Plants acquire carbon through photosynthesis to sustain biomass production, autotrophic respiration and production of non-structural compounds for multiple purposes. The fraction of photosynthetic production used for biomass production, the biomass production efficiency, is a key determinant of the conversion of solar energy to biomass. In forest ecosystems, biomass production efficiency was suggested to be related to site fertility. Here we present a database of biomass production efficiency from 131 sites compiled from individual studies using harvest, biometric, eddy covariance, or process-based model estimates of production. The database is global, but dominated by data from North America and Europe. We show that instead of site fertility, ecosystem management is the key factor that controls biomass production efficiency in terrestrial ecosystems. In addition, in natural forests, grasslands, tundra, boreal peatlands and marshes, biomass production efficiency is independent of vegetation, environmental and climatic drivers. This similarity of biomass production efficiency across natural ecosystem types suggests that the ratio of biomass production to gross primary productivity is constant across natural ecosystems. We suggest that plant adaptation results in similar growth efficiency in high- and low-fertility natural systems, but that nutrient influxes under managed conditions favor a shift to carbon investment from the belowground flux of non-structural compounds to aboveground biomass.

The fraction of gross primary production (GPP) used for biomass production (BP) of terrestrial ecosystems has recently been coined biomass production efficiency (BPE; ref. 2). BPE is typically used as a proxy for the carbon-use efficiency or NPP-to-GPP ratio, where NPP refers to net primary production—that is, BP plus the production of non-structural organic compounds. Current knowledge about BPE is mainly derived from research on forests. Earlier work reported BPE to be conservative across forests, but that nutrient influxes under managed conditions favor a shift to carbon investment from the belowground flux of non-structural compounds to aboveground biomass.

In previous studies, and the impact of N deposition on BPE is largely overlooked. Here, we postulate that the impact of management on BPE is underestimated. In addition to a direct effect on BPE through selection of the most efficient plants, management can indirectly affect BPE through effects on site fertility and related belowground dynamics. Understanding of these dynamics not only will clarify the controls of BPE but also elucidate the human impacts on BPE.

We compiled a new BPE data set comprising 131 sites, including forests, grasslands, croplands, wetlands (temperate marshes and boreal peatlands) and tundras (Methods). All major climatic zones (from polar to tropical) were represented, but managed sites were located almost entirely in the temperate and boreal zone of North America and Europe (Supplementary Fig. 1 and Supplementary Table 1). For each site, our data set also included vegetation characteristics, environmental data and information on anthropogenic impacts such as management and atmospheric N deposition (Supplementary Table 2). With regard to management, we adopted a binary classification (Methods), distinguishing natural sites (pristine sites or sites with a low human impact that remain free of human influence) from managed sites (sites dominated by human activity with impacts that would not occur in nature—for example, newly established and fertilized grasslands). The utility of this classification was tested against more complex classifications.
Annual values of BP and GPP with uncertainty intervals ±n

| Relationship between BP and GPP of natural and managed ecosystems. | BPE of natural and managed ecosystems.
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest nat.</td>
<td>Forest man.</td>
</tr>
<tr>
<td>Grassland nat.</td>
<td>Grassland man.</td>
</tr>
<tr>
<td>Marsh</td>
<td>Peatland</td>
</tr>
<tr>
<td>Tundra</td>
<td>Cropland</td>
</tr>
</tbody>
</table>

| Natural and managed ecosystem types |
| Natural ecosystem types |
| Managed ecosystem type |

Figure 1 | BPE of natural and managed ecosystems. BPE (mean ± 1 s.e.m.) of (a) natural ecosystem types that can be regularly managed, such as forests and grasslands; (b) natural ecosystem types that are not commonly managed, such as temperate marshes, boreal peatlands and tundras; and (c) anthropogenic ecosystem types, such as croplands, that are not in a natural state but are maintained through management. Difference within forest types was significant at p < 0.001 (**), whereas difference within grassland types was significant at p < 0.05 (*). Light grey columns indicate natural (nat.) conditions and dark grey columns managed (man.) conditions.

(Methods), whereas its reproducibility was assured by the definition of several sub-categories within the ‘managed’ and ‘natural’ classes (Supplementary Table 3). The BPE data set, comprising the ancillary site information, is available in Supplementary Data. Our data analysis consisted of multinomial ordered logistic regressions to examine the relationship between fertility and management (code available in Supplementary Information), combined with linear (univariate analysis, multiple linear regressions) and nonlinear approaches (Random Forest) to extract emerging relationships between BPE and its potential predictors (Methods).

The analysis proceeded in five steps, using different subsets of our database. (1) We analysed all natural sites to test whether BPE is driven by natural variation in site fertility. The results showed that this hypothesis was not true. First, BPE did not differ significantly (p = 0.83) among natural ecosystem types of contrasting fertility status—that is, tundra and boreal peatlands (nutrient-poor), temperate marshes (nutrient-rich) and forests and grasslands (with variable but overall intermediate fertility status)—showing an average BPE (and s.e.m.) of 0.46 ± 0.01 (Fig. 1 and Supplementary Table 4). Second, the impact of fertility on the BPE of natural ecosystems remained nonsignificant when accounting for variation in fertility among forests (p = 0.24, n = 43), grasslands (p = 0.72, n = 16) or all natural sites lumped together (p = 0.23, n = 75; Supplementary Fig. 2). (2) We analysed the relationship between fertility and management in natural and managed forests to verify their correlation and disentangle the impact of management on fertility from the fertility status unrelated to management. This analysis confirmed that management was a significant explanatory variable for site fertility (likelihood ratio test of models with and without management as covariate: chi-square = 17.33, p = 0.00017), whereas the relationship between N deposition and fertility was weak (likelihood ratio test: chi-square = 4.80, p = 0.091). This led us to model fertility as a function of management (taking into account that the fertility status was the result of both the impact of management operations on soil nutrient availability and the management choice of which land, for example, high or low fertility, to manage) and to obtain model residuals for each site representing the ‘fertility status not explained by management’—defined hereafter as ‘unexplained natural fertility’ (Methods). (3) Once the effect of fertility and management were disentangled, we evaluated their relative importance as controllers of BPE and compared them to other possible BPE drivers (for example, vegetation and environmental characteristics, N deposition) within the forest data set. This analysis revealed that management was the key determinant of the differences in BPE among forests, N deposition was the second most important driver, and the unexplained natural fertility was insignificant (Supplementary Table 5 and Supplementary Fig. 3). The analysis also showed that age had a significant (negative) impact on BPE, which, however, became negligible when compared to the effect of management and N deposition (Supplementary Table 6). (4) We compared the BPE of key natural and managed ecosystem types (grasses, forests and croplands) that typically share similar environmental characteristics and are regularly converted into one another, and observed that the BPE of managed sites was substantially greater than the BPE of natural sites (Figs 1 and 2 and Supplementary Table 7). (5) Last, we studied the impact of the potential drivers of BPE on all natural ecosystems and found that BPE of natural unmanaged sites was independent not only of the observed site fertility (see above point 1) but also of N deposition, and largely independent of all the vegetation and environmental drivers examined (Supplementary Table 8 and Supplementary Fig. 3). Climate showed an influence on BPE, but this effect was weak (0.05 < p < 0.10) and not consistent across statistical methods (Supplementary Table 8 and Supplementary Fig. 3).

The observed positive impact of management on BPE does not come as a surprise in itself. Rather, the novelty of this study is the finding that management is by far the ‘key’ driver of BPE and more important than any other vegetation or environmental factors. This observation calls for a refinement of the hypothesis, which previously postulated that greater BPE in more fertile sites is related to reduced C allocation to symbiotic fungi, as plants in nutrient-rich conditions invest less in processes facilitating nutrient uptake. Our revised hypothesis relies on the fact that adaptation processes in natural ecosystems could allow plants in
both nutrient-poor and nutrient-rich environments to have similar growth efficiency. However, belowground C transfers to symbionts are not static\(^2\), and the greater nutrient availability caused by management could make root symbiotic associations less important for plants, thus reducing the flux of C from plants to symbionts. This pattern would favour C investment in biomass production, particularly aboveground, as light may become the most limiting resource. This hypothesis is supported by the allocation pattern available for a subset of our forests, showing that management substantially increased allocation to aboveground wood BP (+13%, \(p < 0.001\)) and marginally decreased allocation to fine root BP (−4%, \(p = 0.083\)) (Table 1), and by forest C allocation meta-analyses\(^4\) which reported increased C partitioning to aboveground BP and decreased partitioning to belowground C flux in response to fertilization. Declines in mycorrhizal fungi following fertilization are well known\(^5\). Similarly, thinning can negatively affect the standing crop of mycorrhizal fungi\(^6\) and ectomycorrhizal metabolic activity\(^11,12\), which is consistent with our new interpretation. In addition, the larger BPE in managed ecosystems might also reflect decreased allocation of GPP to autotrophic respiration (Ra), thus lower Ra-to-GPP ratio\(^7\). However, as previous research does not support this hypothesis\(^8,9\) and the variability of the Ra-to-GPP ratio might be small, \(ad\) \(hoc\) experiments combining the assessment of C transfer to mycorrhizal fungi and ecosystem Ra will be needed to ascertain the importance of these dynamics in managed ecosystems. Similarly, further research should explore if the hypothesized reduction in C allocation to mycorrhiza (and exudates) might have a long-term negative feedback on the site nutrient availability where management does not include external input of nutrients, as well as the impact of ecosystem degradation on BPE, especially in tropical areas that are often overexploited.

Nitrogen deposition also seemed to have a positive effect on BPE. Like management, elevated N deposition represents an artificial change in natural fertility and a perturbation of the nutrient cycle. The apparently contrasting evidence that N deposition does not affect BPE of natural ecosystems (when considered separately from the managed ecosystems) is probably related to the intensity of the deposition and the fact that N deposition might influence BPE (like other ecosystem processes\(^13\)) only at higher deposition rates. Natural sites are typically found in less urbanized locations, and in our data set they were characterized by deposition rates 43% lower than those of managed ecosystems. Furthermore, adaptation responses to N deposition are more likely to occur in natural ecosystems where succession is much longer than rotations in managed ecosystems.

Little information was previously available about BPE of non-forest ecosystems\(^4\). Our analysis showed that BPE of natural ecosystems is independent of ecosystem type, vegetation and environmental characteristics (including natural site fertility). The lack of sensitivity of BPE to these potential drivers points to a rather conservative BPE across natural ecosystems. Our study supports the physiological argumentation for a constant ratio between BP and GPP in natural ecosystems\(^3,4\) and provides important constraints for the global models that simulate high variability in BPE or NPP-to-GPP ratio.

Finally, our findings have practical applications, particularly for Europe and North America. First, the quantification of BPE for managed ecosystems can improve yield simulations by models (for example, timber in forests, grains in crops), particularly for algorithms that derive BP as a proportion of GPP (refs 15,16). Second, the land surface component of Earth system models at present does not take into account differences between natural and managed ecosystems which might introduce biases in BP projections. In fact, a case study based on the model ORCHIDEE (ref. 17) showed that taking into account a BPE difference of 8% between natural and managed ecosystems resulted in a 24% increment in BP for Europe (Supplementary Methods). Third, our study indicates new ways to indirectly derive BPE at regional and continental scales from maps of land use and human management. Fourth, whereas C assimilation and BP are extensively studied, the ways to maximize BPE are less explored. However, substantial changes in yield are potentially associated with small changes in BPE. For instance, for a forest with a GPP of 1,500 g C m\(^{-2}\) yr\(^{-1}\), an increase of 12% in BPE (Supplementary Table 7) would enhance BP by 180 g C m\(^{-2}\) yr\(^{-1}\), mainly in wood (Table 1). These examples show that our elucidation of BPE dynamics advances our understanding and quantification of the biomass production of terrestrial ecosystems.

### Methods

Methods and any associated references are available in the online version of the paper.

Received 20 November 2014; accepted 1 September 2015; published online 5 October 2015

### References


### Table 1 | Carbon allocation pattern in natural and managed forests as expressed by the ratio of BP to GPP.

<table>
<thead>
<tr>
<th>BP:GPP ratio</th>
<th>Natural (%)</th>
<th>Managed (%)</th>
<th>(p) difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaves</td>
<td>10 ± 1</td>
<td>10 ± 1</td>
<td>0.91</td>
</tr>
<tr>
<td>Wood</td>
<td>11 ± 1</td>
<td>24 ± 3</td>
<td>0.0001***</td>
</tr>
<tr>
<td>Other aboveground</td>
<td>6 ± 2</td>
<td>7 ± 3</td>
<td>0.61</td>
</tr>
<tr>
<td>Fine roots</td>
<td>12 ± 2</td>
<td>8 ± 2</td>
<td>0.083*</td>
</tr>
<tr>
<td>Coarse roots</td>
<td>3 ± 1</td>
<td>4 ± 1</td>
<td>0.29</td>
</tr>
<tr>
<td>Whole ecosystem (BPE)</td>
<td>41 ± 2</td>
<td>53 ± 3</td>
<td>0.020**</td>
</tr>
</tbody>
</table>

Values are mean ± 1 s.e.m, in percentage; replicates (n): 12 and 19 for natural and managed forests, respectively; other aboveground: reproductive organs and understory; ***\(p < 0.001\); **\(p < 0.05\) and *\(0.05 < p < 0.10\).
Acknowledgements
M.C., S.V. and J.B. are Postdoctoral Fellows of the Research Foundation—Flanders (FWO). S.L. is funded by the European Research Council (ERC) Starting grant 242564 (DOFOCO) and M.F.-M. by the Catalan Government’s FI-2013 grants. We also acknowledge the ERC Synergy grant ERC-2013-SyG-610028 IMBALANCE-P. Ancillary data were provided by: G. Alberti and G. Delle Vedove, C. Bernhofer and T. Grünwald, M. Jackowicz-Korczynski, B. Loubet (Environment and Arable Crops Research Unit, INRA, AgroParisTech), A. E. Suyker, M. Vadeboncoeur and D. Zanotelli. This work used eddy covariance data acquired by FLUXNET and, in particular, by the following networks: AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia and USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, PAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval and Environment Canada, and US Department of Energy; and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California-Berkeley, and University of Virginia. For sites at Kellogg Biological Station, financial support was provided by the DOE Office of Science (DE-FC02-07ER64494) and the Office of Energy Efficiency and Renewable Energy (DE-AC05-76RL01830).

Author contributions
M.C., S.V., S.L. and I.A.J. conceived the paper; M.C. performed the analyses and wrote the text; S.L. provided ORCHIDEE simulations; J.B. developed the multinomial ordered logistic regressions and the statistics; E.C., D.O., D.P., P.F.S., X.W. and T.Z. provided field data or contributed to data collection from external databases and literature; all authors contributed substantially to discussions and revisions.

Additional information
Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.C.

Competing financial interests
The authors declare no competing financial interests.
Methods

Data set. Our analysis required site estimates of biomass production (BP), gross primary production (GPP), and their uncertainty, to derive the biomass production efficiency (BPE) and their uncertainty. The two main sources of uncertainty were the large variation in the availability of site-specific estimates of BP and GPP. Therefore, the data set did not include values obtained from generic algorithms (for example, global models, remote sensing products). BP included above- and belowground growth. In most cases, BP was obtained from harvest or biometric methods (comprising empirical models as, for example, allometric relationships, root growth as function of soil conditions15) and in 5% of the cases from process-based models with site-specific parameterization and/or calibration against growth or biomass data. Model gap-filling was done for BP estimates at some sites (see below). BP methodologies can be divided into broad classes according to method uncertainty (that is, low, medium or high uncertainty1; Supplementary Table 9) related in particular to the approach to determine fine root BP (the component of ecosystem BP most difficult to assess; see Supplementary Methods) or the use of process-based models (Supplementary Table 9). However, additional tests showed that the key results of our analysis were independent of the BP methodology employed (Supplementary Table 10). GPP was mostly estimated from eddy covariance (73% of the cases) or process-based models with site-specific parameterization and/or validation (20% of the cases). Explanation about the preference of these GPP methods instead of other approaches (for example, GPP derived from the sum of all carbon sinks within the ecosystem such as, for example, BP, autotrophic respiration, carbon transfer to mycorrhizal symbionts) is reported extensively in Supplementary Methods. Additional tests showed that the alternative use of eddy covariance- or model-based estimates of GPP did not affect the key results of our analysis (Supplementary Table 10). Detailed information on uncertainty calculations are reported in Statistical analysis.

The integrated data set provided BPE for 96 golden sites, for which BP and GPP were available from the same measuring period (53 forests, 14 grasslands, 24 croplands and 5 additional natural sites for which BP and GPP were both available but not for the same measuring period (16 forests, 6 grasslands, 8 wetlands, 5 tundra). Wetlands were divided into marshes (herbaceous-dominated vegetation of the temperate zone mainly affected by flooding from river, sea or irrigation; 6 in total) and peatlands (ombrotrophic or minerotrophic inland boreal ecosystems rich in herbs, shrubs or mosses; 7 in total). An excerpt from the study data set is shown in Supplementary Table 1, and the geographical distribution of the sites is shown in Supplementary Fig. 1. The key data used in the analysis are provided in Supplementary Data.

Ancillary data such as vegetation characteristics, climate, environmental conditions and anthropogenic impacts were needed for each site to determine the possible effect of these factors on BPE. Such information was retrieved mostly from the literature, open-access databases12,17 or modelling18 (Supplementary Table 2). For N deposition, data for Western Europe and the conterminous USA were retrieved from interpolated gridded maps based on ground observations21, whereas simulated values were used for the rest of the world22,23.

Management classification. Sites were divided into two categories. Natural sites are those characterized by none or low-to-moderate human impact, whereas managed sites are heavily affected by human activity. We defined 'low-to-moderate human impacts' as human activities that largely reproduce naturally occurring processes, for example, low grazing, occasional fire in grasslands, forest regeneration. We considered sites 'heavily affected by human activity' to be those with impacts that would not occur in nature—for example, intense fertilization of poor soils, sowing of cropland monocultures, thinning of healthy trees. The classification was straightforward for marshes, peatlands and tundras (pristine or with minimal human impact except in two managed wetlands) and for croplands (inherently managed) (Supplementary Table 3). For forests and grasslands, the classification included sub-categories for ecosystems rich in herbs, shrubs or mosses; 7 in total). An excerpt from the study data set is shown in Supplementary Table 3, and the geographical distribution of the sites is shown in Supplementary Fig. 1. The key data used in the analysis are provided in Supplementary Data.

Statistical analysis. Analysis overview and data set. Our study consisted of five analyses, using different subsets of our database. (1) We analysed all natural sites (n = 75; managed sites were not considered in this analysis) to test whether BPE is driven by natural variation in site fertility. In particular, we tested whether BPE differs among ecosystem types and sites of contrasting fertility. (2) We analysed the relationship between fertility and management in forests to verify their correlation and disentangle the impact of management on fertility from the fertility status not related to management. This analysis was performed on 53 managed and natural forests for which BP and GPP were measured during the same period. We focused this analysis on forests because they are the ecosystem type best represented in our data set and allow direct comparison with previous studies. (3) The relative importance of fertility, management, and N deposition as controllers of BPE was compared to the importance of other possible BPE drivers. This analysis was performed on the forest data set considered in the second analysis after dissecting the effect of fertility and management. (4) We compared the BPE of key natural and managed ecosystem types (grasslands, forests and croplands) that typically share similar climatic and environmental characteristics and are regularly converted into one another. Only sites with BPE obtained from BP and GPP measured during the same period were used (n = 93). (5) We studied the impact of the potential drivers of BPE in all natural ecosystems (n = 75; this analysis did not include the managed sites).

Gap-filling. Some of the selected sites lacked BP measurements of minor ecosystem biomass components (for example, nonvascular plants, understory) or were affected by minor systematic measurement biases (for example, neglecting litterfall decomposition in tropical forests). These missing BP portions were gap-filled for completeness in analogy to ref. 2. Production of reproductive organs in forests. When missing, this BP component was derived from a relationship between reproductive BP versus aboveground BP (ref. 2) derived from the Global Forest Database1. Leaf biomass production in tropical forests. Estimates of leaf BP in tropical forests are systematically underestimated because of within-canopy decomposition of leaf litter during the collection period. We estimated this missing portion of BP as 12% of the total leaf biomass production. Understory biomass production in forests. BP derived from understory vegetation is significant for boreal forests—thus, boreal forests lacking this BP component were not considered in our analysis2. However, the contribution of understory BP to total ecosystem BP is more limited for temperate and tropical forests2. Thus, we did not discard temperate and tropical forests lacking understory BP, but gap-filled this missing BP component, as done in previous studies1. In particular, understory BP was estimated as a fixed ratio of the forest tree BP: 0.043 for temperate and 0.073 for tropical forests.1

Nonvascular biomass production in tundra. Missing nonvascular BP was derived from a nonvascular productivity ratio (BP-to-biomass ratio, the portion of biomass renewed every year). This ratio was calculated for wet (0.50 yr−1) and mesic tundra (0.42 yr−1) as the average of six observations for each tundra type (Supplementary Table 12). Shrub biomass production due to stem secondary growth in peatland. Missing BP due to unaccounted shrub secondary growth (that is, increase in stem/branch diameter) was estimated to be 29% of the shrub aboveground primary growth (that is, BP due to current-year leaves and stems/branches) from data for subarctic shrubs24.

The gap-filling concerned 31 forests of the 96 golden sites and 17 sites (14 forests, two tundras and one peatland) of the additional 35 natural sites. For 69% of the cases, the gap-filled BP differed by less than 5% than the original BP; for 13% of the cases the gap-filled and original BP differed by 5–10%; whereas for 17% of the cases this difference was 10–15%. Herbivory was not taken into account because it was negligible (for example, for forests1) or because BP measurements were from experiments that excluded large herbivores (for example, for all grasslands examined).

The gap-filling procedure avoided small secondary biases in the analysis but did not alter the primary results (Supplementary Table 13). Overall, original BPE of managed and natural forests (the ecosystem type most affected by gap-filling) was 0.52 ± 0.03 and 0.39 ± 0.02 (mean ± s.e.m), respectively, which was less than 2% smaller than gap-filled BPE (Supplementary Table 7).

We tested the validity of our approach by comparing our binary management classification to a more complex three-level classification. In the latter approach, we considered ‘pristine natural’ the sites that were pristine or with minimal impacts, and ‘semi-natural’ the sites with low-moderate human impacts (these classes were considered jointly in the binary classification as ‘natural’). For forests, we considered as semi-natural the forests that were: unmanaged or with low human impact (for example, understory grazing) in the 50 years before measurement, as well as planted forests without any intervention after planting and at least 10 years old (see above). The statistics of this additional test showed that BPE of pristine natural and semi-natural forests did not differ and that the BPE difference between pristine natural and semi-natural forests was not significant (see Table 11). This confirmed that our standard binary classification is sound. In addition, this exercise revealed that the introduction of more levels in the management classification would not be advantageous. This was evident for grasslands, for which the three-level classification did not alter the BPE pattern but substantially reduced the statistical power (Supplementary Fig. 4 and Supplementary Table 11).
For the analyses 1 and 5, we considered not only the sites for which BP and GPP were measured during the same period but also sites with BP and GPP measured during different (or only partially overlapping) periods (35 out of the 75 sites) to investigate a large set of ecosystem types (for example, from forest to tundra and environmental conditions (for example, climate from tropical to polar, soil from waterlogged to very dry). For sites without management operations (and mostly at mature-old stage) the temporal mismatch in BP and GPP was less crucial, dampened at several sites by multi-year measurements (we used averages of BP and GPP for multi-year observations) and, most importantly, comparative tests revealed that the results of the analyses did not differ when all sites or only sites with temporal match in BP and GPP were considered (for example, Supplementary Table 14).

Relationship between fertility and management. Site fertility and site management are highly correlated factors that are both potentially crucial for BPE. For this study, we wanted to separate both drivers to test for BPE responses to the fertility status induced by management and the fertility status unrelated to management. To disentangle both effects, we applied an approach commonly used to deal with multicollinearity:22 the observed fertility status was modelled as a function of management and the residuals from this model were used as explanatory variables of BPE (instead of the original fertility status). Hence, the residuals reflect the information on fertility not explained by management, which we termed ‘unexplained natural fertility’. Initially, the model also included N deposition as an additional covariate, but we removed it in the final model as the relationship between N deposition and fertility was weak (see ecosystem type in main text).

A multinomial ordered logistic regression model (or ‘proportional odds logistic regression model’23) was fitted with fertility as outcome (ordinal categorical variable with category high, H, medium, M, and low, L) and management (yes/no) as covariate. The model estimates the log odds of falling into or below a fertility category as a function of management:

Logit \( P(\text{fertility} = L) = \text{intercept} + \beta L + \text{management} \)

Logit \( P(\text{fertility} < M) = \text{intercept} + \beta M + \text{management} \)

where intercept, and intercept\(_L\) were \(-2.01\) and \(-0.511\), respectively, and \(\beta L\) and \(\beta M\) were \(2.84\) and \(-0.0488\), respectively. In other words, this model estimates the possible fertility distribution of each site according to its management status (given its management status, the probability to be H, M or L). Also three residuals were obtained for each site, which reflect the deviation of the fertility status of the site from the distribution estimated by the model. The independence of these three residuals on management (unlike the original fertility variable) was verified with t-tests (all p-values > 0.05).

BPE drivers. The relationships between BPE and its potential drivers were explored with three statistical approaches: univariate analysis, multiple linear regressions and Random Forest, which are described below. We used the following predictors: management status, observed natural fertility, climate zone, ecosystem type, growth form (five categorical variables) and N deposition, unexplained natural fertility (the three model residuals described above), soil available water content, annual precipitation and dry months per year (seven continuous variables) (Supplementary Table 2). All analyses were performed with R (Ref. 31).

Univariate analysis tested the significance of the relationships between single predictors and BPE. For continuous variables, this was done with single linear regressions, whereas for categorical variables we used one-way analysis of variance (ANOVAR) with a post hoc Tukey’s HSD test. Normality of residuals was tested with a Shapiro–Wilk’s test and the assumption of homoscedasticity with Levene’s test (for ANOVAs) or Box–Cook–Pagan test (for regressions). For the few cases for which these conditions were not met, data were transformed (for example, log(x), 1/x or \(x^2\)) or treated with alternative methods (Kruskal–Wallis test for non-normality and applications of White method for heteroskedasticity22). Multiple linear regressions allow a comparison of the effect of the potential BPE predictors considering them all together. Whenever a given predictor was significant in the univariate analysis, but not in the multiple linear regressions, this indicated a lower importance of that predictor as compared to other predictors. In practice, we opted for backward stepwise regressions. Accordingly, the best BPE model was determined by starting from the model with all variables and successively removing the least important. The selection was done by comparing the new model (without the removed variable) with the original model (with the original variable) using Likelihood Ratio and Akaike Information Criterion (AIC). In practice, the new model was not accepted if the Likelihood Ratio was significant (p < 0.05) or the AIC increased. Stepwise multiple linear regression was a suitable methodology for our analysis, because it can be applied with both continuous and categorical variables. However, all factors of categorical variables need to be taken into consideration by introducing dummy variables. Prerequisites (or alternatives) for applying linear regressions (for example, residuals normality and homoscedasticity) were tested as described above for univariate analysis.

Random Forest is a partitioning method that we used to produce a large ensemble of regression trees considering always our complete BPE data set but random subsets of predictor variables23. This means that (in contrast to multiple linear regressions) Random Forest accounts also for nonlinear relationships and interactions, and evaluates each predictor variable (even the least important or redundant), providing a ranking of the predictors’ importance. However, this significance label (contrary to linear regressions) analysis. The importance of a given variable is instead indicated by the mean decrease in accuracy (or increase in mean squared error, %InMSE) of model predictions when the value of that given variable was changed (permuted within the data set)23. The more important the variable, the larger the difference between original predictions and new predictions, and the larger the %InMSE. We used the standard Random Forest algorithm24 setting a large number of trees (50,000) to obtain stable results. The response of BPE to N deposition and variables related to the water status (soil available water content, precipitation, dry months per year) could have been confounded by fertilization and irrigation/exceptional soil water conditions, respectively, at some sites. To check for the relevance of confounding factors, the analyses comprising N deposition and the variables related to the water status were performed both on the entire data set and on a subset that excluded sites with fertilization, irrigation, occasional flooding, minerotrophic conditions and permafrost. Overall, the impact of these sites was negligible (Supplementary Table 15)—therefore, they were not removed in the final analyses. Through the analysis, filtering for outliers was minimal and we removed only four sites with unrealistic BPE (0.84–0.94).

Uncertainty. The BP uncertainty for site \(i (\text{s}_{\text{BP},i})\) depended on a typical range of uncertainty \(\text{s}_{\text{BP}}\) (see ecosystem type, the experimental methodology y through a method-specific uncertainty reduction factor \(\text{RF}_{\text{y},i}\)) and the length of the measurement period in years \(\text{t}_{\text{BP},i}\) (Ref. 1):

\[
\text{s}_{\text{BP},i} \text{gapfilled} = \frac{(\text{s}_{\text{BP},i} \times \text{RF}_{\text{y},i})}{(\text{t}_{\text{BP},i})^{0.5}}
\]

where \(\text{s}_{\text{BP},i} \text{gapfilled}\) is the uncertainty of the gap-filled BP estimate. The uncertainty of GPP \(\text{s}_{\text{GPP},i}\) was calculated in the same way as \(\text{s}_{\text{BP},i}\):

\[
\text{s}_{\text{GPP},i} \text{gapfilled} = \left(\frac{(\text{GPP}_{\text{R},i} \times \text{RF}_{\text{y},i})}{(\text{t}_{\text{GPP},i})^{0.5}}\right)^{0.5}
\]

where \(\text{RF}_{\text{y},i}\) is the typical range of GPP uncertainty, \(\text{RF}_{\text{y},i}\) the uncertainty reduction factor dependent on the experimental methodology \(y\) and \(\text{s}_{\text{y},i}\) the length of the measurement period in years. The uncertainty of BPE \(\text{s}_{\text{BPE},i}\) was calculated through error propagation:

\[
\text{s}_{\text{BPE},i} = \left(\frac{\text{BP}_{\text{R},i}}{(\text{BP}_{\text{original}})^{0.5}} + \frac{\text{GPP}_{\text{R},i}}{(\text{GPP}_{\text{original}})^{0.5}}\right)^{0.5}
\]

where \(\text{BP}_{\text{R},i}\) and \(\text{GPP}_{\text{R},i}\) are values of BP and GPP, respectively, for site \(i\) and method \(j\). Values of \(\text{RF}_{\text{y},i}\) and \(\text{RF}_{\text{y},i}\) are reported in Supplementary Table 9 (Ref. 1). For forest ecosystems, values of \(\text{BP}_{\text{R},i}\) and \(\text{GPP}_{\text{R},i}\) were available in the literature, whereas for non-forest ecosystems they were derived from the difference between the ninth and first decile of BP and GPP samples from about 20 to 110 sites according to ecosystem type (Supplementary Table 16).

Code availability. Code available in Supplementary Information.

References

© 2015 Macmillan Publishers Limited. All rights reserved


