Global fire emissions buffered by the production of pyrogenic carbon

Authors:
Matthew W. Jones\textsuperscript{1,2*}, Cristina Santín\textsuperscript{1,3}, Guido R. van der Werf\textsuperscript{4} and Stefan H. Doerr\textsuperscript{1}

Affiliations:
\textsuperscript{1}Geography Department, College of Science, Swansea University, Swansea, UK
\textsuperscript{2}Tyndall Centre for Climate Change Research, University of East Anglia, Norwich, UK
\textsuperscript{3}Biosciences Department, College of Science, Swansea University, Swansea, UK
\textsuperscript{4}Faculty of Science, Vrije Universiteit, Amsterdam, Netherlands
\textsuperscript{*}e-mail: matthew.w.jones@swansea.ac.uk.

Landscape fires burn 3-5 million km\textsuperscript{2} of the Earth’s surface annually. They emit 2.2 Pg carbon per year to the atmosphere while also converting a significant fraction of the burned vegetation biomass to pyrogenic carbon. Pyrogenic carbon can be stored in terrestrial and marine pools for centuries to millennia and therefore its production can be considered a mechanism for long-term carbon sequestration. Pyrogenic carbon stocks and dynamics are not considered in global carbon cycle models, leading to systematic errors in carbon accounting. Here we present a comprehensive dataset of pyrogenic carbon production factors from field and experimental fires and merge this with the Global Fire Emissions Database to quantify the global pyrogenic carbon production flux. We find that 256\textsuperscript{+84}_{-60} Tg of biomass carbon was converted annually to pyrogenic carbon between 1997-2016. Our central estimate equates to 12% of the annual carbon emitted globally by landscape fires, indicating that their emissions are buffered by PyC production. We further estimate that cumulative pyrogenic carbon production was 60 Pg since 1750, or 33-40% of the global biomass carbon lost through land use change in this period. Our results demonstrate
that pyrogenic carbon production by landscape fires could be a significant but overlooked sink for atmospheric CO$_2$.

Globally, landscape fires including wildfires, deforestation fires, and agricultural burns emit approximately 2.2 Pg C year$^{-1}$ to the atmosphere (1997-2016)$^1$. This emission flux includes $\sim$0.4 Pg C year$^{-1}$ due to tropical deforestation and peatland fires, which contribute to net global emissions of carbon due to land use change ($\sim$1.1-1.5 Pg C year$^{-1}$; Figure 1)$^2$-$^4$. The emission fluxes resulting from biomass fires and land use change are outweighed by the re-sequestration flux of carbon to undisturbed and re-growing vegetation ($\sim$5.1 Pg C year$^{-1}$; Figure 1)$^5$-$^8$. Meanwhile, carbon fluxes resulting from non-deforestation fire emissions and related vegetation re-growth are approximately balanced, meaning that these fires have no net influence on atmospheric carbon on decadal timescales$^9$-$^{10}$. These global carbon budget estimates are generated by models that represent the temporally distinct processes of immediate carbon emission from burned areas and decadal-scale sequestration through vegetation (re-)growth in a spatially explicit manner$^1$-$^{11}$-$^{12}$. However, such models routinely overlook the coincident flux of biomass carbon to recalcitrant by-products of fire, which can be stored in terrestrial and marine pools for centuries to millennia, and thus provide a long-term buffer against fire emissions (Figure 1)$^7$-$^{13}$-$^{16}$. Consequently, the legacy effects of fire that operate on the longest timescales are systematically excluded from models of the carbon cycle and from global carbon budgets$^{15}$-$^{17}$.

These legacy effects are due to the incomplete combustion of vegetation during landscape fires, which transforms part of the remaining organic carbon (OC) in biomass to a continuum of thermally-altered products that are collectively termed pyrogenic carbon (PyC)$^{13}$-$^{15}$-$^{18}$. The majority of the PyC produced during landscape fires remains initially on the ground in charcoal particles of varying size and is subsequently transferred to its major global stores in soils$^{19}$-$^{21}$, sediments$^{22}$-$^{23}$ and water bodies$^{24}$-$^{25}$. A smaller fraction of fire-affected
vegetation carbon is emitted as PyC in smoke\textsuperscript{26,27}. PyC includes labile products of depolymerisation reactions as well as aromatic molecules that result from condensation reactions, the latter of which are depleted in functional groups and thus chemically and biologically recalcitrant\textsuperscript{28–30}. The enhanced resistance of PyC to biotic and abiotic decomposition leads to its preferential storage in environmental pools\textsuperscript{16,23} and a residence time that is typically 1-3 orders of magnitude greater than that of its unburnt precursors\textsuperscript{15}. This makes PyC one of the largest groups of chemically discernible compounds in soil with a contribution to soil organic carbon (SOC) stocks of 14\% globally\textsuperscript{19}. A fraction of PyC is also conserved across the land-to-ocean aquatic continuum and thus contributes approximately 10\% of riverine dissolved organic carbon\textsuperscript{31}, 16\% of riverine particulate organic carbon\textsuperscript{32}, and 10-30\% of the organic carbon in ocean sediments\textsuperscript{16,22,33,34}.

A series of reviews and data syntheses have recognised the potential of PyC production to invoke a drawdown (sink) of photosynthetically-sequestered CO\textsubscript{2} to pools that are stable on timescales relevant to anthropogenic climate change and its mitigation\textsuperscript{7,13,15,16,38–43}. Owing to the relative recalcitrance of PyC, the conversion of biomass carbon to PyC represents an extraction of carbon from a pool cycling on decadal timescales to a pool cycling on centennial or millennial timescales\textsuperscript{16,22,23,28,44}. This storage potential contrasts with that of dead vegetation, which degrades on timescales of months to decades or enters soil pools with a shorter residence time than that of PyC\textsuperscript{11,14,28,45,46}. Consequently, post-fire PyC pools emit carbon to the atmosphere over a significantly longer time period than would be the case in the absence of PyC production, meanwhile providing a buffer that moderates atmospheric CO\textsubscript{2} stocks (Figure 1)\textsuperscript{7,15,16}. At present, the fire-enabled vegetation models that are used to make global carbon budget calculations account for short-term fire emissions but routinely exclude fluxes of carbon from biomass to PyC or the delayed emission of carbon from legacy PyC stocks to the atmosphere (Figure 1)\textsuperscript{11,12,17,47,48}. This
introduces systematic errors to global carbon budgets through misrepresentation of modern and historical fire effects on the exchange of carbon between the atmosphere and terrestrial-marine pools\textsuperscript{15–17}.

While PyC has been recognised as a major component of post-fire ecosystem carbon stocks for a number of decades\textsuperscript{13,41}, quantification of its production rate at the global scale has been problematic and estimates vary by roughly an order of magnitude (50-379 Tg C year\textsuperscript{-1})\textsuperscript{15,16,40,42}. A cause of the large range of production estimates is that calculations have previously relied on incomplete information regarding the spatial distribution and type of fires, the allocation of carbon amongst biomass fuel components in burned areas and the specific PyC production factors for these distinct biomass fuel components. To alleviate these issues, we enhanced the Global Fire Emissions Database version 4 with small fires (GFED4s)\textsuperscript{1}, which is one of the principal process-based models used to make estimates of carbon emission from landscape fires\textsuperscript{47,49,50}. Specifically, PyC production was incorporated by following a three-step approach consisting of: (i) the assembly of the most comprehensive global database of PyC production factors (P\textsubscript{PyC}; g PyC g\textsuperscript{-1} C emitted) compiled to date; (ii) the assignment of production factors for individual fuel classes stratified as coarse or fine and as woody or non-woody (Figure 2), and; (iii) the application of production factor (P\textsubscript{PyC}) values to fuel-stratified carbon emissions (CE; g C emitted) modelled by the native fuel consumption model in GFED4s. The output is the first global gridded dataset for monthly PyC production at a resolution of 0.25° × 0.25°, covering the years 1997-2016.

**Global PyC Production**

Our central estimate for global PyC production in the period 1997-2016 was 256 Tg C year\textsuperscript{-1} (Figure 3), with an uncertainty range of 196-340 Tg C year\textsuperscript{-1} (which includes variability in measured P\textsubscript{PyC} and inter-annual variability in global production, but excludes uncertainty
in GFED4s emissions estimates; see methods). Inter-annual variability in global PyC production, expressed as the standard deviation around the mean, was 47 Tg C year\(^{-1}\) and was most strongly associated with variability in woody fuel combustion, including standing wood and coarse woody debris (CWD; supplementary text S1 and Figure S1). Coarse woody fuels produce PyC at a greater rate than finer fuels (Figure 2) and consequently forest fires have disproportionate potential to influence global rates of PyC production (supplementary figure S2).

The El Niño-Southern Oscillation (ENSO) is the primary driver of inter-annual variability in burned area in the tropics\(^5\) and previous analyses conducted with GFED have shown that carbon emissions from tropical forest ecosystems more than doubled on average during positive (El Niño) phases relative to negative (La Niña) ENSO phases\(^5\). Correspondingly, we calculated that global rates of PyC production in tropical forests were 111% greater during the main fire season of El Niño phases than La Niña phases (supplementary Table S1). As rates of PyC production by non-forest fires were not sensitive to ENSO (supplementary Table S1), the major driver of inter-annual variability in total PyC production was variability in tropical forest burned area (Figure 3). The production of PyC was anomalously high in 1997-1998 (366 Tg C year\(^{-1}\)) aligning with a particularly strong positive El Niño phase which promoted extensive burning of (tropical) forests in South and Central America and in Southeast and Equatorial Asia\(^1,5\).

**Major Production Regions**

The PyC production rates modelled by GFED4s+PyC conformed to a latitudinal pattern (Figure 4), with the tropical latitudes clearly dominating production at the global scale. 91% of global production occurred in the tropics and subtropics (0-30° N/S), while temperate
(30-60° N/S) and high-latitude regions (60-90° N) provided small contributions to the global total (8% and 1%, respectively).

The global distribution of PyC production also showed intricate regional patterns driven by variation in both the frequency at which fuel stocks were exposed to fire and the magnitude of the fuel stocks that were combusted during the fires that occurred (supplementary Figures S3 and S4). Fire frequency was ultimately the key determinant of PyC production rate and this explains why the tropics and subtropics were the dominant source regions. Although savannah fires affect low fuel stocks (supplementary text S2), these fires occur frequently and were spatially extensive (supplementary Figure S5 and table S2). They thus made the largest contribution to the global PyC production flux (125 Tg C year⁻¹). Although tropical deforestation fires affected approximately 1% of the area of savannah fires, they affected large stocks of fuel (supplementary table S2) and were thus the second largest driver of global PyC production, contributing 49 Tg C year⁻¹. The area affected by non-deforestation tropical forest fires was more than a factor of 4 larger than that of deforestation fires, however, fuel consumption was relatively low (supplementary table S2). These fires provided the third major component of the global PyC production flux (34 Tg C year⁻¹). Overall, 81% of total global PyC production in the period 1997-2016 occurred in savannahs (49%) and tropical forests (32%).

**Global Carbon Budget Implications**

Here we have quantified the global gross sink of atmospheric carbon caused by the transfer of photosynthetically-sequestered biomass carbon to stocks of PyC during vegetation fires. Our central global PyC production flux estimate (256 Tg C year⁻¹) is nontrivial within the context of the global carbon cycle (Figure 1), equating to 12% of the global carbon emissions flux due to biomass burning and ~8% of the land sink for atmospheric CO₂ (~3.0-
3.2 Tg C year$^{-1}$). The global PyC production flux also equates to 75% of the carbon emitted from tropical deforestation and peat fires, which are the main categories of fire that cause a net loss of carbon to the atmosphere$^{1,7,53}$. The PyC flux modelled here occurs in addition to the smaller global flux of 2 Tg C year$^{-1}$ caused by the emission of PyC in smoke from vegetation fires (according to equivalent estimates made using GFED4s in the years 1997-2016$^1$).

The magnitude of our global estimate for PyC production indicates that the production of PyC during vegetation fires has the potential to significantly influence the atmospheric stock of carbon. A net sink of atmospheric carbon to stocks of PyC can be expected to develop if the flux associated with its production is unmatched by re-mineralisation fluxes from legacy PyC stocks in terrestrial-marine pools (Figure 1). Earth System Models (ESMs) are the most sophisticated tools available to quantify the exchange of carbon between the atmosphere and these pools in time periods for which robust empirical data is sparse or unavailable. Despite foregoing attempts to highlight the importance of PyC production for carbon storage over timescales relevant to anthropogenic climate change and its mitigation$^{40,41,54}$, the absence of the PyC cycle from ESMs has restricted the scope for quantifying its role in the carbon cycle$^{17}$. The method introduced here allows for the routine integration of PyC production into fire-enabled vegetation models in a manner that systematically considers the spatial distribution of fire, the composition of the fuel stocks affected and the specific PyC production factors that apply to individual fuel components. This procedure would be simple to implement in other fire-enabled vegetation models, meaning that the major outstanding challenge to quantifying the net exchange of carbon between the atmosphere and PyC stocks with ESMs will be to improve constraints over its storage and residence time in terrestrial and marine pools (Figure 1)$^{16,17}$. 

We also show that the PyC cycle must be integrated into ESMs if they are to accurately represent the role of fire in Earth’s carbon cycle. The production flux of PyC represents the quantity of carbon that models would otherwise treat either as emitted or as unburned biomass with a residence time in terrestrial pools on the order of months to decades\textsuperscript{11,14,28,45,46,55}. At present, the fate of 11\% of the global biomass carbon stocks affected annually by fire is misrepresented in global models. As PyC dynamics are not represented in the ESMs used to make global carbon budget calculations\textsuperscript{2}, this pool may represent a quantitatively significant missing sink or source of carbon to the atmosphere\textsuperscript{17,56}. Recent estimates suggest that total carbon emissions from biomass burning in the period 1750-2015 amounted to ~500 Pg C (averaging 1.9 Pg C year\textsuperscript{-1})\textsuperscript{47}. Under the assumption that the modern global PyC production flux maintained a constant ratio with the carbon emissions flux throughout this period, we estimate that ~60 Pg C was transferred to PyC stocks since the beginning of the industrial revolution. This value is equivalent to 33-40\% of the carbon lost from biomass pools due to land use change in the same time period (145-180 Pg C)\textsuperscript{4,57}.

Our estimates for modern and historical PyC production incorporate the best of current understanding of PyC production through the combustion of vegetation biomass; however, the limitations of these estimates are worthy of mention. Notably, we do not include the production of PyC through the combustion of organic matter in soils, which may be an important process driving the accumulation of PyC stocks in environments with deep organic layers, particularly peatlands\textsuperscript{58}. We also do not account for the re-combustion of PyC in locations that experience secondary burns, which can drive losses of the PyC that remains exposed at the surface\textsuperscript{60}. PyC mass losses through re-combustion have been reported as <8\% in savannas\textsuperscript{59} and 17-84\% in Boreal forests\textsuperscript{60,61}, however the long fire return intervals in the latter biome typically allow sufficient time for PyC to be protected from re-combustion through its burial in soils\textsuperscript{20}. Our exclusion of re-combustion is deliberate as we consider the
process to be a component of the legacy PyC decomposition flux, which we do not quantify here (Figure 1). Finally, we note that our dataset of PyC production factors cannot provide values for $P_{PyC}$ that are modulated both by fuel class, as in this study (Figure 2), and by the ecosystem properties (e.g. vegetation density) and fire characteristics (e.g. temperature and duration) that are relevant to the formation of PyC$^{42,62,63}$. Continued study of PyC production, with a particular focus on regions with high or rising fire incidence$^{64-66}$ and a range of fire intensities$^{67}$, will facilitate the application of more specific production factors in spatially-explicit global models and thus result in reduced uncertainties in global PyC production.

The production of PyC may become an increasingly important process for global carbon cycling in future centuries. Although global burned area has declined in at least the past two decades due predominantly to the conversion of savannah and grassland to agriculture$^{68,69}$, recent fire modelling studies generally agree that this decline is unlikely to continue past the year 2050$^{64-66}$. It is also likely that a higher fraction of global burned area will be distributed in forests where significant stocks of vegetation carbon are held$^{64,70,71}$. As woody fuels generate more PyC per unit of biomass carbon than other fuels (Figure 2), the spread of fire into forests can be expected to disproportionately enhance global PyC production (supplementary Figure S2). Although it is less clear how fire prevalence will change in tropical and temperate forests owing to a stronger human control over burning in these regions$^{64,68}$, recent increases in fire extent caused by increasing drought frequency in Amazonia are already counteracting reductions in the extent of deforestation fires$^{72}$. Notwithstanding the significant uncertainty that exists in model predictions of future fire regimes, there are strong indications that PyC production rates will increase in some of the Earth’s most carbon-dense regions in response to a changing climate$^{7,11,73}$. This implies that the buffer for atmospheric CO$_2$ emissions resulting from PyC production will grow in future centuries.
References


41. Wei, X., Hayes, D. J., Fraver, S. & Chen, G. Global Pyrogenic Carbon Production During Recent Decades Has Created the Potential for a Large, Long-Term Sink of


Acknowledgements

This work was funded by a Leverhulme Trust Research Project Grant awarded to SD (RPG-2014-095); a Swansea University College of Science Fund awarded to MJ; a Vici grant awarded to GW by the Netherlands organisation for scientific research (NWO), and a European Union Horizon 2020 research and innovation grant awarded to CS (Marie Skłodowska-Curie grant 663830). We thank Cristina Aponte, Claudia Boot, Gareth Clay, Garry Cook, Francesca Cotrufo, Philip Fearnside, Brett Goforth, Robert Graham, Michelle Haddix, Peter Homann, Dale Hurst and Meaghan Jenkins for their assistance during the collation of the global dataset of PyC production factors. We also thank Bill de Groot for his part in securing funding of the Leverhulme Trust Grant.
Author Contributions

MJ, CS and SD designed the study. SD led the Leverhulme Trust Research Project grant that funded the main body of the work. MJ collated the PyC production factor dataset with support from CS. CS and SD provided unpublished PyC production data. GW provided access to the GFED4s code. MJ adapted the GFED4s code to include PyC production with the support of GW. MJ conducted the formal analysis of the production factor dataset and model outputs. All authors contributed to the interpretation of the results. MJ wrote the manuscript text and produced all figures. All authors contributed to the refinement of the manuscript text.

Data Availability

The global dataset of PyC production factors is available as supplementary data file (GlobalPyC_supplementarydataset.xls). This dataset will also be uploaded to the GFED website and updated with new data as it becomes available (http://www.globalfiredata.org). Supplementary text S4 contains full reference to the studies included in the production factor dataset. Burned area and fire emissions data are publicly available at the GFED website (http://www.globalfiredata.org). Additional ancillary data are available from the corresponding author on request.

Materials & Correspondence

Correspondence and material requests should be addressed to MJ.

Financial and non-financial competing interests

The authors declare no competing interests.
**Figure Captions**

**Figure 1:** A schematic of the global carbon cycle including the buffer and legacy roles of PyC. Stock values are expressed in Pg C (1 Pg C = 1 \times 10^{15} \text{ g of carbon}) and flux values are expressed in Pg C year\(^{-1}\). Stocks and fluxes of the global carbon cycle are represented by values from the Global Carbon Budget (GCB) assessment of the decade 2008–2017 (ref. 2) and the IPCC AR5 assessment of the decade 2000-2009 (ref. 4). Fluxes of carbon due to the net land sink are modified from the GCB to exclude non-deforestation fire emissions, while net land use change emissions are modified to exclude deforestation fire emissions. Carbon emissions from deforestation and peat fires and from non-deforestation fires were derived from GFED4s (ref. 1) and relate to the period 1997-2016. PyC production fluxes due to deforestation and non-deforestation fires are based on estimates from GFED4s+PyC (this study). PyC stocks in soils, ocean DOC and ocean sediments are based on representative PyC/OC ratios from references 19, 35, and 16 applied to the estimates of OC stocks and fluxes. PyC fluxes through rivers are the sum of global dissolved and particulate PyC export fluxes (refs. 31 and 32). Residence times shown for soils derive from a meta-analysis of PyC decomposition in space-for-time substitution studies\(^{36}\) and incubation experiment estimates extrapolated to field conditions\(^{28}\). Residence times for oceanic PyC pools derive from references 22 and 37. First-order estimates for legacy PyC decomposition fluxes and their uncertainties are calculated in quadrature for land and ocean pools as the product of PyC stocks and the reciprocal of the residence times for PyC in these pools, assuming that the low- and high- end estimates for each term represent a consistent portion of normally-distributed uncertainty.

**Figure 2:** Box plots showing the distributions of PyC production factor (P\text{PYC}) values for each of the biomass component classes in the production factor dataset. Abbreviations are: CWAGF, coarse woody aboveground fuels; CWSF, coarse woody surface fuels; FWAGF,
fine woody aboveground fuels; FWSF, fine woody surface fuels; NWAGF, non-woody aboveground fuels; NWSF, non-woody surface fuels; CWF, coarse woody fuels (includes both CWSF and CWAGF); FWF, fine woody fuels (includes both FWAGF and FWSF); NWF, non-woody fuels (includes both NWAGF and NWSF). Dots mark the distribution of $P_{PyC}$ values across 1% intervals on the y-axis. Red dots show mean $P_{PyC}$ values while red lines show the bootstrapped 95% confidence interval (see methods). Boxes illustrate the median and interquartile range of values. Letters a and b indicate biomass components with statistically similar $P_{PyC}$ distributions at the 95% confidence level according to Tukey HSD tests. The number of data entries (n) is also shown.

**Figure 3:** Annual global PyC production estimates from GFED4s+PyC. The black line plots the modelled rate of production based on central $P_{PyC}$ ratios (g PyC g$^{-1}$ C emitted) from the global dataset. The shaded area indicates the uncertainty range of modelled values based on the 95% confidence intervals of $P_{PyC}$ values (see Figure 2). The contributions of savannah burning (red line) and tropical forest burning (green line) to global PyC production totals are shown, the latter of which includes deforestation fires (green dashed line).

**Figure 4:** Annual average PyC production rates for the period 1997-2016 from GFED4s+PyC, based on central production factors (see Figure 2). (a) The global distribution of PyC production expressed in g C m$^{-2}$ year$^{-1}$. (b) The total production of PyC (Tg C year$^{-1}$) in 15° latitudinal bands segregated according to the fire type, including: savannah fires (SAVA); non-deforestation tropical forest fires (TROF); tropical deforestation fires (DEFO); agricultural fires (AGRI); temperate forest fires (TEMF); extratropical grassland fires (EXGR), and; boreal forest fires (BORF).
Methods

Global Fuel Consumption Modelling in GFED4s

In GFED4s, carbon emissions to the atmosphere are quantified based on burned area and fuel consumption per unit burned area. Burned area is derived from satellite and fires that are too small to be detected by regular burned area algorithms are derived statistically based on active fire detections and relations with, amongst others, vegetation indices. Fuel consumption is modelled using a satellite-driven biogeochemical model and tuned to match observations. Most of the underlying satellite input datasets have a 500 × 500 m resolution but are aggregated to the model resolution of 0.25° × 0.25°. Total fuel consumption is based on fuel consumption of several fuel components including leaves, grasses, litter, fine woody debris, coarse woody debris (CWD), and standing wood. For more information on the GFED4s modelling approach, the reader is directed to van der Werf et al. (ref. 1).

To calculate PyC production within GFED4s we added a production factor, \( P_{\text{PYC}} \), which quantifies the production of PyC per unit carbon emitted (g PyC g\(^{-1}\) C emitted). Until now, the principle obstacle to performing a global modelling exercise of this type has been the lack of a sufficiently rich and standardised dataset with which to constrain representative values for \( P_{\text{PYC}} \). The remainder of this section details how representative PyC production factors were collated and summarised and subsequently integrated into the fuel consumption model of GFED4s.

Our estimates of uncertainty in annual PyC production relate only to variability in PyC production factors and inter-annual variability in emissions and do not include uncertainty in carbon emission estimates propagating from GFED4s. Uncertainties in GFED4s emissions estimates are discussed at length in refs. 1 and 77 and are predominantly the result of uncertainties in the satellite detection of small fires using thermal anomalies and burn scars.
As carbon emissions and PyC production are co-dependent on burned area, estimation errors relating to fire detection introduce scalar uncertainties. Uncertainty in fuel consumption is an additional component of the overall uncertainty in GFED4s emission estimates and has been reduced from previous versions (e.g. GFED3) through its incorporation of a global dataset of fuel consumption estimates. As discussed in the primary literature relating to the development of the GFED4s (ref. 1), a formal global-scale assessment of the uncertainties in fuel consumption cannot be completed due to a paucity of ground truth data for some input datasets. For the previous version of GFED (GFED3), Monte Carlo simulations accounting for uncertainty in both burned area detection and fuel consumption were used to obtain first-order constraints on the uncertainty in carbon emissions, which were ±20-25% at global, annual scales as a 1 standard deviation (1σ) value. Developments of GFED4s included the incorporation of small fire burned area detection, which led to important reductions in negative bias in emissions estimates; however, small fires are also challenging to detect and a lack of validation data prevents formal investigation of uncertainty in burned area for GFED4s. Hence, the true uncertainty of GFED4s is not known precisely but is likely to be on the same order as GFED3 (1σ = ±20-25%). Nonetheless, uncertainty ranges are likely to be greater in regions where small fires are prevalent or where organic soils are affected (e.g. Central America, Europe and Equatorial Asia).

Regional-scale field studies of fire emissions have served to validate that the GFED modelling framework produces reliable estimates at large scales, for example in Alaska and the tropics. Studies that involve atmospheric tracers have also provided vital diagnostics of the performance of GFED, generally highlighting its proficiency at large scales but revealing some weaknesses in specific regions or during isolated events. Overall, GFED4s is highly suited to the investigation of the effects of fire in global-scale biogeochemical cycles.
and is thus regularly used in global carbon budget assessments and as a reference point for the fire modules of ESMs.

**Collating a Global Dataset of PyC Production Factors**

We compiled a new database of $P_{PYC}$ factors (supplementary dataset) from a global collection of 22 published studies which reported on PyC production in 91 burn units, as well as two new datasets produced by the authors with 23 burn units reported for the first time here, and standardised their reporting. All studies used one of the following two broad approaches to quantify the impacts of fire on the biomass carbon stocks, either: pre-fire and post-fire stocks of biomass carbon and PyC are measured, or; space-for-time substitution is used to constrain burned and unburned stocks of biomass carbon and PyC, which are assumed to be equivalent to pre-fire and post-fire stocks, respectively. Hereafter, the terms “pre-fire” and “post-fire” are used to refer to both types of assessment. Here we focus only on PyC present in charcoal and ash on the ground following fire as well as on charred vegetation. PyC emitted with smoke, transported in the atmosphere and deposited on regional scales area is not included as this process has been studied in separate dedicated studies conducted by atmospheric scientists and represents a relatively small flux in comparison (see main text).

The $P_{PYC}$ values were calculated for each of six classes of widely used biomass components: coarse woody surface fuels (CWSF), including coarse woody debris or downed wood defined by typical diameter thresholds of >7.6 cm or >10 cm; fine woody surface fuels (FWSF), including fine woody debris or any other woody debris with diameters below the thresholds for CWSF; coarse woody aboveground fuels (CWAGF), including trees or branches with diameters greater than the thresholds for CWSF; fine woody aboveground fuels (FWAGF), including material described as shrubs, trees or branches with diameters...
below the thresholds for CWSF; non-woody surface fuels (NWSF), including litter, understory vegetation, grass, root mat and any other form of non-woody material directly in contact with the ground surface, and finally; non-woody aboveground fuels (NWAGF), including foliage, leaves, needles, crown fuels and any other form of non-woody material that attaches to standing wood structures above the ground surface.

For each biomass component, \( P_{\text{PyC}} \) was calculated using the following equation (1):

\[
P_{\text{PyC}} = \frac{C_{\text{Py}}}{C_{\text{PRE}} - C_{\text{POST}}}
\]

where \( C_{\text{Py}} \) is the mass of PyC created during the fire that was attributed to the component, \( C_{\text{PRE}} \) was the pre-fire stock of biomass carbon in the component, and \( C_{\text{POST}} \) was the post-fire stock of biomass carbon in the unburnt component. \( C_{\text{Py}}, C_{\text{PRE}} \) and \( C_{\text{POST}} \) were all expressed in the units g C km\(^{-2}\).

Criteria were applied as filters to the dataset in order to ensure that \( P_{\text{PyC}} \) could be calculated in a consistent and representative manner. Specifically, \( P_{\text{PyC}} \) was calculated if the following conditions were met: first, both pre-fire and post-fire biomass stocks were reported and carbon content (%) was either measured or assumed based on representative values from the literature; second, post-fire stocks of pyrogenic organic matter (charcoal, ash and the charred components of partially-affected vegetation) were reported and their PyC content (%) was either measured or assumed based on representative values from the literature; third, the type of fire that occurred was representative of a widespread regional fire type (e.g. wildfires, slash-and-burn deforestation, and prescribed fire); fourth, in experimental fires, the biomass carbon stock was designed to replicate the density and structure of biomass carbon stocks observed in the field and the burning efficiency was not optimised or adapted as a factor of the study design.
The set of criteria outlined above does not exclude studies that assess the PyC content of charcoal using one of the various chemical or thermochemical techniques available for the separation of pyrogenic carbon from bulk OC\textsuperscript{90,91}. Such techniques are frequently used for the detection of PyC in well-mixed soil, sediment and aquatic matrices. However, we note that none of the studies included in our dataset utilised a chemical or thermochemical approach to separate PyC from non-PyC; instead, these studies consider all carbon in residual products of interest (charcoal, ash and the charred components of partially-affected vegetation) to be PyC. Thus, we highlight that our estimates of $P_{\text{PyC}}$ are free of the inter-method variability in PyC quantification that often confounds the comparison of PyC concentration in environmental matrices across studies and contributes to the notable uncertainty in the magnitude of Earth’s major PyC stocks\textsuperscript{15,16} (Figure 1).

Like biomass carbon, total PyC stocks are distributed across several components including charcoal and ash on the ground, charcoal attached to coarse woody debris, and charcoal attached to aboveground vegetation\textsuperscript{15}. The majority of the studies included in the production factor dataset matched the studied PyC components to individual biomass carbon components from which they were known to derive. However, as some individual components of PyC stocks can have a mixture of sources that are indistinguishable from their location or appearance alone, it was occasionally necessary to make assumptions about the biomass components that were sources of these components. This was done on a study-by-study basis. In cases where the source of each PyC component was not explicitly stated, the following procedural steps were adhered to. On a first basis, the PyC component was assigned to a biomass component according to the most probable source inferred, but not explicitly stated, in the primary literature. Second, where more than one biomass component was inferred to be a source of the PyC stock in the primary literature, the PyC stock was weighted proportionally to the pre-fire stock of carbon present in each of the implicated
biomass components. Otherwise, if no sources of PyC were inferred in the primary literature it was necessary to make independent assumptions about the source of PyC in a manner that was consistent with the other studies included in the dataset and our collective experience of quantifying PyC production in the field.

Summarising Production Factor Values for use in GFED4s+PyC

Our global database suggested that coarse woody surface fuels (CWSF) and aboveground fuels (CWAGF) produce significantly more PyC, relative to carbon emitted, than other fuel classes ($P_{\text{PYC}}$ averaged 0.25 and 0.31 g PyC g$^{-1}$ C emitted, respectively; Figure 2). In contrast, the mean $P_{\text{PYC}}$ values for fine woody surface fuels (FWSF) and fine woody aboveground fuels (FWAGF; 0.12 and 0.076 g PyC g$^{-1}$ C emitted, respectively) did not differ significantly from those of non-woody surface fuels (NWSF) or non-woody aboveground fuels (NWAGF; 0.099 and 0.062 g PyC g$^{-1}$ C emitted, respectively). These results are consistent with previous studies, which suggest that large-diameter woody fuels burn less completely and produce PyC in greater proportions than finer fuels$^{40,92}$.

For each class, the mean PyC production factor was used as the central estimate for $P_{\text{PYC}}$, while the confidence interval around the mean $P_{\text{PYC}}$ was calculated through a bootstrapping procedure. Specifically, the available PyC production factors from the dataset were resampled 50,000 times, the mean $P_{\text{PYC}}$ was calculated for each resample, and the 95% confidence interval was calculated as the middle 95% of the observed 50,000 means (i.e. those ranked 1,250$^{th}$ to 48,750$^{th}$).

According to analysis of variance (ANOVA) with a Tukey Honest Significant Difference post-hoc test, no significant differences in mean $P_{\text{PYC}}$ were observed between the distributions of $P_{\text{PYC}}$ for coarse, fine, and non-woody fuels positioned at the ground surface and those same fuels located above the ground surface. Therefore, the $P_{\text{PYC}}$ values applied
in GFED4s+PyC were based on the distribution of values in three simplified fuel classes (Figure 2): coarse woody fuels (CWF: mean 0.26 g PyC g\(^{-1}\) C; 95% confidence interval 0.18-0.39 g PyC g\(^{-1}\) C), fine woody fuels (FWF: mean 0.096 g PyC g\(^{-1}\) C; 95% confidence interval 0.064-0.15 g PyC g\(^{-1}\) C) and non-woody fuels (NWF: mean 0.091 g PyC g\(^{-1}\) C; 95% confidence interval 0.074-0.11 g PyC g\(^{-1}\) C).

Assigning PyC Production Factors in GFED4s+PyC

\(P_{\text{PyC}}\) values were assigned to each of the native fuel classes of GFED4s\(^1\), which are: leaves; grasses; surface fuels (including litter and fine woody debris); coarse woody debris (CWD), and; standing wood (including trunks, stems and branches). Mean \(P_{\text{PyC}}\) values and bootstrapped confidence interval values for CWF, FWF and NWF from the global dataset were used to define representative \(P_{\text{PyC}}\) values for each of the GFED4s fuel classes (Figure 2). Full details regarding the assignment of \(P_{\text{PyC}}\) values to each GFED4s fuel class are provided in the supplementary material (text S3 and table S3). Briefly: leaf, litter, grass were assigned the relevant \(P_{\text{PyC}}\) values of NWF; fine woody debris and coarse woody debris were assigned the values of FWF and CWF, respectively, and; \(P_{\text{PyC}}\) values for standing wood were applied in a spatially explicit manner as weighted combinations of the \(P_{\text{PyC}}\) values for CWF (for carbon in trunks) and FWF (for carbon in branches). The weighted CWF:FWF ratio was assigned according to empirical relationships defining biomass carbon apportionment to branches and trunks in the various forest types of the GFED4s land cover scheme (supplementary text S3 and table S4)\(^93\).

Quantifying ENSO Impacts on PyC Production

To investigate the influence of pan-tropical climatic variability driven by the El Niño-Southern Oscillation (ENSO) on the production of PyC, we replicated the analysis presented by Chen et al. (ref. \(^52\)) with a focus on PyC production rather than carbon emissions. The pan-
tropics were defined as consisting of Central America (CEAM); Northern Hemisphere South America (NHSA); Southern Hemisphere South America (SHSA); Northern Hemisphere Africa (NHAF); Southern Hemisphere Africa (SHAF); Southeast Asia (SEAS); Equatorial Asia (EQAS), and; Australia (AUST; supplementary Figure S6). PyC production in El Niño and La Niña phases was compared for the major fire season periods defined in each tropical region by Chen et al. (ref. 52); the reader is referred to their study for a thorough explanation of the rationale for selecting these comparison periods. We summed PyC production in the major fire season period of each region and disaggregated this total to forest and non-forest fires according to the dominant land cover type in the GFED4s land cover scheme (based on the MODIS Land Cover Type Climate Modelling Grid product MCD12C1).

**Apportioning Sources of PyC**

Following GFED4s+PyC model runs, PyC production was assigned to specific sources following a method developed previously for use in GFED4s model runs1,77. Specifically, PyC production occurring as a result of non-deforestation fires was disaggregated in each cell to tropical forest, savannah/grassland, boreal forest, temperate forest, and agricultural fires using an existing algorithm that utilises fractional tree cover, climate and fire persistence variables. The reader is referred to ref. 77 for a full discussion of this algorithm. We added an additional latitudinal constraint (30 °N-30 °S) to further disaggregate the savannah compartment, which thus separates tropical savannahs and grasslands from extratropical grasslands.
References only in the Methods


Friedl, M., Sulla-Menashe, D. MCD12C1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V006. (2015). doi:10.5067/MODIS/MCD12C1.006