Manuscript Details

Manuscript number	FORECO_2017_103
Title	Generalized biomass and leaf area allometric equations for European tree species incorporating stand structure, tree age and climate
Article type	Full Length Article

Abstract

Biomass and leaf area equations are often required to assess or model forest productivity, carbon stocks and other ecosystem services. These factors are influenced by climate, age and stand structural attributes including stand density and tree species diversity or species composition. However, such covariates are rarely included in biomass and leaf area equations. We reviewed the literature and built a database of biomass and leaf area equations for 24 European tree species and 3 introduced species. The final dataset contained 973 equations. Most of the equations were site-specific and therefore restricted to the edaphic, climatic and stand structural conditions of the given site. To overcome this limitation, the database was used to develop regional species-specific equations that can be used in a wide range of stands and to quantify the effects of climate, age and stand structure on biomass or leaf area. The analysis showed considerable inter- and intra-specific variability in biomass relationships. The intra-specific variability was related to climate, age or stand characteristics, while the inter-specific variability was correlated with traits such as wood density, specific leaf area and shade tolerance. The analysis also showed that foliage mass is more variable than stem or total aboveground biomass, both within and between species, and these biomass components have contrasting responses to age and changes in stand structure. Despite the large number of published equations, many species are still not well represented. Therefore, generic equations were developed that include species-specific wood density instead of species identity. Further improvements may be possible if future studies quantify the stand structure of individual tree neighbourhoods instead of using the stand means for all trees sampled with the given stand.

Keywords	allometry; forest growth; pseudo-observations; shade tolerance; specific leaf area; wood density
Corresponding Author	David Forrester
Order of Authors	David Forrester, Heloise Tachauer, Peter Annighoefer, Ignacio Barbeito, Hans Pretzsch, Ricardo Ruiz-Peinado Gertrudix, Hendrik Stark, Giorgio Vacchiano, Tzvetan Zlatanov, Tamalika Chakraborty, Somidh Saha, Gudeta Sileshi
Suggested reviewers	David Chojnacky, Dimitris Zianis, Keryn Paul, Aaron Weiskittel

Submission Files Included in this PDF

File Name [File Type]

cover letter3.docx [Cover Letter]

response to reviewers.docx [Response to Reviewers]

maintext2.docx [Manuscript File]

Figure1.pdf [Figure]

Figure2.pdf [Figure]

Figure3.pdf [Figure]

Figure4.pdf [Figure]

Figure5.pdf [Figure]

Figure6.pdf [Figure]

Figure7.pdf [Figure]

highlights.docx [Highlights]

Forrester_etal_AppendixB.docx [Supporting File]

Research Data Related to this Submission

Data set https://data.mendeley.com/datasets/4jytx9s44j/draft?a=5b09bbfa-ca97-4efb-8d64-5fcbc4b34e36

Specific leaf area, biomass equations and leaf area equations for European tree species

The data set contains specific leaf area data, biomass equations and leaf area equations for 24 European tree species and 3 introduced species. This includes 868 equations from the literature and 105 raw data sets that were either from the literature or our own data. The number of biomass or leaf area measurements used to produce the equations or included as raw data was > 60,000. The specific leaf area data was also obtained from the literature. The equations were used to produce pseudo-observations of mass or leaf area and were combined with the raw data sets to develop new “general” equations for each species and biomass component (foliage, branch, stem, root, aboveground) that can account for factors describing stand structure and age by including those factors as independent variables. These general equations are provided in separate worksheets of this dataset. The main findings and methodology are described in Forrester et al. (in review). Forrester, D. I., Tachauer, I. H. H., Annighoefer, P., Barbeito, I., Pretzsch, H., Ruiz-Peinado, R., Stark, H., Vacchiano, V., Zlatanov, T., Chakraborty, T., Saha, S., Sileshi, G. W. in review. Biomass and leaf area allometric equations for European tree species incorporating stand structure and tree age.

David Forrester
Swiss Federal Research Institute WSL
Zürcherstrasse 111
8903 Birmensdorf
Switzerland
david.forrester@wsl.ch

Dear Dr. Fredericksen,

Thank you for the reviewer comments of our paper and for the opportunity to submit a revised version. We would also like to thank the reviewers for providing insightful comments and suggestions for improvement. Many of the comments refer to using actual data instead of pseudo-observations. We fully agree with the reviewers that this would be better and if this had been possible we would have done it. However, this was not feasible in Europe for many species, so no such study is currently possible unless it is greatly restricted in terms of sites and stand characteristics. Instead, the objective of this project was to review existing European biomass and leaf area equations and then to examine patterns in terms of the effects of stand characteristics. Our paper suggests that the next steps are to develop a database containing actual data, and to replace stand variables with crown variables. These recommendations are consistent with the reviewer comments.

On behalf of my co-authors, I am now resubmitting the manuscript after revising it paying particular attention to the reviewers' recommendations.

We look forward to receiving your response.

Best regards,

David Forrester

Ref: FORECO_2017_103
Title: Biomass and leaf area allometric equations for European tree species incorporating stand structure and tree age
Journal: Forest Ecology and Management

Comments from the reviewers:
Reviewer 1

General

This is an interesting, through, and well-written with some important findings. However, I have several significant concerns about the current manuscript. First, I find the use of pseudo-data rather problematic despite all the statistical methods that the authors use to try to make it more like actual data. In particular, I find using pseudo-data to evaluate potential relationship with variables beyond just DBH as problematic because of the high likelihood for spurious findings. I recognize the authors try to avoid that by incorporating variability in the underlying pseudo-data and using cross-validation methods, but I still don't believe that is sufficient when using data that really isn't actual data. Couldn't the authors simply complete the analysis using actual rather than pseudo-data? I recognize this wouldn't be as an extensive of analysis, but there seems to be plenty of existing data and the trends evaluated in this current analysis could still be examined, which I think is the most important contribution in comparison to the extensiveness of the underlying pseudo-data.

Response: We agree that the use of pseudo-observations are problematic and we also prefer to use actual data. However, after a thorough search we have come to the conclusion that for some species and/or sites, generating pseudo-observations is the only option available. Actual data are usually not available (sometimes simply lost or the original authors are not interested in searching for them). Therefore, we settled on generating pseudo-observations as this is better than using a much smaller dataset of actual data because that would defeat our purpose of the generalizability of these equations. That is, if we only use actual data, the range of conditions becomes too small and for many of the species we included.

Second, many of the details that I believe are needed to actually understand this analysis seem to be presented in other manuscripts that are currently under review. I personally believe that each manuscript should be able to stand-alone and this is not the case for this current manuscript. Not sure what the optimal solution is here? Could the necessary information be presented on a website that is fully accessible? I think it is rather non-ideal to rely on having the information in other manuscripts or publications as equal access is not the same across various journals.

Response: We do not have other manuscripts that are currently under review. We believe this comment is a misunderstanding. We strongly agree with the reviewer that all the details required to understand this paper must be available. The only "under review" citation was the data that goes with this paper, which is stored in Mendeley. This data paper will be open access if the main paper is accepted and was simultaneously submitted with the manuscript as a Mendeley data paper. This is the procedure encouraged by Forest Ecology and Management. Our understanding is that all of our pseudo-observations, actual data and equations will be freely available if the paper is published and that these are also available to the reviewer of the manuscript.

Finally, I recognize the authors have provided qualitative justification for the variables examined, but they do seem rather limited in scope and interpreability. Personally, I believe tree height or crown attributes could capture much of the variation explained by stand-level and geographic variables.

Response: We also agree with this comment that crown attributes would have been extremely useful, and probably better than all stand variables. We stated this in the discussion because we suggest that replacing the stand variables with crown variables is an important next step in developing general equations. However, the crown and height information are usually not available. If they were, we would certainly favour them.

I find tree age as difficult to interpret biologically because it can often be confounded by past stand history. Why not use a composite index (DBH2*HT) or tree slenderness (HT/DBH) to account for the effect of tree social position, stand density, and potentially geographic location? Like tree age, I believe the same can be said for latitude.

Response: We did not use any composite variables, especially with height, because they would be automatically correlated with dbh and if we included the composite variable instead of dbh the equations would be difficult to interpret because height to diameter ratios can vary in opposite directions with too many factors (wind, stand density, age, species composition …). We are not suggesting that equations including a height-diameter composite variable would not be as accurate as those we provided, but we do expect that they would be harder to interpret in terms of their biological implications. A second problem with height is that it was usually not available for the corresponding dbh that was available. Therefore including height would require its prediction and this would add unnecessary error to the biomass equations.

Why not use more direct measures of site like elevation and climate, which I believe are much more interpretative and interesting than relatively crude measures like latitude?

Response: We agree that these would probably be better. We had not included them because elevation was generally not provided. Climate was also usually not provided and needed to be calculated from latitude and longitude, which adds error to the models. One the advice of the reviewer, we have now included mean annual precipitation and mean annual temperature in the analyses.

Overall, I believe the manuscript has merit, but the concerns outlined should be addressed prior to publication.

Specific

L135-162: These paragraph seem a bit out of place, unnecessary, and speculative. Seems more of potential Discussion items? The tested variables do make sense, except age, which can introduce additional limitations, particularly for shade-tolerant species capable of surviving long periods of suppression or in stands with past harvesting. I would suggest deleting these paragraphs and revisting these items in the Methods and/or Discussion.

Response: This text has been moved to the methodology section. We agree that age is problematic for some of the species, including the shade tolerant species. This is why we did not force any of the variables (including age) to be in the equations and only used explanatory variables that were significant and passed all of our statistical tests.

L182: Seems a description of the Study Area is needed? Is this work applicable to all of Europe or just certain locations? Either way, I do think a description of the tree species and typical forest types is warranted.

Response: Two sentences have been added to the first paragraph of the methodology section to indicate the generality of the equations and to refer to Table 1 that shows the ranges of site and stand variables for each species.

L231-249: The idea of "pseduo-data" continues to allure despite its multitudes of limitations. I strongly believe that any conclusions drawn from "pseduo-data" are mostly speculative and not easy to verify. Most importantly, they overly underestimate the observed variance, which I believe can make relationship spurrious. So I think better highlighting the limitations of the approach is warranted, which I see is down below this paragraph. However, I believe these points can't be overemphasized. Again, I think the analysis would be stronger using just actual rather than pseudo-data.

Response: As mentioned in a previous response, we fully agree that actual data is strongly preferred to pseudo-observations and we specifically state this in the methodology section. Unfortunately, these data appear to be much more difficult to accumulate in Europe compared with other continents (containing fewer countries). To prevent the underestimation of variance we use the MSE of the published equations to calculate pseudo-observations that took account of this error. This approach is described in the methods section and discussed in the discussion. We certainly don't suggest that this completely solves the problem, but it has been demonstrated to be satisfactory in the papers that we cited in the methods section (e.g. de-Miguel et al., 2014). We also discuss some of the implications in the discussion section.

L244: How were the 33 estimates actually generated? Random diameters or using the observed range of diameters with various quantiles? The latter seems it would be most effective. I see this is discussed a bit more in the preceding paragraph, but it is still unclear the actual approach used in this analysis.

Response: We have added a sentence explaining that "… 33 pseudo-observations were calculated for each equation that were evenly distributed between the smallest and largest diameter used to develop the given equation." Unfortunately the observed distributions of diameters and quantiles were almost never provided (mentioned in the discussion section), otherwise this would have been used.

L248-249: Should define the % of studies that this actually represented. I suspect it would be a rather small value. Did you see what data was available Ecology, 96(5), 2015, 1445 (http://onlinelibrary.wiley.com/doi/10.1890/14-1889.1/epdf)?

Response: This was mentioned in the results and is now also mentioned in this part of the text as well (868 equations and 105 raw data sets).

L300-301: I remain confused on how things like age and basal area were determined if they weren't available from the original data and you are using psdeuo-data. If you have the latitude, why not use more direct measures of site like climate?

Response: Stand variables were only ever taken from the papers, and never calculated when no information was provided. Climate has now been included in the analyses.

L355-360: This is nice and worthwhile, but seems a bit overkill given the lack of actual data. At best, I would call it verification.

Response: The section has been renamed by replacing the word validation with verification.

L365: I believe inclusion of DBH and HT is important in biomass equations. Why not use a composite index like DBH2HT or tree slenderness (HT/DBH) ?

Response: We responded to this comment when it was listed in the general comments from this Reviewer.

The issues with such composite variables are described in more detail in Sileshi, G. W., 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. Forest Ecology and Management 329, 237-264

L381: Was this done using just the fixed-effects or the fixed+random effects? I believe the former is more telling than the latter.

Response: We only used mixed models, so therefore the latter. We have now explained this in the methods section.

L399-403: Likely a regression-based equivalence test would be more effective here. See Tree Phy. 25: 903-913. I believe reporting the area of equivalence is more effective than testing whether the slope and intercept are statistically significant.

Response: This is an interesting study. We have noted it for the next step of this work but have not applied it in this study.

L420-422: Seems a bit odd to citing tables from other manuscripts or publications in the Results. Seems best to make each publication stand-alone.

Response: There appears to have been a misunderstanding. This citation is for the data paper that accompanies this manuscript, which appears to be the preferred approach of Forest Ecology and Management, otherwise we would be happy to provide it as an Excel table as supplementary information.

Table 1: Probably be helpful to report HT. I think it would also be important to distinguish between pseduo- and raw data. Seems you could do that by providing both the number of equations and number of raw datasets by species.

Response: We are not sure what is meant by HT. In the case that it is height: We agree that this is a useful variable to provide. However, we did not include height because it is not often reported and most of the heights in the database were predicted (as explained in the methods section) for the published equations that required height (in order to calculate pseudo-observations). Therefore our height data is probably not representative enough of the majority of equations to be useful in this table and that is the reason why we did not include it in the table.

The number of equations and raw data sets is now included in Table 1.

Table 2: Not sure I understand the caption, "excludes the equations where all broadleaved species, coniferous species or all species were grouped together." If that's the case, what does the equation actually include? Seems worthwhile to split it by species type as well as present overall coefficients.

Response: This text has now been removed to avoid confusion.

Table 3: Be helpful to provide some actual fit statistics (R2, RMSE, MB, BIC, etc.) for these various equations.

Response: In the caption we mention that the fit statistics for these equations are provided in a table of the supplementary information (Mendeley data paper).

Figure 2: Probably be more useful to display the authors' of the four equations rather than simply labeling them 1-4. I recognize that they are defined in the other manuscript, but I strongly believe each publication (especially Tables and Figures) should be able to stand alone.

Response: The citations for the papers have been added. As mentioned above, the data is in an accompanying paper, it is not really a separate paper.

Figure 3: How were the equations actually selected for display? Are they all equations for each species? Seems it would be more effective to choose for a wide range of observed geographic distribution. Couldn't you also put uncertainty bands around these to highlight the underlying variability?

Response: Yes, these are all equations for each species. They are used to illustrate the variability. We did not include the fitted equations or uncertainty bands to avoid making these figures too complicated and difficult to read and also because they don't help to make the point.

Figure 6: Should define what is actually meant by "low" and "high".

Response: An explanation has been added to the figure caption (low = mean – 1 standard deviation, high = mean + 1 standard deviation).

Reviewer 2

General comments
The topic of the submitted manuscript is within the scope of Forest Ecology and Management, it is highly relevant and very interesting but should not be published in the present form. There are two main reasons to reach such a decision:
i) the paper does not address the most important aspects of current debate on allometry i.e., the usefulness of linear regression on log-transformed values (Packard and Birchard 2008 J. Exp. Biol. 211:3581–3587; Xiao, et al. 2011; Ecology 92:1887–1894) and the usefulness of Bayesian analysis for developing tree biomass allometry (Zapata-Cuartas et al. 2012 For. Ecol. Manage. 277:173–179; Zianis et al. 2017 For. Sci. 62:247–259). I think that the major contribution to tree biomass allometric studies should include a comparison between log-linear and non-linear model as well as the performance of classical statistics vs bayesian analysis. My suggestion is that the authors should explicitly state these important limitations in their MS.

Response: While the use of non-linear regression as argued by Pickard may be applicable for fitting a simple allometry (our Model 3), non-linear regression cannot be used for the various models we explored. We agree that comparison of the results with Bayesian analysis is certainly an interesting topic but we view this as an alternative objective to those of our paper

herehereherehere

and not prerequisites for new papers about allometry. There are several debates in the field of allometry but these methodology-related debates are beyond the scope of our paper. All of our data will be freely available as a Mendeley database (if this paper is accepted) and could therefore be used to address these methodological debates in future studies.

ii) inconsistencies in the methodological and statistical structure of the manuscript (both in the scientific and linguistic context) can easily confound and disorient the reader.

Response: We agree that inconsistencies are problematic in papers. But after reading all the comments from this reviewer, we don't understand what they are referring to.

Large part of the Discussion section should be rewritten since the conclusions about several structural-functional characteristics of the compiled species are not directly derived from the analysis (e.g., no validation is presented against the derived predictions)

Response: In the discussion we describe the statistically significant patterns that were found in this set of equations (and raw data sets). Patterns that were not examined, or significant, are not discussed. There was a validation in terms of the k-fold method for the equations. This is described in section “2.5 Verification and quality of fit”. Therefore we don’t understand what the reviewer would like us to improve.

A non-exhaustive list of specific comments is following.

Specific comments

The title is misleading since the analysis of the MS is based on generalized equations (derived from pseudo-observations) and not on actual raw data collected across several stands differing in structure and tree age. Thus the word generalized should be added before Biomass.

Response: We have added the word generalized to the title.

Line 75-86: Should be removed; not relevant to the main analysis which should have been based on several regression techniques in order to reject or accept a theoretical value.

Response: We agree that this forms a small part of our analysis. But the values of allometric exponents are a topic of fierce debate and it is likely that if we do not at least provide a sentence acknowledging the competing theories our introduction will appear as an inadequate representation of the current state of knowledge. Therefore we have retained two sentences that name the competing hypotheses.

Line 142-153: It should be reported how much of biomass variability across the diameter range is explained by age, TPH, basal area etc. It is most common that the diameter alone explains more than 90%, so further experimental justification for the use of more than one variable is needed.

Response: Following this comment and a comment from reviewer 1, these sentences have been moved to the methods section. The variability that is explained by the different equation forms (Equations 3-13) and therefore the different variables is provided in the supplementary information (e.g. Table A.4) where the statistics for each equation are described in detail. We also included Figures 6 and 7 to indicate the importance of each variable on each biomass component.

Line 165:…develop a database..” to my understanding the authors presented pseudo-observations derived from the compiled equations in Table A.2; Table A.3 is a reference list of the compiled equations; Tables A.4-6 present parameters for the eqs 3-13. Thus, not a database of allometric biomass equations is presented in the Excel file.

Response: Table A.2 contains all the equations (in column R) as well as all the pseudo-observations or raw data. In the figure caption we now specifically state where the equations are listed. We don`t understand what the reviewer means when suggesting a large set of equations or a large set of pseudo-observations or a long list of papers is not a database. We fully acknowledge that this database is composed of only a small amount (about 10%) of raw data, but we are not familiar with the definition that databases only contain raw data.

Line 169: Even thought the authors report that "The review of the literature resulted in a total of 973 equations, including raw data sets obtained from tables…", table A.2 contains 982 equations. Pseudo-observations should not be mixed with raw data. Pseudo-observations are used to derive generalized equations (according to Pastor et al terminology), and raw data are used to derive original empirical equations. Raw data should be used for validating generalized equations and NOT mixed with pseudo-observations to develop generalized equations since they are much more 'noisy' than the pseudo-observations. In any case, the raw data should be reported separately in Table A.

Response: It is not clear which "Pastor et al" the reviewer is citing. However, the papers we mentioned above, including Pastor et al did not define generalized equations as the reviewer defines them. Our description, and that of Paster et al., are consistent with the other studies that used pseudo-observations and/or generalized equations (e.g. from our reference list: Muukkonen 2007, de-Miguel et al 2014, Jenkins et al., 2003, Chojnacky et al., 2014, Pastor et al., 1984, Paul et al 2016). Our understanding from these papers is that the term generalized equations refers to fitting equations after combining data sets from a wider range of conditions (than each individual data set), not whether pseudo-observations were used or not.

The database includes a column (column AB) that indicates whether the data is raw data or pseudo-observations.

Line 179: "….and can therefore be used in a wide range of forest types." For aboveground compartment of Fagus sylvatica only 9 equations were compiled. In 3 of them, age, basal area and TPH are missing. So, 6 equations could not be used as a basis to predict aboveground biomass for all the beech "forest types". It should be restated.

Response: This sentence, in our introduction, was describing our objective, not the results. We have reworded it to say "wider" instead of "wide". In the 3rd paragraph of the discussion we state that the equations should only be used for stands that fall within the range of stand or climate conditions where the published equations were developed.

Lines 195-197: It could be argued that this is not a statistically sound approach since H-D allometry is not that strong (R^2 around or less than 80%). Further justification should be provided and potential ramifications should be reported.

Response: We agree that H-D allometry is often not very strong, especially for larger trees. Therefore, we only used the shape of the height functions and applied this to the minimum and maximum height of the measured trees. This constrains the predictions much more and minimised the potential problem. This has been explained in the text.

Line 198: Was there a statistical test for outliers?

Response: No test was required. The “incorrect parameters” that we mention in this sentence refer to equations where the parameters are obviously incorrect: e.g. the predicted biomass of a tree might be only a few grams, even though the diameter is 10 cm. We have used the term “implausible” to make this clearer.

Lines 212-214: Issues arising from the additivity property should be reported.

Response: We avoided constructing additive equations to avoid these issues. The calculation of pseudo-observations while considering the MSE of the equations was used to provide a more accurate estimate of the error, this has been described in the methods section.

Line 222: the performance of the equation for the studied species should be reported.

Response: The MSE has been added for each equation in Figure B.1. The equations have been provided in Figure B.1 and the raw data is provided in Table A.1.

Line 231: the word published should be changed to compiled.

Response: This has been changed.

Lines 241-243: Not clear; rephrase.

Response: This sentence has been reworded. “Equations with large numbers of pseudo-observations (due to large diameter ranges) would bias the data set towards the values of the age and stand structural variables associated with the given equation.”

Lines 242-244: It is not clear how the age, basal area and trees per hectare calculated for the pseudo-observations.

Response: A new sentence has been added to Figure 1, which this comment is referring to.

Lines 248-249: the authors fail to distinguish between generalized equations (derived from pseudo-observations) and pooled equations (derived from compiled raw dataset).

Response: As explained above, the definition used in the paper is the same as that used in previous studies such as Muukkonen 2007, de-Miguel et al 2014, Jenkins et al., 2003, Chojnacky et al., 2014, Pastor et al., 1984, Paul et al 2016.

Lines 272-290: References about the validity of the approach for generating pseudo-observations should be reported. Why Monte Carlo simulation (or any other Data Generating Process) was not used?

Response: We agree that additional methodological comparisons would be an interesting use for the data set. This was beyond the scope of this study but could be done using the freely available and open access data set. Also, in the paragraph before these lines, the approach is described and we cite a paper that validated this approach (de-Miguel et al 2014). The first reviewer also commented that the validation approach used was already almost “overkill”.

Lines 306-316: A comparison of linear regression on log-transformed data to non-linear models could provide an “added-value” in the analysis. Rejecting a priori nonlinear models should be justified by the authors.

Response: We agree that it would be interesting to use the data generated in our review to test different methods of fitting the equations. This was not the objective of this paper and did not fit within this paper but could be done with the data, which will be freely available if the paper is published. The use of nonlinear models was not rejected, we just chose to use linear models.

Lines 410-411: already reported in line 169.

Response: Of the 5 numbers provided here (numbers of trees, species, equations, raw data sets and total data sets), only the latter was provided on line 169 (the species a few lines earlier), and it is provided here again because helps to put the others into context.

Line 419: “…the sample sizes were small..” the sample size is related to the approach used. If Bayesian analysis is used then 30 trees is large sample (see Zapata-Cuartas et al. 2012 For. Ecol. Manage. 277:173–179).

Response: This is a good point. We described it as small because it is small in relation to the approach we used.

Lines 422-426: what if non-linear regression or reduced major axis is used? What is the value of the derived exponent?

Response: As mentioned above, comparisons of methods was beyond the scope of this study, but we have provided all the data required to do this in future studies.

Lines 443-445: is it implied that in (534-317)/534 = 40% of the equations these variables did not influence the biomass-diameter relationship? If this is the case then it should be reported.

Response: No, we specifically did not say this because some of the 40% of equations could not contain the variables because some studies simply didn’t provide them for the equations they published.

Lines 448-449: Not clear; rephrase.

Response: This sentence has been reworded.

Lines 452-453: how much of biomass variability is explained by these variables?

Response: This was not specified because it varies between species and components.

Lines 515-518: already reported in lines 135-153.

Response: This sentence has been deleted.

Fig. B1: what do sigmas and t stand for?

Response: They refer to the parameters in Equation 2. This has now been stated.

Fig. B4: It is not common to present graphs without data or predictions. Empty figures should be removed.

Response: These figures have been simplified. We used the same format for all of the figures showing the pseudo-observations and when no data were available for the species-component combination, we retained the empty panels to make this clear. This can be changed even further but we thought it would be clearer this way.

Generalized biomass and leaf area allometric equations for European tree species incorporating stand structure, tree age and climate

David I. Forrester[1,2*], I. Heloise H. Tachauer[3], Peter Annighoefer[4], Ignacio Barbeito[5], Hans Pretzsch[3], Ricardo Ruiz-Peinado[6,7], Hendrik Stark[2], Giorgio Vacchiano[8], Tzvetan Zlatanov[9], Tamalika Chakraborty[10], Somidh Saha[2,11], Gudeta W. Sileshi[12]

[1]Swiss Federal Institute of Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland
[2]Chair of Silviculture, Albert-Ludwigs-Universität Freiburg, Germany
[3]Chair for Forest Growth and Yield Science, Technische Universität München, Germany
[4]Abteilung Waldbau und Waldökologie der gemäßigten Zonen, Georg-August-Universität Göttingen, Germany
[5]LERFoB, AgroParisTech, INRA, F-54000, Nancy, France
[6]Department of Silviculture and Forest Systems Management, INIA-CIFOR Forest Research Centre
[7]Sustainable Forest Management Research Institute (IuFOR), University of Valladolid & INIA
[8]European Commission, Joint Research Centre, Directorate D, Sustainable Resources - Bio-Economy Unit, Ispra, Italy
[9]Department of Silviculture, Forest Research Institute, Sofia, Bulgaria
[10]Chair of Site Classification and Vegetation Science, Albert-Ludwigs-Universität Freiburg, Germany
[11]Institute for Technology Assessment and Systems Analysis, Karlsruhe Institute of Technology, Germany
[12]Plot 1244 Ibex Hill, Lusaka, Zambia

*Corresponding author: WSL Swiss Federal Institute of Forest, Snow and Landscape Research, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland, Email: david.forrester@wsl.ch, Phone: +41 (0) 44 739 2869

Abstract

Biomass and leaf area equations are often required to assess or model forest productivity, carbon stocks and other ecosystem services. These factors are influenced by climate, age and stand structural attributes including stand density and tree species diversity or species composition. However, such covariates are rarely included in biomass and leaf area equations. We reviewed the literature and built a database of biomass and leaf area equations for 24 European tree species and 3 introduced species. The final dataset contained 973 equations. Most of the equations were site-specific and therefore restricted to the edaphic, climatic and stand structural conditions of the given site. To overcome this limitation, the database was used to develop regional species-specific equations that can be used in a wide range of stands and to quantify the effects of climate, age and stand structure on biomass or leaf area. The analysis showed considerable inter- and intra-specific variability in biomass relationships. The intra-specific variability was related to climate, age or stand characteristics, while the inter-specific variability was correlated with traits such as wood density, specific leaf area and shade tolerance. The analysis also showed that foliage mass is more variable than stem or total aboveground biomass, both within and between species, and these biomass components have contrasting responses to age and changes in stand structure. Despite the large number of published equations, many species are still not well represented. Therefore, generic equations were developed that include species-specific wood density instead of species identity. Further improvements may be possible if future studies quantify the stand structure of individual tree neighbourhoods instead of using the stand means for all trees sampled with the given stand.

Keywords: allometry; forest growth; pseudo-observations; shade tolerance; specific leaf area; wood density

1. Introduction

Allometric relationships are critical for quantifying many aspects of ecology and forestry including the prediction of tree and stand variables to assess productivity, carbon stocks and other ecosystem services at the tree, stand, landscape or regional levels (Henry *et al.*, 2013;

Chave *et al.*, 2014; Paul *et al.*, 2016). They are also required when quantifying or modelling forest functioning, such as how light, water, nutrient and carbon pools and fluxes respond to changes in climate or management.

Allometric relationships are often expressed in the form of Equation 1, implying a 1% change in variable X will result in a b% change in variable Y.

$$Y = aX^{b_{y,x}} \tag{1}$$

The value of the exponent b has been hotly debated (Sileshi, 2014) and hypothesised to relate to mechanical constraints that prevent trees from buckling (Greenhill, 1881; McMahon, 1973), hydraulic constraints (Ryan *et al.*, 2006) and biophysical constraints. Contributions regarding the biophysical constraints include geometric scaling (Yoda *et al.*, 1963; Gorham, 1979; Pretzsch *et al.*, 2012), which suggests proportionality between different linear dimensions; linear tree dimensions (e.g., diameter) are related to quadratic or area-related dimensions (e.g., leaf area) as linear $\propto$ quadratic$^{1/2}$ and to cubic variables (e.g., biomass) as linear $\propto$ cubic$^{1/3}$ or quadratic $\propto$ cubic$^{2/3}$. In contrast, the metabolic scaling theory describes resource distribution along hierarchical branching networks (West *et al.*, 1999; West *et al.*, 2009) and predicts that $b_{\text{biomass, diameter}} = 8/3$, $b_{\text{leaf area, diameter}} = 4/3$ (Pretzsch *et al.*, 2012). However, b is usually not invariant for these relationships and the frequency distribution of b is not necessarily centred on the value of b predicted by the geometric or metabolic scaling theories (Coomes, 2006; Pretzsch, 2006; Ducey, 2012; Lines *et al.*, 2012; Pretzsch and Dieler, 2012; Pretzsch *et al.*, 2012; Pretzsch *et al.*, 2013; Sileshi, 2014). Therefore, while the general allometric exponents may be useful for rough scaling they are less useful for modelling stand growth dynamics or for developing biomass and leaf area equations to upscale from tree measurements.

The variability in the exponent *b* is related to the fact that allometric relationships reflect current and past environmental conditions and provide information about within-tree carbon partitioning, which affects a trees' ability to acquire and compete for resources. Therefore, allometric relationships between diameter and biomass (foliage, stems or roots) or leaf area can vary with age (Wirth *et al.*, 2004; Genet *et al.*, 2011; Shaiek *et al.*, 2011), stand density (Monserud and Marshall, 1999), species mixing (Laclau *et al.*, 2008) and site characteristics (Wirth *et al.*, 2004; Russell *et al.*, 2015). As a result, equations developed using trees sampled from a single stand may be unbiased and precise for that situation but they are unlikely to be suitable for other ages or stands that differ in structure, climate or site characteristics (Muukkonen, 2007). Despite this, variables describing age, site and stand structural characteristics such as density, species composition or diversity are rarely included in biomass equations (Zianis *et al.*, 2005) because this would require a larger sample of trees from a range of ages and site conditions.

In a recent review, only about 24% of equations were found to contain more than one independent variable, usually diameter (Henry *et al.*, 2011). Nevertheless, for some species there are already many published biomass equations (Zianis *et al.*, 2005) and the suitability of each equation for use in different stands can be determined, for example, by sampling some trees and comparing the measured biomass with the biomass predicted by the published site-specific equations (Freese, 1960; Pérez-Cruzado *et al.*, 2015). However, this requires destructive biomass sampling in each target stand. It also requires that there is a published equation suitable for that stand, for which the likelihood declines as the number of published equations declines. An alternative approach is to use all of the published site-specific equations to develop new "regional" allometric equations that include independent variables such as climate, age, stand density and any other important site characteristics.

Several studies have developed regional species-specific or even generic (species independent) biomass equations (Pastor *et al.*, 1984; Wirth *et al.*, 2004; Lambert *et al.*, 2005; Case and Hall, 2008; Seidl *et al.*, 2010; Shaiek *et al.*, 2011; Chave *et al.*, 2014; de-Miguel *et al.*, 2014; Paul *et al.*, 2016). These often combine raw data from many different studies, but such data do not exist for many species or regions, or biomass data that was used to develop site-specific equations has been lost or is unavailable. Therefore, some studies have used pseudo-observations calculated from published equations, such as predicted biomass values for each 1-cm or 5-cm diameter class (Jenkins *et al.*, 2003; Muukkonen, 2007; Chojnacky *et al.*, 2014) or a given number of pseudo-observations between the range of diameters sampled to produce the given site-specific equation (Pastor *et al.*, 1984). Regardless of the approach used, most of the resulting regional or generic equations have included only tree-level variables (e.g., diameter, height) and/or species-level variables (e.g., wood density) and therefore average out or group the variability in tree biomass that might otherwise be explained by age, climate, soils, stand density or species mixing (Wirth *et al.*, 2004; Chojnacky *et al.*, 2014; Weiskittel *et al.*, 2015). Such variables could facilitate the development of biomass equations that are applicable to a wider range of sites and stands, and can be used to examine the effects of these factors on stand growth and biomass stocks.

Despite the large number of published equations, many European species are still not well represented. Therefore, the first objective of this study was to develop a database containing biomass and leaf area equations for 24 European tree species and 3 introduced species (*Pseudotsuga menziesii, Robinia pseudoacacia* and *Prunus serotina*) that are currently considered important by European foresters. The review of the literature resulted in a total of

973 equations, including raw data sets obtained from tables in publications or from our previous work. These data were used to test the hypotheses that: (1) foliage or branch mass are more variable than stem, coarse root or total aboveground biomass; (2) age, trees per hectare, basal area and climate all influence the relationships between tree diameter and biomass or leaf area; (3) these variables have contrasting effects on different biomass components; (4) there are significant differences between species in terms of their response to age, trees per hectare, basal area and climate, and these differences vary in relation to traits such as specific leaf area, wood density and shade tolerance. Our second objective was to develop regional equations for each species, or species group, and each biomass component or leaf area, which include the independent variables age, trees per hectare, basal area, mean annual precipitation or mean annual temperature and can therefore be used in a wider range of forest types.

2. Material and methods

2.1 Selection of equations

A literature search was used to find biomass and leaf area equations for 27 species (and several species groups) summarised in Table 1. For most species the equations included a wide range of sites across the current species distributions within Europe and are therefore assumed to be representative of the given species within Europe. These ranges, for each species, are indicated in Table 2 as ranges of stand and site characteristics. Species selection was based on the availability of equations, but also reflects the economic importance of the species. Equations were excluded when the diameter range or number of sample trees was not provided. They were also excluded when they did not include diameter at 1.3 m as an independent variable. Equations based on root collar diameter (often 0.1 m) were also excluded. Height-diameter relationships for a given species can vary with age and site characteristics (Watt and Kirschbaum, 2011). Therefore to avoid adding bias that may result from height predictions, equations with height were only used when no diameter-only equations were available from the same study. If height was included in the equation, it was predicted using the diameter-height equation provided in the given study or published diameter-height equations (Merganič *et al.*, 2011; Pretzsch *et al.*, 2013; Özcelik *et al.*, 2014). The predicted range in height was forced to match the height range reported for the sample trees by multiplying all values by a constant, thereby preserving the shape of the height-diameter relationship. Equations with incorrect parameters (identified by their implausible biomass or leaf area predictions) were also not used. While several general equations have been developed for some of the species in Table 1 (Wirth *et al.*, 2004; Muukkonen, 2007), we did not use those general equations and instead aimed to use the equations that were used to develop those general equations.

Many different biomass components can be defined for trees, but for a given component the definitions can vary between studies. For example, roots may be sampled down to minimum diameters of 1 cm, 0.5 cm or 0.2 cm (Jenkins *et al.*, 2003; Wirth *et al.*, 2004), stem can begin at the base of the tree and extend up to a minimum diameter of 5 cm (Canadell *et al.*, 1988) or it may only include the straight part of the stem (Annighöfer *et al.*, 2012). This will add an error to the regional or generic equations that are produced. In addition to leaf area, we searched for equations that predict the biomass components of foliage, branch wood (total, live or dead, all including the bark), stem (including wood and bark), stem and branch (including stem + branch wood and bark), aboveground (including stem + branch wood and bark + foliage) and roots. Where the studies divided the components even further (e.g., stem wood and stem bark), we used both equations to calculate the stem or we combined the aboveground components to get aboveground mass.

When a study provided equations for leaf mass but not leaf area, or vice versa, the specific leaf area (SLA, $m^2\ kg^{-1}$) was used to predict the leaf area from the leaf mass, or vice versa. Here, leaf area strictly refers to one-sided projected area. SLA can decline with tree age and this relationship was obtained for each species using published SLA-age data that was fitted to Equation 2.

$$SLA(t) = SLA_1 + (SLA_0 - SLA_1)e^{-(ln2)(t/t_{SLA})^2} \tag{2}$$

where SLA_0 and SLA_1 are the values of SLA in young and older stands, respectively, and t_{SLA} is the age at which $SLA = ½(SLA_0 + SLA_1)$ (Sands and Landsberg, 2002). The SLA data are provided as supplementary information Table A.1 in the data paper associated with this study (Forrester *et al.*, in review) and SLA-age relationships are shown in Figure B.1 of appendix B.

2.2 Calculation of pseudo-observations

The compiled equations were used to generate pseudo-observations for use in fitting the regional allometric models. Although raw data are strongly favoured above pseudo-observations, raw data were not available for many of the equations in this study. Restricting the data set only to raw data would therefore have severely restricted the ranges of stand conditions, climates, ages and tree sizes available and thus greatly limited our ability to examine the effects of these factors on biomass relationships. Therefore, pseudo-observations were generated for many of the published equations. The equations and their pseudo-observations are provided in Table A.2 of Forrester *et al.* (in review). Several approaches have been used to generate pseudo-observations in biomass studies. One method involves calculating one or more pseudo-observations per 1-cm (or n-cm) diameter class (Muukkonen, 2007; Chojnacky *et al.*, 2014; de-Miguel *et al.*, 2014). Another method involves the use of an equal number of pseudo-observations per equation (Pastor *et al.*, 1984; Jenkins *et al.*, 2003). In this study, equal numbers of pseudo-observations were used for each equation instead of pseudo-observations at each n-cm class. This was necessary because some equations had a much larger diameter range than others and they would have had a correspondingly higher number of pseudo-observations. Equations with large numbers of pseudo-observations (due to large diameter ranges) would bias the data set towards the values of the climate, age and stand structural variables associated with the given equation. The distributions of diameter, age, basal area and trees per hectare are shown in Figure 1. The median sample size for all equations was 33. Therefore, 33 pseudo-observations were calculated for each equation that were evenly distributed between the smallest and largest diameter used to develop the given equation. This is larger than the n in some previous studies (e.g., Pastor *et al.*, 1984; Jenkins *et al.*, 2003; Chojnacky *et al.*, 2014). The larger n was used to reduce the possibility of overestimating the parameters (and hence biomass), which can result from small sample sizes (Duncanson *et al.*, 2015). Where studies included the raw data in published tables or the raw data were provided by the authors, the raw data were used instead of pseudo-observations (10%; 105 raw data sets but 868 equations).

As is often the case, there was a bias towards the smaller diameter classes (Figure 1h). To reduce the influence of tree size distribution, Duncanson *et al.* (2015) binned data by calculating the median value of the dependent variable in equally spaced independent variable bins. However, this approach can underestimate the variability of the equations (Jucker *et al.*, 2017). Furthermore, binning could not be done in this study because there was more than one independent variable and binning would require a complete matrix of diameter, age, climate and stand variables, which does not exist in this data set.

While the pseudo-observation datasets could potentially be used to examine the variability between equations, they do not include the variability within a given equation (or site) and therefore will lead to underestimation of errors of regional equations (Case and Hall, 2008; de-Miguel *et al.*, 2014). The within-equation error includes the uncertainty of the parameter estimates as well as the residual error of the equation (Breidenbach *et al.*, 2014). The latter can be approximated by randomly generating pseudo-observations by assuming a normal distribution of residuals with a mean of zero and the variance (mean square error of the residuals; MSE) reported for the given equation (de-Miguel *et al.*, 2014). This recovers the within- and between-equation variability of the original field measurements (de-Miguel *et al.*, 2014). In order to illustrate this we show the pseudo-observations calculated from all (four) published equations for estimation of aboveground biomass of the first species listed in our database, *Abies alba* (Figure 2).

It was possible to obtain usable variance values for only 14% of the published equations and this includes MSE values that were calculated by fitting Equation 3 to each of the raw data sets that were included in the database. This scarcity of model variance information resulted because many studies did not report model errors or it was not clear what the error terms actually were. For example, the names given to the errors varied widely and we generally used values that were called standard deviation, MSE, Root MSE, Sy,x, se, residual error, residual standard error, residual deviation and mean error. To ensure compatibility, we only used the MSE values from equations fitted to ln-transformed biomass components or leaf area. The MSE values for the 86% of equations for which we did not have a model error value were estimated based on the observation that the MSE values appeared to vary between biomass components and also between broad-leaved and coniferous species. Therefore, we filled in the missing MSE values using the mean MSE values for each of these combinations (Figure B.2). The MSE was only available for 10 leaf area equations and therefore the mean MSE was used for all other leaf area equations. From here on, Dataset 1 refers to the dataset with pseudo-observations that ignore within-equation variability, and Dataset 2 refers to the dataset containing the pseudo-observations that were calculated by taking into account the MSE values. Both datasets contained the raw biomass or leaf area data (105 of the 973 equations). The regional equations were fitted using Dataset 1, while Dataset 2 was only used to estimate a more realistic MSE of the regional equations ($\mathrm{MSE}_{\mathrm{real}}$).

2.3 Candidate independent variables

The selection of candidate independent variables for biomass equations should be based on clear hypotheses and theory (Sileshi, 2014). Furthermore, when using published equations or published data sets, the independent variables are restricted to variables that were described in the publications and that are measured in a standard/comparable way in each study. This therefore restricts the variables to stem diameter, tree age, stand basal area, trees per hectare, latitude, longitude (and therefore climate) and species proportions. Tree size, in terms of diameter or height, is a clear candidate variable and is included in the vast majority of equations. It indirectly provides information about the trees' history. Age is another strong candidate variable because many studies have shown that allometry and biomass partitioning are affected by age (Wirth *et al.*, 2004; Litton *et al.*, 2007; Genet *et al.*, 2011; Shaiek *et al.*, 2011). The number of trees per hectare and mean diameter are allometrically related as described by the self-thinning law (Reineke, 1933; Yoda *et al.*, 1963), so trees per hectare is also a good candidate variable. For a given number of trees per ha, a stand can have a wide range of basal area values. Therefore, basal area or mean diameter may also be required to more adequately describe the stand structure. Biomass relationships may also vary between sites due to differences in climate and soils or because the provenances and genotypes of the given species vary (Litton *et al.*, 2007; Poorter *et al.*, 2012; Reich *et al.*, 2014). Mean annual temperature and mean annual precipitation were predicted from latitude and longitude using the WorldClim database (Hijmans *et al.*, 2005). Provenance and genotype information are often not provided and it was assumed that latitude may approximate some of these effects, and it is provided in nearly all studies.

Variables such as tree species proportion are too specific to be useful unless the equations are only designed for a specific forest type. This is because there is an extremely high number of potential species combinations and proportions where any given species can occur, and this variability in species composition cannot be quantified using a single stand variable (Forrester *et al.*, in press). Therefore, alternative variables have been proposed that quantify the horizontal (basal area, trees per ha) or vertical stand structure (tree dominance status) and that are weighted, if necessary, by traits assumed to indicate species-specific competitive ability, such as wood density, specific leaf area and shade tolerance (Forrester *et al.*, in press).

2.4 Regional equations

Regional equations (Equations 3-21) for each species and component were fitted using linear mixed models. The pseudo-observations (or raw data) belonging to a given published equation were given a unique number to identify the published equation from which they were obtained (this number ranged from 1 to 982). This unique equation number was included as the random variable to account for within study correlation (pseudo-observations from a given published equation are not independent of each other). The fixed effects were the independent variables including tree diameter (*d* in cm) at 1.3 m, age (*A* in years), stand basal area (*BA* in m^2 ha^{-1}), the number of trees per hectare (*TPH*), latitude (in °), mean annual temperature (° C) and mean annual precipitation (mm). The biomass, leaf area, diameter, age and *TPH* were all ln-transformed to reduce homoscedasticity and to produce linear relationships with biomass or leaf area. The following functional forms were compared:

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \varepsilon \tag{3}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 BA + \varepsilon \tag{4}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(A) + \varepsilon \tag{5}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(TPH) + \varepsilon \tag{6}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 Latitude + \varepsilon \tag{7}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 Prec + \varepsilon \tag{8}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 Temp + \varepsilon \tag{9}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 BA + \beta_3\ln(A) + \varepsilon \tag{10}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 BA + \beta_3\ln(TPH) + \varepsilon \tag{11}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 BA + \beta_3 Latitude + \varepsilon \tag{12}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 BA + \beta_3 Prec + \varepsilon \tag{13}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2 BA + \beta_3 Temp + \varepsilon \tag{14}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(A) + \beta_3\ln(TPH) + \varepsilon \tag{15}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(TPH) + \beta_3 Prec + \varepsilon \tag{16}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(TPH) + \beta_3 Temp + \varepsilon \tag{17}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(TPH) + \beta_3 Latitude + \varepsilon \tag{18}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(A) + \beta_3 Prec + \varepsilon \tag{19}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(A) + \beta_3 Temp + \varepsilon \tag{20}$$

$$\ln(Y) = \ln\beta_0 + \beta_1\ln(d) + \beta_2\ln(A) + \beta_3 Latitude + \varepsilon \tag{21}$$

where ε is an error term that is assumed to be normally distributed with a mean (μ) of 0 and standard deviation of σ. In order to correct for the bias that results from back-transforming predictions of *Y*, a correction factor (CF) was calculated from the mean of the pseudo-observations divided by the mean of the values predicted by the regional equation (Snowden, 1991).

To compare the effect of each independent variable on each biomass component and to determine whether these responses were related to species traits, Equations 22-24 were used after standardising all independent variables by subtracting the mean from the variable value and then dividing it by the standard deviation of the given variable. The standardised independent variables then each have a mean (μ) of 0 and a standard deviation (σ) of 1. If the independent variable was ln-transformed, then the transformed values were standardised to μ = 0 and σ = 1. Therefore, the slope parameters (hereafter called beta weights) in Equations 22-24 were used to provide the effect size for each variable. Beta weights (standardised coefficients) are scale free, and quantify how many standard deviations the dependent variable (*Y*) will change per standard deviation increase in the independent variable (*X*). Thus they can be used as effect size estimates (Peterson and Brown, 2005; Nieminen *et al.*, 2013). In this analysis, the effect size of all independent variables was compared in order to identify those that have the largest influence on biomass or leaf area.

$$\ln(Y) = \ln\beta_0 + \beta_1 \ln(d) + \beta_2 BA + \beta_3 \ln(A) + \beta_4 \ln(TPH) + \beta_5 Prec + \beta_6 Temp + \beta_7 SLA + \beta_8 \ln(d) \times SLA + \beta_9 BA \times SLA + \beta_{10} \ln(A) \times SLA + \beta_{11} \ln(TPH) \times SLA + \beta_{12} Prec \times SLA + \beta_{13} Temp \times SLA + \varepsilon \quad (22)$$

$$\ln(Y) = \ln\beta_0 + \beta_1 \ln(d) + \beta_2 BA + \beta_3 \ln(A) + \beta_4 \ln(TPH) + \beta_5 Prec + \beta_6 Temp + \beta_7 \rho + \beta_8 \ln(d) \times \rho + \beta_9 BA \times \rho + \beta_{10} \ln(A) \times \rho + \beta_{11} \ln(TPH) \times \rho + \beta_{12} Prec \times \rho + \beta_{13} Temp \times \rho + \varepsilon \quad (23)$$

$$\ln(Y) = \ln\beta_0 + \beta_1 \ln(d) + \beta_2 BA + \beta_3 \ln(A) + \beta_4 \ln(TPH) + \beta_5 Prec + \beta_6 Temp + \beta_7 ST + \beta_8 \ln(d) \times ST + \beta_9 BA \times ST + \beta_{10} \ln(A) \times ST + \beta_{11} \ln(TPH) \times ST + \beta_{12} Prec \times ST + \beta_{13} Temp \times ST + \varepsilon \quad (24)$$

where *d, BA, A, TPH, Temp, Prec* and *SLA* are as defined earlier for Equations 3-21, ρ is the wood density (oven dry mass per fresh volume; g cm^{-3}) (Zanne *et al.*, 2009) and *ST* is the shade tolerance from Niinemets and Valladares (2006) (Table B.3). *SLA* is the mean *SLA* for the given species obtained from the published values in Table A.1 of Forrester *et al.* (in review). Equation 23, which contains wood density, was also fitted to Datasets 1 and 2 without standardising the independent variables. This was done to provide generic (species independent) models that can be applied to species that are not well represented in the data base. These generic equations were fitted to all data in Dataset 1. Equations 3-21 were fitted using the *nlme* package in R (Pinheiro et al., 2013) and using the REML method, which maximises the restricted log-likelihood.

2.5 Verification and quality of fit

Verification or validation of the final models based on Equations 3-21 was done based on the K-fold or leave-one-out methods of cross-validation (Arlot and Celisse, 2010; Picard *et al.*, 2012) whereby Equations 3-21 were fitted to the pseudo-observations (or raw data sets) from all but one of the published equations, and then the predictions of the resulting regional model were compared with the pseudo-observations (or raw data sets) of the excluded published equation. This was repeated until all of the published equations had been used as validation data.

To check for collinearity between independent variables, variance inflation factors (VIF) were calculated using the *corvif* function in the *AED* package in R (Zuur *et al.*, 2009). VIF above 5 indicate high multicollinearity between independent variables (Sileshi, 2014). While many biomass studies include variables that are highly correlated (e.g., diameter and height), we avoided this to ensure that the parameter estimates represented causal relationships as closely as possible.

The uncertainty of the model parameters was examined using the percent relative standard errors PRSE (Equation 25; Sileshi, 2014).

$$PRSE\ (\%) = 100\frac{parameter\ standard\ error}{parameter\ estimate} \tag{25}$$

PRSE values more than 50% were considered unreliable. Akaike's information criterion (AIC) was used to examine whether any of the regional equations were over-parameterised. However, Equations 3-21 were often fitted to different subsets of data because different published equations provided different combinations of the independent variables. Therefore the Bayesian information criterion (BIC) was also used because it takes the sample size into account.

The precision of the regional equations was quantified using the square root of the mean square error (RMSE) using Equation 26 calculated from the ln-transformed data.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \tag{26}$$

where O are the observed values and P are the predicted values. The precision was also quantified using the mean absolute percentage error (MAPE) using Equation 27 and the back-transformed predicted values that had been corrected for the bias using the correction factor (CF) described above from Snowden (1991).

$$MAPE\ (\%) = \frac{100}{n}\sum_{i=1}^{n}\frac{|O_i - P_i|}{O_i} \tag{27}$$

The bias was quantified using Equation 28 (Sileshi, 2014), also using the back-transformed predicted values that had been corrected for the bias using the CF described above from Snowden (1991).

$$Bias\ (\%) = \frac{100}{n} \sum_{i=1}^{n} \frac{P_i - O_i}{O_i} \tag{28}$$

Linear regression of the observed (y-axis) against the predicted values (all ln-transformed) were used to examine how well the models fit the observed data. This was used to test whether the slope was significantly different from 1 and whether the intercept was significantly different from 0 (Sileshi, 2014) by calculating 95% confidence intervals (standard error of parameter × 1.96) for the slope and intercept. During the validation, the mean PRSE, RMSE, MAPE and Bias were calculated from all the K-fold validations for the given general equation. All analyses were performed using R software version 3.3.1 (R Core Team, 2016).

3. Results

The final dataset contained pseudo-observations for 27 species from 868 equations and 105 raw data sets (973 in total) (Table 1). This included 60,294 biomass or leaf area samples from the reviewed studies. The equations covered a broad range of stand characteristics, with basal areas ranging from < 5 to > 75 $m^2\ ha^{-1}$ and stand densities ranging from < 200 to 70,000 trees per hectare (Figure 1a, b). However, there was a clear skew towards the smaller tree sizes and younger ages (Figure 1c, h). The equations were distributed across the latitudinal range of Europe with a peak in the middle, at latitudes of about 45 - 50° (Figure 1d). While many of the sample sizes were small (< 30 trees) there were also many that contained > 50 trees (Figure 1e).

All validation equations and final equations are listed in Table A.4 of Forrester *et al.* (in review). The final equations that were valid (e.g., $P < 0.05$ for all variables, VIF < 5 and PRSE < 50%) are listed in Table A.5 of Forrester *et al.* (in review). Focusing only on equations in the form of Equation 3, the confidence intervals for the exponent β_1, averaged across all species, often contained the value predicted by the metabolic scaling theory (although it was usually close to the upper limit) and often did not contain the value predicted by the geometric scaling theory (Table 3).

3.1 Intra-specific variability

In agreement with our first hypothesis, the foliage mass, and to a lesser extent branch mass, was generally more variable for a given diameter than stem mass (Figure 3, see also Figures B.3-B.25). This is indicated by the greater mean $\mathrm{RMSE}_{\mathrm{real}}$ for foliage (0.42) and total branch mass (0.54) compared with stem (0.33), stem + branch mass (0.28) and aboveground mass (0.18), with intermediate RMSE for root mass (0.37) (from Table A.5 of Forrester *et al.*, in review).

Using the RMSE when calculating the pseudo-observations, as shown in Figure 2, resulted in $\mathrm{RMSE}_{\mathrm{real}}$ that were on average 76% larger (0.404 compared with 0.230) than those calculated when the RMSE of the individual equations was not considered when calculating the pseudo-observations.

For many species, there were enough equations available to test whether the biomass components were influenced by age, basal area, trees per ha, climate or latitude. These variables were significant in 387 out of the 576 regional equations listed in Table A.5 of

Forrester *et al.* (in review), which is consistent with our second hypothesis that these variables influence diameter-biomass relationships. The parameter estimates in Table A.5 show that climate, age and the stand variables have contrasting effects on different biomass components and species. Stem, stem + branch and aboveground mass often tended to increase (for a given diameter) with all of these independent variables (except climate and latitude where effects were more variable). In contrast, foliage mass and leaf area sometimes declined (for a given diameter) with increasing age, basal area and trees per ha. For most species or species groups (24 out of 31) equations that contained diameter only (Equation 3) had, on average, similar or higher RMSE, MAPE and Bias than the other equations (Table 4). There was no clear trend as to whether age, basal area, trees per ha, precipitation or temperature was the best independent variable to add to the equations.

3.2 Inter-specific variability

Wide inter-specific variability was observed in biomass components, especially foliage mass (Figure 4). The biomass components for which there was a lot of intra-specific variability also showed a lot of inter-specific variability. For example, for a diameter of 60 cm, the predicted foliage mass (kg per tree) varied by 767% from about 15 kg in *C. betulus* to 130 kg in *P. abies*, while stem mass varied by 160% from about 1000 kg in *P. pinaster* to 2600 kg in *C. betulus*.

The inter-specific variability in the biomass – diameter relationship was generally high for foliage mass and lower for stem mass (Figures 4 & 5). Examination of the effect sizes (beta-weights) showed that diameter always had the greatest influence on biomass (Figures 6 & 7). Increasing basal area generally reduced leaf area, leaf mass and branch mass. Increasing age generally increased stem mass but reduced root mass; for a given diameter older trees had less root mass than younger trees (Figure 6). Foliage biomass or leaf area increased with temperature or precipitation for low trait values (SLA, wood density and shade tolerance) but decreased for high trait values. For the other component-independent variable combinations, there were often contrasting responses ranging from increases to decreases in biomass. This is consistent with our third hypothesis that stand characteristics and age have contrasting effects on different biomass components.

Wood traits (wood density, SLA or shade tolerance) were significantly correlated with inter-specific variability for more than half of all possible combinations of biomass components and independent variables, consistent with our fourth hypothesis. For example, the effect of diameter on leaf mass declined as all wood traits increased (Figure 6).

Equation 23 was also fitted for all species before the data were standardised. This enabled the development of a generic equation for each biomass component and leaf area that could be used for any species in our data set. However, it should be noted that the species-specific regional models are expected to be the most accurate and the generic model fitted to Equation 23 is intended to be used only for rarer species until those species are better represented.

4. Discussion

4.1 Intra-specific variability

For a given diameter, there was considerable intra-specific variability and this was greater for the shorter lived components such as foliage and branches than for longer lived components like stems. This finding is consistent with our first hypothesis and with previous studies

(Wirth *et al.*, 2004; Saint-André *et al.*, 2005; António *et al.*, 2007; Genet *et al.*, 2011; Xiang *et al.*, 2011; Clough *et al.*, 2016; Xiang *et al.*, 2016). Shorter lived tissues such as foliage, branches and fine roots may turn over more readily and rapidly as conditions become more or less favourable thereby making them more variable for a given diameter than structural support components such as stems.

It follows therefore that the intra-specific variability not only reflects tree age and genetics but that it can also indicate how trees have responded to previous growing conditions. When biomass equations are developed, it is common to average out this variability or to group it all together (Sileshi, 2014; Weiskittel *et al.*, 2015) by developing regional or generic models that include a single independent variable (e.g., dbh or a compound variable that includes dbh; Chave *et al.*, 2014) or by assigning the variability to the random variables in mixed models (de-Miguel *et al.*, 2014). Alternatively, this study shows that such intra-specific variability can be partitioned to some of the independent variables that are causing this variability such as age, basal area, trees per ha, climate and latitude, instead of to random effects, consistent with our second hypothesis. An advantage of this approach is that the effects of climate, age, stand basal area and trees per ha on biomass stocks can be predicted and the size and direction of their effects are clearly evident from the model parameters.

It is worth noting that the difference in biomass predictions for a given diameter often differed by more than 50% for the same species and component, and sometimes by more than 100% (e.g., Figure 3). This highlights the risk of applying species-and-site specific published biomass equations to stands where they were not developed. Even for the regional equations developed in this study, it is critical to note the ranges in all independent variables, because these should include the values of the target stands to avoid extrapolating beyond the range of data used to develop the equations.

Not only the magnitude, but also the direction of the response to these independent variables differed between components, consistent with our third hypothesis and with previous meta-analyses (Litton *et al.*, 2007; Poorter *et al.*, 2012). For example, foliage mass declined with increasing basal area while stem mass did not change (Figure 6). In relatively open and low basal area stands, biomass partitioning can be more to foliage to maximise light absorption because competition for soil resources is relatively low and there is less incentive for height (and stem) growth to outcompete neighbours. On the other hand, in dense stands there is less space for expanding crowns or fine-root systems and more biomass may be partitioned to the stems in order to grow taller to maintain a competitive position within the canopy (Poorter *et al.*, 2012). Stand density or species composition have also been found to influence biomass relationships in other studies (Monserud and Marshall, 1999; Wirth *et al.*, 2004; Laclau *et al.*, 2008). Roots were more variable in terms of their response to basal area, but it was not possible to divide this component into fine, medium and coarse roots, which prevents the patterns in root mass from being examined in detail.

The response to age was also consistent with previous studies that have found declines in foliage mass or leaf area and increases in stem mass as trees age (Porte *et al.*, 2000; Porté *et al.*, 2002; Saint-André *et al.*, 2005; Wutzler *et al.*, 2008; Genet *et al.*, 2011; Shaiek *et al.*, 2011). Age had a varied effect on root mass, consistent with previous work where central European *F. sylvatica* fine medium and coarse roots were not influenced by age (Genet *et al.*, 2011) but increased with age in another study (Wirth *et al.*, 2004; Wutzler *et al.*, 2008). Fine roots of *Eucalyptus* were also reduced with age for a given diameter (but not medium or coarse roots) (Saint-André *et al.*, 2005). The variability in this study is likely to reflect the variable definition of roots used in the reviewed studies and that different components (fine, medium and coarse roots) do not necessarily respond to age or stand structure in the same way, similar to the contrasting aboveground responses of stem mass versus foliage mass (Saint-André *et al.*, 2005).

4.2 Inter-specific variability

The considerable inter-specific variability in dbh – biomass (or crown or height) relationships has been found to correlate with species traits such as wood density, such that for a given diameter, biomass increases (or height declines) for species with higher wood density (Ketterings *et al.*, 2001; Chave *et al.*, 2005; Dietze *et al.*, 2008; van Breugel *et al.*, 2011; Ducey, 2012; Chojnacky *et al.*, 2014). This study shows that this pattern not only exists for dbh-biomass relationships, but that relationships between biomass and age, basal area, trees per ha and climate also vary between species in relation to the species-specific wood density, shade tolerance or specific leaf area. For example, increasing basal area reduced foliage mass more for low wood density species but increased it for high wood density species. This may reflect a lower tolerance to competition of the low wood density species (Kunstler *et al.*, 2016). The increase in stem mass with increasing age was greater for species with low wood density or low SLA. Since wood density is globally negatively correlated with a species' competitive effect on neighbours as well as its ability to tolerate competition (Kunstler *et al.*, 2016), this may reflect a higher partitioning of biomass to stems (or height) as trees age in response to competition. Similarly, the contrasting responses of height, crown length or crown diameter to stem diameter or stand structural characteristics were correlated with wood density or shade tolerance for 17 European species (Forrester et al. in press).

The inclusion of wood density in biomass equations can make them generic (e.g., Equation 15), so that they can also be used for rarer species (e.g., *Prunus avium, Tilia, Sorbus* or *Pinus cembra*) that are not well represented in datasets (Chave *et al.*, 2004; Dietze *et al.*, 2008; van

Breugel *et al.*, 2011). While the species-specific regional equations are likely to be more accurate for the range of stand conditions where they were developed, the generic equations where species identity is quantified as wood density may be a short term solution for rare species until they are better represented with biomass samples. Of the three generic equations used in this study (Equations 22-24), our focus was on Equation 23, which was based on wood density, because this trait is easier to obtain than shade tolerance (Equation 24) and is assumed to be less variable within a species than SLA (Equation 22). However, even wood density can vary between individuals of the same species (e.g., Niinemets, 2015).

4.3 Important considerations about the dataset

A source of error for the general equations produced in this study was the variability in definitions for a given biomass component. Different studies typically have different questions and therefore good reasons for varying the definitions. However, this would have inflated the error of the general equations.

The large dataset used for this study helps to make the equations in Tables A.5 and A.6 (in Forrester *et al.*, in review) more widely applicable. However, despite the wide range in conditions these equations are still subject to at least three sources of uncertainty. The first is the inherent variability in the population (e.g., RMSE), which we aimed to partition, at least partly, to stand variables, climate and age. The second is the sample size (van Breugel *et al.*, 2011; Roxburgh *et al.*, 2015). Many of the published equations were developed from small sample sizes (Figure 1e), which can lead to overestimates of biomass due to the overestimation of parameters (Duncanson *et al.*, 2015). While the median sample size was 33 observations for the published equations used in this study, previous studies have recommended at least 40-50 or more, depending on the type of tree species (Roxburgh *et al.*, 2015). Therefore, the pseudo-observations derived from the published equations that were based on small data sets may have been slightly overestimated and there is no way of knowing how this could have influenced the regional or generic equations developed in this study. The accuracy of biomass models is inherently dependent on the accuracy of the raw data, or in our case, the pseudo-observations.

The third source of uncertainty is related to the diameter distributions because biomass predictions are most precise near the mean stem diameter that was used to fit the equation (van Breugel *et al.*, 2011). If the equation is applied to a population with a different diameter distribution, then the precision will be reduced (Snowdon *et al.*, 2002; van Breugel *et al.*, 2011). That is, the uncertainty in the predictions depends on the characteristics of the population where the biomass equations are applied and not only on the sample size and characteristics of the population used to develop the equation (Roxburgh *et al.*, 2015). Therefore, when deciding which trees to sample for biomass, it is pertinent to consider the diameter distribution of the population where the final equation will be applied (Roxburgh *et al.*, 2015). While the diameter distributions of the pseudo-observations are shown in Figure 1h, and could also be calculated for each species-component using the data in Table A.2 of Forrester *et al.* (in review), the studies where the equations were obtained rarely described the diameter distributions of their sample trees or of the population where the equations were expected to be applied.

The MSE_{real} of the general models, which were developed after incorporating the error of the published equations (e.g., Figure 2) were probably overestimates of the error. This is because many studies did not provide information about age, basal area or trees per ha. Therefore, the variability in biomass resulting from those variables could not be partitioned to those variables in the fitted equations, thereby inflating their MSE_{real}. However, even when the stand variables are provided, they are means for the whole stand and are therefore not specific to individual sample trees. In reality, within-stand variability probably results in individual trees experiencing different neighbourhood basal areas, trees per ha and species compositions. Therefore, the MSE_{real} in Tables A.5 and A.6 of Forrester *et al.* (in review) are provided only as estimates of the model errors in case they are required for calculating the prediction errors. To improve the MSE_{real}, and to account for within-stand variability, future studies may benefit from using a neighbourhood index approach to quantify the structural variability within single stands so that instead of assigning the same value of each structural variable to all sample trees from a given stand, each tree is assigned the actual structural information for its neighbourhood (e.g., Thorpe *et al.*, 2010; Forrester, 2015). This would also make better use of the biomass data, which is typically very expensive and time consuming to collect relative to the collection of neighbourhood index information.

4.4 Next steps

Vertical stand structure has been shown to be one of the most important variables for predicting live-crown lengths in a range of European tree species (Forrester *et al.*, in press) and may also be worth considering when predicting leaf mass or leaf area. Vertical structure can be quantified using the relative height, calculated as the height of the sample tree divided by the mean height of the stand (or neighbourhood). In the same study the stand basal area variable was weighted by the wood density of the species within the plots while assuming that the wood density approximated any differences in the species competitive abilities (Forrester *et al.*, in press). This was done so that the equations could be applied to stands with contrasting species compositions without directly including the species composition in the equations. A similar approach may be required for biomass equations if it is found that species interactions influence biomass relationships (e.g., Laclau *et al.*, 2008) and those effects of species interactions are not adequately explained using stand variables such as relative height, basal area and trees per ha.

Alternatively, instead of using stand variables, a single tree variable could be used that is influenced by climate, species mixing and all stand structural characteristics. For example, crown diameter and height were found to be adequate for predicting biomass for a wide range of species from a wide range of forest types (Jucker *et al.*, 2017) and other studies have also found that crown characteristics were good for predicting leaf area (Ledermann and Neumann, 2006). The reason for the strong correlations between crown dimensions and tree biomass are that crown dimensions are sensitive to many stand structural and climatic characteristics and can therefore potentially replace all of those variables in biomass equations. Also, the crown and height variables probably reflect the past growing conditions experienced by the tree better than the current stand structural conditions. Unfortunately, crown dimensions are time consuming to measure directly, are measured in many different ways and are rarely available for all trees in a stand, which prevents them from being a current practical solution. However, this may change in future as methods such as terrestrial laser scanning (Seidel *et al.*, 2011) and remote sensing of crowns and heights becomes more practical. Therefore, while there are increasingly more variables being added to equations (Henry *et al.*, 2011), we suggest that these should be variables known to be sensitive to stand structure.

It is worth considering that the more independent variables an equation contains, the higher the probability that it cannot be used in other studies because some of those independent variables are unavailable. Many of the studies reviewed provided only the “optimal” equation, although the combination of independent variables in those equations varied between species or studies. Therefore, this study focused on relatively simple equations (Equations 3-21) and all equations for the given species and component that were valid are provided (not only the best equations) so that potential users can select the equations based on their available data and objectives.

In conclusion, there was considerable inter- and intra-specific variability in biomass relationships. Intra-specific variability could be explained using age or stand characteristics while the inter-specific variability was related to traits such as wood density, specific leaf area and shade tolerance. Using this information, regional biomass equations were developed for 27 species and several species groups that summarise much of the information about European biomass equations in the literature, and that can be used in a wider range of forest types than the individual site-specific equations in the publications. However, while nearly 1000 equations or raw datasets were found for the species in this study, many individual species were not well represented. This was in terms of the range of ages or stand conditions where the data was collected, but it can also be illustrated by considering the number of equations per land area where the given species occurs (Table 1). Based on this calculation, many of the species with a high number of equations (*P. sylvestris, F. sylvatica, P. abies*) have a low value (< 10 equations per land area) because they are so widely distributed. Our data set has been provided as Supplementary information with this study (Table A.2 in Forrester *et al.*, in review) so that regional species-specific or generic equations can be refined for specific regions or any other purposes that have not been considered in this study. While this dataset can easily be updated as new equations are published, a much more valuable database could be developed by adding more of the historic raw data on which the original published equations were based, as currently done in the US (Weiskittel *et al.*, 2015) and elsewhere (Falster *et al.*, 2015), and by targeting new biomass sampling to fill in the gaps of the database in terms of the distributions of independent variables (diameter, stand structural variables etc.), species representation relative to their abundance and sample distributions that are as close as possible to the diameter distributions in the target stands.

5. Acknowledgements

D.I.F. was funded by a Heisenberg Fellowship (FO 791/4-1) from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). We would like to thank Lars Drössler and Kamil Bielak for their help in obtaining publications that we did not have access to and also with translations. Lars Sprengel helped to extract some of the equations for *P. abies*. We would also like to thank the people who provided information that was not available in published equations, Evy Ampoorter, Emil Cienciala and Rainer Joosten, and also to R. Moshammer for obtaining some of the raw data . Finally, we would particularly like to thank the authors of the many publications that contained the biomass and leaf area equations that were used in this study.

6. References

Annighöfer, P., Mölder, I., Zerbe, S., Kawaletz, H., Terwei, A., Ammer, C., 2012. Biomass functions for the two alien tree species *Prunus serotina* Ehrh. and *Robinia pseudoacacia* L. in floodplain forests of Northern Italy. European Journal of Forest Research 131, 1619-1635.

António, N., Tomé, M., Tomé, J., Soares, P., Fontes, L., 2007. Effect of tree, stand, and site variables on the allometry of *Eucalyptus globulus* tree biomass. Canadian Journal of Forest Research 37, 895-906.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Statistic Surveys 4, 40-79.

Breidenbach, J., Antón-Fernández, C., Petersson, H., McRoberts, R.E., Astrup, R., 2014. Quantifying the Model-Related Variability of Biomass Stock and Change Estimates in the Norwegian National Forest Inventory. Forest Science 60, 25-33.

Canadell, J., Riba, M., Andres, P., 1988. Biomass equations for *Quercus ilex* L. in the Montseny Massif, Northeastern Spain. Forestry 61, 137-147.

Case, B.S., Hall, R.J., 2008. Assessing prediction errors of generalized tree biomass and volume equations for the boreal forest region of west-central Canada. Canadian Journal of Forest Research 38, 878-889.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B.W., Ogawa, H., Puig, H., Riera, B., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 87-99.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., Perez, R., 2004. Error propagation and scaling for tropical forest biomass estimates. Philosophical Transactions of the Royal Society B 359, 409-420.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology 20, 3177-3190.

Chojnacky, D.C., Heath, L.S., Jenkins, J.C., 2014. Updated generalized biomass equations for North American tree species. Forestry 87, 129-151.

Clough, B.J., Russell, M.B., Domke, G.M., Woodall, C.W., 2016. Quantifying allometric model uncertainty for plot-level live tree biomass stocks with a data-driven, hierarchical framework. Forest Ecology and Management 372, 175-188.

Coomes, D.A., 2006. Challenges to the generality of WBE theory. Trends in Ecology and Evolution 21, 593-596.

de-Miguel, S., Mehtatalo, L., Durkaya, A., 2014. Developing generalized, calibratable, mixed-effects meta-models for large-scale biomass prediction. Canadian Journal of Forest Research 44, 648-656.

Dietze, M.C., Wolosin, M.S., Clark, J.S., 2008. Capturing diversity and interspecific variability in allometries: A hierachical approach. Forest Ecology and Management 256, 1939-1948.

Ducey, M.J., 2012. Evergreenness and wood density predict height-diameter scaling in trees of the northeastern United States. Forest Ecology and Management 279, 21-26.

Duncanson, L., Rourke, O., Dubayah, R., 2015. Small sample sizes yield biased allometric equations in temperate forests. Scientific Reports 5, 17153.

Falster, D.S., Duursma, R.A., Ishihara, M.I., Barneche, D.R., FitzJohn, R.G., Vårhammar, A., Aiba, M., Ando, M., Anten, N., Aspinwall, M.J., Baltzer, J.L., Baraloto, C., Battaglia, M., Battles, J.J., Bond-Lamberty, B., Breugel, M.v., Camac, J., Claveau, Y., Coll, L., Dannoura, M., Delagrange, S., Domec, J.-C., Fatemi, F., Feng, W., Gargaglione, V., Goto, Y., Hagihara, A., Hall, J.S., Hamilton, S., Harja, D., Hiura, T., Holdaway, R., Hutley, L.S., Ichie, T., Jokela, E.J., Kantola, A., Kelly, J.W.G., Kenzo, T., King, D., Kloeppel, B.D., Kohyama, T., Komiyama, A., Laclau, J.-P., Lusk, C.H., Maguire, D.A., Maire, G.l., Mäkelä, A., Markesteijn, L., Marshall, J., McCulloh, K., Miyata, I., Mokany, K., Mori, S., Myster, R.W., Nagano, M., Naidu, S.L., Nouvellon, Y., O'Grady, A.P., O'Hara, K.L., Ohtsuka, T., Osada, N., Osunkoya, O.O., Peri, P.L., Petritan, A.M., Poorter, L., Portsmuth, A., Potvin, C., Ransijn, J., Reid, D., Ribeiro, S.C., Roberts, S.D., Rodríguez, R., Saldaña-Acosta, A., Santa-Regina, I., Sasa, K., Selaya, N.G., Sillett, S.C., Sterck, F., Takagi, K., Tange, T., Tanouchi, H., Tissue, D., Umehara, T., Utsugi, H., Vadeboncoeur, M.A., Valladares, F., Vanninen, P., Wang, J.R., Wenk, E., Williams, R., Ximenes, F.d.A., Yamaba, A., Yamada, T., Yamakura, T., Yanai, R.D., York, R.A., 2015. BAAD: a Biomass And Allometry Database for woody plants. Ecology 96, 1445-1445.

Fattorini, L., Gasparini, P., Nocetti, M., Tabacchi, G., Tosi, V., 2004. Above-ground tree phytomass prediction and preliminary shrub phytomass assessment in the forest stands of Trentino. Studi Trent. Sci. Nat., Acta Biol. 81, Suppl. 1, 75-121.

Forrester, D.I., 2015. Transpiration and water-use efficiency in mixed-species forests versus monocultures: Effects of tree size, stand density and season. Tree Physiology 35, 289-304.

Forrester, D.I., Benneter, A., Bouriaud, O., Bauhus, J., in press. Diversity and competition influence tree allometry - developing allometric functions for mixed-species forests. Journal of Ecology.

Forrester, D.I., Tachauer, I.H.H., Annighoefer, P., Barbeito, I., Pretzsch, H., Ruiz-Peinado, R., Stark, H., Vacchiano, G., Zlatanov, T., Chakraborty, T., Saha, S., Sileshi, G.W., in review. Specific leaf area, biomass equations and leaf area equations for European tree species. Mendeley data.

Freese, F., 1960. Testing accuracy. Forest Science 6, 139-145.

Gasparini, P., Nocetti, M., Tabacchi, G., Tosi, V., 2006. Biomass equations and data for forest stands and shrublands of the Eastern Alps (Trentino, Italy). In: Reynolds, K.M. (Ed.), IUFRO Conference "Sustainable Forestry in Theory and Practice" USDA General Technical Report PNW-GTR 688. 5-8 April, 2008 Edinburgh, Scotland.

Genet, A., Wernsdörfer, H., Jonard, M., Pretzsch, H., Rauch, M., Ponette, Q., Nys, C., Legout, A., Ranger, J., Vallet, P., Saint-André, L., 2011. Ontogeny partly explains the apparent heterogeneity of published biomass equations for *Fagus sylvatica* in central Europe. Forest Ecology and Management 261, 1188-1202.

Gorham, E., 1979. Shoot height, weight and standing crop in relation to density of monospecific plant stands. Nature 279, 148-150.

Greenhill, G., 1881. Determination of greatest height consistent with stability that a vertical pole or mast can be made, and the greatest height to which a tree of given proportions can grow. Proc. Cambridge Phil. Soc. 4, 65-73.

Henry, M., Bombelli, A., Trotta, C., Alessandrini, A., Birigazzi, L., Sola, G., Vieilledent, G., Santenoise, P., Longuetaud, F., Valentini, R., Picard, N., Saint-André, L., 2013. GlobAllomeTree: international platform for tree allometric equations to support volume, biomass and carbon assessment. iForest 6, 326-330.

Henry, M., Picard, N., Trotta, C., Manlay, R.J., Valentini, R., Bernoux, M., Saint-André, L., 2011. Estimating tree biomass of Sub-Saharan African forests: a review of available allometric equations. Silva Fennica 45, 477-569.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965-1978.

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., Birdsey, R.A., 2003. National-scale biomass estimators for United States tree species. Forest Science 49, 12-35.

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., Ewijk, K.Y.v., Forrester, D.I., Haeni, M., Higgins, S.I., Holdaway, R.J., Iida, Y., Lorimer, C., Marshall, P.L., Momo, S., Moncrieff, G.R., Ploton, P., Poorter, L., Rahman, K.A., Schlund, M., Sonké, B., Sterck, F.J., Trugman, A.T., Usoltsev, V.A., Vanderwel, M.C., Waldner, P., Wedeux, B.M.M., Wirth, C., Wöll, H., Woods, M., Xiang, W., Zimmermann, N.E., Coomes, D.A., 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programs. Global Change Biology 23, 177-190.

Ketterings, Q.M., Coe, R., Noordwijk, M.v., Ambagau, Y., Palm, C.A., 2001. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. Forest Ecology and Management 146, 199-209.

Köble, R., Seufert, G., 2001. Novel maps for forest tree species in Europe. In, A changing atmosphere, 8th European Symposium on the Physicochemical behaviour of atmospheric pollutants, 17-20 September, Torino.

Kunstler, G., Falster, D., Coomes, D.A., Hui, F., Kooyman, R.M., Laughlin, D.C., Poorter, L., Vanderwel, M., Vieilledent, G., Wright, S.J., Aiba, M., Baraloto, C., Caspersen, J., Cornelissen, J.H.C., Gourlet-Fleury, S., Hanewinkel, M., Herault, B., Kattge, J., Kurokawa, H., Onoda, Y., Peñuelas, J., Poorter, H., Uriarte, M., Richardson, S., Ruiz-Benito, P., Sun, I.-F., Ståhl, G., Swenson, N.G., Thompson, J., Westerlund, B., Wirth, C., Zavala, M.A., Zeng, H., Zimmerman, J.K., Zimmermann, N.E., Westoby, M., 2016. Plant functional traits have globally consistent effects on competition. Nature 529, 204-207.

Laclau, J.-P., Bouillet, J.-P., Gonçalves, J.L.M., Silva, E.V., Jourdan, C., Cunha, M.C.S., Moreira, M.R., Saint-André, L., Maquère, V., Nouvellon, Y., Ranger, J., 2008. Mixed-species plantations of *Acacia mangium* and *Eucalyptus grandis* in Brazil 1. Growth dynamics and aboveground net primary production. Forest Ecology And Management 255, 3905-3917.

Lambert, M.-C., Ung, C.-H., Raulier, F., 2005. Canadian national tree aboveground biomass equations. Canadian Journal of Forest Research 35, 1996-2018.

Ledermann, T., Neumann, M., 2006. Biomass equations from data of old long-term experimental plots. Austrian Journal of Forest Science 123, 47-64.

Lines, E.R., Zavala, M.A., Purves, D.W., Coomes, D.A., 2012. Predictable changes in aboveground allometry of trees along gradients of temperature, aridity and competition. Global Ecology and Biogeography 21, 1017-1028.

Litton, C.M., Raich, J.W., Ryan, M.G., 2007. Carbon allocation in forest ecosystems. Global Change Biology 13, 2089-2109.

McMahon, T.A., 1973. Size and shape in biology. Science 179, 1201-1204.

Merganič, J., Fabrika, M., Merganičová, K., 2011. Submodel of height-diameter function for climatic-site strata of the model of natural regeneration in the growth simulator SIBYLA. Acta Facultatis Forestalis Zvolen 53, 155-168.

Monserud, R.A., Marshall, J.D., 1999. Allometric crown relations in three northern Idaho conifer species. Canadian Journal of Forest Research 29, 521-535.

Muukkonen, P., 2007. Generalized allometric volume and biomass equations for some tree species in Europe. European Journal of Forest Research 126, 157-166.

Nieminen, P., Lehtiniemi, H., Vähäkangas, K., Huusko, A., Rautio, A., 2013. Standardised regression coefficient as an effect size index in summarising findings in epidemiological studies. Epidemiology Biostatistics and Public Health 10, e 8854-8851.

Niinemets, U., 2015. Is there a species spectrum within the world-wide leaf economics spectrum? Major variations in leaf functional traits in the Mediterranean sclerophyll *Quercus ilex*. New Phytologist 205, 79-96.

Niinemets, Ü., Valladares, F., 2006. Tolerance to shade, drought and waterlogging of temperate northern hemisphere trees and shrubs. Ecological Monographs 76, 521-547.

Nord-Larsen, T., Nielsen, A.T., 2015. Biomass, stem basic density and expansion factor functions for five exotic conifers grown in Denmark. Scandinavian Journal of Forest Research 30, 135-153.

Özcelik, R., Yavuz, H., Karatepe, Y., Gürlevik, N., Kris, R., 2014. Development of ecoregion-based height-diameter models for 3 economically important tree species of southern Turkey. Turkish Journal of Agriculture and Forestry 38, 399-412.

Pastor, J., Aber, J.D., Melillo, J.M., 1984. Biomass prediction using generalized allometric regressions for some northeast tree species. Forest Ecology and Management 7, 265-274.

Paul, K.I., Roxburgh, S.H., Chave, J., England, J.R., Zerihun, A., Specht, A., Lewis, T., Bennett, L.T., Baker, T.G., Adams, M.A., Huxtable, D., Montagu, K.D., Falster, D.S., Feller, M., Sochacki, S., Ritson, P., Bastin, G., Bartle, J., Wildy, D., Hobbs, T., Larmour, J., Waterworth, R., Stewart, H.T.L., Jonson, J., Forrester, D.I., Applegate, G., Mendham, D., Bradford, M., O'Grady, A., Green, D., Sudmeyer, R., Rance, S.J., Turner, J., Barton, C., Wenk, E.H., Grove, T., Attiwill, P.M., Pinkard, E., Butler, D., Brooksbank, K., Spencer, B., Snowdon, P., O'Brien, N., Battaglia, M., Cameron, D.M., Hamilton, S., McAuthur, G., Sinclair, J., 2016. Testing the generality of above-ground biomass allometry across plant functional types at the continent scale. Global Change Biology 22, 2106-2124.

Pérez-Cruzado, C., Fehrmann, L., Magdon, P., Cañellas, I., Sixto, H., Kleinn, C., 2015. On the site-level suitability of biomass models. Environmental Modelling & Software 73, 14-26.

Peterson, R.A., Brown, S.P., 2005. On the use of beta coefficients in meta-analysis. Journal of Applied Psychology 90, 175-181.

Picard, N., Saint-André, L., Henry, M., 2012. Manual for building tree volume and biomass allometric equations: from field measurement to prediction. Food and Agricultural Organization of the United Nations, Rome, and Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Montpellier, 215 pp. In.

Poorter, H., Niklas, K.J., Reich, P.B., Oleksyn, J., Poot, P., Mommer, L., 2012. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. New Phytologist 193, 30-50.

Porte, A., Bosc, A., Champion, I., Loustau, D., 2000. Estimating the foliage area of Maritime pine (Pinus pinaster Ait.) branches and crowns with application to modelling the foliage area distribution in the crown. Annals of Forest Science 57, 73-86.

Porté, A., Trichet, P., Bert, D., Loustau, D., 2002. Allometric relationships for branch and tree woody biomass of Maritime pine (Pinus pinaster Aït.). Forest Ecology and Management 158, 71-83.

Pretzsch, H., 2006. Species-specific allometric scaling under self-thinning: evidence from long-term plots in forest stands. Oecologia 146, 572-583.

Pretzsch, H., Daubert, E., Biber, P., 2013. Species-specific and ontogeny related stem allometry of European forest trees: evidence from extensive stem analyses. Forest Science 59, 290-302.

Pretzsch, H., Dieler, J., 2012. Evidence of variant intra- and interspecific scaling of tree crown structure and relevance for allometric theory. Oecologia 169, 637-649.

Pretzsch, H., Matthew, C., Dieler, J., 2012. Allometry of Tree Crown Structure. Relevance for Space Occupation at the Individual Plant Level and for Self-Thinning at the Stand Level. In: Matyssek, R., Schnyder, H., Oßwald, W., Ernst, D., Munch, J.C., Pretzsch, H. (Eds.), Growth and Defence in Plants. Springer, Berlin Heidelberg, pp. Ecological Studies 220, 287-310.

R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. In.

Reich, P.B., Luo, Y., Bradford, J.B., Poorter, H., Perry, C.H., Oleksyn, J., 2014. Temperature drives global patterns in forest biomass distribution in leaves, stems, and roots. Proceedings of the National Academy of Sciences 111, 13721-13726.

Reineke, L.H., 1933. Perfecting a stand-density index for even aged forests. Journal of Agricultural Research 46, 627-638.

Roxburgh, S.H., Paul, K.I., Clifford, D., England, J.R., Raison, R.J., 2015. Guidelines for constructing allometric models for the prediction of woody biomass: How many individuals to harvest? Ecosphere 6, 38.

Ruiz-Peinado, R., Rio, M.d., Montero, G., 2011. New models for estimating the carbon sink capacity of Spanish softwood species. Forest Systems 20, 176-188.

Russell, M.B., Domke, G.M., Woodall, C.W., D'Amato, A.W., 2015. Comparisons of allometric and climate-derived estimates of tree coarse root carbon stocks in forests of the United States. Carbon Balance and Management 10, 20.

Ryan, M.G., Phillips, N., Bond, B.J., 2006. The hydraulic limitation hypothesis revisited. Plant Cell And Environment 29, 367-381.

Saint-André, L., M'Bou, A.T., Mabiala, A., Mouvondy, W., Jourdan, C., Roupsard, O., Deleporte, P., Hamel, O., Nouvellon, Y., 2005. Age-related equations for above- and below-ground biomass of a *Eucalyptus* hybrid in Congo. Forest Ecology and Management 205, 199-214.

Sands, P.J., Landsberg, J.J., 2002. Parameterisation of 3-PG for plantation grown *Eucalyptus globulus*. Forest Ecology and Management 163, 273-292.

Seidel, D., Leuschner, C., Müller, A., Krause, B., 2011. Crown plasticity in mixed forests - Quantifying asymmetry as a measure of competition using terrestrial laser scanning. Forest Ecology and Management 261, 2123-2132.

Seidl, R., Rammer, W., Bellos, P., Hochbichler, E., Lexer, M.J., 2010. Testing generalized allometries in allocation modeling within an individual-based simulation framework. Trees 24, 139-150.

Shaiek, O., Loustau, D., Trichet, P., Meredieu, C., Bachtobji, B., Garchi, S., Aouni, M.H.E., 2011. Generalized biomass equations for the main aboveground biomass components of maritime pine across contrasting environments. Annals of Forest Science 68, 443-452.

Sileshi, G.W., 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. Forest Ecology and Management 329, 237-264.

Snowden, P., 1991. A ratio estimator for bias correction in logarithmic regressions. Canadian Journal of Forest Research 21, 720-724.

Snowdon, P., Raison, R.J., Keith, H., Ritson, P., Grierson, P., Adams, M., Montagu, K., Bi, H., Burrows, W., Eamus, D., 2002. Protocol for sampling tree and stand biomass. National Carbon Accounting System Technical Report No. 31. The Australian Greenhouse Office, Canberra, Australia. In.

Tabacchi, G., Di Cosmo, L., Gasparini, P., Morelli, S., 2011. Stima del volume e della fitomassa delle principali specie forestali italiane. Equazioni di previsione, tavole del volume e tavole della fitomassa arborea epigea. Consiglio per la Ricerca e la sperimentazione in Agricoltura, Unità di Ricerca per il Monitoraggio e la Pianificazio ne Forestale. In, p. 412.

Thorpe, H.C., Astrup, R., Trowbridge, A., Coates, K.D., 2010. Competition and tree crowns: A neighborhood analysis of three boreal tree species. Forest Ecology and Management 259, 1586-1596.

van Breugel, M., Ransijn, J., Craven, D., Bongers, F., Hall, J.S., 2011. Estimating carbon stock in secondary forests: Decisions and uncertainties associated with allometric biomass models. Forest Ecology and Management 262, 1648-1657.

Watt, M.S., Kirschbaum, M.U.F., 2011. Moving beyond simple linear allometric relationships between tree height and diameter. Ecological Modelling 222, 3910-3916.

Weiskittel, A.R., MacFarlane, D.W., Radtke, P.J., Affleck, D.L.R., Temesgen, H., Woodall, C.W., Westfall, J.A., Coulston, J.W., 2015. A call to improve methods for estimating tree biomass for regional and national assessments. Journal of Forestry 113, 414-424.

West, G.B., Brown, J.H., Enquist, B.J., 1999. A general model for the structure and allometry of plant vascular systems. Nature 400, 664-667.

West, G.B., Enquist, B.J., Brown, J.H., 2009. A general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106, 7040-7045.

Wirth, C., Schumacher, J., Schulze, E., 2004. Generic biomass functions for Norway spruce in Central Europe - A meta-analysis approach toward prediction and uncertainty estimation. Tree Physiology 24, 121-139.

Wutzler, T., Wirth, C., Schumacher, J., 2008. Generic biomass functions for Common beech (*Fagus sylvatica*) in Central Europe: predictions and components of uncertainty. Canadian Journal of Forest Research 38, 1661-1675.

Xiang, W., Liu, S., Deng, X., Shen, A., Lei, X., Tian, D., Zhao, M., Peng, C., 2011. General allometric equations and biomass allocation of *Pinus massoniana* trees on a regional scale in southern China. Ecological Research 26, 697-711.

Xiang, W., Zhou, J., Ouyang, S., Zhang, S., Lei, P., Li, J., Deng, X., Fang, X., Forrester, D.I., Bauhus, J., 2016. Species-specific and general allometric equations for estimating tree biomass components of subtropical forests in southern China. European Journal of Forest Research 135, 963-979.

Yoda, K., Kira, T., Ogawa, H., Hozami, K., 1963. Self thinning in overcrowded pure stands under cultivated and natural conditions. Journal of Biology Osaka City University 14, 107-129.

Zanne, A.E., Lopez-Gonzalez, G., Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L., Miller, R.B., Swenson, N.G., Wiemann, M.C., Chave, J., 2009. Global wood density database. Dryad. Identifier: http://hdl.handle.net/10255/dryad.235.

Zianis, D., Muukkonen, P., Mäkipää, R., Mencuccini, M., 2005. Biomass and stem volume equations for tree species in Europe. Silva Fennica Monographs 4, 63.

Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R. Springer.

Table 1. The number of published equations or data sets found for each component.

Species	Equations per species distribution*	Number of equations/raw data sets								
		Foliage mass or leaf area	Live branch	Dead branch	Live & dead branch	Stem	Stem & branch	Aboveground	Root	Total***
Abies alba	6.7	1				3	1	4	1	0/10
Acer pseudoplatanus	(>110)	3			5	5	7	1	1	1/21
Alnus glutinosa	48.6	3			3	4	3	4	1	0/18
Alnus incana	(>100)	4			4	3	4	4	1	0/20
Betula pendula	8.7	2			5	5	5	3		6/14
Betula pubescens	5.4	3	1	1	4	4	4	8	1	0/26
*B. pendula or B. pubescens***	10.1	7	3	3	13	13	17	13	3	12/60
Carpinus betulus	14	1			4	3	4	1	1	0/14
Castanea sativa	27.3	4			5	6	5	9	1	0/30
Fagus sylvatica	9.9	10	2	1	13	12	14	9	10	3/68
Fraxinus excelsior	58.7	7			7	7	10	4	2	3/34
Larix decidua	36.1	4	3	3	5	5	4	5	1	0/30
Picea abies	6	28	12	12	19	15	16	18	7	16/111
Pinus cembra	(>10.0)							2		0/2
Pinus nigra	8.6	2			2	3	1	5	2	3/12
Pinus pinaster	25	14	1	1	10	14	6	14	5	6/59
Pinus sylvestris	3.5	22	7	7	16	15	18	15	9	2/107
Populus alba	(>30)	1			1	1	1	1	1	0/6
Populus tremula	25.4	3			3	4	3	3		0/16
*P. alba, P. tremula and P. tremula hybrids***		5			10	11	11	6	1	12/31
Prunus avium	(>10)	2								0/2
Prunus serotina		1			2	2	3	1		0/9
Pseudotsuga menziesii	342.6	40	17	14	23	26	18	19	4	15/138
Quercus ilex	13.5	5			4	5	5	7	3	5/24
Quercus petraea	9.1				4	5	11		1	13/8
Quercus robur	9	3			6	6	8	2	2	13/14
*Q. petraea or Q. robur***	11.3	3	1		13	13	24	2	4	27/33
Robinia pseudoacacia	51.9	3			6	6	5	4	3	0/27
Sorbus aucuparia	(>25)	1			1	1	1	1		0/5
Tilia cordata or platyphyllas	15					1	1	1		0/3

* The number of equations divided by the percentage of Europe's forest area where that species occurs (from Köble and Seufert, 2001). The numbers in parentheses are for species that had no area available in Köble and Seufert (2001), so their area was assumed to occupy < 0.2%, which would result in the numbers in parentheses.

**Several species from the same genera were grouped because some studies did not differentiate the species. For such groups, the data from each species was combined (not only the studies that did not differentiate the species).

***Number of raw data sets / number of equations.

Table 2. Characteristics of the species and the mean (minimum/maximum) characteristics of the stands where those equations or data sets were obtained.

Species	Wood density (g cm^{-3})*	Shade tolerance*	Specific leaf area (m^2 kg^{-1})**	Diameter (cm)	Age (years)	Basal area (m^2 ha^{-1})	Trees per ha	Latitude (°)	Longitude (°)	Mean annual temperature (°C)##	Mean annual precipitation (mm)##
Abies alba	0.35	4.6	9.08	34 (5.7/80)	32.7 only			47.0 (42.7/56.0)	8 (0.8/15.7)	8.4 (6.3/11.3)	999 (718/1144)
Acer pseudoplatanus	0.51	3.73	19.51	8.4 (0.1/88.2)	13.4 (4.8/143.8)	20.4 (17.7/30.6)	21179 (8202/39870)	50.9 (45.6/56.0)	6.8 (-87.3/19.3)	7.6 (5.6/8.6)	810 (697/1287)
Alnus glutinosa	0.44	2.71	14.93	14.9 (1/47.3)	26.3 (11/49)	22.3 (12.2/32.5)	8756 (1114/16399)	56.1 (42.4/59.5)	11.9 (1.6/16)	6.5 (5.8/7.8)	712 (569/1011)
Alnus incana	0.44	2.3	14.93	9.3 (0.7/24.6)	19.5 (6.5/40)	27.6 (10.3/33.4)	13730 (1854/27719)	60.4 (57.3/61.8)	16.8 (10.8/26.1)	4.4 (2.9/6.7)	678 (630/753)
Betula pendula	0.53	2.03	9.84	7.7 (1/25.5)	12.4 (6.5/20)	14.9 (7.7/19.8)	13202 (466/22684)	57.8 (49.9/62.0)	12.1 (5.3/26.1)	6 (3/9.5)	717 (630/809)
Betula pubescens	0.53	1.85	9.84	4.8 (0.5/15)	9.2 (8/12)	14.6 (10.6/21.3)	28919 (3060/69527)	61.3 (52.0/69.7)	22.8 (10.4/27.3)	3.3 (-1.1/8.2)	576 (414/723)
B. pendula or B. pubescens#				8.5 (0.5/38)	17.5 (6.5/44)	15.5 (7.7/21.3)	20170 (466/69527)	60.7 (49.9/69.7)	17.7 (-3.1/27.3)	3.8 (-1.1/9.5)	659 (414/1287)
Carpinus betulus	0.71	3.97	19.83	24 (0.1/75)	14.9 (12/32)	23.3 (21.3/35.1)	7921 (168/15466)	45.5 (36.4/52.0)	26.5 (7.6/52)	9.7 (1.4/13.7)	672 (418/1523)
Castanea sativa	0.46	3.15	14.34	21.4 (1/64.2)	57.4 (12.3/100)	30.9 (28.4/43.2)	3002 (1597/3970)	40.7 (37.5/46.1)	-4.4 (-7.6/15)	12.3 (6.8/15.8)	740 (468/1275)
Fagus sylvatica	0.59	4.56	21.54	27.1 (0.5/84)	64.2 (8/145)	24.1 (2.9/73.8)	3726 (196/16815)	49.4 (42.3/56.0)	6.2 (-4.2/16.2)	8.9 (1.4/11.3)	766 (537/1523)
Fraxinus excelsior	0.56	2.66	17.08	7.7 (0.1/69.3)	12.2 (4.6/114.2)	21.4 (17.7/28.8)	17082 (1000/35620)	50.9 (45.2/59.5)	11.4 (-3.3/19.3)	8.3 (6.7/10.9)	802 (549/1325)
Larix decidua	0.47	1.46	12.18	30.4 (4/90.1)	24 (20/28)	28.6 (28.6/28.6)	2390 (2280/2500)	46.8 (43.9/50.6)	-12.6 (-91.9/15.7)	6.7 (2.9/11.3)	856 (744/1126)
Picea abies	0.37	4.45	5.76	22.8 (0.4/82)	47.9 (17/152)	34.9 (2.4/62)	4301 (308/17388)	54.1 (42.7/67.0)	1.2 (-91.9/26.7)	5.6 (-0.3/11.3)	769 (521/1280)
Pinus cembra	0.42	2.87	4.6	32.2 (7.7/56.3)				46.1 (46.1/46.1)	11.1 (11.1/11.1)	11.3 (11.3/11.3)	874 (874/874)
Pinus nigra	0.42	2.1	***	29.5 (6/77.3)	21.1 (13/62)	39.9 (39.9/39.9)	2500 (2500/2500)	41.9 (40.6/51.2)	20.6 (-1.6/33.7)	10.9 (9.1/13.4)	727 (522/931)
Pinus pinaster	0.41	2.21	2.97	24.9 (1/64)	26.9 (7.6/64)	38.8 (17.1/66.7)	2183 (223/23614)	41.2 (32.8/44.7)	10.6 (-8.1/117)	12.8 (10/15.9)	818 (381/1298)
Pinus sylvestris	0.42	1.67	4.28	18.4 (1/76)	48.7 (10/165)	25.2 (16/37.6)	4645 (223/17388)	55.9 (41.2/69.3)	14.1 (-4.2/26.7)	6.2 (0.8/11.3)	702 (492/1340)
Populus alba	0.35	2.3	9.84	5.4 (0.1/10.7)	8 (8/8)			44.5 (44.5/44.5)	26.7 (26.7/26.7)	11 (11/11)	550 (550/550)
Populus tremula	0.37	2.22	9.84	11.3 (1/40)	18.4 (10/46)	32.6 (26.1/37.5)	9549 (1246/17805)	58.7 (51.2/61.0)	12.5 (5.3/17)	5.7 (3.1/9.5)	712 (622/809)
P. alba, P. tremula and P. tremula hybrids#				13.5 (0.1/44.7)	16.5 (7/46)	30.3 (10.6/37.5)	5834 (440/17805)	55.1 (44.5/61.0)	13.2 (7.5/26.4)	8 (5.1/11)	672 (550/920)
Prunus avium	0.47	3.33	12.43	5.5 (1/10)			167 (167/167)	43.8 (43.8/43.8)	1.7 (1.7/1.7)	12.9 (12.9/12.9)	724 (724/724)
Prunus serotina	0.68	2.46	***	23.7 (0.9/49.6)	31.9 (27.5/45)	37.7 (37.7/37.7)		43.9 (39.6/45.3)	-13.3 (-80/8.9)	12.4 (10.9/12.9)	1012 (993/1069)
Pseudotsuga menziesii	0.43	2.78	5.98	26.6 (1/163)	40.2 (10/262)	46.4 (0.5/75)	3893 (235/12000)	46.1 (35.3/56.0)	-108.6 (-123.9/11.7)	7.7 (4.2/10.9)	1121 (347/2315)
Quercus ilex	0.82	3.02	4.83	17.9 (1/90)	78.7 (31/133)	27.2 (16.3/38.8)	6271 (1814/10865)	37.9 (32.9/43.7)	3.6 (-5.4/15)	16.2 (8.2/21.7)	456 (59/760)
Quercus petraea	0.56	2.73	14.81	20.4 (3/77.1)	61 (20/198)	19.9 (7.7/35.1)	4523 (1419/10700)	51.5 (48.7/54.4)	3 (-3.1/8)	7.7 (6.8/9.5)	987 (745/1287)
Quercus robur	0.56	2.45	14.81	29 (4/67.5)	75.2 (20/122)	32.4 (14.3/41.9)	1105 (375/6137)	46.6 (42.7/51.3)	-1.3 (-8/7.5)	10.3 (7.8/11.7)	953 (746/1284)
Q. petraea or Q. robur#				24 (1/77.1)	64.2 (10/198)	26.8 (7.7/41.9)	2235 (163/10700)	48.9 (42.7/54.4)	3 (-8/15.3)	8.9 (1.4/11.7)	936 (626/1523)
Robinia pseudoacacia	0.68	1.72	24.84	11.3 (0.1/40.1)	15.3 (8/27.5)	31.2 (24.8/37.7)	3551 (3551/3551)	43.1 (35.1/51.2)	32.3 (-83.4/109)	10.4 (8.6/12.9)	712 (462/1915)
Sorbus aucuparia	0.63	2.73	***	5.5 (1/10)	11 (11/11)			59.5 (59.5/59.5)	10.8 (10.8/10.8)	6.7 (6.7/6.7)	753 (753/753)
Tilia cordata or platyphyllas	0.42	4.18	22.11	26.6 (4/58.6)	73.6 (73.6/73.6)			34.6 (24.8/54.4)	8.1 (-3.1/13.7)	6.9 (6.8/6.9)	870 (662/1287)

*See Table B.3 for more information (Niinemets and Valladares, 2006; Zanne *et al.*, 2009).

**Means calculated from data in Table A.1 of Forrester *et al.* (in review)

*** Specific-leaf area data was not available for these species

#Several species from the same genera were grouped because some studies did not differentiate the species. For such groups, the data from each species was combined (not only the studies that did not differentiate the species).

The climate data was obtained from Hijmans *et al.* (2005).

Table 3. Mean β_1 of Equation 3 and its 95% confidence interval for all of the final equations in Table A.5 of Forrester *et al.* (in review).

Component	Mean (95% confidence interval)
Aboveground	2.41 (2.12 / 2.70)#
Stem and branch mass	2.37 (2.09 / 2.65)* #
Stem mass	2.39 (2.08 / 2.70) #
Live branch mass	2.15 (0.16 / 4.14)
Total branch mass	2.31 (1.72 / 2.89) #
Foliage mass	1.83 (1.13 / 2.53)* #
Root mass	2.27 (1.74 / 2.80) #

*interval does not contain the β_1 predicted by the metabolic scaling theory (2.67)
#interval does not contain the β_1 predicted by the geometric scaling theory (3)

Table 4. Mean values of statistics for Equation 3 (with only diameter as an explanatory variable) compared with all other equations, which also included stand and climatic variables (Equations 4-21). PRSE = percent relative standard errors (Equation 25), AIC = Akaike's information criterion, BIC = Bayesian information criterion, RMSE = square root of the mean square error (Equation 26), MAPE = mean absolute percentage error (Equation 27) and Bias = bias calculated using Equation 28.

Species	Equation	Mean AIC	Mean BIC	Mean RMSE	Mean MAPE	Mean Bias
All species	All except 3	539	570	0.264	22	11.4
	3	955	979	0.279	23.2	12.6
Broadleaved species	All except 3	497	525	0.296	34.1	23.6
	3	501	522	0.286	24.3	13.6
Coniferous species	All except 3	-134	-105	0.231	16.5	6.4
	3	227	250	0.259	20.8	9.8
Abies alba	3	-211	-200	0.085	5.4	-0.3
Acer pseudoplatanus	All except 3	54	70	0.284	20.3	1
	3	32	44	0.255	18.3	2.5
Alnus glutinosa or *Alnus incana*	All except 3	-64	-44	0.159	12.2	1.1
	3	131	148	0.223	18	6.6
Alnus glutinosa	All except 3	-92	-80	0.13	10.2	0.3
	3	-40	-29	0.204	16.8	4.5
Alnus incana	All except 3	-311	-298	0.047	3.7	-0.2
	3	-138	-127	0.135	9.6	-0.3
B. pendula or B. pubescens	All except 3	-41	-27	0.869	39.5	7.8
	3	-8	5	0.285	31.5	22.8
Betula pendula	All except 3	-140	-127	0.097	7.2	0.8
	3	-43	-33	0.2	17.9	7
Betula pubescens	All except 3	-213	-201	0.062	5.5	-0.9
	3	-195	-183	0.138	10.6	0.5
Carpinus betulus	All except 3	-144	-131	0.142	12.1	-3.1
	3	-59	-48	0.197	16.2	-2.8
Castanea sativa	All except 3	6	20	0.283	23.5	5.1
	3	12	25	0.257	17.5	-3.1
Fagus sylvatica	All except 3	-91	-73	0.306	36	21.6
	3	-56	-41	0.205	19	9.2
Fraxinus excelsior	All except 3	114	130	0.322	25.1	3.5
	3	157	170	0.313	25.6	-6.5
Larix decidua	All except 3	-137	-126	0.13	11.2	2.2
	3	-92	-80	0.181	13.5	1.2
Picea abies	All except 3	-87	-66	0.202	14.5	5.5
	3	-84	-67	0.219	15.9	7.1
Pinus cembra	3	-190	-181	0.05	4.2	0.6
Pinus nigra	3	-56	-47	0.134	8.4	0.5
Pinus pinaster	All except 3	-159	-140	0.168	13.2	6.2
	3	-105	-92	0.195	14.7	5.8
Pinus sylvestris	All except 3	-155	-133	0.195	14.9	6.1
	3	-58	-42	0.218	17	5.3
P. alba, P. tremula and P. tremula hybrids	All except 3	3	19	0.22	18.9	11.7
	3	-168	-155	0.147	12.6	6.9
Prunus avium or Prunus serotina	3	-51	-42	0.164	14.3	4.1
Prunus avium	3	-31	-22	0.16	13.3	1.4
Prunus serotina	3	-56	-47	0.17	15.8	7.4
Pseudotsuga menziesii	All except 3	-3	19	0.262	21.8	8.2
	3	92	109	0.262	21.7	7.2
Quercus ilex	All except 3	190	209	0.212	15.4	4
	3	96	109	0.203	15	1.8

Quercus petraea	All except 3	-355	-339	0.085	6.3	-0.6
	3	-162	-153	0.16	14	1.9
Q. petraea or Q. robur	All except 3	-122	-105	0.182	19.3	13.3
	3	-145	-133	0.157	14.3	5.7
Quercus robur	All except 3	-111	-98	0.128	10.1	5.5
	3	-73	-63	0.147	11.7	3.8
Robinia pseudoacacia	All except 3	-151	-138	0.097	8.2	3.8
	3	-84	-73	0.153	11.7	4.1

Figure captions

Figure 1. Frequency distributions of stand variables, including basal area (a), trees per ha (b), age (c) and latitude (d), as well as sample characteristics including the sample size (e), the minimum diameter (dbh) sampled (f), the maximum diameter sampled (g), the distribution of diameter pseudo-observations (h) and the mean square error (i). The frequency indicates the number of equations or raw data sets that included the value on the x-axis, except for (h) which includes all individual pseudo-observations or raw data points. This figure summarizes the data provided in Table A.2 of Forrester *et al.* (in review). In addition to the pseudo-observations, the raw data was also included when available. Note that for clarity, the x-axes do not extend to their maximum for (b) maximum = 70000, (e) maximum = 4213 and (h) maximum = 163.

Figure 2. The pseudo-observations calculated from four published equations for *Abies alba* predicting aboveground biomass (kg/tree) without (a) or with (b) the model error taken into account. The equation numbers in the legend refer to the equations in Table A.2 of Forrester *et al.* (in review). In the legends 1= (Fattorini *et al.*, 2004; Gasparini *et al.*, 2006), 2=(Nord-Larsen and Nielsen, 2015), 3=(Ruiz-Peinado *et al.*, 2011), 4=(Tabacchi *et al.*, 2011).

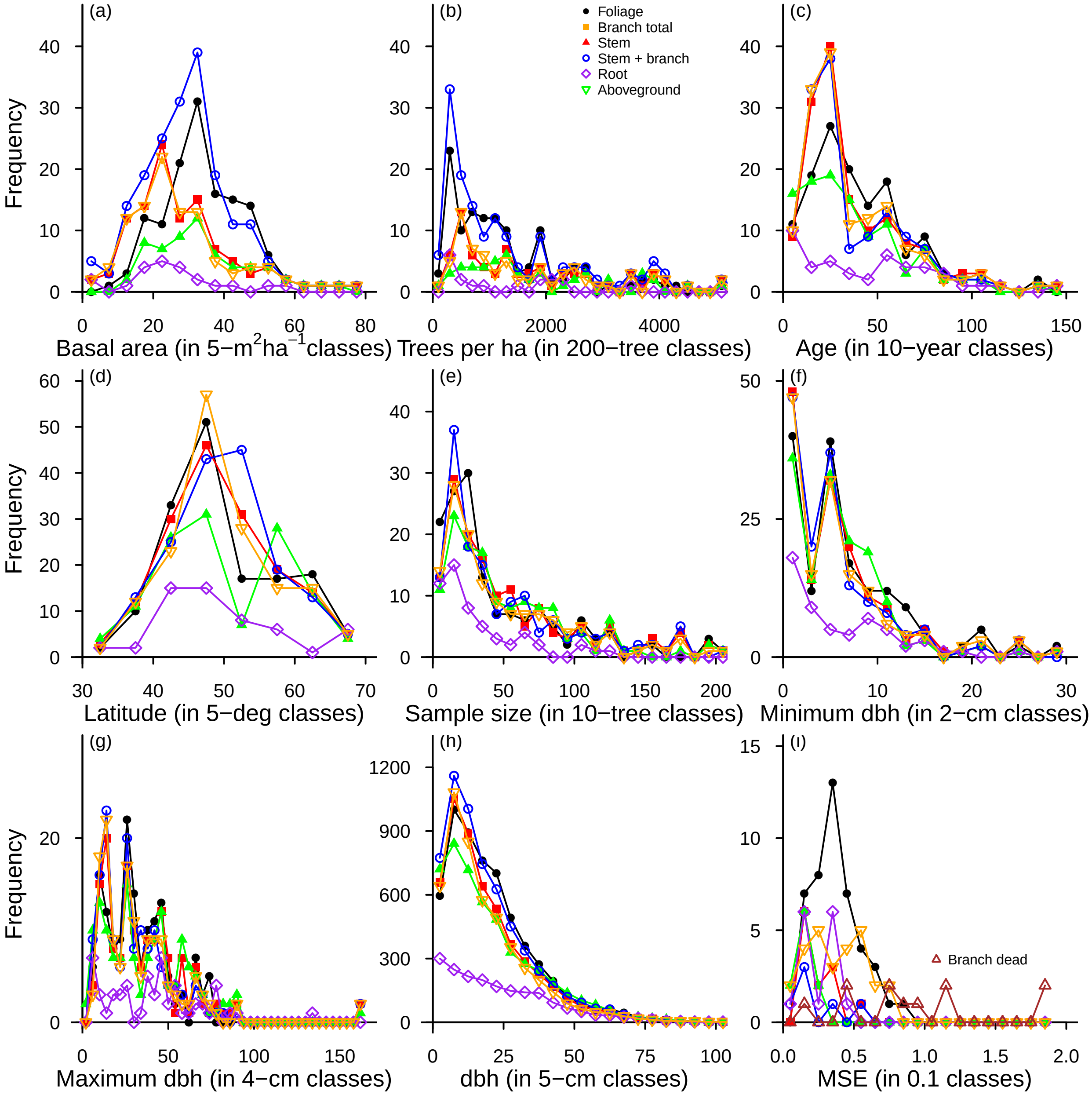
Figure 3. The pseudo-observations for individual tree foliage mass (a,b) and stem mass (c,d) for two contrasting species that were well represented in the data set, *Fagus sylvatica* (a,c, deciduous, broad-leaved) and *Pinus sylvestris* (b,d, evergreen conifer). The numbers in the legends represent the equation number in the database (Table A.2 in Forrester *et al.*, in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. One raw data set was excluded from (b) to improve clarity.

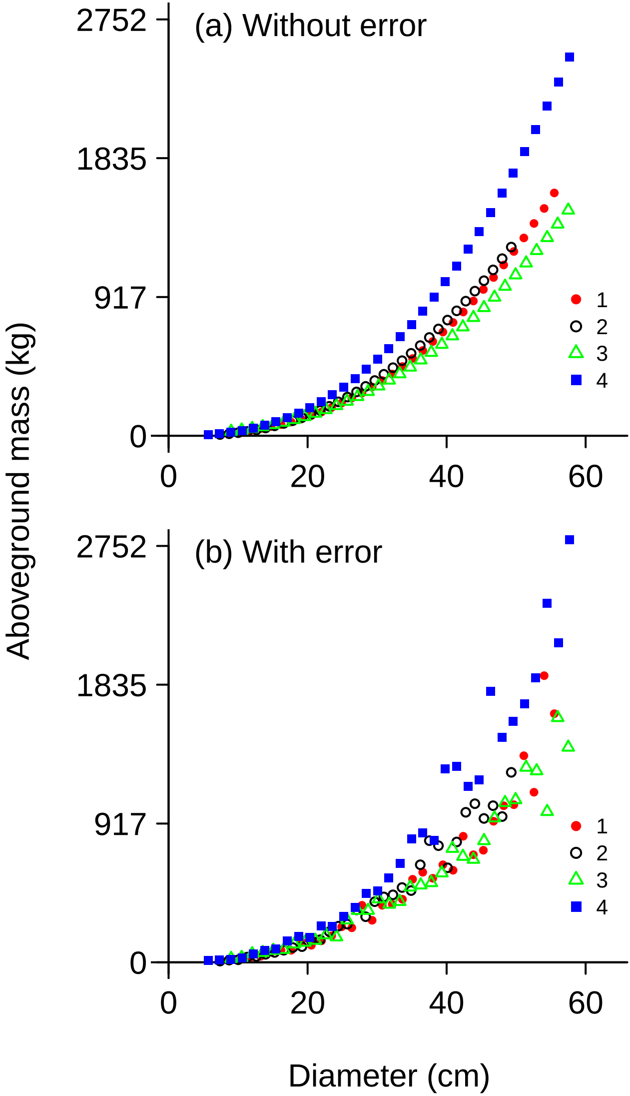
Figure 4. The relationships between foliage mass and diameter for various European species. The lines are Equation 3 fitted to all available data for the given species (Table A.2) and are provided in Table A.5 (Forrester *et al.*, in review). The dotted lines show the same relationships after dividing the data into broad-leaved species (dotted blue line), conifers (dotted orange line) or all species (dotted green line). The inset is the same relationship for species with small ranges of pseudo-observations.

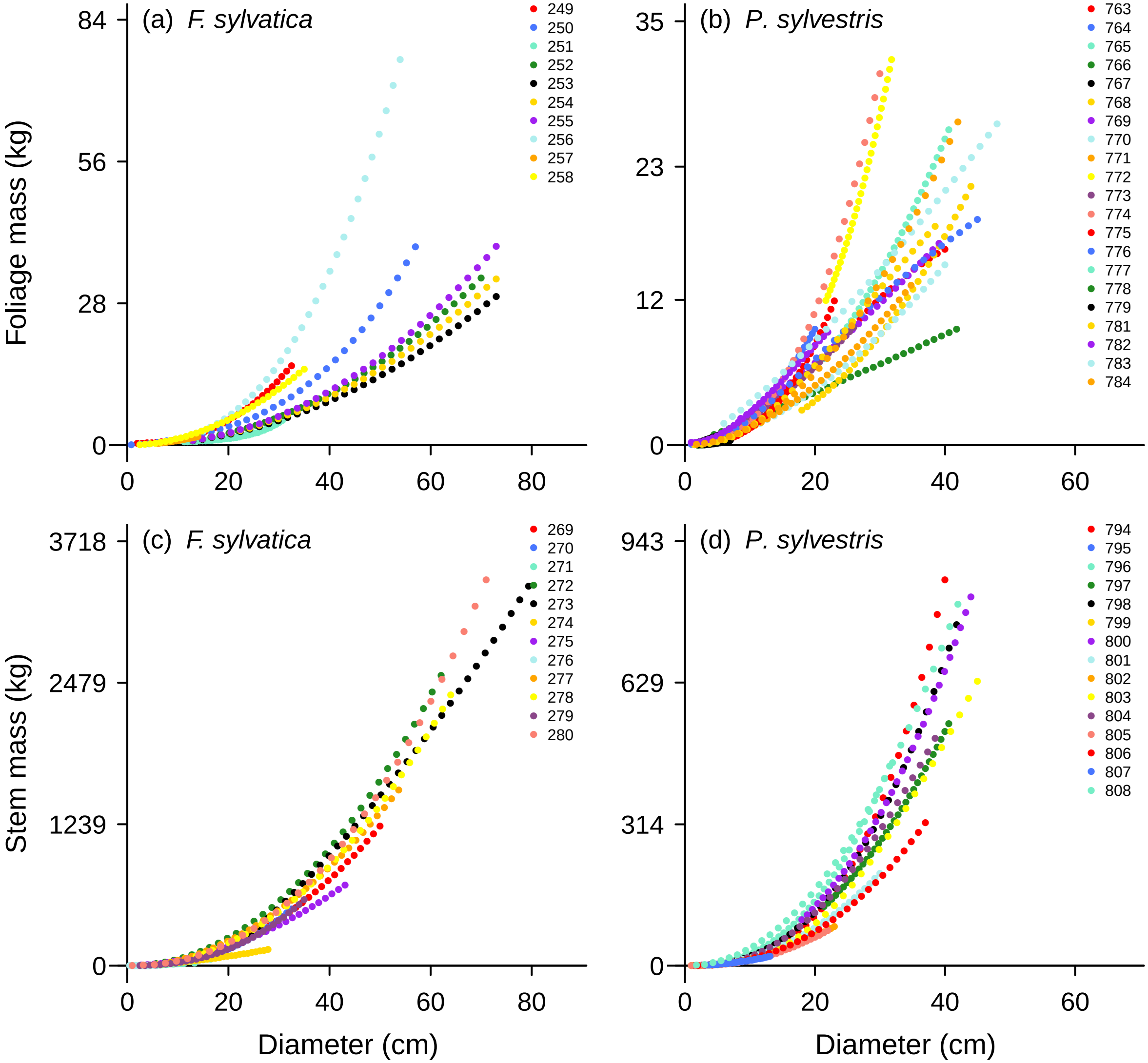
Figure 5. The relationships between stem mass and diameter for various European species. The lines are Equation 3 fitted to all available data for the given species (Table A.2) and are provided in Table A.5 (Forrester *et al.*, in review). The dotted lines show the same relationships after dividing the data into broad-leaved species (dotted blue line), conifers (dotted orange line) or all species (dotted green line). The inset is the same relationship for species with small ranges of pseudo-observations.

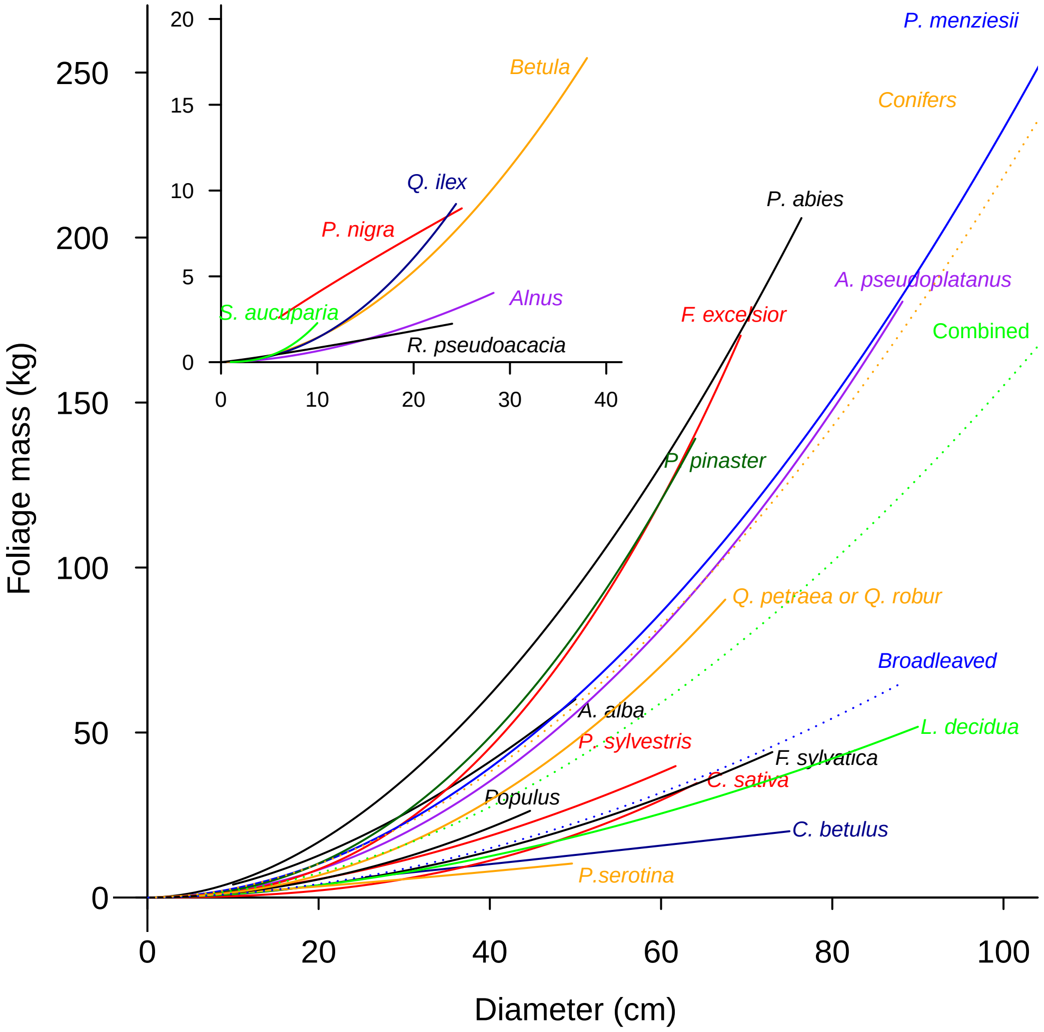
Figure 6. The effect sizes (beta weights) of independent variables used to predict foliage mass (a-c), stem mass (d-f) and root mass (g-i) using Equations 22-24. The effect sizes are quantified as the slopes of the variables on the x-axis, all of which have been standardised (mean = 0 and standard deviation = 1). “I-ns” indicates that there was no significant interaction between the given x-axis variable and the trait (specific leaf area (SLA), wood density, shade tolerance). “ns” indicates that the x-axis variable was not significant in the model. The statistical information is in Table A.6 (Forrester *et al.*, in review). For the trait levels, Low = Mean – 1 standard deviation, Mean = Mean of all species, High = Mean + 1 standard deviation.

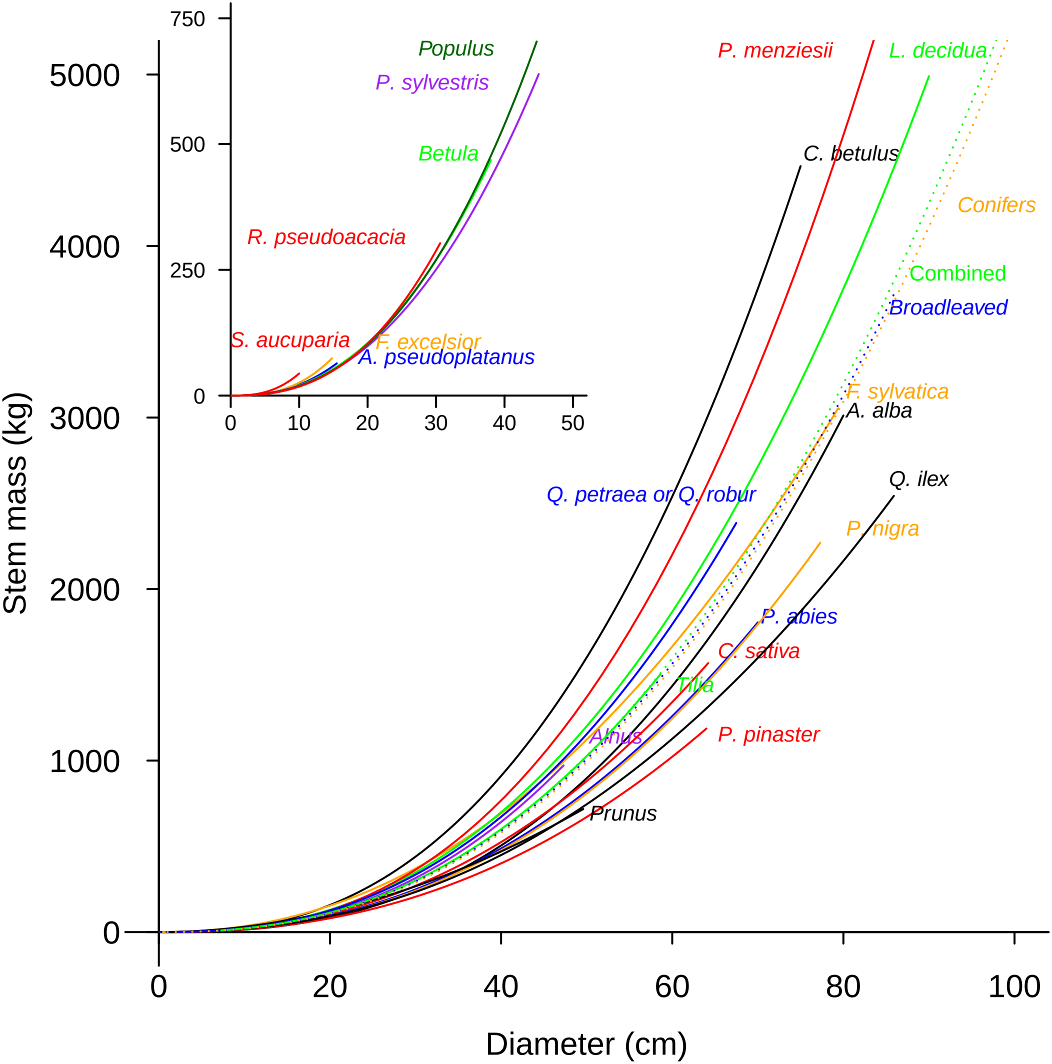
Figure 7. The effect sizes (beta weights) of all independent variables used to predict aboveground mass (a-c), stem + branch mass (d-f) and total branch mass (g-i) using Equations 22-24. The effect sizes are quantified as the slopes of the variables on the x-axis, all of which have been standardised (mean = 0 and standard deviation = 1). “I-ns” indicates that there was no significant interaction between the given x-axis variable and the trait (specific leaf area (SLA), wood density, shade tolerance). “ns” indicates that the x-axis variable was not significant in the model. The statistical information is in Table A.6 of Forrester et al. (in review). For the trait levels, Low = Mean – 1 standard deviation, Mean = Mean of all species, High = Mean + 1 standard deviation.

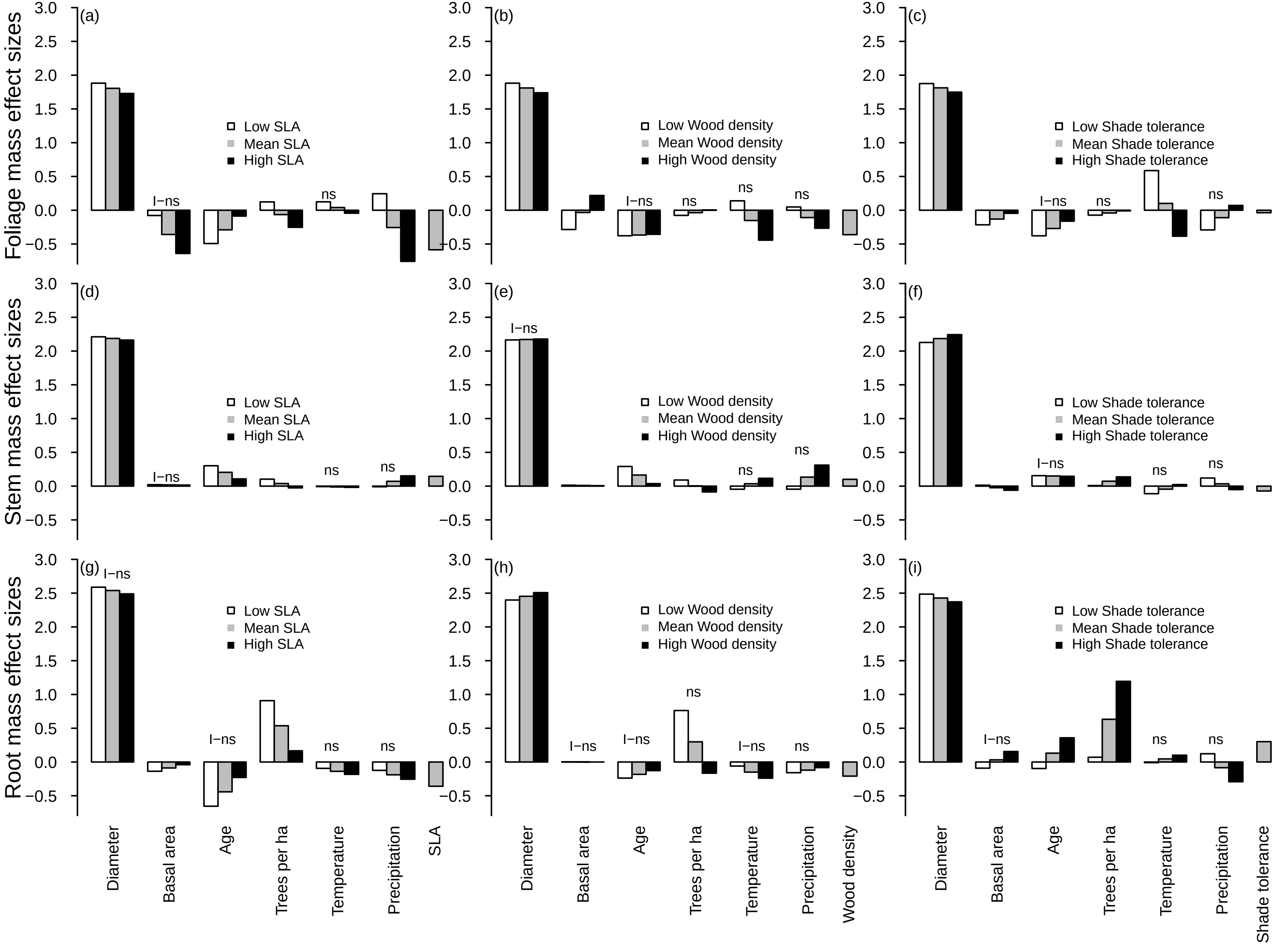

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
● Foliage
■ Branch total
▲ Stem
○ Stem + branch
◇ Root
▽ Aboveground
△ Branch dead
Frequency
Basal area (in 5−m^{2}ha^{-1}classes)
Trees per ha (in 200−tree classes)
Age (in 10−year classes)
Latitude (in 5−deg classes)
Sample size (in 10−tree classes)
Minimum dbh (in 2−cm classes)
Maximum dbh (in 4−cm classes)
dbh (in 5−cm classes)
MSE (in 0.1 classes)
0
10
20
30
40
50
60
0
20
40
60
80
0
2000
4000
0
50
100
150
30
40
50
60
70
200
25
300
600
900
1200
0.0
0.5
1.0
1.5
2.0
5
15
75

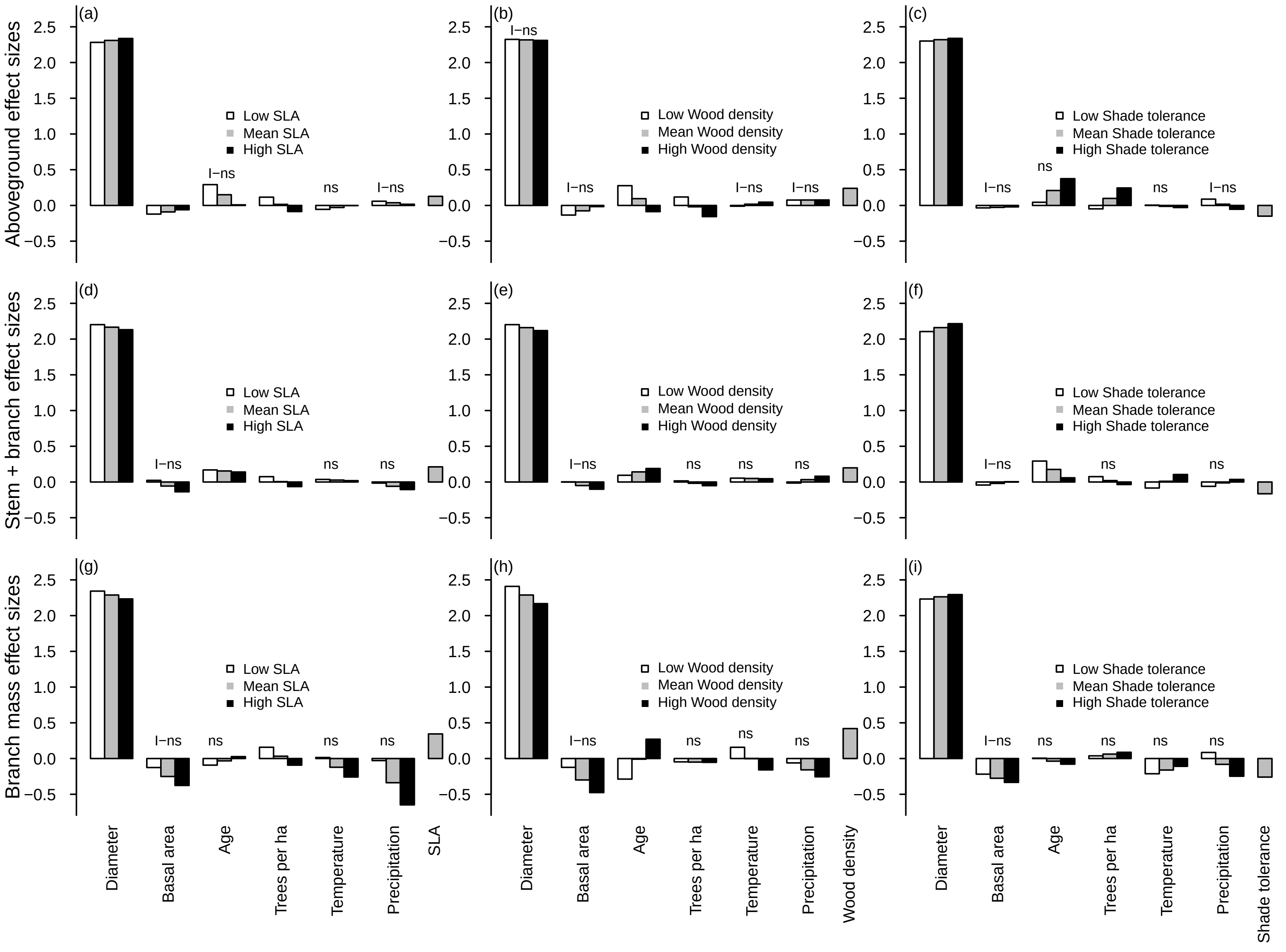

(a) Without error
2752
1835
917
0
0
20
40
60
1
2
3
4
(b) With error
2752
1835
917
0
0
20
40
60
1
2
3
4
Aboveground mass (kg)
Diameter (cm)


(a) F. sylvatica
(b) P. sylvestris
(c) F. sylvatica
(d) P. sylvestris
Foliage mass (kg)
Stem mass (kg)
Diameter (cm)
Diameter (cm)
0
28
56
84
0
12
23
35
0
1239
2479
3718
0
314
629
943
0
20
40
60
80
0
20
40
60
249
250
251
252
253
254
255
256
257
258
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
781
782
783
784
269
270
271
272
273
274
275
276
277
278
279
280
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808


Foliage mass (kg)
250
200
150
100
50
0
0
20
40
60
80
100
Diameter (cm)
20
15
10
5
0
0
10
20
30
40
Betula
Q. ilex
P. nigra
Alnus
S. aucuparia
R. pseudoacacia
P. menziesii
Conifers
P. abies
A. pseudoplatanus
F. excelsior
Combined
P. pinaster
Q. petraea or Q. robur
Broadleaved
A. alba
L. decidua
P. sylvestris
F. sylvatica
C. sativa
Populus
C. betulus
P.serotina


750
500
250
0
0
10
20
30
40
50
Populus
P. sylvestris
Betula
R. pseudoacacia
S. aucuparia
F. excelsior
A. pseudoplatanus
P. menziesii
L. decidua
C. betulus
Conifers
Combined
Broadleaved
F. sylvatica
A. alba
Q. ilex
Q. petraea or Q. robur
P. nigra
P. abies
C. sativa
Tilia
Alnus
P. pinaster
Prunus
5000
4000
3000
2000
1000
0
Stem mass (kg)
0
20
40
60
80
100
Diameter (cm)


Foliage mass effect sizes
Stem mass effect sizes
Root mass effect sizes
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
Low SLA
Mean SLA
High SLA
Low Wood density
Mean Wood density
High Wood density
Low Shade tolerance
Mean Shade tolerance
High Shade tolerance
I−ns
ns
Diameter
Basal area
Age
Trees per ha
Temperature
Precipitation
SLA
Wood density
Shade tolerance


Aboveground effect sizes
Stem + branch effect sizes
Branch mass effect sizes
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
Low SLA
Mean SLA
High SLA
Low Wood density
Mean Wood density
High Wood density
Low Shade tolerance
Mean Shade tolerance
High Shade tolerance
I−ns
ns
Diameter
Basal area
Age
Trees per ha
Temperature
Precipitation
SLA
Wood density
Shade tolerance

A database containing nearly 1000 European biomass equations was developed
Biomass and leaf area allometry were influenced by stand structure
Species traits were correlated with interspecific differences in responses to stand structure

Biomass and leaf area allometric equations for European tree species incorporating stand structure, tree age and climate
David Forrester et al.

Appendix B

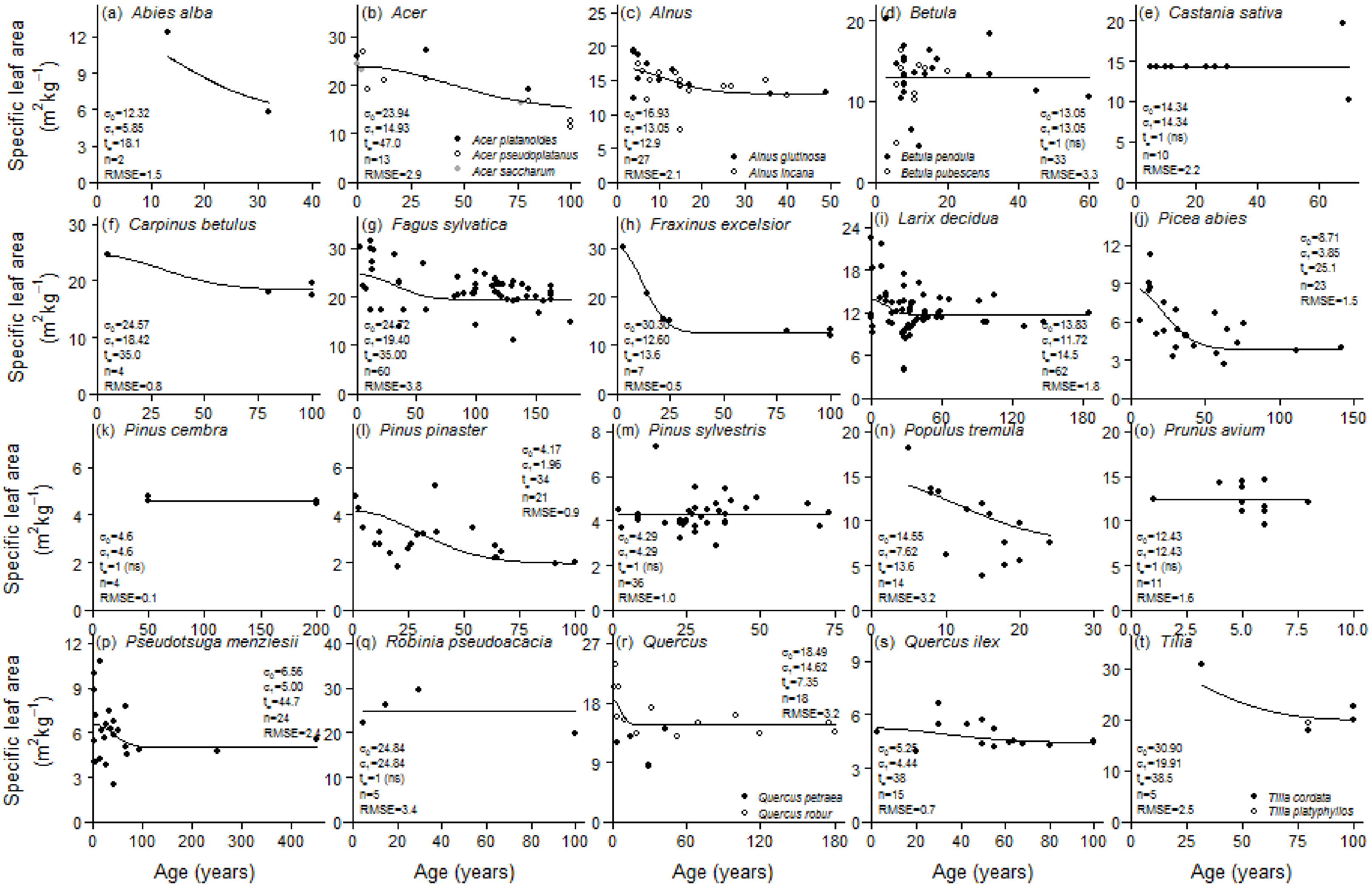


Figure B.1. Relationship between specific leaf area and age. Fitted lines and the parameters shown in each panel (σ_0, σ_1, t_σ; see Equation 2) are based on Equation 2 and the data collected from the literature that is provided in Table A.1 in Forrester et al. (in review). n is the sample size and RMSE is the Root Mean Square Error (Equation 26).

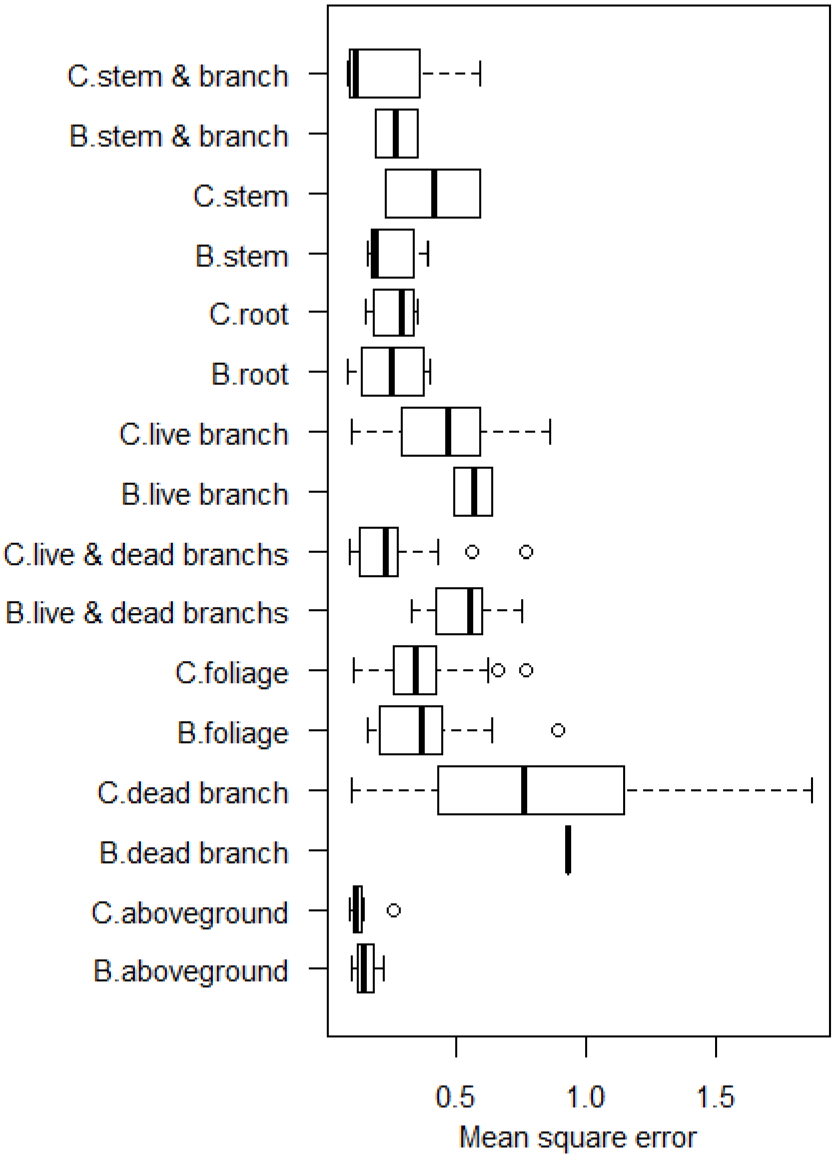


Figure B.2. Boxplot of the mean square error (MSE) of the published equations in Table A.2 in Forrester et al. (in review), when available. The MSE for all raw data sets were also calculated (and used in this figure) by fitting the raw data to Equation 3. On the y-axis C = conifer and B = broadleaved.

Table B.3. Shade tolerance, specific leaf area and wood density values for each species used in this study, and their sources.

Species	Shade tolerance (Niinemets and Valladares, 2006)	Specific leaf area (m^2 kg^{-1}) (from Table A.1 in Forrester et al., in review)	Wood density (g cm^{-3})*	Source for wood density within Global wood density database*
Abies alba	4.6	9.08	0.353	(Brzeziecki and Kienast, 1994)
Acer pseudoplatanus	3.73	19.51	0.508	(Forestry Compendium; Brzeziecki and Kienast, 1994)
Alnus glutinosa	2.71	14.93	0.439	(Brzeziecki and Kienast, 1994)
Alnus incana	2.3	14.93	0.439	(Brzeziecki and Kienast, 1994)
Betula pendula	2.03	9.84	0.525	(Brzeziecki and Kienast, 1994)
Betula pubescens	1.85	9.84	0.525	(Brzeziecki and Kienast, 1994)
Carpinus betulus	3.97	19.83	0.706	(Brzeziecki and Kienast, 1994; Schütt et al., 1994)
Castanea sativa	3.15	14.34	0.463	(Gutierrez Oliva and Plaza Pulgar, 1967; Lavers et al., 1983; Rijsdijk and Laming, 1994)
Fagus sylvatica	4.56	21.54	0.585	(Brzeziecki and Kienast, 1994)
Fraxinus excelsior	2.66	17.08	0.560	(Brzeziecki and Kienast, 1994)
Larix decidua	1.46	12.18	0.474	(Brzeziecki and Kienast, 1994)
Picea abies	4.45	5.76	0.370	(Brzeziecki and Kienast, 1994)
*Pinus cembra***	2.87	4.6	0.417	
Pinus nigra	2.1	na	0.417	(Gutierrez Oliva and Plaza Pulgar, 1967; Lavers et al., 1983)
*Pinus pinaster****	2.21	2.97	0.412	(Gutierrez Oliva and Plaza Pulgar, 1967; Lavers et al., 1983)
Pinus sylvestris	1.67	4.28	0.422	(Brzeziecki and Kienast, 1994)
Populus alba	2.3	9.84	0.353	(Brzeziecki and Kienast, 1994)
Populus tremula	2.22	9.84	0.374	(Forestry Compendium; Brzeziecki and Kienast, 1994)
Prunus avium	3.33	12.43	0.474	(Brzeziecki and Kienast, 1994)
Prunus serotina	2.46	na	0.684	(Annighöfer et al., 2012)
Pseudotsuga menziesii	2.78	5.98	0.428	(Forestry Compendium; Alden, 1997)
Quercus ilex	3.02	4.83	0.820	(Gutierrez Oliva and Plaza Pulgar, 1967)
Quercus petraea	2.73	14.81	0.560	(Brzeziecki and Kienast, 1994)
Quercus robur	2.45	14.81	0.560	(Brzeziecki and Kienast, 1994)
Robinia pseudoacacia	1.72	24.84	0.675	(Cheng et al., 1992; Alden, 1995)
Sorbus aucuparia	2.73	na	0.629	(Brzeziecki and Kienast, 1994)
Tilia cordata	4.18	22.11	0.422	(Brzeziecki and Kienast, 1994)

*The sources for wood density values were obtained from the wood density data base (Zanne et al., 2009)

***P. cembra* was assumed to have the mean wood density of the other *Pinus* used in this study.

****P. pinaster* was assumed to have the mean shade tolerance of the other *Pinus* used in this study.

An “na” in the specific leaf area column indicates that specific leaf area data could not be obtained for these species.

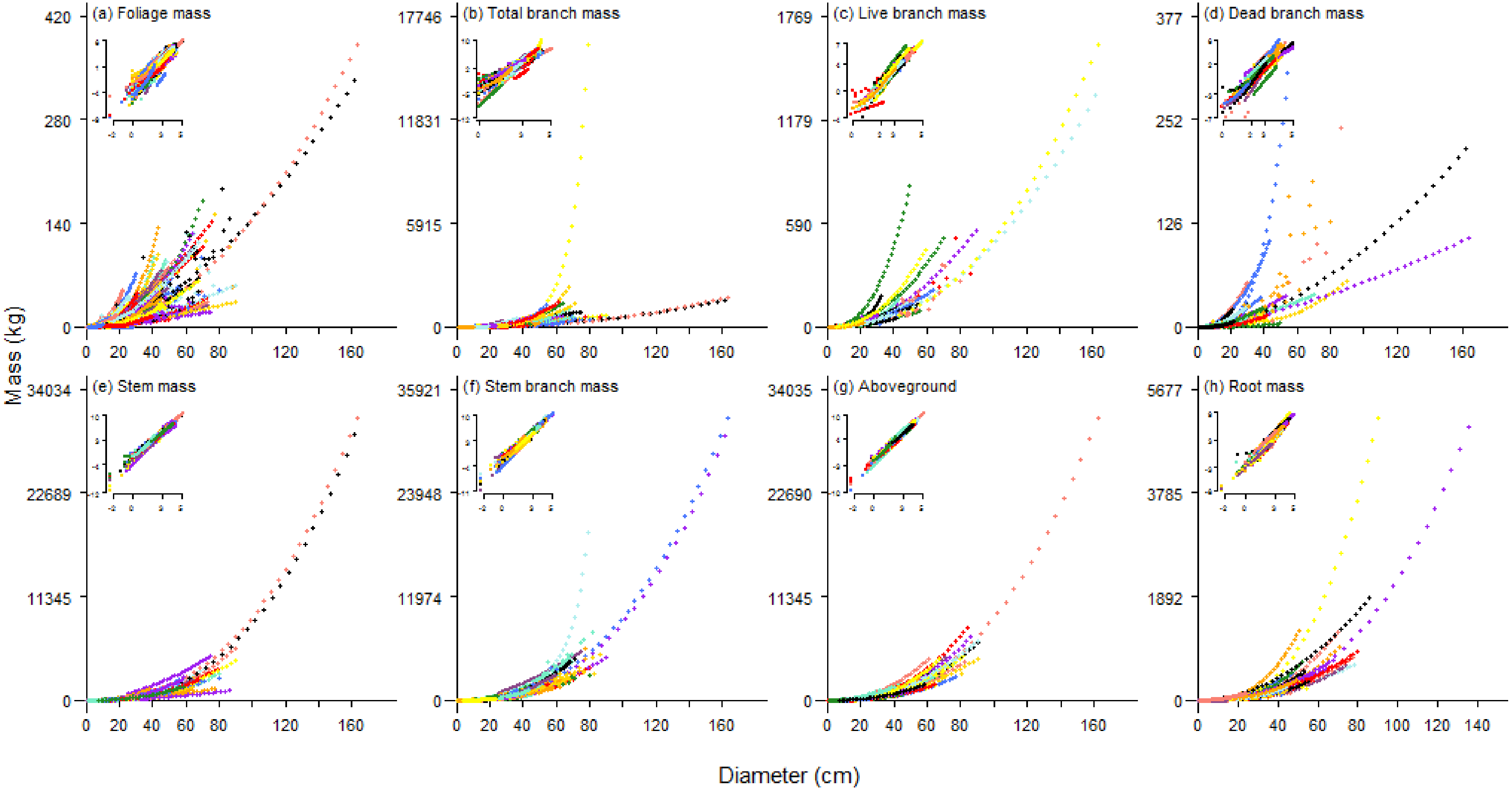


Figure B.3. The pseudo-observations for all species and all biomass components against stem diameter. The data is provided in the database (Table A.2 in Forrester et al. in review), which also identifies the sources of the published equations used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

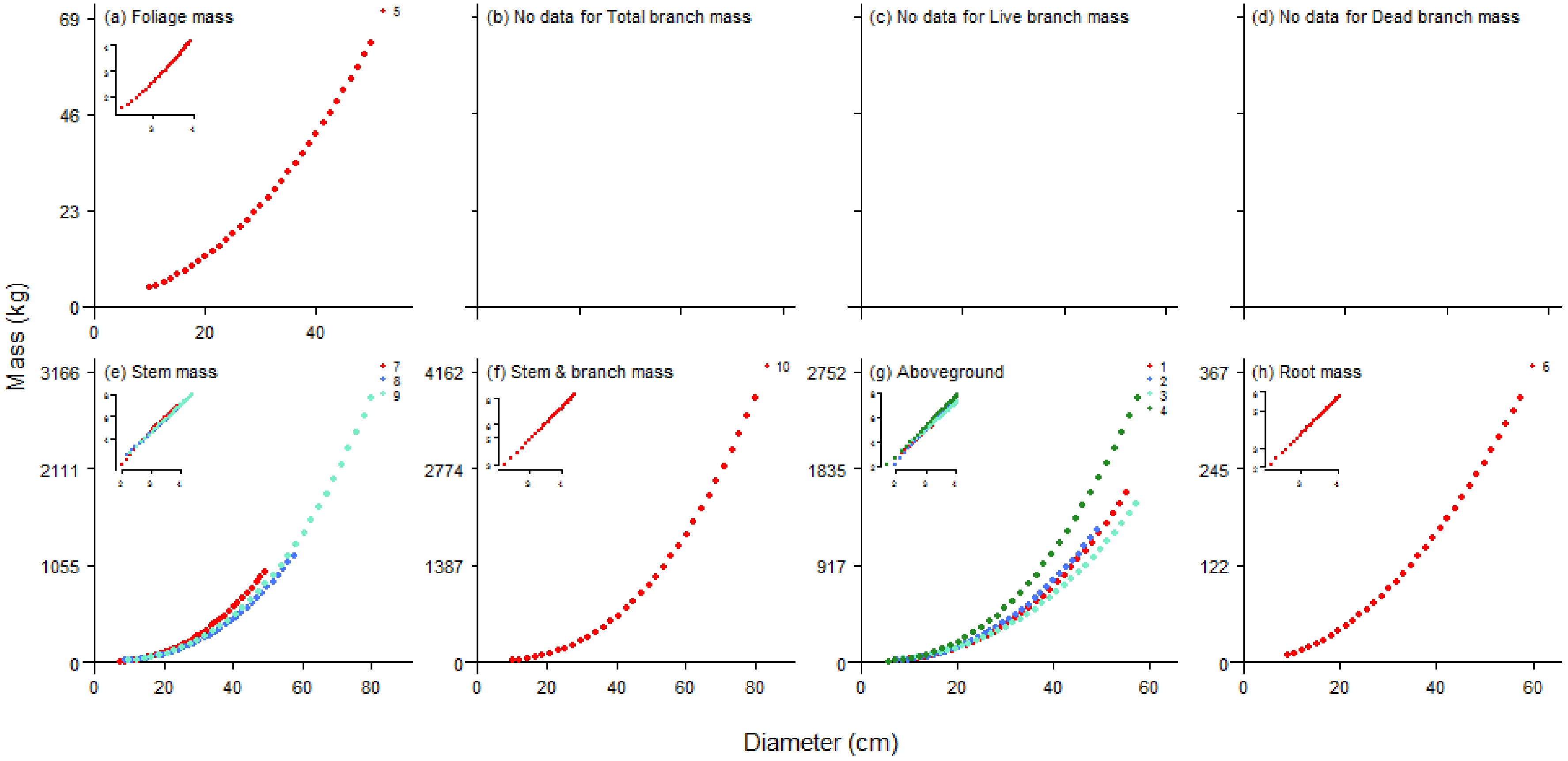


Figure B.4. The pseudo-observations for ***Abies alba*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (b,c,d) indicate that no data were available.

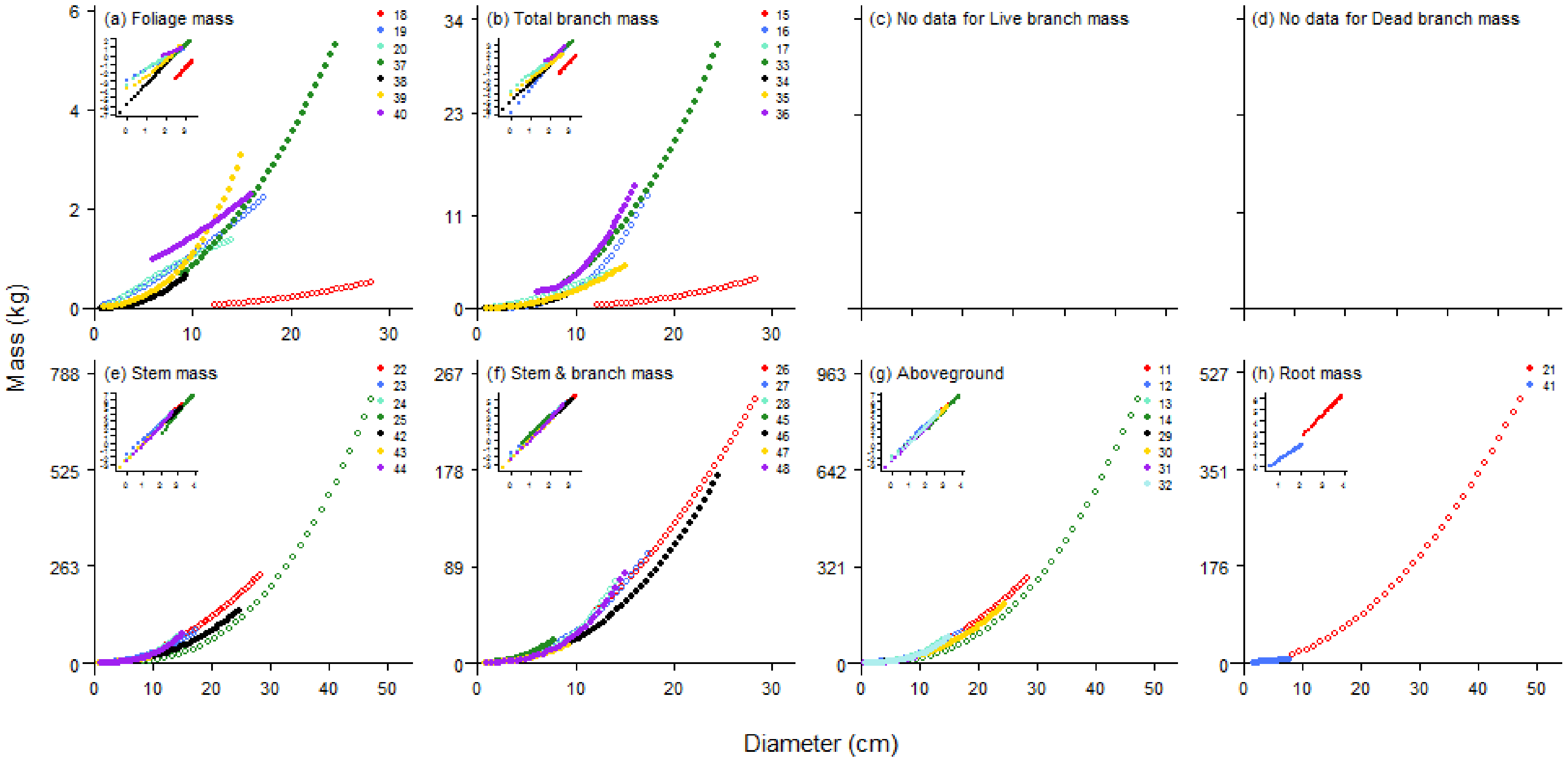


Figure B.5. The pseudo-observations for ***Alnus glutinosa*** (open circles) and ***Alnus incana*** (closed circles) biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

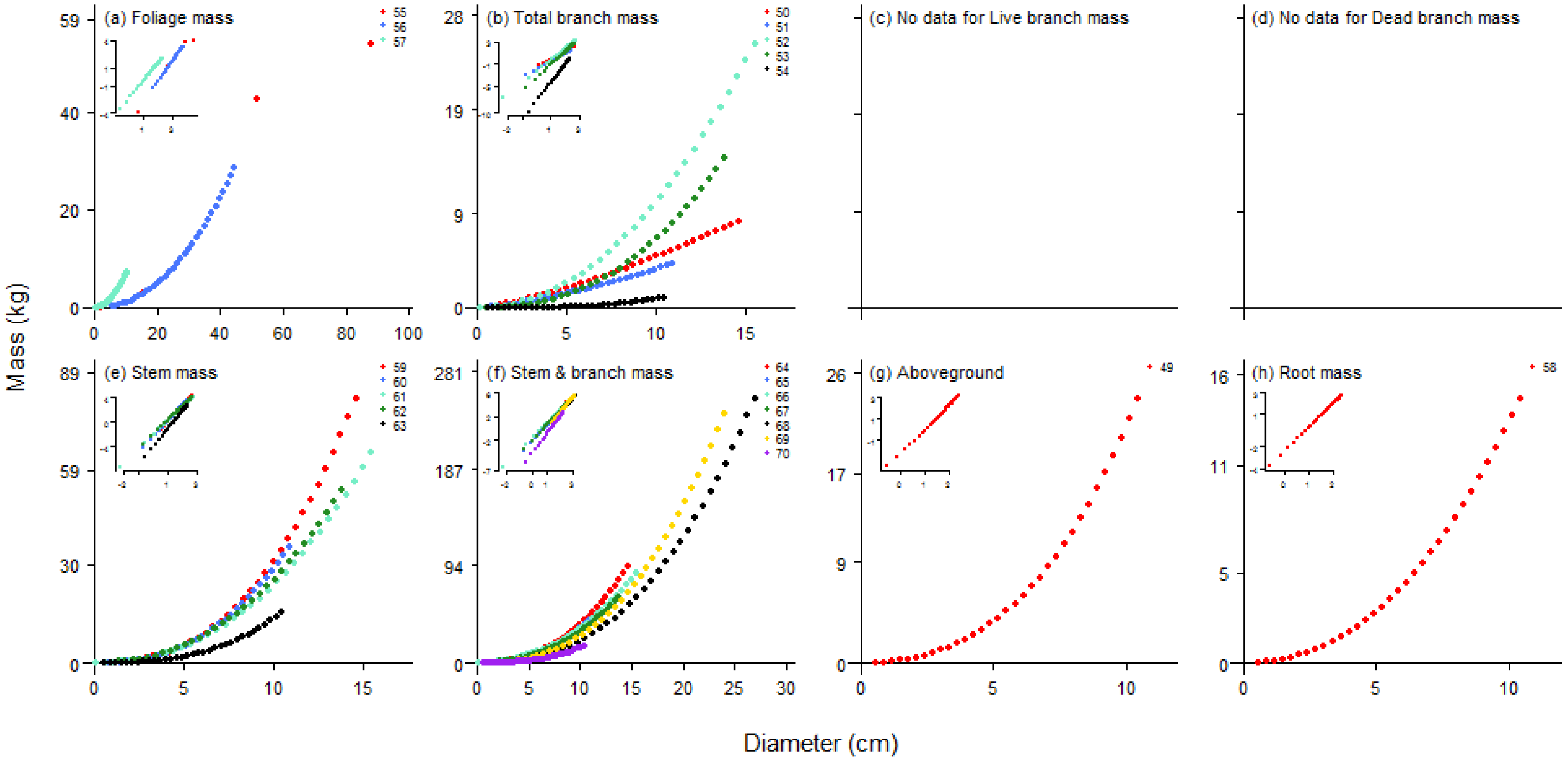


Figure B.6. The pseudo-observations for ***Acer pseudoplatanus*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

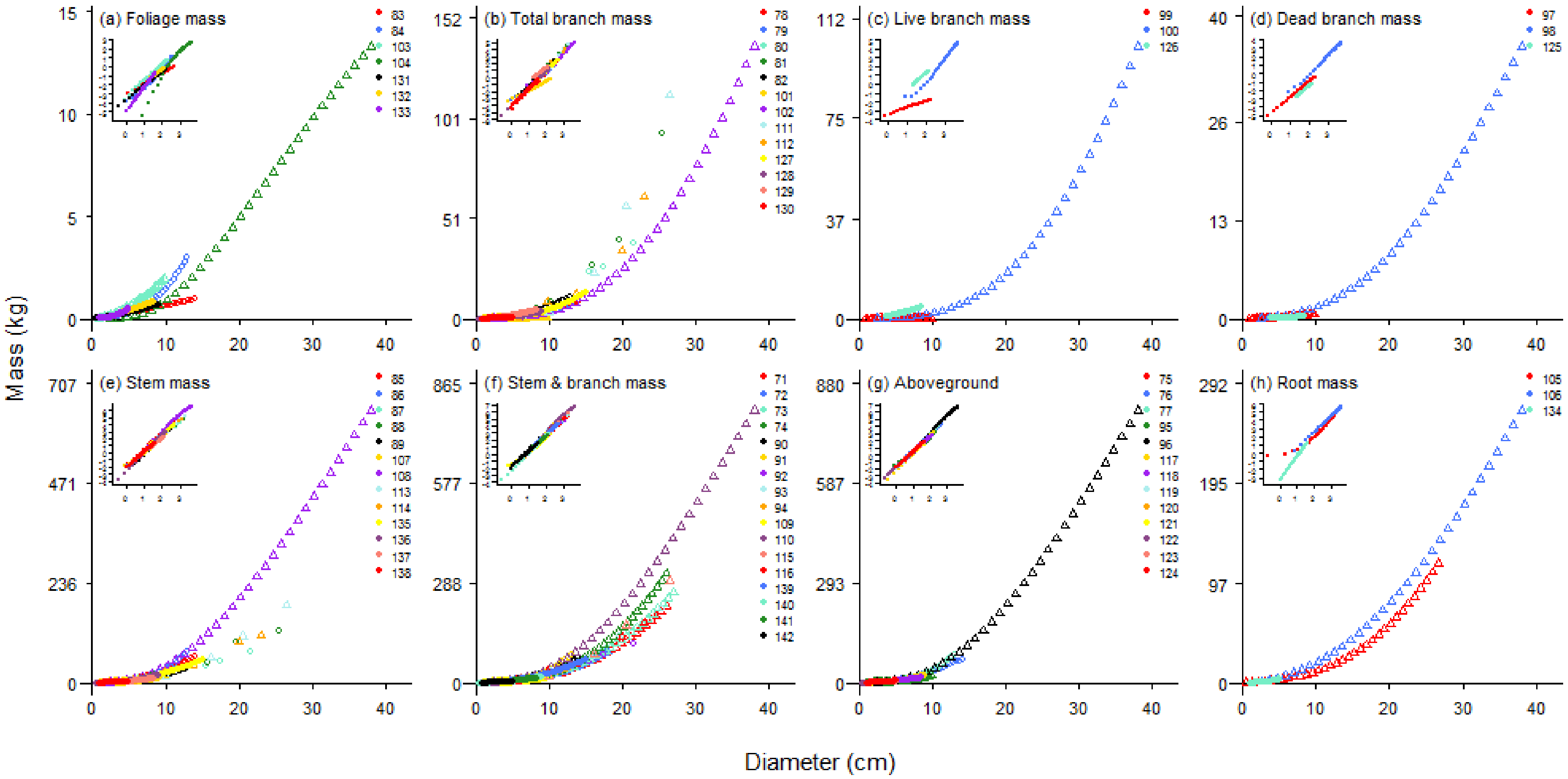


Figure B.7 The pseudo-observations for the biomass components of ***Betula pendula*** (open circles), ***Betula pubescens*** (closed circles) and sources that did not differentiate between these two species (triangles) plotted against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

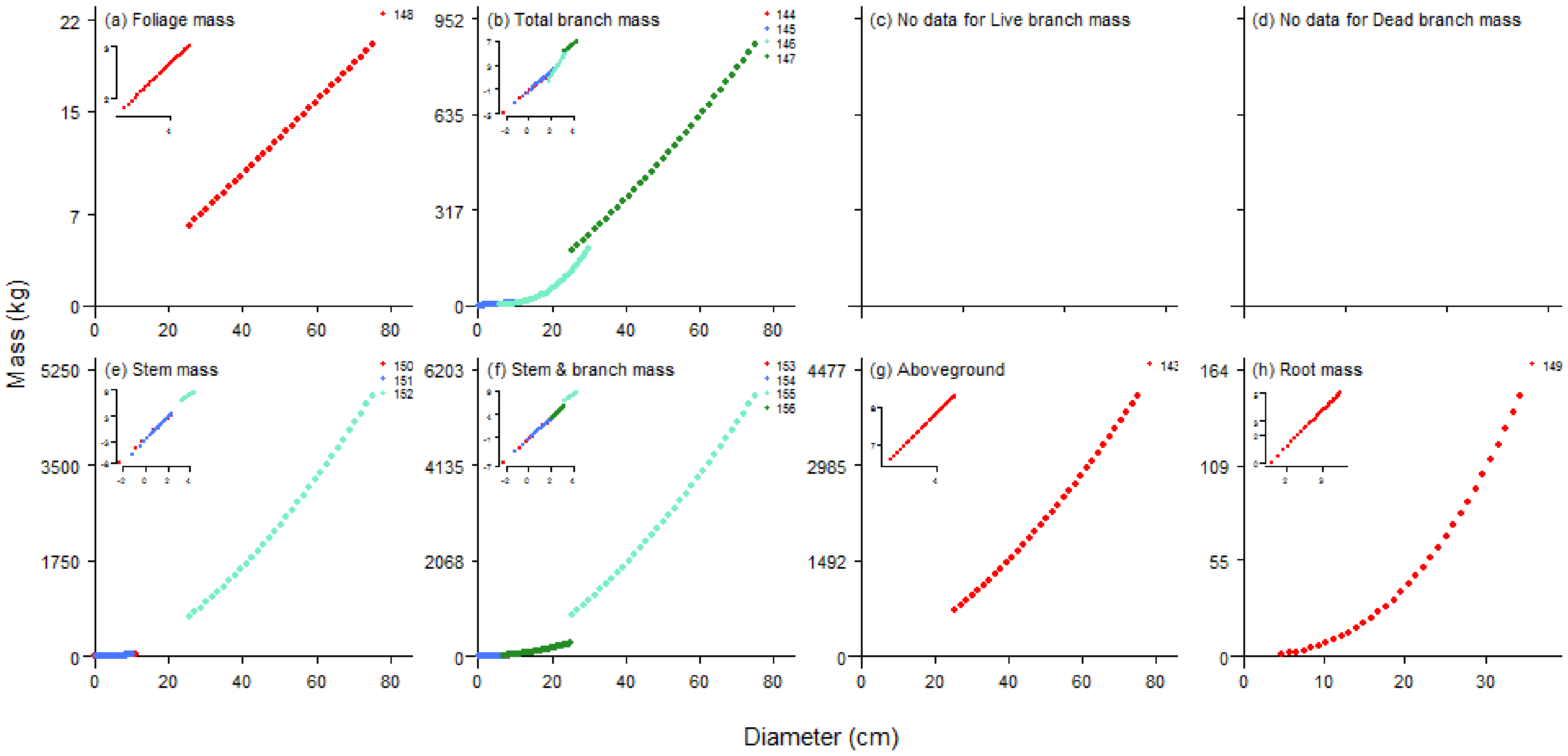


Figure B.8. The pseudo-observations for ***Carpinus betulus*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

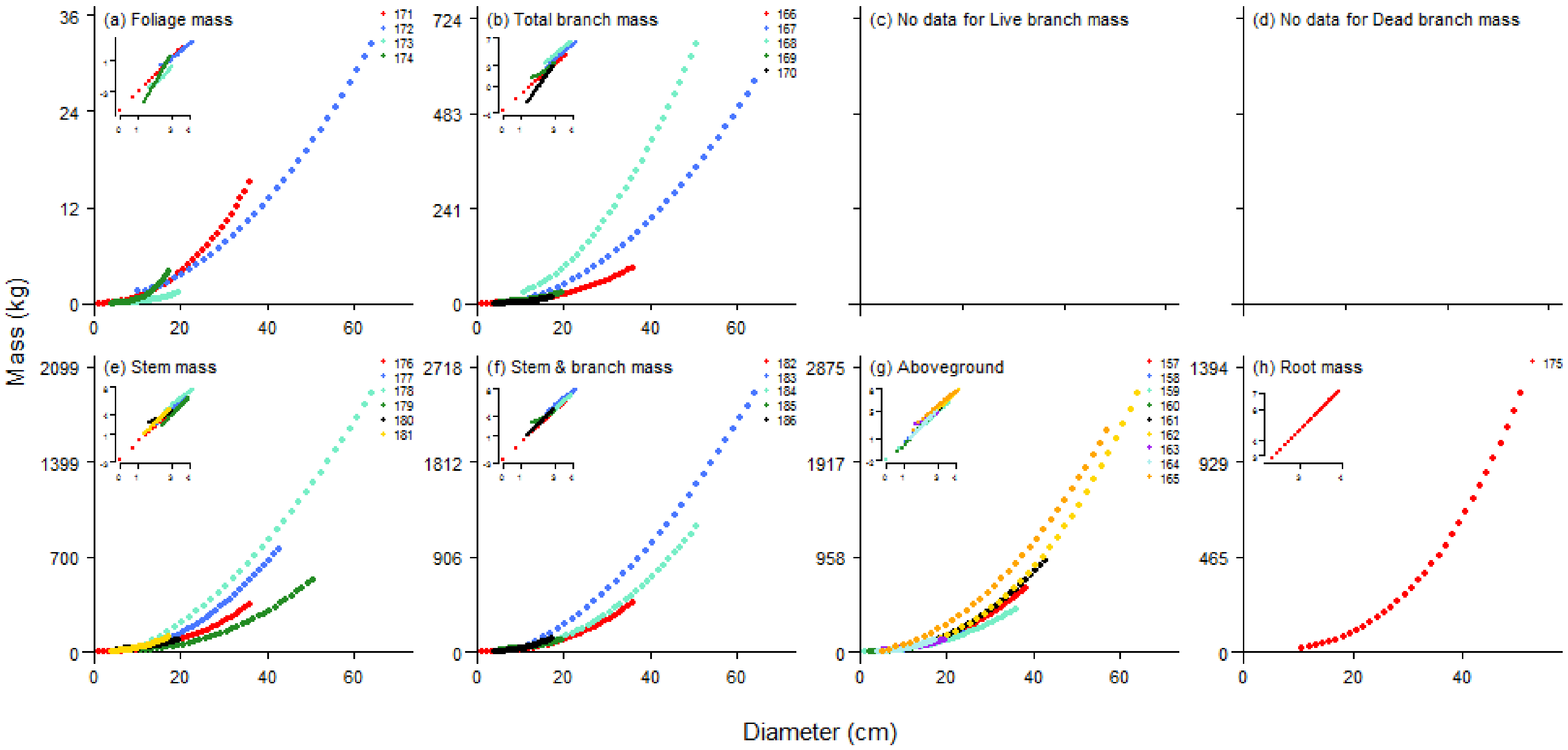


Figure B.9. The pseudo-observations for ***Castania sativa*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

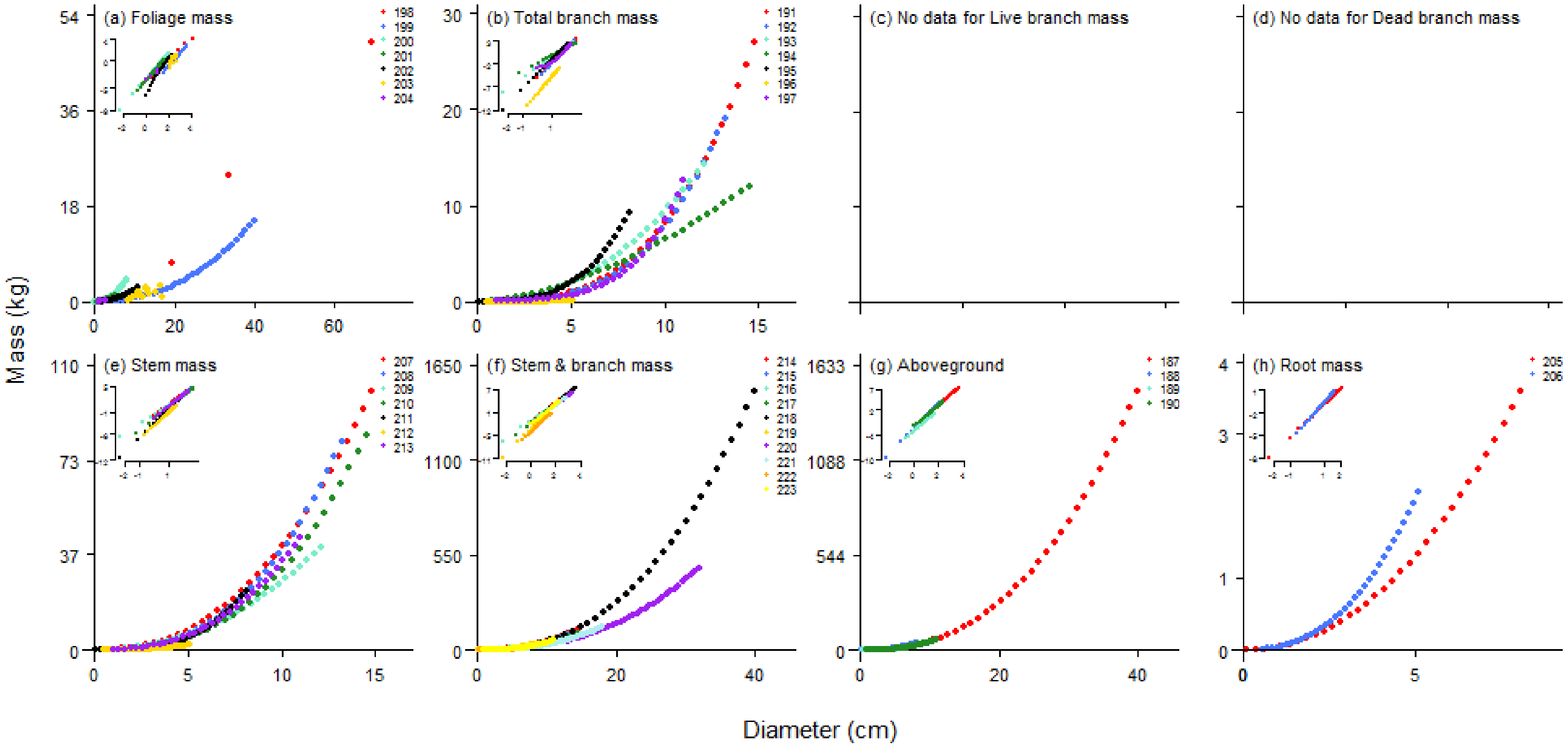


Figure B.10. The pseudo-observations for ***Fraxinus excelsior*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

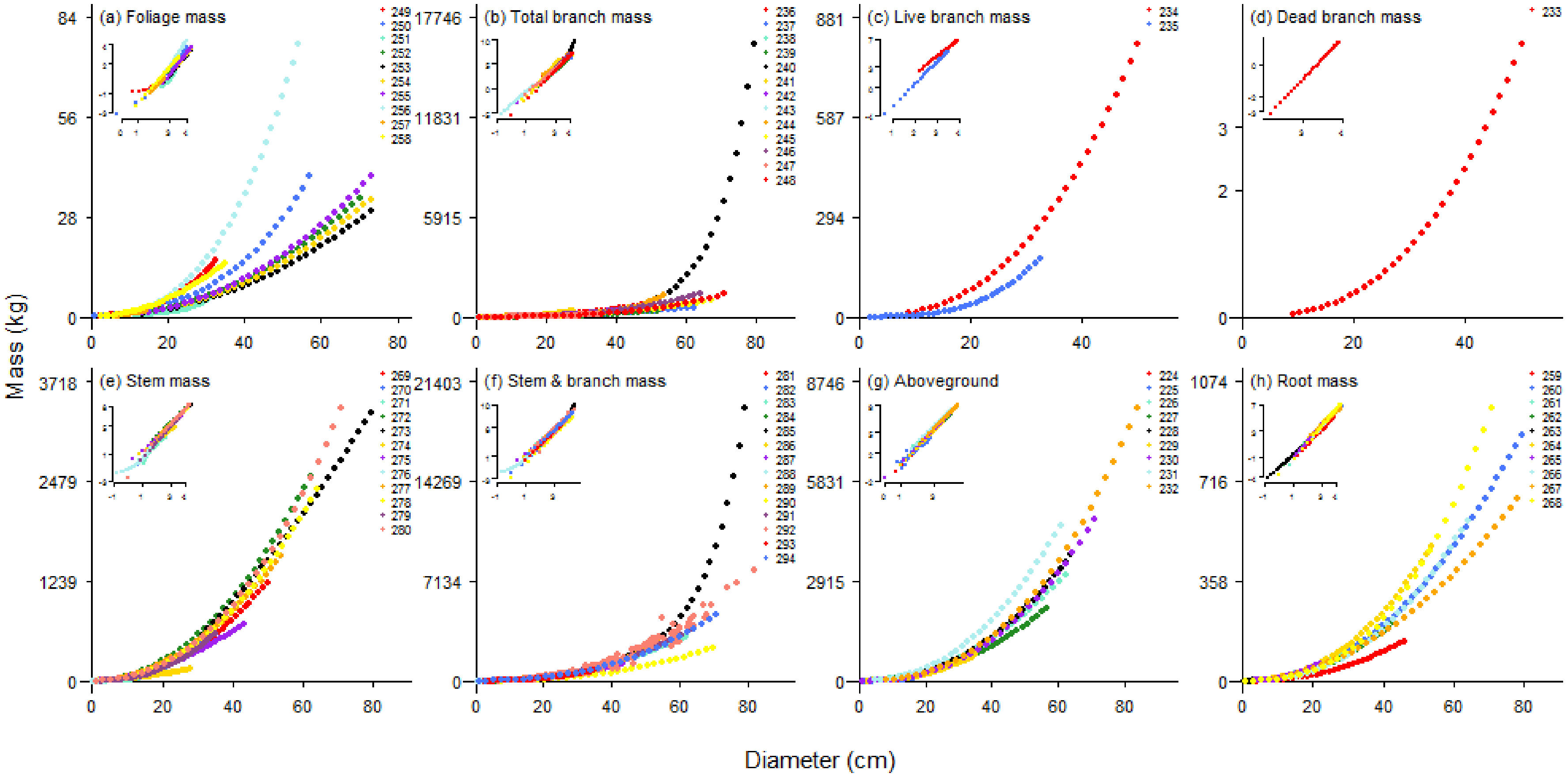


Figure B.11. The pseudo-observations for ***Fagus sylvatica*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

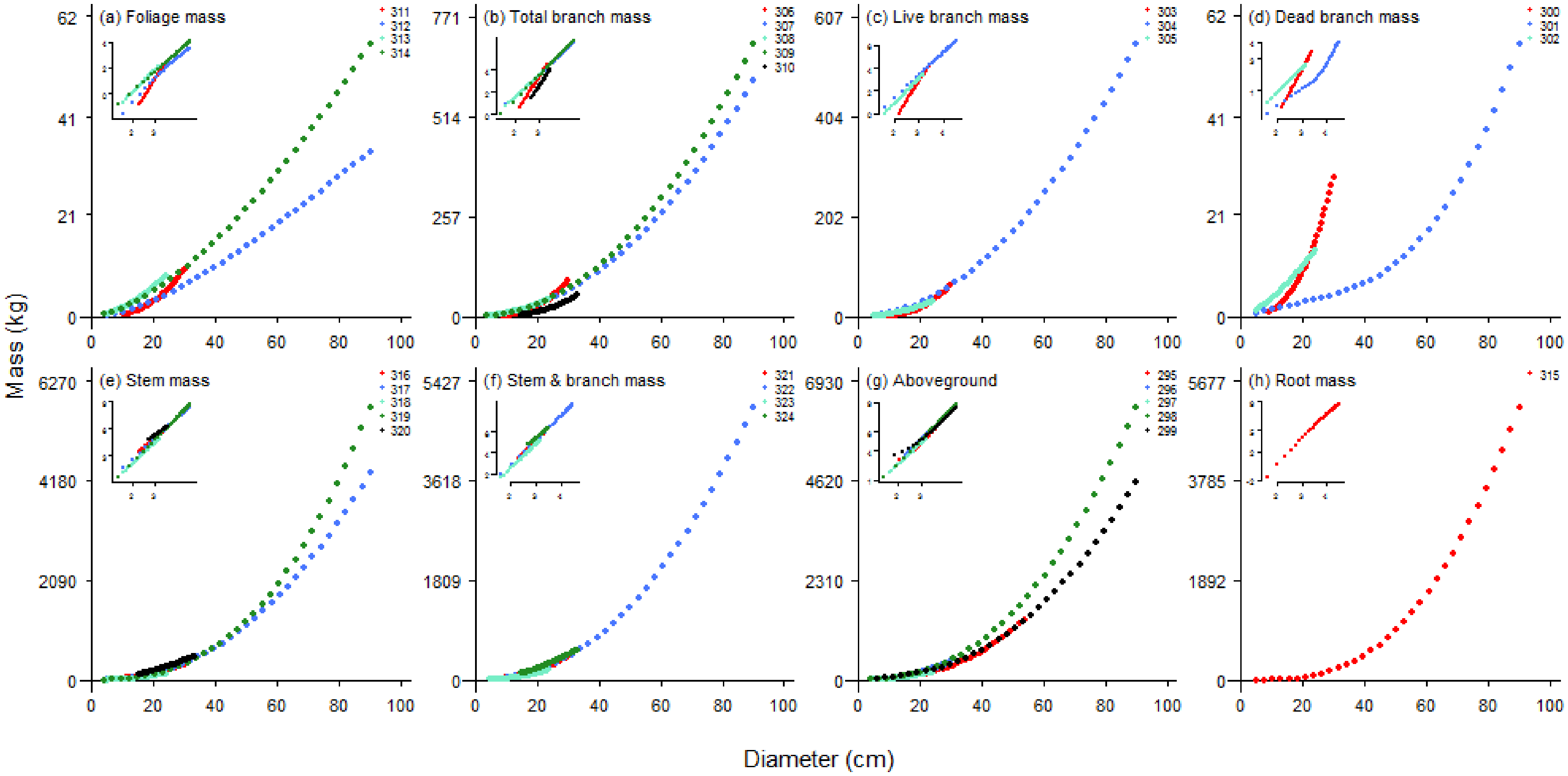


Figure B.12. The pseudo-observations for ***Larix decidua*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

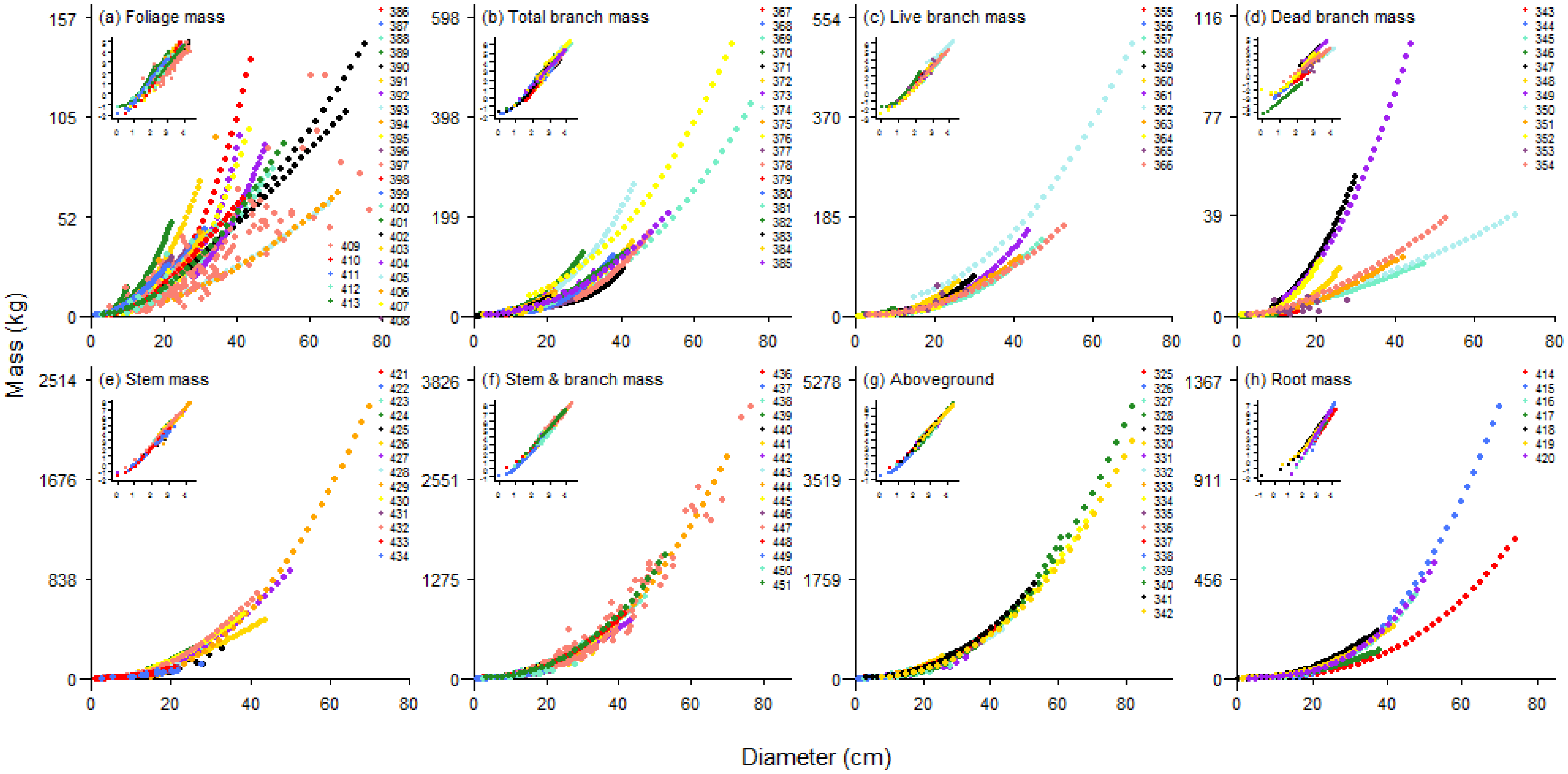


Figure B.13. The pseudo-observations for ***Picea abies*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

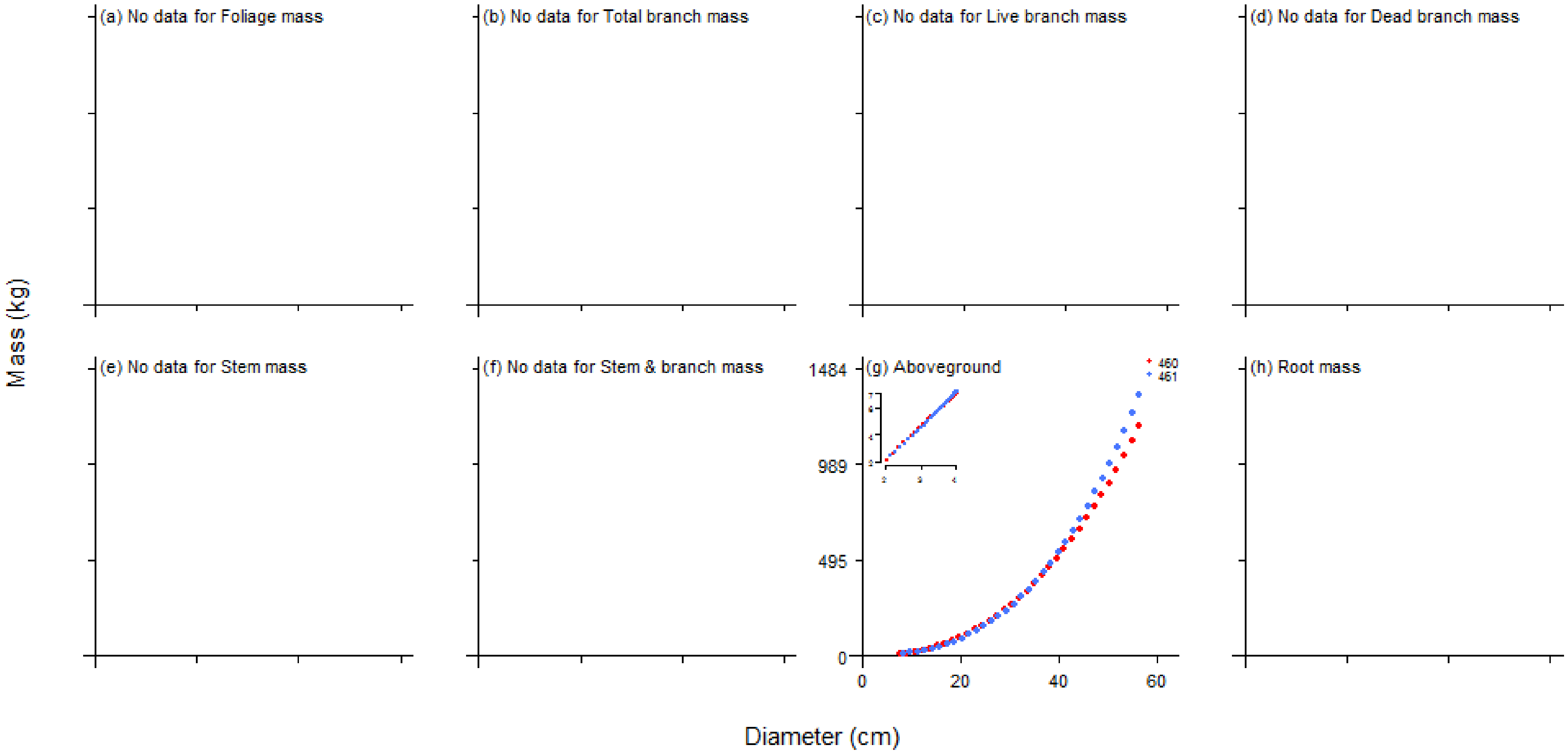


Figure B.14. The pseudo-observations for ***Pinus cembra*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (all except g) indicate that no data were available.

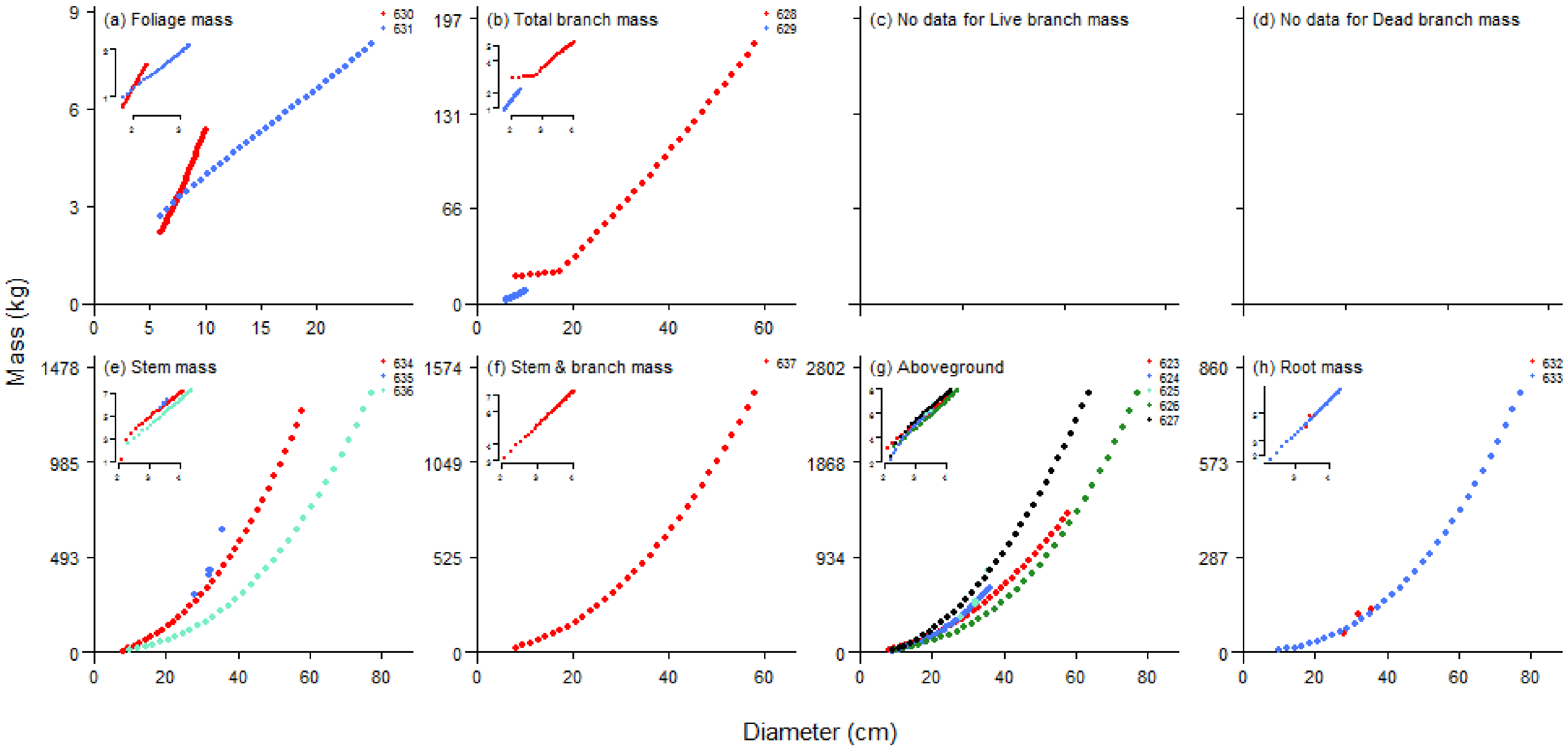


Figure B.15. The pseudo-observations for ***Pinus nigra*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

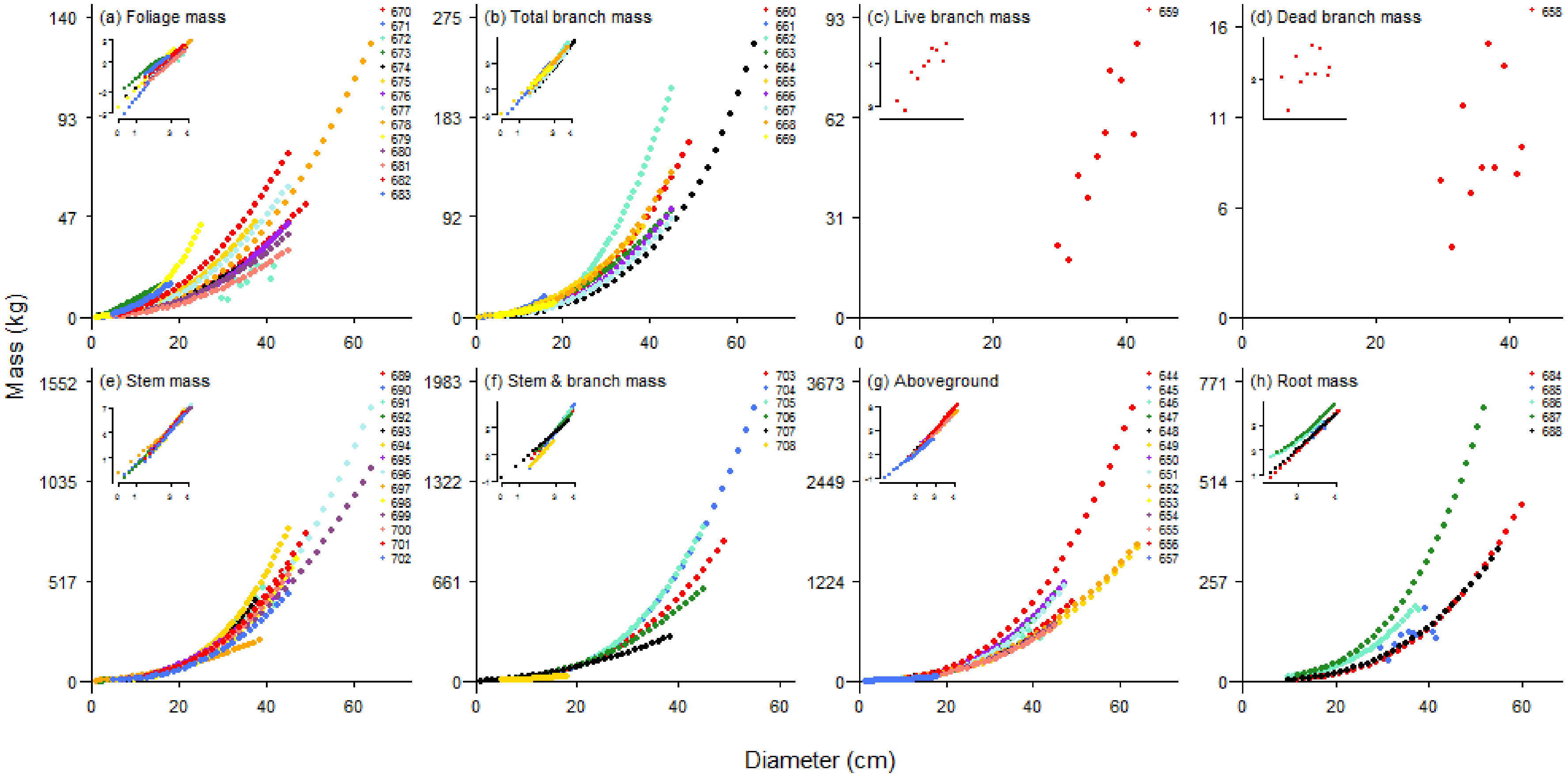


Figure B.16. The pseudo-observations for ***Pinus pinaster*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

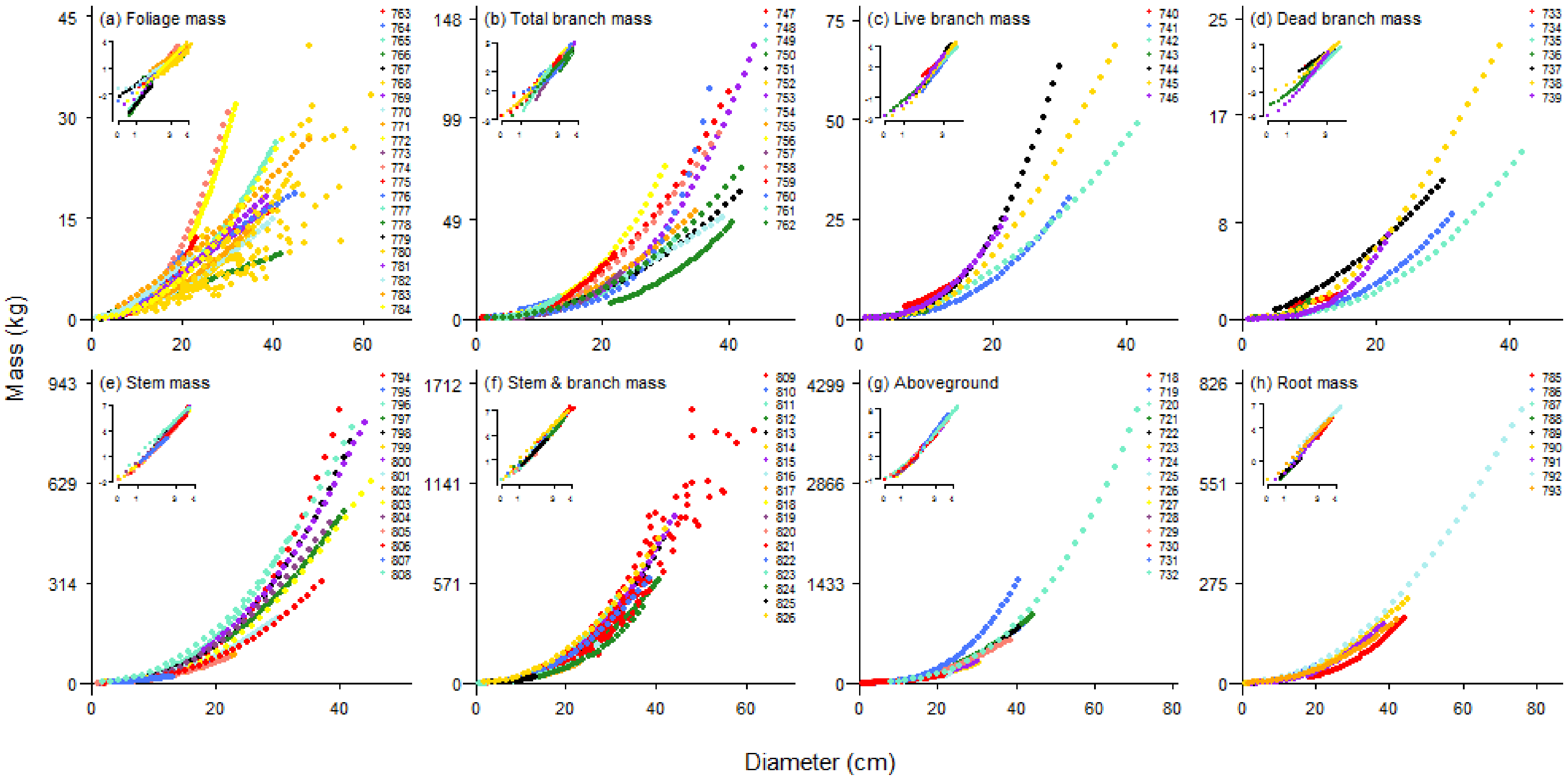


Figure B.17. The pseudo-observations for ***Pinus sylvestris*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

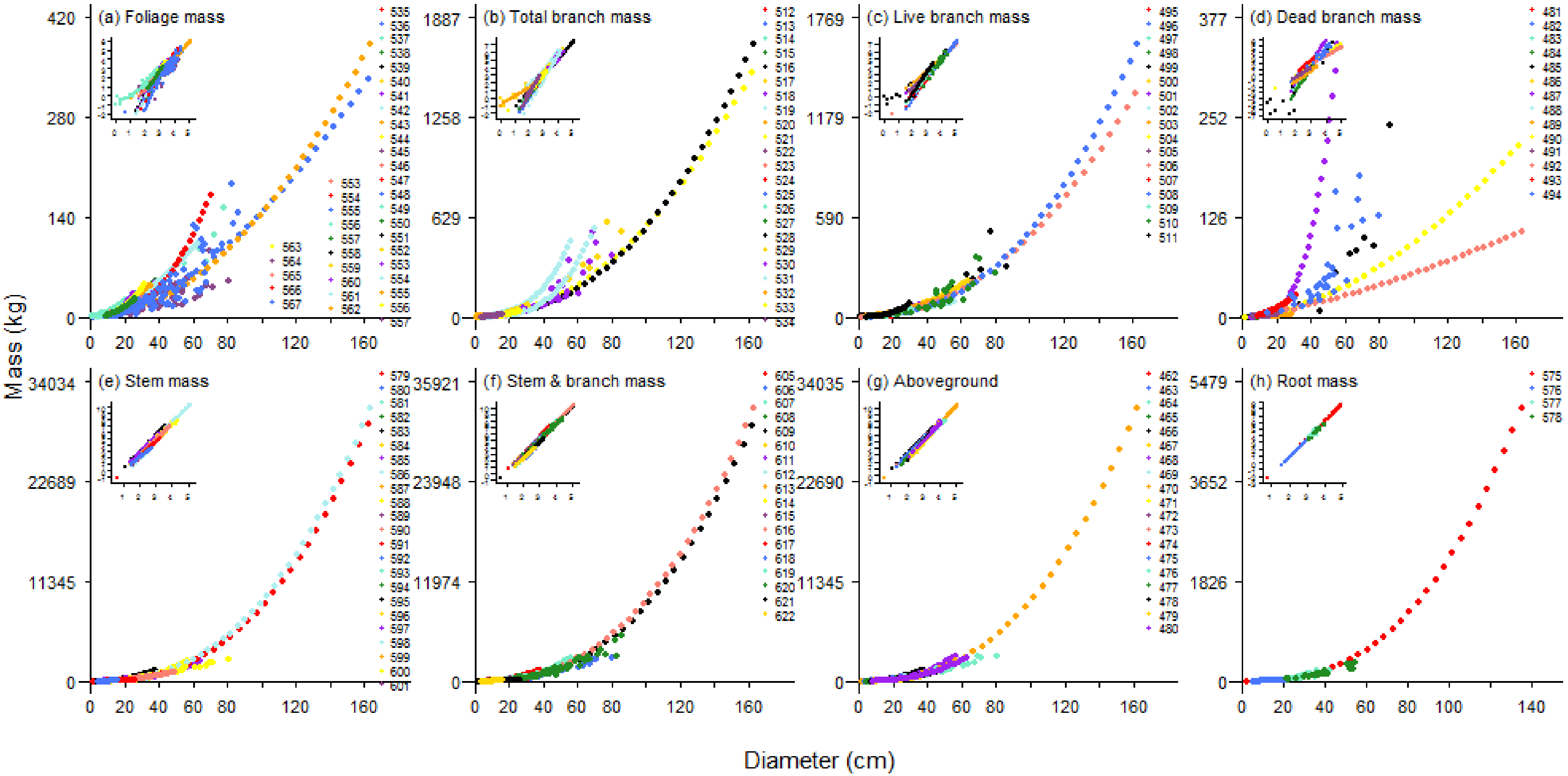


Figure B.18. The pseudo-observations for ***Pseudotzuga menziesii*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale.

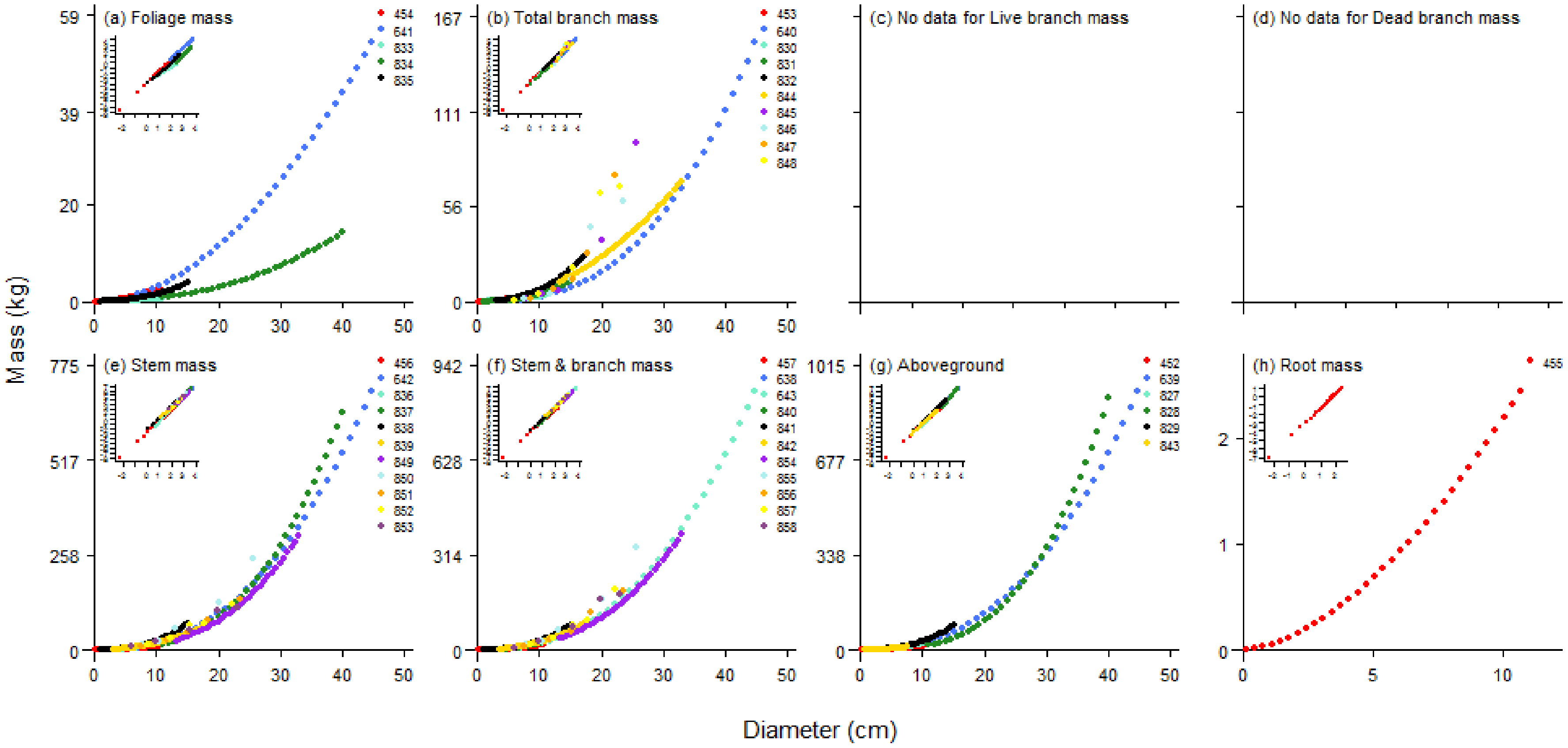


Figure B.19. The pseudo-observations for ***Populus (P. nigra, P. tremula, P. tremula x tremuloides)*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

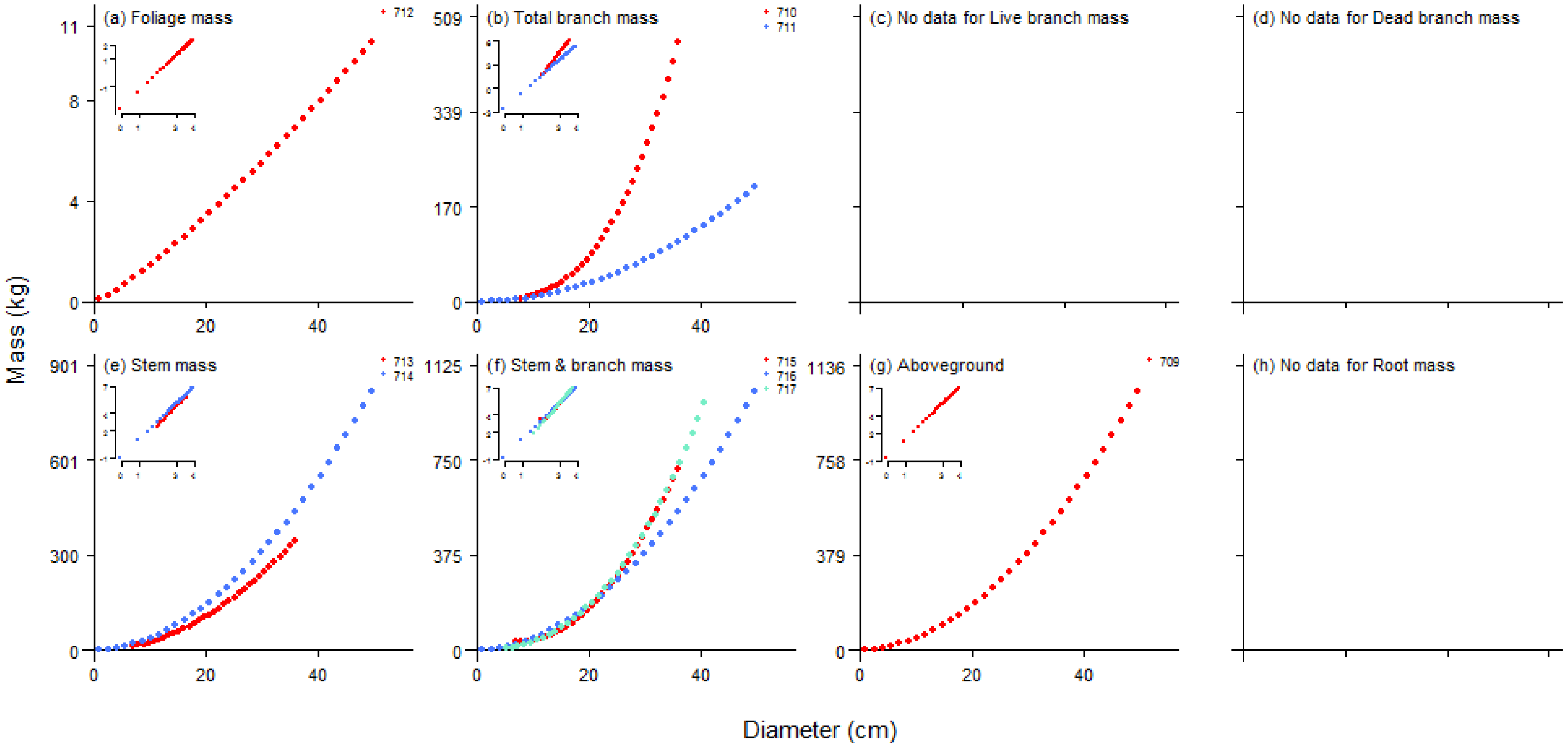


Figure B.20. The pseudo-observations for ***P. serotina*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d,h) indicate that no data were available.

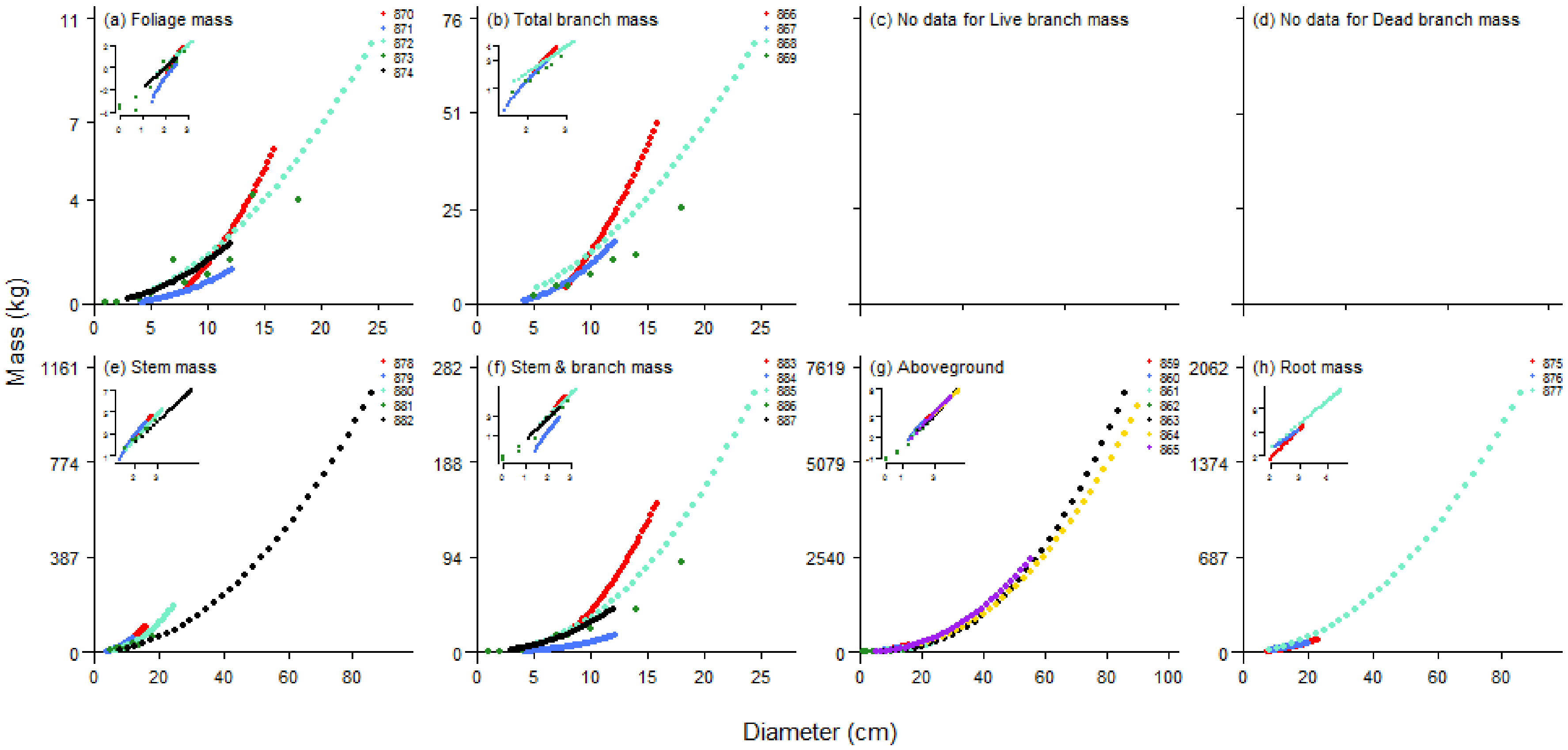


Figure B.21. The pseudo-observations for ***Quercus ilex*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

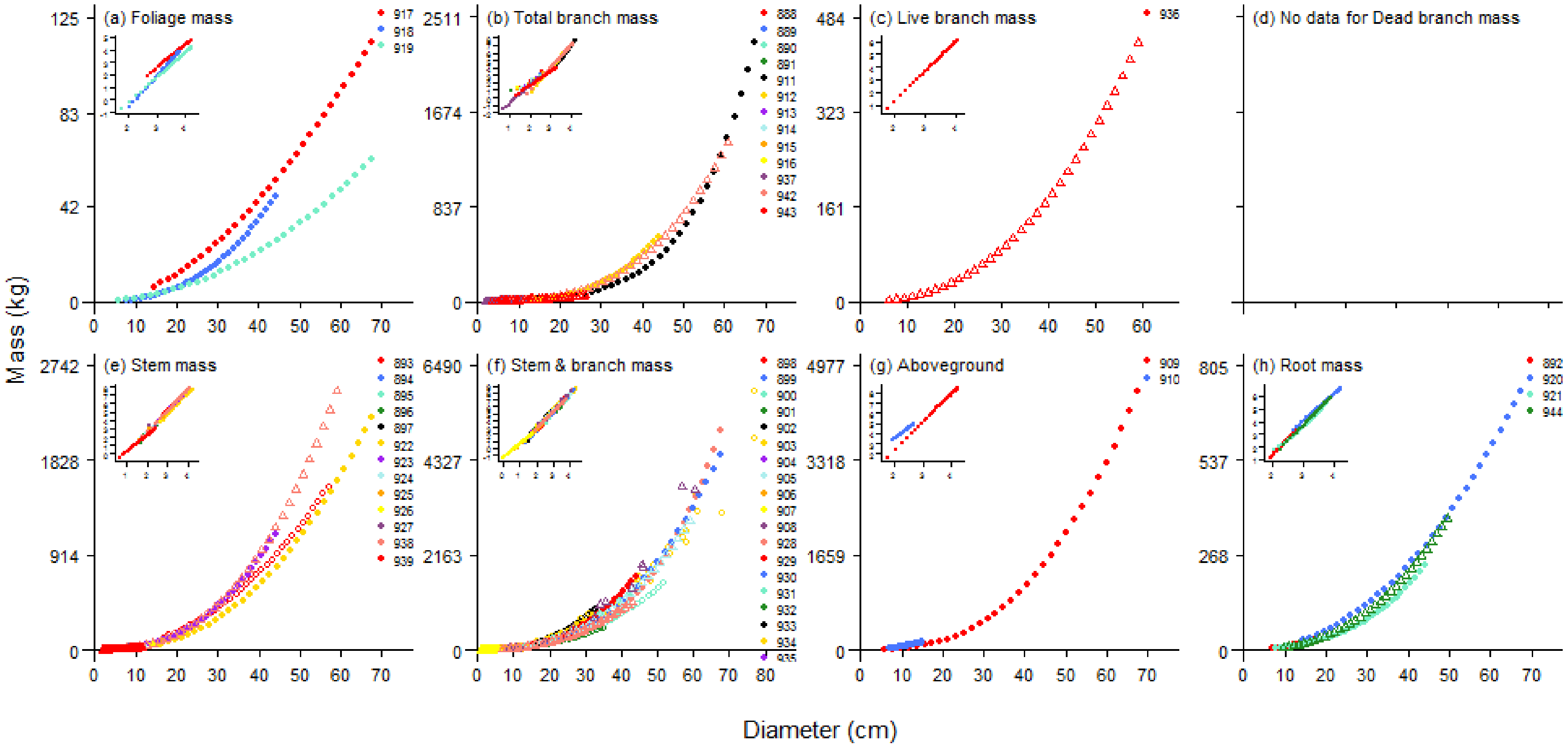


Figure B.22. The pseudo-observations for the biomass components of ***Quercus petraea*** (open circles), ***Quercus robur*** (closed circles) and sources that did not differentiate between these two species (triangles) plotted against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. The empty pane (d) indicates that no data were available.

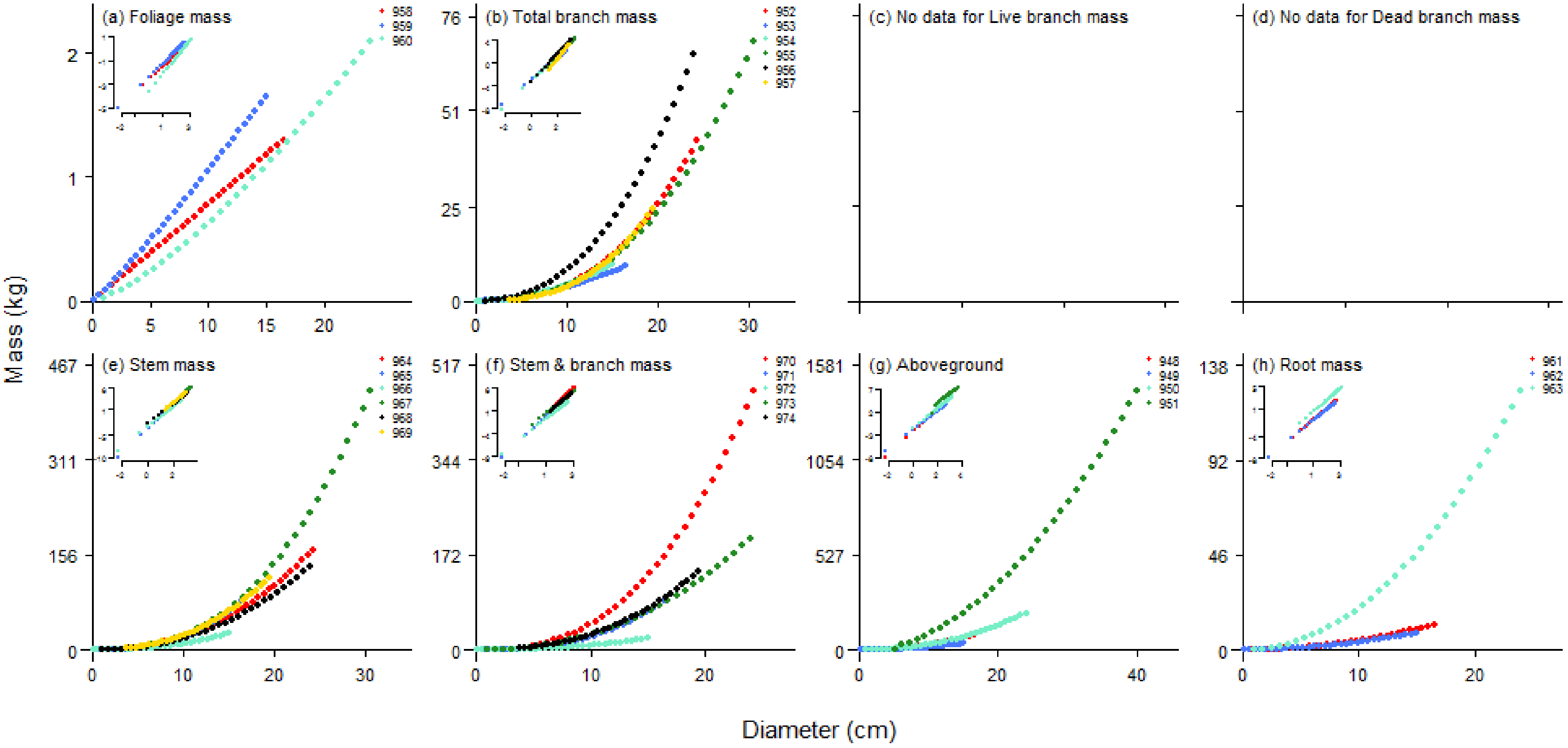


Figure B.23. The pseudo-observations for ***Robinia pseudoacacia*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d) indicate that no data were available.

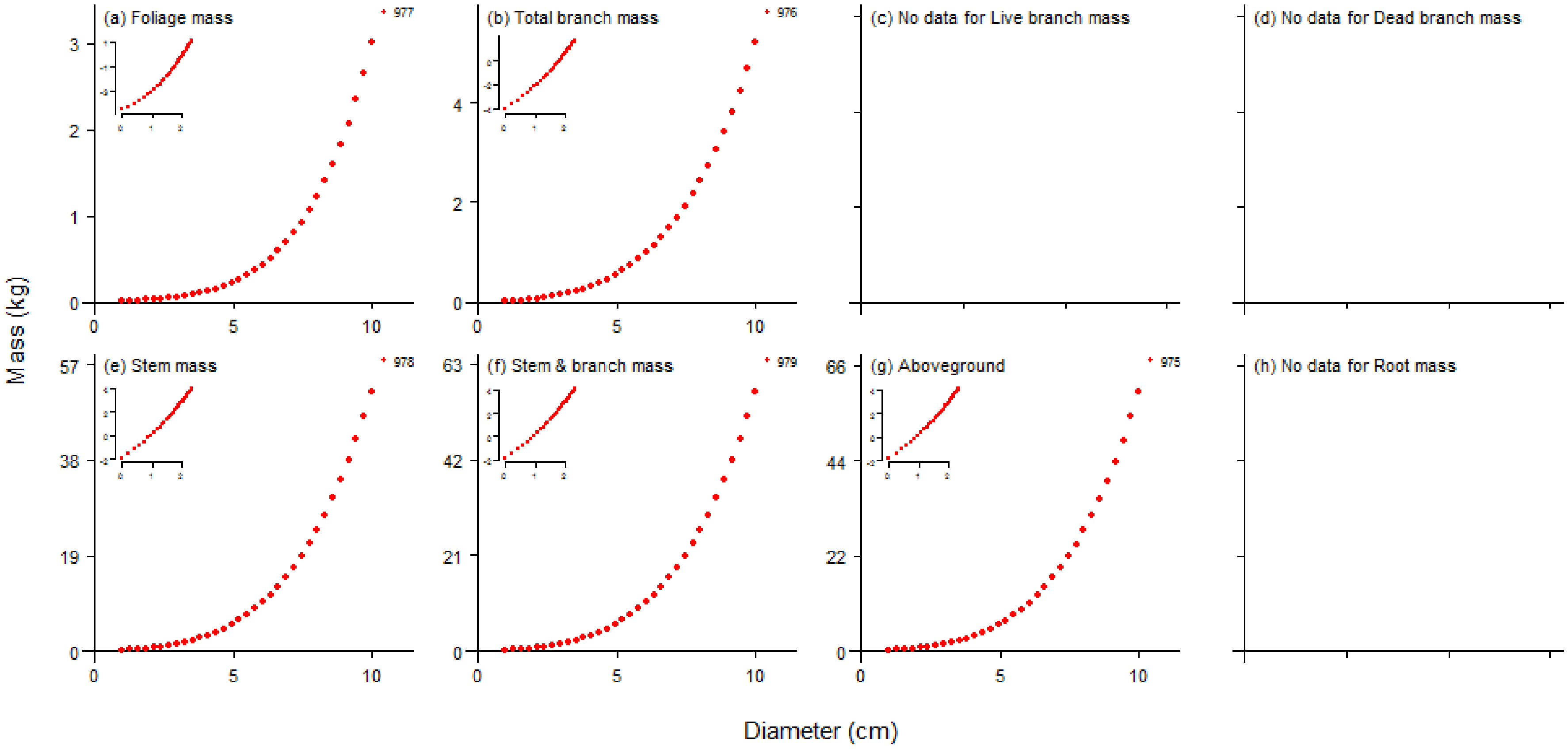


Figure B.24. The pseudo-observations for ***Sorbus aucuparia*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes (c,d,h) indicate that no data were available.

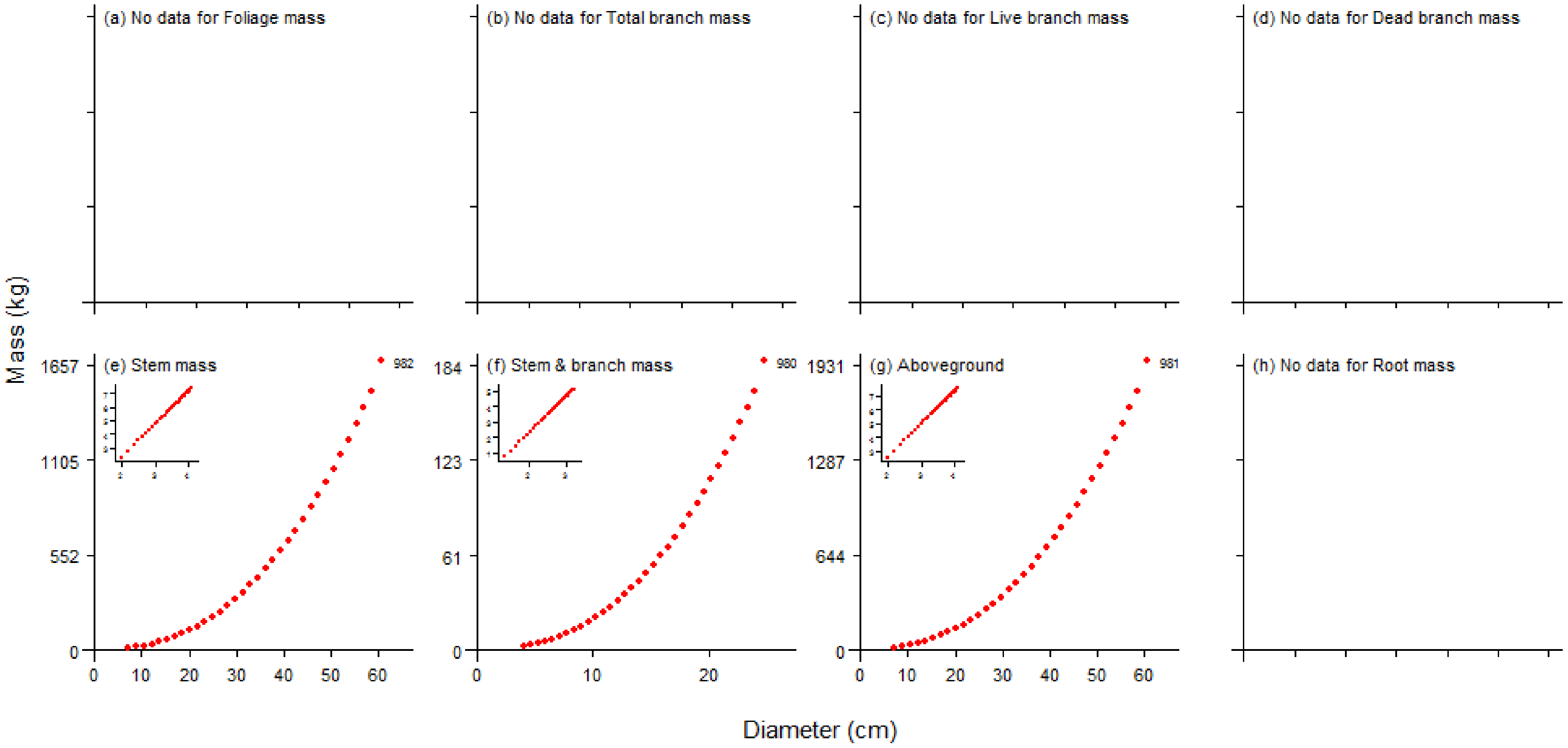


Figure B.25 The pseudo-observations for ***Tilia cordata*** or ***Tilia platyphyllos*** biomass components against stem diameter. The numbers in the legends represent the equation number in the database (Table A.2 in Forrester et al., in review) that identifies the source of the published equation used to produce the pseudo-observations. Insets show the same data on the ln – ln scale. Empty panes indicate that no data were available.

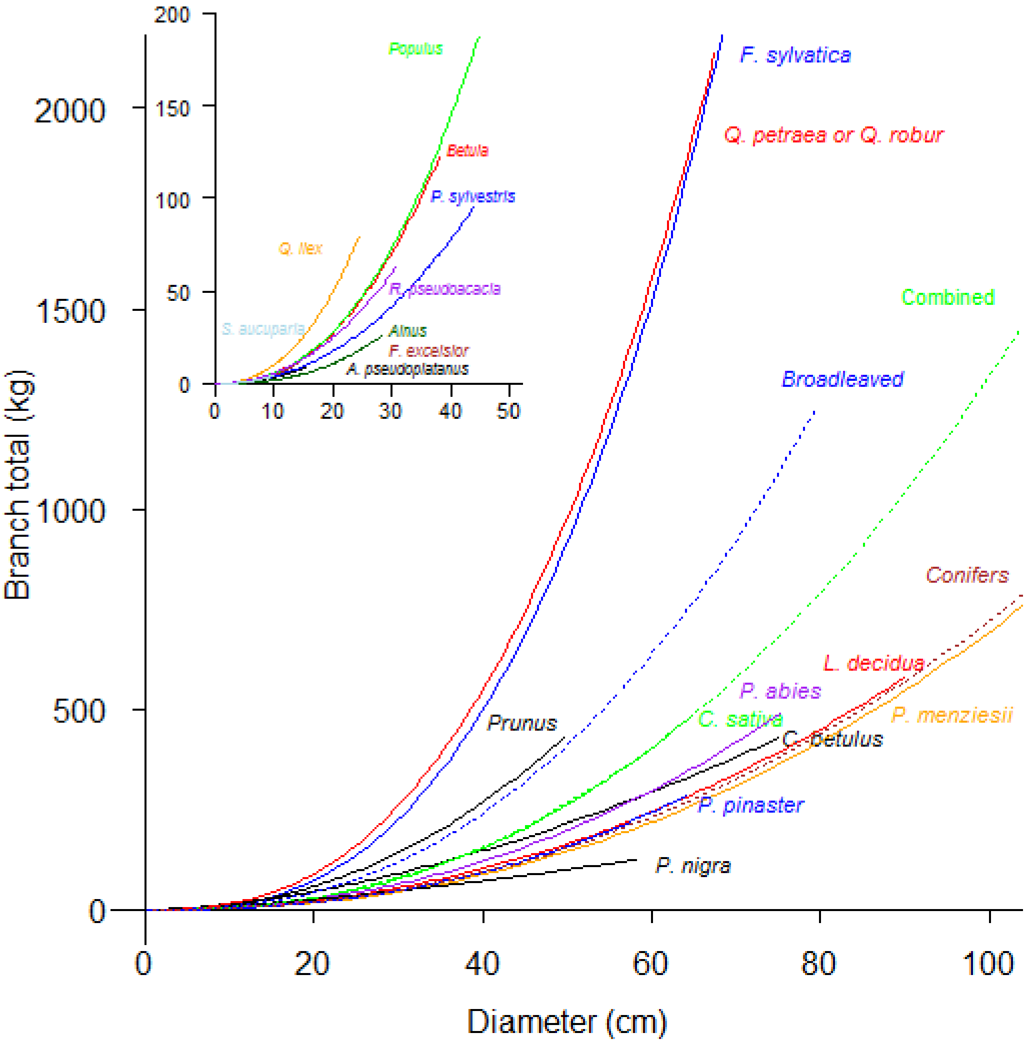


Figure B.26. The relationships between branch mass (live & dead) and diameter for various European species. The lines are Equation 3 fitted to all available data for the given species (Table A.2) and are provided in Table A.5 in Forrester et al. (in review). The dotted lines show the same relationships after dividing the data into broad-leaved species (dotted blue line), conifers (dotted orange line) or all species (dotted green line). The inset is the same relationship for species with small ranges of pseudo-observations.

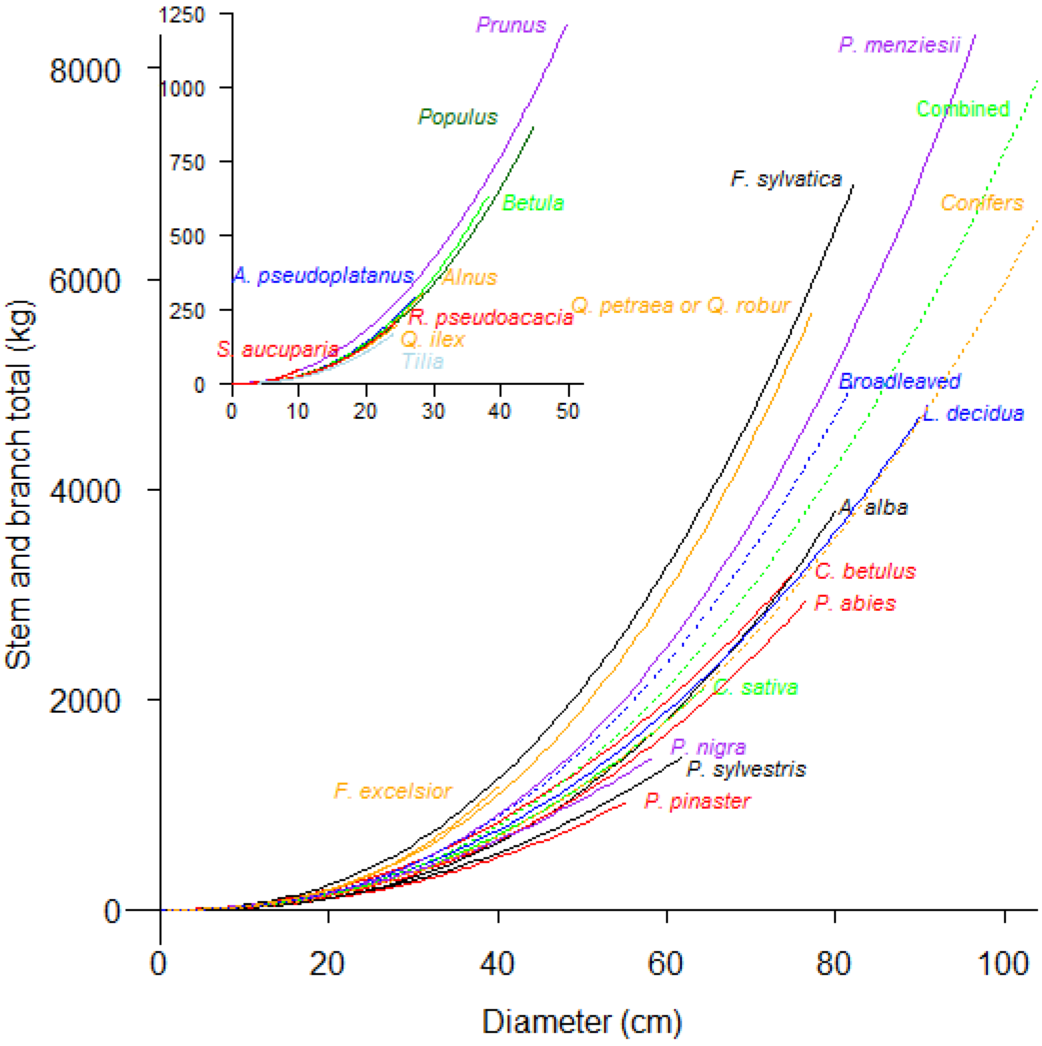


Figure B.27. The relationships between stem + branch mass and diameter for various European species. The lines are Equation 3 fitted to all available data for the given species (Table A.2) and are provided in Table A.5 in Forrester et al. (in review). The dotted lines show the same relationships after dividing the data into broad-leaved species (dotted blue line), conifers (dotted orange line) or all species (dotted green line). The inset is the same relationship for species with small ranges of pseudo-observations.

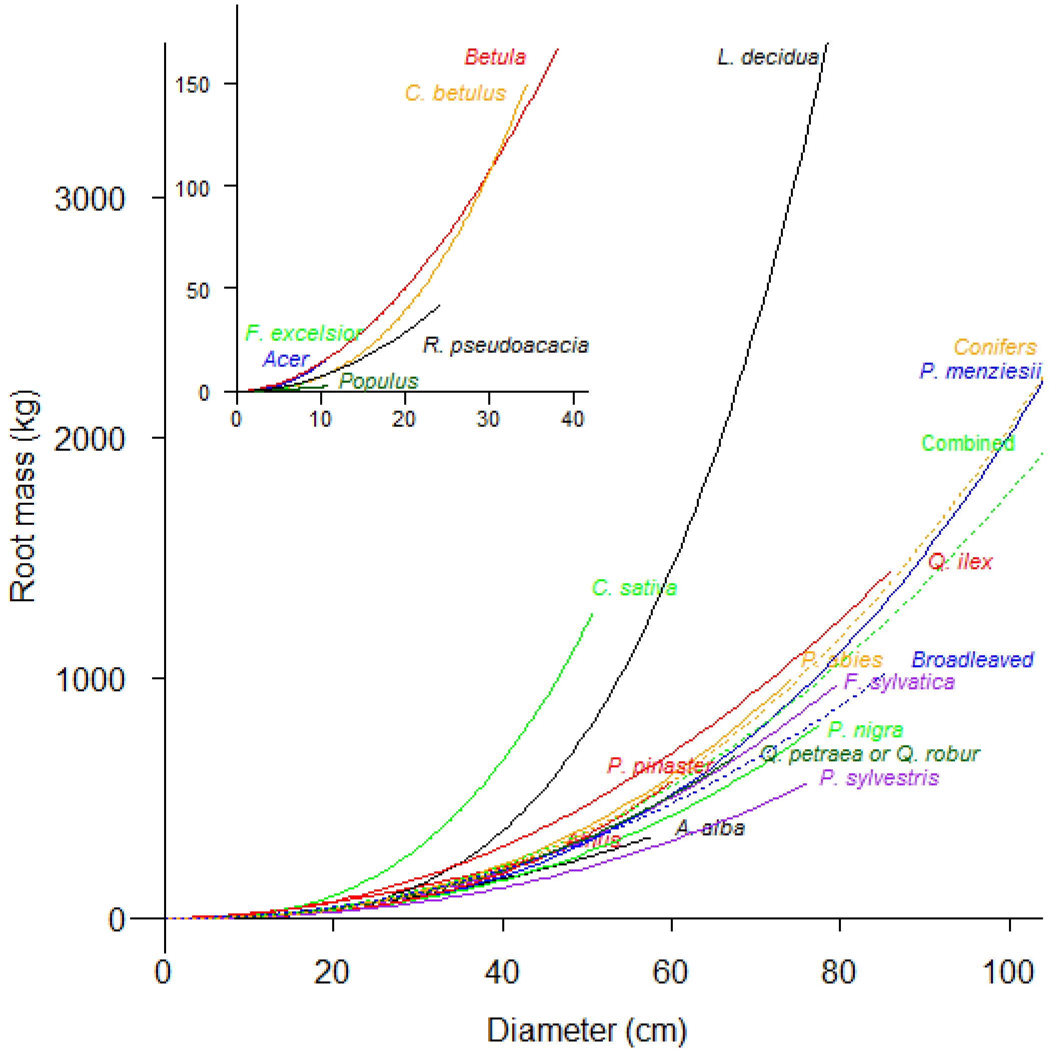


Figure B.28. The relationships between root mass and diameter for various European species. The lines are Equation 3 fitted to all available data for the given species (Table A.2) and are provided in Table A.5 of Forrester et al. (in review). The dotted lines show the same relationships after dividing the data into broad-leaved species (dotted blue line), conifers (dotted orange line) or all species (dotted green line). The inset is the same relationship for species with small ranges of pseudo-observations.

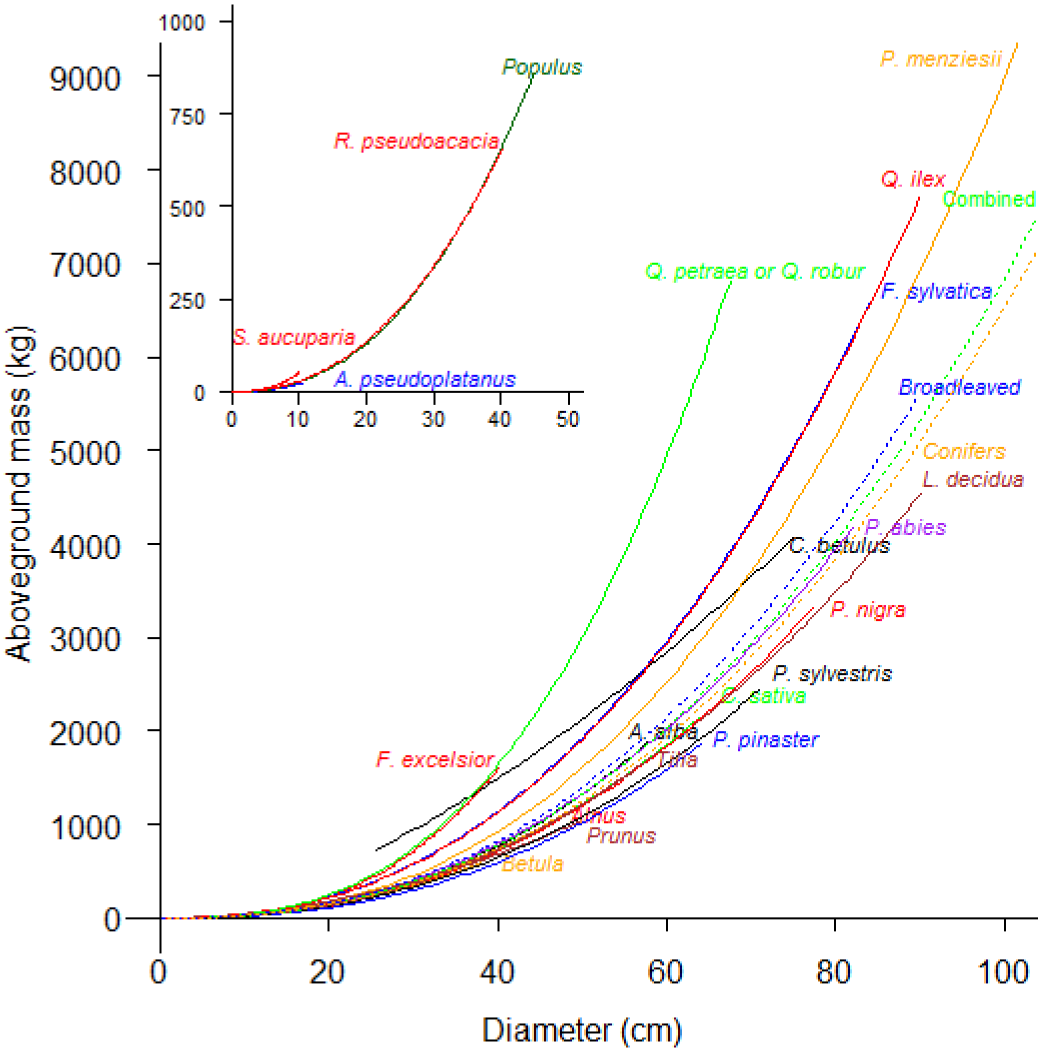


Figure B.29. The relationships between aboveground mass and diameter for various European species. The lines are Equation 3 fitted to all available data for the given species (Table A.2) and are provided in Table A.5 of Forrester et al. (in review). The dotted lines show the same relationships after dividing the data into broad-leaved species (dotted blue line), conifers (dotted orange line) or all species (dotted green line). The inset is the same relationship for species with small ranges of pseudo-observations.

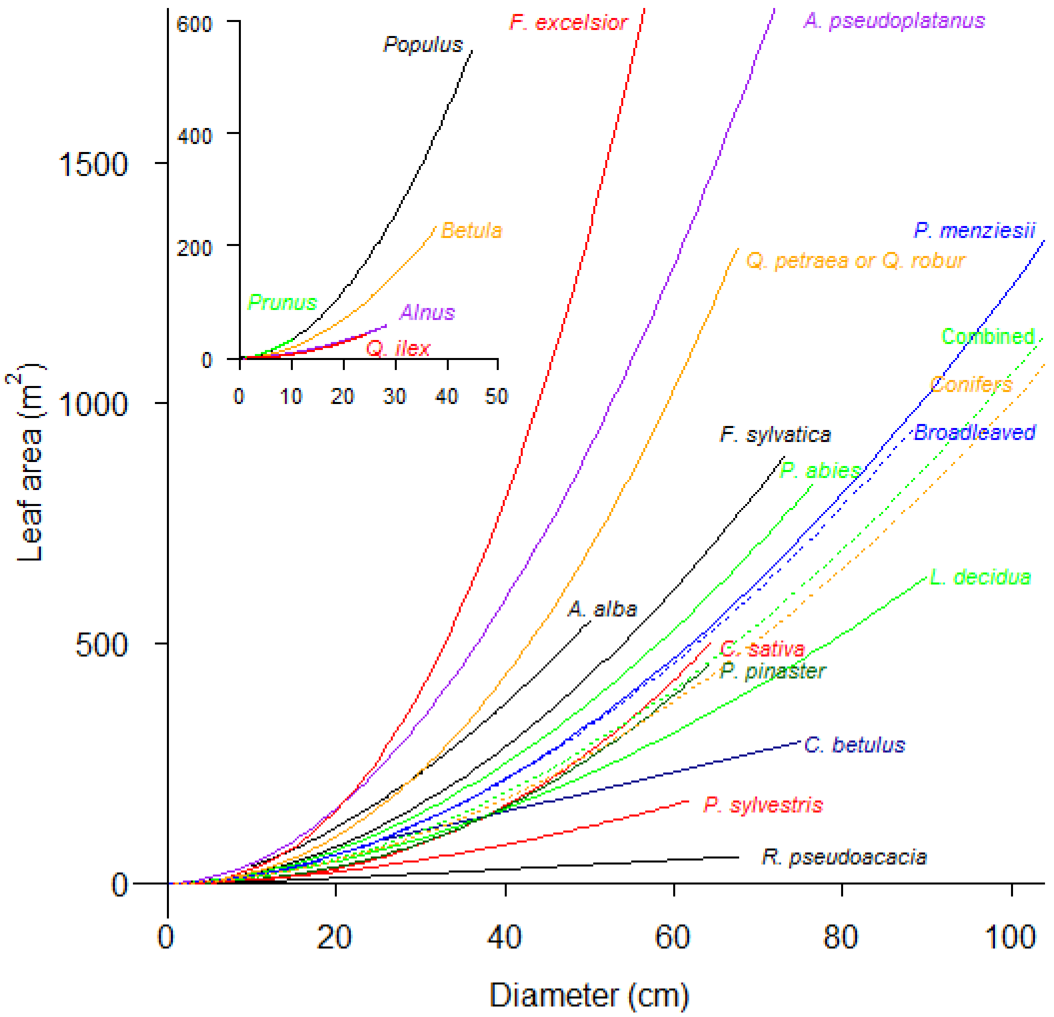


Figure B.30. The relationships between leaf area and diameter for various European species. The lines are Equation 3 fitted to all available data for the given species (Table A.2) and are provided in Table A.5 of Forrester et al. (in review). The dotted lines show the same relationships after dividing the data into broad-leaved species (dotted blue line), conifers (dotted orange line) or all species (dotted green line). The inset is the same relationship for species with small ranges of pseudo-observations.

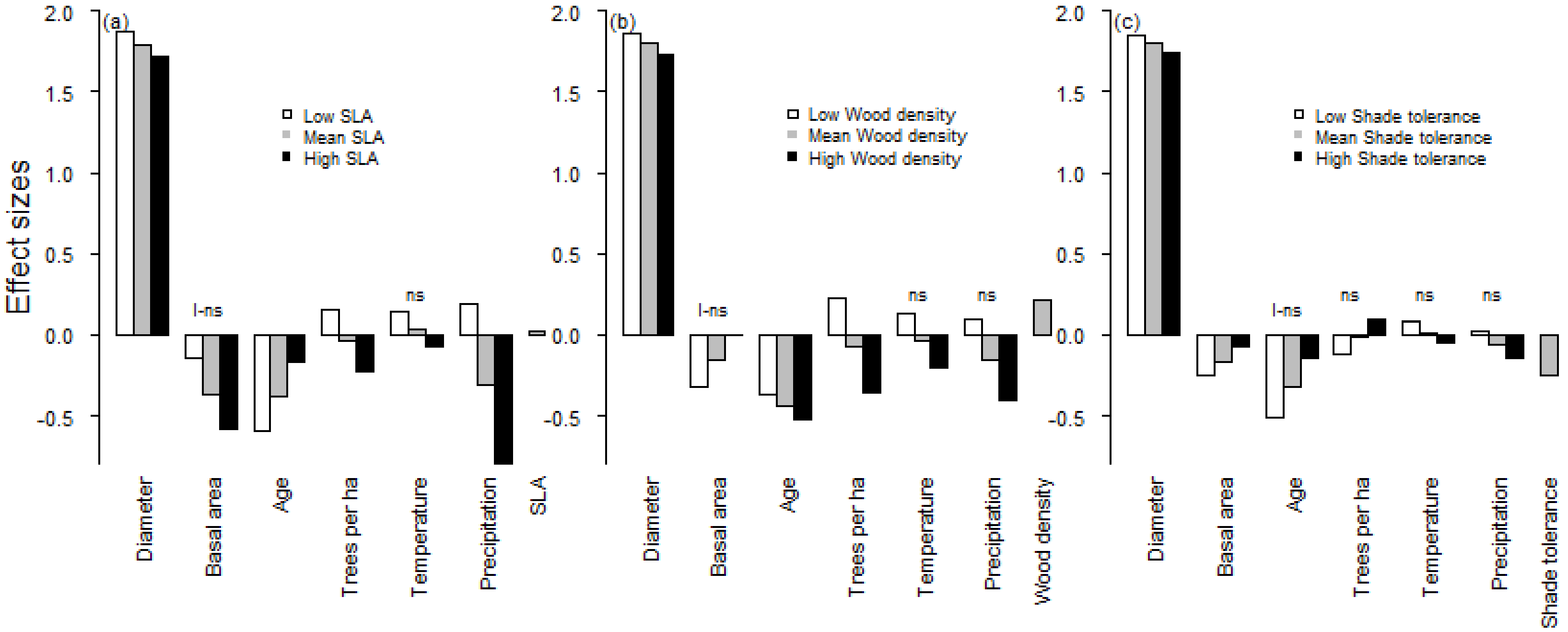


Figure B.31. The effect sizes (beta weights) of all independent variables used to predict leaf area using Equations 22-24. The effect sizes are quantified as the slopes of the variables on the x-axis, all of which have been standardised (mean = 0 and standard deviation = 1). “I-ns” indicates that there was no significant interaction between the given x-axis variable and the trait (specific leaf area (SLA), wood density, shade tolerance). “ns” indicates that the x-axis variable was not significant in the model. The statistical information is in Table A.6 of Forrester et al. (in review). For the trait levels, Low = Mean – 1 standard deviation, Mean = Mean of all species, High = Mean + 1 standard deviation.

References

Alden, H. 1995. Hardwoods of North America. United States Department of Agriculture, Forest Service, Forest Products Laboratory. Gen. Tech. Report FPL-GTR-83, http://www2.fpl.fs.fed.us/TechSheets/hardwood.html.

Alden, H. 1997. Softwoods of North America. United States Department of Agriculture, Forest Service, Forest Products Laboratory. Gen. Tech. Report FPL-GTR-102.

Annighöfer, P., Mölder, I., Zerbe, S., Kawaletz, H., Terwei, A., Ammer, C., 2012. Biomass functions for the two alien tree species *Prunus serotina* Ehrh. and *Robinia pseudoacacia* L. in floodplain forests of Northern Italy. European Journal of Forest Research 131, 1619-1635.

Brzeziecki, B., Kienast, F., 1994. Classifying the life-history strategies of trees on the basis of the Grimian model. Forest Ecology and Management 69, 167-187.

Cheng, J. C., Yang, J., Liu, P., 1992. Anatomy and Properties of Chinese Woods. Chinese Forestry Publishing, Beijing, China.

Forestry Compendium. CAB International. http://www.cabi.org/compendia/fc/.

Forrester, D. I., Tachauer, I. H. H., Annighoefer, P., Barbeito, I., Pretzsch, H., Ruiz-Peinado, R., Stark, H., Vacchiano, V., Zlatanov, T., Chakraborty, T., Saha, S., Sileshi, G. W. in review. Biomass and leaf area allometric equations for European tree species incorporating stand structure and tree age. Forest Ecology and Management.

Gutierrez Oliva, A., Plaza Pulgar, F. 1967. Caracteristicas fisico-mecanicas de las maderas espanolas. Ministerio de Agricultura, Direccion general de montes, caza y pesca fluvial, Instituto Forestal de Investigaciones y Experiencias. Madrid.

Lavers, G. M., Moore, G. L., Building Research Establishment, 1983. The strength properties of timbers. Garston, Watford. London.

Niinemets, Ü., Valladares, F., 2006. Tolerance to shade, drought and waterlogging of temperate northern hemisphere trees and shrubs. Ecological Monographs 76, 521-547.

Rijsdijk, J. F., Laming, P. B., 1994. Physical and Related Properties of 145 Timbers. Kluwer Academic Publishers, Dordrecht, Netherlands.

Schütt, P., Schuck, H. J., Aas, G., Lang, U. A., 1994. Enzyklopädie der Holzgewächse. Handbuch und Atlas der Dendrologie. Ecomed, Landsberg am Lech, Germany.

Zanne, A. E., Lopez-Gonzalez, G., Coomes, D. A., Ilic, J., Jansen, S., Lewis, S. L., Miller, R. B., Swenson, N. G., Wiemann, M. C., Chave, J., 2009. Global wood density database. Dryad. Identifier: http://hdl.handle.net/10255/dryad.235.