Response to Referee comments:
Interactive comment on “Inferring the effects of sink strength on plant carbon balance processes from experimental measurements” by Mahmud et al.

Response to Referee #1:

Overall Review
The manuscript uses a data-assimilation technique to combine an essential but well thought model of carbon balance and plant growth with observations of photosynthesis, maintenance respiration, changes in biomass and NSC obtained in seedlings of Eucalyptus tereticornis planted in containers of different volumes and freely in the soil. The original experiment of Campany et al 2017 reproduced indeed different degrees of sink limitations. The data-assimilation technique allows the authors to infer time dynamics of model parameters (e.g., allocation fractions) and to quantify the relative importance of different processes in downregulating plant-growth under sink limitations. The article shows that the reduction of photosynthesis rate due to sink limitation is not sufficient to explain the reduction of plant growth since other adjustments in NSC utilization, allocation dynamics, and modified respiration and leaf turnover rates are playing an important role. The inclusion of a NSC storage pool and the capability to account for sink limitations emerge as key model components if results of the experiment are to be reproduced. While the path toward modeling plant growth under sink-limitations in mature ecosystems and under various environmental conditions remain long, this contribution is surely an important advancement in the right direction. The article is very well written and presented and most important it is very novel with comparison to the existing literature. As far as I know, it is the first time evidence of carbon sink-limitations is presented so markedly and modeled in a realistic context. In summary, I am very positive concerning the content and conclusions of the article. I think the manuscript is making a very important contribution to the field and I sincerely congratulate the authors for this nice piece of work. In the following, I just have some minor comment that can be helpful to improve further the presentation of this work, especially the comments: P.16 Line 352-358 and P.19. Line 416.

We appreciate the reviewer’s comments and careful reading of our manuscript.

Minor comments

We had intended to cite Bonan’s book “Ecological Climatology” but in line with the reviewer’s suggestion we have now chosen something more process-based about the modelling of forest growth: a review by Mäkelä et al. (2000) in Line 42.

P.2 Line 53. I would suggest to add also Paul and Foyer 2001, very relevant here.

The reference is added in Line 53.

P.3. Line 67. It must be “Fatichi et al 2014”.

This was a typo, is now corrected (Line 70).

P.5. Line 146. Maybe this is not a case that has been encountered in this article. However, how does the model work when NPP is negative and therefore maintenance respiration is larger than carbon assimilation? Is maintenance respiration generally subtracted by the non-structural storage or is it done for each of the tissue separately?

We modify the C balance model (Figure 1) to clarify this situation of having negative NPP. The arrow representing the total maintenance respiration, $R_{m,\text{tot}}$, is relocated to connect with the non-structural C pool ($C_n$). So, the total daily inputs of GPP directly enter into $C_n$ pool and daily $R_{m,\text{tot}}$ is then subtracted from this pool before utilizing the reminder for growth. The text is adjusted accordingly in Line 207-210.

P.5 Line 171. The relative amount of NSC in the roots appear to be a very small number, 9% of the total, while generally one would expect a significant amount of non-structural C-storage in roots especially for seedlings and grasses. Do you have any explanation for this?

There are not a lot of data on NSC of Eucalyptus seedling roots. We used data from an experiment with a related species, *Eucalyptus globulus*, grown in small pots until four months of age (Duan et al. 2013). There’s no obvious reason why these plants would have exceptionally small root NSC contents. We have access to observations from *Eucalyptus parramattensis* saplings in another experiment, which show even lower fine root NSC fraction, and those plants were considerably bigger and older (~17 months). Therefore, it is possible that these very fast-growing Eucalypts do not start to accumulate root reserves until they are well-established. We have introduced this discussion in the manuscript to explain the low root NSC in Line 612-618.

P.8. Line 202-206. How are you dealing with the heterogeneity in photosynthetic properties among leaves and among plants? Were they significant? I know that you wrote in Line 211, that you use the mean for each treatment; is this the mean of how many replicates? Did you average the photosynthetic and stomatal model parameters ($V_{\text{cmax}}, g_1, \ldots$) or the A-Ci and stomatal conductance values?

Measurements of photosynthesis were made fortnightly throughout the experiment on one fully expanded leaf per plant (Campany et al. 2017). A-Ci curves were also measured twice during the experiment. Treatment effects on photosynthesis were detected immediately on newly produced (fully expanded) leaves and we did not observe variation over time in photosynthetic rates. Hence, the photosynthesis parameters were assumed not to change over time, but were specific for each treatment. This was also queried by reviewer 2. We have added text to clarify this point in Line 181-183, 192-195.
P.9 Line 222-223. If I am reading correctly there are 18 (3x6) coefficient to determine for each treatment, maybe this can be written explicitly to compare with number of measurements (44 points) in line 231-233. This allows some redundancy even in the case of separating each container size.

The total counts of parameters and measurements are added to the manuscript to show the actual number of redundancy (Line 247-249). There are in fact 5 parameters to determine for each treatment group (as root allocation, \(a_r = 1 - a_f - a_w\)), so a total of 15 (3x5) coefficients to determine for each treatment group, compared with total 44 data measurements available for each treatment.

P. 9. Line 243-247. Please explain better this part of the data assimilation methodology, as it is not very clear to me.

We have elaborated on this section to provide more in depth idea of the DA algorithm (Line 259-272).

P. 14. Line 319. There are not “bold values” in Table 3. Probably a formatting issue.

Yes, it was a formatting issue, and is now corrected in Line 332.

P.16 Line 352-358. The lowest utilization rate in seedling in small containers would theoretically lead to an accumulation of NSC, at least in relative terms, which is something we do not see in Figure 2 and Figure 4. The explanation for such non-intuitive results is only provided in the discussion (Line 505-510) and justified as a temporal effect, where NSC first accumulates in seedling in small containers but then they are depleted by the higher respiration costs and leaf turnover rates. I think it would be quite interesting to see in Figure 2, NSC (\(C_{n,f}\)) reported as fraction of total C mass in foliage (\(C_{t,f}\)), e.g., \(C_{n,f}/C_{t,f}\). This would serve the double purpose of explaining such a different dynamic in the use of NSC as the season progress and will provide the percentage of NSC in leaves that can be compared with other studies. This will likely highlight a higher concentration in seedling in small containers at the beginning of the season but a lower concentration at the end (as in Figure 4).

We have replaced the NSC total plot with an NSC concentration plot in Figure 2, and moved the NSC total plot to supplementary material (Supplementary Figure S1). The text is also adjusted accordingly (Line 358-360).

P. 18. Line 401-403. The authors for some reason never refer to the concept of Carbon Use Efficiency (CUE), but I suggest it would be useful here to explain the results. Substantially what they are saying is that CUE = \((1 - (R_{m,tot} + R_g)/GPP)\) is higher in free seedling and it is reduced by sink limitations. Maybe, a figure showing the temporal evolution of CUE for the various treatments would be also an interesting piece of information.

Carbon use efficiency (CUE), calculated as seedling biomass per unit total photosynthesis, was shown in Campany et al. (2017) and we had tried to avoid duplicating results that were already published. However, we have added numbers to this part of the results (Line 429-432). The temporal evolution of CUE for various treatment groups is also interesting and is added to the supplementary material (Supplementary Figure S2).
I am not sure if parameters were changed one at a time resetting the previous parameter to the original value or if effects are added up (which seems more the case from the presentation of results). If parameters are really changed “one at a time”, this does not allow to reproduce all the interactions among parameters and can inflate the role of certain parameters. Therefore the total effect of Table 4 (54.8 gC plant-1) does not necessarily correspond with the real total effect, which is not reported for comparison. If the parameters are switched on sequentially then you will obtain the total effect but the importance of certain parameters will not be clearly separated, since it will depend on the adopted sequence of switchers. For instance, the role of “k” would be likely smaller when the interplay with the other parameters is considered. Now, I am not asking to running simulations with interactions among parameters since they would be an extremely high number (going factorial) and they will not add much to the overall discussion on the model results. However, this simplification and the specific method needs to be stated explicitly and the difference in the total effect between the simulations changing parameters “one at a time” and the total observed effect needs to be mentioned, since it can provide an idea of the importance of the interaction among parameters.

The reviewer makes a good point. We suggest that we keep the attribution analysis shown in Figure 5 where the effects of various parameters are sequentially added up to get the total effect over the duration of the experiment, but that we modify Table 4 to present the effect of the parameters changing individually from the value of the smallest container treatment (5 L) to that of free seedling (FS) and other way around, resetting the previous parameter to the baseline value. The new table also lists the total C for different sequential attribution scenarios, that showed the total observed effect due to sink limitation. The final biomass values indicate both the contribution (+/-) of each individual parameter separately (Table 4, 2nd and 3rd columns) and sequentially (Table 4, 4th column). This fulfils both purposes, eventually showing the contribution of each individual parameter separately. All these are adjusted with explicit text to clarify section 2.4.3 and 3.6 (Table 4, Line 312-319, 445-454).

Following my previous comment, I wonder if it is not better to show the effect of each parameter independently rather than the sum of the effects of the parameters on the final biomass. I think it would be better to show the effect of each parameter by itself on the baseline rather than what is shown now. In any case, a clearer explanation of what is shown would be necessary.

According to the previous response, we have modified Table 4 to present the effect of the parameters changing one at a time resetting the previous parameter to the baseline value, which illustrates the contribution of each individual parameter separately.

The reference is added in Line 502.

P.25. Line 550-552. Another way to say the same concept is that CUE is higher in free seedlings.
    Yes, indeed. We have mentioned CUE while discussing the C utilization rate (Line 537-538).

    The reference is added in Line 741-742.

P. 27. Line 605-610. I agree with the authors, but there is still an important challenge of dealing with sink limitations in ecosystems encompassing tall-trees and heterogeneous vegetation types and for which observations for data-assimilation may not be available.
    We agree with the reviewer that there are challenges to apply this approach at ecosystem scale due to data availability. However, the main focus of our paper is to build up a foundational step towards understand plant functioning rather than solving ecosystem problems. Moreover, one of the important features of DA is that it does not need all the data streams from every individual C stocks and fluxes. There are several successful examples of DA being applied to forest growth, albeit without a focus on storage (e.g. Bloom et al. (2016), Van Oijen (2008), Williams et al. (2005)). Since this was also a question raised by the second reviewer, we have enhanced the discussion on the potential for applying DA to investigate storage at ecosystem scale in Line 631-641.

P. 29. Line 675-677. I thank the authors for referencing to my work, but this article is completely irrelevant for the current paper and indeed is just quoted by mistake.
    This was a typo, and is now removed.
Response to Referee #2:

Overview
The paper by Mahmud et al. presents a data assimilation exercise where data from a manipulative experiment on small trees had been ‘assimilated’ by a carbon balance model (CBM). The manipulation aimed to reduce root sink strength by constraining growth space of the root systems with varying pot sizes (5 – 30 L, in 5 L steps, 35L, and a ‘free’ treatment where trees were grown without limitation). On these trees biomass pools (structural biomass of leaves, wood and roots, non-structural carbohydrates, NSC, in leaves) were measured with different frequencies and used to constrain the CBM which simulated GPP based on parameters derived from punctual measurements of assimilation and respiration. Model runs with different structures (with/without a NSC ‘storage’ pool) were performed to test how important such a carbon buffer is for CBM simulations. The set of parameters of a best suited model (parametrized with three sink strength classes) was discussed with respect to plant carbon (allocation) dynamics in response to sink limitation. In addition, an attribution analysis was performed which aimed to provide information of the underlying mechanisms responsible for changes in biomass from sink limitation. The authors highlight the need for including a ‘storage’ component in vegetation models and the usefulness of their approach for further investigations to ‘develop appropriate representations of sink-limited growth in terrestrial biosphere models’.

We appreciate the reviewer’s comments and careful reading of our manuscript.

General comments
This is a very nice project as it combines experimental manipulations with a data assimilation procedure. During the last years I have been running several experiments to manipulate the plant carbon balance. I have been thinking repeatedly about such a data assimilation approach to learn more about plant carbon dynamics and the underlying mechanisms. This study here does exactly this and I applaud the authors for making this progress.

We appreciate the reviewer’s positive comments, and are hopeful that he will find our approach useful in future. To aid uptake of our approach, we have made all code freely available as a Git repository:
https://github.com/kashifmahmud/DA_Sink_limited_experiment

That being said, I think that the interpretation of the data and the general presentation of the manuscript can be improved to increase its impact. For example, one of the main findings of the study, i.e. the importance of a storage component for (more) realistic simulations of plant functioning, is a strawman. Plants do have a storage component and of course models that specifically include carbon storage will be more realistic, in particular in situation where NSC may accumulate due to sink limitations.

Here we must respectfully disagree with the reviewer. Although in reality plants do have a storage component, it is not necessarily the case that including such a storage component in the model will lead to model improvements. Making a model more and more complex to “better represent reality” is unwarranted in many situations – the idea of a model is to abstract reality, not reproduce it. Hence, it is actually quite
important to test whether or not adding the storage component improves the performance of the model enough to justify the additional complexity. To address this comment, we have inserted few sentences to the introduction in Line 110-113 to explain why we feel it necessary to test the “strawman” null hypothesis.

In my opinion, the merits of the study are not the particular findings but rather the documentation of the potential of the data assimilation approach. The findings have to be taken with caution as the constraints from measurements are simply not sufficient to allow deeper insights into plant functioning. For example, measurements of assimilation and respiration have been done twice only and the photosynthetic parameters derived from these two measurements were used to estimate GPP over the whole season. How robust are these parameters for that purpose? Similarly, leaves were sampled every second week for NSC measurements and the structural biomass of stem and roots was determined only for the fourth months or at the end of the experiment, respectively. Given these limitations, how relevant are your inferences, for example, that sink limitation has led to reductions in photosynthetic rates or enhanced respiratory losses?

We are grateful that the reviewer recognises the potential of the DA approach to provide insight into C balance processes. However, we must also respectfully disagree with his comment that the measurements “are simply not sufficient to allow deeper insights into plant functioning”. Although some measurements we’d like to have had are missing – notably root and stem NSC, growth respiration, turnover, and root exudation – we believe that this particular experiment is sufficiently well-constrained to draw the inferences that we have.

The reviewer mentions photosynthesis. Although we only used two sets of A-Ci data to estimate the parameters, net leaf photosynthesis was measured fortnightly throughout the experiment (Campany et al. 2017) and we can be confident from the lack of trend over time that these measurements are robust and representative. We have added information on this in Line 181-183, 192-195.

The height, diameter and leaf area were measured fortnightly throughout the experiment, and root biomass is also constrained at the end. As we note in the discussion, there is some question over whether our “Y” value represents growth respiration or other carbon losses from the plant, such as root exudation, but we have reasonable confidence, from the combination of measurements available, in our inference that the C loss term is increased with sink limitation.

No experiment can perfectly quantify the C balance – there will always be data gaps. We are enthusiastic to see our approach applied to other experiments, such as those described by the reviewer, but there are likely also to be gaps in these datasets to constrain the key C balance processes. Fortunately, the DA approach does not need to have all of the data streams from each of the C stocks or fluxes, and can estimate some missing C stocks from other measurements, although of course the precision of model predictions increases with data availability. We already covered the impact of
missing data streams at some length in the discussion (Line 591-606); we have now refined this discussion to more directly addressing the reviewer’s comment in Line 631-638.

Additionally, NSC were measured in leaves only and their distribution among plant organs estimated with fixed parameters. For a study that specifically aims to highlight the role of NSC storage in plant functional processes, this is a critical shortcoming. Within the experimental period, there may have been substantial shifts in the NSC distribution across organs and this could have a substantial impact on the simulated carbon dynamics.

The sink-limited container experiment only measured foliage NSC, and therefore to estimate the partitioning of the non-structural C among different organs, we used data from a different experiment on similar-sized seedlings of a related species (Eucalyptus globulus) (Duan et al. 2013). In that experiment, foliage, wood and root NSC were measured repeatedly over two months and there was no statistically significant difference in the NSC distribution. We thus believe this is a justifiable assumption. We have added this information in Line 176-179, and acknowledged this as an uncertainty.

That being said, I think you should rather discuss the approach, its potential but also its limitations. One aspect, for example, is how well a study on seedlings can ‘develop appropriate representations of sink-limited growth in terrestrial biosphere models’. Such models usually simulate mature trees, not seedlings. We have recently published a paper addressing this particular topic: how to make use of seedling studies for inferences on mature trees and modeling of vegetation dynamics (see reference from EEB below). Instead of too many inferences I would like to see a critical evaluation of your method, including an assessment of what data are needed to get better constraints for the model. I have done experiments with small trees in growth chambers where GPP and several components of the carbon balance have been assessed continuously or at high temporal resolution. Applying sink limitation (I used source limitation but also drought, which is also a form of sink limitation) in such an experimental setting would allow making much more robust inferences that with the data set used here. Hence, my suggestion is to move away from the current focus of interpreting plant functional responses and instead concentrate on presenting the approach as a promising avenue for how to gain insights into plant functioning.

We have added a discussion of the implications of this study for larger plants in Line 631-641, drawing on the review paper mentioned by the reviewer. We already have a section where we discuss potential and limitations of this approach and also identify the data needed to better constrain the model framework in article 4.2 (Line 577-606). However, we are reluctant to remove our focus on drawing inferences from the experimental study here. As described above, we believe that our inferences are valid - to the extent that we already describe in the Discussion - and help to demonstrate the utility of the approach. Hence, we respectfully suggest, that we should continue to do both: interpret the responses and present the approach.

I hope my comments can help increasing the paper’s impact.
Specific comments:

Abstract: Please add what species you have been working with.
Added in Line 17.

L 20: processes affected by growth? That doesn’t make sense.
The clause read “processes affected by growth under sink limitation”. We have now shortened this to “processes affected by sink limitation” (Line 18-19).

L 21: What do you mean by ‘component processes’?
We meant various C cycle processes contributing to growth e.g. photosynthesis, respiratory losses, utilization of NSC, allocation pattern, turnover rates as mentioned in the following sentence. There is not space in the abstract to further expand on this; we have elaborated the processes at line 107 in the Introduction.

L27-29: Not much content in this sentence.
The sentence is now restructured (Line 25-28).

Introduction in general: The structure of several paragraphs is not ideal and reduces the logical flow. For example, on L 55 you start a paragraph by asking how to include source and sink limitations in models but then you move to NSC storage in models. I understand that storage allows buffering discrepancies in source and sink activity but this is not strictly related to the limitations. A more logical flow would be to say that there is ongoing discussion about realistic implementations of NSC in vegetation models and that, because of their multiple roles in plant functioning, such an implementation also provides a buffer against discrepancies in source and sink activity.
We have reorganized the paragraph with the addition of referee’s observation in Line 54-57 on ongoing discussion about the realistic implementations of NSC in vegetation models.

L74-76: Quantify growth by manipulating rooting volume? That also does not make sense.
We have omitted “to quantify growth” from this sentence (Line 77-78).

L 88: Very good point!
We are not quite sure which point the reviewer is referring to here. However, as it is positive, no revision is needed.

We have added a reference to this paper in Line 90-93 as it’s very relevant.

L 105-123: The presentation of the hypothesis is very awkward. Could you present this please in a more accessible and appealing way? This is not a funding proposal but a text
intended for keeping readers keen on reading on. Please rephrase and restructure to make this a flowing text.

The presentation of hypotheses is evidently a matter of taste. We felt that, given we had several different hypotheses to test, it made sense to lay them all out clearly and concisely in this way. We ask for the editor’s advice here: we can turn this part into regular paragraphs if he feels it will improve the communication of our ideas.

L137-141: Relocate further up to L 129 (after Australia).
   Relocated as suggested (Line 137-142).

L 142: I suggest presenting the data first, then the model.
   We have swapped the subsection 2.2 and 2.3 to follow reviewer’s comment.

Table 1: Why is there no hypothesis for simulation set C?
   Simulation set C aimed to quantify the effect of changes individual parameters on overall seedling growth, and that is why this section does not come with any hypothesis.

Results: The results are presented in a very uncommon form. The text repeats the hypotheses (not useful) and reads more like a discussion than a presentation of results. I suggest adapting a more formal style so the reader knows to differentiate between results and interpretation of results.

Again, we refer to the editor for advice. It is true that for many papers, it is appropriate to present the results without reference to the hypotheses, and only return to the hypotheses in the discussion. However, here we felt it was necessary to address the hypotheses directly in the results to better guide the reader through the outcomes of the different model analyses. For example, consider section 3.1, where we present the BIC values for different simulation sets. We could here reduce the paragraph to just lines 299-303:

“Table 3 (Simulation Set A) shows the results for model fits with the optimal grouping strategy (three treatment groups). BIC values were consistently lower for the model including the storage pool; the improvement in model fit is most noticeable for the containerized seedlings.”

and move the sentences describing how the hypothesis is tested to the discussion. Our personal feeling is that this makes this result paragraph fairly impenetrable. However, we are happy to take the editor’s decision on this.

L 329: You mean Fig. 2. Please correct figure numbering for the following figures also.
   We have checked the Figure numbers on the pdf version from BGD library and believe that all figure numbers are correct.

Figure 1 (actually Fig. 2): Add title to each panel (leaf, wood, root, NSC).
We have added the titles to each panel of Figure 2.

L 352: Shouldn’t this section be presented before the modeling outcome? Parameters first, then the modeled pools?

When writing the manuscript, we experimented with both orders of presentation. We found this order of presentation to be more logical. Showing the data first allows the reader to see how the DA works in terms of data fit, before moving on to examine the model parameters.

L 413: Is this a sensitivity analysis?

No, it’s something a bit different from a sensitivity analysis, which is why we call it an attribution analysis. A sensitivity analysis aims to quantify the sensitivity of model outputs to a change in one or more input parameters. What we are doing is trying to work out (attribute) how much of the observed change in growth is due to the observed change in each parameter.

L 418: This belongs into the methods section.

We have already mentioned the procedure in methods (Line 312-315); this sentence was intended to remind the readers that the analysis attributes the change in biomass from the smallest container to free seedling. This sentence is now removed.

L 422: And this should go into the figure caption.

Moved to figure caption (Line 477-479).

Table 4: Most of this information has been reported in Fig. 4 already.

According to the response to reviewer 1, we have modified Table 4 to present the effect of the parameters changing individually from the value of the smallest container treatment (5 L) to that of free seedling (FS) and other way around, resetting the previous parameter to the baseline value. The new table also lists the total C for different sequential attribution scenarios, that showed the total observed effect due to sink limitation. So, the table in its recent version does not repeat the numbers showed in Figure 4.

L 460-465: The emphasis here is on inferences on processes that are poorly constrained. See my general comments.

The beauty of the DA approach is that it is possible to make inferences about processes that are individually poorly constrained, but contribute to overall growth, and hence can be constrained against total growth. Here, as mentioned above, we have 15 coefficients to determine against 44 mean data values, each of which is determined by 7 replicates. This total amount of data enables us to fully constrain the model. We have added this explanation (Line 247-249, 603-604).

L 466-467 (and at beginning of other paragraphs): Please avoid restating the hypotheses.
Again, this is a matter of taste, and we will defer to the editor. The reviewer’s point is that we have already stated the hypotheses in the introduction, so there is no need to re-state them here. Our preference is to re-state them, otherwise the reader will need to keep flipping back to the introduction to check which one “H2” was again, and we think that could be annoying.

Discussion in general:
I’d relocate the focus to discuss the potential of the approach and move away from interpreting the model outcome with respect to plant functioning. The discussion is somewhat lengthy and verbose. Please be more concise and to the point.

As explained above, we prefer to keep both emphases in the discussion. We have gone through the discussion and attempted to reduce the length and verbosity.

A few suggestions from my own work which are based on whole-plant assessments of the C balance:
3. Hartmann H, Trumbore S. 2016. Understanding the roles of nonstructural carbohydrates in forest trees – from what we can measure to what we want to know. New Phytologist 211, 386-403.

We have read through the suggested papers and cited three of them with the relevancy of the text (Line 63, 68 and 633).

References:
Campany, C.E., Medlyn, B.E. and Duursma, R.A. (2017) Reduced growth due to belowground sink limitation is not fully explained by reduced photosynthesis. Tree Physiol 37(8), 1042-1054.
Inferring the effects of sink strength on plant carbon balance processes from experimental measurements

Kashif Mahmud\textsuperscript{1}, Belinda E. Medlyn\textsuperscript{1}, Remko A. Duursma\textsuperscript{1}, Courtney Campany\textsuperscript{1,2}, Martin G. De Kauwe\textsuperscript{3}

\textsuperscript{1}Hawkesbury Institute for the Environment, Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia
\textsuperscript{2}Department of Biology, Colgate University, NY 13346, USA
\textsuperscript{3}ARC Centre of Excellence for Climate Extremes, University of New South Wales, Sydney, NSW 2052, Australia

Correspondence to: Kashif Mahmud (k.mahmud@westernsydney.edu.au)

Abstract

The lack of correlation between photosynthesis and plant growth under sink-limited conditions is a long-standing puzzle in plant ecophysiology that currently severely compromises our models of vegetation responses to global change. To address this puzzle, we applied data assimilation to an experiment where sink strength of \textit{Eucalyptus tereticornis} seedlings were manipulated by restricting root volume. Our goals were to infer which processes were affected by sink limitation, and to attribute the overall reduction in growth observed in the experiment, to the effects on various carbon (C) component processes. Our analysis was able to infer that, in addition to a reduction in photosynthetic rates, sink limitation reduced the rate of utilization of non-structural carbohydrate (NSC), enhanced respiratory losses, modified C allocation and increased foliage turnover. Each of these effects was found to have a significant impact on final plant biomass accumulation. We also found that inclusion of a NSC storage pool was necessary to capture seedling growth over time, particularly for sink limited seedlings. Our approach of applying data assimilation to infer C balance processes in a manipulative experiment enabled us to extract new information on the timing, magnitude, and direction of the internal C fluxes from an existing dataset. We suggest this approach could, if used more widely, be an invaluable
tool to develop appropriate representations of sink-limited growth in terrestrial biosphere models.

**Keywords:** Non-structural carbohydrate, carbon allocation, data assimilation, mass-balance, photosynthesis, plant growth, sink regulation

### 1 Introduction

Almost all mechanistic models of terrestrial vegetation function are based on the carbon (C) balance: plant growth is represented as the difference between C uptake (through photosynthesis) and C loss (through respiration and turnover of plant parts). This approach to modeling plant growth dates back to early crop and forest production models (McMurtrie and Wolf, 1983; de Wit and van Keulen, 1987; de Wit, 1978) and now provides the fundamental quantitative framework to integrate our scientific understanding of plant ecosystem function (Makela et al., 2000).

However, C balance models have been criticized for being “source-focused” (Fatichi et al., 2014). Most C balance models predict growth from the environmental responses of photosynthesis (“source limitation”). In contrast to this assumption, many experimental studies demonstrate that growth is directly limited by environmental conditions (“sink limitation”) rather than the availability of photosynthate. For example, growth is more sensitive to water limitation than is photosynthesis (Bradford and Hsiao, 1982; Müller et al., 2011; Mitchell et al., 2014); low temperatures are considerably more limiting to cell division than to photosynthesis (Körner et al., 2014); nutrient limitation may slow growth without reducing photosynthesis (Reich, 2012; Crous and Ellsworth, 2004); and, physical sink-limitation may reduce growth with a decline in photosynthetic capacity and an accumulation of leaf starch (Arp, 1991; Campany et al., 2017; Poorter et al., 2012a; Paul and Foyer, 2001).

How can we move to models that include both source- and sink-limitation? There is ongoing discussion about realistic implementations of non-structural carbohydrates (NSC) in vegetation models and because of their multiple roles in plant functioning, such an implementation provides a buffer against discrepancies in source and sink activity. Some C balance models...
include a “storage” pool of NSC (Running and Gower, 1991; Bossel, 1996; Thornley and Cannell, 2000), but most of these models make the assumption that the NSC pool acts merely as a buffer between C sources and sinks, balancing out seasonally or at least over several seasons (Fatichi et al., 2014; Friend et al., 2014; De Kauwe et al., 2014; Schiestl-Aalto et al., 2015). There is mounting evidence that the NSC play s a more active role in tree physiology (Buckley, 2005; Sala et al., 2012; Wiley and Helliker, 2012; Hartmann et al., 2015). For example, NSC accumulation can lead to down-regulation of photosynthesis (Nikinmaa et al., 2014). Therefore, the need to quantify the NSC pool and to better understand the prioritisation of storage vs. growth is of great importance.

An understanding of the dynamics of storage is also essential to correctly represent the C balance in models (Hartmann and Trumbore, 2016). If, for example, a direct growth limitation is implemented into models, how should the surplus of accumulated photosynthates be treated? In their proof-of-concept sink-limited model, Fatichi et al. (2014) allowed reserves to accumulate indefinitely. Alternatively, some models (e.g. CABLE (Law et al., 2006), O-CN (Zaehle and Friend, 2010)) increase respiration rates when excess labile C accumulates. Both approaches can be seen as model-oriented solutions to maintain C balance that are unsatisfactory because they are not based on empirical data. Experiments where sink strength is manipulated may provide the key to improve our understanding of C balance processes under direct growth limitation.

Efforts have been made to understand the physiological and morphological changes in response to belowground C sink limitation by manipulating rooting volume in tree seedlings (Arp, 1991; Campany et al., 2017; Poorter et al., 2012a). These experiments often reveal photosynthetic down-regulation and accumulation of leaf starch, and reductions in growth (Arp, 1991; McConnaughay and Bazzaz, 1991; Gunderson and Wullschleger, 1994; Sage, 1994; Poorter et al., 2012a; Robbins and Pharr, 1988; Maina et al., 2002; Campany et al., 2017). In a recent study with Eucalyptus seedlings, Campany et al. (2017) showed that the reduction in seedling growth when rooting volume was restricted could not be completely explained by the negative effects of sink limitation on photosynthesis, suggesting that other components of the C balance were affected in the process. However, Campany et al. (2017) could not accurately quantify all components of tree C balance, i.e. photosynthesis, carbohydrate storage, biomass partitioning and respiration.
Quantifying all components of C balance is not an easy task, given that not all processes are measured with equal fidelity, and data gaps will always occur. Klein and Hoch (2015) used a C mass balance approach with a tabular process flowchart to decipher C components and provide a full description of tree C allocation dynamics. Here, rather than using a manual process, we used a data assimilation (DA)-modelling framework, which has been proven to be a powerful tool in analyzing complex C balance problems (Williams et al., 2005; Richardson et al., 2013). For example, Richardson et al. (2013) use DA to discriminate among alternative models for the dynamics of non-structural carbon (NSC), finding that a model with two NSC pools, fast and slow, performed best; Rowland et al. (2014) applied DA to experimental observations of ecosystem C stocks and fluxes to infer seasonal shifts in C allocation and plant respiration in an Amazon forest; and Bloom et al. (2016) used DA to constrain a C balance model with satellite-derived measurements of leaf C, to simulate continental-scale patterns in C cycle processes.

Our goal in this paper was to use DA to quantify the impact of sink limitation on C balance processes. We utilized data from an experiment in which sink limitation was induced by restricting the rooting volume of *Eucalyptus tereticornis* seedlings over the course of 4 months (Campany et al., 2017). We assimilated photosynthesis and growth measurements from the experiment into a simple C balance model, to infer the effects of sink limitation on the main C balance processes, namely: respiration, carbohydrate utilization, allocation, and turnover. We first tested two null hypotheses:

H1: There is no need to consider storage in the model: growth can be adequately predicted from current day photosynthate. Although in reality plants do have a storage component, it is not necessarily the case that including such a storage component in the model leads to model improvements. Hence, it is important to test whether or not adding the storage component improves the performance of the model enough to justify the additional complexity.

H2: There is no effect of sink limitation on C balance processes other than via a reduction of photosynthesis.

We were then interested to test the following specific hypotheses about the impact of sink limitation on C balance:
H3: We hypothesized that the rate of utilization of carbohydrate for plant growth would be lower under sink limitation, causing growth rates to slow and non-structural carbohydrate to accumulate.

H4: We hypothesized that under sink limitation a larger proportion of C would be lost to growth respiration and less used for production. We have dubbed this the “wasteful plant” hypothesis; this hypothesis corresponds to the assumption embedded in some models that respiration is up-regulated when labile C accumulates e.g. CABLE, O-CN (Law et al., 2006; Zaehle and Friend, 2010).

H5: We hypothesized that foliage and root C allocation fractions would be reduced, in favour of wood allocation. Sink limitation induced by nutrient and/or water stress often results in a shift in C allocation away from foliage and towards fine roots (Poorter et al., 2012b). However, for this experiment, the physical restriction of root growth limits the potential for root allocation. Hence, we predicted that both foliage and fine root allocation would decrease.

2 Materials and Methods

2.1 Experiment description

The site and experimental setup have been described in detail by Campany et al. (2017), so we only provide a brief description here. The experiment was carried out at the Hawkesbury Forest Experiment site (33°37'S 150°44'E) in Richmond, NSW, Australia. The site is located in the sub-humid temperate region and experiences warm summers and cool winters. The seedlings were planted on 21st January 2013 (mid-summer) and harvested on 21st May 2013 (late autumn). Mean daily temperatures ranged from 22.8 to 46.4 °C (monthly mean of 32.1 °C) in January 2013, which was the warmest month of the year, and cooled down in May 2013 with an average of 21 °C (BoM, 2017).

Twenty-week old *Eucalyptus tereticornis* seedlings in tube stock were chosen from a single local Cumberland plain cohort. Ten seedlings were harvested at the start of the experiment to measure initial leaf area and dry mass of foliage, woody components and roots. Forty-nine seedlings were used in the main experiment, allocated to seven treatments. The plants were grown in containers of differing volume set into the ground (5, 10, 15, 20, 25 or 35 L), or were
planted directly into soil (free seedlings, used as the control). All plants were grown in the open
under field conditions, but were watered regularly to avoid moisture stress.

2.2 Experimental data acquisition

Full details of all measurements are given in Campany et al. (2017). The mass of each pool
(foliage, wood, root, storage) was estimated over time as follows. The initial dry mass of leaves,
woods and roots was measured for 10 seedlings at the start of the experiment using the
harvesting procedure described in Campany et al. (2017). The dry mass of all experimental
plants was measured at the end of the experiment following the same procedure. Seedling
growth was tracked during the four months of the experiment, by measuring stem height (h),
diameter at 15 cm height (d) and number of leaves on a weekly basis. These measurements
were used to estimate the time course of wood and foliage biomass: for root total C we used
only initial and final harvest measurements. Initial root C was estimated by averaging all 10
harvested seedlings.

We estimated weekly total C in wood (C_{s,w}) from the measurements of stem height and
diameter, by using an allometric model fitted to initial and final harvest data.

\[
\log(C_{s,w}) = b_1 + b_2 \log(d) + b_3 \log(h)
\]

(1)

For each seedling, the total leaf area (LA) and foliage total C (C_{f,t}) over time (t) were estimated
based on harvested data (T = time of harvest) and weekly leaf counts (LC) over time.

\[
LA(t) = \frac{LA(T)}{LC(T)} \times LC(t)
\]

(2)

\[
C_{f,t}(t) = \frac{M_f(T)}{LC(T)} \times LC(t)
\]

(3)

Fully expanded new leaves were sampled for total non-structural carbohydrate (NSC)
concentration on a fortnightly basis. These concentrations were multiplied by leaf biomass to
estimate the foliage TNC pool (C_{n,f}) at each time point. The partitioning of the non-structural
C amongst foliage, wood and root tissues, according to empirically-determined fractions, was
then used to estimate the wood and root components of the total TNC pool. Structural C mass
for each component was estimated by subtracting non-structural C mass from total C mass.
Only foliage non-structural C ($C_{n,f}$) was measured, so to estimate the partitioning of the non-structural C among different organs, we used data from a different experiment on similar-sized seedlings of a related species ($Eucalyptus globulus$), which were grown in 5L pots until four months of age (Duan et al., 2013). We used data from the ambient well-watered control treatments. In that experiment, foliage, wood and root NSC were measured repeatedly over two months. There was no statistically significant change over time in the NSC distribution, so we used the mean distribution for mass-specific $C_n$ over time, which was calculated to be a ratio of 75:16:9 among foliage, wood and root pools.

We estimated daily GPP from leaf-level gas exchange measurements and a simple canopy scaling scheme as described in Campany et al. (2017), and summarized below. Measurements of photosynthesis were made fortnightly throughout the experiment on one fully expanded leaf per plant (Campany et al., 2017). Photosynthetic CO$_2$ response (AC$i$) curves and leaf dark respiration rates ($R$) were measured on two occasions, 13-14$^{th}$ March 2013 (when new leaves were first produced) and 14-15$^{th}$ May 2013 (prior to the final harvest). The AC$i$ curves were used to estimate photosynthetic parameters (the maximum rate of Rubisco carboxylation, $V_{c_{max}}$ and the maximum rate of electron transport for RuBP regeneration under saturating light, $J_{max}$) using the biochemical model of Farquhar et al. (1980) and fit with the ‘plantecophys’ package (Duursma, 2015) in R. The parameter $g_1$, reflecting the sensitivity of stomatal conductance ($g_s$) to the photosynthetic rate, was estimated by fitting the optimal stomatal conductance model of Medlyn et al. (2011) to measured stomatal conductance data.

Treatment effects on photosynthesis were detected immediately on newly produced (fully expanded) leaves and Campany et al. (2017) did not observe variation over time in photosynthetic rates. Hence, the photosynthesis parameters were assumed not to change over time but were specific for each treatment. Therefore, daily net C assimilation per unit leaf area ($C_{day}$) was estimated by using a coupled photosynthesis–stomatal conductance model (Farquhar et al., 1980; Medlyn et al., 2011) using mean photosynthetic parameters ($J_{max}$, $V_{c_{max}}$, $g_1$ and $R_d$) for each treatment and meteorological data from the onsite weather station. The daily GPP was estimated by multiplying $C_{day}$, total leaf area ($LA$) and a self-shading factor. The self-shading factor, which is a linear function of LA, is calculated by via simulation with a detailed radiative transfer model, the ‘YplantQMC’ R package of Duursma (2014) for individual treatment. The leaf maintenance respiration rate ($R_m$, g C g$^{-1}$ C plant d$^{-1}$) was calculated for each seedling by
scaling the measured rate (R) to air temperature using a $Q_{10}$ value of 1.86 (Campany et al., 2017).

### 2.3 Carbon Balance Model (CBM)

We used a DA-modelling framework, similar to that used by Richardson et al. (2013). This approach uses a simple carbon balance model shown in Figure 1. The model is driven by daily input of gross primary production (GPP), which directly enters into a non-structural C pool ($C_n$). The daily total maintenance respiration, $R_{m,tot}$, (calculated as a temperature-dependent respiration rate, $R_m$, multiplied by plant biomass), is subtracted from $C_n$ pool. The pool is then utilized for growth at a rate $k$ (i.e. $kC_n$). Of the utilization flux, a fraction $Y$ is used in growth respiration ($R_g$), and the remaining fraction $(1-Y)$ is allocated to structural C pools ($C_s$): among foliage, wood and root ($C_{s,f}$, $C_{s,w}$, $C_{s,r}$). The foliage pool is assumed to turn over with rate $s_f$. We assume there is neither wood or root turnover as the seedlings in the experiment were young.

**Figure 1:** Structure of the Carbon Balance Model. Pools, shown as boxes: $C_n$, non-structural storage C; $C_{s,f}$, structural C in foliage; $C_{s,r}$, structural C in roots; $C_{s,w}$, structural C in wood. Fluxes, denoted by arrows, include: GPP, gross primary production; $R_{m,tot}$, total maintenance respiration; $R_g$, growth respiration; $C_{t,lit}$, structural C in leaf litterfall. Fluxes are governed by six key parameters: $k$, storage utilization coefficient; $Y$, growth respiration fraction; $a_f$, allocation to foliage; $a_w$, allocation to wood; $a_r$, allocation to roots; $s_f$, leaf turnover rate.

The dynamics of the four carbon pools are described by four difference equations:
\[
\Delta C_n = GPP - R_m (C_{t,f} + C_{t,w} + C_{t,r}) - k C_n
\]  
\[\Delta C_{s,f} = k C_n (1 - Y) a_f - s_f C_{s,f}\]  
\[\Delta C_{s,w} = k C_n (1 - Y) a_w\]  
\[\Delta C_{s,r} = k C_n (1 - Y) (1 - a_f - a_w)\]  

Where GPP is the gross primary production (g C plant\(^{-1}\) d\(^{-1}\)); \(R_m\) is the maintenance respiration rate (g C g\(^{-1}\) C d\(^{-1}\)); \(C_{t,f}, C_{t,w}, \text{ and } C_{t,r}\) are the total C in foliage, wood and root respectively (g C plant\(^{-1}\)); \(k\) is the storage utilization coefficient (g C g\(^{-1}\) C d\(^{-1}\)); \(Y\) is the growth respiration fraction; \(a_f, a_w, a_r\) are the allocation to foliage, wood and root respectively; and \(s_f\) is the leaf turnover rate (g C g\(^{-1}\) C d\(^{-1}\)).

The non-structural (storage) C pool (\(C_n\)) is assumed to be divided amongst foliage, wood and root tissues (\(C_{n,f}, C_{n,w}, C_{n,r}\)) according to an empirically-determined ratio. Total carbon in each tissue (\(C_t\)) is then calculated as the sum of non-structural carbon (\(C_n\)) and structural carbon (\(C_s\)) for that tissue.

\[C_{t,f} = C_{n,f} + C_{s,f}\]  
\[C_{t,w} = C_{n,w} + C_{s,w}\]  
\[C_{t,r} = C_{n,r} + C_{s,r}\]

### 2.4 Application of Data Assimilation (DA) algorithm

DA was used to estimate the six parameters (\(k, Y, a_f, a_w, a_r, s_f\)) of the CBM for this experiment. All parameters were allowed to vary quadratically with time, i.e. each parameter was represented as:

\[p = p_1 + p_2 t + p_3 t^2\]

Quadratic variation over time was found to yield significantly better model fits than either constant parameter values or linear variation over time (see supplementary section S1). We executed three distinct sets of model simulations (Table 1), with the goals of (1) testing the need for a storage pool; (2) determining the effect of sink limitation on model parameters; and (3) attributing the overall effect of sink limitation on growth to the change in individual parameters.
For each set of model simulations, GPP and R\text{m} were used as inputs to the DA framework, and the measurements of total C mass of each of the plant components and foliage NSC concentrations were used to constrain the parameter values. The set of constraints included 18 measurements of C\text{t,f} and C\text{t,w}, two measurements of C\text{t,r} (start and end of the experiment), and six measurements of foliage NSC. There were 5 quadratically-varying parameters to determine for each treatment (as root allocation, a\text{r} = 1 – a\text{f} - a\text{w}), summing to a total of 15 (3x5) coefficients to determine, compared with total 44 data measurements available, for each treatment.

We used the Metropolis algorithm (Metropolis et al., 1953) as implemented by Zobitz et al. (2011), with broad prior Probability Density Functions (PDFs) for the parameters (Table 2). Values of k, a\text{f}, a\text{r} and s\text{f} were allowed to vary within the maximum possible range, while parameter Y was constrained according to the literature on growth respiration (Villar and Merino, 2001). Parameter a\text{r} was calculated from a\text{f} and a\text{w} with a check on a\text{r} to ensure that it had reasonable values (0 < a\text{r} < 1). Standard Error (SE) was used as an estimate of uncertainty on the assimilated data (Rowland et al., 2014; Richardson et al., 2010), and was calculated based on six replicate measurements. When combining errors, the errors were assumed to be uncorrelated (Hughes and Hase, 2010).

Model parameters were assumed to be real, positive and to have a lognormal probability distribution (Rowland et al., 2014). Therefore, all processes of parameter selection, and acceptance and rejection of parameters in relation to prior ranges were performed in lognormal space (Knorr and Kattge, 2005). We performed the first iteration starting from the prior set of parameters. To generate subsequent values for each parameter, a new point was generated by varying all vector elements by some step, chosen with a Gaussian distributed random number generator having a mean of 0 and a SD of 0.005 in log-normal space. We adjusted the step length for each parameter to values which lead to an average acceptance rate of the new points around 35–40%. We confirmed the chain convergence, having 3000 iterations to adequately explore the posterior parameter space, by visual inspection of the trace plots of different parameters as suggested by Van Oijen (2008). The trace plots show how the chain moves through parameter space for each individual parameter. The parameter vectors sampled during the first phase of the chain were not representative and therefore the first 10% of the chain was discarded from the posterior sample.

Table 1: Summary of the three model simulation sets
Simulation Set | Goal | Features | Addressing hypothesis
---|---|---|---
A | Test importance of storage pool | • DA applied to estimate parameters for model without storage pool and model with storage pool | H1
| | | • Three treatment groups | |
| | | • Not constrained with NSC data | |
| | | • No leaf area feedback | |
B | Identify effect of sink limitation on model parameters | • DA applied to estimate parameters for model with storage pool | H2-H5
| | | • Data divided into one, two, three or seven treatment groups | |
| | | • Constrained with NSC data | |
| | | • No leaf area feedback | |
C | Attribute overall effect on growth to changes in individual parameters | • Forward model runs to quantify impact of individual processes on overall plant growth | |
| | | • 5L & free seedlings considered | |
| | | • Parameters changed one at a time | |
| | | • Leaf area feedback | |

Table 2: Prior parameter PDFs (with uniform distribution) and the starting point of the iteration for all parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Starting value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>$Y$</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>$a_t$</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>$a_w$</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_t$</td>
<td>0</td>
<td>0.01</td>
<td>0.005</td>
</tr>
</tbody>
</table>

$a_t = 1 - (a_t + a_w)$, where $0 < a_t < 1$
2.4.1 Importance of storage pool

We tested the hypothesis (H1) on the importance of including a non-structural C storage pool in CBM by contrasting fits of the full model with fits of a simplified model without the non-structural C pool (Simulation Set A, Table 1). The simplified model omits the non-structural C pool (Cₜ) from the full model (Figure 1) and assumes that all available C is utilized for growth each day. We applied the DA framework to both model options and calculated the Bayesian Information Criterion, BIC (Schwarz, 1978) to determine the better model structure. BIC measures how well the model predicts the data based on a likelihood function and compare model performance taking into account the number of fitted parameters, with the lowest BIC number indicating the best model setting. For this comparison, both models were fit to the biomass data only, not leaf NSC data, in order to ensure that both models were fit to the same number of data points.

2.4.2 Effects of sink limitation on model parameters

The effects of sink limitation on C balance were investigated by applying the DA framework to data from all treatments combined, and then subsets of treatments (Simulation Set B, Table 1). Considering all treatments pooled together gives same parameters for all the treatments and effectively assumes no effect of sink limitation. On the other hand, taking more subsets of treatments produces more parameter sets (one for each subset) and allows for parameters to vary according to the degree of sink limitation. We first fitted the model to all data, ignoring treatment differences; then considered 2 treatment groups (free seedling / 5-35 L containerized seedlings), 3 groups (free / 5–15 L / 20–35 L) and 4 groups (free / 5-10 L / 15-20 L / 25-35 L). We also fitted the model to each of the 7 treatments individually, where the parameter set for each treatment is unique. The BIC values were compared across treatment groupings.
2.4.3 Attribution analysis

We performed a sensitivity analysis to quantify the impact of the response of each individual process to sink limitation on overall plant growth (Simulation Set C, Table 1). This analysis consisted of forward runs of the model, including a leaf area feedback to GPP. That is, rather than taking GPP based on measured LA (Eq. 9) as input, in this version of the model we calculated daily GPP using the modelled LA, the photosynthesis rate and corresponding self-shading factor. Adding the LA feedback to the model was necessary to quantify how the treatment effect on individual model parameters affects final seedling biomass through its effect on foliage mass, and consequently GPP, over time.

LA in each time step is estimated from NSC-free specific leaf area (SLA\textsubscript{nonsc}) and the predicted foliage structural carbon (C\textsubscript{s,f}) in that time step. SLA\textsubscript{nonsc} is calculated at harvest discarding the foliage NSC portion and is assumed to be constant for a given treatment throughout the experiment.

\[
LA = SLA_{\text{nonsc}} \times C_{s,f}
\]  

Once the LA feedback was implemented in the CBM, we ran the model with the inputs and modelled parameters from the smallest pot seedling (5 L), then changed the parameters to those for the free seedling sequentially in order to quantify the effect of each parameter on the final seedling biomass. The parameters we considered for this attribution analysis were: daily photosynthetic rate per unit leaf area (C\textsubscript{day}), maintenance respiration rate (R\textsubscript{m}), C allocation fractions to biomass (\(a_t\), \(a_w\), \(a_e\)), growth respiration rate (Y), foliage turnover rate (s\textsubscript{f}) and utilization coefficient (k). We additionally carried out a sensitivity analysis in which we varied each parameter from its baseline value separately.
3 Results

3.1 Importance of storage pool

First, we tested the null hypothesis (H1) that there is no need for a non-structural carbohydrate storage pool in the carbon balance model. We compared BIC values for model structures with and without a storage pool. Table 3 (Simulation Set A) shows the results for model fits with the optimal grouping strategy (three treatment groups). BIC values were consistently lower for the model including the storage pool; the improvement in model fit is most noticeable for the containerized seedlings. This analysis demonstrates that the model does need to include a storage pool to correctly represent the experimental data. In all remaining analyses, the full CBM (with non-structural C pool) is applied to data from all four plant C pools (NSC, foliage, wood and root biomass).

Table 3: BIC values from model fits. The lowest BIC values indicate the best performing parameter settings for any particular simulation. Note that, for Sim A, leaf NSC data were not used to constrain either model, to ensure that both models were fit to the same dataset, resulting in lower BICs compared to Sim B. Treatment groups are: ‘Small’ - 5 L, 10 L and 15 L containers; ‘Large’ - 20 L, 25 L and 25 L containers; ‘Free’ – freely rooted seedlings; ‘All’ - all data; ‘Containerized’ - all plants in containers.

<table>
<thead>
<tr>
<th>Simulation Set</th>
<th>Model Setting</th>
<th>Treatment groups</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sim A</td>
<td>Model without storage pool</td>
<td>Small</td>
<td>459</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Large</td>
<td>550</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Free</td>
<td>182</td>
</tr>
<tr>
<td></td>
<td>Model with storage pool</td>
<td>Small</td>
<td>215</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Large</td>
<td>338</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Free</td>
<td>167</td>
</tr>
<tr>
<td>Sim B</td>
<td>7 treatments combined</td>
<td>All</td>
<td>2768</td>
</tr>
<tr>
<td></td>
<td>2 groups</td>
<td>Containerized</td>
<td>1813</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Free</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1983</td>
</tr>
<tr>
<td></td>
<td>3 groups</td>
<td>Small</td>
<td>683</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Large</td>
<td>457</td>
</tr>
<tr>
<td></td>
<td></td>
<td>free</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1310</td>
</tr>
<tr>
<td></td>
<td>7 treatments individually</td>
<td>5 L</td>
<td>85</td>
</tr>
</tbody>
</table>
### 3.2 Sink limitation effect on C balance processes

We addressed our second null hypothesis (H2), that there is no effect of sink limitation on carbon balance processes, by comparing BIC values obtained for model fits when all treatments were combined vs separating the treatments into sub-groups. If there was no effect of sink limitation, the BIC value when all treatments are fit together would be similar to that obtained when treatments are separated into groups. The BIC values shown in Table 3 (Simulation Set B) decrease strongly as number of treatment groups increases, indicating a clear effect of sink limitation on carbon balance processes. Although the BIC values continue to decrease as more treatment groups are considered, we also found that interpreting parameter changes became more difficult as the number of groups increased. Hence, further analyses in this paper used unique parameter sets for three treatment groups: small containers, large containers, and free seedlings.

### 3.3 Analysis of carbon stock dynamics

Figure 2 shows the correspondence between modeled C pools and data. The model reproduced the key features of biomass growth over time in response to treatment. Biomass growth (Figure 2A, B and C) and the foliage storage pool (Figure 2D) were very clearly impacted by sink limitation: biomass growth was strongly reduced for containerised seedlings, which was very well mimicked by the model. Foliage growth in the free seedlings slowed towards the end of the experiment. Wood and root growth continued throughout the experiment in freely-rooted seedlings but slowed down during the second half of the experiment in containerized seedlings. NSC concentrations ($C_{n,f}/C_{t,f}$) in seedlings in small containers were higher compared those in free seedlings at the beginning of the season but all treatments had similar concentrations after four months (Figure 2D). In March, at the time of the first leaf NSC measurements, the foliage storage pool (Supplementary Figure S1) was similar in size across all treatments, but it increased over time in the free seedlings as these plants continued to grow, and decreased over time in the plants in small containers.
Modelled C stocks for all 7 treatments closely tracked their corresponding observations (Figure 2) as most of the predicted biomass values were within one standard error of the measurements. The exception is the 35 L container treatment, which is underestimated slightly because the grouping of 20, 25 and 35 L treatments into one group makes it difficult for the model to fit all treatments in this group.

**Figure 2:** Modelled C stocks (lines) with optimum parameter settings and corresponding observations (symbols): (A) total C mass in foliage $C_{t,f}$, (B) total C mass in wood $C_{t,w}$, (C) total C mass in root $C_{t,r}$ and (D) foliage NSC concentration $(C_{n,f}/C_{t,f})$. Note that the carbon pools and foliage NSC concentration (y-axes) are plotted on log scale to visualize the changes at the beginning of the experiment. Error bars (1 SE, $n = 6$) are shown for each observation.
3.4 Parameter estimates

Data assimilation indicated significant treatment effects on all five fitted parameters (Figure 3). There was a large effect of sink limitation on the utilization coefficient ($k$). In agreement with our hypothesis H3, the free seedling had the highest $k$, and the seedlings in small containers (most sink limited) had the lowest $k$ (Figure 3A). As the experiment progressed, the utilization rate of free seedlings began to decrease (Figure 3A). In contrast to the free seedlings, the potted seedlings had relatively low utilization rates initially ($k$ close to 0.5) and the utilization rates slowed down abruptly with time, most significantly in the smallest container treatments (Figure 3A).

In agreement with hypothesis H4, the estimated growth respiration rate ($Y$) varied according to the sink strength of the treatment groups, and was highest in the lowest sink strength treatments (Figure 3B). Moreover, $Y$ did not vary significantly over time for the sink limited treatment groups. However, the rate of growth respiration for the free seedling slowed down over time.

The data assimilation process also indicated that the growth allocation fractions vary among treatments and over time. Consistent with hypothesis H5, wood allocation fraction was highest in the smallest container treatments, and lowest in the free seedlings (Figure 3D). For the free seedlings, allocation was initially highest to foliage and roots (Figure 3C-E); over time, the plants reduced allocation to foliage and increased it to wood and roots. In the containerized seedlings, allocation was initially highest to wood and foliage; over time, foliage allocation decreased to almost zero and root allocation increased.

The estimated leaf turnover rate, $s_f$ was also notably higher for sink-limited treatments compared to free seedlings (Figure 3F). The large value of modelled leaf litterfall for sink-limited treatments is consistent with observations during the experiment that containerized seedlings had relatively large leaf litterfall, beyond normal senescence. Estimated $s_f$ increased over time for all treatment groups (most notably in free seedlings), due to a combination of ontogeny, seasonal change, and growth restriction in the sink-limited seedlings.
Figure 3: Modelled final parameters for three groups of treatments during the experiment period (21st Jan to 21st May 2013): (A) storage utilization coefficient, \(k\); (B) growth respiration fraction, \(Y\); (C) allocation to foliage, \(a_f\); (D) allocation to wood, \(a_w\); (E) allocation to roots, \(a_r\); and (F) leaf turnover rate, \(s_f\). The grey shaded area shows the 95% confidence intervals of modelled parameters.

3.5 Carbon budget

The model was used to partition total GPP (g C plant\(^{-1}\)) from the entire experiment period into different C pools (growth respiration, maintenance respiration, non-structural carbon, structural foliage, wood, and root carbon, and litterfall) for all 7 treatments (Figure 4). Total GPP was
considerably lower for the containerized seedlings, owing to lower photosynthetic rates per unit leaf area, $C_{\text{day}}$ (Figure 5A), and lower total leaf area (LA) per plant. Though starting with the same total LA of 0.016 m$^2$, the 5 L containerized and free seedlings had total LA of 0.031 and 0.516 m$^2$ respectively after four months of treatment. Simultaneously, the partitioning of GPP changed considerably across different treatments.

Small container seedlings (5, 10, 15 L) had a higher fraction of GPP lost in leaf litterfall compared to other seedlings (Figure 4), consistent with observations during the experiment. The proportion of GPP in final foliage mass was extremely low in sink limited treatments (also shown in Figure 2A). Allocation of GPP to final foliage and root biomass were highest in the free seedlings, although interestingly allocation to final wood biomass was similar across treatments. The final allocation to storage was also higher in free seedlings. The sink limited seedlings had a higher proportional C lost through maintenance respiration. Tissue specific respiration rates were similar in free and containerized seedlings, so the ~35% reduction in photosynthetic rate for the smallest containerized seedling, led to a higher overall $R_{m,\text{tot}}/\text{GPP}$ fraction. In summary, the estimated total respiration ($R_{m,\text{tot}} + R_g$) to GPP ratio was considerably lower for the free seedlings compared to the sink limited treatments. The carbon use efficiency (CUE) remained relatively constant and high over time for free seedlings (~0.65), whereas CUE in the smallest container treatments showed a sharp reduction over time down to ~0.25 (Supplementary Figure S2).
Figure 4: Proportional C partitioning for the whole experimental period. The total accumulated GPP (g C plant\(^{-1}\)) for individual treatments is shown (in red) at the top of each column. Free stands for free seedling. Different C partitions are in the colour legend: total litterfall, C\(_{t,lit}\); foliage structural C, C\(_{s,f}\), wood structural C, C\(_{s,w}\), root structural C, C\(_{s,r}\); non-structural C pool, C\(_n\); total maintenance respiration, R\(_{m,tot}\) and growth respiration, R\(_g\).

3.6 Attribution analysis

Sink limitation affected biomass growth via a range of processes, namely reduction in photosynthesis, and variation in the utilization rate, growth respiration, leaf litterfall, and C allocations to foliage, wood and root across various treatment groups. We quantified the contribution of each of these process responses separately by running the CBM with parameter inputs changing one at a time. Figure 5 shows how biomass (M\(_f\), M\(_w\) and M\(_r\)) is predicted to change when each parameter is changed sequentially from 5 L container (gray line, Figure 5) to free seedling (red line, Figure 5). Table 4 presents the effect of the parameters changing individually from the value of the smallest container treatment (5 L) to that of free seedling (FS) and other way around, resetting the previous parameter to the baseline value. The final biomass values indicate both the contribution (+/-) of each individual parameter separately.
(Table 4, 2nd and 3rd columns) and sequentially (Table 4, 4th column). Photosynthetic capacity had the largest individual effect on total plant growth (+15.28 and -71.9 g C) compared to the rest of the parameters. However, allocation pattern and the utilization rate also had a sizeable effect on final biomass (Table 4).

Daily net C assimilation per unit leaf area (C_{day}), which was 30% higher for free seedling compared to 5 L container treatment (Figure 5A), had a large impact on plant growth (final total biomass was increased by 11 g, Table 4 and Figure 5G-I, gray to orange). Maintenance respiration rate (R_m) did not vary significantly across treatments (Figure 5B), in line with the data presented in Campany et al. (2017), and consequently its impact was insignificant (the final total biomass is reduced by only 0.24 g, Table 4 and Figure 5G-I, orange to light blue).

The modelled biomass allocation fractions (a_f, a_w and a_r) in Figure 5C had important, but mixed, effects on C stocks. The final foliage mass was increased from 3.4 g to 9.6 g due to the increase in C allocation to foliage (Figure 5G, light blue to green), which has a positive feedback on GPP. Concomitant changes in C allocation to wood and root resulted in smaller changes to these biomasses as shown in Figure 5H-I (2.5 g and 7.0 g rise respectively). Overall, the change in allocation pattern resulted in an increase in final total biomass by 15.74 g (Table 4). Growth respiration rate (Y) was ~20% lower in free seedlings (Figure 5D), which had a considerable impact on C budgets (the final total biomasses were increased by 9.56 g, Table 4 and green to yellow, Figure 5G-I). Leaf turnover, s_f was low in the free seedlings compared to the 5 L container treatment (Figure 5E) which had a large positive effect on final C pools (yellow to blue, Figure 5G-I). The foliage mass was increased by 5.6 g; the wood and root masses were also further increased (3.4 g and 5.8 g respectively) due to the increase in GPP when foliage is retained for longer. Finally, the utilization coefficient, k was higher in free seedlings (Figure 5F) causing a 20-30% positive feedback on C budgets (total biomass increased by 23.08 g, Table 4 and blue to red, Figure 5G-I).
**Figure 5:** Attribution analysis. Left column (A-F): changes in input parameters; Right column (G-I): associated impacts on C budgets (right column, G-I). Different colours in the figure indicate the parameter shifts (left column, A-F) and their associated impacts on C budgets (right column, G-I). Legend: 5L, highly sink-limited treatment with container size of 5 L; FS, Free Seedling without any sink limitation. Note that the orange line is overlaid by the light blue line: the small change in maintenance respiration results in a very minor effect on biomass growth.

**Table 4:** Estimates of final biomass due to parameter change (individual and sequential), showing the contribution of each parameter separately and successively to biomass changes. All values in g C plant⁻¹. +/- indicates biomass increase or decrease due to particular parameter change. The final column corresponds to the changes shown in Figure 5.

<table>
<thead>
<tr>
<th>Parameter change</th>
<th>Individually</th>
<th>Sequentially</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Cₚ</td>
<td>5.81</td>
<td>83.99</td>
</tr>
<tr>
<td>Cₚday</td>
<td>+15.28</td>
<td>-71.9</td>
</tr>
<tr>
<td>Rₚ</td>
<td>-0.08</td>
<td>+1.1</td>
</tr>
<tr>
<td>(a_r + a_w + a_r)</td>
<td>+1.53</td>
<td>-45.5</td>
</tr>
<tr>
<td>Y</td>
<td>+0.41</td>
<td>-19.22</td>
</tr>
<tr>
<td>sᵣ</td>
<td>+1.13</td>
<td>-19.17</td>
</tr>
<tr>
<td>k</td>
<td>+0.44</td>
<td>-23.08</td>
</tr>
<tr>
<td>FS total observed Cₚ</td>
<td></td>
<td>83.99</td>
</tr>
</tbody>
</table>

### 4 Discussion

#### 4.1 Effects of sink limitation on C balance

Our DA-model analysis of this root volume restriction experiment provided significant new insights in the response of key C balance processes to sink limitation. We were able to infer that, in addition to a reduction in photosynthetic rates, sink limitation reduced NSC utilization rates, increased growth respiration, modified allocation patterns and enhanced senescence. Our attribution analysis indicated that all of these process responses contributed significantly to the overall reduction in biomass observed under low rooting volume.

We first tested the null hypothesis (H1) that seedling growth rates could be adequately predicted from current-day photosynthate. This hypothesis was rejected, with a storage pool being necessary to simulate growth, particularly for containerized seedlings (Sim A, Table 3).
The approach of simulating growth from current-day photosynthate is commonly used in models, particularly for evergreen plants (e.g. (Jain and Yang, 2005; Law et al., 2006; Thornton et al., 2007)), but several authors have proposed the need for a storage pool to balance the C sources and sinks in the short term, as well as simulate the effects of photosynthetic down-regulation in the long-term (Pugh et al., 2016; Richardson et al., 2013; Fatichi et al., 2016). Our results support the need for an NSC pool in CBMs.

We then tested the second null hypothesis (H2) that there was no effect of treatment on the parameters of the C balance model. This hypothesis was also rejected: fitting the DA-model framework simultaneously to all treatments with one set of parameters (ignoring sink limitation effect) gave a low goodness of fit (Sim B, Table 1). This result is consistent with the finding of Campany et al. (2017) that the observed effects of sink limitation on photosynthesis in this experiment were insufficient to explain the large reduction in biomass. Instantaneous photosynthetic rates were reduced 20-30% by sink limitation. Our DA analysis indicated that several other processes contributed to the reduction in biomass growth, including carbohydrate utilization, growth respiration, allocation patterns, and turnover.

Our results suggested a significant effect of sink limitation on the carbohydrate utilization rate, \( k \) (Figure 3A). The modelled \( k \) values were approximately twice as large in free seedlings compared to the small containers. This result supports the hypothesis (H3) that plants would have the lowest utilization rate under sink-limited conditions. At the start of the measurement period, the free seedlings were utilizing almost all C produced immediately in growth (\( k \) close to 1.0, Figure 3A). The utilization coefficient of the free seedlings decreased over time, causing a build-up in C storage (Figure 2D). This decrease in utilization rate could potentially be an ontogenetic effect, with free seedlings initially allocating all carbon to growth during establishment but increasing storage with increasing size. However, ontogenetic effects are confounded with season in this experiment, such that decreasing utilization rates over time could also be a result of decreasing temperatures moving into autumn. There is a real need to quantify how the carbohydrate utilization rate varies with environmental conditions and ontogeny; data assimilation of experiments in which photosynthesis and growth rates have been monitored over time offer one means to do so.

Although the carbohydrate utilization rate was highest in the free seedlings, leaf carbohydrate concentrations were not lower in these plants at the end of the experiment. As shown in the final C budget analysis (Figure 4), there was a higher total C allocation to the NSC pool in free
seedlings than sink-limited treatments. Final carbohydrate storage was high in free seedlings despite high $k$ because the carbohydrate pool was recharged throughout the experiment (Figure 2D), as the free seedlings had high photosynthetic rates but no higher maintenance respiration requirement. In contrast, NSC was depleted for the smallest pot treatments after mid-March (Figure 2D) when demand exceeded supply due to both limited production of photoassimilates and enhanced leaf litterfall (Figure 3F).

The modelled rate for growth respiration, $Y$ was larger for sink limited treatments than the free seedling (Figure 3B). Overall, there was lower C utilization (i.e. CUE) in plant structural growth in sink-limited treatments (~45%) compared to free seedlings (~60%). This finding supports the “wasteful plant” hypothesis H4. Inferred $Y$ remained constant over time for the containerized treatments, implying a fixed portion of C loss due to growth respiration despite seasonal variation. However, a reduction in $Y$ over time was inferred for the free seedling, suggesting a possible ontogenetic effect. However, it is important to note that we have inferred growth respiration from the CBM framework. Therefore, these estimates could possibly also include C losses via other pathways. Direct measurements of growth respiration rates would be useful to confirm the inferred effects of sink limitation and investigate potential underlying mechanisms.

We also demonstrated that the allocation fractions among organs change in sink-limited conditions, with sizeable consequences for plant growth rates. Previous analyses of pot-size experiments have generally only been able to estimate changes in final biomass partitioning (e.g. Poorter et al. 2012a). Campany et al. (2017) analysed final biomass partitioning in the experiment and did not find any significant difference in biomass partitioning in sink-limited seedlings compared to free seedlings, once ontogenetic drift was taken into account. Our analysis adds to that of Campany et al. (2017) by calculating the dynamics of allocation over time and taking estimated foliage loss into account. Our analysis showed that modelled allocation fractions vary significantly over time (Figure 3C, D and E). In the free seedlings, allocation to foliage decreased, and allocation to both wood and roots increased, reflecting the ontogenetic effects mentioned by Campany et al. (2017). However, our analysis also highlights significant variations among treatments in the modelled C allocation fractions to foliage, wood and root that are not ontogenetic. At the beginning of the experiment, foliage allocation fractions were similar for all treatment groups, but wood allocation was higher, and root allocation lower, in the containerized seedlings compared to the free seedlings. For the
containerized seedlings, changes over time also differed from those in the free seedlings: wood allocation decreased marginally, rather than increasing, foliage allocation declined steeply over time, and root allocation increased steeply. These allocation patterns supported our hypothesis H5 that sink limitation due to root restriction would favour allocation to wood over foliage or fine roots. Calculating dynamic allocation patterns over the course of an experiment thus provides additional insights beyond analysis of the final biomass outcome.

4.2 Application of DA to infer C balance processes

We have demonstrated that the DA approach can be an invaluable tool for quantifying C fluxes in experimental systems, enabling us to extract important new information from existing datasets to inform carbon balance models, such as the rate and timing of the transfer of photosynthate to and from storage pools. The DA-modelling approach is able to draw together the experimental data to estimate all the components of C balance, including photosynthesis, respiration, NSC, biomass partitioning and turnover. This approach could readily be applied to other experiments to derive new information allowing better representation of C balance processes in vegetation models.

Applying this approach requires a range of measurements to constrain the key C balance processes. Here, we used estimated daily C assimilation and maintenance respiration rate as model inputs and constrained the model with measurements of biomass pools (foliage, wood, root) and foliage NSC concentrations. We used fortnightly foliage and wood biomass measurements; the DA framework would work with fewer data observations, but parameters would be estimated with less accuracy. Informal exploration of the model suggested that measurements of foliage turnover would have been particularly useful to better constrain the model. Any experiment having estimates of GPP, maintenance respiration, and structural biomass could potentially be investigated with this framework. However, additional measurements of storage and turnover would be highly beneficial for the performance of the simulation. Repeated observations over time are also useful, particularly for young plants, to account for variations in parameter values over time. We found significant changes in parameter values during the course of the 4-month experiment, which may be linked to both ontogeny and seasonal variation in temperature.

One major caveat on our results is that below-ground carbon cycling processes were not well characterized. For practical reasons, processes such as root growth, respiration, turnover, and
exudation are rarely well quantified in empirical studies. Here, we had access to initial and final estimates of root biomass. Root respiration was estimated; root turnover and exudation were assumed to be zero. There is evidence that stress can increase rates of root exudation: for example, Karst et al. (2016) demonstrate increased exudation rates in seedlings exposed to cold soils. They also showed that stressed plants may exude C beyond that predicted by simple concentration gradients in NSC between root and soil. The loss of C independent of NSC in roots suggests that exudation may be actively enhanced once plant growth is limited (Hamilton et al., 2008; Karst et al., 2017). As our CBM does not include this process, it would attribute any C loss through root exudation to another process removing C from the plant, such as growth respiration. The increase in growth respiration that we inferred may thus potentially include root exudation. We have reasonable confidence, from the combination of measurements available, in our inference that the C loss term was increased with sink limitation. However, direct measurements of one or both processes would be required to determine the role of root exudation.

In addition, we did not have access to estimates of root or wood NSC. We used data measured in a previous experiment on 4-month old E. globulus seedlings (Duan et al. 2013) to estimate these values from foliage NSC. It would have been useful to obtain these values, particularly since wood and root tissue can act as storage organs, and the timing of storage development would be extremely useful to quantify. The concentration of NSC in plant roots measured by Duan et al. (2013) was relatively small compared to that of foliage (mean 2.15%). However, fine root NSC values in a nearby experiment on 17-month-old E. parramattensis saplings were even lower (0.78%) (Morgan E. Furze et al. unpublished data). It is possible that these very fast-growing Eucalypt species only start to accumulate root reserves when they are established. Further research is needed to quantify the trade-off between allocation to growth and storage during establishment.

4.3 Implications for modelling plant growth under sink limited conditions

The goal of our study was to examine how carbon balance models should be modified to represent sink limitation of growth, whilst maintaining mass balance. Our results demonstrate that several process representations need to be modified. Firstly, we demonstrate a clear need to incorporate a carbohydrate storage pool, with a dynamic utilization rate for growth. We demonstrate that the utilization rate is slowed by sink limitation, and may also vary with ontogeny. Targeted experimental work is needed to better quantify this variation in utilization...
rates. Secondly, in addition to a feedback on photosynthetic rates, other plant processes including growth respiration, turnover and allocation are also affected by sink limitation. Applying a DA-modelling framework to experimental data with rooting volume restriction has allowed us to quantify these effects in this experiment. Applying this approach more broadly would potentially allow us to identify general patterns that could then be formulated for inclusion into models.

The inferences on carbohydrate dynamics from seedling studies could be used to infer mature tree responses that can subsequently be integrated at ecosystem level and beyond using the concepts of Hartmann et al. (2018). We are enthusiastic to see the approach applied to other experiments, but there are likely to be gaps in the datasets to constrain the key C balance processes. Fortunately, the DA approach does not require continuous measurements of all of the C stocks and fluxes. In the absence of measurements, the model can be relied upon to project the time evolution of missing stocks and fluxes, although of course, the precision of model estimates and insights that can be gained, increases with data availability. DA can also be applied at ecosystem scale. There are several successful examples of DA being applied to forest growth, albeit without a focus on storage (e.g. Van Oijen (2008); Williams et al. (2005); Bloom et al. (2016); Quaife et al. (2008); Pinnington et al. (2016)). Overall, this approach provides important insights into the regulation of carbohydrate storage and would significantly advance our ability to predict the impacts of environmental changes on plant growth and vulnerability to stress.

**Data availability**

The raw data are freely available on Figshare (doi: https://doi.org/10.6084/m9.figshare.5125087.v3). The R source code to perform all the data processing and analysis to replicate the figures is freely available as a Git repository (https://github.com/kashifmahmud/DA_Sink_limited_experiment).

**Author contribution**

KM analyzed the data, developed the model code, performed the simulations and wrote the paper. BEM conceived the idea and helped in data analysis. RAD and CC provided the experimental data. BEM, RAD, CC and MGD provided in-depth editing of the manuscript.

**Competing interests**
The authors declare that they have no conflict of interest.

Acknowledgements
This research was supported by the Australian Research Council (Discovery, DP
DP160103436), the Hawkesbury Institute for the Environment, and Western Sydney
University. The authors wish to thank Burhan Amiji for his technical assistance and all
individuals from Hawkesbury Institute for the Environment who helped during the
experimental harvest. We thank Mathew Williams for advice on implementing the data
assimilation framework.

References
Arp, W. J.: Effects of source sink relations on photosynthetic acclimation to elevated carbon

state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools,
and residence times, Proceedings of the National Academy of Sciences, 113, 1285-1290,

BoM: Climate Data Online (Station 067105), Bureau of Meteorology Melbourne.

Bossel, H.: treedyn3 forest simulation model, Ecol. Model., 90, 187-227,
https://doi.org/10.1016/0304-3800(95)00139-5, 1996.

Bradford, K. J., and Hsiao, T. C.: Stomatal behavior and water relations of waterlogged
tomato plants, Plant Physiol, 70, 1508-1513, 1982.

Buckley, T. N.: The control of stomata by water balance, New Phytologist, 168, 275-292,
Campany, C. E., Medlyn, B. E., and Duursma, R. A.: Reduced growth due to belowground sink limitation is not fully explained by reduced photosynthesis, Tree Physiol, 37, 1042-1054, 10.1093/treephys/tpx038, 2017.


Jain, A. K., and Yang, X.: Modeling the effects of two different land cover change data sets on the carbon stocks of plants and soils in concert with CO2 and climate change, Global Biogeochemical Cycles, 19, n/a-n/a, 10.1029/2004GB002349, 2005.


Rowland, L., Hill, T. C., Stahl, C., Siebicke, L., Burban, B., Zaragoza-Castells, J., Ponton, S.,
Bonal, D., Meir, P., and Williams, M.: Evidence for strong seasonality in the carbon storage
and carbon use efficiency of an Amazonian forest, Glob. Change Biol., 20, 979-991,

Running, S. W., and Gower, S. T.: FOREST-BGC, A general model of forest ecosystem
processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets1,

Sage, R. F.: Acclimation of photosynthesis to increasing atmospheric CO2: the gas-exchange


dynamic model for predicting intra-annual sink demand and interannual growth variation in

Schwarz, G.: Estimating the Dimension of a Model, Ann. Statist., 6, 461-464,

Thornley, J. H. M., and Cannell, M. G. R.: Modelling the Components of Plant Respiration:

carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate
variability, Global Biogeochemical Cycles, 21, n/a-n/a, 10.1029/2006GB002868, 2007.

Van Oijen, M.: Bayesian Calibration (BC) and Bayesian Model Comparison (BMC) of

Villar, R., and Merino, J.: Comparison of leaf construction costs in woody species with
differing leaf life-spans in contrasting ecosystems, New Phytologist, 151, 213-226,

