Dear Dr. Ke  
  
Thank you for your submission of your Tansley review manuscript. I have now received three reviews, which were somewhat mixed in their assessments. I have also read your manuscript with interest. Like reviewers 1 and 2, I found many ideas interesting. However, I also agree with the assessment of reviewer 2 that "I think the ms is not nearly mature enough to publish, many arguments are simply repeated throughout the ms, without gaining in depths as I read on". Both reviewers 2 and 3 rated the manuscript as being of only average scientific importance, and reviewer 3 had a particularly negative overall assessment.

Based on the reviews, the manuscript is unfortunately falling substantially short of the standards for a Tansely review as reflected in 2 of the 3 reviews. Therefore, despite some very interesting aspects, I am unfortunately unable to recommend acceptance of the manuscript.  
  
I regret not having a more positive outcome from this review, but I hope that you will find the reviews helpful. From my perspective, it feels like some of the most novel ideas in the review could form the basis of a strong but much shorter "Tansley Insight". If you choose this route, I believe we could treat your submission as a new manuscript.  
  
Sincerely,  
Ian Dickie  
Editor, New Phytologist  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
Decision: Reject  
  
Referee: 1

Comments to the Author  
The authors address several important and timely topics of plant-soil feedback that are previously neglected or underexplored. They focus on the temporal dynamics of plant-soil feedback, and the effect of feedback on demographic rates of plants other than biomass. The authors then outline several future directions that could guide theoretical and empirical studies. This well-written review paper covers many topics comprehensively yet concisely, and it fits well with the scope and format of a Tansley Review.  
  
Overall, I do not have many comments on the content and organization of the manuscript as the authors did an excellent job, but there are several places that I wish could be elaborated/clarified. Some are rather minor but might interrupt the flow of the paper.  
  
Line 10-11: “These experimental designs…” reads repetitively from the previous sentence in line 7, “This experimental design…”, and could confuse readers on what “these” refers to.  
  
Line 105-107, 366-370, and 401-406: In addition to plant phenology and their regulations from microbes mentioned in these lines, soil microbes also have distinctive phenology (reviewed by Rudgers et al. 2020), and shifts in both plant and soil microbial phenology could lead to mismatches with fitness consequences. For example, if plants germinate late, they might be at a smaller stage that is more vulnerable to pathogens. In line 105 the temporal variation of soil microbial community and plants’ response seems to be a good introduction to the potential phenological mismatches, although the subsections focus more on the temporal developments of interactions within a year rather than the starting time of the interactions. This phenological mismatch is worth mentioning especially since part of the initiatives of this review is to discuss temporal processes underlying plant-soil feedback and steering towards a less plant-centric view of plant-soil feedback.  
  
Line 128-129: Is there a need of emphasizing taxonomic composition?  
  
Line 168-169: Why does this result potentially indicate functional redundancy?  
  
Line 243-244: From the previous subsections I don't sense the lack of empirical information but a lack of consensus: there are plenty of studies that show the temporal dimension of plant-microbe interactions, but they are highly context-dependent and often have contrary results. Therefore, this concluding section should spend some time summarizing the different factors that could contribute to this context dependency, e.g., functional group and life history of the conditioning and response plants, and groups of microbes involved, and make this part of the recommendations for future experiments.  
  
Line 331-333: Long and complicated sentence.  
  
Line 412-415: This statement reads a bit overly pessimistic, especially in a section that talks about future directions.  
  
Line 428: This section talks about two potential paths of modifying current plant-soil feedback models using patch-occupancy models and letting soil microbes affect demographic rates explicitly. While these are important and correspond to the previous two sections of the paper, this section did not mention a family of models that I argue would be fundamental to understanding the temporal dynamics of plant-soil feedback: discrete-time models. First, patch occupancy model is spatially discrete (which was also not emphasized in this section), and the paper has already covered how it could help to advance the classic models. Given that one of the paper's main objectives is to advocate for temporally explicit studies of plant-soil feedback, it is rather surprising that the paper did not discuss discrete models in the temporal dimension. Discrete-time models usually do not allow for timescale separation, which “forces” the modeler to explicitly consider the temporal dynamics of soil microbiomes. It is also suitable for many realistic scenarios mentioned in the paper, such as the gaps between plant generations, decay of soil microbiome, temporal turnover of microbial community structure, and phenological mismatches between plants and microbes in my previous comment. Although discrete-time models with microbial dynamics would require much more effort to parameterize, so do many demographic models mentioned in the paper. I understand that this is not the focus of the paper (and there’s likely not much to review given the lack of such models in literature), but it would still be nice to mention this just to complete the narrative.  
  
Line 567: This is not very precise. The intrinsic growth rate usually refers to the population growth rate (i.e., births minus deaths), but most plant-soil feedback experiments in the greenhouse do not grow plants for more than one generation. Some collect seeds to measure fecundity, others may just collect biomass (which was mentioned previously in the paper). In the latter case, they can only measure the physiological growth rate/developmental rate and maybe productivity, but not the intrinsic growth rate.  
  
Line 567-569: This is also not very precise. If models are parameterized by observational CNDD data in the field then that reflects all possible mechanisms happening in the field, not only microbial effects. Unless the data is collected from controlled studies (e.g., removal of litter, herbivore exclusion, or application of biocide) it cannot be used to imply the role of microbes.  
  
Box 1. I like the idea, but I am confused by the figure. What do the transparent circles mean and how are they different from the scattered pie charts? Also, the pie charts and circles in the “field” panel are all overlapping and hard to see.

Referee: 2  
  
Comments to the Author  
Po-Ju Ke et al contribute a well written ms about an interesting and relevant topic to the plant soil feedback literature, which in turn has relevance for our understanding of the plant communities in our natural and agroecosystems. They add that soil conditioning and plant responses to altered microbial composition and functioning are dynamics in time and need to be included in PSF models and our conceptual thinking. I think that is a good idea, although I think the ms is not nearly mature enough to publish, many arguments are simply repeated throughout the ms, without gaining in depths as I read on. In addition, I don’t care too much about novelty in general, but the need to assess different demographic rates in PSF has been called out in the literature before (also cited in the ms), and it is not clear to me what new ideas you are adding to the existing literature (e.g. L399-401). I think the temporal aspect has a clear potential to be of publishable value to the journal, but I think it needs to be worked out in more depth.  
  
For one thing, I don’t think the classic two stage design really contradicts the idea of dynamic (re)conditioning and plants responding (L40 and elsewhere). It is a simple tool to get some idea of PSF and you are right to point out it may be too simple, but I think many people in the PSF field have this constant reconditioning in mind when thinking about plant-microbiome interactions - although indeed they might think a bit more explicitly about it. Also, you say the classic approach assumes instantaneous feedback development (e.g. L60-61), but I think the assumption is that the development time is short (i.e. weeks-months), which makes sense since we are finding measurable PSF on those timescales. Your point that it is not constant (or develops linearly or so) is much more interesting, but needs to be made more precise. Based on the (limited) data and literature we do have, what do the dynamics look like? And if we feed that into model, what does that mean for the (long term) outcome? How wrong are our current predictions? In the intro and elsewhere you need to better argue why it is important to take a more dynamic view to PSF (e.g. L52-54, 70-72, 135-138). Why is it essential we do that? What will we get wrong if we don’t? by what margin of error? All models are wrong, but some useful – why is the classic PSF model not (sufficiently) useful? That is the case you have to make.  
  
For instance with respect to the timelag between conditioning phase harvest and feedback phase setup I am not so convinced it will matter a lot. E.g. in section II.2 several examples are given where PSF generated from soil collected under life or dead trees did in fact not differ. I think those examples argue against inclusion of timedelays in models. Also, it is left unclear how frequently and for what duration soil patches are ‘empty’ of living roots upon plant death in the field. I think that will be the case only during cold or dry periods for more than a few weeks for most systems, and it is not clear to me if the induced feedback really changes over such periods of generally low biological activity.  
  
I advise that you incorporate a bit more detailed info on the changes happening in microbial community composition and plant responses, qualitative and quantitative, over the different timescales (e.g. L105-108, 129), that helps readers to get a better sense of the importance. Just knowing something ‘changes’ does not mean the change is in any way important or meaningful.  
  
Regarding the proposed models. You propose to work with an annual plant population model (box 3) to judge different demographic rates, but can that account for the short term (intra-annual) conditioning dynamics (L122-126), and also can it account for perennial species? Is that model flexible enough? Do you need shorter timesteps then years? How do the time-varying PSF interaction strengths enter into your proposed model? From Box3 it seems the microbial effects are also not dynamic in time, i.e. you talk of an increase of 40% only, how does that capture the temporal complexity you discuss in the main paper? Regarding patch and demographic models, is one approach more useful than the other do you think? I think in the last part of the paper you mention some examples where this is (partly) done – can you give that more prominence? Can you maybe add that to Box3 and include a model for perennials with time-varying effects of PSF? In that vein I was also wondering, would it be possible to adapt integral population projection models to this case?  
  
The ‘separation of timescales’ is mentioned a few times in the ms (e.g. L592) as a simplifying approach to temporal microbial dynamics, but it is never worked out what it is. I think it could be one of the central organizing ideas in the paper. Maybe you could make a table of ‘simplifying assumptions’, the risks of ignoring them, and the potential gains when we include them in models.  
  
Detailed comments and concerns are listed below. I hope you can revise along these lines as I think it will really enhance the value of reading your paper and thus its potential impact.  
  
Minor  
I am missing the perspective that plants during soil conditioning also change their soil abiotic environment in the paper (c.f. abstract, introduction).  
L5- note that soil conditioning by plants also changes abiotic aspects of the soil  
L10 – change to “measuring just plant biomass” to make it more clear  
L12 – can you name those processes to make it concrete?  
L28: actually Jim introduced it first in Bever 1994 Ecology doi 10.2307/1941601 as far as I know, the first maths cam in the 97 paper.  
L29 also include the Bever et al 2015 Annual reviews paper (doi 10.1146/annurev-ecolsys-112414-054306) in your thinking.  
L37: that ‘immediate’ is maybe an open question, I don’t see it often reported how much time went between the harvest of phase 1 and the setup op phase 2.  
L46 can you change comport to a more simple word, I had to look it up.  
L49 that ‘see also other studies’ part is useless to a reader – where should they look?  
L50 shapes or should reshape?  
L53 I get annoyed when people talk about the ‘effect of time’, time does not do anything of itself, it only provides the temporal space for processes to happen.  
L56 why would you neccesarily need a different feedback duration per se for longer lived plants? Yes I can see that if you want to know eventual seed set, but what if you want to know about changes in competitive ability due to microbes over a year for two plant species that co-occur?  
L69 but is that really unreasonable? I guess microbes change over a winter period, but then in the spring new seeds emerge and the feedback and reconditioning go on again. That temporal delay is maybe not that long, particularly considering that we can trace microbial legacies for years-decades?  
L71 I am still not too convinced by the criticalness by the present text so far – I agree with you from what I know, but you have to sharpen the argument I think.  
L73-75: I don’t get this sentence, people usually focus on one demographic aspect only – so what is now necessitated?  
L84-92: this section is repeating the same ideas as above – here is a place to dive into things a bit deeper still and convince me why the Bever model is not enough. That demograhic rates could be different does not mean we get the coexistence wrong or does it?  
L94-96 this humility sentence is a bit unnecessary, just say what you focus on.  
L98-99: that is a lot more work, justify that time investment to me.  
L146-7: does that not support the idea that short term assessment capture the critical part of PSF, when the plants are still sensitive.  
L139-154: this para is meandering a bit and it is not apparent what the take away message is, please rewrite.  
L162: please use the word critical only when something really is critical, you are not being so convincing in this way.  
L172: for how many of the papers in those two meta-analyses do we know the time that elapsed between conditioning phase harvest and feedback phase setup (seeding/planting)?  
L176: I doubt that in many natural ecosystems soil will be ‘empty’ of roots for much time really. Is there work done on this in the field that can give us a sense of the frequency and duration of soil ‘emptyness’?  
L179: yes, but do these effects occur in real-world conditions too?  
L180-181: if life and dead tree soils give the same feedback, does that not argue against your point that time since conditioning needs to be accounted for?  
L182: that makes me think of Casper et al 2003 plants zone of influence paper – would those timelags differ a lor for species that form monospecific stand and species that blend in the multispecies matrix more (like prunus serotina)?  
L187: yeah, key in that sense are I think microbial seedbanks, see Lennon & Jones 2011 doi 10.1038/nrmicro2504.  
L189: as before, does this example not argue against inclusion of time delays in our models since life and dead tree soil give the same PSF?  
L191-194: yes, but I don’t think they need to stay active, they can just go dormant and wait until a suitable (indirect) hosts comes along again. Many of these microbes can persist for years or decades in soil without being active (what is it 80% or so of cells in soils are inactive at any one time).  
L204: yes, but do those phyllosphere microbes change the PSF very much. I guess you have a point for litter-mediated feedbacks, but how strong are those compared to direct PSFs induced by root-associated microbes?  
L209: can you indicate what mechanisms are involved in that pathogen suppression?  
L212-213: you have not yet convinced me that we actually need too incorporate those based on this section. Some more work is needed.  
L224-6: you have not really discussed that evidence yet. That PSF change with nutrients and competition, does not mean that feedback over the plants lifetime is dynamic – you need to flesh that out more.  
L227-241: okay this is a good paragraph, and I think this may need to come first in the section II.3. Then in a next para you should review the available evidence for the (relative) magnitude of the impact of temporal dynamics in feedback on plants.  
L248: and across growing seasons? are they also well reflected do you think?  
L250: it is not clear to me why we necessarily need to understand within-year dynamics to project multi-year responses? Can we not give an aggregate effect for a given year, a net result of all that nitty-gritty variation?  
L254: yeah, the drought (or dry period) part here is key I think, it is not so much time per se that we need to include, but time under the env conditions that matter during that time.  
L256-8: but that is being done right? So what insight did we get from those approaches so far? What is the untapped potential still around?  
L260-2: yeah your study there is one of the few nice examples I think, can you use that to underline the potential importance of the dynamic PSF perspective a bit more in the ms?t  
L269-70: since conditioning and responding of and to PSF are continuous in nature, should we not explore more non-destructive measures of growth too? E.g. use minirhizotrons timeseries?  
L307: where do you base the ‘often’ on? Can you quantify that in some way?  
L308-10: I think the main thing missing here is to give the reader a sense of how important these things are. How often and to what degree do microbial effects on different demographic rates vary? To what extend does that limit prediction of longer term consequences? How wrong will those predictions be?  
L316: and does that grouping not give us net compound effects, is that not enough? Why do we need to know exactly when the seedling died, is it not enough to know it died somewhen in those first weeks/months?  
L325: write ‘integrated’ instead of ‘integrative’ – I also like compound microbial effect.  
L340: This read like the start of a new topic, and makes me expect a new paragraph, please rephrase so the reading flow is more continuous.  
L344: what aspect of location did they study?  
L346-8 there for sure are a lot more papers on microbial effects on trees in the tropics… Do they not need to be included?  
L352: yeah, although I think there are also many tropical plants that rely more on resprouting from stems and root (fragments) then regeneration from seed. I would make a comment about the relative frequency of these ways of plant propagation.  
L358: replace ecological setting with natural setting – agroecosystems are just as ecological as natural ones.  
L387-9: this is a bit repetitive from before.  
L407: what herbivores? Ungulates? Insects?  
L421: would it not be logical to have the models precede the data, indeed to inform what data is important to collect?  
L424: what do you mean by ‘remains a critical research direction’?  
L440: I think the Bever et al 1997 model also includes competitive effects among plants? So microbes are not the only mechanism.  
L457-8: this timescale separation needs to be explained.  
L468-9 why is it needed to now these traits for parametrization? Make that explicit.  
L471-2 could patch model not model compound microbial effects from seed-to-seed depending on the data used for parameterization?  
L482 but Box3 only uses an annual plant model, extend it to include the perennial case please.  
L526-8: it is still not clear to me how much including those gaps is likely to matter for our predictions.  
L541-4: it is not clear to me how that is fundamentally different from have a patch model with multiple plant life stages? Why is patch occupancy frequency so different from density?  
L546-7: yes and if you can provide guidance on those choices that would be more helpful. In the examples that follow, how should we model that?  
  
Box 3: why the Beverton-Holt model? Are the relevant others? How does this play out for perennial plants? Check the subscripts in the equation, should i and j not be 1 and 2 (as in N1 and N2 etc). Regarding live-sterilized comparison, would a conspecific-heterospecific soil comparison not be more appropriate? You consider only 100 timesteps in your sensitivity analysis, does that choice matter for the importance of the different processes?  
  
Fig 2. I don’t think people believed the effect are instantaneous, but sure they develop in a matter of weeks.  
Box Fig 1, I think a logarithmic timescale would allow us to better see the short-term majority of studies. Then you can also probably fit in the outlier study.  
Box Fig 2: what is N2? How does +/-5% randomly imply weaker microbial effects for non-focal process? All of them could even go up in some replicates right?  
  
  
  
  
Referee: 3  
  
Comments to the Author  
I read the review manuscript about temporal and demographic contexts of plant-soil microbe interactions. The review topic is very interesting, and I did like some of the ideas proposed by the authors to better integrate microbial-mediated impacts on feedback phase plant responses in PSF experiments. I, however, disagree with the central premise of this review that biomass-based approaches should be replaced by demographic measurements to understand microbial-mediated PSFs better. I further think that some of the experimental design ideas to incorporate demographic components of a plant rather complicate the PSF experiment approach and may even deviate from the key objectives of several PSF studies. Finally, the review on temporal dimensions of microbial effects (section II) is not convincing enough to shift from current PSF approaches to what the authors propose. I have detailed my comments in a pdf file (and below).  
  
Lines 7-8/lines 10-11: there is no such assumption in PSF literature that plant-microbial interactions follow a simple temporal trajectory. Since the history of PSF research, conditioning duration, type of soil inoculum (hence the microbes), life history of conditioning and feedback phase plants (and thus variation in plant development) have been, to the least, implicitly considered. I also disagree that there is any such implicit assumption in contemporary PSF studies that plant biomass ‘sufficiently’ captures the consequences of microbial impacts on plant growth patterns. With the possibility of microbial functional and taxonomic analyses, there are heaps of studies (e.g., Bezemer et al. 2006, Rigg et al. 2011, Hu et al. 2018, Wang et al. 2020, Friman et al. 2021, Steinauer et al. 2023: I can go on and on; coincidently, none of these is cited in the current review) directly relating microbial responses to plant responses during the response phases. I, therefore, am not convinced by the basic premise of this review piece.  
Line 36: Bever et al. 1997 is indeed one of the original citations of the PSF approach; however, there are several recent syntheses on PSF methodologies. It won’t harm to add some of those recent ones.  
Line 37: it’s too simplistic to say that plants modify the soil microbial communities. Plants modify a lot of various things in the soil, including microbial communities (sometimes those, even indirectly through root exudation and modifying physical properties in the soil).  
Lines 98-101: after reading most of the introduction, I am not sure if the authors have convinced a reader like me that the plant biomass approach should be replaced by plant demographic measurements. The authors need to provide a strong argument for how biomass approaches have failed to predict microbial-mediated PSFs. Moreover, being an empirical person, I started to wonder if the growth rate (and other demographic parameters) remains constant during the plant’s development at the conditioning phase. If one needs to establish a strong link between a plant’s ontogeny and the subsequent effect on soil microorganisms, it will require several stages of conditioning phases to capture the plant’s ontogeny, which I think gets extremely challenging in terms of logistics. Furthermore, different ontogenetic stages of a plant would differentially affect pathogen and mutualist microorganisms. How should one capture this?  
Lines 301-310: I am not entirely sure how microbial impacts on various demographic factors could not be reflected in plant biomass, particularly if both shoot and root biomass are considered, and along with some of the plant’s morphological traits are incorporated.  
Lines 330-395: after reading these paragraphs, my general feeling is that the authors are advocating the importance of seed survival and, more importantly, the successful germination of response plants in PSF experiments. It’s a fair point, and I agree that systems, where plant germination is highly constrained by soil-borne pathogens, should certainly consider germination success as the measure of the PSF effect. However, the number of seeds produced in natural setting could also play a role here, and having that many individual plants during the conditioning phase to obtain comparable seed numbers would pose another logistical challenge. Furthermore, where should those seeds come from? The same maternal plants used during the conditioning phase? This would  
then affect the duration of the conditioning phase if one were testing the PSF with perennial and/or  
long-living plants.  
Lines 397-406: many PSF studies run the response phase to capture soil microbial effects at the earlygrowth  
performance of plants. To my knowledge, this is usually 4-5 weeks, of course, depending on  
the life form of a plant (e.g., woody vs. herbaceous). It is, therefore, tricky to examine reproductive  
differences explained by the conditioning phase soil microorganisms in the test phase. The authors  
should consider that such early-growth phase performance (measured in biomass) when differs  
substantially between, let’s say, conspecific and heterospecific soils, it would then affect their longterm  
performance, such as their reproductive phenology, which may not even be dependent on soil  
microorganisms.  
Line 432: what is a site’s microbial legacy? How to quantify this? How are these independent of plant  
communities of the site?  
Lines 470-515: Regarding the use of competition models using demographic parameters of plants to  
predict microbial-mediated PSF, I wonder how much feasible this would be in PSF experiments where  
many microbes jointly affect the survival and/or performance of test-phase plants. I recon that  
finding a single or two pathogens (or mutualist), as in Mordecai 2013 paper, may not always be  
feasible in PSF experiments.  
Figure 1: I think the sequential harvesting approach makes PSF experiments quite challenging  
logistically. Yes, this might be feasible for providing model parameterization with a few focal plants  
and then extrapolating those model outputs with a larger set of plants. Although, I worry if that is  
really the purpose of PSF experiments. There is quite an understanding of what kind of plants exhibit  
stronger negative conspecific feedback and some of the underlying factors (Bezemer et al, 2006,  
Cortois et al. 2016, Hu et al. 2018), and I think the combined approach of theory and experiments,  
can help unravel how these findings are valid in the real world. The approaches proposed by this  
review, I doubt, complicate the simple and often powerful classical approach of PSF experiments.  
Figure 2: One of the key aspects of this review is that soil microbes change throughout the lifetime of  
a plant. Sure!! But this does not necessarily mean that all microbial changes have the same  
importance in terms of determining the strength of feedback. A key premise of PSF studies is to  
identify the net microbial effects (which are the result of several stages a plant goes through) that are  
likely to stay in the soil, arguably when the local soil is well homogenized/engineered by an individual  
or a plant population. This is one of the reasons why the conditioning phase, at least in several pot  
experiments, is carried out until the soil is well covered by the roots of the conditioning phase plant.  
In fact, what is rather more interesting, that the strength of feedback changes temporally as the testphase  
plant starts to steer the conditioned phased soils (e.g., Steinauer et al. 2023).