Review of Homeier and Leuschner Biogeosciences.

Factors controlling the productivity of tropical Andean forests: Climate and soil are more important than tree diversity

This paper is a very exciting summary of a large number of 400 m2 forest plots distributed over a large elevation gradient in the Andes. It addresses questions relating to productivity, species diversity and environment. It is rare for a dataset of this size to be assembled, and that alone justifies support for publication, after revision. The focus of the paper starts with a discussion of the diversity-function debate and then moves to a discussion on the determinants of low productivity at high elevation in tropical montane forests. Both subjects are dealt with usefully but I think both need some attention. The theoretical importance of the diversity-function debate is mentioned, but if this line of argument is to be maintained, I think it needs further justification. The debate has practical importance, but the theoretical importance, whilst it exists, receives a fair bit of skepticism in the literature. I don't think this point is central to the main strength of the paper, but I mention it, as the authors may choose to soften the stance on the theoretical importance or to support it more fully. Further on this point, I note that some large observational datasets are referred to, in order to address diversity-function relationships across biome types. I am supportive of this; all approaches to the diversity-function question have some unavoidable flaws in relation to this question but the experimental ones seem most prone to it. On the other hand some key experimental work in the tropics is not mentioned – surely it is of interest to place this discussion in the context of the Sabah Biodiversity Experiment – a large scale experiment in lowland tropical forest? There are few such comparisons in tropical forest to make use of and this feels like a gap.

On the question of causes/predictors of reduced productivity at high elevation, the comparative and interpretive analysis seems a little limited, even though the data are impressive and wide-ranging. The authors note the work of Fyllas 2017 (Ecology Letters), where a modelling approach was taken using annual estimates of GPP along a Peruvian (ie also in the Andes) elevation gradient as validation data. However, Homeier et al miss the first mechanistic modelling study of tropical montane forest productivity differences by elevation, presented by van der Weg et al. 2014 (Ecosystems). The 2014 paper validates model estimates of productivity using fine-scale mechanistically-related detailed sap flux data, whilst the 2017 paper validates mainly against impressive annual GPP estimates based on summed measurements of NPP and respiration. They come to different conclusions. Based on Amax values measured above 20 deg C, Fyllas et al highlight the importance of variation in Amax with elevation and variation in radiation with elevation, whilst van der Weg et al using Vcmax/Jmax, a stomatal model and measured leaf temperature (shown to be frequently well below 20 C in the high elevation site), conclude that variations in temperature and radiation are the most important drivers. The Fyllas conclusion is attractive as it suggests that despite high species turnover, overall Amax 'responds' by increasing at lower temperatures, suggesting a degree of 'optimisation to environment', filtered by species turnover. On the other hand, the mechanistic validation in the van der Weg paper, and its use of real leaf temperatures well below 20 C suggests a key role for temperature not strongly evident in the Fyllas analysis. The paper here (Homeier et al) could contribute strongly to this overall discussion with independent data and analysis. Whilst the density of measurements is not the same as presented in Fyllas, and there is no modelling (which is not necessarily a problem), there is very detailed edaphic information, as well as productivity and information on species identity. It is also clear that there is a huge range in productivity at each elevation among the different forest plots. This may be noise related (smallish plots…but lots of them!) or it may be environmental; I note that the data Fyllas et al use for Amax values suggest a wide range of photosynthetic capacities at each elevation….ie a similar pattern as found here, of much variation at each elevation. In sum, this paper has the potential to make a bigger contribution to this discussion than it currently does. I hope these comments are of use; it is worth expending effort on in a revision because the question is about fundamental tropical ecology (or indeed montane-to-lowland ecology), and had remained in the realm of 'many explanatory factors but we don't know which' until relatively recently.

Detailed points

Line 28. What about reference to the Sabah Biodiv Expt? Can you make more of the comparison of the diversity-function relationship at high vs low diversity? Also, is it useful to refer to Sullivan 2016 Scientific Reps (biomass and diversity…very slightly different question as biomass is not 'function' but it does discuss plot size)

Line 37-45. Missing the van der Weg 2014, which used data and modelling in early trop montane forest data+modelling analysis. It shows temp and radiation dependence mechanistically, and water use is validated using sap flux data. It also demonstrates importance of low leaf temps affecting function. The analysis in these lines mentions some key comparative flux and parameter data (eg Girardin 2014 Malhi 2017), but omits the Fyllas 2017 paper mentioned above. A bigger discussion is needed somewhere here to set this paper up more comprehensively.

Line 47. I note the use of the Chisholm 2013 reference, but I think this analysis/discussion needs to take into account dynamics too, if only briefly. ABG does always reflect productivity - see Baker 2004, Galbraith et al. 2013, Malhi et al. 2015); residence time is important, as is recruitment. So the point made here needs to be made in relation to this wider discussion on determinants of ABG.

Line 59. The choice and effectiveness of small plots needs a fuller discussion than reference only to Chisholm. For example, if diversity effects are only likely to emerge at small scale, what does this mean for their fundamental role?

Line 135. It's great to see a careful path analysis approach being taken here to distinguish different drivers. However, can you explain how the original model structure affected the ultimate outcomes? Might you have had a different outcome had your starting point (structure) been different or have your methods fully accounted for this? (apologies if I've missed this point).

Line 160. Are there any recruitment data to advance the C dynamics analysis?

Line 164. LAI measurements are really important but difficult to make. Are you sure the differences you see in LAI are related to leaf area and not a change in stem density/canopy structure? High stem density would increase Plant Area Index (ie leaves and wood) even if LAI did not increase. I know this is hard to separate, but some comment/discussion/caveat would be useful.

Fig 2. There are strong signals of variation in mean values with elevation in some of the key metrics (eg WP, Stem density). But there is also very large variation at each elevation. is this discussed? The variation by elevation is larger than the overall mean signal in the regression; this has also been observed in ecophys measurements elsewhere (eg Bahar et al. 2016 New Phytologist).

Line 193. As per the second paragraph above, the Fyllas 2017 paper notes that there is large turnover in species but argues that there are directional changes in mean trait values with elevation and these become determining of productivity along with radiation....how can we link these different findings?

Line 237. The role of low leaf temperatures needs further consideration in affecting rates of carbon gain, not just radiation levels.

Line 243. This soils dataset is very substantive and provides detailed driver information for the path analysis. It may be possible to use this to help the contrast with or discussion of preceding analyses on this general productivity/elevation subject.

Line 250. Do you have soil respiration data or root productivity data to back this up (ie higher allocation of C to root production/symbionts)?

Line 265. I wonder if an analysis of productivity vs biomass would help here too – ie productivity does not determine biomass in all circumstances because of other fluxes/processes affecting C residence time.

Line 268, it seems natural to consider a comparison with the effects of fertlisation in the Andes reported by Fisher et al. 2013, Oecologia, as well as this 1989 reference.

Line 274 – as before, please consider LAI vs PAI differences, causative factors.

Line 280. I wonder whether this section could be given a bit more depth by including a discussion on the relationship between trait diversity and species diversity? Might we expect a stronger response at low species diversity because trait diversity may increase rapidly as you add species at first, but if trait diversity is ultimately lower than species diversity we might expect the function-diversity graph to saturate more quickly using traits? Also of course there is the wider discussion on how the relationship (with species) varies under harsh and less harsh environments (eg. Paquette et al. 2011, Glob Ecol and Biogeog).

Line 291. Again, might some discussion on traits be useful here too?