AGU ADVANCING EARTH AND SPACE SCIENCE

Earth's Future

RESEARCH ARTICLE
10.1029/2022EF002805

Diffuse Radiation Forcing Constraints on Gross Primary Productivity and Global Terrestrial Evapotranspiration

T. Chakraborty[1,2], X. Lee[2], and D. M. Lawrence[3]

[1]Pacific Northwest National Laboratory, Richland, WA, USA, [2]Yale School of the Environment, New Haven, CT, USA, [3]National Center for Atmospheric Research, Boulder, CO, USA

Key Points:

- Diffuse radiation fertilization effect is understudied and hard to quantify globally using observations, necessitating model simulations
- Response of terrestrial carbon and energy budget simulated by Community Land Model to a realistic range of diffuse fraction forcing tested
- Large differences in carbon budget (small for water budget) due to range of forcing; systematic evaluations across models important

Supporting Information:

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

Correspondence to:

T. Chakraborty,
tc.chakraborty@pnnl.gov

Citation:

Chakraborty, T., Lee, X., & Lawrence, D. M. (2022). Diffuse radiation forcing constraints on gross primary productivity and global terrestrial evapotranspiration. *Earth's Future*, *10*, e2022EF002805. https://doi.org/10.1029/2022EF002805

Received 26 MAR 2022
Accepted 10 JUL 2022

Author Contributions:

Conceptualization: T. Chakraborty
Data curation: T. Chakraborty
Formal analysis: T. Chakraborty
Funding acquisition: X. Lee
Investigation: T. Chakraborty
Methodology: T. Chakraborty, X. Lee, D. M. Lawrence
Project Administration: X. Lee
Resources: X. Lee, D. M. Lawrence
Supervision: X. Lee, D. M. Lawrence
Validation: T. Chakraborty
Visualization: T. Chakraborty
Writing – original draft: T. Chakraborty, X. Lee, D. M. Lawrence

Abstract

The diffuse radiation fertilization effect—the increase in plant productivity in the presence of higher diffuse radiation ($K_{\downarrow,d}$)—is an important yet understudied aspect of atmosphere-biosphere interactions and can modify the terrestrial carbon, energy, and water budgets. The $K_{\downarrow,d}$ fertilization effect links the carbon cycle with clouds and aerosols, all of which are large sources of uncertainties for our current understanding of the Earth system and for future climate projections. Here we establish to what extent observational and modeling uncertainty in sunlight's diffuse fraction (k_d) affects simulated gross primary productivity (GPP) and terrestrial evapotranspiration (λE). We find only 48 eddy covariance sites with simultaneous sufficient measurements of $K_{\downarrow,d}$ with none in the tropical climate zone, making it difficult to constrain this mechanism globally using observations. Using a land modeling framework based on the latest version of the Community Land Model, we find that global GPP ranges from 114 Pg C year^{-1} when using k_d forcing from the Modern-Era Retrospective analysis for Research and Applications, version 2 reanalysis to a ~7% higher value of 122 Pg C year^{-1} when using the Clouds and the Earth's Radiant Energy System satellite product, with especially strong differences apparent over the tropical region (mean increase ~9%). The differences in λE, although smaller (−0.4%) due to competing changes in shaded and sunlit leaf transpiration, can be greater than regional impacts of individual forcing agents like aerosols. Our results demonstrate the importance of comprehensively and systematically validating the simulated k_d by atmosphere modules as well as the response to differences in k_d within land modules across Earth System Models.

Plain Language Summary

Due to clouds and small particles present in the atmosphere, some part of sunlight changes its direction, known as diffuse radiation. Leaves that are normally in the shadow of upper leaves can absorb this diffuse sunlight and then take part in photosynthesis, which also increases water released from them. The global strength of this effect—the diffuse radiation fertilization effect—is difficult to calculate using observations because most measurements are not in places where this effect might be strongest (like tropical forests). So, we commonly use computer models to calculate this. Here we first consider all sites that have the required measurements to study this effect to show that they are not suitable for global calculations. Then, we run a computer land model using different global datasets that give us a realistic range of diffuse radiation. We find that the change in photosynthesis due to this range has larger than expected effects on the carbon absorbed by the Earth's plants during photosynthesis in this model. The effects are less important for water released from leaves. Since different computer models calculate this effect differently, we need to test how other models react to similar ranges of diffuse radiation in the future.

1. Introduction

Clouds and the carbon cycle represent two large sources of uncertainty in our understanding of the Earth system, particularly relevant for the inter-model spread in future climate projections (Arias et al., 2021; Friedlingstein et al., 2014; Lawrence et al., 2016; Webb et al., 2017). An important and currently understudied mechanism that links cloud cover and the terrestrial carbon budget is the diffuse radiation fertilization effect (Mercado et al., 2009; Rap et al., 2018). The presence of scattering agents like clouds and aerosols in the atmosphere can change the direction of a portion of the total solar radiation ($K_{\downarrow}$), thus exposing normally shaded leaves to sunlight. By absorbing this diffuse radiation ($K_{\downarrow,d}$), these leaves can then contribute to photosynthesis, increasing carbon uptake by vegetation, enhancing evapotranspiration, and lowering surface and air temperature (Chakraborty et al., 2021; Knohl & Baldocchi, 2008; Mercado et al., 2009; Rap et al., 2018).

Writing – review & editing: T. Chakraborty, X. Lee, D. M. Lawrence

The $K_{\downarrow,d}$ fertilization effect is difficult to quantify and constrain with observations due to the dearth of simultaneous in situ measurements of $K_{\downarrow,d}$ and carbon and energy fluxes (Chakraborty & Lee, 2021a; Emmel et al., 2020; Steiner et al., 2013; Zhou et al., 2021). Consequently, to estimate the impact of $K_{\downarrow,d}$ fertilization effect on climate, we have to rely on global models, which of course have multiple sources of uncertainties. In atmospheric models, accurate estimates of $K_{\downarrow,d}$ depend on adequate parameterizations for clouds, radiation transfer, and aerosols, all of which vary widely between models (Chakraborty & Lee, 2021b; Pincus et al., 2016). Unfortunately, most models taking part in the Coupled Model Intercomparison Project do not publicly archive the diffuse component of $K_{\downarrow}$. For the few current-generation global reanalysis and satellite-derived products that do provide $K_{\downarrow,d}$, large differences in $K_{\downarrow,d}$ are seen, which is at least partly due to differences in cloud cover (Chakraborty & Lee, 2021a). On the land modeling side, capturing the response of surface climate to $K_{\downarrow,d}$ depends strongly on how the leaf-to-canopy upscaling process is represented, another major source of inter-model variability (Chakraborty et al., 2021; Luo et al., 2018).

Recent modeling evidence suggests that even when the total $K_{\downarrow}$ stays the same, changes in the diffuse fraction (k_d) affects gross primary productivity (GPP) and latent heat flux (λE) (Chakraborty et al., 2021). However, to reduce uncertainty associated with disparate representation of the $K_{\downarrow,d}$ fertilization effect requires improvements in multiple model components. Current generation inter-model comparisons have not focused on this aspect of atmosphere-biosphere interactions. For instance, for the Radiative Forcing Model Intercomparison project, the focus, naturally, is on the total radiative effect of climate forcers, but the partitioning of $K_{\downarrow}$ into $K_{\downarrow,d}$ and its direct beam component ($K_{\downarrow,b}$) (Pincus et al., 2016) is not considered. For the biosphere component, two relevant MIPs, the Land Surface, Snow and Soil moisture Model Intercomparison Project (LS3MIP) (van den Hurk et al., 2016) and Coupled Climate–Carbon Cycle Model Intercomparison Project (Jones et al., 2016), are not focused on the impact of $K_{\downarrow,d}$ on the carbon or energy cycle. None of the land-only forcing datasets used in the LS3MIP or Trends in the land carbon cycle (Sitch et al., 2015) projects provide k_d, meaning the partitioning of $K_{\downarrow}$ into $K_{\downarrow,d}$ and $K_{\downarrow,b}$ is left at the discretion of the land component, which also varies between models (Clark et al., 2011; Wozniak et al., 2020; Zhang et al., 2020).

Here we quantify the $K_{\downarrow,d}$ fertilization effect across a network of flux tower sites and then use a modeling framework with different global estimates of k_d to illustrate the important role of this inter-product k_d forcing spread on estimates of the terrestrial carbon and energy budgets. Our results demonstrate the need to comprehensively and systematically examine the simulated k_d by the atmosphere components and as well as the $K_{\downarrow,d}$ fertilization effect across land components in Earth System Models (ESMs).

2. Materials and Methods

2.1. Processing Site-Level Observations

We obtained publicly-available data from all AmeriFlux (Novick et al., 2018) (Table S1 in Supporting Information S1) and FLUXNET (Baldocchi et al., 2001) (Table S2 in Supporting Information S1) sites that include observations of $K_{\downarrow,d}$ (Figure 2a). Since the data structures from these two observation networks are different, their data were processed separately. The hourly FLUXNET measurements were subset based on quality control flags for the relevant variables, namely $K_{\downarrow,d}$, $K_{\downarrow}$, reflected shortwave radiation ($K_{\uparrow}$), λE, and GPP. The GPP field used was the one calculated using the daytime partitioning method (Lasslop et al., 2010). All hourly observations that were measured, gap-filled with high quality, or could be downscaled from reanalysis data were used. Finally, nighttime values and measurements corresponding to when the diffuse fraction ($k_d = K_{\downarrow,d}/K_{\downarrow}$) was greater than 1 or lower than 0 (both theoretically impossible) were removed.

For the AmeriFlux measurements, nighttime and physically impossible k_d values were first omitted. For multiple observations of $K_{\downarrow}$, $K_{\downarrow,d}$, or $K_{\uparrow}$ at a single site, the unweighted mean of the observations were used. AmeriFlux sites do not include the separated GPP field, so the net ecosystem exchange (NEE) columns were examined instead. All data points were binned based on absorbed radiation ($K_{abs} = K_{\downarrow} - K_{\uparrow}$) into 100 W m^{-2} bins between 100 and 600 W m^{-2}. K_{abs} is more relevant for estimating the available energy for photosynthesis at the canopy-scale than $K_{\downarrow}$, but similar results are seen when using $K_{\downarrow}$ bins (not shown). For each bin, low ($k_d < 0.35$) and high ($k_d > 0.65$) k_d regimes are defined, following Davin and Seneviratne (2012), and the variables of interest (moisture and carbon fluxes) were compared. Note that not all sites have sufficient (or any observations) in all bins and k_d regimes.

2.2. Simulating Meteorological and Default Radiative Forcing Data

Our modeling framework consists of generating climatological forcing data by running the Community Atmosphere Model (CAM) (Neale et al., 2010) and then simulating the surface energy and carbon budget by running the Community Land Model (CLM) (Lawrence et al., 2019). The latest version of CAM (CAM version 6) was first run with a slab ocean model, prescribed sea ice, and present-day distribution of aerosols for the period 2001–2003 at a spatial resolution of 0.9375 × 1.25°. Among other improvements, CAM6 uses a new cloud macrophysics parameterization for better performance while simulating boundary layer clouds and also captures cloud-aerosol interactions (indirect effect) in its default configuration (Gettelman et al., 2019). The atmospheric variables simulated by CAM that were used to force CLM include the direct beam radiation ($K_{\downarrow,b}$), $K_{\downarrow,d}$, incoming longwave radiation ($L_{\downarrow}$), and precipitation at surface and air temperature, specific humidity, wind speed, and atmospheric pressure at screen height.

2.3. Generating Monthly-Climatology-Adjusted Diffuse Fraction Forcing Data

In order to examine the sensitivity of model-simulated carbon and energy fluxes to a realistic spread of k_d, we extracted $K_{\downarrow,d}$ and $K_{\downarrow}$ at the surface for the 2001–2003 period from five global data products that publicly archive $K_{\downarrow,d}$ or $K_{\downarrow,b}$ (in addition to the CAM-simulated values). These data products are: (a) NOAA-CIRES-DOE – Twentieth Century Reanalysis version 3 (Slivinski et al., 2019) from National Oceanic and Atmospheric Administration, Cooperative Institute for Research in Environmental Science, and the Department of Energy, (b) National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) – 50-year Reanalysis (Kistler et al., 2001) from NCEP and NCAR, (c) MERRA-2 – Modern-Era Retrospective analysis for Research and Applications, version 2 (Randles et al., 2017) from National Aeronautics and Space Administration (NASA), (d) ERA5 – Fifth Generation Reanalysis (Hersbach et al., 2020) from the European Centre for Medium-Range Weather Forecasts, and (e) CERES – latest version of the Clouds and the Earth's Radiant Energy System product from NASA (CERES_EBAF_Ed4.1) (Rutan et al., 2015). Of these, $K_{\downarrow,d}$ is derived as the sum of diffuse photosynthetically active radiation (PAR_d) and diffuse near-infrared radiation (NIR_d) for MERRA-2 and as the difference between $K_{\downarrow}$ and $K_{\downarrow,b}$ for ERA5. Since these datasets have different spatial resolution (Chakraborty & Lee, 2021), all the datasets were interpolated to a regular 0.5 × 0.5° grid – the forcing resolution used for the subsequent land model runs – using nearest-neighbor interpolation.

The climatological state at the diurnal scale will not necessarily be consistent across all these products, partly because unlike assimilated surface meteorology, atmospheric constituents like clouds are modeled and aerosols are not explicitly represented in most of these products (except MERRA-2; Randles et al., 2017). Since the meteorological forcing and total $K_{\downarrow}$ are specific to the CAM-simulated (not assimilated) climatology and same for all the simulations, we adjusted the k_d for the other forcing data based on their monthly k_d. This monthly adjustment was done because unlike diurnal climatology, the monthly k_d climatology do show similar intra-annual patterns (but with large differences in magnitude; Chakraborty & Lee, 2021b). Thus:

$$K_{\downarrow,d,a} = k_{d,m} K_{\downarrow,h} \tag{1}$$

where $K_{\downarrow,d,a}$ is the monthly-climatology-adjusted three-hourly $K_{\downarrow,d}$ for a particular product, $k_{d,m}$ is monthly mean k_d for that month for that product, and $K_{\downarrow,h}$ is the three-hourly $K_{\downarrow}$ from the CAM simulations. Then $K_{\downarrow,b,a}$ (monthly-climatology-adjusted $K_{\downarrow,b}$) is the difference between $K_{\downarrow,h}$ and $K_{\downarrow,d,a.}$ Similarly, instead of using the three-hourly k_d simulated by CAM when generating the final CAM forcing data, we adjusted $K_{\downarrow,d}$ based on the average k_d for each simulation month for consistency with the result of the simulations. Another important reason for this adjustment is because NCEP/NCAR and NOAA-CIRES-DOE $K_{\downarrow,d}$ are only available at the monthly scale.

2.4. Land Model Simulations

The meteorological and $L_{\downarrow}$ forcing data from CAM and six sets of $K_{\downarrow,d}$ and $K_{\downarrow,b}$ fields (from NCEP/NCAR, NOAA-CIRES-DOE, MERRA-2, ERA5, CERES, and CAM after monthly-climatology adjustment) were used to run the latest version of the Community Land Model (CLM version 5; Lawrence et al., 2019) with biogeochemistry turned on. The biogeochemistry module allows for prognostic vegetation and helps us examine feedback on the canopy state due to the $K_{\downarrow,d}$ fertilization effect (and its inter-product spread). Since the differences in forcing are small (only due to changes in k_d), we allowed enough time for the model to adjust to the different forcing sets

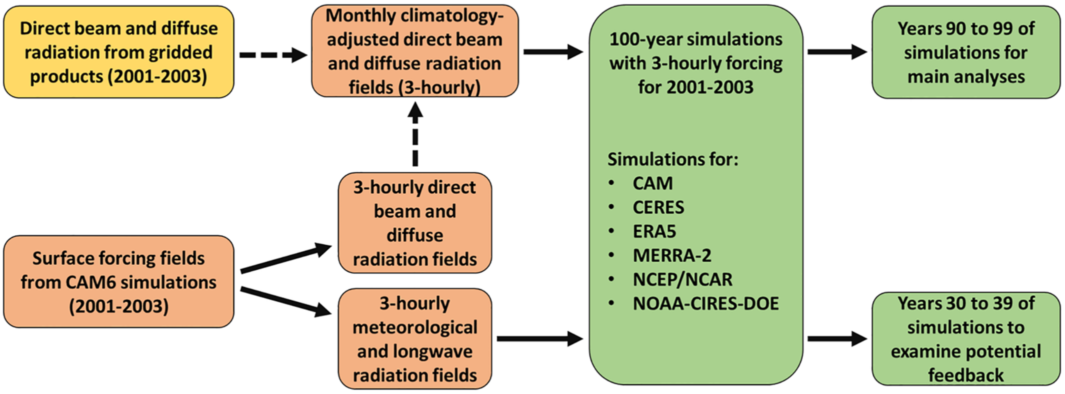

Figure 1. Schematic of the overall modeling framework.

by looping over the same forcing (2001–2003) for 100 years initiated for the year 2001. The results from years 90–99 of the simulations are presented as by then, all components of the carbon budget, including soil carbon would equilibrate to the forcing differences. To examine possible feedback, we also analyzed data for years 30–39 of the same simulations. The model outputs are for every month at a spatial resolution of 0.9375 × 1.25°. Note that CLM takes $K_{\downarrow,\mathrm{d}}$ and $K_{\downarrow,\mathrm{b}}$ separately as forcing fields and all model results presented are after calculating k_d from the monthly output fields for consistency with the model simulations. Figure 1 gives an overview of the modeling framework of this study.

In addition to the GPP, sensible heat flux (H), and λE, we examined how their sub-components respond to the inter-product spread in k_d. The ecosystem respiration (ER) was estimated as the difference between GPP and net primary productivity (NEP). The total λE can be further sub-divided into evaporation from ground (λE_g), evaporation from canopy (λE_c), and transpiration (λE_t), while the sensible heat flux H can be from the ground (H_g) or vegetation (H_v). All of these terms were simulated by CLM. We modified the CLM code to separately output the total λE_t from sunlit ($\lambda E_{\mathrm{t,sun}}$) and shaded leaves ($\lambda E_{\mathrm{t,sha}}$). These modifications are based on the internal implementation of the two big-leaf model of evapotranspiration in CLM (Oleson et al., 2013) and given by:

$$\lambda E_{\mathrm{t,sun}} = \frac{\frac{\mathrm{LAI_{sun}}}{r_\mathrm{b}+r_\mathrm{s}^{\mathrm{sun}}}}{\frac{\mathrm{LAI_{sun}}}{r_\mathrm{b}+r_\mathrm{s}^{\mathrm{sun}}}+\frac{\mathrm{LAI_{sha}}}{r_\mathrm{b}+r_\mathrm{s}^{\mathrm{sha}}}}\lambda E_\mathrm{t} \tag{2}$$

and

$$\lambda E_{\mathrm{t,sha}} = \frac{\frac{\mathrm{LAI_{sha}}}{r_\mathrm{b}+r_\mathrm{s}^{\mathrm{sha}}}}{\frac{\mathrm{LAI_{sun}}}{r_\mathrm{b}+r_\mathrm{s}^{\mathrm{sun}}}+\frac{\mathrm{LAI_{sha}}}{r_\mathrm{b}+r_\mathrm{s}^{\mathrm{sha}}}}\lambda E_\mathrm{t} \tag{3}$$

Here, $\mathrm{LAI_{sun}}$ and $\mathrm{LAI_{sha}}$ are the leaf area index for sunlit and shaded leaves, respectively, $r_\mathrm{s}^{\mathrm{sun}}$ and $r_\mathrm{s}^{\mathrm{sha}}$ are the stomatal resistances for sunlit and shaded leaves, respectively, and r_b is the leaf boundary layer resistance.

2.5. Regions of Interest

Land area weighted means of the variables of interest were calculated using the CLM surface data set. Additionally, the CLM grids were also separated into the Koppen-Geiger climate zones (Rubel & Kottek, 2010), namely tropical, arid, temperate, boreal, and polar (Figure 2a) and similar weighted means were calculated for these zones. These climate zones represent distinct classes of surface characteristics and atmospheric forcing (Chakraborty & Lee, 2019; Rubel & Kottek, 2010).

2.6. Impact of Monthly-Climatology Adjustment on Model Simulations

In its default configuration (CAM forcing to drive CLM), this modeling framework has been extensively evaluated against both gridded and point measurements in a previous study (Chakraborty et al., 2021). Additionally,

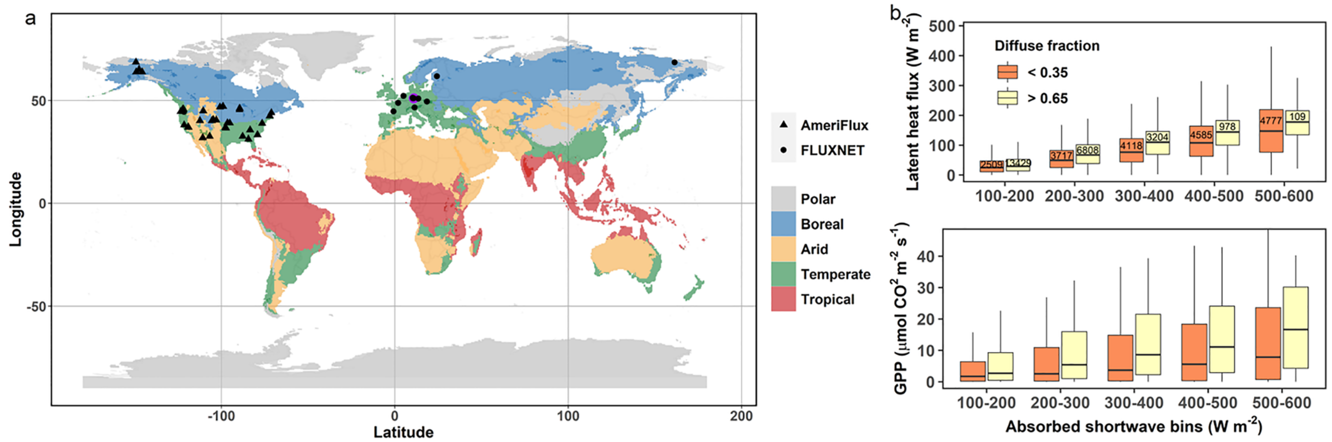

Figure 2. Diffuse radiation fertilization effect at the site scale. Sub-figure (a) shows the locations of the measurement sites with simultaneous measurements of diffuse radiation, carbon fluxes, and energy fluxes considered in this study. The background colors represent the extent of the Koppen-Geiger climate zones used to examine regional trends. Sub-figure (b) illustrates the latent heat flux and gross primary productivity (GPP) for the Gebesee FLUXNET site in Germany (site with the most available data points; location denoted by concentric purple circle in sub-figure (a)) for high and low regimes of diffuse fraction in different absorbed shortwave radiation bins (similar results when using incoming shortwave radiation bins; not shown). The number of hourly observations in each bin is noted and all differences are statistically significant ($p < 0.01$). Results for the rest of the sites are summarized in Tables S3–S6 of Supporting Information S1.

the $K_{\downarrow,\mathrm{d}}$ and $K_{\downarrow}$ fields of all the datasets have been compared with in situ measurements (Chakraborty et al., 2021; Chakraborty & Lee, 2021b). The monthly-climatology adjustment would have some impact on the simulations though (Zhang, Ciais et al., 2021), since k_d varies both during the month and even during the day and the bias-adjustment thus overestimates the k_d slightly. To quantify the impact of this simplification on model simulations, we compared the relevant variables (GPP, λE, H) for two simulations – one using the original CAM-simulated $K_{\downarrow,\mathrm{d}}$ ($K_{\downarrow,\mathrm{d,h}}$) and another using monthly-climatology-adjusted values ($K_{\downarrow,\mathrm{d,a}}$). The results are summarized in Table 1 for global land surfaces and each climate zone. Overall, the spatial patterns are virtually identical ($r^2 = 0.99$) in all cases with small biases. The biases are greatest for GPP at −2.26% for global surfaces, which are smaller than the overall perturbations we see between the products. Although the differences in GPP between the simulations could partly be the result of this artifact, we would not expect the direction of

Table 1
Evaluation of the Sensible Heat Flux, Latent Heat Flux, and Gross Primary Productivity Simulated by CLM Using the Original CAM $K_{\downarrow,d}$ ($K_{\downarrow,d,h}$) Forcing and the Monthly-Climatology-Adjusted Values ($K_{\downarrow,d,a}$) for the World's Land Surfaces and for Each Climate Zone

		Regions of interest					
Variable	Case	Global land	Tropical	Arid	Temperate	Boreal	Polar
Sensible heat flux (W m^{-2})	CAM $K_{\downarrow,\mathrm{d,h}}$	32	42.02	56.16	40.67	19.71	−9.84
	CAM $K_{\downarrow,\mathrm{d,a}}$	31.94	42.12	56.08	40.61	19.5	−9.91
	r^2	0.99	0.99	0.99	0.99	0.99	0.99
	MPE (%)	−0.19	0.24	−0.14	−0.15	−1.07	0.71
Latent heat flux (W m^{-2})	CAM $K_{\downarrow,\mathrm{d,h}}$	37.4	80.06	24.24	51.8	27.83	7.3
	CAM $K_{\downarrow,\mathrm{d,a}}$	37.36	79.76	24.25	51.65	27.96	7.36
	r^2	0.99	0.99	0.99	0.99	0.99	0.99
	MPE (%)	−0.11	−0.37	0.04	−0.29	0.47	0.82
Gross primary productivity (Pg C year^{-1})	CAM $K_{\downarrow,\mathrm{d,h}}$	119.73	58.11	12.93	23.37	23.01	2.4
	CAM $K_{\downarrow,\mathrm{d,a}}$	117.02	56.7	12.83	22.62	22.57	2.41
	r^2	0.99	0.99	0.99	0.99	0.99	0.99
	MPE (%)	−2.26	−2.43	−0.77	−3.21	−1.91	0.42

Note. The top two rows for each variable show the grid-area weighted mean for the two cases (grid-area weighted sum for gross primary productivity). The statistical parameters for model evaluation are the coefficient of determination (r^2) and mean percentage error (MPE).

the bias to be different across the simulations when using $K_{\downarrow,d,h}$ instead of $K_{\downarrow,d,a}$. Additionally, we can compare these perturbations against the results of a previous study using a similar modeling framework for the aerosol impact on surface processes that used actual three-hourly k_d differences based on radiation diagnostic simulations (Chakraborty et al., 2021). That study showed that an increase in the global k_d over land from 0.27 (comparable to an aerosol-free atmosphere) to 0.34 would increase GPP by 2.2 Pg C y^{-1} (1.8%). Linearly extrapolating to the range of k_d used here (0.35–0.60) would lead to a change in GPP of 7.8 Pg C y^{-1} versus the 7.6 Pg C y^{-1} found here (see Results). Although we do see some consistency here for CLM, the difference in simulated GPP between a monthly-climatology-adjusted k_d and diurnally varying k_d is also dependent on model structure. For example, a 13 Pg C y^{-1} difference in GPP was found when using climatologically averaged monthly k_d instead of 6-hourly k_d in the recently developed ORCHIDEE-DF land-surface model (LSM) though using a different study design (Zhang, Boucher, et al., 2021), suggesting the need for a standardized framework to compare the response to $K_{\downarrow,d}$ across LSMs.

2.7. Statistical Analysis

For the in situ AmerFlux and FLUXNET observations, two-sampled t-tests were used to confirm whether the GPP (or NEE) and λE are statistically different ($p < 0.01$) between the low and high k_d regimes in each bin. For the global study, we examined the inter-product spread at the grid level by calculating standard deviation (σ) from the six simulations with the six k_d forcing data (Figure 4). Since standard deviation would be impacted by the baseline values, we also calculated the coefficient of variation (CV), which is unitless and scale independent, to get the relative dispersion around the mean. Coefficient of variation is given by:

$$\mathrm{CV} = \frac{\sigma}{\mu} \tag{4}$$

where μ is the six-product or six-simulation mean. Note that the CV values for both the energy budget components and GPP can be very large in polar regions since the denominator is close to zero. As such, the CV is marked for only regions within a reasonable threshold (see Figure 4 and Section 3.2).

The global and regional mean variables of interest (and their subcomponents) were also linearly regressed against the k_d across the respective simulations to examine sensitivities of the variables to the inter-product k_d spread. Since the response of GPP to $K_{\downarrow,d}$ has been shown to be non-linear in past studies at the site level (Mercado et al., 2009; Zhou et al., 2021), we also used a logarithmic fit of the form:

$$y = a + b\log(x) \tag{5}$$

for tropical and temperate climate, where the GPP (and λE) response to $K_{\downarrow,d}$ is expected to be stronger. Here y is the variable of interest, x is forcing k_d and a and b are the model coefficients.

3. Results

3.1. Observational Evidence of the Diffuse Radiation Fertilization Effect at the Site Scale

To illustrate the dearth of observational constraints on the $K_{\downarrow,d}$ fertilization effect, we processed all the AmeriFlux (Novick et al., 2018) and FLUXNET (Baldocchi et al., 2001) site data with measurements of $K_{\downarrow,d}$. These data came from 12 FLUXNET sites and 36 AmeriFlux sites, with the majority located in evergreen needleleaf forests (16), deciduous broadleaf forests (9), and grasslands (9; Tables S1 and S2 in Supporting Information S1). Importantly, none of these sites are located in tropical rain forests, where the $K_{\downarrow,d}$ fertilization effect is expected to be the strongest (Chakraborty et al., 2021, Figure 2a).

The $K_{\downarrow,d}$ fertilization effect can be seen by identifying low (<0.35) and high (>0.65) k_d regimes and comparing GPP (or NEE) and λE during these two regimes. Almost all the sites show a clear $K_{\downarrow,d}$ fertilization effect, with λE and GPP being higher (NEE is lower) for the high k_d regime across bins and especially at high absorbed shortwave levels (Figure 2b, Tables S3–S6 in Supporting Information S1). These results are generally consistent with the commonly held hypothesis that plant photosynthesis, and thus transpiration, increase under diffuse conditions (Davin & Seneviratne, 2012; Mercado et al., 2009). Of note, the impacts of the $K_{\downarrow,d}$ fertilization effect is more clearly visible for the FLUXNET sites compared to the Ameriflux sites (Tables S3–S6 in Supporting

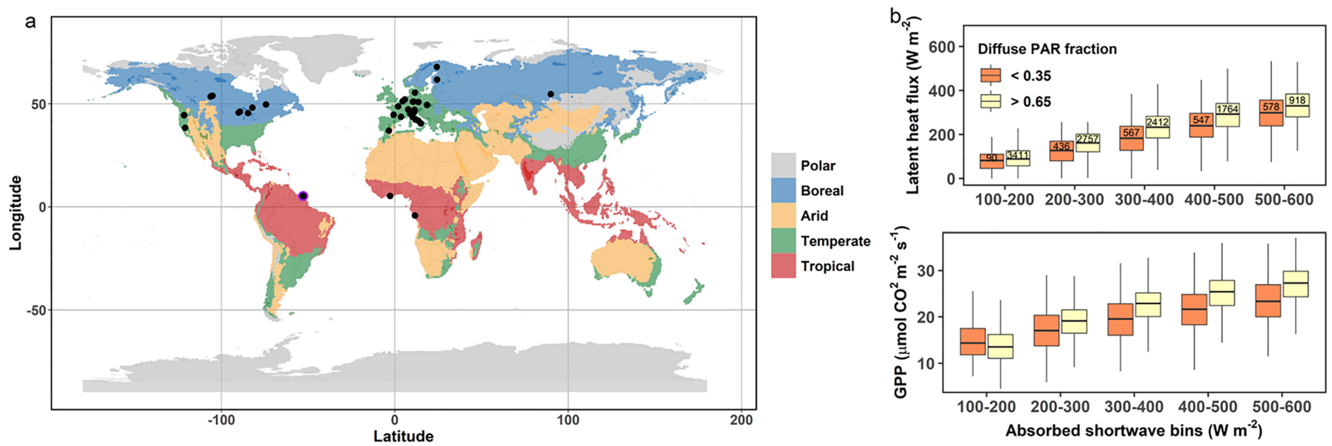

Figure 3. Photosynthetically active diffuse radiation measurements and associations with carbon and moisture flux. Sub-figure (a) shows the locations of the FLUXNET sites with simultaneous measurements of photosynthetically active diffuse radiation (PAR_d), carbon fluxes, and energy fluxes. The background colors represent the extent of the Koppen-Geiger climate zones. Sub-figure (b) illustrates the latent heat flux and gross primary productivity (GPP) for the Guyaflux FLUXNET site in French Guiana (location denoted by concentric purple circle in sub-figure (a)) for high and low regimes of diffuse PAR fraction in different absorbed shortwave radiation bins. The number of hourly observations in each bin is noted.

Information S1). The stronger $K_{\downarrow,d}$ fertilization signal in the FLUXNET sites may be partly because Ameriflux measurements are more intermittent and generally have far fewer available data points.

Although there are other flux tower networks throughout the world, including some in tropical forests (Restrepo-Coupe et al., 2021; Sarangi et al., 2022), few have continuous measurements of $K_{\downarrow,d}$ (Zhou et al., 2021). For instance, there are several (39) FLUXNET towers with measurements of PAR_d (Figure 3a), which is more relevant than total $K_{\downarrow,d}$ for GPP and λE (Mercado et al., 2009). However, only three of these sites are in the tropics. Among them, two did not have enough measurements to replicate the analysis in Figure 2b. The Guyaflux FLUXNET, which has also been examined in previous studies (O'Sullivan et al., 2021; Rap et al., 2015), also show higher λE and GPP for the high diffuse PAR fraction (PAR_d/PAR) regime in most absorbed shortwave bins (Figure 3b). Note that CLM, which is used in the subsequent sections to examine the $K_{\downarrow,d}$ fertilization effect at a global scale, is driven by total k_d, with the partitioning between PAR_d and NIR_d done internally by the model. The observational results presented here (Table S3–S6 in Supporting Information S1) are consistent with other existing site-based estimates (Davin & Seneviratne, 2012; Emmel et al., 2020; Ezhova et al., 2018; Sarangi et al., 2022; Wang et al., 2018; Yue & Unger, 2017) and demonstrate the $K_{\downarrow,d}$ fertilization effect at the site-scale. However, the tower site results cannot be used to provide global estimates due to both the sampling biases (e.g., lack of representation of tropical and other ecosystems) and lack of complete annual temporal coverage after quality-control, especially because the magnitude of the $K_{\downarrow,d}$ fertilization effect is relatively small compared to the uncertainties in the measured fluxes. Although in situ measurement networks have been used to create global estimates of the surface energy budget, these studies have access to many more observation stations with multi-year data and still find larger errors in tropical areas, including South America and Africa (Chakraborty & Lee, 2021a; Jung et al., 2019).

3.2. Global Spatial Distributions of Inter-Product Variability

Since models are frequently used to examine the $K_{\downarrow,d}$ fertilization effect to avoid the spatiotemporal sampling issues of in situ observations (Chakraborty et al., 2021; Mercado et al., 2009; Oliveira et al., 2011; O'Sullivan et al., 2021; Rap et al., 2018; Zhang, Ciais et al., 2021), we examine how simulated GPP and λE would vary for a realistic range of atmospheric k_d forcing. The meteorological forcing data are from the latest version of CAM, while the k_d is derived from monthly-climatology-adjusted current-generation data products, namely the NCEP/NCAR (Kistler et al., 2001), NOAA-CIRES-DOE (Slivinski et al., 2019), ERA5 (Hersbach et al., 2020), and MERRA-2 (Randles et al., 2017) reanalysis and the CERES (Rutan et al., 2015) product, as well as the default CAM outputs. Larger differences in k_d across these datasets are found in the mid-latitudes and high-latitudes, probably due to the higher baseline k_d in these regions (Figure 4a). We account for this difference in baseline

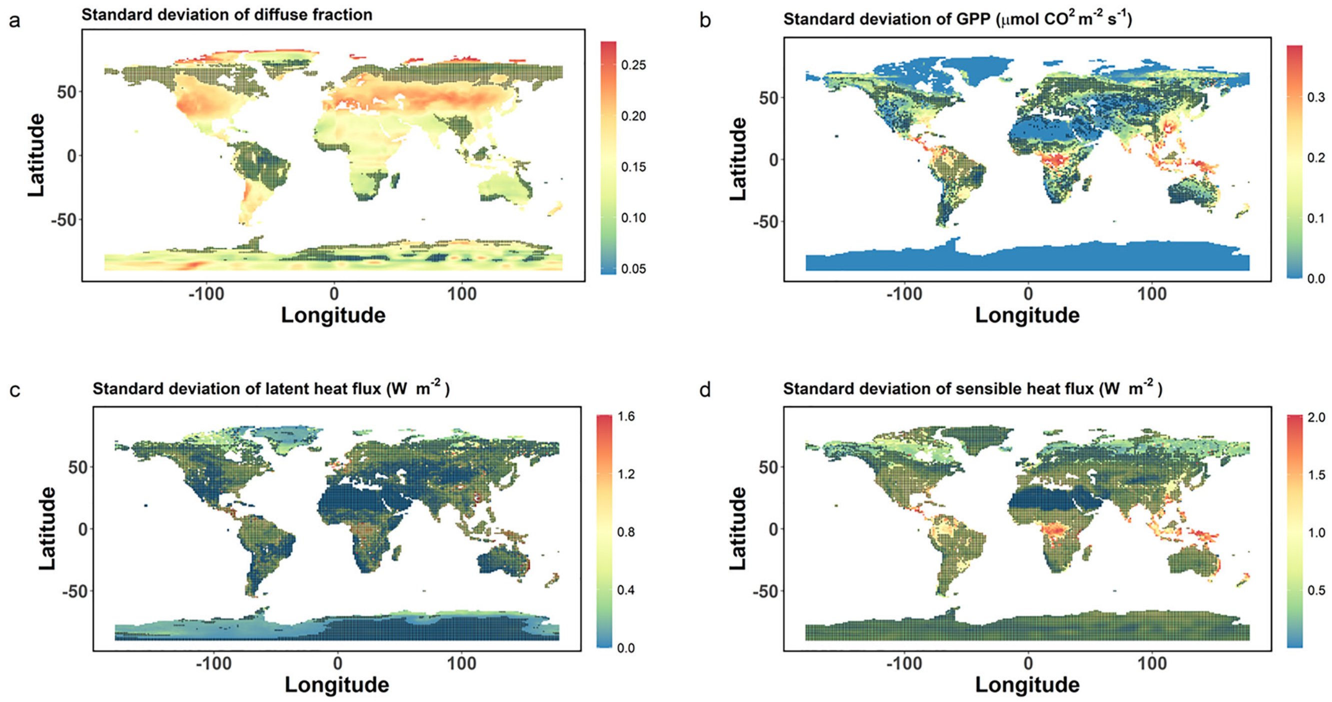

Figure 4. Spatial patterns of inter-product variability. Global distribution of the standard deviation in (a) diffuse fraction from the six products (NCEP/NCAR, NOAA-CIRES-DOE, ERA5, MERRA-2, CERES, and CAM) considered here and simulated (b) gross primary productivity (GPP), (c) latent heat flux, and (d) sensible heat flux from Community Land Model simulations that differ only in their diffuse fraction as defined by the six products. Grids with a coefficient of variation of less than 3% (<30% for diffuse fraction) are marked with + signs to represent regions with stronger agreement.

by also calculating the CV (regions where CV is less than 30% are marked with + signs in Figure 4). Most of the high latitudes fall within this zone, but the CV exceeds 30% for the rest of the Earth's surface, except for the Amazon and parts of eastern China. These forcing data, with all variables except for k_d being identical, are then used to run the latest version of CLM (Lawrence et al., 2019).

The standard deviation and CV in the simulated surface energy budget components (λE and sensible heat flux H) and GPP are lower than that for k_d (a CV threshold of only 3% is used for these). This is expected since the six simulations are forced with identical meteorological data, except for their k_d values, which provides a strong constraint on simulated GPP, λE, and H. Gross primary productivity shows the greatest variability (Figure 4b), with higher CV values seen over the Congo Basin, Southeastern US, and large parts of South and South-East Asia. Interestingly, even though the $K_{\downarrow,d}$ fertilization effect directly affects λE, there are regions with higher CV values for H (Figure 4d).

3.3. Impact of Diffuse Fraction Forcing on the Terrestrial Carbon and Energy Budget

The global mean k_d over land varies between 0.35 for MERRA-2 to 0.6 for CERES, with the true climatological mean expected to be around 0.42 based on the recent Bias-adjusted RADiation data set (BaRAD; Chakraborty & Lee, 2021a). A recent study found that the spread in both multi-year average $K_{\downarrow,d}$ and $K_{\downarrow,b}$ among global gridded datasets (other than NCEP/NCAR) were strongly correlated with the corresponding spread in cloud cover (Chakraborty & Lee, 2021a). Thus, the differences in k_d forcing may largely stem from dissimilar cloud parameterizations, since cloud cover is essentially a simulated, and not assimilated, variable in these products (Wright et al., 2020). Here, we demonstrate that the spread in simulated GPP is strongly associated with this inter-product k_d spread, not only globally but also for most climate zones (Figure 5). Among these, tropical and temperate areas show the greatest sensitivity of annual GPP to k_d (15.2 and 4.5 Pg C per unit change in k_d, respectively) and the polar region shows a weak relationship ($r^2 = 0.04$). The global GPP simulated by CLM using the default CAM forcing is close to upscaled FLUXNET-based estimates (118 Pg C year^{-1}; Jung et al., 2011), but varies from 114 Pg C year^{-1} when using MERRA-2 k_d as forcing to a ~7% higher value of 122 Pg C year^{-1} when

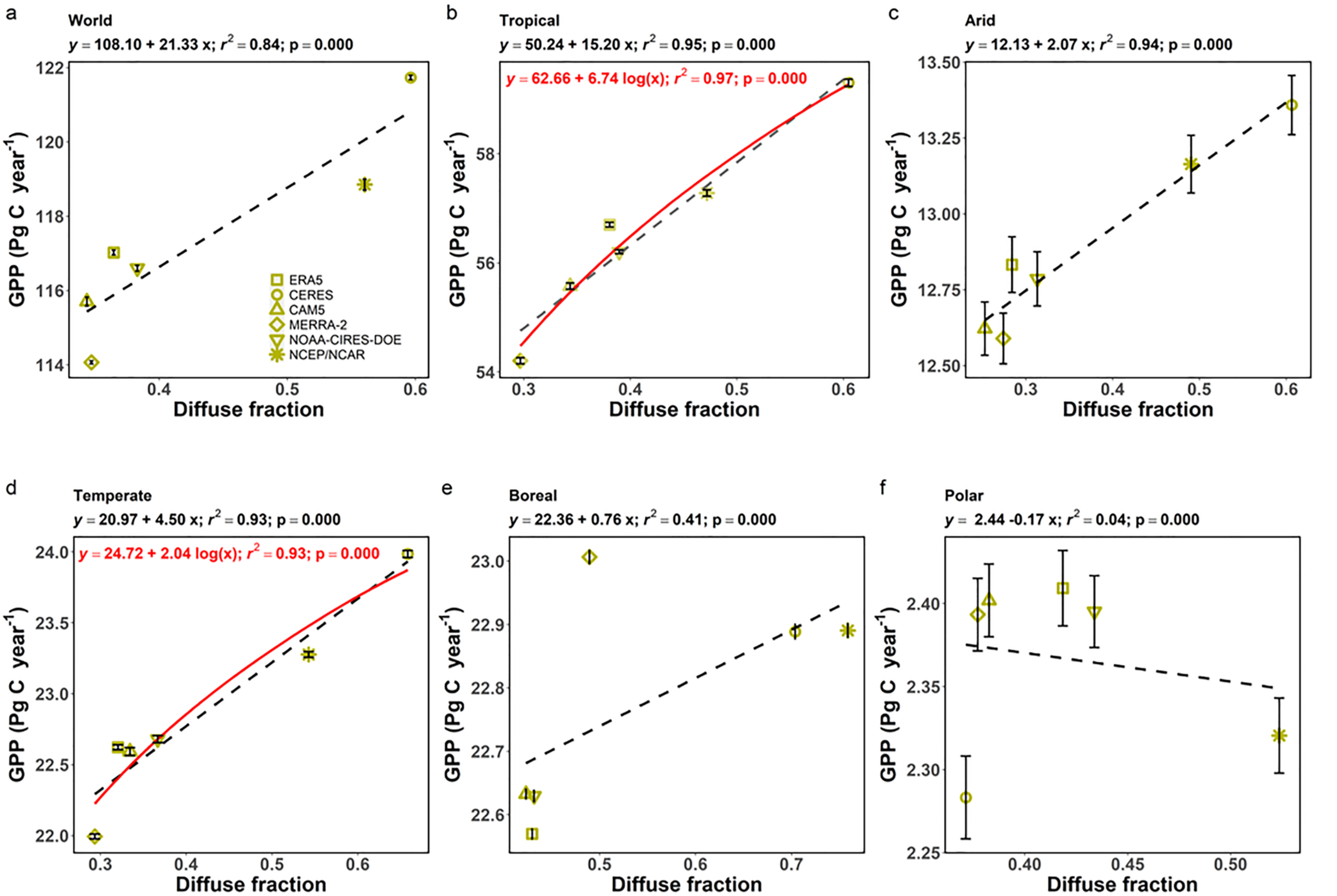

Figure 5. Response of gross primary productivity to inter-product diffuse fraction spread. Associations between gross primary productivity (GPP) and diffuse fraction (k_d) across different land model simulations forced using k_d from the six products (NCEP/NCAR, NOAA-CIRES-DOE, ERA5, MERRA-2, Clouds and the Earth's Radiant Energy System CERES, and CAM; represented using different symbols) considered here for (a) all terrestrial surfaces, (b) tropical climate, (c) arid climate, (d) temperate climate, (e) boreal climate, and (f) polar climate. The lines of best fit and the linear regression equations, with coefficient of determination r^2 and p-values are noted. For tropical and temperate climate, logarithmic fits and associated equations are also noted (in red). The vertical error bars show the inter-annual standard error for the 10-year period.

using CERES k_d. By comparison, Chen et al. (2017) found a standard deviation of global GPP across eight biome models biome models using the same climate forcing of 13 Pg C y^{-1}, with the inter-quartile range approaching 25 Pg C $year^{-1}$. The inter-product spread in GPP of ~8 Pg C $year^{-1}$ found here is also much higher than mean (from nine dynamic global vegetation models) global land carbon sink (−2.4 Pg C $year^{-1}$), a dominant source of uncertainty in our understanding of the carbon cycle (Sitch et al., 2015). The tropical annual GPP varies from 54 to 59 Pg C (9.3% higher) when switching from MERRA-2 to CERES k_d forcing. When examining the sensitivity of NEP and ER to the inter-product spread in k_d, similar positive correlations are seen for all cases other than for polar climate (Figures S1 and S2 in Supporting Information S1). Note that although site-level analyses have shown non-linear and somewhat asymptotic response of GPP to $K_{\downarrow,d}$ (Mercado et al., 2009; Zhou et al., 2021), when examining climate-zone-scale perturbations of GPP due to the inter-product k_d spread, the associations are practically linear, as illustrated by the comparisons with the logarithmic regressions for tropical and temperate climate (Figures 4b and 4d).

The sensitivities of the surface energy budget components to the inter-product k_d spread are generally weaker than that for GPP (Figure 6; Figures S3–S5 in Supporting Information S1). This is because the turbulent heat fluxes are more strongly constrained by available energy than GPP, as also seen in a previous study on aerosol-induced global dimming versus $K_{\downarrow,d}$ fertilization (Chakraborty et al., 2021). Globally, λE increases by only ~0.4% (from 37.24 to 37.38 W m^{-2}) and H decreases by ~3.0% (32.15–31.19 W m^{-2}) for the range of k_d considered. As such, the Bowen ratio ($\beta = H/\lambda E$) decreases globally and for all climate zones (Figs 6c; Figure S5 in Supporting Information S1).

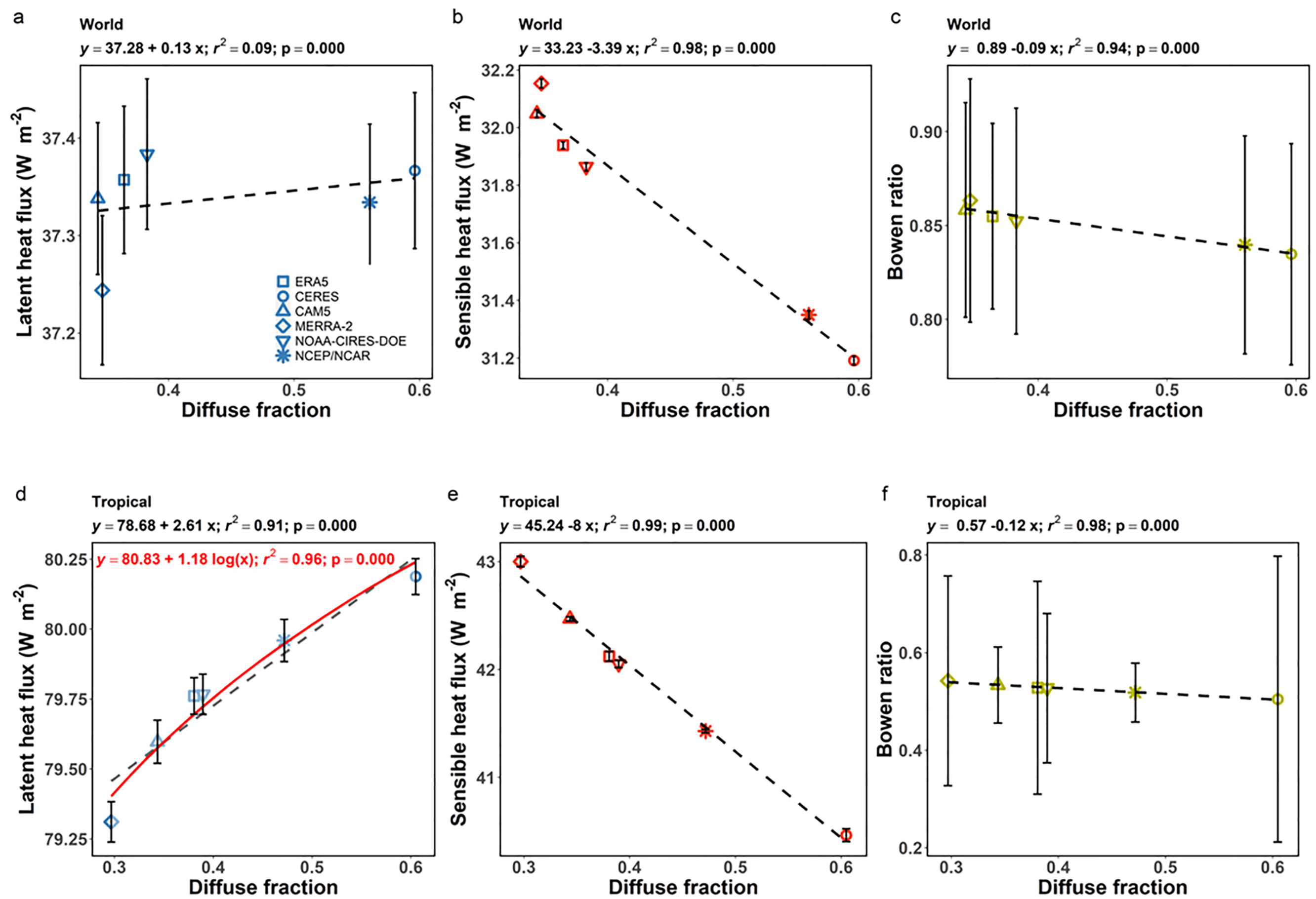

Figure 6. Response of surface energy budget components to inter-product diffuse fraction spread. Associations between (a) latent heat flux, (b) sensible heat flux, and (c) Bowen ratio and diffuse fraction (k_d) across different land model simulations forced using k_d from the six products (NCEP/NCAR, NOAA-CIRES-DOE, ERA5, MERRA-2, CERES, and CAM; represented using different symbols) considered here for all terrestrial surfaces. Sub-figures (d, e, and f) are similar, but for tropical climate. The lines of best fit and the linear regression equations, with coefficient of determination r^2 and p-values are noted. For tropical climate, a logarithmic fit and the associated equation is also noted for latent heat flux (in red). The vertical error bars show the inter-annual standard error for the 10-year period.

For tropical regions, the changes are slightly stronger, with λE increasing by ~1.1% (79.31–80.19 W m^{-2}) and H decreasing by ~5.9% (43–40.46 W m^{-2}). As the case with GPP, the improvements when using a logarithmic fit instead of a linear fit are marginal (r^2 increases from 0.91 to 0.96; Figure 6d; also see Figure S3b in Supporting Information S1 for temperate climate). The range of simulated λE and H due to different k_d forcing is smaller than the standard deviation across CMIP6 (3.5 W m^{-2} for λE and 2.7 W m^{-2} for H) and CMIP5 (3.9 W m^{-2} for λE and 2.6 W m^{-2} for H) models (Wild, 2020). To examine further, we separate λE and H into its sub-components. Globally and across most climate zones, the $\lambda E_{t,sha}$ and λE_c increased, while $\lambda E_{t,sun}$ and λE_g decreased (Figures S6 and S7 in Supporting Information S1; Figures 7 and 8). This increase in $\lambda E_{t,sha}$, due to additional illumination of the vegetation canopy under more diffuse conditions, is compensated by a decrease in $\lambda E_{t,sun}$, leading to minor increases in total λE. Global and regional decreases in H_g for the increasing k_d runs (around 5.5% globally; Figure S8a in Supporting Information S1) is only slightly compensated for by the increase in H_v (roughly 2% globally, but contrasting patterns across climate zones; Figure S9 in Supporting Information S1). This explains the larger spread in H (compared to λE) due to k_d forcing across the six simulations also seen in Figure 4d.

Even though both total $K_\downarrow$ and total $L_\downarrow$ are kept constant in all model simulations, there is a slight reduction in the sum of H and λE with the increase in k_d (from 69.39 to 68.57 W m^{-2} globally). This suggests that there were adjustments to the other components of the surface energy budget due to the change in k_d, which we examine using the six model simulations (Figure 9). Overall, we do see a reduction in net incoming radiation (Figure 9c) with increasing k_d (from 70.17 W m^{-2} when using MERRA-2 forcing to 69.32 W m^{-2} when using CERES forcing), which explains the overall reduction in the sum of turbulent heat fluxes. In parallel, the emitted longwave

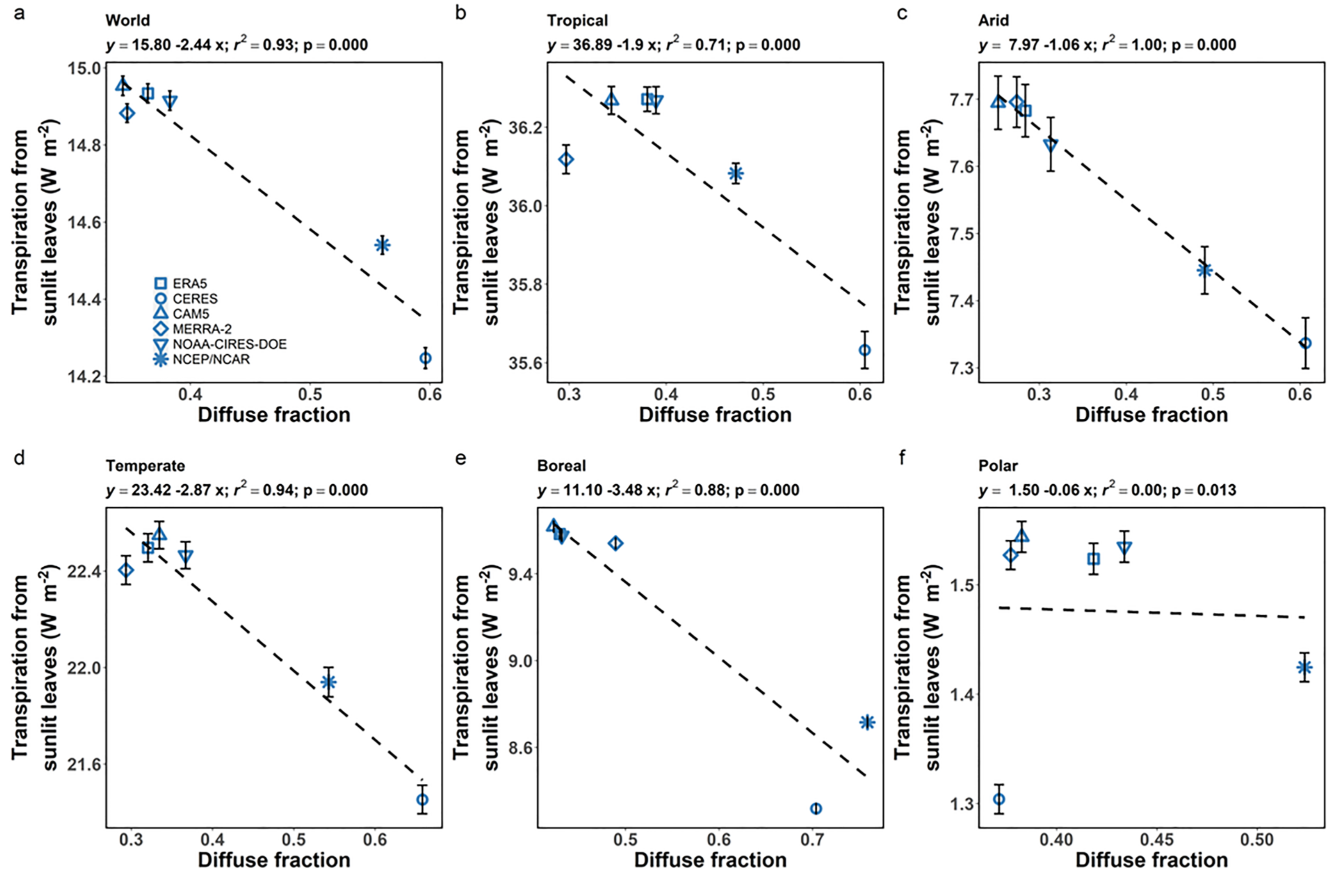

Figure 7. Response of transpiration from sunlit leaves to inter-product diffuse fraction spread. Associations between transpiration from sunlit leaves and diffuse fraction (k_d) across different land model simulations forced using k_d from the six products (NCEP/NCAR, NOAA-CIRES-DOE, ERA5, MERRA-2, CERES, and CAM; represented using different symbols) considered here for (a) all terrestrial surfaces, (b) tropical climate, (c) arid climate, (d) temperate climate, (e) boreal climate, and (f) polar climate. The lines of best fit and the linear regression equations, with coefficient of determination r^2 and p-values are noted. The vertical error bars show the inter-annual standard error for the 10-year period.

decreases (Figure 9b), partly related to surface cooling via evaporation due to the $K_{\downarrow,d}$ fertilization effect seen in Chakraborty et al. (2021), and the reflected shortwave increases (Figure 9a). The reflected shortwave increase is larger than the other changes in Figure 8 (from 49.08 to 50.19 W m^{-2}) and is dependent on surface albedo. Community Land Model uses separate albedos for $K_{\downarrow,d}$ and $K_{\downarrow,b}$, with an evident increase in all-sky albedo under higher k_d. Finally, there is an almost negligible decrease in ground heat flux with increasing k_d (Figure 9d).

4. Discussion

Since both $K_\downarrow$ and $K_{\downarrow,d}$ vary in these gridded products, we would expect the effect of variations in $K_\downarrow$ to overwhelm that of changes in $K_{\downarrow,d}$ (Chakraborty & Lee, 2019; Wild et al., 1998; Winter & Eltahir, 2010). The differences between datasets are also larger than perturbation signals seen for many individual atmospheric components (Chakraborty et al., 2021; Matsui et al., 2008; Oliveira et al., 2011; O'Sullivan et al., 2021). A couple of cases are discussed here. For eastern United States during the summer, Matsui et al. (2008) showed an average decrease in $K_\downarrow$ of 15.4 W m^{-2} and an increase in k_d by 3.48% for the 2000–2001 period on removing all aerosols. For the LSM used in that study, these aerosol-induced perturbations led to decreases in λE and H by over 2% and 11%, respectively. In comparison, the difference in annual average $K_\downarrow$ over the entire United States between CERES and NCEP/NCAR is 41.3 W m^{-2}, while the k_d varies from 0.24 in CERES to 0.45 in MERRA-2. Therefore, the effect of switching between gridded products of k_d to force an LSM will be potentially larger than the effect of removing all aerosols from the atmosphere. Oliveira et al. (2011) showed that for Europe and eastern United States, a roughly 7 W m^{-2} solar dimming between 1960 and 1990 decreased λE by 1.5 W m^{-2} and increased surface

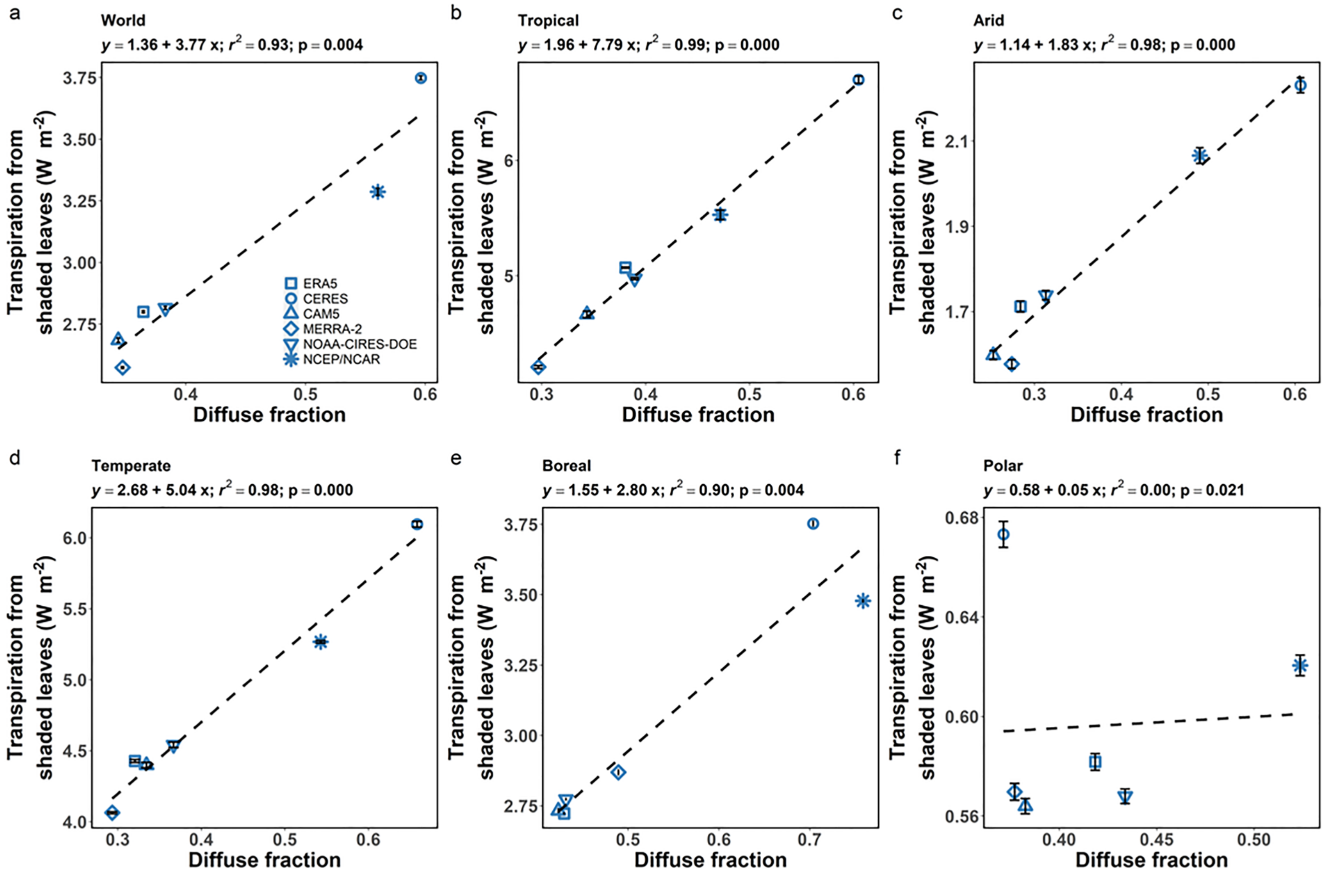

Figure 8. Response of transpiration from shaded leaves to inter-product diffuse fraction spread. Associations between transpiration from shaded leaves and diffuse fraction (k_d) across different land model simulations forced using k_d from the six products (NCEP/NCAR, National Oceanic and Atmospheric Administration, Cooperative Institute for Research in Environmental Science, and the Department of EnergyNOAA-CIRES-DOE, Fifth Generation Reanalysis (ERA5), Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2), Clouds and the Earth's Radiant Energy System (CERES), and CAM; represented using different symbols) considered here for (a) all terrestrial surfaces, (b) tropical climate, (c) arid climate, (d) temperate climate, (e) boreal climate, and (f) polar climate. The lines of best fit and the linear regression equations, with coefficient of determination r^2 and p-values are noted. The vertical error bars show the inter-annual standard error for the 10-year period.

runoff by ~5%. Similarly, the subsequent solar brightening between 1990 and 2004 of 6 W m^{-2} increased λE by 3 W m^{-2} and decreased surface runoff by 7% and 10% for the two regions. For the gridded products considered here, $K_\downarrow$ changes by 46.6 W m^{-2} between ERA5 and NCEP/NCAR for Europe and by 41.3 W m^{-2} over the United States, both perturbations being substantially larger than the temporal change in that study. Oliveira et al. (2011) also found that higher k_d (from 0.3 to 0.35) between 1960 and 1990 increased evapotranspiration in the tropics by 2.5 W m^{-2}. In comparison, the mean k_d over the tropical grids varies from ~0.30 when using MERRA-2-based forcing versus 0.6 for CERES-based forcing; 6 times that range.

Since the focus here is on the $K_{\downarrow,d}$ fertilization effect, we keep the total $K_\downarrow$ constant across model simulations to isolate the impact of changing k_d on carbon and energy fluxes. Gross primary productivity shows a stronger sensitivity to k_d than λE, which is in line with recent results for only the aerosol-induced changes in k_d (Chakraborty et al., 2021). Since we use a dynamic vegetation scheme with canopy state responding to the atmospheric forcing, we find that this sensitivity remains essentially the same for years 90–99 of the simulations compared to years 30–39 globally and across most climate zones (Figures 5 and 6, Figures S3, S10, and S11 in Supporting Information S1). For global land for instance, GPP increases by 6.1% (λE decreases by 0.35%) in years 30–39, versus +7% (GPP) and −0.37% (λE) for years 90–99 of the simulations. These small changes (less than a percent for GPP) over the roughly 80-year span suggest we should be cautious when linearly extrapolating the results from perturbation studies. For instance, taking the sensitivities from the feedback loop between increases in k_d due to emissions of Biogenic Volatile Organic Compounds and GPP enhancement proposed by Rap et al. (2018) and

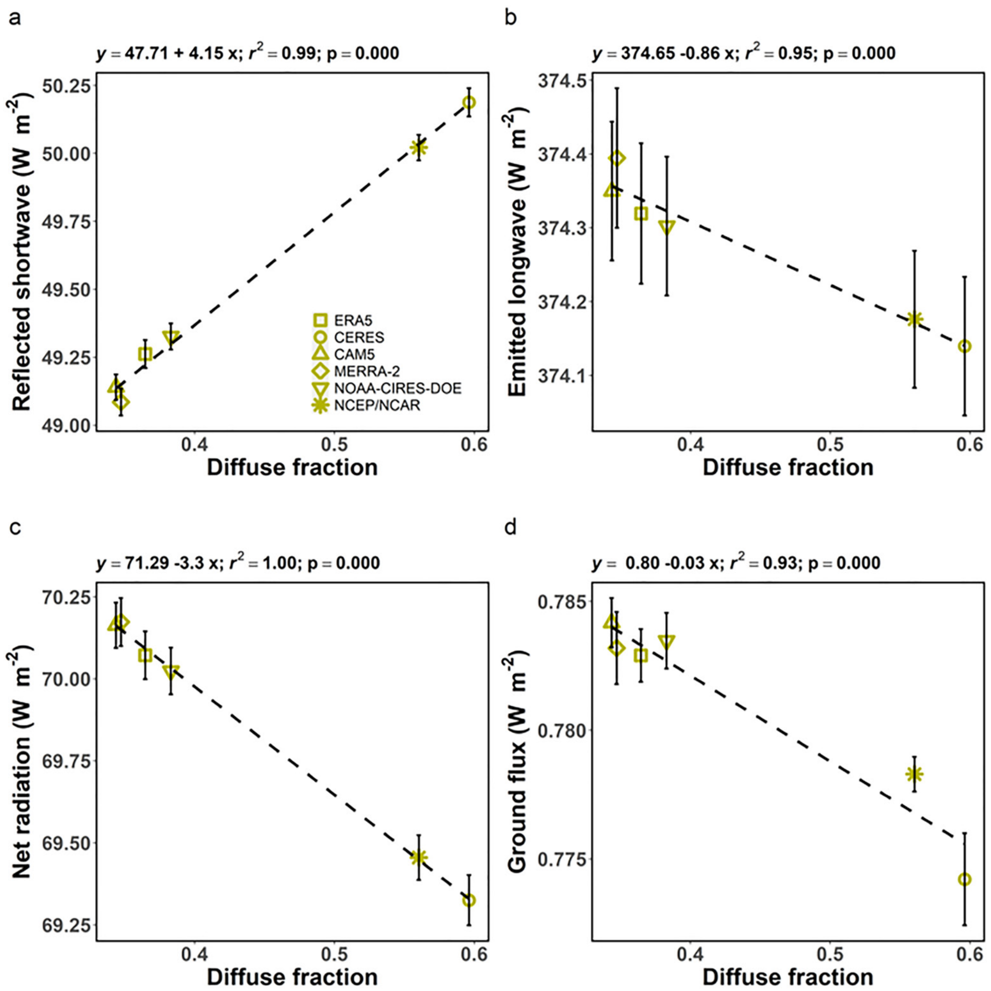

Figure 9. Response of other components of the surface energy budget to inter-product diffuse fraction spread. Associations between (a) reflected shortwave, (b) emitted longwave, (c) Bowen ratio and diffuse fraction (k_d) across different land model simulations forced using k_d from the six products (NCEP/NCAR, NOAA-CIRES-DOE, ERA5, Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2), Clouds and the Earth's Radiant Energy System (CERES), and CAM; represented using different symbols) considered here for all terrestrial surfaces. Sub-figure (d–f) are similar, but for tropical climate. The lines of best fit and the linear regression equations, with coefficient of determination r^2 and p-values are noted. For tropical climate, a logarithmic fit and the associated equation is also noted for latent heat flux (in red). The vertical error bars show the inter-annual standard error for the 10-year period.

implementing it between the total k_d values in MERRA-2 and CERES would yields a 5.7% increase in global terrestrial GPP due to the feedback alone. In reality, the actual changes would be mediated by other negative feedback loops (Rap, 2019; B. Wang et al., 2019). One such feedback is surface cooling (and thus GPP decrease) (Zhu et al., 2016), including cloud-induced cooling, with Ban-Weiss et al. (2011) showing a global surface temperature reduction of 0.54 K due to an increase in evaporative fraction (EF = $\lambda E/(\lambda E + H)$; by 0.014) via increased cloudiness. The change in EF when switching from MERRA-2 to CERES k_d forcing is 0.008; roughly half of that. Note however that these estimates of potential feedback (both in Rap et al., 2018 and here) are modeled and thus dependent on the accuracy with which the models can capture the response to $K_{\downarrow,d}$. For the summertime GPP simulated by the uncoupled multi-layer implementation of CLM, for instance, there is evidence that the response to $K_{\downarrow,d}$ is overestimated for a temperate deciduous forest site (Wozniak et al., 2020).

Although inter-model spread in $K_{\downarrow}$ has been examined across CMIP6 and CMIP5 models (Wild, 2020), similar analysis for $K_{\downarrow,d}$ (and thus k_d) are missing, partly because this variable is not always publicly archived. Although we do not expect the variability in k_d in current ESMs to be much larger than the range considered here, it is important to examine the spread across the radiative transfer modules used in CMIP6 models to identify potential reasons for the discrepancies. A bigger limitation is that we use a single LSM. Even with the same forcing data, different LSMs can show wide ranges in simulated carbon and moisture fluxes due to different implementations of model physics, land use representations, canopy architecture, presence or absence of dynamic vegetation, topography, etc (Bonan et al., 2021; Hao et al., 2021; Lawrence et al., 2016; Wild, 2020; Yao et al., 2014). However, CLM is a good starting point since different versions of it have been incorporated in multiple ESMs that are participating in CMIP6 (Chakraborty et al., 2021). Given the large response of the terrestrial GPP and evapotranspiration to the inter-product spread in k_d forcing seen here, it is critical to systematically examine these sensitivities across land modules in current ESMs. In parallel, we need to gather more simultaneous observations of $K_{\downarrow,d}$ and carbon and energy fluxes to benchmark model performance and improve our understanding of this aspect of atmosphere-biosphere interactions. Of particular interest are the tropics, where current measurements are insufficient and the largest responses to k_d forcing are seen. Given the distinct treatment of the $K_{\downarrow,d}$ fertilization effect across LSMs from various model lineages (Chakraborty et al., 2021), the participation of different modeling groups using a standardized approach to compare the response to $K_{\downarrow,d}$ across models would be an effective contribution to CMIP7.

5. Conclusions

Clouds, aerosols, and the carbon budget are large sources of uncertainty in our understanding of the Earth system and how it will change in the future. The diffuse radiation fertilization effect links these three components and remains a relatively understudied aspect of atmosphere-biosphere interactions with global estimates relying on model simulations. Here we first demonstrate the sampling bias in existing flux tower networks to observationally constrain this effect and then examine the impact of a realistic spread in diffuse fraction forcing, derived from global gridded products, on components and subcomponents of the terrestrial carbon and energy budgets simulated by the latest version of the CLM. Large differences are seen in gross primary productivity (GPP; around ~7% globally) for this inter-product spread with larger differences (~9%) in tropical regions. Overall, simulated GPP due to inter-product diffuse fraction spread in CLM is roughly a third of the inter-quartile GPP spread seen previously across biome models. Changes in terrestrial evapotranspiration are smaller due to contrasting changes in shaded and sunlit leaf transpiration but greater than regional impacts of individual forcing agents. No current Model Intercomparison Project, whether focusing on the atmosphere or the biosphere, explicitly accounts for the diffuse radiation or its impacts. Our results demonstrate the importance of systematically examining the simulated diffuse radiation by atmosphere modules and response to the same in land modules across ESMs. Doing so can identify potential deficiencies in current-generation models, inform future model development, and better constrain land carbon uptake and its potential feedback in future climate change assessments.

Acknowledgments

PNNL is operated for the Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830. The authors are also grateful to Computational and Information Systems Lab at NCAR for the computing grants necessary to run the model simulations and to Yale Center for Earth Observation and Yale Center for Research Computing for additional computational resources. The CAM and CLM models were supported primarily by the National Science Foundation (NSF). The data analysis was supported in part by NSF through grant AGS1933630 (to X.L.) and by travel funding from U.S. Department of Energy and Yale Institute for Biospheric Studies (to T.C.). This material is based upon work supported by NCAR, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. D.M.L. was supported in part by the RUBISCO Scientific Focus Area, which is sponsored by the Regional and Global Climate Modeling Program in the Climate and Environmental Sciences Division of the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science. The authors thank Keith Oleson at the National Center for Atmospheric Research (NCAR) for suggestions on the CLM model modifications and Lina Mercado at the University of Exeter for inputs on sources of diffuse radiation measurements in the tropics.

Data Availability Statement

The Community Earth System Model is a public domain software and its releases are accessible through this GitHub repository: https://github.com/ESCOMP/CESM. The CERES data were obtained from the NASA Langley Research Center CERES ordering tool (https://ceres.larc.nasa.gov/). The NOAA–CIRES–DOE and NCEP–NCAR reanalysis datasets were downloaded from the PSL website (https://psl.noaa.gov/). The MERRA-2 reanalysis data set can be found on NASA's website (https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/). The ERA5 reanalysis data were downloaded from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/). Other data sets used and generated for this study are available from the authors upon request.

References

Arias, P., Bellouin, N., Coppola, E., Jones, R., Krinner, G., Marotzke, J., et al. (2021). Climate change 2021: The physical science basis. contribution of working group14 I to the sixth assessment report of the intergovernmental panel on climate change; Technical Summary.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., et al. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*(11), 2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:fantts>2.3.co;2

Ban-Weiss, G. A., Bala, G., Cao, L., Pongratz, J., & Caldeira, K. (2011). Climate forcing and response to idealized changes in surface latent and sensible heat. *Environmental Research Letters*, *6*(3), 034032. https://doi.org/10.1088/1748-9326/6/3/034032

Bonan, G. B., Patton, E. G., Finnigan, J. J., Baldocchi, D. D., & Harman, I. N. (2021). Moving beyond the incorrect but useful paradigm: Reevaluating big-leaf and multilayer plant canopies to model biosphere-atmosphere fluxes–a review. *Agricultural and Forest Meteorology*, *306*, 108435. https://doi.org/10.1016/j.agrformet.2021.108435

Chakraborty, T., & Lee, X. (2019). Land cover regulates the spatial variability of temperature response to the direct radiative effect of aerosols. *Geophysical Research Letters*, *46*(15), 8995–9003. https://doi.org/10.1029/2019gl083812

Chakraborty, T., & Lee, X. (2021a). Large differences in diffuse solar radiation among current-generation reanalysis and satellite-derived products. *Journal of Climate*, 1–52. https://doi.org/10.1175/JCLI-D-20-0979.1

Chakraborty, T., & Lee, X. (2021b). Using supervised learning to develop BaRAD, a 40-year monthly bias-adjusted global gridded radiation dataset. *Scientific Data*, *8*(1), 1–10. https://doi.org/10.1038/s41597-021-01016-4

Chakraborty, T., Lee, X., & Lawrence, D. M. (2021). Strong local evaporative cooling over land due to atmospheric aerosols. *Journal of Advances in Modeling Earth Systems*, *13*(5). https://doi.org/10.1029/2021MS002491

Chen, M., Rafique, R., Asrar, G. R., Bond-Lamberty, B., Ciais, P., Zhao, F., et al. (2017). Regional contribution to variability and trends of global gross primary productivity. *Environmental Research Letters*, *12*(10), 105005. https://doi.org/10.1088/1748-9326/aa8978

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., et al. (2011). The Joint UK Land Environment Simulator (JULES), model description–Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, *4*(3), 701–722. https://doi.org/10.5194/gmd-4-701-2011

Davin, E. L., & Seneviratne, S. I. (2012). Role of land surface processes and diffuse/direct radiation partitioning in simulating the European climate. *Biogeosciences*, *9*(5), 1695–1707. https://doi.org/10.5194/bg-9-1695-2012

Emmel, C., D'Odorico, P., Revill, A., Hörtnagl, L., Ammann, C., Buchmann, N., & Eugster, W. (2020). Canopy photosynthesis of six major arable crops is enhanced under diffuse light due to canopy architecture. *Global Change Biology*, *26*(9), 5164–5177. https://doi.org/10.1111/gcb.15226

Ezhova, E., Ylivinkka, I., Kuusk, J., Komsaare, K., Vana, M., Krasnova, A., et al. (2018). Direct effect of aerosols on solar radiation and gross primary production in boreal and hemiboreal forests. *Atmospheric Chemistry and Physics*, *18*(24), 17863–17881. https://doi.org/10.5194/acp-18-17863-2018

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, *27*(2), 511–526. https://doi.org/10.1175/JCLI-D-12-00579.1

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., et al. (2019). High climate sensitivity in the community Earth system model version 2 (CESM2). *Geophysical Research Letters*, *46*(14), 8329–8337. https://doi.org/10.1029/2019gl083978

Hao, D., Bisht, G., Gu, Y., Lee, W.-L., Liou, K.-N., & Leung, L. R. (2021). A parameterization of sub-grid topographical effects on solar radiation in the E3SM Land Model (version 1.0): Implementation and evaluation over the Tibetan Plateau. *Geoscientific Model Development*, *14*(10), 6273–6289. https://doi.org/10.5194/gmd-14-6273-2021

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). *The ERA5 global reanalysis*. Quarterly Journal of the Royal Meteorological Society.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., et al. (2016). C4MIP – The coupled climate–carbon cycle model intercomparison project:experimental protocol for CMIP6. *Geoscientific Model Development*, *9*(8), 2853–2880. https://doi.org/10.5194/gmd-9-2853-2016

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, *6*(1), 74. https://doi.org/10.1038/s41597-019-0076-8

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., et al. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research*, *116*(G3), G00J07. https://doi.org/10.1029/2010jg001566

Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., et al. (2001). The ncep–ncar 50-year reanalysis: Monthly means cd-rom and documentation. *Bulletin of the American Meteorological Society*, *82*(2), 247–268. https://doi.org/10.1175/1520-0477(2001)082<0247:tnnyrm>2.3.co;2

Knohl, A., & Baldocchi, D. (2008). Effects of diffuse radiation on canopy gas exchange processes in a forest ecosystem. *Journal of Geophysical Research*, *113*(G2). https://doi.org/10.1029/2007jg000663

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., et al. (2010). Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: Critical issues and global evaluation. *Global Change Biology*, *16*(1), 187–208. https://doi.org/10.1111/j.1365-2486.2009.02041.x

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018ms001583

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., et al. (2016). The land use model intercomparison project (LUMIP) contribution to cmip6: Rationale and experimental design. *Geoscientific Model Development*, *9*, 2973–2998. https://doi.org/10.5194/gmd-9-2973-2016

Luo, X., Chen, J. M., Liu, J., Black, T. A., Croft, H., Staebler, R., et al. (2018). Comparison of big-leaf, two-big-leaf, and two-leaf upscaling schemes for evapotranspiration estimation using coupled carbon-water modeling. *Journal of Geophysical Research: Biogeosciences*, *123*(1), 207–225. https://doi.org/10.1002/2017jg003978

Matsui, T., Beltran-Przekurat, A., Niyogi, D., Pielke, R. A., & Coughenour, M. (2008). Aerosol light scattering effect on terrestrial plant productivity and energy fluxes over the eastern United States. *Journal of Geophysical Research*, *113*(D14), D14S14. https://doi.org/10.1029/2007jd009658

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M., & Cox, P. M. (2009). Impact of changes in diffuse radiation on the global land carbon sink. *Nature*, *458*(7241), 1014–1017. https://doi.org/10.1038/nature07949

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., et al. (2010). Description of the NCAR community atmosphere model (CAM 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, *1*(1), 1–12.

Novick, K. A., Biederman, J. A., Desai, A. R., Litvak, M. E., Moore, D. J., Scott, R. L., & Torn, M. S. (2018). The AmeriFlux network: A coalition of the willing. *Agricultural and Forest Meteorology*, *249*, 444–456. https://doi.org/10.1016/j.agrformet.2017.10.009

O'Sullivan, M., Zhang, Y., Bellouin, N., Harris, I., Mercado, L. M., Sitch, S., et al. (2021). Aerosol–light interactions reduce the carbon budget imbalance. *Environmental Research Letters*, *16*(12), 124072. https://doi.org/10.1088/1748-9326/ac3b77

Oleson, K., Lawrence, D. M., Gordon, B., Flanner, M. G., Kluzek, E., Peter, J., et al. (2013). Technical description of version 4.5 of the community land model (CLM) (No. NCAR/TN).503.

Oliveira, P. J. C., Davin, E. L., Levis, S., & Seneviratne, S. I. (2011). Vegetation-mediated impacts of trends in global radiation on land hydrology: A global sensitivity study. *Global Change Biology*, *17*(11), 3453–3467. https://doi.org/10.1111/j.1365-2486.2011.02506.x

Pincus, R., Forster, P. M., & Stevens, B. (2016). The radiative forcing model intercomparison project (RFMIP): Experimental protocol for cmip6. *Geoscientific Model Development*, *9*(9), 3447–3460. https://doi.org/10.5194/gmd-9-3447-2016

Randles, C., da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., et al. (2017). The merra-2 aerosol reanalysis, 1980 onward. Part i: System description and data assimilation evaluation. *Journal of Climate*, *30*(17), 6823–6850. https://doi.org/10.1175/jcli-d-16-0609.1

Rap, A., Scott, C. E., Reddington, C. L., Mercado, L., Ellis, R. J., Garraway, S., et al. (2018). Enhanced global primary production by biogenic aerosol via diffuse radiation fertilization. *Nature Geoscience*, *11*(9), 640–644. https://doi.org/10.1038/s41561-018-0208-3

Rap, A., Scott, C. E., Reddington, C. L., Mercado, L., Ellis, R. J., Garraway, S., et al. (2019). Reply to: Complexities between plants and the atmosphere. *Nature Geoscience*, *12*(9), 695. https://doi.org/10.1038/s41561-019-0437-0

Rap, A., Spracklen, D. V., Mercado, L., Reddington, C. L., Haywood, J. M., Ellis, R. J., et al. (2015). Fires increase Amazon forest productivity through increases in diffuse radiation. *Geophysical Research Letters*, *42*(11), 4654–4662. https://doi.org/10.1002/2015gl063719

Restrepo-Coupe, N., da Rocha, H. R., Hutyra, L. R., de Araujo, A. C., Borma, L. S., Christoffersen, B., et al. (2021). Large scale biosphere-atmosphere experiment (LBA-ECO)LBA-ECO CD-32 flux tower network data compilation, Brazilian Amazon: 1999-2006, V2 [CSV]. 0 MB. https://doi.org/10.3334/ORNLDAAC/1842

Rubel, F., & Kottek, M. (2010). Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift*, *19*(2), 135–141. https://doi.org/10.1127/0941-2948/2010/0430

Rutan, D. A., Kato, S., Doelling, D. R., Rose, F. G., Nguyen, L. T., Caldwell, T. E., & Loeb, N. G. (2015). Ceres synoptic product: Methodology and validation of surface radiant flux. *Journal of Atmospheric and Oceanic Technology*, *32*(6), 1121–1143. https://doi.org/10.1175/jtech-d-14-00165.1

Sarangi, C., Chakraborty, T., Tripathi, S., Krishnan, M., Morrison, R., Evans, J., & Mercado, L. M. (2022). Observations of aerosol–vapor pressure deficit–evaporative fraction coupling over India. *Atmospheric Chemistry and Physics*, *22*(5), 3615–3629. https://doi.org/10.5194/acp-22-3615-2022

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., et al. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, *12*(3), 653–679. https://doi.org/10.5194/bg-12-653-2015

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., et al. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the twentieth century reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876–2908. https://doi.org/10.1002/qj.3598

Steiner, A. L., Mermelstein, D., Cheng, S. J., Twine, T. E., & Oliphant, A. (2013). Observed impact of atmospheric aerosols on the surface energy budget. *Earth Interactions*, *17*, 14–22. https://doi.org/10.1175/2013ei000523.1

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., et al. (2016). LS3MIP (v1.0) contribution to CMIP6: The land surface, Snow and soilmoisture model intercomparison project – Aims, setup and expected outcome. *Geoscientific Model Development*, *9*(8), 2809–2832. https://doi.org/10.5194/gmd-9-2809-2016

Wang, B., Shugart, H. H., & Lerdau, M. T. (2019). Complexities between plants and the atmosphere. *Nature Geoscience*, *12*(9), 693–694. https://doi.org/10.1038/s41561-019-0413-8

Wang, X., Wu, J., Chen, M., Xu, X., Wang, Z., Wang, B., et al. (2018). Field evidences for the positive effects of aerosols on tree growth. *Global Change Biology*, *24*(10), 4983–4992. https://doi.org/10.1111/gcb.14339

Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., et al. (2017). The cloud feedback model intercomparison project (CFMIP) contribution to CMIP6. *Geoscientific Model Development*, *10*(1), 359–384. https://doi.org/10.5194/gmd-10-359-2017

Wild, M. (2020). The global energy balance as represented in CMIP6 climate models. *Climate Dynamics*, *55*(3–4), 553–577. https://doi.org/10.1007/s00382-020-05282-7

Wild, M., Ohmura, A., Gilgen, H., & Morcrette, J.-J. (1998). The distribution of solar energy at the Earth's surface as calculated in the ECMWF re-analysis. *Geophysical Research Letters*, *25*(23), 4373–4376. https://doi.org/10.1029/1998gl900175

Winter, J. M., & Eltahir, E. A. (2010). The sensitivity of latent heat flux to changes in the radiative forcing: A framework for comparing models and observations. *Journal of Climate*, *23*(9), 2345–2356. https://doi.org/10.1175/2009jcli3158.1

Wozniak, M. C., Bonan, G. B., Keppel-Aleks, G., & Steiner, A. L. (2020). Influence of vertical heterogeneities in the canopy microenvironment on interannual variability of carbon uptake in temperate deciduous forests. *Journal of Geophysical Research: Biogeosciences*, *125*, 8. https://doi.org/10.1029/2020jg005658

Wright, J. S., Sun, X., Konopka, P., Krüger, K., Legras, B., Molod, A. M., et al. (2020). Differences in tropical high clouds among reanalyses: Origins and radiative impacts. *Atmospheric Chemistry and Physics*, *20*(14), 8989–9030. https://doi.org/10.5194/acp-20-8989-2020

Yao, Y., Liang, S., Li, X., Hong, Y., Fisher, J. B., Zhang, N., et al. (2014). Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations. *Journal of Geophysical Research: Atmospheres*, *119*(8), 4521–4545. https://doi.org/10.1002/2013jd020864

Yue, X., & Unger, N. (2017). Aerosol optical depth thresholds as a tool to assess diffuse radiation fertilization of the land carbon uptake in China. *Atmospheric Chemistry and Physics*, *17*(2), 1329–1342. https://doi.org/10.5194/acp-17-1329-2017

Zhang, Y., Bastos, A., Maignan, F., Goll, D., Boucher, O., Li, L., et al. (2020). Modeling the impacts of diffuse light fraction on photosynthesis in ORCHIDEE (v5453) land surface model. *Geoscientific Model Development*, *13*(11), 5401–5423. https://doi.org/10.5194/gmd-13-5401-2020

Zhang, Y., Boucher, O., Ciais, P., Li, L., & Bellouin, N. (2021a). How to reconstruct aerosol-induced diffuse radiation scenario for simulating GPP in land surface models? An evaluation of reconstruction methods with ORCHIDEE_DFv1. 0_DFforc. *Geoscientific Model Development*, *14*(4), 2029–2039. https://doi.org/10.5194/gmd-14-2029-2021

Zhang, Y., Ciais, P., Boucher, O., Maignan, F., Bastos, A., Goll, D., et al. (2021b). Disentangling the impacts of anthropogenic aerosols on terrestrial carbon cycle during 1850–2014. *Earth's Future*, *9*(7). https://doi.org/10.1029/2021EF002035

Zhou, H., Yue, X., Lei, Y., Zhang, T., Tian, C., Ma, Y., & Cao, Y. (2021). Responses of gross primary productivity to diffuse radiation at global FLUXNET sites. *Atmospheric Environment*, *244*, 117905. https://doi.org/10.1016/j.atmosenv.2020.117905

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., et al. (2016). Greening of the Earth and its drivers. *Nature Climate Change*, *6*(8), 791–795. https://doi.org/10.1038/nclimate3004