Comments and responses VG and IB

**Introduction**

Moved pollination paragraph up

Model structure paragraph – change focus to forecasting – why LME good – for overlap / pseudo-replication

PGLS = body size definitively linked to phylogeny

**Method**

Pre-existing equations

* I agree with moving to methods and appendix – extending analyses to body length will take too long and make the paper long-winded
* we don’t have body length from USA or Switzerland or Belgium

Specimen collection

Add average per species, range, also range in specimen age – we have collection dates but not measurement dates so mode is difficult to calculate. The reason older specimens were included was because I decided, after discussion with two different people, that there shouldn’t be an influence on weight for properly curated specimens.

Model structure

Get feedback on how to best present model formula and if needed

-fix random factors

*Data analysis: Incorporating phylogeny*

NB: Why? I guess is to avoid the random factor species, but this makes models hard to compare with LME. Maybe it’s ok,

* I took out sex as I couldn’t code a PGLS that could include sex – i.e. with two tips per species – if this is doable, lets do it

NB: If only this two species appear twice, maybe the LME can be used without species as random and also exclude this two? Sex may be even kept there… just a though.

* This is a poor explanation on my part – for LME I averaged per country so we have species overlap within Europe i.e. Bombus from Britain/Ireland/Germany/Spain
* When averaged by region alone, only honeybee and Halictus rubicundus are present in more than one region – one naturally (Halictus), one unnaturally

VG: Why? How would leaving them affect the models?

I couldn’t formulate a PGLS with multiple representatives of each species – perhaps this is an oversight on my behalf

*Data analysis: Cross-validation*

NB: So you just use 1/5 of the data for training? This is too low for your sample size. I’ll do 4/5 training, 1/5 testing and permutate that…

It was 4/5 training and 1/5 testing – ive tried to make it clearer in the MS

NB: CV r2 with LME and random factors: I followed this <http://rpubs.com/ledongnhatnam/241926>

Why not the other equations from the literature!?

I chose to just use IT as I thought including body length as well would overcomplicate the manuscript and we also don’t have body length from everywhere i.e. USA, Switzerland

Maybe we need a small bit about how IT is better performing than BL??

Correct me if im wrong but I don’t think you need the phylogeny to run predict on a PGLS object, i.e. LME and PGLS are as practical as one another