

20th US-China Carbon Consortium

Xining China

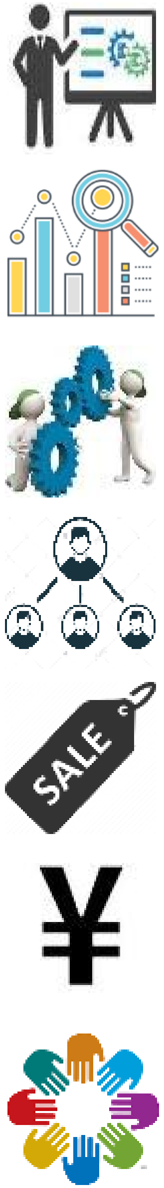
# Estimation of terrestrial gross primary production: current progresses and challenges

Prof. Dr. Tiexi CHEN / Dr. Xin CHEN

Qinghai University of Technology

Nanjing University of Information Science and Technology

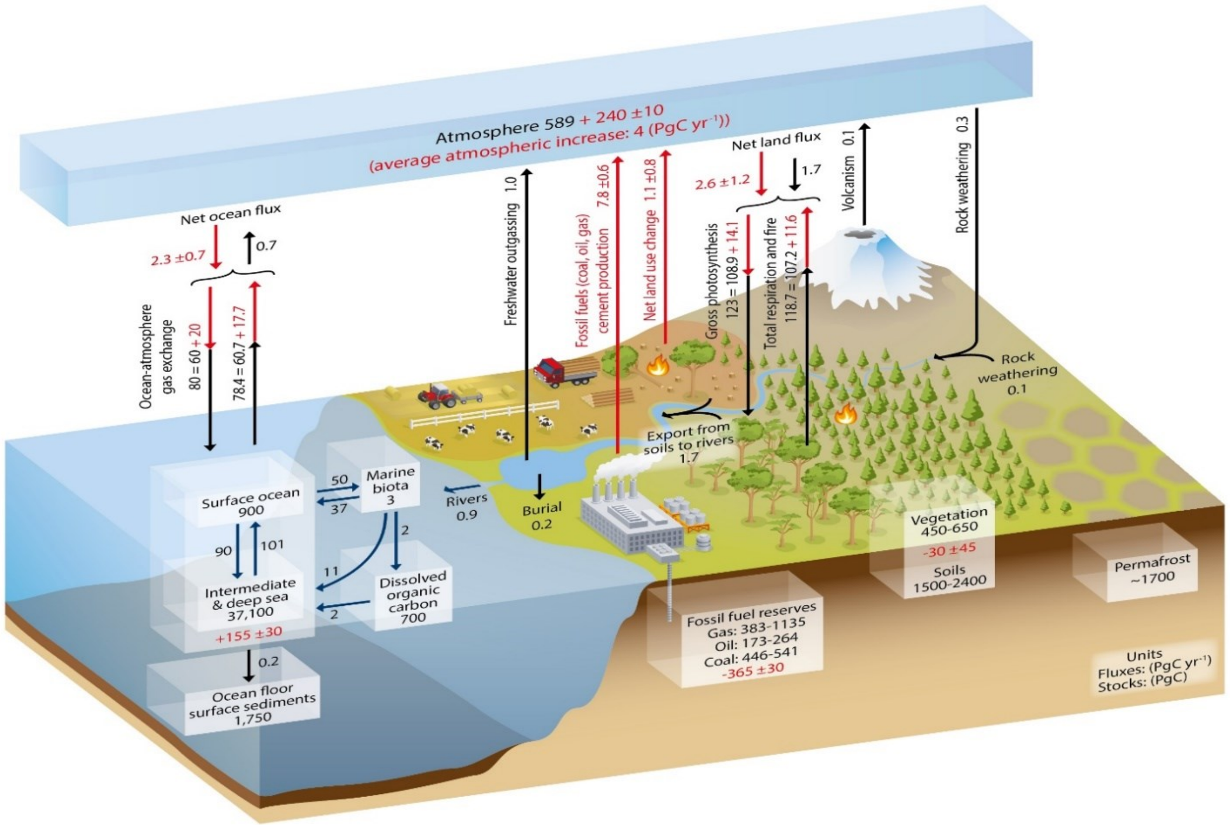
2024.07.17

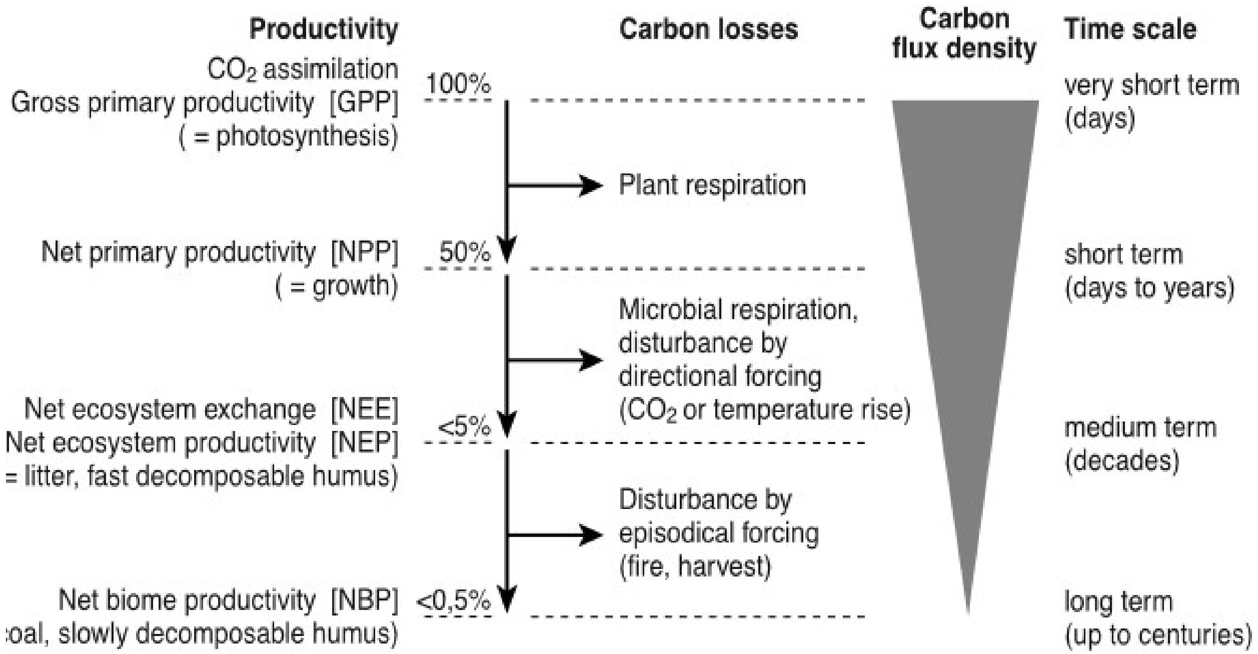



1. Background
2. GPP estimation model and its improvement
3. Uncertainty in GPP estimates
4. Sources of uncertainty in GPP estimates
5. Conclusion and prospect

# 1. Background



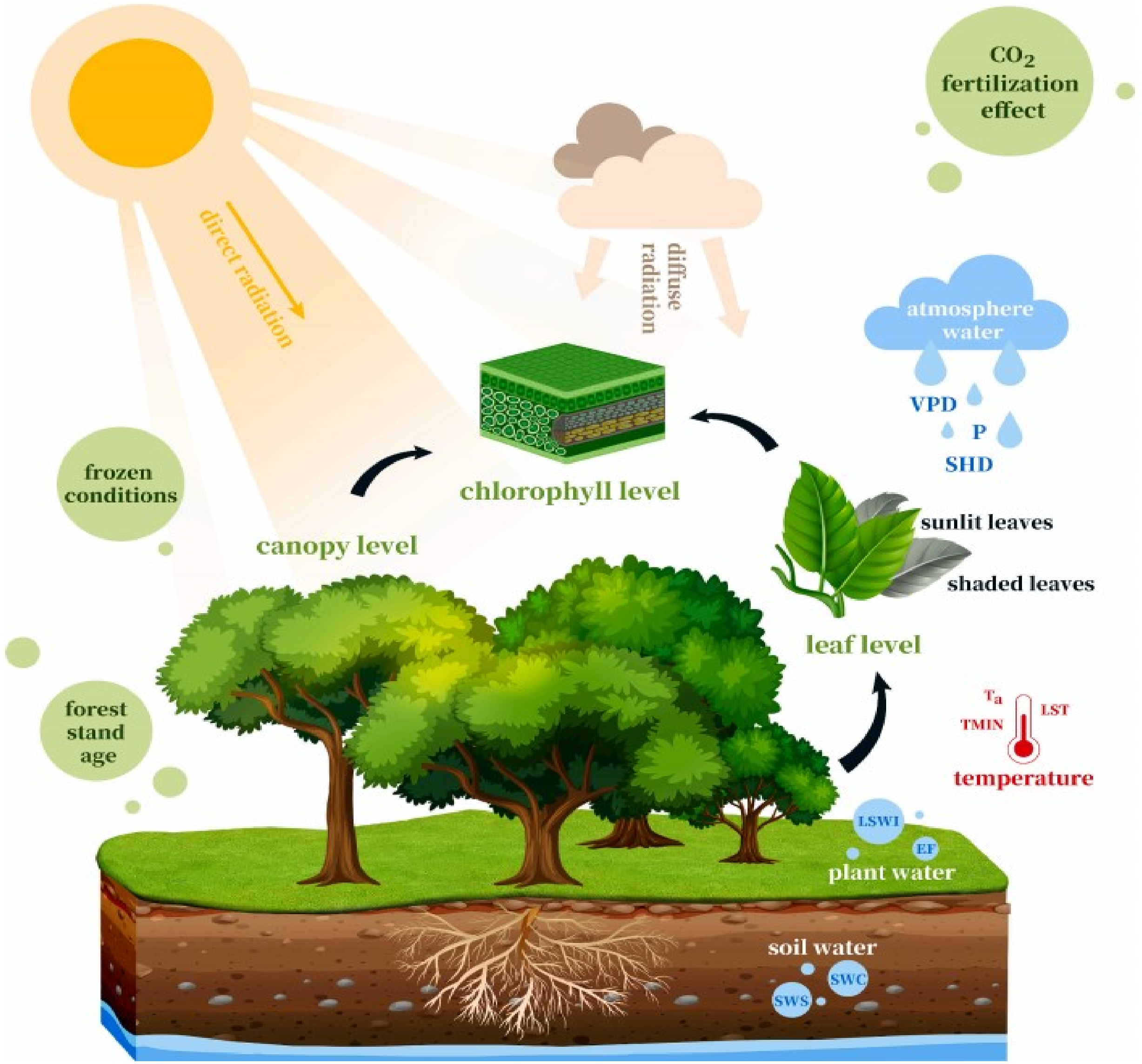


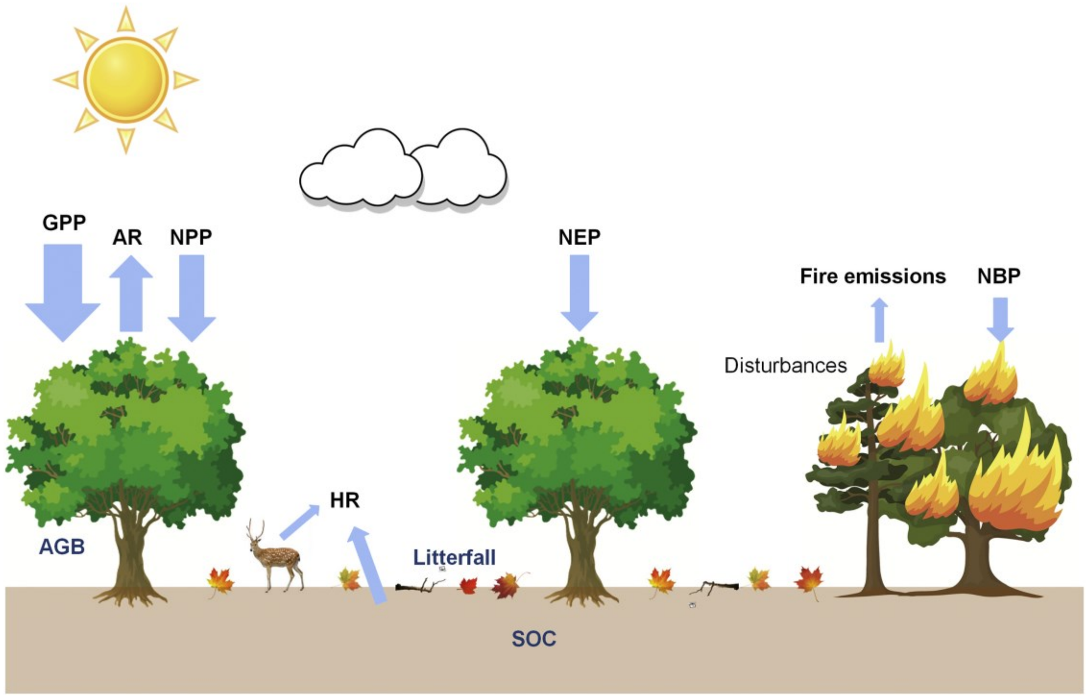

**Gross Primary Productivity (GPP) refers to the $CO_2$ absorbed by vegetation through photosynthesis, and GPP is the largest carbon flux between terrestrial ecosystems and the atmosphere, playing a key role in the global carbon cycle.**

# 1. Background





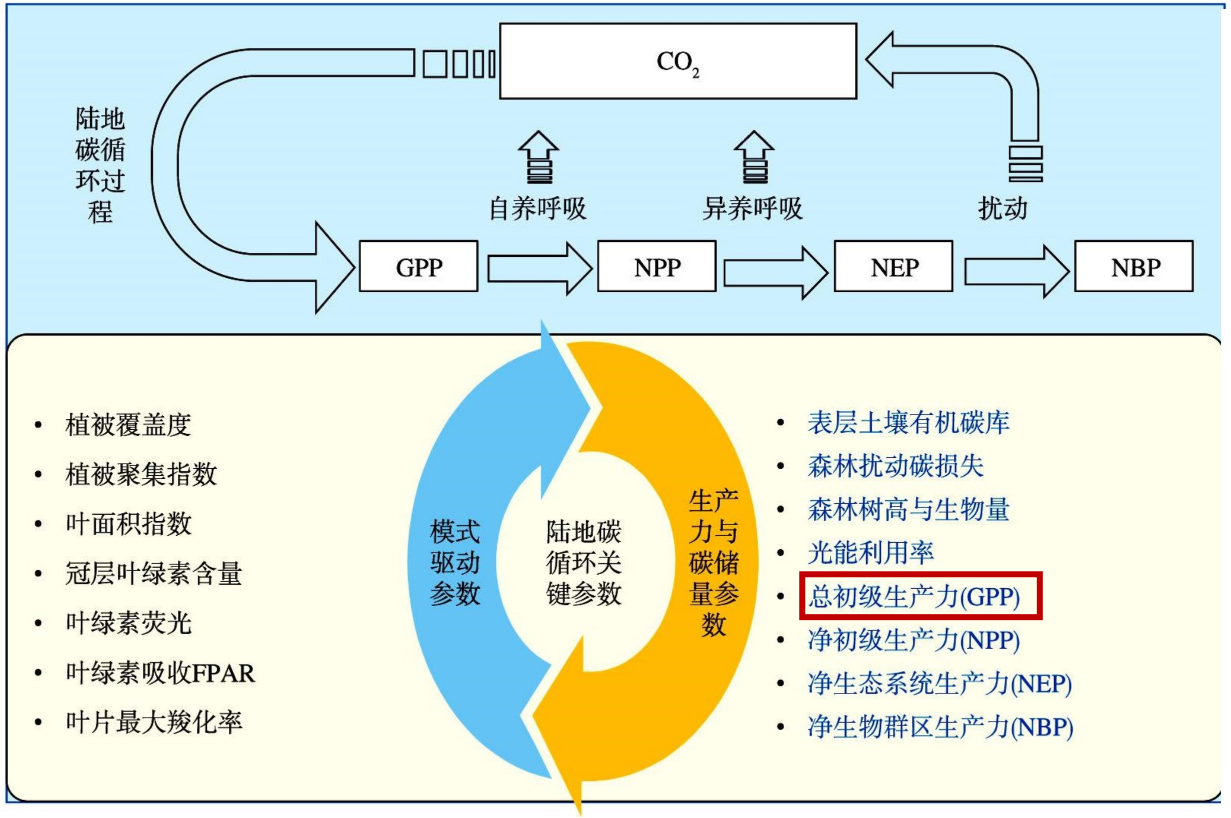

Factors affecting photosynthesis（Pei 2022）



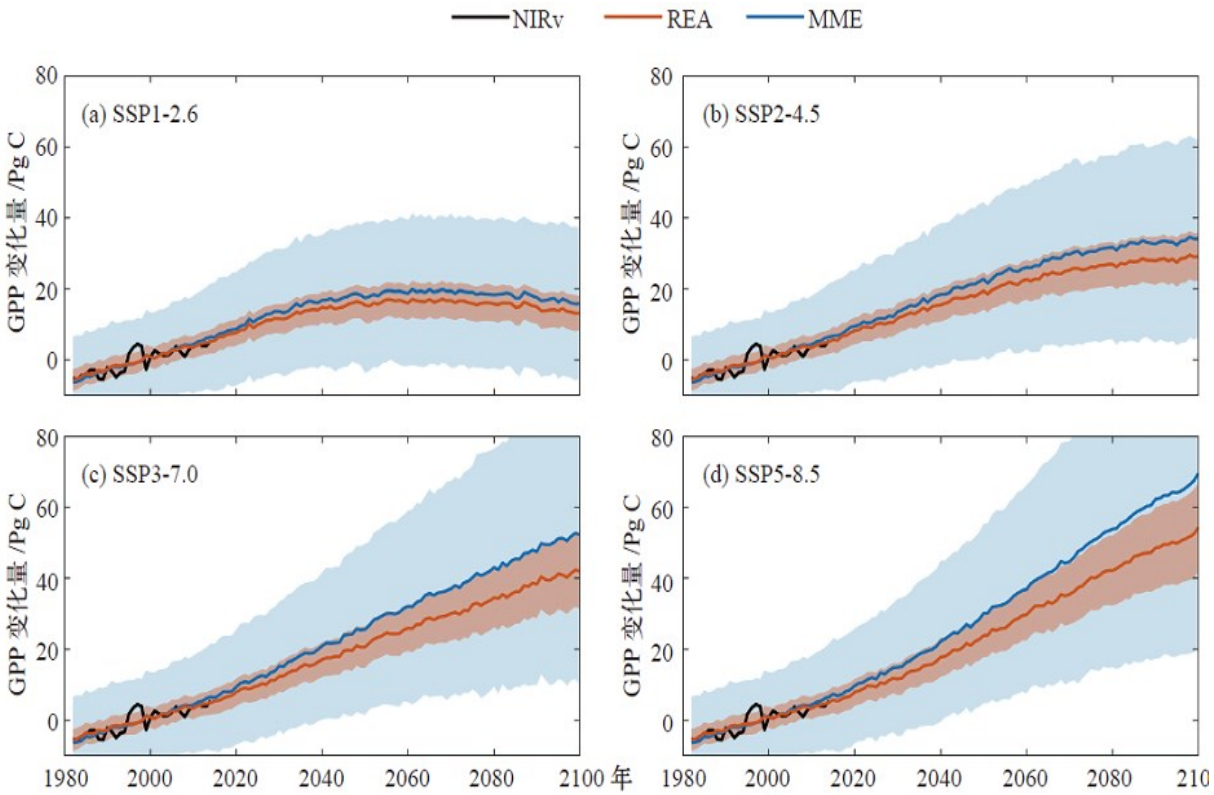

Carbon flux conversion process（Xiao 2019）

# 1. Background

NUIST




Key parameters of carbon cycle in terrestrial ecosystems and their contribution to carbon source and sink estimation（Liu 2022）

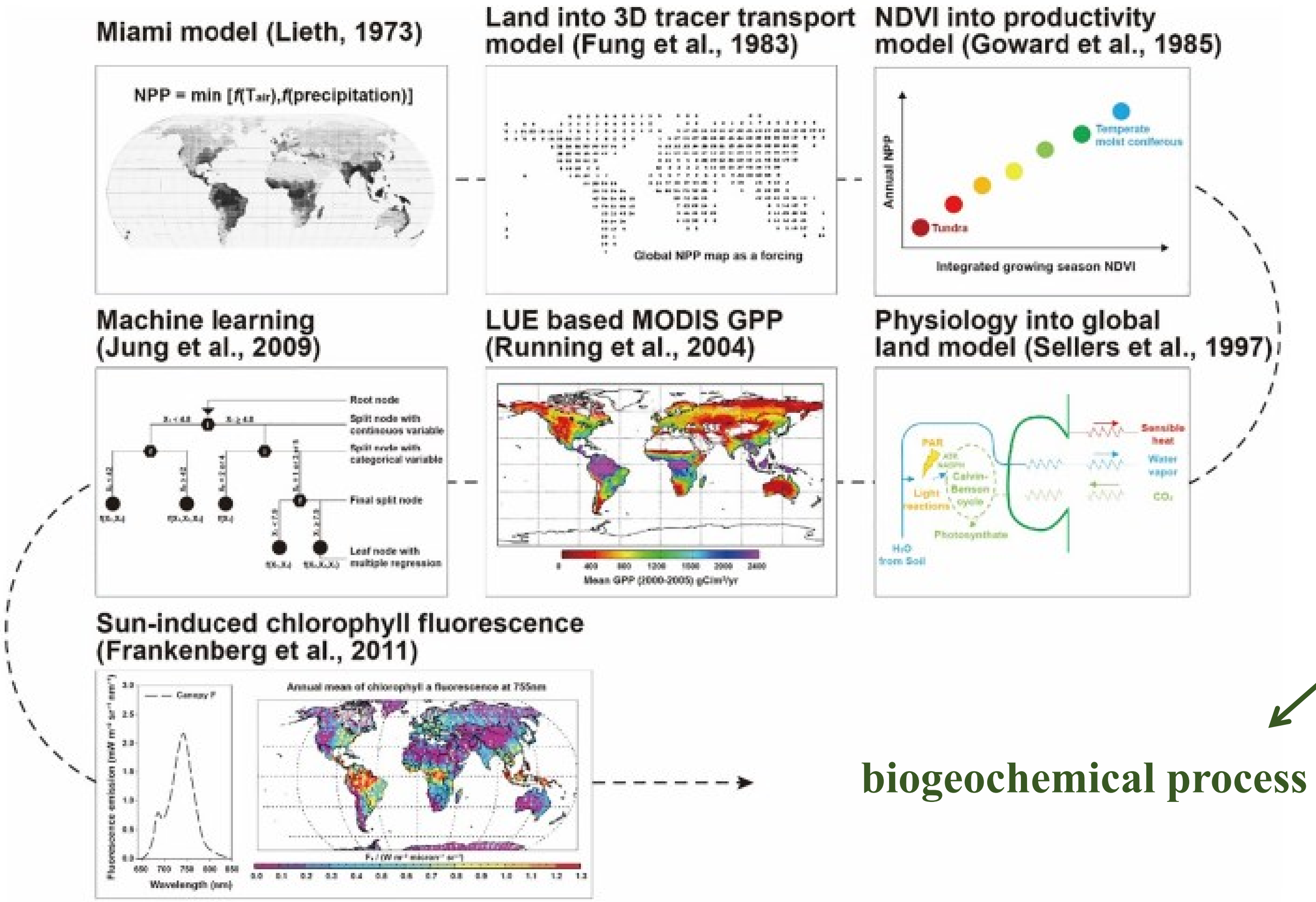



Global GPP changes in REA and MME method index under different SSP scenarios from 1982 to 2100 (Huang 2021)

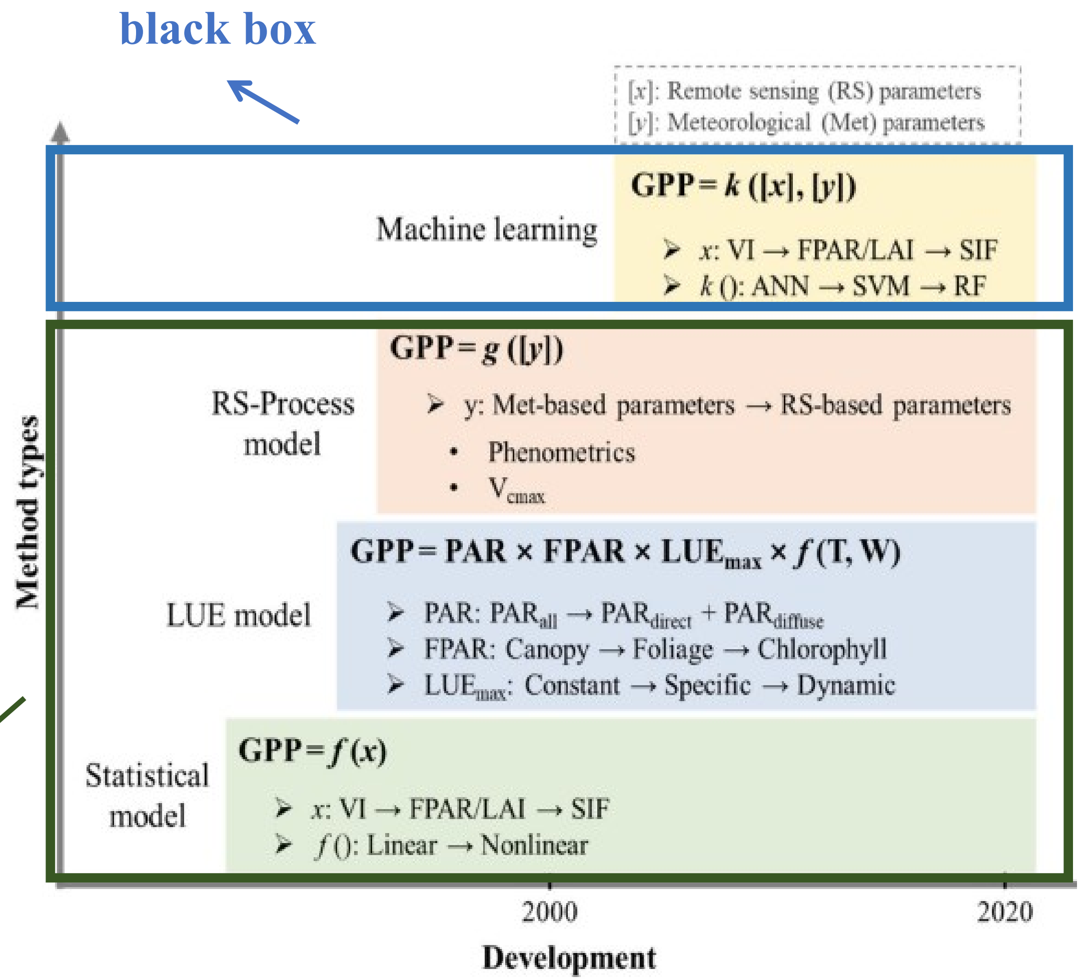
**The importance of accurately estimating GPP:**

**(1) Estimates of other carbon fluxes, such as terrestrial carbon sinks (2) estimates of future GPP**

# 1. Background

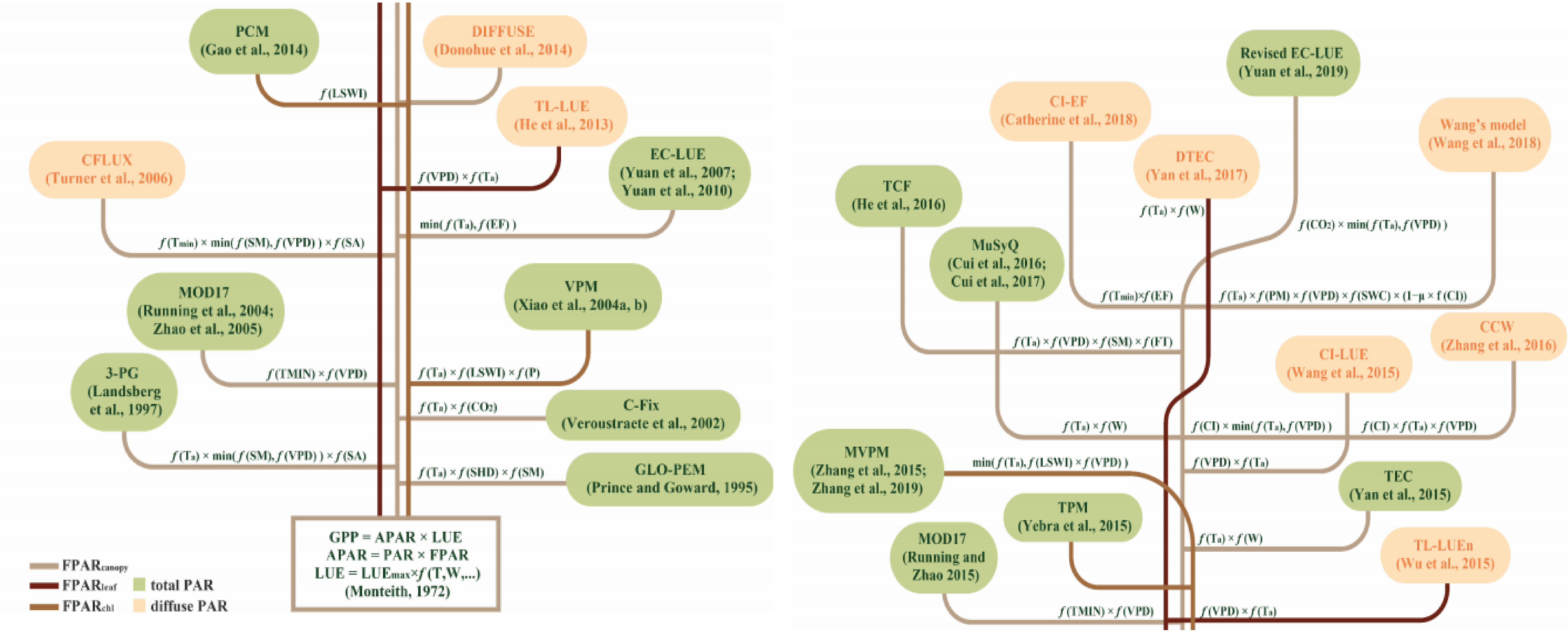



Representative results of estimates of total terrestrial primary productivity（Ryu 2019）



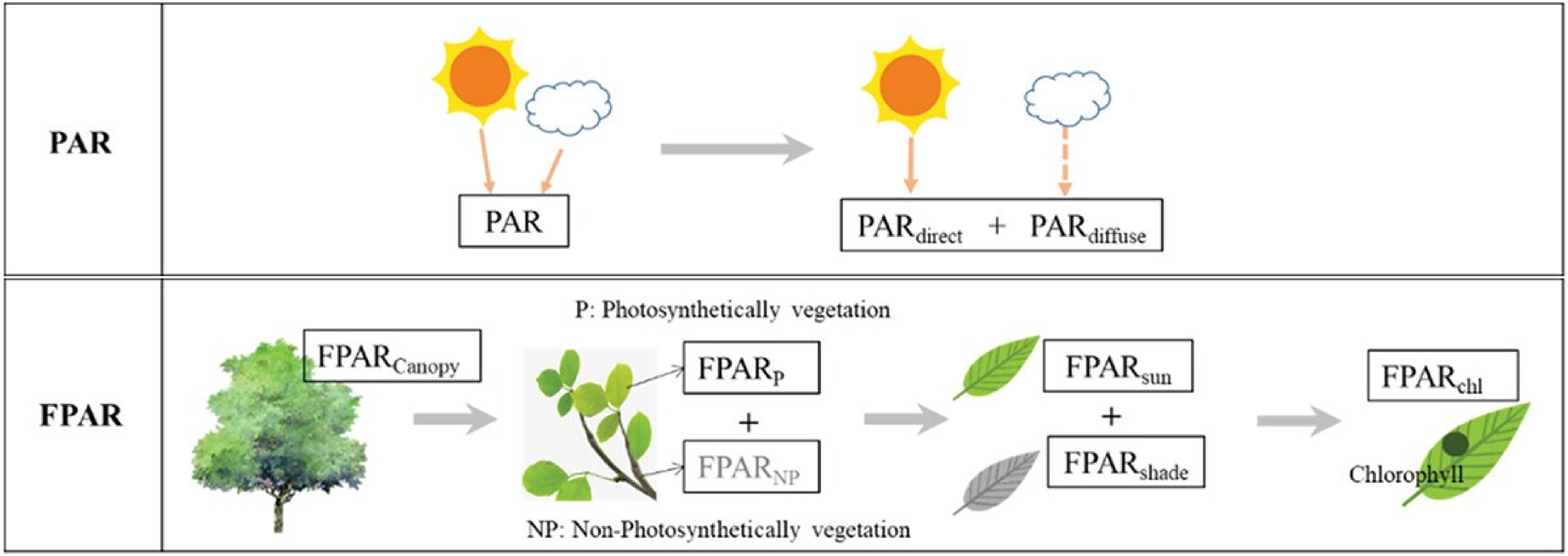

Development of GPP estimation methods（Zhu 2024）

**After years of development, GPP estimation methods have been relatively mature compared with other carbon fluxes**




The development of LUE model（Pei 2022）

**In recent decades, numerous LUE models have been developed to estimate GPP**




**An improved method of LUE model（Zhu 2024）**

**Big Leaf Medel**

$$GPP = \varepsilon \times FPAR \times PAR \times f(W) \times f(T)$$

**Two-leaf model**

$$GPP = (\varepsilon_{sun} \times APAR_{sun} + \varepsilon_{shd} \times APAR_{shd}) \times f(W) \times f(T)$$

Remote Sensing of Environment 295 (2023) 113696

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

ELSEVIER

DA

## BESSv2.0: A satellite-based and coupled-process model for quantifying long-term global land–atmosphere fluxes

Bolun Li [a], Youngryel Ryu [a,b,*], Chongya Jiang [c], Benjamin Dechant [a,d,e], Jiangong Liu [a], Yulin Yan [f], Xing Li [a]

[a] Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, South Korea
[b] Department of Landscape Architecture and Rural Systems Engineering, Seoul National University, Seoul, South Korea
[c] College of Agricultural, Consumer and Environmental Sciences, University of Illinois at Urbana Champaign, Urbana, IL, USA
[d] German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Leipzig, Germany
[e] Leipzig University, Leipzig, Germany
[f] Interdisciplinary Program in Agricultural and Forest Meteorology, Seoul National University, South Korea

ARTICLE INFO

Edited by Jing M. Chen

ABSTRACT

Recent remote-sensing-based global carbon, water and energy budgets over land still include considerable uncertainties. Most existing flux products of terrestrial carbon, water and energy components were developed individually, despite the inherently coupled processes among them. In this study, we present a new set of global daily surface downwelling shortwave radiation (SW), net radiation ($R_{net}$), evapotranspiration (ET), gross primary productivity (GPP), terrestrial ecosystem respiration (TER) and net ecosystem exchange (NEE) datasets at 0.05° resolutions from 1982 to 2019, by improving a satellite-based and coupled-process model—the Breathing Earth System Simulator (BESS). The new version of BESS (v2.0) integrated a newly developed ecosystem respiration module, an optimality-based maximum carboxylation rate ($V_{cmax}$) model, and extended the temporal coverage of flux datasets from 1982 to 2019. We evaluated BESS products against the FLUXNET2015 dataset at the site scale

at t

JGR Biogeosciences

Research Article | Free Access

## Drought Risk of Global Terrestrial Gross Primary Productivity Over the Last 40 Years Detected by a Remote Sensing-Driven Process Model

Qiaoning He, Weimin Ju, Shengpei Dai, Wei He, Lian Song, Songhan Wang, Xinchuan Li, Guangxiong Mao

First published: 02 June 2021 | https://doi.org/10.1029/2020JG005944 | Citations: 25

SECTIONS PDF TOOLS SHARE

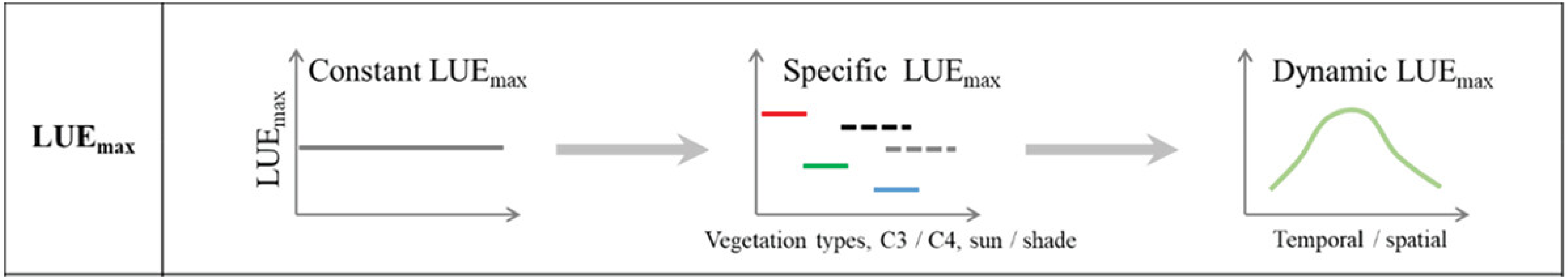
### Abstract

Gross primary productivity (GPP) is the largest flux in the global terrestrial carbon cycle. Drought has significantly impacted global terrestrial GPP in recent decades, and has been projected to occur with increasing frequency and intensity. However, the drought risk of global terrestrial GPP has not been well investigated. In this study, global terrestrial GPP during 1981–2016 was simulated with the process-based Boreal Ecosystem Productivity Simulator model. Then, the drought risk of GPP was quantified as the product of drought probability and reduction of GPP caused by drought, which was determined using the standardized precipitation evapotranspiration index. During the study period, the drought risk of GPP was high in the southeastern United States, most of South America, southern Europe, central and eastern Africa, eastern and southeastern Asia, and eastern Australia. It was low at some high latitudes of the Northern Hemisphere and in part of tropical South America, where terrestrial GPP increased slightly in drought years. The drought risk of terrestrial GPP was greater during 2000–2016 than during 1981–1999 in 21 out of 24 climatic zones. The global mean drought risk of GPP increased from

**Many current GPP products use the structure of the two-leaf model**

# 2. GPP estimation model and its improvement

## The model parameters need to be further optimized




**An improved method of LUE model（Zhu 2024）**

***Table 2.2.*** *Biome-Property-Look-Up-Table (BPLUT) for MODIS GPP/NPP algorithm with NCEP-DOE reanalysis II and the Collection5 FPAR/LAI as inputs. The full names for the University of Maryland land cover classification system (UMD_VEG_LC) in MCDLCHKM dataset (fieldname: Land_Cover_Type_1) are, Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (EBF), Deciduous Needleleaf Forest (DNF), Deciduous Broadleaf Forest (DBF), Mixed forests (MF), Closed Shrublands (CShrub), Open Shrublands (OShrub), Woody Savannas (WSavanna), Savannas (Savanna), Grassland (Grass), and Croplands (Crop).*

| UMD_VEG_LC | ENF | EBF | DNF | DBF | MF | CShrub | OShrub | WSavanna | Savanna | Grass | Crop |
|---|---|---|---|---|---|---|---|---|---|---|---|
| LUEmax (KgC/m$^2$/d/MJ) | 0.000962 | 0.001268 | 0.001086 | 0.001165 | 0.001051 | 0.001281 | 0.000841 | 0.001239 | 0.001206 | 0.000860 | 0.001044 |
| Tmin_min (C) | -8.00 | -8.00 | -8.00 | -6.00 | -7.00 | -8.00 | -8.00 | -8.00 | -8.00 | -8.00 | -8.00 |
| Tmin_max (C) | 8.31 | 9.09 | 10.44 | 9.94 | 9.50 | 8.61 | 8.80 | 11.39 | 11.39 | 12.02 | 12.02 |
| VPD_min (Pa) | 650.0 | 800.0 | 650.0 | 650.0 | 650.0 | 650.0 | 650.0 | 650.0 | 650.0 | 650.0 | 650.0 |
| VPD_max (Pa) | 4600.0 | 3100.0 | 2300.0 | 1650.0 | 2400.0 | 4700.0 | 4800.0 | 3200.0 | 3100.0 | 5300.0 | 4300.0 |
| SLA (LAI/KgC) | 14.1 | 25.9 | 15.5 | 21.8 | 21.5 | 9.0 | 11.5 | 27.4 | 27.1 | 37.5 | 30.4 |
| $Q_{10}$* | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 |
| froot_leaf_ratio | 1.2 | 1.1 | 1.7 | 1.1 | 1.1 | 1.0 | 1.3 | 1.8 | 1.8 | 2.6 | 2.0 |
| livewood_leaf_ratio | 0.182 | 0.162 | 0.165 | 0.203 | 0.203 | 0.079 | 0.040 | 0.091 | 0.051 | 0.000 | 0.000 |
| leaf_mr_base | 0.00604 | 0.00604 | 0.00815 | 0.00778 | 0.00778 | 0.00869 | 0.00519 | 0.00869 | 0.00869 | 0.0098 | 0.0098 |
| froot_mr_base | 0.00519 | 0.00519 | 0.00519 | 0.00519 | 0.00519 | 0.00519 | 0.00519 | 0.00519 | 0.00519 | 0.00819 | 0.00819 |
| livewood_mr_base | 0.00397 | 0.00397 | 0.00397 | 0.00371 | 0.00371 | 0.00436 | 0.00218 | 0.00312 | 0.00100 | 0.00000 | 0.00000 |

*: The constant $Q_{10} = 2.0$ is applied to fine roots and live wood, while for leaves, a temperature acclimation $Q_{10}$ value is used as described in Equation.

**Table 3.** Optimized parameters ($\varepsilon_{msu}$, $\varepsilon_{msh}$, $\varphi$, and $VPD_0$) of the revised EC-LUE model for different vegetation types.

| Vegetation types | DBF | ENF | EBF | MF | GRA | CRO-$C_3$ | CRO-$C_4$ | SAV | SHR | WET |
|---|---|---|---|---|---|---|---|---|---|---|
| $\varepsilon_{msu}$ (g C MJ$^{-1}$) | 1.28 ± 0.36 | 1.72 ± 0.42 | 1.67 ± 0.85 | 1.38 ± 0.21 | 1.16 ± 0.15 | 1.25 ± 0.42 | 2.46 ± 0.78 | 2.24 ± 0.68 | 1.21 ± 0.25 | 1.34 ± 0.26 |
| $\varepsilon_{msh}$ (g C MJ$^{-1}$) | 3.59 ± 0.66 | 3.87 ± 0.58 | 4.35 ± 0.72 | 3.29 ± 0.63 | 1.91 ± 0.46 | 2.46 ± 0.52 | 5.64 ± 1.02 | 4.26 ± 0.95 | 2.71 ± 0.52 | 2.62 ± 0.49 |
| $\varphi$ (ppm) | 32 ± 8.25 | 25 ± 7.59 | 20 ± 6.36 | 49 ± 11.25 | 57 ± 11.85 | 43 ± 9.56 | 54 ± 15.36 | 54 ± 12.23 | 34 ± 7.59 | 36 ± 10.32 |
| $VPD_0$ (k Pa) | 1.15 ± 0.25 | 1.34 ± 0.26 | 0.57 ± 0.15 | 0.62 ± 0.14 | 1.69 ± 0.35 | 1.02 ± 0.19 | 1.53 ± 0.31 | 1.65 ± 0.26 | 1.34 ± 0.21 | 0.62 ± 0.12 |

| Vegetation Type | DBF | EBF | ENF | MF | CRO | GRA | OSH | SAV | WET | WSA |
|---|---|---|---|---|---|---|---|---|---|---|
| $\varepsilon_{msh}$(gC MJ$^{-1}$) | 3.75 ± 0.52 | 3.26 ± 0.93 | 3.40 ± 1.19 | 3.00 ± 0.66 | 4.80 ± 1.94 | 4.57 ± 1.67 | 3.10 ± 0.42 | 4.65 ± 0.64 | 2.53 ± 1.02 | 2.70 |
| $\varepsilon_{msu}$(gC MJ$^{-1}$) | 0.92 ± 0.29 | 1.44 ± 0.64 | 0.89 ± 0.49 | 0.80 ± 0.41 | 1.43 ± 0.75 | 1.16 ± 0.45 | 0.65 ± 0.07 | 3.45 ± 0.64 | 1.23 ± 0.92 | 2.60 |
| $VPD_{max}$(kPa) | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 |
| $VPD_{min}$(kPa) | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 |
| $T_{opt}$(℃) | 23.1 | 25.8 | 19.7 | 24.5 | 23.5 | 20.9 | 22.3 | 25.8 | 24.2 | 26.2 |
| albedo(α)[49,95,96] | 0.18 | 0.18 | 0.15 | 0.17 | 0.23 | 0.23 | 0.16 | 0.18 | 0.23 | 0.23 |
| Clumping index(Ω)[97] | 0.8 | 0.8 | 0.6 | 0.7 | 0.9 | 0.9 | 0.8 | 0.8 | 0.9 | 0.8 |

# 2. GPP estimation model and its improvement

NUIST

## The spatial distribution of parameters is used in the GPP estimation model

JGR Biogeosciences

RESEARCH ARTICLE

10.1029/2020JG005651

**Key Points:**

- Global spatial maps of vegetation optimum growth temperature and maximum light use efficiency
- A new global gross primary production data set with significantly improved accuracy
- Optimum growth temperature and maximum light use efficiency can indicate the interaction between plants and the environment

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

X. Feng,
fengxm@rcees.ac.cn

**Citation:**

Chen, Y., Feng, X., Fu, B., Wu, X., & Gao, Z. (2021). Improved global maps of the optimum growth temperature, maximum light use efficiency, and gross primary production for vegetation. *Journal of Geophysical Research: Biogeosciences*, *126*, e2020JG005651. https://doi.org/10.1029/2020JG005651

Received 18 JAN 2020
Accepted 22 FEB 2021

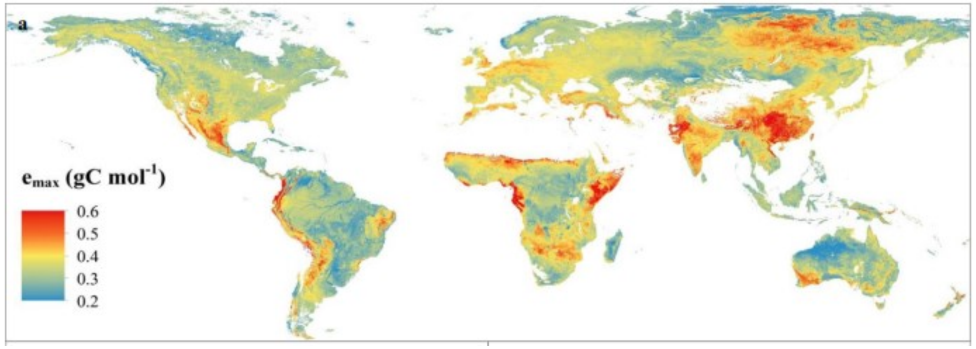
### Improved Global Maps of the Optimum Growth Temperature, Maximum Light Use Efficiency, and Gross Primary Production for Vegetation

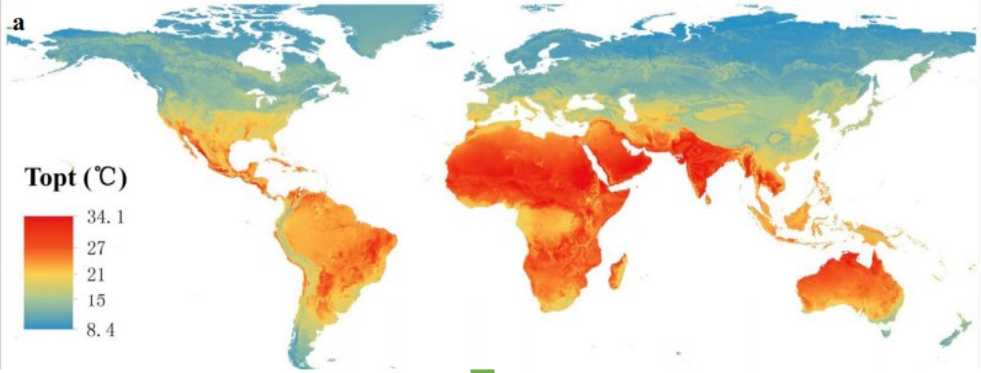
Yongzhe Chen[1,2], Xiaoming Feng[1], Bojie Fu[1,2], Xutong Wu[3], and Zhen Gao[1]

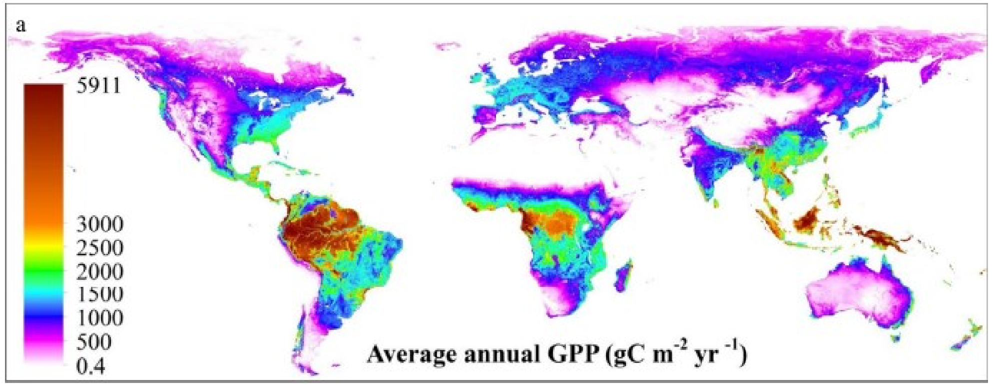
[1]State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]College of Urban and Environmental Sciences, Peking University, Beijing, China

**Abstract** The optimum growth temperature ($T_{opt}$) and maximum light use efficiency ($\varepsilon_{max}$) of terrestrial vegetation are closely related to plant photosynthesis in current and future Earth environments, yet little is known about their spatial distributions at the global scale. This study derived global maps of $T_{opt}$ and $\varepsilon_{max}$ separately, under the light use efficiency (LUE) model framework by utilizing FLUXNET measurements and satellite-observed solar/sun-induced chlorophyll fluorescence (SIF), as well as multiple regression and neural network regression based on environmental and biological factors. $T_{opt}$ is found to be positively correlated with annual mean temperature ($T$), except in cold areas with $T < 9°C$, where $T_{opt}$ stays within the range of 10°C–15°C. $T_{opt}$ is equal to $T$ in tropical areas with $T \geq 25°C$, but is obviously higher than $T$ in other regions. $\varepsilon_{max}$ is high in regions with a large amount of diffuse radiation and increases significantly with water stress. The maps of $T_{opt}$ and $\varepsilon_{max}$ improved the global gross primary production (GPP) estimation ($R^2 = 0.83$, RMSE = 1.38 g C $m^{-2}$ $d^{-1}$ against flux observations). The average annual GPP was 126 ± 1.5 PgC $yr^{-1}$, with a trend of 0.6 ± 0.1 PgC $yr^{-2}$ during 2001–2016, faster than most previous estimates. Our study suggests that the positive anthropogenic impacts on GPP were underestimated in existing products, including cropland expansion in southern Brazil and afforestation/forest protection efforts in China and western Europe. This study also provides a potential method for unified GPP modeling under the LUE framework.

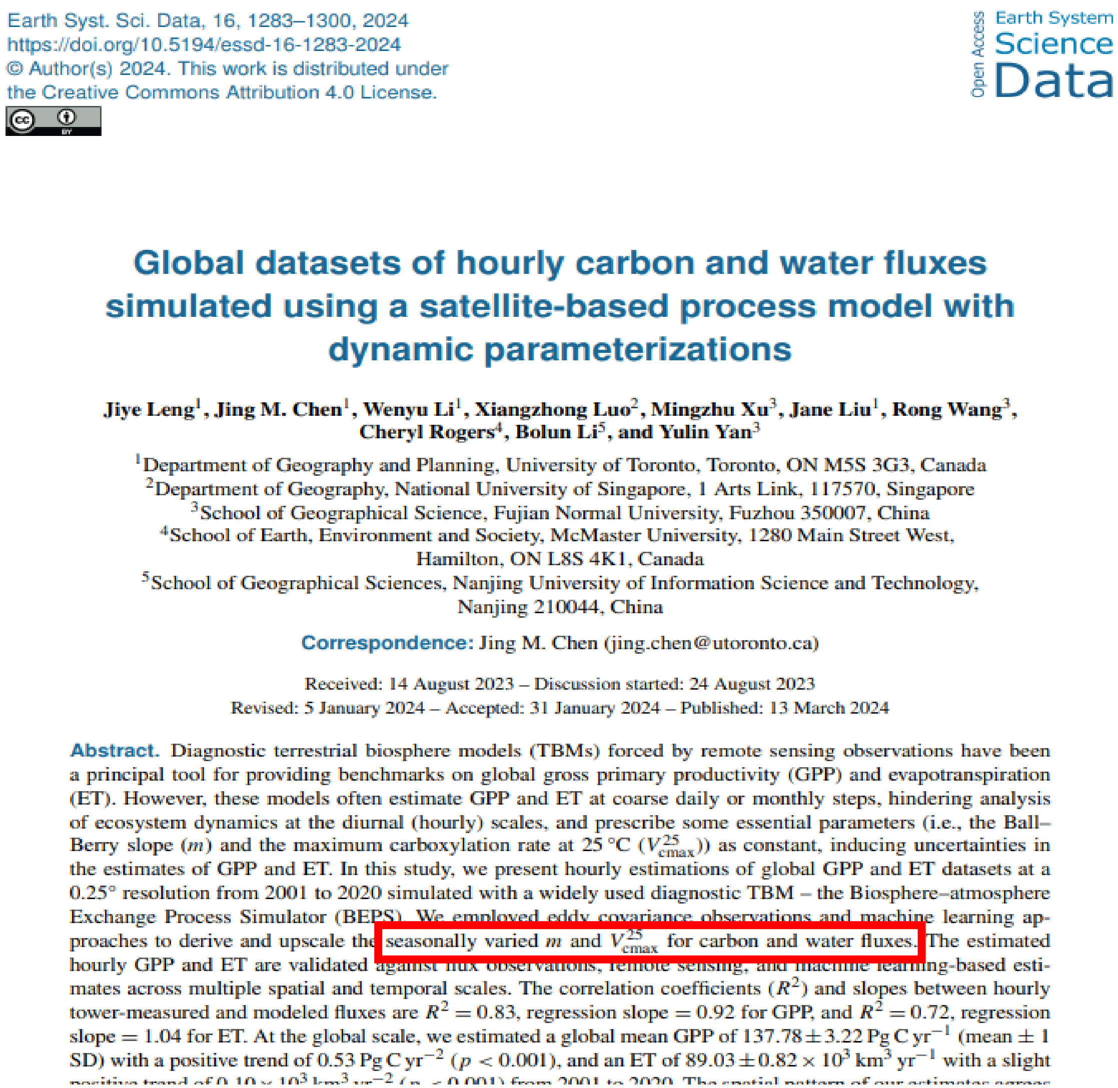
**Plain Language Summary** Light use efficiency models are commonly applied to simulate global terrestrial gross primary production (GPP)—carbon uptake by terrestrial vegetation through



# Dynamic parameters that vary in time and space are used in the model


Earth Syst. Sci. Data, 16, 1283–1300, 2024
https://doi.org/10.5194/essd-16-1283-2024

Open Access Earth System Science Data

## Global datasets of hourly carbon and water fluxes simulated using a satellite-based process model with dynamic parameterizations

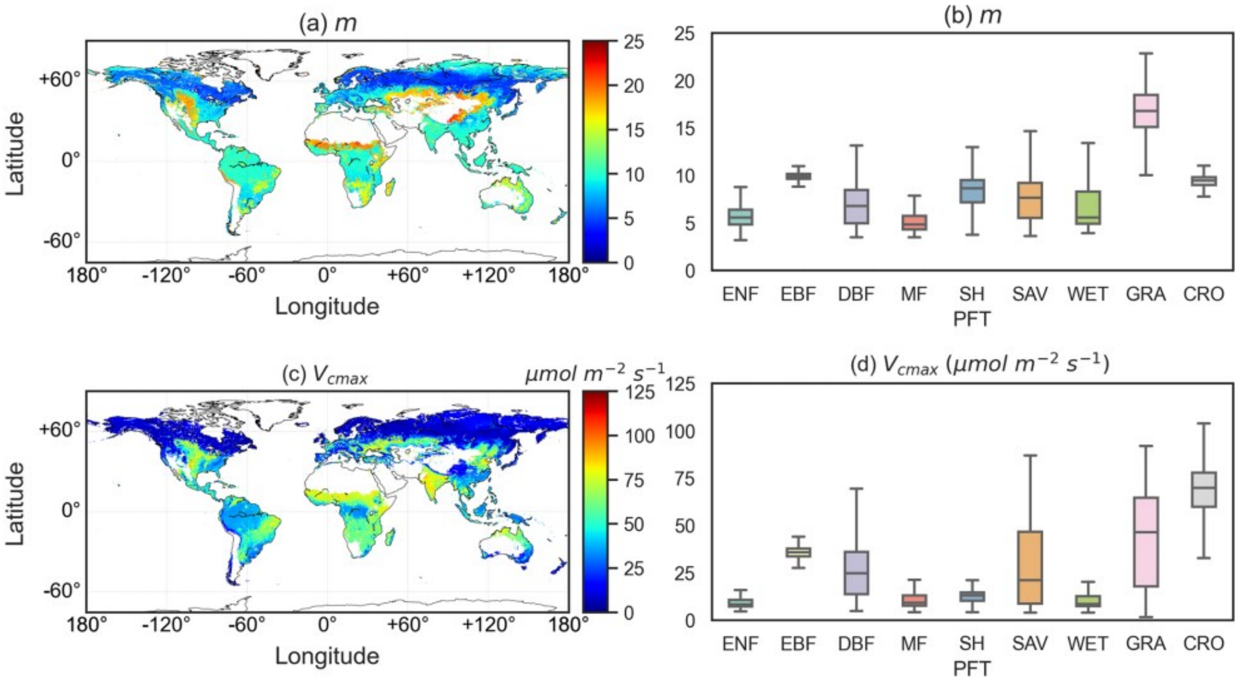
Jiye Leng[1], Jing M. Chen[1], Wenyu Li[1], Xiangzhong Luo[2], Mingzhu Xu[3], Jane Liu[1], Rong Wang[3], Cheryl Rogers[4], Bolun Li[5], and Yulin Yan[3]

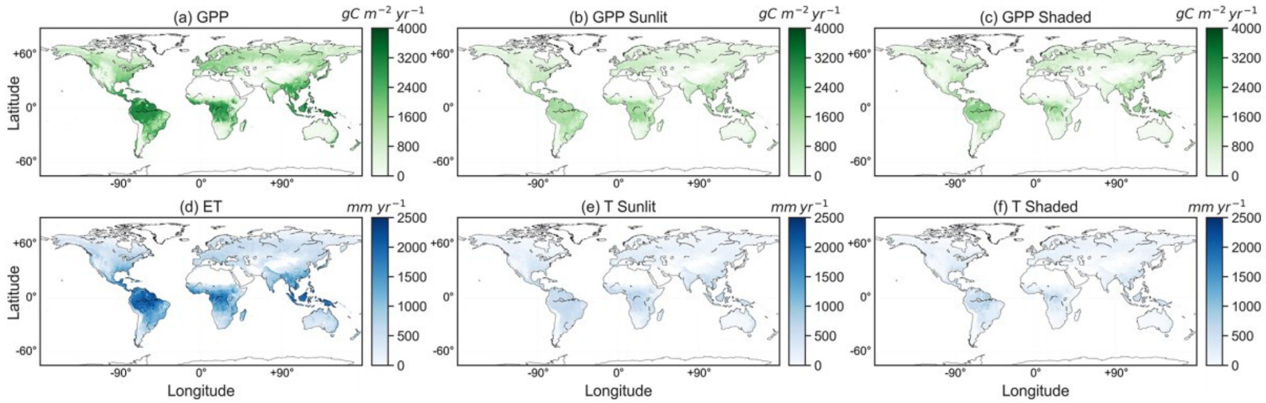
[1]Department of Geography and Planning, University of Toronto, Toronto, ON M5S 3G3, Canada
[2]Department of Geography, National University of Singapore, 1 Arts Link, 117570, Singapore
[3]School of Geographical Science, Fujian Normal University, Fuzhou 350007, China
[4]School of Earth, Environment and Society, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
[5]School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China

**Correspondence:** Jing M. Chen (jing.chen@utoronto.ca)

Received: 14 August 2023 – Discussion started: 24 August 2023
Revised: 5 January 2024 – Accepted: 31 January 2024 – Published: 13 March 2024

**Abstract.** Diagnostic terrestrial biosphere models (TBMs) forced by remote sensing observations have been a principal tool for providing benchmarks on global gross primary productivity (GPP) and evapotranspiration (ET). However, these models often estimate GPP and ET at coarse daily or monthly steps, hindering analysis of ecosystem dynamics at the diurnal (hourly) scales, and prescribe some essential parameters (i.e., the Ball–Berry slope ($m$) and the maximum carboxylation rate at 25 °C ($V_{cmax}^{25}$)) as constant, inducing uncertainties in the estimates of GPP and ET. In this study, we present hourly estimations of global GPP and ET datasets at a 0.25° resolution from 2001 to 2020 simulated with a widely used diagnostic TBM – the Biosphere–atmosphere Exchange Process Simulator (BEPS). We employed eddy covariance observations and machine learning approaches to derive and upscale the seasonally varied $m$ and $V_{cmax}^{25}$ for carbon and water fluxes. The estimated hourly GPP and ET are validated against flux observations, remote sensing, and machine learning-based estimates across multiple spatial and temporal scales. The correlation coefficients ($R^2$) and slopes between hourly tower-measured and modeled fluxes are $R^2 = 0.83$, regression slope $= 0.92$ for GPP, and $R^2 = 0.72$, regression slope $= 1.04$ for ET. At the global scale, we estimated a global mean GPP of $137.78 \pm 3.22$ Pg C $yr^{-1}$ (mean ± 1 SD) with a positive trend of 0.53 Pg C $yr^{-2}$ ($p < 0.001$), and an ET of $89.03 \pm 0.82 \times 10^3$ $km^3$ $yr^{-1}$ with a slight



# 2. GPP estimation model and its improvement

NUIST

Biogeosciences, 17, 1343–1365, 2020
https://doi.org/10.5194/bg-17-1343-2020


Biogeosciences
Open Access
EGU

## Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach

**Martin Jung[1], Christopher Schwalm[2], Mirco Migliavacca[1], Sophia Walther[1], Gustau Camps-Valls[3], Sujan Koirala[1], Peter Anthoni[4], Simon Besnard[1,5], Paul Bodesheim[1,6], Nuno Carvalhais[1,7], Frédéric Chevallier[8], Fabian Gans[1], Daniel S. Goll[9], Vanessa Haverd[10], Philipp Köhler[11], Kazuhito Ichii[12,13], Atul K. Jain[14], Junzhi Liu[1,15], Danica Lombardozzi[16], Julia E. M. S. Nabel[17], Jacob A. Nelson[1], Michael O'Sullivan[18], Martijn Pallandt[19], Dario Papale[20,21], Wouter Peters[22], Julia Pongratz[23,17], Christian Rödenbeck[19], Stephen Sitch[18], Gianluca Tramontana[20,3], Anthony Walker[24], Ulrich Weber[1], and Markus Reichstein[1]**

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, 07745 Jena, Germany
[2]Woods Hole Research Center, Falmouth, MA 02540-1644, USA
[3]Image Processing Laboratory (IPL), Universitat de València, Paterna, 46980, Spain
[4]Institute of Meteorology and Climate Research – Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology, 82467 Garmisch-Partenkirchen, Germany
[5]Laboratory of Geo-Information Science and Remote Sensing, Wageningen University and Research, Wageningen, 6708 PB, the Netherlands
[6]Computer Vision Group Mathematics and Computer Science, Friedrich Schiller University Jena, 07743 Jena, Germany
[7]Departamento de Ciências e Engenharia do Ambiente (DCEA), Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, Caparica, 2829-516, Portugal
[8]Laboratoire des Sciences du Climat et de l'Environnement (LSCE/IPSL), Université Paris-Saclay, Gif-sur-Yvette, 91198, France
[9]Department of Geography, University of Augsburg, 86159 Augsburg, Germany
[10]Department Continental Biogeochemical Cycles, CSIRO Oceans and Atmosphere, Canberra, 2601, Australia
[11]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[12]Center for Environmental Remote Sensing (CEReS), Chiba University, Chiba, 263-8522, Japan
[13]Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, 305-8506, Japan

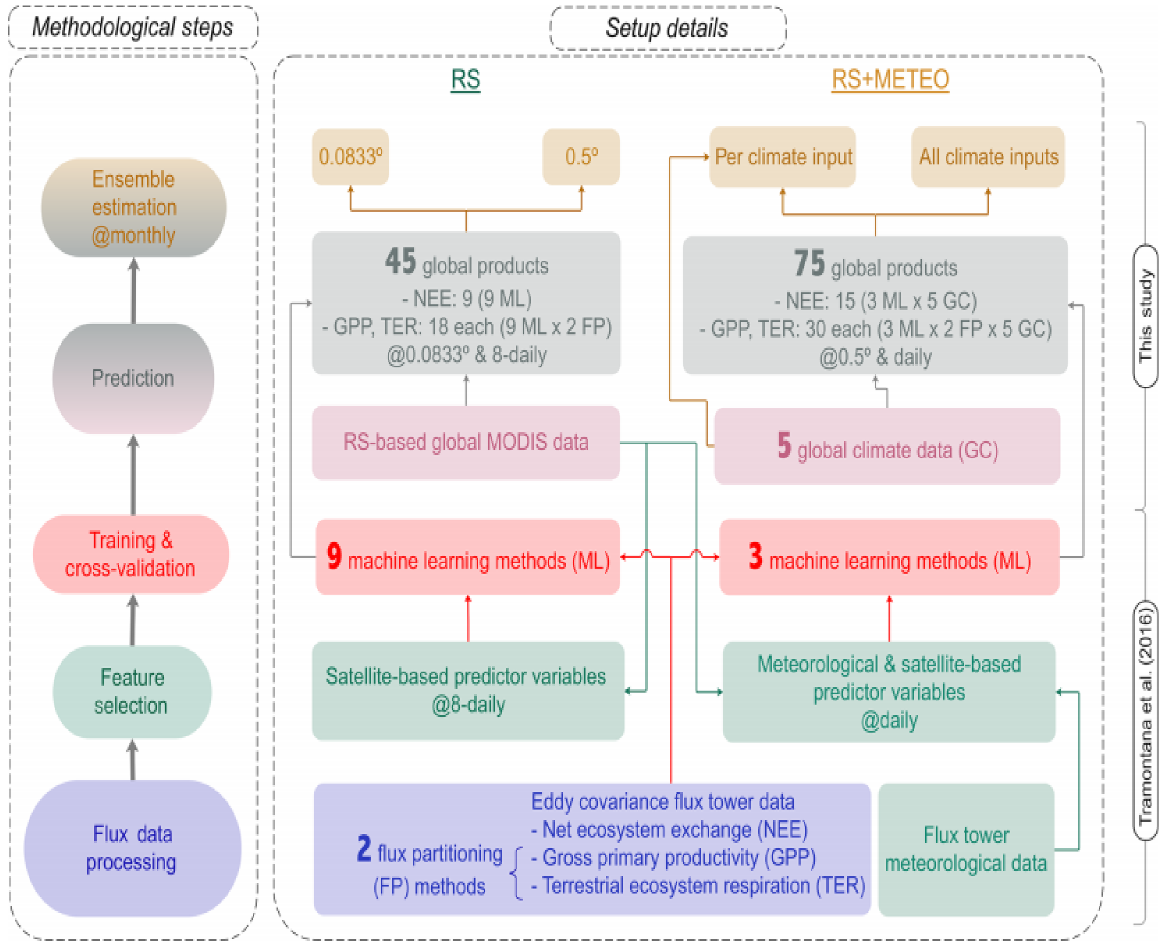
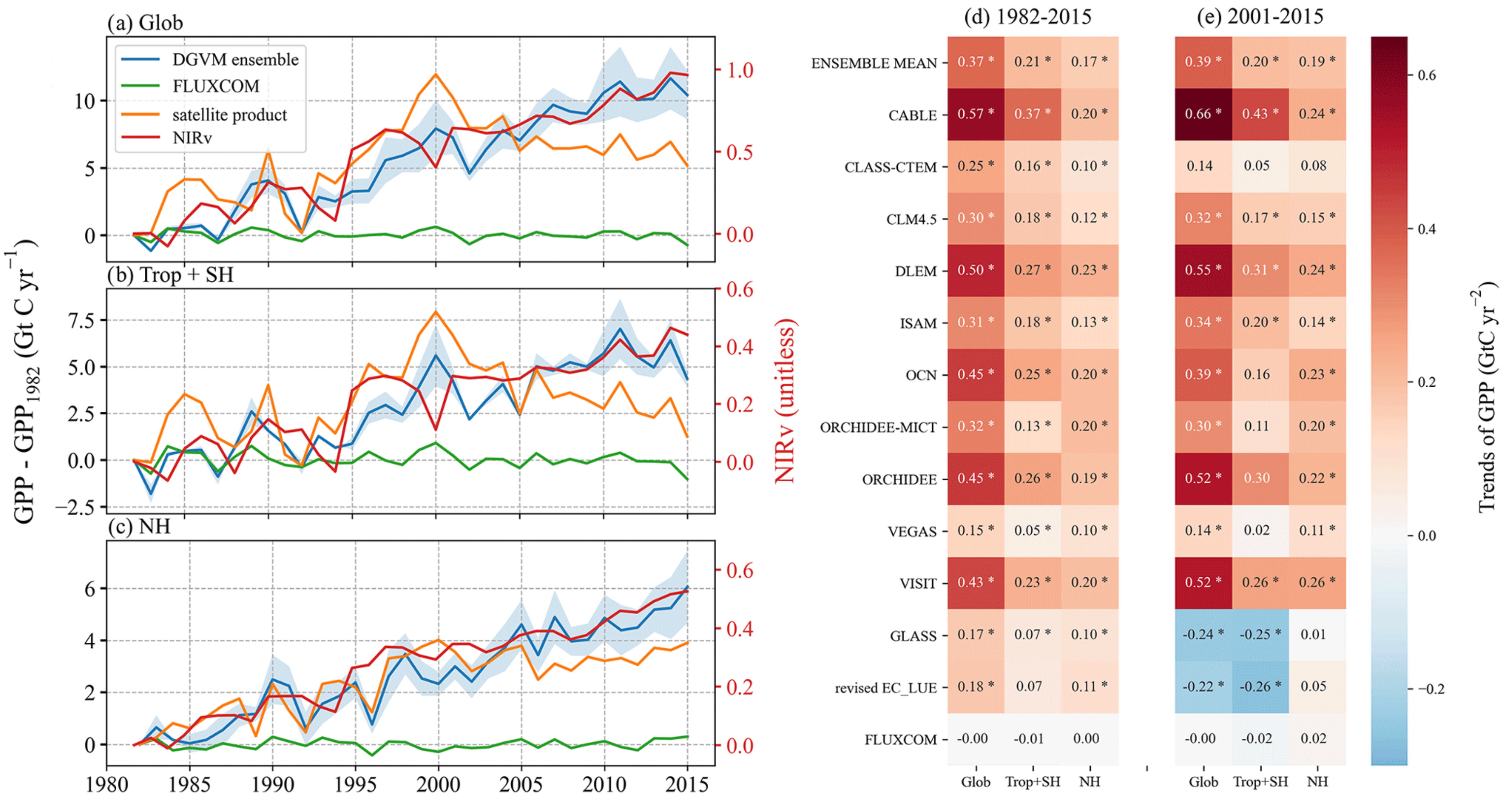



**Figure 1.** Schematic overview of the methodology and data products from the FLUXCOM initiative. The flow diagram shows the methodological steps for the remote sensing (RS, left) and the remote sensing and meteorological data (RS+METEO, right) FLUXCOM products. Final monthly ensemble products for NEE, GPP and TER from RS are available at 0.0833° and at 0.5° spatial resolution. Ensemble products from RS+METEO are available per climate forcing (GC) data set as well as a pooled ensemble at 0.5° spatial resolution. All ensemble products encompass ensemble members of different machine learning methods (ML, nine for RS, three for RS+METEO) and flux partitioning methods (FP, two for GPP and TER).

# 2. GPP estimation model and its improvement

NUIST




**Long-term trends in GPP products Yang 2022**

## Train machine learning models from the perspective of different vegetation types


JGR Biogeosciences

RESEARCH ARTICLE

10.1029/2022JG007100

**Special Section:**
Understanding carbon-climate feedbacks


**Key Points:**

- The accuracy of gross primary production (GPP) estimation can be improved by distinguishing plant functional types, especially for C3 and C4 crops
- Significant increasing trend is found in this random forest-based data set
- Leaf area index plays a leading role in both the average state and long-term trend of GPP

**Supporting Information:**

Supporting Information may be found in the online version of this article.


**Correspondence to:**

T. Chen,
txchen@nuist.edu.cn



**Citation:**

Guo, R., Chen, T., Chen, X., Yuan, W., Liu, S., He, B., et al. (2023). Estimating global GPP from the plant functional type perspective using a machine learning approach. *Journal of Geophysical Research: Biogeosciences*, *128*, e2022JG007100. https://doi.org/10.1029/2022JG007100


## Estimating Global GPP From the Plant Functional Type Perspective Using a Machine Learning Approach

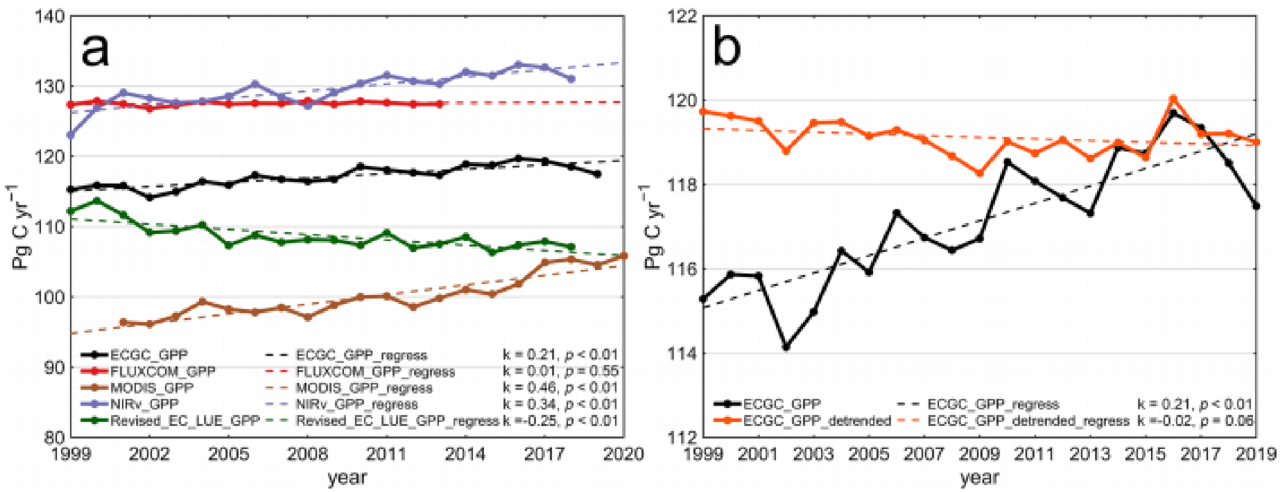

Renjie Guo[1], Tiexi Chen[1,2,3], Xin Chen[1], Wenping Yuan[4], Shuci Liu[5], Bin He[6], Lin Li[7,8], Shengzhen Wang[2,3], Ting Hu[9], Qingyun Yan[9], Xueqiong Wei[1], and Jie Dai[1]

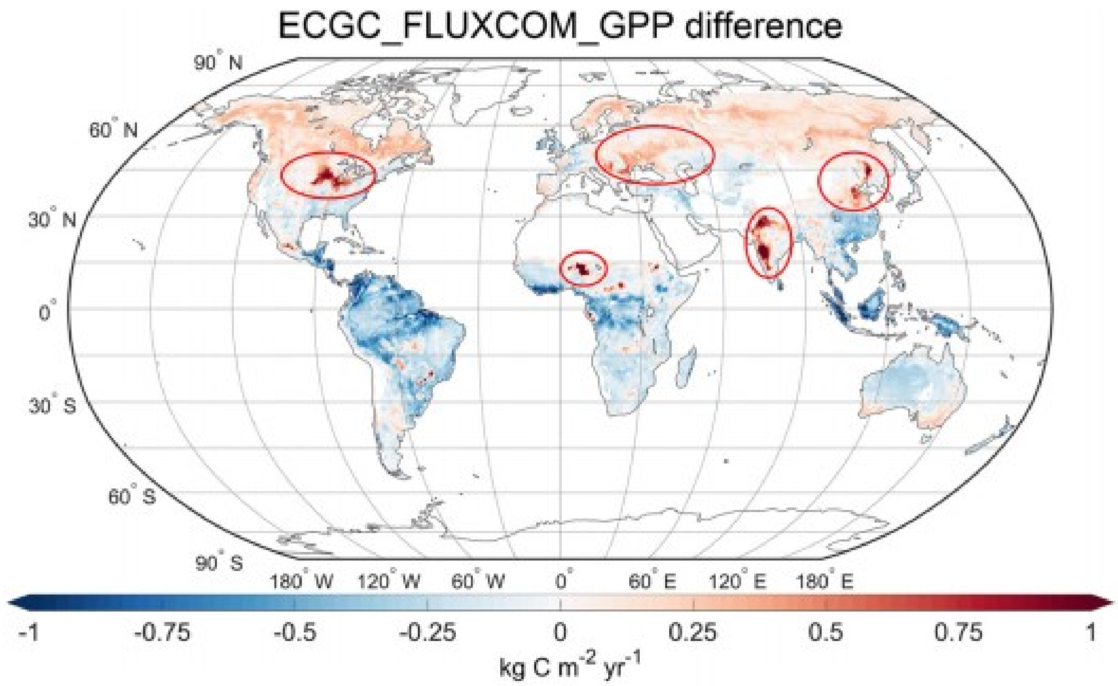
[1]School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing, China, [2]School of Geographical Sciences, Qinghai Normal University, Xining, China, [3]School of Ecology and Environmental Science, Xining University, Xining, China, [4]School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, China, [5]Department of Environment and Science, Queensland Government, Brisbane, Australia, [6]College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [7]Computer Department, Qinghai Normal University, Xining, China, [8]Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining, China, [9]School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing, China



**Abstract** The long-term monitoring of gross primary production (GPP) is crucial to the assessment of the carbon cycle of terrestrial ecosystems. In this study, a well-known machine learning model (random forest, RF) is established to reconstruct the global GPP data set named ECGC_GPP. The model distinguished nine functional plant types, including C3 and C4 crops, using eddy fluxes, meteorological variables, and leaf area index (LAI) as training data of RF model. Based on ERA5_Land and the corrected GEOV2 data, global monthly GPP data set at a 0.05° resolution from 1999 to 2019 was estimated. The results showed that the RF model could explain 74.81% of the monthly variation of GPP in the testing data set, of which the average contribution of LAI reached 41.73%. The average annual and standard deviation of GPP during 1999–2019 were 117.14 ± 1.51 Pg C $yr^{-1}$, with an upward trend of 0.21 Pg C $yr^{-2}$ ($p < 0.01$). By using the plant functional type classification, the underestimation of cropland is improved. Therefore, ECGC_GPP provides reasonable global spatial pattern and long-term trend of annual GPP.

**Plain Language Summary** Accurate estimation of gross primary production (GPP) is critical for understanding the terrestrial ecosystem carbon cycle. There are a variety of GPP data sets based on different methods, but huge differences validated by the GPP measured values of flux observation towers still exist. At present, a large amount of GPP measured data provides us with the opportunity to use machine learning models to estimate global GPP. This paper presents a new global GPP data set (ECGC_GPP) with 0.05° and

# 2. GPP estimation model and its improvement

NUIST

## Incorporation of $CO_2$ fertilization effects into machine learning models

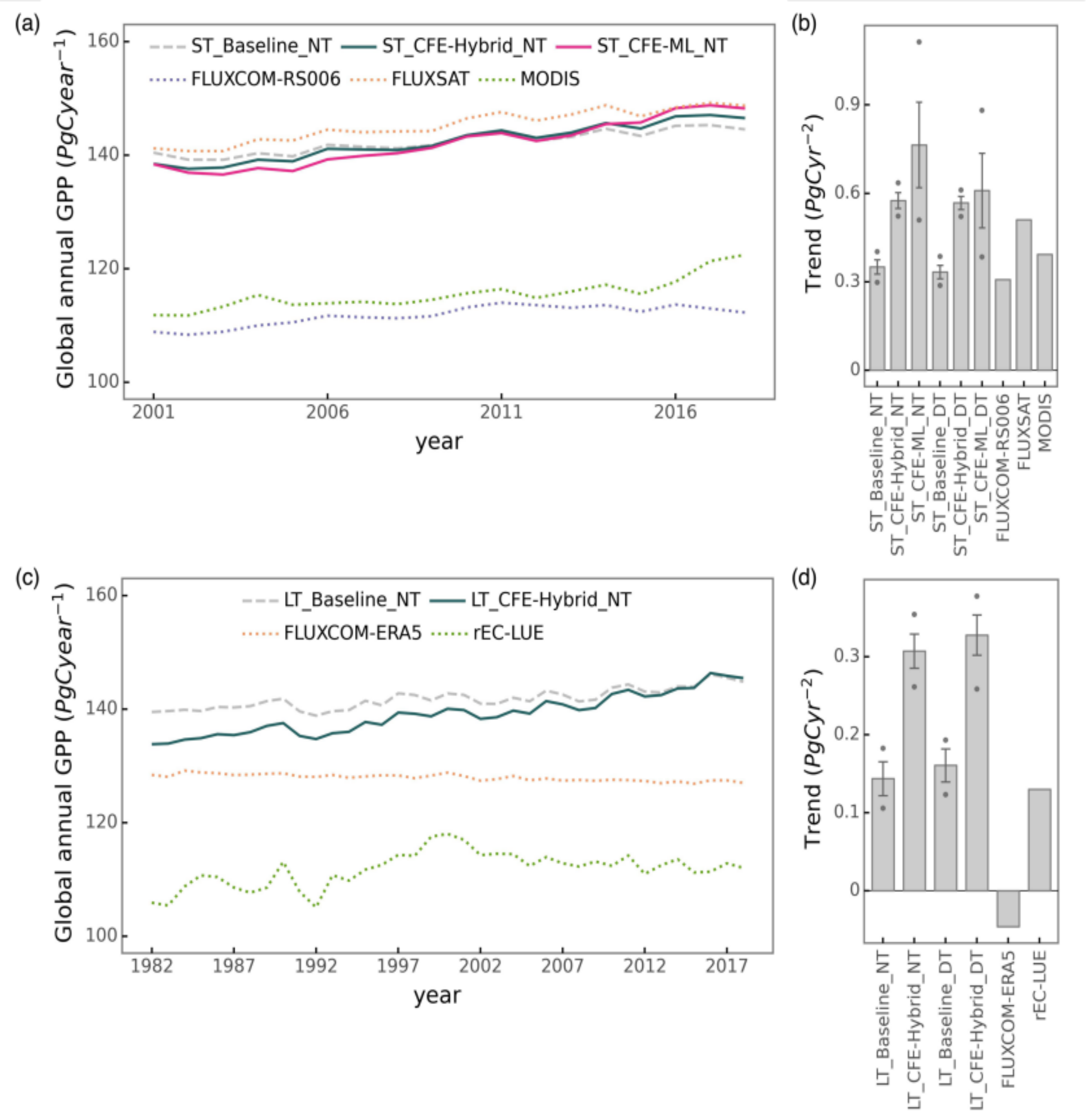
### CEDAR-GPP: spatiotemporally upscaled estimates of gross primary productivity incorporating $CO_2$ fertilization

Yanghui Kang, Max Gaber, Maoya Bassiouni, Xinchen Lu, and Trevor Keenan

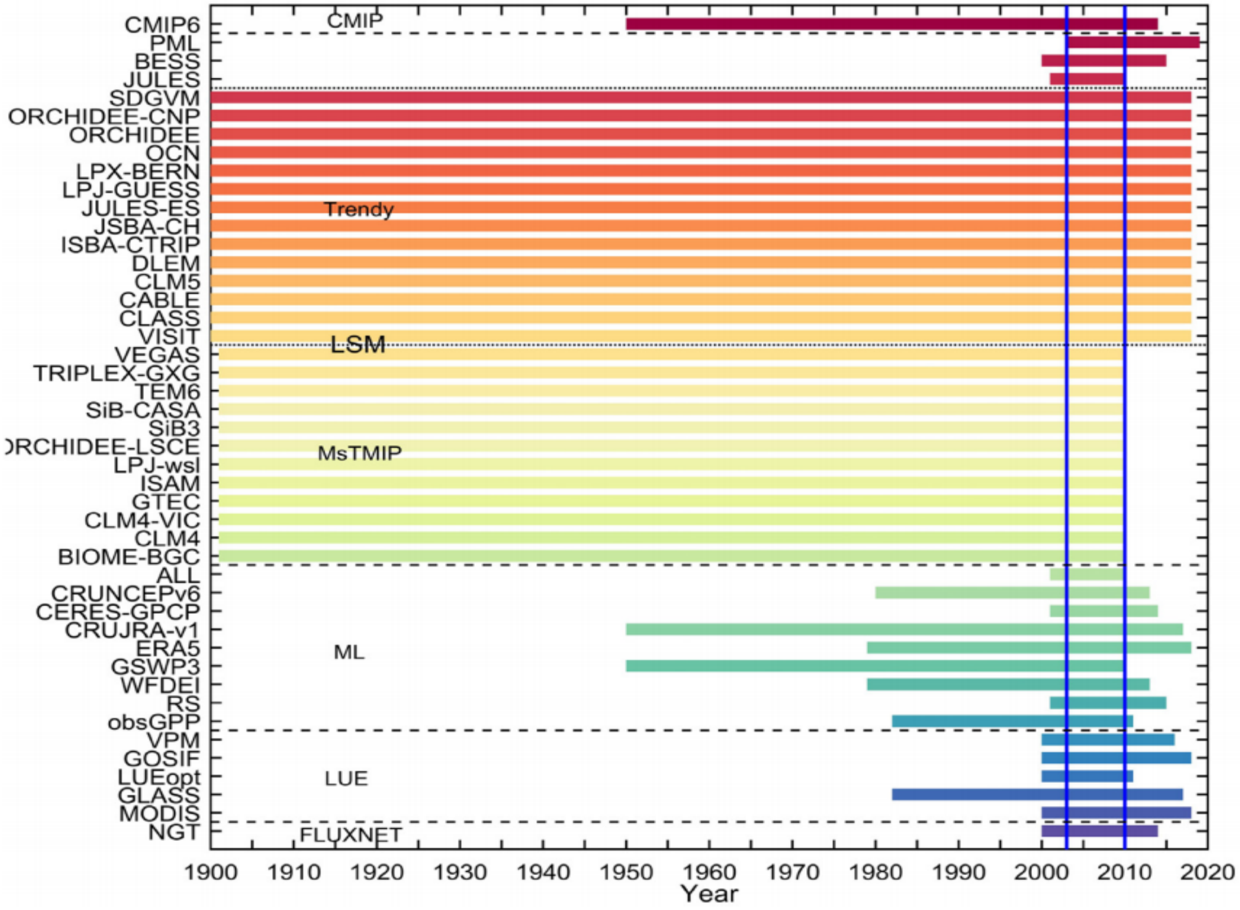
**Abstract.** Gross primary productivity (GPP) is the largest carbon flux in the Earth system, playing a crucial role in removing atmospheric carbon dioxide and providing the sugars and starches needed for ecosystem metabolism. Despite the importance of GPP, however, existing estimates present significant uncertainties and discrepancies. A key issue is the underrepresentation of the $CO_2$ fertilization effect, a major factor contributing to the increased terrestrial carbon sink over recent decades. This omission could potentially bias our understanding of ecosystem responses to climate change.

Here, we introduce CEDAR-GPP, the first global upscaled GPP product that incorporates the direct $CO_2$ fertilization effect on photosynthesis. Our product is comprised of monthly GPP estimates and their uncertainty at 0.05° resolution from 1982 to 2020, generated using a comprehensive set of eddy covariance measurements, multi-source satellite observations, climate variables, and machine learning models. Importantly, we used both theoretical and data-driven approaches to incorporate the direct $CO_2$ effects. Our machine learning models effectively predicted monthly GPP ($R^2$ ~ 0.74), the mean seasonal cycles ($R^2$ ~ 0.79), and spatial variabilities ($R^2$ ~ 0.67). Incorporation of the direct $CO_2$ effects substantially improved the models' ability to estimate long-term GPP trends across global flux sites. While the global patterns of annual mean GPP, seasonality, and interannual variability generally aligned with existing satellite-based products, CEDAR-GPP demonstrated higher long-term trends globally after incorporating $CO_2$ fertilization, particularly in the tropics, reflecting a strong temperature control on direct $CO_2$ effects. CEDAR-GPP offers a comprehensive representation of GPP temporal and spatial dynamics, providing valuable insights into ecosystem-climate interactions. The CEDAR-GPP product is available at https://doi.org/10.5281/zenodo.8212707 (Kang et al., 2023).

**How to cite.** Kang, Y., Gaber, M., Bassiouni, M., Lu, X., and Keenan, T.: CEDAR-GPP: spatiotemporally upscaled estimates of gross primary productivity incorporating $CO_2$ fertilization, Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2023-337, in review, 2023.

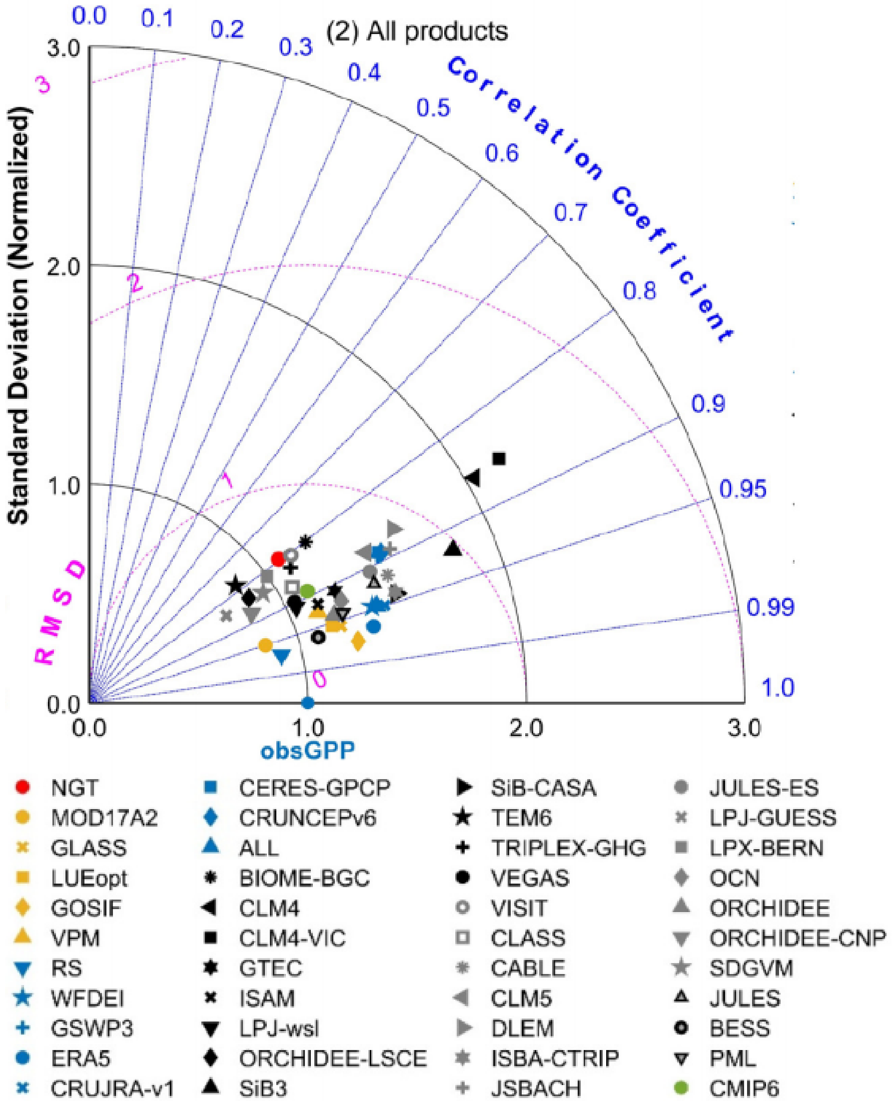


# 3. Uncertainty in GPP estimates




**Released GPP products（Zhang 2021）**

**Several GPP products have been released based on different approaches**




**Comparison with FLUXCOM-GPP（Zhang 2021）**

# 3. Uncertainty in GPP estimates

NUIST

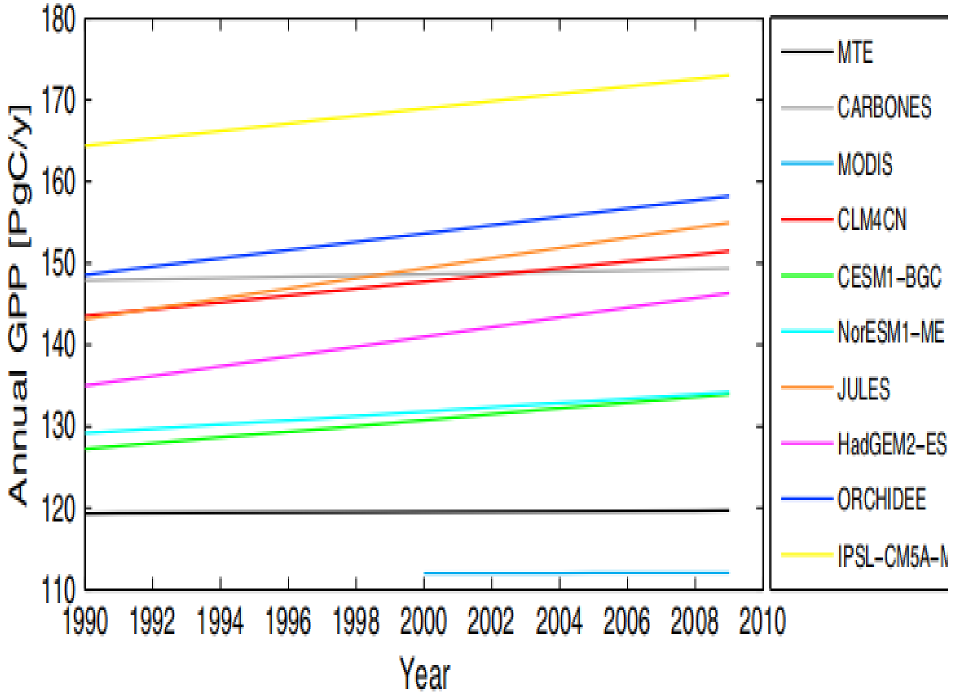



**Table 3.** Synthesis of Global Mean Annual GPP, Interannual Variability, and Trend

| Data Set | Mean (Pg C $yr^{-1}$) | IAV (Pg C $yr^{-1}$) | Trend (Pg C $yr^{-2}$) |
|---|---|---|---|
| MTE | 119 | 1.3 | 0.0181 |
| MODIS | 112 | 0.8 | 0.005 |
| CARBONES | 148 | 2.45 | 0.078 |
| CLM4CN | 147 | 2.87 | 0.417 |
| CESM1-BGC | 130 | 2.44 | 0.353 |
| NorESM1-ME | 131 | 1.76 | 0.262 |
| JULES | 149 | 4.4 | 0.621 |
| HadGEM2-ES | 140 | 3.89 | 0.598 |
| ORCHIDEE | 153 | 3.23 | 0.508 |
| IPSL-CM5A-MR | 169 | 3.26 | 0.454 |

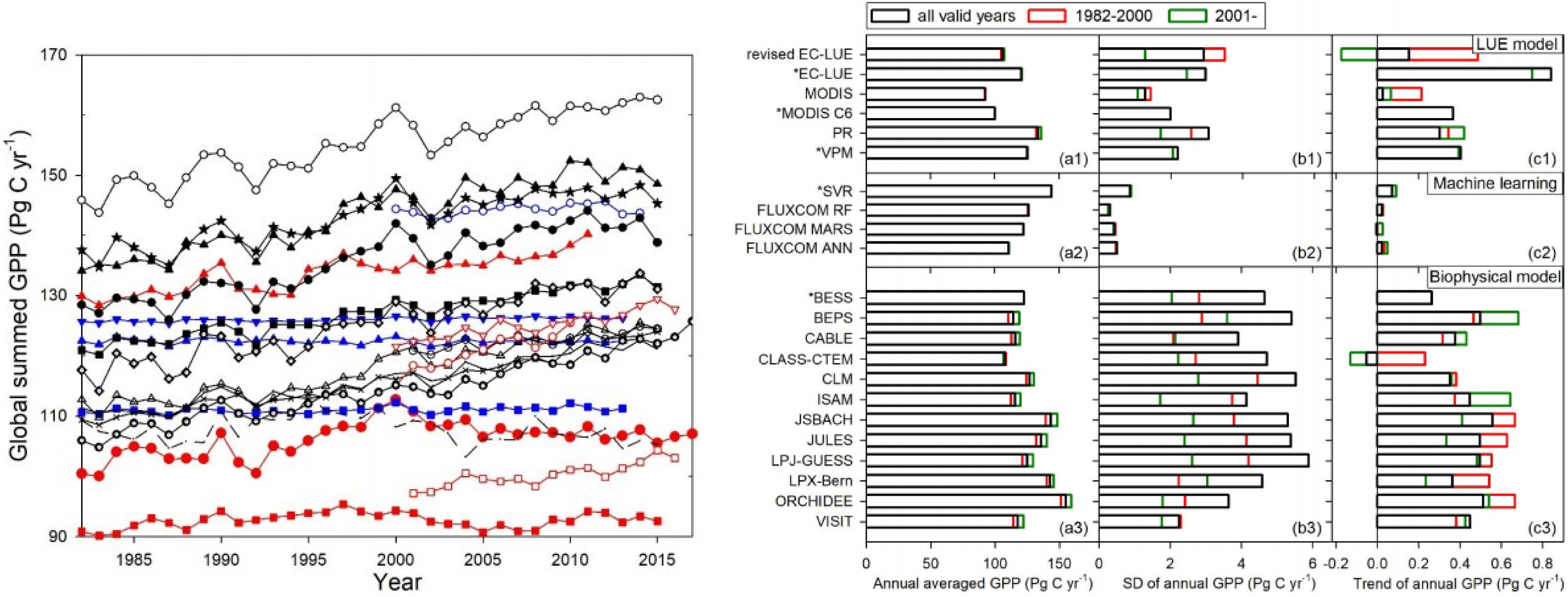
**Characteristics of different GPP products（Anav 2015）**

**The results of Anav et al. (2015) highlight the uncertainty in global GPP estimates.**

**Annual total: 119-169 Pg C $yr^{-1}$ Interannual variability: 0.8-3.89 Pg C $yr^{-1}$ Long-term trend: 0.0181-0.621 Pg C $yr^{-2}$**

# 3. Uncertainty in GPP estimates

NUIST



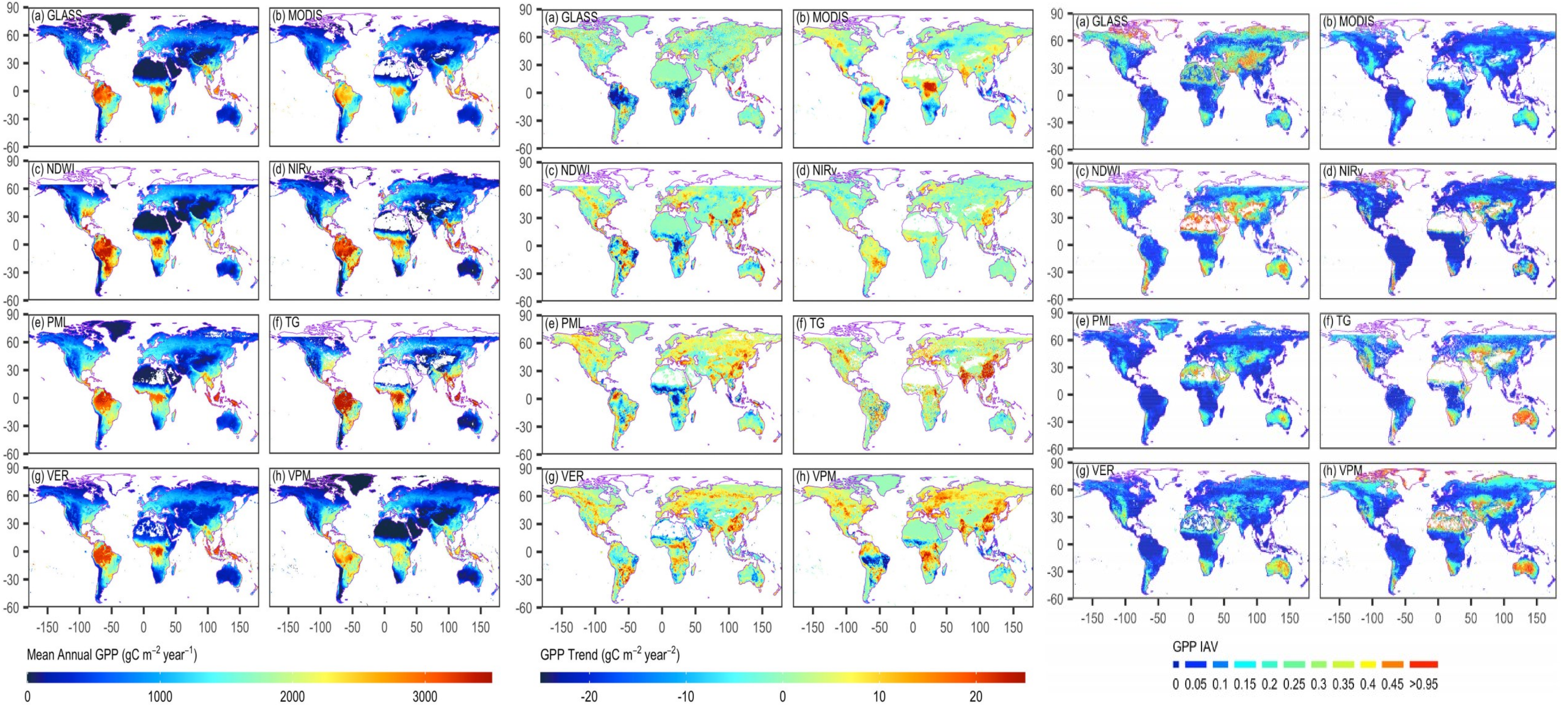

**Characteristics of different GPP products（Zheng 2020）**

**Based on the study by Zheng et al. (2020), the uncertainty in the GPP estimate is further expanded**

**Annual total: 92.7-168.7 Pg C yr$^{-1}$ Interannual variation: 0.32-5.89 Pg C yr$^{-1}$ Long-term trend: -0.05-0.84 Pg C yr$^{-2}$**

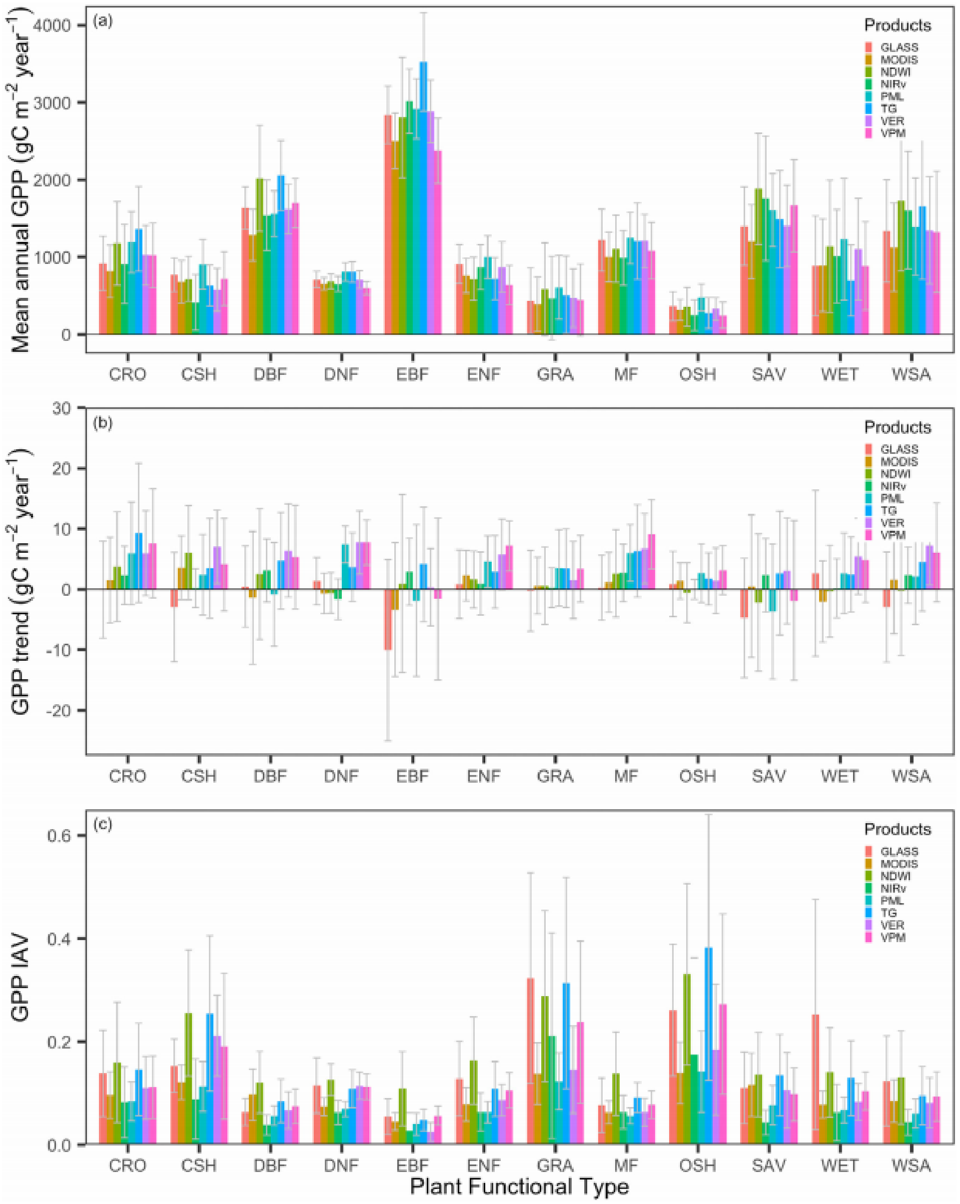
# 3. Uncertainty in GPP estimates

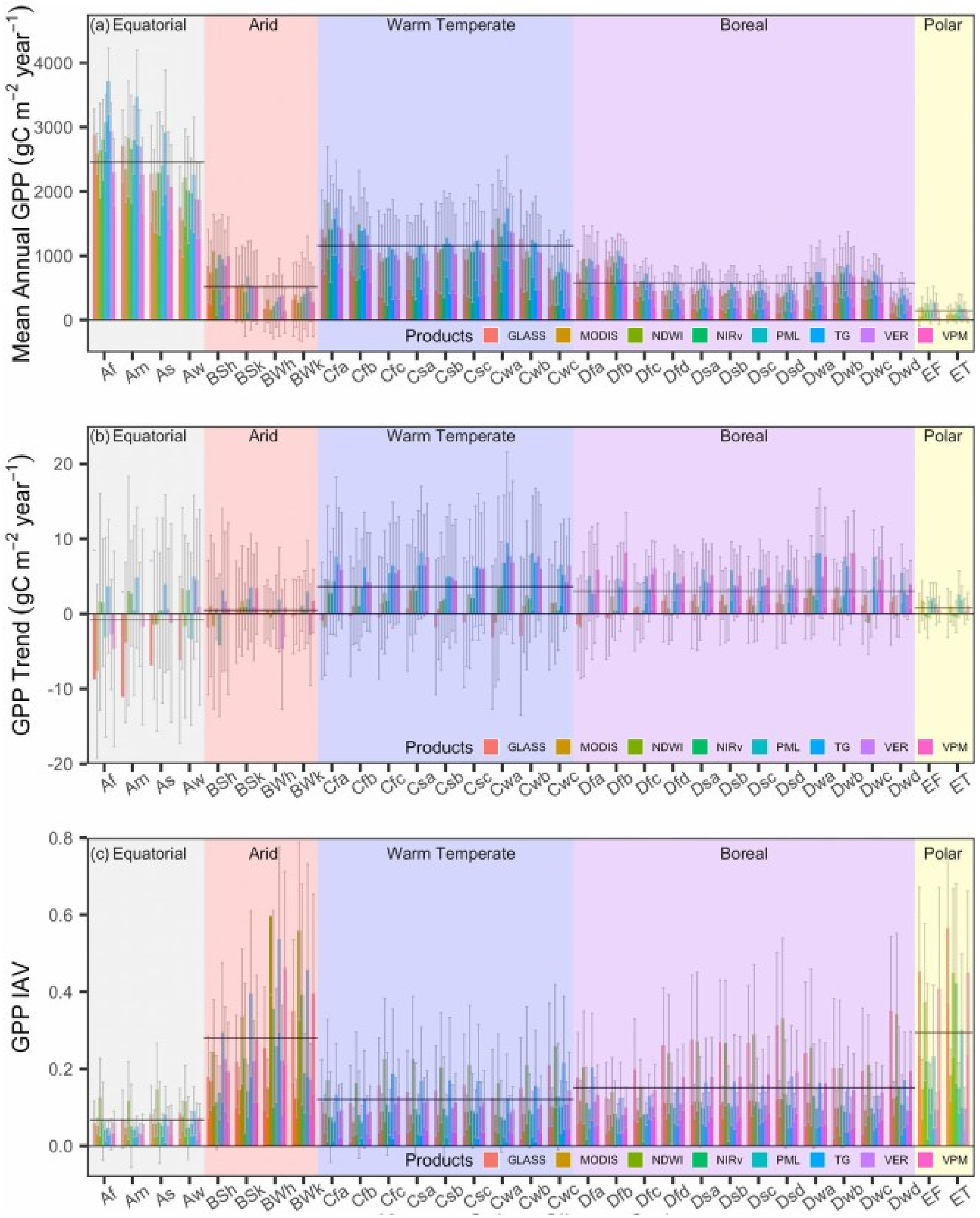





**Spatial distribution, trends and interannual variability (coefficient of variation) of GPP simulated by different models（Dong 2022）**

# 3. Uncertainty in GPP estimates

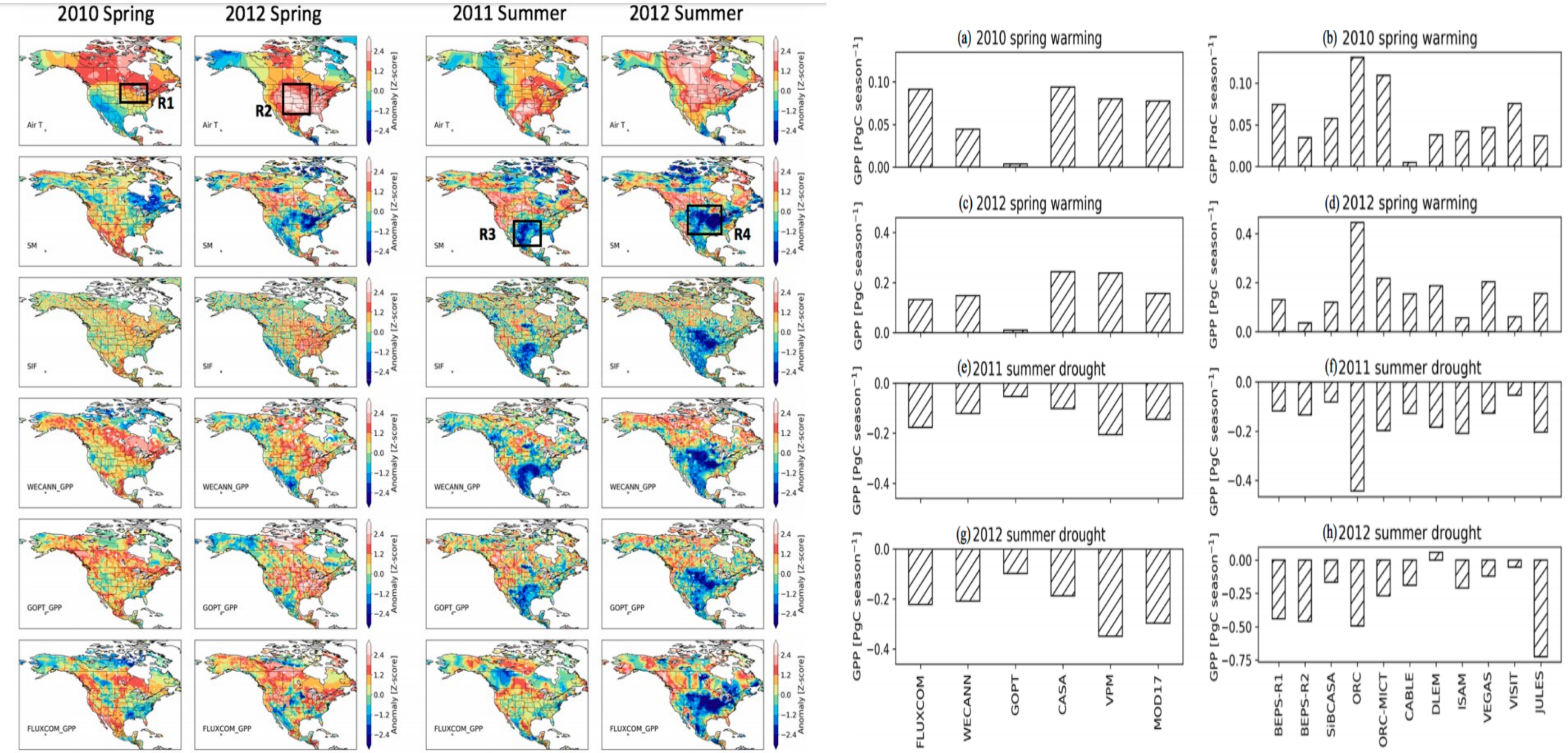



**The difference of GPP simulated by different models in different vegetation types and climatic zones（Dong 2022）**

# 3. Uncertainty in GPP estimates

NUIST




**Performance of GPP products of different models during drought（Dong 2022）**

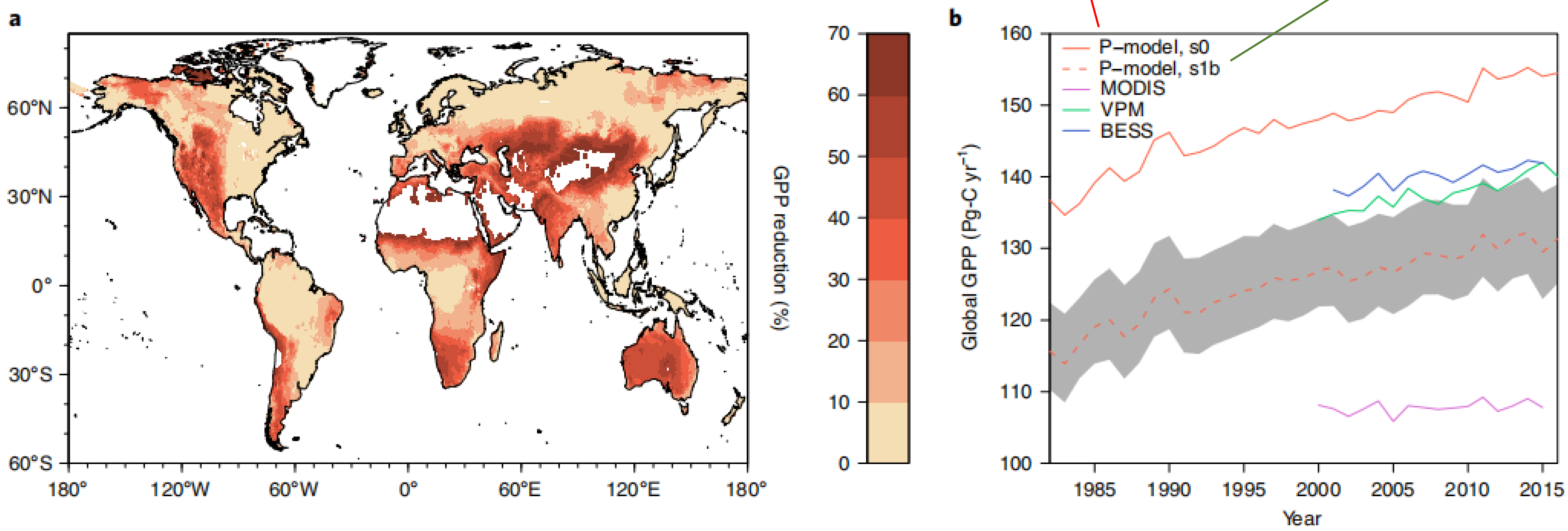
# 4. Sources of uncertainty in GPP estimates



**Differences in processes included in different GPP estimation models (model structure)**

Regardless of soil moisture

Consider soil moisture




**Effects of soil moisture on GPP estimates, where a represents the spatial difference between GPP estimates with and without soil moisture (Stocker 2019)**

**In a well-functioning model, it is often necessary to recalibrate the parameters when changing areas or using data from other sensors (model parameters)**

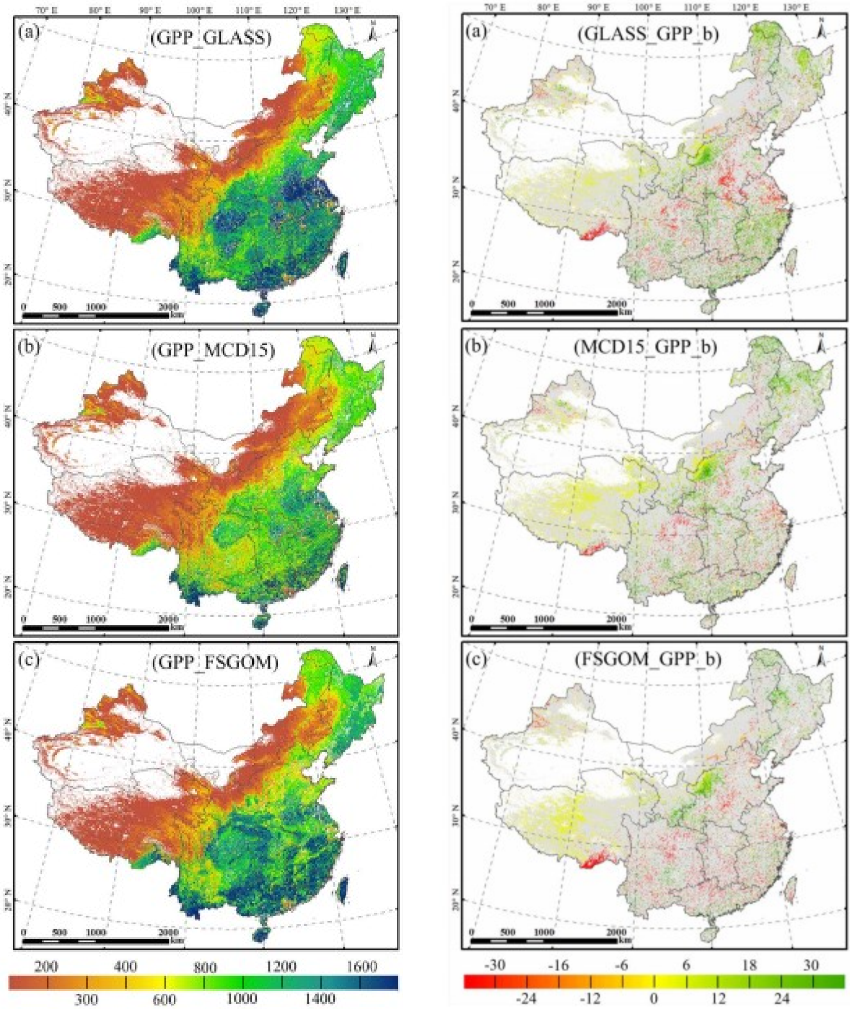
| Biome Type | Annual Mean (g C/$m^2$/yr) | | | Difference | | Biome area ($10^7$ $km^2$) | Annual Total (Pg C/yr) | |
|---|---|---|---|---|---|---|---|---|
| | $GPP_{EC}$ | $GPP_{VPM-b-LT}$ | $GPP_{VPM-sy-EVI}$ | GPP (g /$m^2$/yr) | Ratio | | $GPP_{VPM-b-LT}$ | $GPP_{VPM-sy-EVI}$ |
| CRO | 1308.08 | 1048.86 | 1234.83 | 185.96 | 17.73% | 1.01 | 10.60 | 12.48 |
| CSH | 1358.83 | 1282.76 | 1234.59 | -48.17 | -3.76% | 0.03 | 0.42 | 0.41 |
| DBF | 1567.88 | 1543.76 | 1572.00 | 28.23 | 1.83% | 0.23 | 3.53 | 3.59 |
| EBF | 2412.94 | 1662.70 | 1807.88 | 145.18 | 8.73% | 0.82 | 13.56 | 14.74 |
| ENF | 1372.88 | 846.79 | 883.74 | 36.95 | 4.36% | 0.27 | 2.28 | 2.38 |
| GRA | 1191.89 | 988.57 | 1172.73 | 184.15 | 18.63% | 2.71 | 26.79 | 31.78 |
| MF | 1414.49 | 1205.11 | 1216.54 | 11.42 | 0.95% | 0.49 | 5.95 | 6.01 |
| OSH | 307.43 | 224.36 | 301.19 | 76.83 | 34.25% | 1.41 | 3.16 | 4.24 |
| SAV | 1078.44 | 1053.15 | 1066.70 | 13.55 | 1.29% | 1.43 | 15.11 | 15.30 |
| WET | 1093.06 | 1091.54 | 1124.37 | 32.84 | 3.01% | 0.10 | 1.08 | 1.11 |
| WSA | 1103.08 | 1033.81 | 1063.38 | 29.57 | 2.86% | 1.08 | 11.20 | 11.52 |
| Annual Total GPP for Selected Biomes | | | | | | | 93.68 | 103.57 |

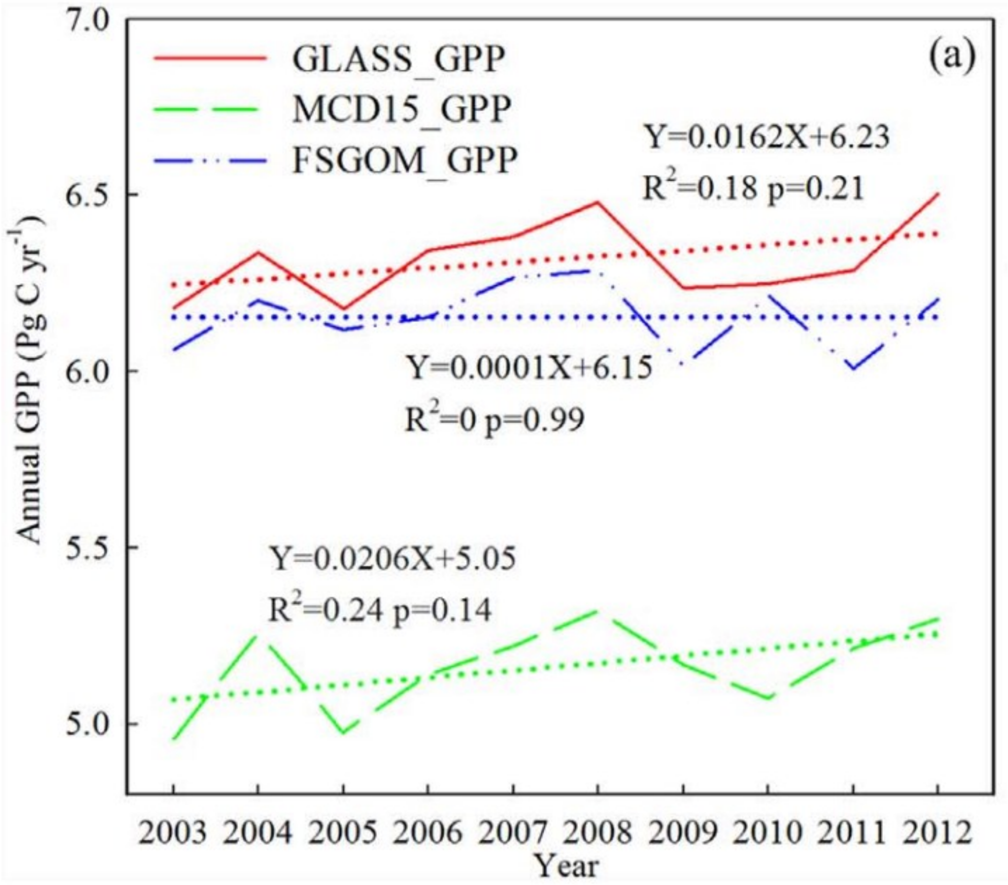
**The difference of GPP estimated by different model parameters in different vegetation types（Chang 2021）**

# 4. Sources of uncertainty in GPP estimates

NUIST

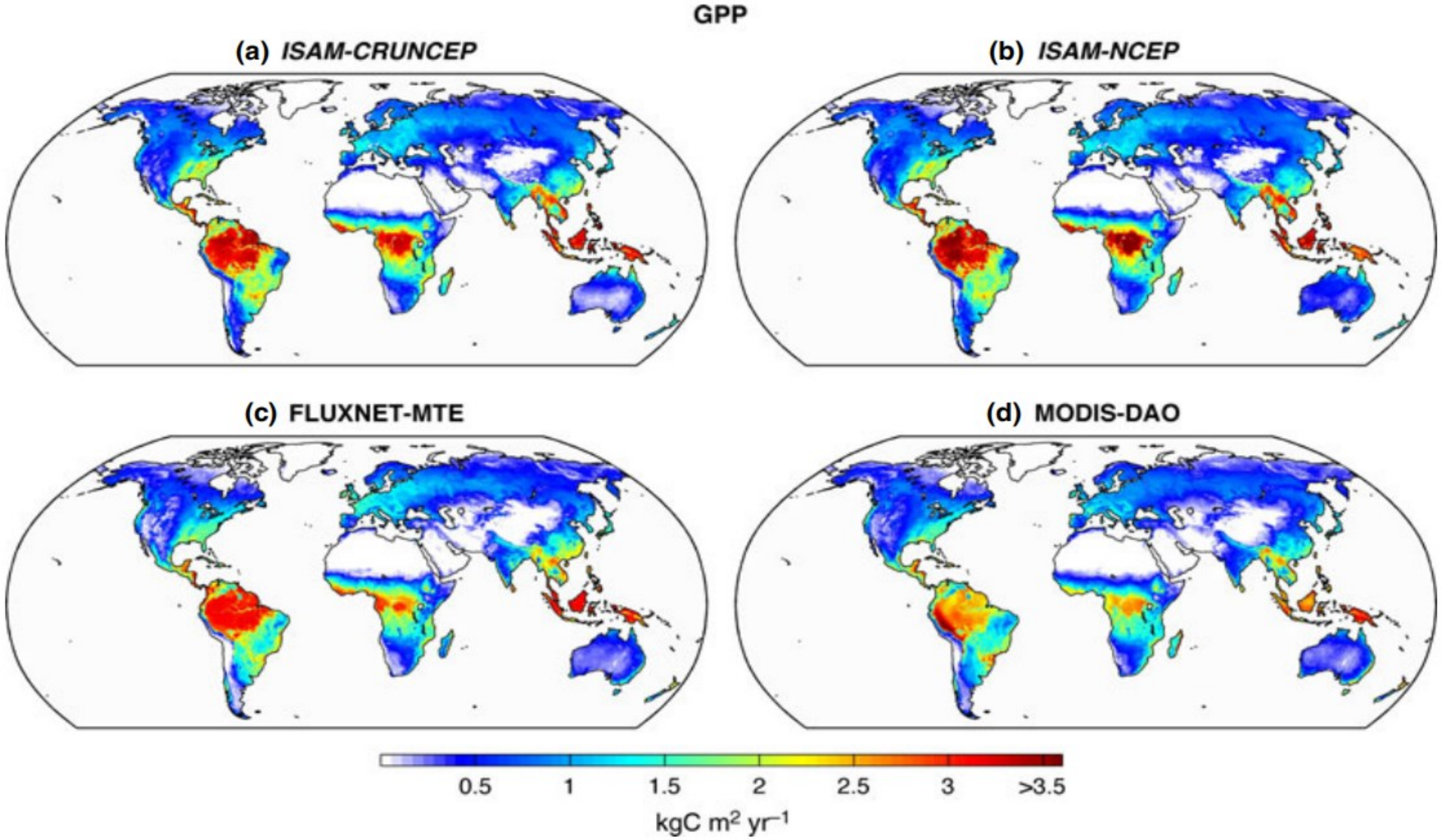
**When conducting GPP simulations on a global scale, global vegetation datasets need to be used, and vegetation datasets come from many sources (input datasets).**

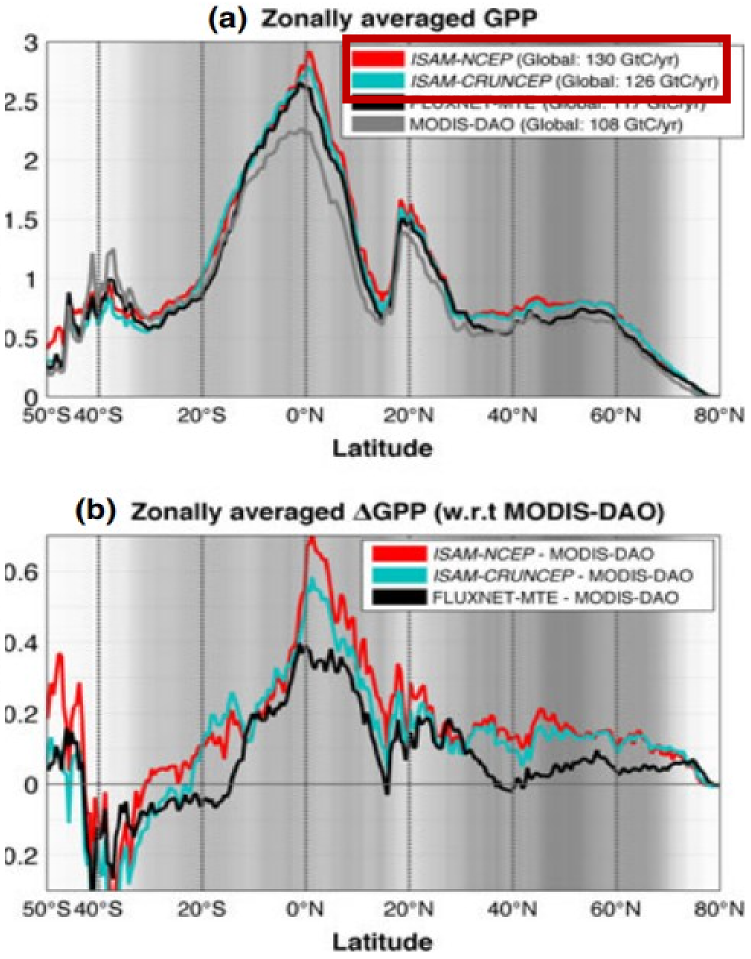



**Differences in GPP simulations based on different LAI（Liu 2018）**

**When conducting GPP simulations at the global scale, reconstructed meteorological data sets need to be used, and meteorological data sets come from many sources (input data sets).**




Differences in GPP driven by different meteorological data（Barman 2014）

In recent decades, the development of remote sensing and ground observation technology has made important progress in the theoretical research of GPP estimation. However, accurately quantifying and understanding the spatio-temporal pattern of GPP remains a serious challenge.

- The model structure, model parameters and input data sets all bring great uncertainty to GPP estimation.
- Generally speaking, This uncertainty is mainly reflected in the annual total of global GPP (92.7-169 Pg C $yr^{-1}$), the long-term trend (-0.05-0.84 Pg C $yr^{-2}$), and the interannual variability (0.32-5.89 Pg C $yr^{-1}$) and response to drought events (varying degrees of negative anomaly).

Therefore, it is important to emphasize that while the GPP estimates are relatively mature compared to other carbon fluxes, they are far from the point where they can slow down exploration. We must deeply understand existing theories and models, combining data from increasingly advanced satellite datasets with observations from flux towers to evaluate and constrain GPP estimation models to improve global GPP estimates.

# *Thanks*