30,31]. The use of point data to validate gridded climate products is therefore not straightforward, and some systematic variation between the station observations and pixel estimates should be expected [22]. CMAP is not included in the station assessment as it was impractical to compare point observations to such a large area (each cell in CMAP is approximately 62 500 km^2) and it was not possible to aggregate multiple stations within a pixel due to non-overlapping records from adjacent stations.

Rainfall gauge observations and near-surface air temperature observations were compiled from the Global Historical Climatology Network Monthly (GHCN-M) database [32] (https://gis.ncdc.noaa.gov/maps/ncei/cdo/monthly), the Climate Research Unit Temperature V. 4 (CRUTEM4) database [33] (https://crudata.uea.ac.uk/cru/data/temperature), the CRU TS v 4.01 database [34] (https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.01/), the UK Met Office Integrated Data Archive System (MIDAS) database (http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0) and the Global Ecosystem Monitoring (GEM) Network (http://gem.tropicalforests.ox.ac.uk/). Each of these datasets is subject to a range of quality control measures that are conducted by each of the institutions listed. The meteorological station records over tropical forests are notoriously patchy so all data available within the time period in question were retained regardless of the continuity of the record. Over Amazonia and Southeast Asia, the shortest station record contained 12 months of climate data over the 1980–2016 period, while over Africa, the shortest station record contained only three months of climate data over the same interval. Generally, the number of observations included in the statistical analysis is sufficient to overwhelm any seasonal bias introduced through the inclusion of partial length time-series. This is not the case, however, for the assessment of rainfall over Africa during the 2015–2016 El Niño, which only included 23 observations. Hence, the results for the station analysis over Africa during the 2015–2016 El Niño should be viewed with caution. Locations of the stations used in this analysis are shown in the electronic supplementary material, figure S1, and table 2 through 5 show the number of station observations used for each assessment.

The model assessment process used the following statistical metrics to evaluate the quality of each climate dataset: the Coefficient of Determination (R^2), Mean Absolute Error (MAE), the refined Index of Agreement (d_r) [35] and the relative proportion of systematic to unsystematic Mean-Squared Error (MSE_s/MSE and MSE_u/MSE). Some assert it is better to rely on absolute measures of error rather than squared differences when evaluating models; however, computing MSE allows for deducing the type of error (systematic or random) [36]. d_r indicates the error of the model predictions relative to the observed deviations about the observed mean. It varies between −1.0 and 1.0, where a d_r equal to 1 would indicate perfect agreement between observed and predicted values. A d_r of 0.5 indicates the sum of the errors is half the sum of the perfect-model-deviation and observed-deviation magnitudes [35]. A negative d_r indicates the sum of the errors is larger than those from using the observed mean. d_r is used here in place of Willmott, Ackleson [36]'s original index of agreement as the refined index is less sensitive to large outliers and is also more useful at differentiating models that both show close agreement [35]. Use of these statistics allows for assessing the closeness of agreement between on-the-ground observations and the estimates generated by the rainfall and temperature products, the magnitude and direction of deviations between observations and model estimates, and the type of error these estimates produce. The following equations define each of the statistical metrics employed:

$$R^2 = \left[\frac{\sum_{i=1}^{n} (O_i - \bar{O})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{n} (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^{n} (S_i - \bar{S})^2}} \right], \tag{2.1}$$

$$\mathrm{MAE} = n^{-1} \sum_{i=1}^{n} |O_i - S_i|, \tag{2.2}$$

$$\mathrm{MSE}_s = n^{-1} \sum_{i=1}^{n} (\hat{S}_i - O_i)^2, \tag{2.3}$$

$$\mathrm{MSE}_u = n^{-1} \sum_{i=1}^{n} (S_i - \hat{S}_i)^2 \tag{2.4}$$

and

$$d_r = \begin{cases} 1 - \dfrac{\sum_{i=1}^{n} |S_i - O_i|}{2 \sum_{i=1}^{n} |O_i - \bar{O}|}, & \text{if } \sum_{i=1}^{n} |S_i - O_i| \le 2 \sum_{i=1}^{n} |O_i - \bar{O}| \\ \dfrac{2 \sum_{i=1}^{n} |O_i - \bar{O}|}{\sum_{i=1}^{n} |S_i - O_i|} - 1, & \text{if } \sum_{i=1}^{n} |S_i - O_i| > 2 \sum_{i=1}^{n} |O_i - \bar{O}| \end{cases} \tag{2.5}$$

where O_i is the observed monthly mean temperature or monthly total rainfall, S_i is the simulated temperature or rainfall from the corresponding month, $\bar{O}$ is the observed climatological mean and $\hat{S}_i$ is the predicted temperature or rainfall as given by the least square regression model, $\hat{S}_i = a + bSi$ between S and O.

3. Results

(a) Climate anomalies during the 2015–2016 El Niño

(i) Temperature

Across all three tropical continents, the reanalysis temperature datasets show little agreement on the magnitude and spatial distribution of temperature anomalies during the 2015–2016 El Niño. Figure 1*a*,*b* shows the spatial distribution of the mean and maximum near-surface temperature anomalies for each of the four reanalysis products during the 2015–2016 El Niño, respectively. Figure 1*c* shows the continent-averaged time-series of these anomalies, and table 1 displays the zonal mean statistics for the plots in figure 1*a*,*b*.

Averaged across all of Amazonia, MERRA2 showed the largest positive temperature anomaly with mean monthly temperatures 1.13°C above average, with large regions of eastern Amazonia reaching mean temperature anomalies throughout the peak El Niño period of 1.5–2.5°C. Maximum temperatures anomalies in MERRA2 reach 3.5–5.5°C over wide regions of the basin (these are anomalies based on monthly *mean* temperatures). All the reanalysis products show a positive temperature anomaly over Amazonia, with MERRA2 and ERA-Interim (0.87°C) presenting significantly larger and more widespread anomalies than JRA55 (0.47°C) and CFSR (0.33°C). Peak temperature anomalies over Amazonia were a prolonged event that occurred from September 2015 until January 2016, followed by a secondary smaller peak in mid-2016 (figure 1*c*).

Over Africa, there is essentially no agreement between the reanalysis products on the spatial pattern of temperature anomalies. ERA-Interim shows a continent-wide mean temperature anomaly of 0.66°C, with some areas of the northeast Congo and coastal west Africa recording a mean anomaly higher than 1°C. Mean temperature anomalies over Africa in MERRA2 (0.39°C) and JRA55 (0.40°C) were less uniform and less severe than in ERA-Interim, and maximum temperature anomalies were located over the central Congo Basin rather than the northeast. CFSR records, on average, a temperature anomaly of −0.33°C throughout the peak 2015–2016 El Niño period. The temporal signature of anomalies over Africa shows a much greater level of coherence than the spatial patterns. All four reanalyses record a peak, short-lived positive temperature anomaly in February 2016, with ERA-Interim, MERRA2 and JRA55 recording the magnitude of this peak as 1.7°C above the long-term average.

In Southeast Asia, there is again modest spatial coherence between reanalyses. All analyses show strong positive maximum anomalies in peninsular southeast Asia (Vietnam-Thailand), and weaker and less coherent anomalies in insular Asia. MERRA2 and JRA55 record nearly average conditions prevailing over large parts of the maritime continent, while ERA-Interim records a widespread positive temperature anomaly that, averaged over all forest areas, reaches 0.7°C. CFSR shows some localized small positive temperature anomalies around southern Borneo, but otherwise records the El Niño period as a relatively mild event (0.37°C). The temporal signatures show a high degree of covariance and reveal a double peak in temperature anomalies, the first occurring in December 2016, and the second in April 2016.

(ii) Precipitation

In contrast to the temperature datasets, there is generally good agreement in the spatial pattern of mean precipitation anomalies, with three of the four rainfall datasets showing substantial rainfall deficits in the northeast of Amazonia, over eastern Borneo, portions of Sumatra and southern New Guinea (between −50 and −100 mm/month). Mean and maximum precipitation anomalies for the 2015–2016 El Niño are presented in figure 1*d*,*e*, and the zonally averaged time-series of precipitation anomalies for each continent are shown in figure 1*f*. PERSIANN-CDR is the only product to depart from this trend, modelling instead an extensive and substantial negative precipitation anomaly over the south and central portions of Amazonia and a neutral-to-positive rainfall anomaly over northeast Amazonia. Africa appears to have largely avoided any negative precipitation anomaly during this El Niño event with the mean anomaly ranging from −7.3 mm/month according to CHIRPS, to −0.8 mm/month according to PERSIANN-CDR (table 1). Maximum precipitation anomalies during the El Niño period show greater spatial variability between products than the mean anomaly, with large regions of Amazonia and Africa showing greater than 75 mm/month differences in total precipitation for the driest month. However, when these differences are averaged over each continent, with the exception of PERSIANN-CDR, the magnitude of maximum precipitation anomalies broadly align (table 1).

The time-series in figure 1*f* reveal that each of the rainfall datasets strongly covaries. South America experienced a significant and prolonged dry spell from August 2015 until January 2016, African conditions were approximately neutral throughout the El Niño period and Southeast Asia experienced overall drier conditions from February 2015 until November 2015. Driven by an apparently very large negative precipitation deficit over the southern half of Amazonia, PERSIANN-CDR simulated an extreme negative precipitation anomaly in January, February and March of 2016.

(b) Station assessment

(i) 2015–2016 El Niño

To evaluate which products most reliably capture the weather of the 2015–2016 El Niño in the tropical forest biome, station observations were filtered to the 2015–2016 El Niño period and compared to the estimates from each climate product. Tables 2 and 3 present the agreement statistics for temperature and precipitation, respectively (corresponding scatterplots are available in electronic supplementary material, figure S2).

The near-surface air temperature results indicate that ERA-interim best modelled monthly mean temperatures during the 2015–2016 El Niño. On all three tropical continents, ERA-Interim records the lowest MAE over Africa and Southeast Asia, it achieves the highest level of agreement of the four products ($d_r = 0.46$ and 0.42, respectively) and has a similar level of agreement to MERRA2 over South America ($d_r = 0.52$ versus $d_r = 0.55$). Most products appear to have a

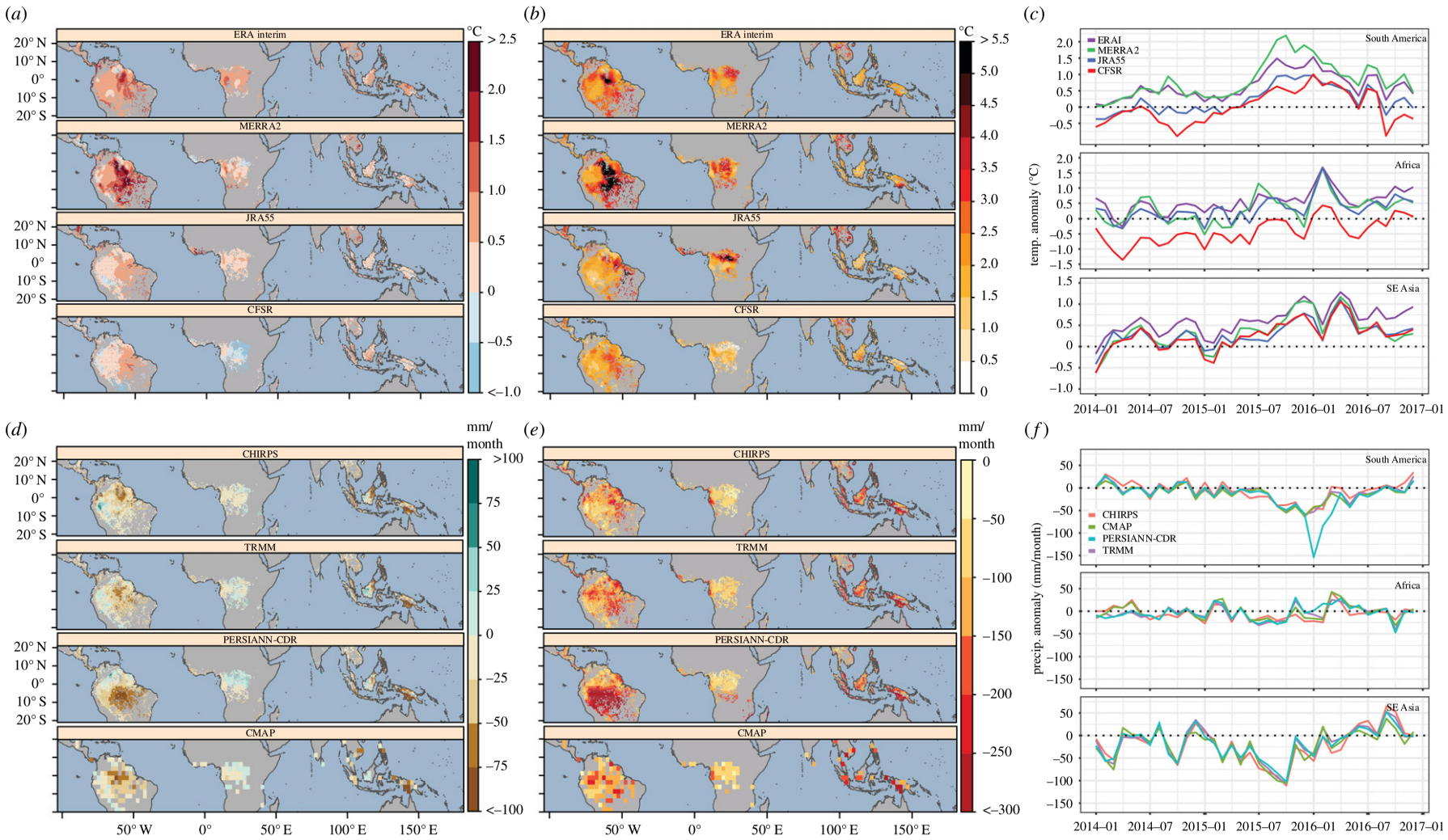

Figure 1. Spatio-temporal pattern of climate anomalies during the 2015–2016 El Niño (defined here as January 2015 to May 2016) from four state-of-the-art reanalyses products and four rainfall products. (*a*) Mean monthly temperature anomalies for each pixel, (*b*) the maximum monthly temperature anomaly for each pixel, (*c*) time-series of zonally averaged monthly mean temperature anomalies from January 2014 to December 2016, (*d*) mean monthly precipitation anomalies for each pixel, (*e*) the maximum monthly precipitation anomalies for each pixel, (*f*) time-series of zonally averaged monthly precipitation anomalies from January 2014 to December 2016. All temperature figures are based on monthly mean temperatures from the near-surface (2 m), and anomalies are calculated against a 1980–2016 base period. All precipitation figures are based on total monthly precipitation and a reference period of 1998–2016 is used.

Table 1. Zonal statistics for the mean and maximum temperature and precipitation anomalies during the 2015–2016 El Niño. These values correspond to the mean of all pixels within each continental region shown in figure 1*a,b,d,e* delineating the lowland tropical forest biome.

	South America		Africa		Southeast Asia	
product	mean	max	mean	max	mean	max
temperature (°C)						
ERAI	0.87	2.23	0.66	1.86	0.70	1.66
MERRA2	1.13	3.00	0.39	2.04	0.51	1.72
JRA55	0.47	1.75	0.40	2.03	0.39	1.58
CFSR	0.33	1.53	−0.33	0.78	0.37	1.40
precipitation (mm/month)						
TRMM	−22.7	−137.6	−3.3	−93.0	−32.8	−186.1
CHIRPS	−20.8	−123.7	−7.3	−78.3	−41.8	−187.1
PERSIANN-CDR	−34.9	−190.2	−0.8	−94.0	−33.8	−169.9
CMAP	−24.9	−125.3	−2.3	−77.1	−38.3	−187.7

systematic cold bias of approximately 1–2°C with the slopes of the least-squares regression commonly less than 1 and scatter points routinely below the 1 : 1 line, apart from MERRA2, which appears to have substantially over-estimated temperatures at the high end of the temperature distribution (slope = 1.27). Of the four products evaluated, CFSR performed the worst on all statistical metrics across each continent, indicating that its ability to model near-surface temperatures during the 2015–2016 El Niño is severely limited.

The results for total monthly rainfall indicate that, over South America and Southeast Asia, TRMM achieved the closest level of agreement and the lowest mean error, with results closely conforming to observations (table 3). CHIRPS also closely tracked the rain gauge data from Southeast Asia with d_r, R^2 and error scores similar to TRMM. Rainfall records in Africa during the El Niño period are extremely limited, with only 23 observations recorded across six stations, with half of our data coming from a single station in the Lopé National Park, Gabon. Hence, the assessment statistics should be viewed with caution. Nevertheless, on the basis of these results CHIRPS performs best, with a d_r score of 0.78 and an MAE 10 mm/month lower than TRMM. PERSIANN-CDR performed substantially worse than TRMM and CHIRPS on all three continents.

(ii) 1980–2016

The performance of the four reanalysis temperature datasets over the full satellite observation period is presented in table 4. As with the results of the station assessment during the 2015–2016 El Niño, ERA-Interim and MERRA2 perform substantially better than JRA55 and CFSR. On all three tropical continents, MERRA2 achieves the lowest MAE (1.22°C in South America, 1.41°C in Africa and 1.47°C in Southeast Asia) and MERRA2 and ERA-interim achieve a very similar level of agreement. Again, there is predominately a cold bias in these temperature datasets on the order of 1–2°C, with the cold bias greatest over Southeast Asia. Over South America, MERRA2 and CFSR both substantially overestimate temperatures above the 95th percentile (see Q–Q plots in electronic supplementary material, figure S4). Most of the errors in these datasets are systematic rather than random, except for MERRA2 over South America where 61% of the MSE is attributable to random errors.

Results for the precipitation datasets are shown in table 5. CHIRPS performs substantially better than both TRMM and PERSIANN-CDR over South America and Africa, achieving a high level of agreement ($d_r = 0.77$ and 0.80, respectively) and relatively low mean error (less than 50 mm/month). Over Southeast Asia, TRMM and CHIRPS achieve very similar scores across all statistical metrics. PERSIANN-CDR performs similarly to the other products over South America but performs substantially worse over Africa and Southeast Asia. All the products tend to considerably underestimate rainfall in high rainfall months, and marginally overestimate lower monthly rainfall totals (electronic supplementary material, figure S4). Most values depart from the 1 : 1 line at the 90th percentile, in accordance with findings elsewhere [3,18,37]. Over Africa, PERSIANN-CDR grossly underestimates rainfall beyond the 75th percentile. The errors in these datasets are generally systematic, except for PERSIANN-CDR in Southeast Asia where the systematic and unsystematic errors are approximately equal.

4. Discussion and conclusion

(a) Characterizing the 2015–2016 El Niño

(i) Temperature

The results of the El Niño period station assessment indicate that ERA-Interim and MERRA2 (up to the 95th percentile) provide the best overall guide for characterizing the spatio-temporal pattern of tropical temperature anomalies during the 2015–2016 El Niño, and we take these products as our best description of the climate characteristics of the El Niño in tropical forest regions. Over South America, a widespread positive temperature anomaly of at least 1°C occurred over the eastern half of Amazonia, while smaller regions in the northeast and southeast probably reached a mean temperature anomaly closer to 1.5°C. During late 2015 and early 2016, the northeast of Amazonia probably experienced more extreme conditions with monthly mean temperatures greater than 3°C above average. The period of peak positive

Table 2. Agreement statistics between observed and reanalysis temperatures during the 2015–2016 El Niño. Bold typeface indicates the highest-performing product for each metric. Scatterplots for these statistics are included in the electronic supplementary material, figure S2.

	South America				Africa				Southeast Asia			
	ERAI	MERRA2	JRA55	CFSR	ERAI	MERRA2	JRA55	CFSR	ERAI	MERRA2	JRA55	CFSR
n	209				246				134			
obs. mean and s.d. (°C)	27.09 ± 1.03				26.51 ± 1.71				27.81 ± 1.13			
mean (°C)	26.33	26.95	25.53	25.41	25.11	25.09	25.04	24.35	26.75	26.50	26.04	26.18
s.d. (°C)	1.11	1.95	1.08	1.64	1.14	1.26	1.42	1.38	0.87	1.23	1.41	1.38
slope	0.77	1.27	0.84	1.13	0.38	0.35	0.28	0.28	0.49	0.55	0.74	0.78
R^2	0.62	0.51	**0.65**	0.54	**0.43**	0.24	0.14	0.13	**0.41**	0.25	0.35	0.40
MAE (°C)	**0.90**	1.00	1.56	1.72	**1.61**	1.71	1.96	2.37	**1.16**	1.35	1.81	1.68
d_r	0.52	**0.55**	0.37	0.38	**0.46**	0.45	0.39	0.37	**0.42**	**0.42**	0.34	0.34
MSE_s/MSE (%)	63.2	5.3	85.7	69.9	86.0	75.6	74.3	82.0	76.4	63.7	71.4	70.8
MSE_u/MSE (%)	36.8	94.7	14.3	30.1	14.0	24.4	25.7	18.0	23.6	36.3	28.6	29.2

Table 3. Agreement statistics between station and remotely sensed precipitation during the 2015–2016 El Niño. Bold typeface indicates the highest-performing product for each metric. Scatterplots for these statistics are included in the electronic supplementary material, figure S2.

	South America			Africa			Southeast Asia		
	TRMM	CHIRPS	PERSIANN	TRMM	CHIRPS	PERSIANN	TRMM	CHIRPS	PERSIANN
n	90			23			143		
obs. mean & s.d. (mm/month)	239.6 ± 144.4			230.9 ± 182.4			245.4 ± 207.8		
mean (mm/month)	213.0	208.4	201.9	175.5	197.0	156.2	223.0	226.7	236.0
s.d. (mm/month)	132.5	124.0	142.5	126.3	152.9	107.4	181.3	157.8	139.4
slope	0.7	0.6	0.7	0.6	0.7	0.4	0.8	0.7	0.5
R^2	**0.66**	0.52	0.49	0.63	**0.70**	0.39	**0.85**	0.79	0.46
MAE (mm/month)	**54.5**	61.6	72.0	78.4	**68.4**	105.1	**59.7**	61.0	98.4
d_r	**0.76**	0.71	0.68	0.73	**0.78**	0.64	**0.79**	0.78	0.62
MSE_s/MSE (%)	19.4	32.3	18.8	56.1	28.1	69.4	28.2	47.4	56.0
MSE_u/MSE (%)	80.6	67.7	81.2	43.9	71.9	30.6	71.8	52.6	44.0

Table 4. Agreement statistics between station and reanalysis temperatures from 1980 to 2016. Bold typeface indicates the highest-performing product for each metric on each continent. Scatterplots for these statistics are included in the electronic supplementary material, figure S3.

	South America				Africa				Southeast Asia			
	ERAI	MERRA2	JRA55	CFSR	ERAI	MERRA2	JRA55	CFSR	ERAI	MERRA2	JRA55	CFSR
n	6814				4437				2496			
obs. mean and s.d. (°C)	26.71 ± 1.15				25.61 ± 1.64				27.52 ± 0.76			
mean (°C)	25.5	25.9	25.2	25.3	24.0	24.5	24.3	24.3	26.0	26.1	25.9	25.5
s.d. (°C)	1.09	1.44	1.04	1.72	1.21	1.44	1.38	1.50	0.95	0.91	0.92	1.03
slope	0.54	0.60	0.38	0.76	0.47	0.41	0.31	0.32	0.60	0.61	0.65	0.68
R^2	**0.37**	0.25	0.20	0.28	**0.42**	0.26	0.17	0.14	0.23	0.25	**0.30**	0.26
MAE (°C)	1.26	**1.22**	1.64	1.72	1.51	**1.41**	1.59	1.63	1.52	**1.47**	1.66	2.01
d_r	0.43	**0.45**	0.33	0.38	**0.48**	0.47	0.42	0.43	0.27	**0.28**	0.25	0.22
MSE_s/MSE (%)	71.1	39.0	78.6	50.6	76.2	60.2	66.4	59.2	76.2	77.0	82.2	83.4
MSE_u/MSE (%)	28.9	61.0	21.4	49.4	23.8	39.8	33.6	40.8	23.8	23.0	17.8	16.6

Table 5. Agreement statistics between station and remotely sensed precipitation from 1980–2016 (1998–2016 for TRMM). Bold typeface indicates the highest-performing product for each metric on each continent. Scatterplots for these statistics are included in the electronic supplementary material, figure S3.

	South America			Africa			Southeast Asia		
	TRMM	CHIRPS	PERSIANN	TRMM	CHIRPS	PERSIANN	TRMM	CHIRPS	PERSIANN
n	2328	34448		350	976		2301	4431	
obs. mean & s.d. (mm/month)	213.3 ± 129.5	193.5 ± 134.1		196.9 ± 178.1	172.1 ± 165.7		249.8 ± 177.4	243.0 ± 173.3	
mean (mm/month)	214.5	196.7	206.0	164.3	166.7	143.3	249.7	256.6	234.7
s.d. (mm/month)	123.6	117.4	119.4	131.0	141.8	102.0	150.8	148.2	121.9
slope	0.79	0.76	0.72	0.64	0.79	0.44	0.73	0.73	0.50
R^2	0.70	**0.75**	0.66	0.66	**0.82**	0.49	**0.73**	**0.73**	0.51
MAE (mm/month)	51.7	**47.3**	58.1	62.9	**46.4**	67.9	67.6	**65.4**	88.1
d_r	0.75	**0.77**	0.72	0.75	**0.80**	0.67	0.73	**0.74**	0.62
MSE_s/MSE (%)	17.9	23.7	24.3	36.5	22.5	63.1	26.5	28.2	50.8
MSE_u/MSE (%)	82.1	76.3	75.7	63.5	77.5	36.9	73.5	71.8	49.2

temperature anomalies in South America began in July of 2015 and extended until mid-2016.

The African temperature patterns are harder to resolve as the divergence between products here was the greatest. Again though, ERA-Interim provides the best guide and it is likely that some regions of the Congo basin reached a mean anomaly of 0.5–1°C. Temporally, the products all agree that African tropical forests experienced a short-lived temperature spike in February 2016, but the spatial extent of this anomaly is difficult to resolve and underscores the need for a more extensive on-the-ground climate monitoring network.

In Southeast Asia, regions around Laos and Cambodia probably experienced a mean positive temperature anomaly approaching 1°C. The magnitude of other positive temperature anomalies shows a high degree of inconsistency between products, making it difficult to conclude which regions experienced the most temperature stress. However, averaged over the entire region, Southeast Asia probably experienced a positive temperature anomaly of 0.5–1°C. Positive temperature anomalies occurred from November 2015 until May 2016, with a period of recovery to nearly average temperatures during February 2016.

(ii) Precipitation

In contrast to the reanalysis temperature datasets, three of the four precipitation datasets were relatively consistent, and the station assessment suggests close agreement with rainfall gauge observations. Substantial rainfall deficits occurred in the northeast and southeast of Amazonia, over eastern Borneo, portions of Sumatra and southern New Guinea. Much of Central Africa appears to have largely avoided any widespread or prolonged precipitation deficit during this El Niño, but moderate deficits were observed in coastal West and Central Africa. Results for Central Africa come with the caveat that very few gauge observations are used for bias correction, so the uncertainty of the products is probably high over this region. The timing of these rainfall deficits occurred from August 2015 until January 2016 in South America, and Southeast Asia experienced overall drier conditions from February 2015 until November 2015. In Southeast Asia, the negative precipitation anomaly occurred before the positive temperature anomaly, with rainfall largely recovering by the time temperatures peaked. Over South America, the temperature and precipitation anomalies both temporally and spatially coincide, thereby probably exacerbating the negative impact of El Niño conditions on the tropical forests of this region. Given the spatio-temporal coherence of climate anomalies over northeastern Amazonia during the 2015–2016 El Niño, this is likely to be a region where impacts on the tropical forest carbon cycle are most pronounced.

(b) Quality of gridded climate products in the tropics

The findings from the longer-term station assessment indicate that ERA-Interim and MERRA2 are the most reliable near-surface reanalysis air temperature products over lowland tropical forests, though even these datasets generally have a relatively low level of agreement with on-the-ground observations. JRA55 and CFSR both perform significantly worse and are not recommended for use in the tropics. Borrowing from conclusions reached elsewhere [11,38], CFSR probably performs poorly due to the reanalysis being executed via six parallel-runs covering different periods, which were then joined to produce the final dataset. JRA55 excludes all atmospheric pressure data over South America to counter a dry bias over the Amazon Basin that plagued their earlier reanalysis, JRA25, and undoubtedly this contributes to its relatively poor performance in this region. The systematic cold bias seen in all reanalyses may be closely related to the different estimates of rainfall in the reanalysis products. For example, relative to rainfall in the Global Precipitation Climatology Project (GPCP) dataset, ERA-Interim significantly overestimates tropical rainfall [10]. The extra cloud-cover associated with enhanced precipitation reduces modelled solar radiation reaching the surface and hence suppresses tropical temperatures [39]. Where these biases are systematic and uniform, it may be possible to perform a simple bias correction using a 1–2°C offset to bring reanalysis temperatures into line with station observations within a region of interest.

Three high-resolution combined satellite-gauge rainfall products, TRMM 3B42 V7 Daily, CHIRPS V2.0 and PERSIANN-CDR were evaluated to reveal which of these products best-estimate rainfall over tropical forests at the continental scale. Point-to-pixel analysis between satellite estimates and station measurements reveal that CHIRPS performs best over South America and tropical Africa, and TRMM and CHIRPS perform similarly in Southeast Asia. PERSIANN-CDR generally performs significantly worse than the other products, and in South America, where it performs comparably with TRMM and CHIRPS, it provides a spurious precipitation deficit during the 2015–2016 El Niño. It is therefore not recommended that PERSIANN-CDR be used for analysis over tropical forests. All the products tend to significantly underestimate rainfall totals in high rainfall months and slightly overestimate rainfall in very low rainfall months. These datasets should therefore be used with caution when attempting to characterize droughts through metrics such as the Climatological Water Deficit, as rainfall estimates may fail to capture heavy rainfall events that reset moisture deficits.

In conclusion, we have evaluated a range of global temperature and precipitation products over the tropical forest biome, compared these products against ground observation and used the best performing of these products to draw the most likely general spatial and temporal characteristics of the 2015–2016 El Niño over the tropical forest biome. Our analysis demonstrates that caution is needed in the use and interpretation of these products, but overall conclusions about the spatio-temporal footprint of the El Niño can be drawn.

Data accessibility. Code and data are available from this link: https://github.com/cbur24/Burton_et_al_2018_data-code.

Authors' contributions. C.B. performed statistical analysis and geospatial analysis, and drafted the paper. S.R. contributed to the conception, design and interpretation of the article, collected datasets and helped draft the article. Y.M. contributed to the conception, design and interpretation of the article and helped draft the article. All authors gave final approval for publication.

Competing interests. C.B. and S.R. have a professional relationship with Y.M., who is a guest editor on this issue of *Philosophical Transactions of the Royal Society B*.

Funding. This work was supported by UK Natural Environment Research Council [grant no. NE/P001092/1] to Y.M., a grant from the Nature Conservancy supporting S.R. and a European Research Council Advanced Investigator Award (GEM-TRAIT, # 321131) to Y.M. Y.M. is also supported by the Jackson Foundation.

Acknowledgements. We thank Dr Sihan Li, Dr Sebastian Engelstaedter and Dr Karsten Haustein for their advice and guidance in compiling the climate datasets used here and interpreting the results.

References

1. Fujiwara M *et al*. 2017 Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *ASTM Spec. Tech. Publ.* **17**, 1417–1452. (doi:10.5194/acp-17-1417-2017)
2. Dee D. 2017 The climate data guide: ERA-Interim. Last modified 21 July 2017. See https://climatedataguide.ucar.edu/climate-data/era-interim.
3. Beck HE *et al*. 2017 Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.* **21**, 6201. (doi:10.5194/hess-21-6201-2017)
4. Trenberth KE, Stepaniak DP, Hurrell JW, Fiorino M. 2001 Quality of reanalyses in the tropics. *J. Clim.* **14**, 1499–1510. (doi:10.1175/1520-0442(2001)014<1499:QORITT>2.0.CO;2)
5. Lorenz C, Kunstmann H. 2012 The hydrological cycle in three state-of-the-art reanalyses: Intercomparison and performance analysis. *J. Hydrometeorol.* **13**, 1397–1420. (doi:10.1175/JHM-D-11-088.1)
6. Wang J, Zeng N, Wang M. 2016 Interannual variability of the atmospheric CO_2 growth rate: roles of precipitation and temperature. *Biogeosciences* **13**, 2339–2352. (doi:10.5194/bg-13-2339-2016)
7. Wang W *et al*. 2013 Variations in atmospheric CO_2 growth rates coupled with tropical temperature. *Proc. Natl Acad. Sci. U.S.A.* **110**, 13 061–13 066. (doi:10.1073/pnas.1219683110)
8. Qian H, Joseph R, Zeng N. 2008 Response of the terrestrial carbon cycle to the El Niño-Southern Oscillation. *Tellus B* **60**, 537–550. (doi:10.1111/j.1600-0889.2008.00360.x)
9. Zeng N, Mariotti A, Wetzel P. 2005 Terrestrial mechanisms of interannual CO_2 variability. *Global Biogeochem. Cycles* **19**, GB002273. (doi:10.1029/2004GB002273)
10. Dee D *et al*. 2011 The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. R. Meteorol. Soc.* **137**, 553–597. (doi:10.1002/qj.828)
11. Saha S *et al*. 2010 The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* **91**, 1015–1057. (doi:10.1175/2010BAMS3001.1)
12. Saha S *et al*. 2014 The NCEP climate forecast system version 2. *J. Clim.* **27**, 2185–2208. (doi:10.1175/jcli-d-12-00823.1)
13. Gelaro R *et al*. 2017 The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454. (doi:10.1175/JCLI-D-16-0758.1)
14. Harada Y *et al*. 2016 The JRA-55 reanalysis: representation of atmospheric circulation and climate variability. *J. Meteorol. Soc. Jpn. Ser. II* **94**, 269–302. (doi:10.2151/jmsj.2016-015)
15. Kobayashi S *et al*. 2015 The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteorol. Soc. Jpn. Ser. II* **93**, 5–48. (doi:10.2151/jmsj.2015-001)
16. Decker M, Brunke MA, Wang Z, Sakaguchi K, Zeng X, Bosilovich MG. 2012 Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *J. Clim.* **25**, 1916–1944. (doi:10.1175/JCLI-D-11-00004.1)
17. Trenberth KE, Fasullo JT, Mackaro J. 2011 Atmospheric moisture transports from ocean to land and global energy flows in reanalyses. *J. Clim.* **24**, 4907–4924. (doi:10.1175/2011JCLI4171.1)
18. Trejo FJ. P., Barbosa HA, Peñaloza-Murillo MA, Moreno MA, Farías A. 2016 Intercomparison of improved satellite rainfall estimation with CHIRPS gridded product and rain gauge data over Venezuela. *Atmósfera* **29**, 323–342. (doi:10.20937/ATM.2016.29.04.04)
19. Dinku T, Ceccato P, Grover-Kopec E, Lemma M, Connor S, Ropelewski C. 2007 Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.* **28**, 1503–1526. (doi:10.1080/01431160600954688)
20. Tan ML, Ibrahim AL, Duan Z, Cracknell AP, Chaplot V. 2015 Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sens.* **7**, 1504–1528. (doi:10.3390/rs70201504)
21. Melo DDC, Xavier AC, Bianchi T, Oliveira PT, Scanlon BR, Lucas MC, Wendland E. 2015 Performance evaluation of rainfall estimates by TRMM multi-satellite precipitation analysis 3B42V6 and V7 over Brazil. *J. Geophys. Res. Atmos.* **120**, 9426–9436. (doi:10.1002/2015JD023797)
22. Dembélé M, Zwart SJ. 2016 Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. *Int. J. Remote Sens.* **37**, 3995–4014. (doi:10.1080/01431161.2016.1207258)
23. Liu Z, Ostrenga D, Teng W, Kempler S. 2012 Tropical rainfall measuring mission (TRMM) precipitation data and services for research and applications. *Bull. Am. Meteorol. Soc.* **93**, 1317–1325. (doi:10.1175/BAMS-D-11-00152.1)
24. Funk C *et al*. 2015 The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes. *Sci. Data* **2**, 150066. (doi:10.1038/sdata.2015.66)
25. Ashouri H, Hsu K-L, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, Nelson BR, Prat OP. 2015 PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* **96**, 69–83. (doi:10.1175/BAMS-D-13-00068.1)
26. Xie P, Arkin PA. 1997 Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Am. Meteorol. Soc.* **78**, 2539–2558. (doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2)
27. Rayner N, Parker DE, Horton E, Folland C, Alexander L, Rowell D, Kent E, Kaplan A. 2003 Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res. Atmos.* **108**, 4407. (doi:10.1029/2002JD002670)
28. Hansen MC *et al*. 2013 High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853. (doi:10.1126/science.1244693)
29. Kharin VV, Zwiers FW, Zhang X. 2005 Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations. *J. Clim.* **18**, 5201–5223. (doi:10.1175/JCLI3597.1)
30. Rudolf B, Rubel F. 2005 *Global precipitation*, pp. 11.11–11.53. Berlin, Germany: Springer.
31. Franchito SH, Rao VB, Vasques AC, Santo CM, Conforte JC. 2009 Validation of TRMM precipitation radar monthly rainfall estimates over Brazil. *J. Geophys. Res. Atmos.* **114**. (doi:10.1029/2007JD009580)
32. Lawrimore JH, Menne MJ, Gleason BE, Williams CN, Wuertz DB, Vose RS, Rennie J. 2011 An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *J. Geophys. Res. Atmos.* **116**. (doi:10.1029/2011JD016187)
33. Osborn T, Jones P. 2014 The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. *Earth Syst. Sci. Data* **6**, 61–68. (doi:10.5194/essd-6-61-2014)
34. Harris I, Jones P, Osborn T, Lister D. 2014 Updated high-resolution grids of monthly climatic observations – the CRU TS3. 10 Dataset. *Int. J. Climatol.* **34**, 623–642. (doi:10.1002/joc.3711)
35. Willmott CJ, Robeson SM, Matsuura K. 2012 A refined index of model performance. *Int. J. Climatol.* **32**, 2088–2094. (doi:10.1002/joc.2419)
36. Willmott CJ, Ackleson SG, Davis RE, Feddema JJ, Klink KM, Legates DR, O'donnell J, Rowe CM. 1985 Statistics for the evaluation and comparison of models. *J. Geophys. Res. Oceans* **90**, 8995–9005. (doi:10.1029/JC090iC05p08995)
37. Toté, C., Patricio D, Boogaard H, van der Wijngaart R, Tarnavsky E, Funk C. 2015 Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique. *Remote Sens.* **7**, 1758–1776. (doi:10.3390/rs70201758)
38. Wang A, Zeng X. 2012 Evaluation of multireanalysis products with in situ observations over the Tibetan Plateau. *J. Geophys. Res. Atmos.* **117**.
39. Betts AK, Ball JH, Viterbo P, Dai A, Marengo J. 2005 Hydrometeorology of the Amazon in ERA-40. *J. Hydrometeorol.* **6**, 764–774. (doi:10.1175/JHM441.1)