The impact of the 2023-2024 drought on intact Amazon forests' productivity

Felicien Meunier

felicien.meunier@gmail.com

Ghent University https://orcid.org/0000-0003-2486-309X

Pascal Boeckx

Ghent University https://orcid.org/0000-0003-3998-0010

Santiago Botía

Max Planck Institute for Biogeochemistry

Marijn Bauters

Ghent University https://orcid.org/0000-0003-0978-6639

Wout Cherlet

Ghent University

Philippe Ciais

Laboratoire des Sciences du Climat et de l'Environnement https://orcid.org/0000-0001-8560-4943

Steven De Hertog

Ghent University

Michael Dietze

Boston University

Marc Peaucelle

INRAE, Université de Bordeaux https://orcid.org/0000-0003-0324-4628

Thomas Sibret

Ghent University

Stephen Sitch

University of Exeter https://orcid.org/0000-0003-1821-8561

Wei Li

Tsinghua University https://orcid.org/0000-0003-2543-2558

Hans Verbeeck

Ghent University https://orcid.org/0000-0003-1490-0168

Article

Keywords:

Posted Date: August 26th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4705191/v1

Additional Declarations: There is **NO** Competing Interest.

Abstract

In the Amazon, the dry season of 2023 as well as the beginning of the wet season in 2024 were marked by unprecedented high temperatures and large precipitation deficits. While the tropical forests in the Amazon play a crucial role in the global carbon cycle and are a biodiversity hotspot, they were also shown to suffer from El-Niño related droughts in the past, leading to legitimate concerns about the ecological consequences of the recent climate conditions. To this day, while there is a growing effort to make remote sensing products available close to real-time, land surface models that are critical tools to understand the interactions between the biosphere and the environment have lagged behind the present due to the complexity to run and process large model ensembles. In this study, we employed advanced machine learning models trained on state-of-the-art remote sensing and dynamic global vegetation model estimates of gross primary productivity (GPP). The models provide near real-time GPP estimates, revealing significant productivity reductions during the 2023/2024 drought. Negative GPP anomalies were more widespread across the Amazon than during any other recent major drought event. The Climate-GPP relationships that emerged from the models suggest that future temperature increases and changes in precipitation will severely challenge Amazon forest resilience.

Introduction

The Amazon hosts the World's largest contiguous tropical forest[1,2], and the largest expanse of intact tropical forests (i.e. forests unaffected by disturbances caused by human activities in the recent past)[3]. Amazon tropical forests play a crucial role in global biogeochemical cycles[4], e.g. being responsible for 25% of the global land carbon sink and storing approximately 120 petagrams of carbon[2,5], a large fraction of which is in aboveground tree biomass[6]. Intact forests' critical contribution to climate change mitigation has been frequently acknowledged as they are among the most carbon dense[7] and productive[8] ecosystems on Earth. They are also critical biodiversity hotspots[9].

However, Amazon intact forests are under pressure[10] and have experienced a persistent decline in their sink capacity in the past decades[11], potentially attributable to drought-induced increases in tree mortality[12,13]. Climate change, including the increasing intensity and severity of extreme events, may play an important role in this progressive shift of intact Amazon forests from a carbon sink to a carbon source, as extreme productivity and mortality anomalies notably occur during or directly after extremely hot and dry events such as the large-scale drought of 2015/2016[14,15]. In 2023, the Amazon was struck by an unprecedented drought event and extreme heat wave which started in the dry season[16], of which the ecological consequences on tropical forests are yet to be determined.

Multiple methods exist to monitor the productivity of tropical forests[17]. Plot (re-)censusing is the most direct way to coincidently measure forest woody productivity and mortality[17] but requires extensive networks of large plots to be accurately representative of a vast biome like the Amazon rainforests, which makes this technique particularly labour-intensive and time-consuming. In addition, plot

recensusing has typically a low periodicity (e.g. 5 years) which prevents observations of inter-annual variability[14]. Eddy-covariance flux towers directly measure in situ the net CO_2 exchange between the atmosphere and the ecosystem, but their spatial and temporal coverage in the Tropics is limited due to the cost and the technical difficulties of their implementation and maintenance. Remote sensing[18] (RS), inverse methods[19,20], land data assimilation[21], and dynamic global vegetation models[20] (DGVM) can provide large-scale estimates of land carbon fluxes.

While more and more remote sensing products are made available close to real time, significant delays remain for new model output release due to the lengthy process of new data acquisition, processing, computing, and formatting even when no new developments are required. This delay is further exacerbated when multiple RS products or DGVMs are combined in ensembles to estimate uncertainties and produce an ensemble mean likely closer to the ground truth[22]. Those technical workflows typically translate to months or years of lag between the occurrence of a climate event like the 2023 drought and the quantification of its ecological impact by DGVMs. This hinders near real time assessment of extreme events such as the extreme climatic conditions experienced in 2023/2024, which is necessary for early-warning systems to inform policy-makers, accelerate research and make it more relevant to society[23].

Here, we used a new near-real time approach to achieve a very fast evaluation of large-scale impacts of climate events. We trained advanced gradient-boosted trees (GBT) models[24] to emulate multiple RS products and DGVM estimates of gross primary productivity (GPP) using climatic variables from state-of-the-art reanalyses (ERA5, JRA-55) as explanatory features. We focused our analysis on intact forests as defined by Potapov et al. (2008)[25] in the Amazon basin[26] and used recent DGVM simulations made for the global carbon assessment (TRENDY[27]) as training data. The DGVM outputs used for training were only available until 2021 but the GBT models allowed us to extend the time series until near real-time (May 2024), revealing the ecological consequences of the 2023 El-Niño year for the Amazon rainforests in near real-time. This opens new avenues for a broad range of vegetation ecology applications, including early warning signalling systems.

Main

The Amazon drought of 2023/2024 for intact Amazon forests

According to ERA5 re-analyses, the period between July 2023 and April 2024 was the hottest in at least the past 30 years over intact Amazon forests (see methods for definition and mask of intact forests of the Amazon) with an average air temperature of 26.7°C (Fig. 1), 1.4°C above the the past 30 years average (25.3°C) and 1.3°C above the average value for the July-May window of the last three decades (25.4°C). July 2023 - April 2024 was on average 0.5°C and 0.1°C above the El Niño related drought events of September 1997 - April 1998 (26.2°C) and August 2015 - March 2016 (26.6°C). Over the 10-month time period between July 2023 and April 2024, the precipitation was also below average with a monthly mean of 142 mm $month^{-1}$, which is 47 mm $month^{-1}$ below its average for that period of the year (Fig. 1a

and c, supplementary Figure S1). The precipitation anomaly was larger than during previous drought events, which reached − 30 mm month^{-1} between August 2015 and March 2016 and − 29 mm month^{-1} between September 1997 and April 1998 (Fig. 1a and c). The hot and dry episode of 2023–2024 was also characterised by exceptionally high monthly minimum (23.1°C) and maximum (31.9°C) temperature (0.9°C and 1.8°C, respectively, above the 1994–2023 averages).

In October 2023, the monthly air temperature reached its peak with an average 27.9°C, 1.8°C above the monthly mean of the last three decades (Fig. 1b and d). This 1.8°C anomaly corresponded to 3.4 standard deviations of the monthly mean temperature (supplementary Figure S1). In September 2023, the precipitation dropped to an at least 30-years time low with 78 mm, 91 mm below (or 3.2 monthly standard deviations) its monthly average. Both temperature and precipitation anomalies were widespread throughout the Amazon biome (supplementary Figure S2).

The impact of drought on intact Amazon forest productivity

GBT models trained on RS products and TRENDY simulations could successfully reproduce the seasonality, and the long-term trends, both at the regional level and at the local scale, for each and every RS product and DVGM (supplementary Figures S3). GBT models could also accurately predict the impacts of an unprecedented drought like 2015–2016 on Gross Primary Productivity (GPP), when trained with pre-2015 data only.

While GPP has steadily increased over the last three decades according to both RS-trained and TRENDY-trained GBT models (p-value < 0.001 for the slope of the linear models in Fig. 2a), it showed noticeable negative anomalies during the past (1997–1998, 2015–2016) and recent (2023–2024) dry and hot episodes (Fig. 2b). Since July 2023, GBT models forecasted consistently lower GPP compared to the average of the past three decades (Fig. 2a and b). Between July 2023 and April 2024, GBT models projected a GPP reduction of 1.2 Mg C ha^{-1} year^{-1} (RS-trained GBT models) and 1.5 Mg C ha^{-1} year^{-1} (TRENDY-trained GBT models) compared to the seasonal average for all intact forests in the Amazon, which corresponded to an average decline of 2.6 and 2.2 monthly standard deviations (Fig. 2b, supplementary Figure S5). The productivity of 2023–2024 reached its lowest level in October 2023 with RS and TRENDY trained GBT models predicting a respective mean GPP value of 30.7 and 27.9 Mg C ha^{-1} year^{-1}, respectively 2.1 and 3.4 Mg C ha^{-1} year^{-1} lower than their expected value in the absence of drought. It was the least productive October of the past 30 years according to both RS and TRENDY models. This sharp decrease corresponds to a reduction equivalent to 3.9 and 3.4 monthly standard deviations, a three-decades minimum only surpassed during the month of January 2016 for TRENDY-trained GBT models (3.6 monthly standard deviations then, but for a smaller absolute GPP reduction: 2.2 Mg C ha^{-1} year^{-1} in January 2016).

Negative anomalies were distributed all across the Amazon during the 2023–2024 dry and hot episode: 98.2% (RS models) or 98.6% (TRENDY models) of all intact Amazon forest grid cells experienced negative GPP anomalies, on average, over the July 2023 - April2024 period (Fig. 3). It was the most

spatially widespread drought anomaly, surpassing the 1997–1998 drought according to both RS- and TRENDY-trained GBT models (respectively 97.7% and 96.3% of grid cells had negative anomalies then). During the month of October 2023, 78.8%/62.0% (RS-trained GBT models) or 49.3%/42.3% (TRENDY-trained GBT models) of all intact forest grid cells in the Amazon had a negative GPP anomaly corresponding to at least two/three monthly standard deviations.

ERA5 mean temperatures and precipitation, as well as their anomalies, showed strong correlations with those of JRA-55 which also marked October 2023 exceptionally dry and hot (supplementary Figure S3). Hence, repeating the analyses with JRA-55 as climate drivers of the GBT models led to very similar conclusions about both the magnitude of the 2023 dry and hot episode and its spatial distribution (supplementary Figures S6 and S7).

Climate-GPP relationships of the intact Amazon forests

Both RS- and TRENDY-trained GBT models exhibited significant, nonlinear relationships between monthly GPP and mean temperature anomalies (p-value < 0.001, Fig. 4), as well as with VPD and precipitation anomalies (supplementary Figure S8). In the next 30 years (2024–2053), air temperature is expected to increase almost linearly over the study area, with warming rates varying between 0.19°C decade^{-1} (SSP1-2.6) and 0.44°C decade^{-1} (SSP5-8.5), according to a weighted average of CMIP6 model simulation made to reproduce ERA5 temperatures (supplementary Figures S9). In the next three decades, temperature anomalies similar to the 2023–2024 dry and hot episodes are expected to become common for all climatic scenarios (Fig. 4c). By the end of this century, such temperature anomalies are expected to remain the norm for SSP1-2.6 and become cold outliers for all other scenarios (Supplementary Figure S10). These future high temperatures will generate correspondingly high VPD while precipitation is expected to decrease by the end of this century (Supplementary Figure S10), which should further aggravate GPP anomalies.

Discussion

According to the novel model projection tool that we developed, the 2023–2024 dry and hot atmospheric conditions will lead to exceptional ecological consequences for the intact forests of the Amazon. Driven by abnormally high temperatures and precipitation deficits, intact forests' GPP has likely hit a record low anomaly in the month of October 2023. This event could reverberate in the upcoming months and years as drought has been unequivocally linked to tree mortality[13,28–31]. TRENDY (Supplementary Figure S11) and other models[32], in line with field inventory data[14,33], predict significant reductions of the intact Amazon forest carbon sink for previous large-scale drought events, which is likely to repeat given the 2023–2024 drought magnitude.

Regardless of the future emission scenario, temperatures are expected to increase in the next three decades and Amazon intact forest GPP will likely be affected by this temperature increase, the related increase in VPD, and the possible reduction in precipitation. Hence, the CO_2 fertilisation[34], also observed

here in the RS- and TRENDY-trained GBT models, might be outpaced in the near future by the future increase in frequency and severity of droughts, even in the absence of CO_2 saturation for photosynthesis[35]. This applies to GPP and potentially to net ecosystem productivity (NEP), which has shown no increasing trends (p-value 0.19) and strong drought negative responses, according to both TRENDY models (Supplementary Figure S12) and field data[12].

The new method presented in this article is highly flexible, robust (supplementary Figure S3), fast, and easy to use. The models were trained on a specific region but could easily be extended to include more (human-disturbed) biomes, RS products or DGVMs, features (e.g., including sub-monthly variability such as climate extremes as explanatory variables or more features), or even novel types of data (e.g., eddy covariance fluxes, field inventories, or soil moisture observations that were shown to critically impact land carbon fluxes[36]). After training, the GBT models require no expert knowledge and only need the location (latitudes and longitudes), time (years and months), and climate of the region of interest to simulate GPP. It took less than one second on a regular 8-cores laptop with 16 GB of RAM to run any of the RS or TRENDY surrogates for a full year for the entire intact Amazon forests, contrasting with the large (but very model-specific) number of CPU-hours required to run a typical TRENDY model .

Despite these advantages, it is unclear whether the models developed in this study can accurately predict land fluxes under unprecedented climatic circumstances. On the one hand, the emulators were able to reproduce the extreme 2015–2016 dry and hot episode once trained on earlier GPP estimates (supplementary Figure S4). On the other hand, Amazon rainforest tipping points have been postulated[10,37–39] if drought intensity and/or frequency increases beyond some threshold, but so far such threshold responses have not been observed at large scales and hence are not part of the training data. As new RS products or model outputs are released, our GBT model projections can be confirmed or if needed, those newly available data, as well as other covariates, can feed the GBT models as additional training data to improve their robustness. Hybrid AI[40], as well as DGVM future projection[41] under hotter and drier conditions, will help overcome this potential issue.

The methodology suggested here opens the door for a wide range of other potential applications. Because climate variables needed to force those GBT models are widely available including for the future, it makes the development of near-real time ecological forecasting[23] or early warning systems[42] for land fluxes based on RS and DGVM possible. Such a tool could become incredibly useful for decision-makers and conservation policies overall. Such approaches can also help identify (the reasons behind) discrepancies between DVGM outputs and RS products, prioritise necessary model development[43] or refine part of RS algorithms for land flux retrieval.

Methods

Study area

The Amazon basin (https://github.com/gamamo/AmazonBasinLimits) delimited our study area[26]. We further restricted the analyses to intact forests using the intact forest map of Potapov et al. (2008)[25] (http://www.intactforests.org/data.ifl.html) for the year 2020.

TRENDY model simulations

We used simulations of a set of 16 Dynamic Global Vegetation Models (DGVMs), which participated in the past global carbon budgets assessment[27], named TRENDY-v11. Because we focus on intact forests, we extracted model outputs from TRENDY-v11 scenario S2 in which CO_2 and climate vary with time but which uses a time-invariant "pre-industrial" land use mask. In TRENDY, all models are forced with the same climate forcing, namely CRUJRA, see[44–46]. We mainly extracted GPP monthly outputs from TRENDY-v11 but we also pulled monthly averages of NPP and heterotrophic respiration to compute NEP from their difference for every model that provided those variables as well. NEP values were only used as is (i.e. did not serve for GBT model training) to investigate their correlation with GPP. TRENDY-v11 model outputs are available for the time period 1901-2021.

Remote sensing estimates

To benchmark TRENDY model outputs, we used four recent GPP products, hereafter referred to as Li and Xiao (2019)[8], Bi et al. (2022)[47], Wild et al. (2022)[48], and Wang et al. (2021)[49]. Briefly, Li and Xiao (2019) developed a global, high temporal (8-days) and spatial (0.05°) resolutions GPP product based on Orbiting Carbon Observatory-2 SIF data, MODIS and meteorological re-analyses through a data-driven approach over the 2000-2022 period[8]. Bi et al. (2022) generated a global 0.05°, 8-day dataset for GPP (1992-2020) with a two-leaf light use efficiency model, driven by CRUJRA reanalysis, ESA-CCI land cover and leaf area index from GLOBMAP[47]. Wild et al. (2022) utilised microwave remote sensing of vegetation optical depth to derive global GPP at moderate spatial (0.25°) and high temporal (8-days) resolutions[48] for the period 1988-2020. Finally, Wang et al. (2021) correlated eddy-covariance GPP and AVHRR near-infrared reflectance from the Land Long Term Data Record[50] to generate a global long-term (1982-2018) time series at high spatial resolution (0.05°) of monthly GPP. We aggregated all four datasets at the monthly timescale for further analyses.

AI models

We used the extreme gradient boosting (XGBoost[51]) approach, as implemented in the R-package *xgboost*[52], to reproduce time series of GPP as estimated from DGVMs or from remote sensing. XGBoost is an advanced machine learning algorithm reputed to be highly efficient in terms of model performance and computational speed. Individual GBT models were trained for each DGVM/RS estimate of GPP using the longest possible time series and the largest possible area. We splitted all the data into training (60%),

validation (20%) and test (20%) for each grid cell and year. So on average, 2.4 months per year served as validation and test for each grid cell while 7.2 months per year were used for training. The months were randomly attributed to training, validation and test dataset. We optimised/tuned GBT model (hyper)parameters on validation data with the R-package *caret*[53] and 8 resampling iterations, focusing on the maximum decision tree depth and the number of decision trees to grow.

Training variables included monthly averages of climate data (monthly mean, minimum, and maximum temperature, monthly precipitation, VPD, total incoming short- and long-wave radiation), atmospheric CO_2 concentration (constant spatially), as well as temporal (year/month) and spatial (latitude/longitude) information. The latter variables were important to include to account for model-specific spatial variability in e.g., soil maps and plant functional type spatial distribution. We computed VPD from sub-daily air temperatures and specific humidity using the R-package *Pecan.data.atmosphere*[54] before taking the monthly average. Similarly, we used sub-daily air temperature values to compute the daily minimum and maximum temperatures and finally calculate the monthly minimum and maximum temperature as the average of all daily minimum and maximum temperatures of a given month.

The CRUJRA meteorological reanalysis, which is used as the driver for TRENDY models, is only updated once a year (typically between April and August), which prevents its use for near-real time forecasting. Therefore we decided to use ERA5 and JRA-55 reanalysis as input data for the GBT models. Doing so we internalised the bias between reanalysis sources into the GBT models, as CRUJRA is constructed by adjusting data from JRA-55 where possible to align with the CRU TS data. We downloaded all ERA5 and JRA-55 climatic drivers available in early June 2024, i.e. until May 2024 and December 2023 respectively.

We performed the same model training and validation described above, independently with ERA5 and JRA-55 reanalysis to check the influence of climatic forcings on the model predictions. We also checked how absolute values and anomalies correlated for both reanalysis sources for the region of interest. For ERA5, we used the 2-metre air temperature variable for the air temperature.

Final GBT model performance was evaluated on test data only and quantified using the root mean square deviation and the square of the Pearson correlation coefficient (R^2) between reference (DGVM outputs or RS estimates) and predicted values.

Analyses

Given the different resolutions of the individual TRENDY models and the remote sensing GPP estimates, we reprojected the GBT model features (climate forcings) on each DGVM/RS grid before training with the R-package *raster*[55]. To generate spatial ensemble means of GPP, we also reprojected all rasters to a common grid (ERA5) using bilinear interpolations.

When reprojection was needed for the intact forest map, we used bilinear interpolations, introducing decimal numbers. We then limited the analyses to all grid cells with interpolated values larger than 50%

intact forest cover. Unless otherwise stated, we limited our analyses to the past 30 years (1994-2023).

For generating model ensemble means, we applied simple means to all RS products and DGVM outputs, giving equal weight to each ensemble member. We used the GBT models for extending RS time series to recent years (e.g., 2023/2024) but also to older times (from 1994 to the actual start of the product time series) to avoid discontinuities when computing ensemble means when not all products were available.

We compared the recent dry and hot episode of 2023-2024 with previous major drought events, namely August 2015-March 2016, and September 1997-April 1998. We delimited those drought events of the past from their precipitation and air temperature anomalies in ERA5.

To assess whether the GBT models were able to accurately predict GPP in case of unprecedented events, we verified how the density distribution of the residuals compared for the test data of the dry and hot episodes of 1997-1998 and 2015-2016 and all other climatic conditions. We also repeated the model training and validation described above using pre-2015 data only and tested the model performance specifically for the 2015-2016 dry and hot episode.

In both RS estimates and DGVM simulations, GPP presents both a seasonality and an increasing long-term trend due to CO_2 fertilisation. We detrended the GPP time series by subtracting to the raw data the best linear model estimates from the 1994-2023 time period and its average seasonal variation. We used a linear model for detrending as GPP has been steadily increasing the past decades due to CO_2 fertilisation. We detrended the climate variables by subtracting the overall variable average and its seasonal mean to the raw time series. The anomalies were normalised using the monthly standard deviation of the respective variables for the overall time period (1994-2023). For spatial analyses, we applied the detrending method described above to the grid cells individually, while we first averaged the time series for biome-wide analyses. Anomalies were computed on the model ensemble mean and the average of RS products.

To smooth monthly time series, we applied rolling averages with a centred 6-month time window. We tested the effect of this choice by changing the window size between 1 month (no average) and one year and verified how it changed the correlations between variables.

For future trends, we downloaded monthly temperature and precipitation outputs from Earth system models (ESM) available in January 2024 for the models participating in the Coupled Model Intercomparison project 6 (CMIP6)[56]. We downloaded model outputs for the historical period and the following shared socioeconomic pathways[57]: SSP1-2.7, SSP2-4.5, SSP3-7.0, and SSP5-8.5. We restricted the analysis to the models which had model outputs available for every shared socioeconomic pathway. We then compared the regional monthly mean temperature and the mean annual precipitation time series predicted by each ESM for the recent historical period (1985-2014) with those of ERA5, and only kept the 10 with the lowest root mean square error for each variable. We then averaged those 10 models using the inverse of the squared bias as weight for both the historical period and the future simulations. Temperature and precipitation anomalies were defined using 1985-2014 or 1994-2023 as reference for

the historical period or the future scenarios, respectively, since CMIP6 historical simulations stop in 2014.

Temperature anomaly outliers were identified following the classical outlier 1.5 interquartile (IQR) rule[58]: any observation that was smaller than 1.5 IQR below the first quartile or larger than 1.5 IQR above the third quartile was considered as a cold or hot outlier, respectively.

Declarations

Acknowledgements

FM was funded by the FWO as a senior postdoc and is thankful to this organisation for its financial support (FWO grant no.1214723N). Michael C. Dietze was supported by funding from NSF Grant 1638577 and NASA Carbon Monitoring System (80NSSC17K0711). SDH was supported by the BELSPO project DAMOCO (B2/223/P1/DAMOCO). MP acknowledges the financial support of the European Research Council (grant No. 101117001 LEAFPACE).

References

1. Hoang NT, Kanemoto K (2021) Mapping the deforestation footprint of nations reveals growing threat to tropical forests. Nat Ecol Evol 5:845–853
2. Saatchi SS et al (2011) Benchmark map of forest carbon stocks in tropical regions across three continents. *Proc. Natl. Acad. Sci.* 108, 9899–9904
3. Schulze K, Malek Ž, Verburg PH (2019) Towards better mapping of forest management patterns: A global allocation approach. Ecol Manag 432:776–785
4. Lewis SL, Edwards DP, Galbraith D (2015) Increasing human dominance of tropical forests. Science 349:827–832
5. Malhi Y et al (2006) The regional variation of aboveground live biomass in old-growth Amazonian forests. Glob Change Biol 12:1107–1138
6. Feldpausch TR et al (2012) Tree height integrated into pantropical forest biomass estimates. Biogeosciences 9:3381–3403
7. Spawn SA, Sullivan CC, Lark TJ, Gibbs HK (2020) Harmonized global maps of above and belowground biomass carbon density in the year 2010. Sci Data 7:112
8. Li X, Xiao JA, Global (2019) 0.05-Degree Product of Solar-Induced Chlorophyll Fluorescence Derived from OCO–2, MODIS, and Reanalysis Data. Remote Sens 11:517
9. de Lima RAF et al (2020) The erosion of biodiversity and biomass in the Atlantic Forest biodiversity hotspot. Nat Commun 11:6347
10. Flores BM et al (2024) Critical transitions in the Amazon forest system. Nature 626:555–564
11. Brienen RJW et al (2015) Long-term decline of the Amazon carbon sink. Nature 519:344–348

12. Hubau W et al (2020) Asynchronous carbon sink saturation in African and Amazonian tropical forests. Nature 579:80–87
13. Tavares JV et al (2023) Basin-wide variation in tree hydraulic safety margins predicts the carbon balance of Amazon forests. Nature 617:111–117
14. Bennett AC et al (2023) Sensitivity of South American tropical forests to an extreme climate anomaly. Nat Clim Change 13:967–974
15. Jiménez-Muñoz JC et al (2016) Record-breaking warming and extreme drought in the Amazon rainforest during the course of El Niño 2015–2016. Sci Rep 6:33130
16. Wagner FH et al (2024) Amazon's 2023 Drought: Sentinel–1 Reveals Extreme Rio Negro River Contraction. Preprint at http://arxiv.org/abs/2401.16393
17. ForestPlots.net et al (2021) Taking the pulse of Earth's tropical forests using networks of highly distributed plots. Biol Conserv 260:108849
18. Heinrich VHA et al (2021) Large carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. Nat Commun 12:1785
19. Frankenberg C (2024) *Data Drought in the Humid Tropics: How to Overcome the Cloud Barrier in Greenhouse Gas Remote Sensing*. /users/528524/articles/717473-data-drought-in-the-humid-tropics-how-to-overcome-the-cloud-barrier-in-greenhouse-gas-remote-sensing?commit=c1c370fe625963945ecaea0faa78b3b911e10e05 doi:10.22541/essoar.170923255.57545328/v1
20. Rosan TM et al (2024) Synthesis of the land carbon fluxes of the Amazon region between 2010 and 2020. Commun Earth Environ 5:1–15
21. Dokoohaki H et al (2022) Development of an open-source regional data assimilation system in PEcAn v. 1.7.2: application to carbon cycle reanalysis across the contiguous US using SIPNET. Geosci Model Dev 15:3233–3252
22. Li L et al (2023) The optimization of model ensemble composition and size can enhance the robustness of crop yield projections. Commun Earth Environ 4:1–11
23. Dietze MC et al (2018) Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proc. Natl. Acad. Sci.* 115, 1424–1432
24. Greener JG, Kandathil SM, Moffat L, Jones DT (2022) A guide to machine learning for biologists. Nat Rev Mol Cell Biol 23:40–55
25. Potapov P et al The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. Sci Adv 3, e1600821
26. Boulton CA, Lenton TM, Boers N (2022) Pronounced loss of Amazon rainforest resilience since the early 2000s. Nat Clim Change 12:271–278
27. Friedlingstein P et al (2023) Global Carbon Budget 2023 Earth Syst Sci Data 15:5301–5369
28. Williamson GB et al (2000) Amazonian Tree Mortality during the 1997 El Niño Drought. Conserv Biol 14:1538–1542

29. et al.Forest fluxes and mortality response to drought: model description (ORCHIDEE-CAN-NHA,r7236) and evaluation at the Caxiuanã drought experiment. *Geosci. Model Dev. Discuss.*1–38 (2021) doi:10.5194/gmd–2021–362
30. Rowland L et al (2015) Death from drought in tropical forests is triggered by hydraulics not carbon starvation. Nature 528:119–122
31. McDowell NG (2011) Mechanisms Linking Drought, Hydraulics, Carbon Metabolism, and Vegetation Mortality. Plant Physiol 155:1051–1059
32. Tian H et al (1998) Effect of interannual climate variability on carbon storage in Amazonian ecosystems. Nature 396:664–667
33. Feldpausch TR et al (2016) Amazon forest response to repeated droughts. Glob Biogeochem Cycles 30:964–982
34. Chen C, Riley WJ, Prentice IC, Keenan TF (2022) CO2 fertilization of terrestrial photosynthesis inferred from site to global scales. *Proc. Natl. Acad. Sci.* 119, e2115627119
35. Wang S et al (2020) Recent global decline of CO2 fertilization effects on vegetation photosynthesis. Science 370:1295–1300
36. Yang H et al (2022) Climatic and biotic factors influencing regional declines and recovery of tropical forest biomass from the 2015/16 El Niño. *Proc. Natl. Acad. Sci.* 119, e2101388119
37. Cox PM et al (2004) Amazonian forest dieback under climate-carbon cycle projections for the 21st century. Theor Appl Climatol 78:137–156
38. Parry I, Ritchie P, Cox P (2022) *Evidence of Amazon Rainforest Dieback in CMIP6 Models*. 10.5194/egusphere–2022–82
39. Lovejoy TE, Nobre C (2018) Amazon Tipping Point. Sci Adv 4:eaat2340
40. Koppa A, Rains D, Hulsman P, Poyatos R, Miralles D (2022) G. A deep learning-based hybrid model of global terrestrial evaporation. Nat Commun 13:1912
41. Frieler K et al (2024) Scenario setup and forcing data for impact model evaluation and impact attribution within the third round of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP3a). Geosci Model Dev 17:1–51
42. Boulton CA, Good P, Lenton TM (2013) Early warning signals of simulated Amazon rainforest dieback. Theor Ecol 6:373–384
43. Fisher RA, Koven CD (2018) Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *J. Adv. Model. Earth Syst.* 12, eMS001453 (2020)
44. Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci Data 7:109
45. Harris I, Jones Pd, Osborn Tj, Lister D (2014) h. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. Int J Climatol 34:623–642
46. Kobayashi S et al (2015) The JRA–55 Reanalysis: General Specifications and Basic Characteristics. J Meteorol Soc Jpn Ser II 93:5–48

47. Bi W et al (2022) A global 0.05° dataset for gross primary production of sunlit and shaded vegetation canopies from 1992 to 2020. Sci Data 9:213
48. Wild B et al (2022) VODCA2GPP – a new, global, long-term (1988–2020) gross primary production dataset from microwave remote sensing. Earth Syst Sci Data 14:1063–1085
49. Wang S, Zhang Y, Ju W, Qiu B, Zhang Z (2021) Tracking the seasonal and inter-annual variations of global gross primary production during last four decades using satellite near-infrared reflectance data. Sci Total Environ 755:142569
50. Pedelty J et al (2007) Generating a long-term land data record from the AVHRR and MODIS Instruments. in. *IEEE International Geoscience and Remote Sensing Symposium* 1021–1025 (IEEE, Barcelona, Spain, 2007). 10.1109/IGARSS.2007.4422974
51. Chen T, Guestrin C, XGBoost: (2016) A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794ACM, San Francisco California USA, 10.1145/2939672.2939785
52. Chen T et al (2024) xgboost: Extreme Gradient Boosting
53. A Short Introduction to the caret Package https://cran.r-project.org/web/packages/caret/vignettes/caret.html
54. PEcAn functions used for managing climate driver data. https://pecanproject.github.io/modules/data.atmosphere/docs/index.html
55. Hijmans RJ et al (2023) raster: Geographic Data Analysis and Modeling
56. Eyring V et al (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958
57. Riahi K et al (2017) The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. Glob Environ Change 42:153–168
58. Vinutha HP, Poornima B, Sagar BM (2018) Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset. in *Information and Decision Sciences* (eds. Satapathy, S. C., Tavares, J. M. R. S., Bhateja, V. & Mohanty, J. R.) 511–518Springer, Singapore, 10.1007/978-981-10–7563–6_53

Figures

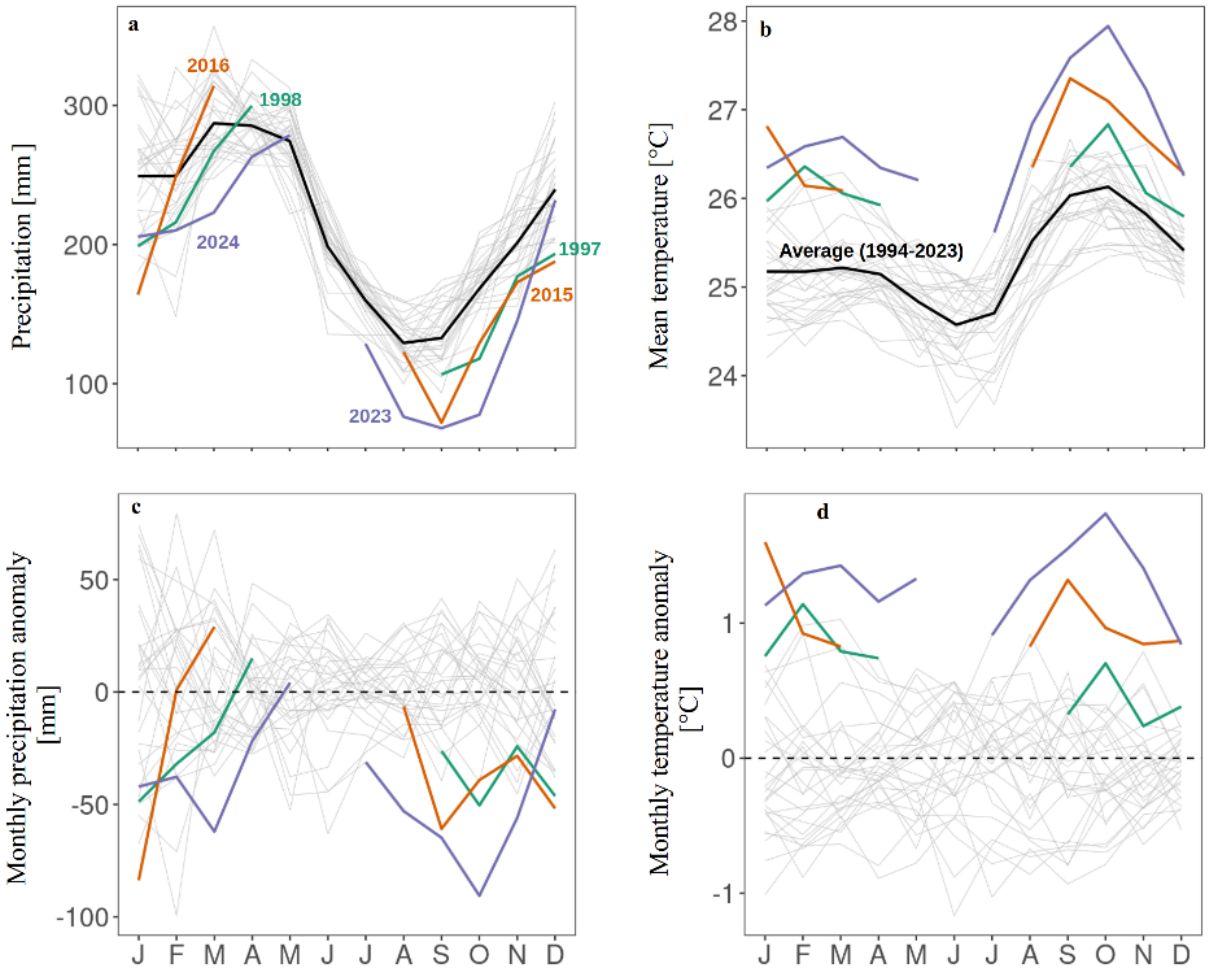

Figure 1

Seasonal cycle of the precipitation and the average air temperature (subplots a-b) and their anomalies (subplots c-d) over the intact Amazon forests according to ERA5 reanalyses. The individual years of the 1994-2023 time period are represented by the thin grey lines (subplots a-d) and their averages, by the thick black line (subplots a-b). Major dry and hot periods (1997-1998, 2015-2016, 2023-2024) are highlighted with the coloured lines and are discontinuous because they overlap two calendar years.

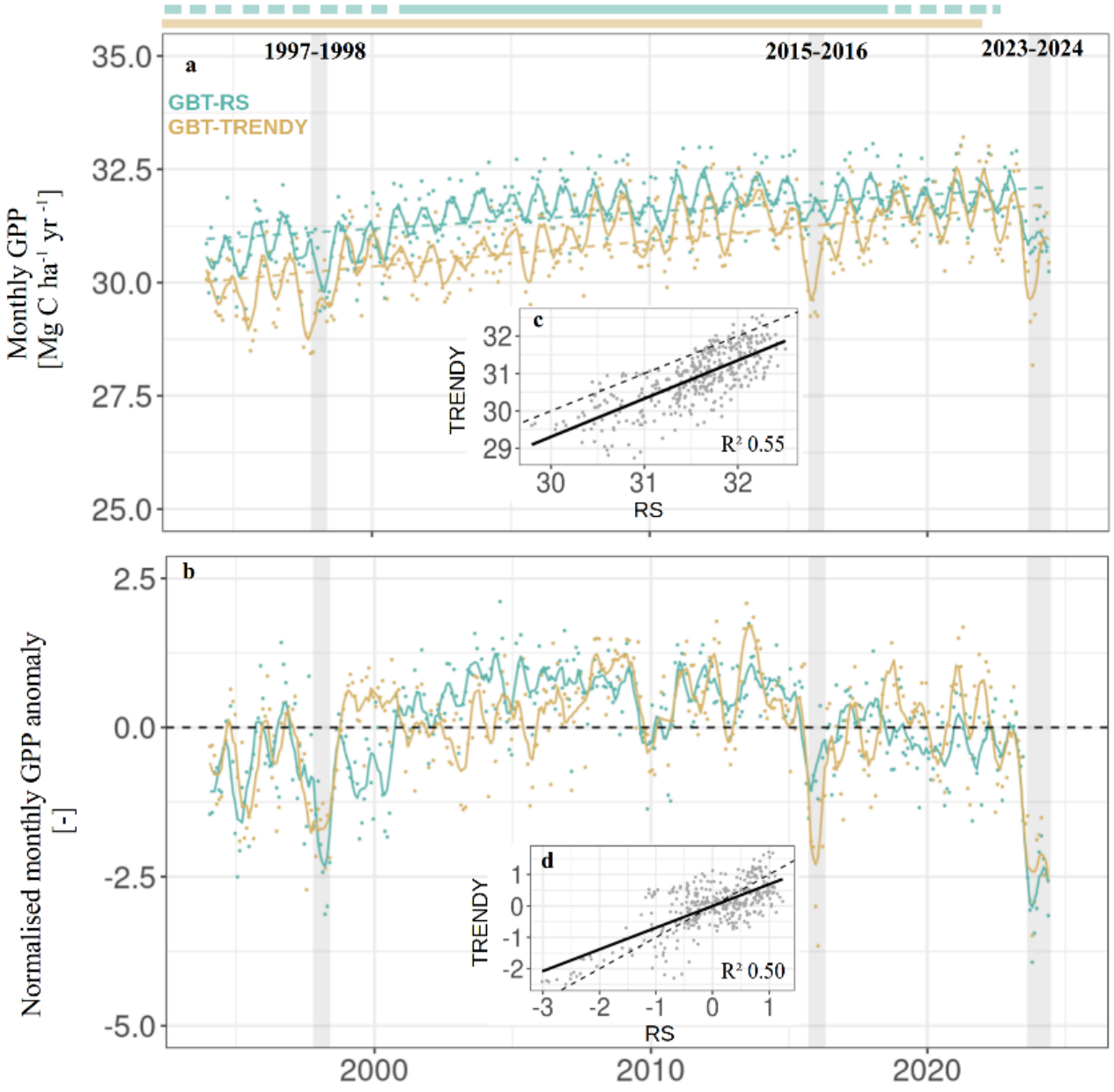

Figure 2

Time series of the monthly GPP (subplot a) and its normalised anomalies (subplot b) over the intact Amazon forests for the 1994-2024 time period, as generated by GBT models trained on TRENDY simulations (yellow, TRENDY) or remote sensing products (green, RS) with ERA5 re-analyses as model features. The anomalies were normalised by the standard deviations of the respective months and sources. Each data point represents the average of multiple vegetation models or multiple remote sensing products. The coloured dots are the monthly values while the curves are rolling averages over a 6-month time window. In subplot a, the dashed lines represent the respective linear increases used to detrend GPP. The insets (subplots c and d) show the correlations between the TRENDY model ensemble

mean and the average of the remote sensing products. In subplot a, the R^2 of the linear models are 0.20 and 0.28 for RS and TRENDY models, respectively (p-value < 0.001). The coloured rectangles above the figure represent the time coverage of the training data (until the end of 2021 for all TRENDY models, varying between 1992-2001 and 2018-2022 for the different RS products, see methods).

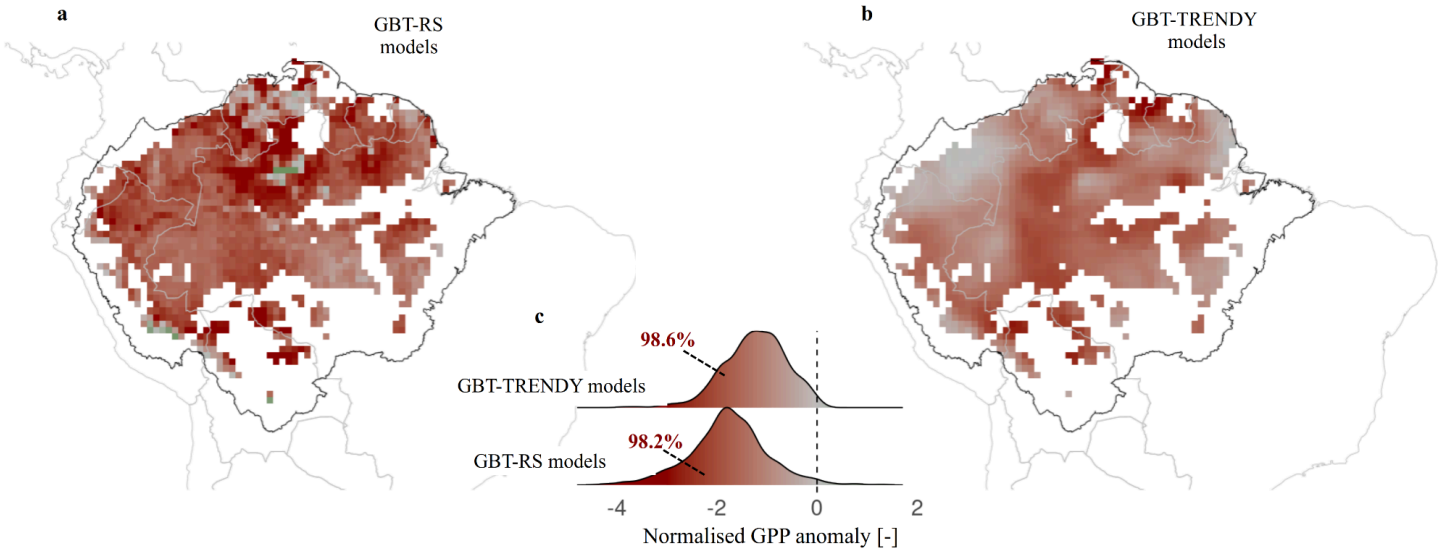

Figure 3

Spatial distribution of the average normalised GPP anomaly for the 2023-2024 dry and hot period as predicted from GBT models trained on remote sensing products (subplot a), or TRENDY models (subplot b). In subplots a and b, the black line delineates our study area and only intact forests were included in the analysis (white grid cells within the Amazon basin are mostly not intact forests). Subplots c shows the marginal density distribution of the anomalies in both cases.

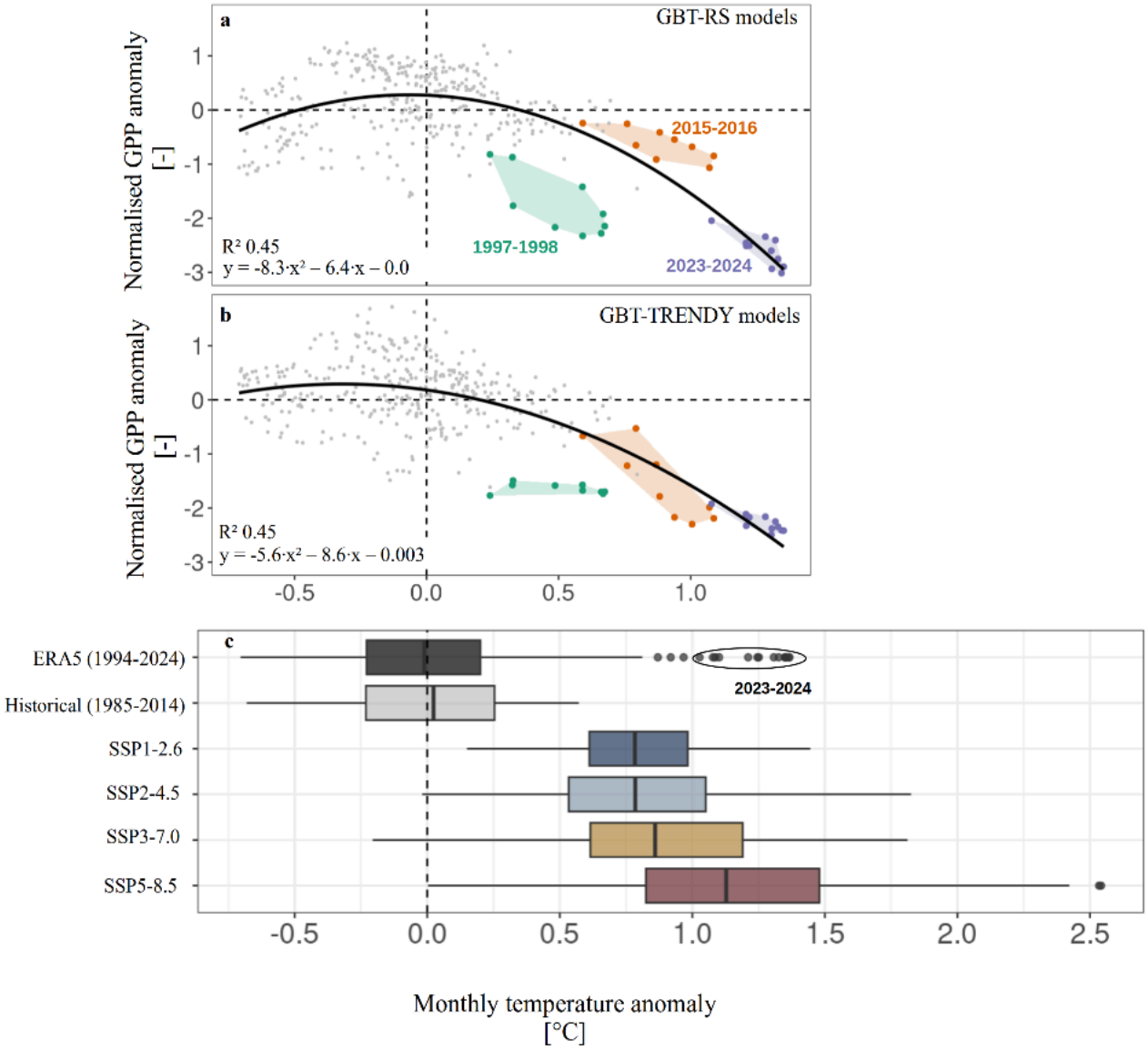

Figure 4

Relationships between normalised monthly GPP anomalies and monthly mean air temperature anomalies for the intact Amazon forests, based on GBT models trained on remote sensing (subplot a) or TRENDY models estimates of GPP (subplot b). Each data point represents the monthly average of multiple vegetation models (subplot a) or multiple remote sensing products (subplot b). In subplots a-b, temperature and GPP anomalies were smoothed using a moving average over a 6-month time period and the coloured dots and zones correspond to monthly values of recent major dry and hot episodes. The monthly temperature anomalies predicted from CMIP6 simulations are shown in subplot c for the historical period (1985-2014) or the next 30 years (2024-2053) according to the different shared

socioeconomic pathways. In subplot c, the temperature anomalies of the 2023-2024 dry and hot episode are highlighted for the ERA5 reanalysis. P-values of the second-order polynomial models are < 0.001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplement.docx