Oikos

OIK-07022

Cid, B., Carbone, C., Fernandez, F. A. S., Jansen, P. A., Rowcliffe, J. M., O'Brien, T., Akampurira, E., Bitariho, R., Espinosa, S., Gajapersad, K., Santos, T. M. R., Gonçalves, A. L. S., Kinnaird, M. F., Lima, M. G. M., Martin, E., Mugerwa, B., Rovero, F., Salvador, J., Santos, F., Spironello, W. R., Wijntuin, S. and Oliveira-Santos, L. G. R. 2020. On the scaling of activity in tropical forest mammals. - Oikos doi: 10.1111/oik.07022

Appendix 1

Potential bias in activity range estimation

Camera traps can show bias favoring data on larger animals (Burton et al. 2015) because of increased detection range and day range (Rowcliffe et al. 2014, 2016). If there was bias favoring the capture of larger animals in our dataset, it could also bias our analysis and results. To check if this was a potential problem in our analysis, we 1) calculated the distribution of body size classes among the mammalian species analysed in this study; 2) used a linear model to understand the relationship between body mass and the number of independent records for the 249 populations in our dataset; 3) simulated the sensibility of the activity range estimation to the number of independent records, using five species with different activity patterns as models; and 4) included the number of independent records in the model to understand it influence as a covariate.

We have a balanced representation of different body size classes in our dataset: of 128 species analysed in this study, we had 20% of small species (< 1 kg), 44% medium-sized species (> 1 kg and < 10 kg) and 36% large species (> 10 kg).

The linear model showed no relationship between body mass and the number of independent records in a natural logarithm scale ($F_{1,247}$ = 0.337; r = 0.001; p = 0.561; Figure A1.1), evidencing

that there is no sampling bias towards larger animals associated with the use of camera traps, at least in our case.

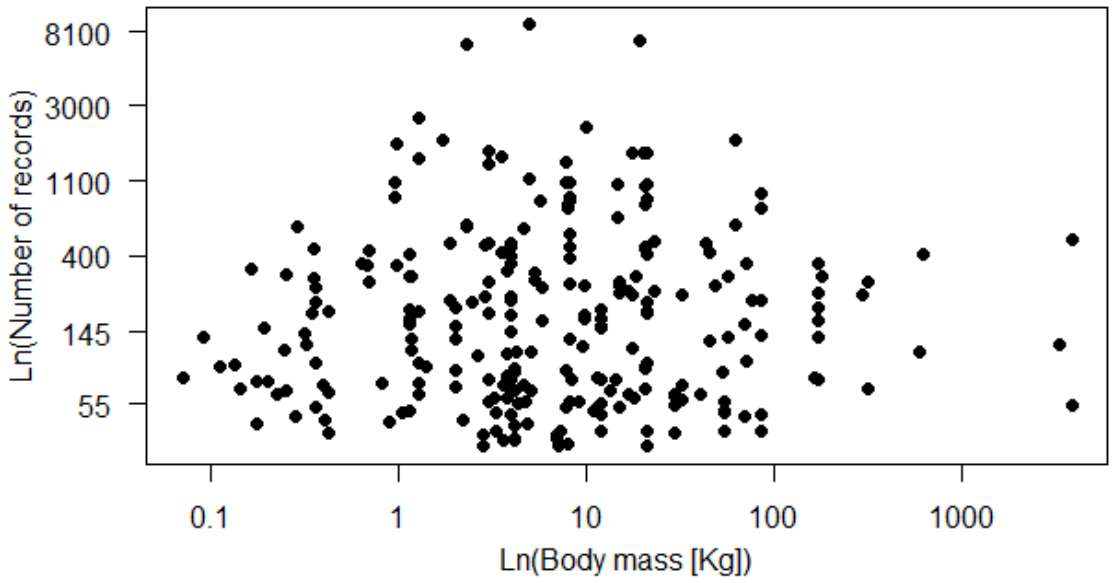
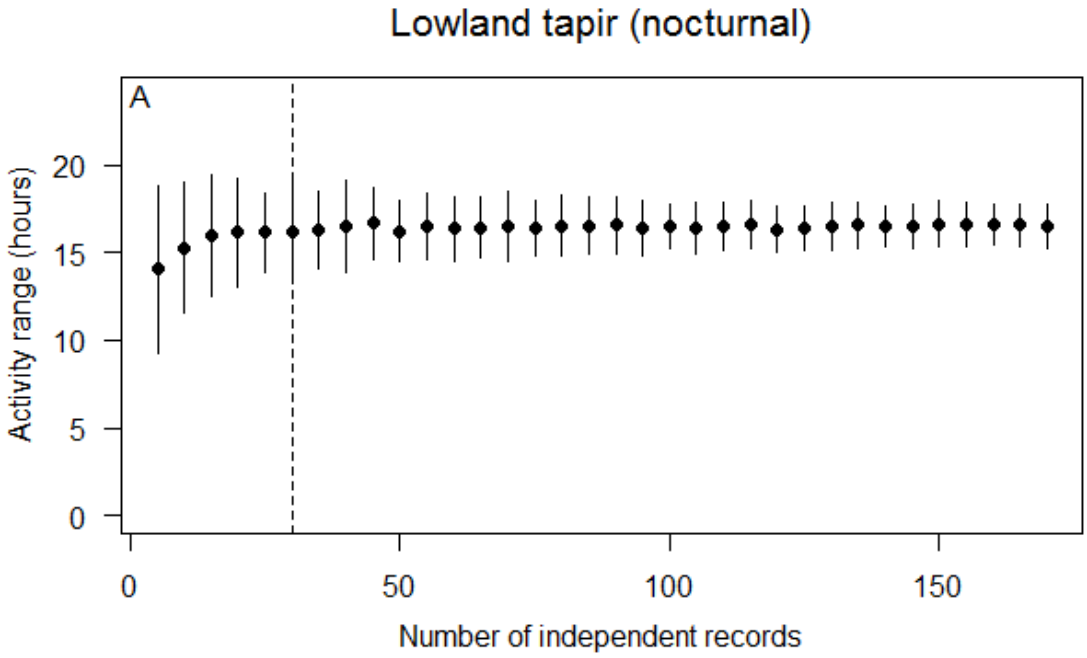
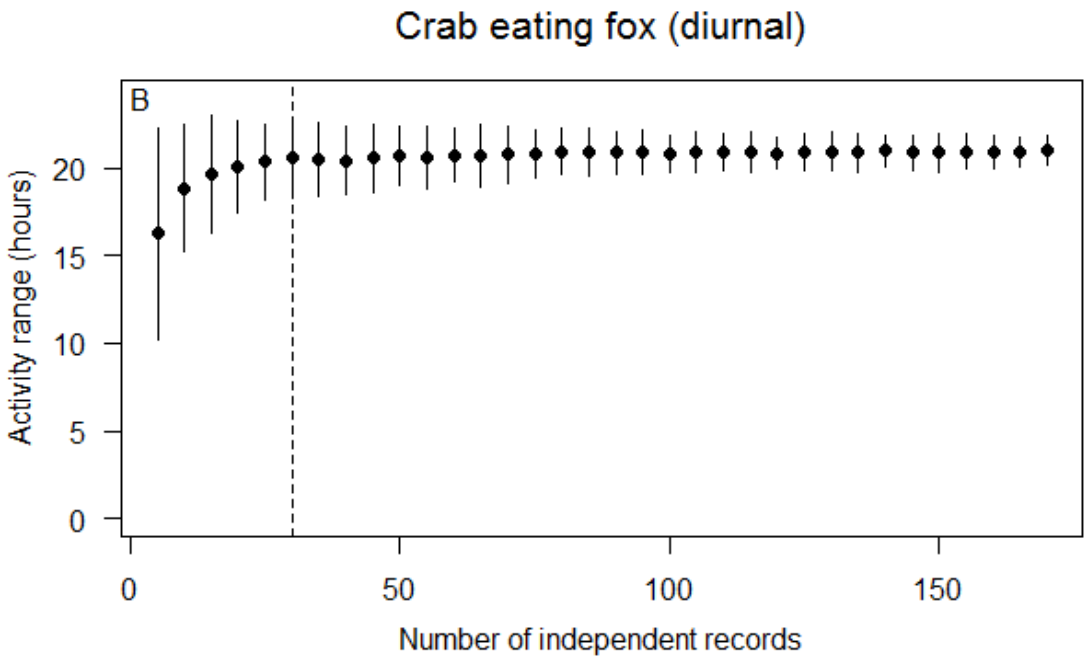


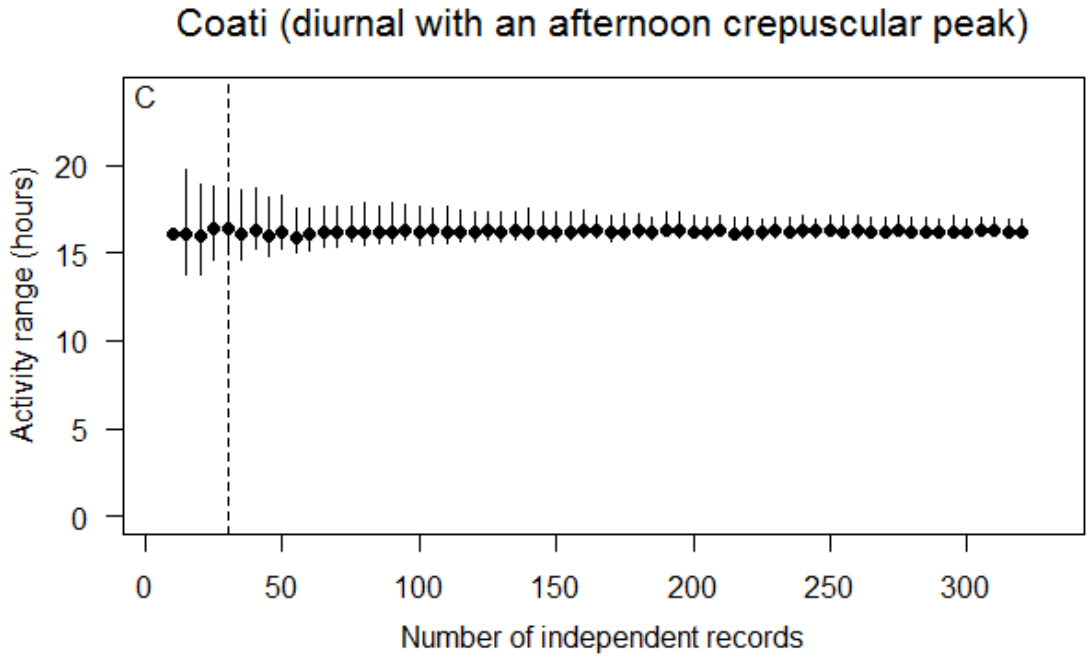
Figure A1.1. Overall natural log-log relationship between body mass and the number of independent records for the 249 populations analyzed in this study. Both axes are in natural log scale but real values were printed.

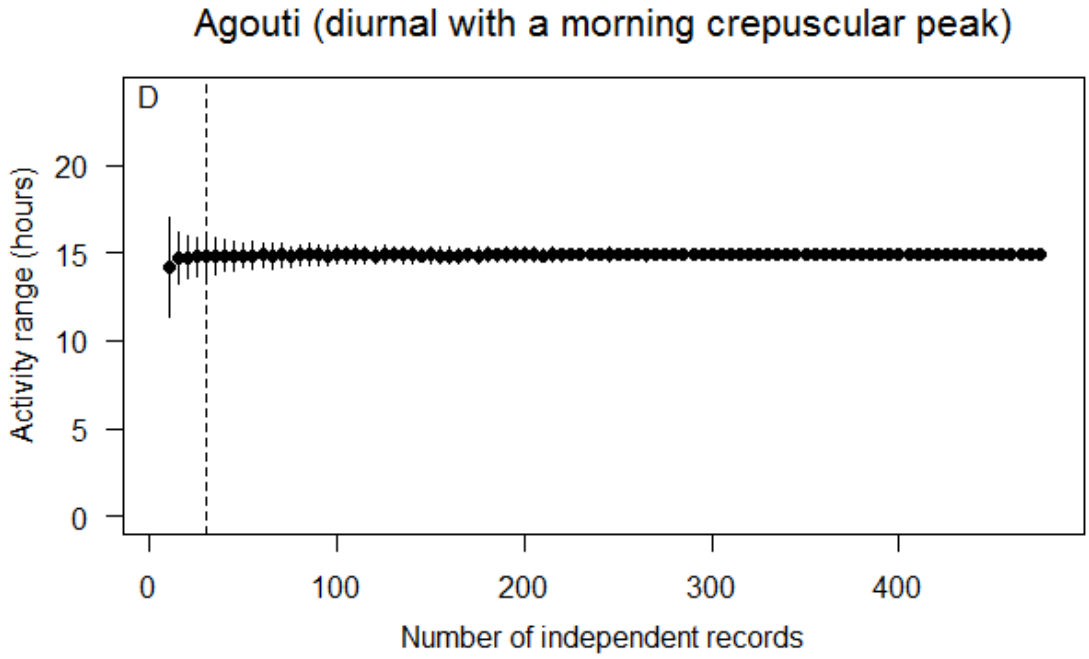
To understand if the activity range value was sensible to the number of independent records used to estimate it, we performed a simulation. We selected five mammal species, each one with a different activity pattern as models: the lowland tapir (nocturnal), the crab eating fox (diurnal), the coati (diurnal with an afternoon crepuscular peak), the agouti (diurnal with a morning crepuscular peak) and the collared peccary (cathemeral). To perform the simulation, the number of records for each species was trimmed to meet the minimum of 1-hr interval between consecutive records of the same species at the same camera site to reduce temporal dependence (Gómez et al. 2005). Then, for each species, we sub-sampled (with replacement) the number of independent records by progressive interval classes of five records (data n = 5, 10, 15, …, n total). For each n, we estimated the activity

range 100 times with replacement (using a bootstrapping technique; Figure A1.2. A-E).






Coati (diurnal with an afternoon crepuscular peak)
C
Activity range (hours)
0
5
10
15
20
0
50
100
150
200
250
300
Number of independent records


Agouti (diurnal with a morning crepuscular peak)
D
Activity range (hours)
0
5
10
15
20
0
100
200
300
400
Number of independent records

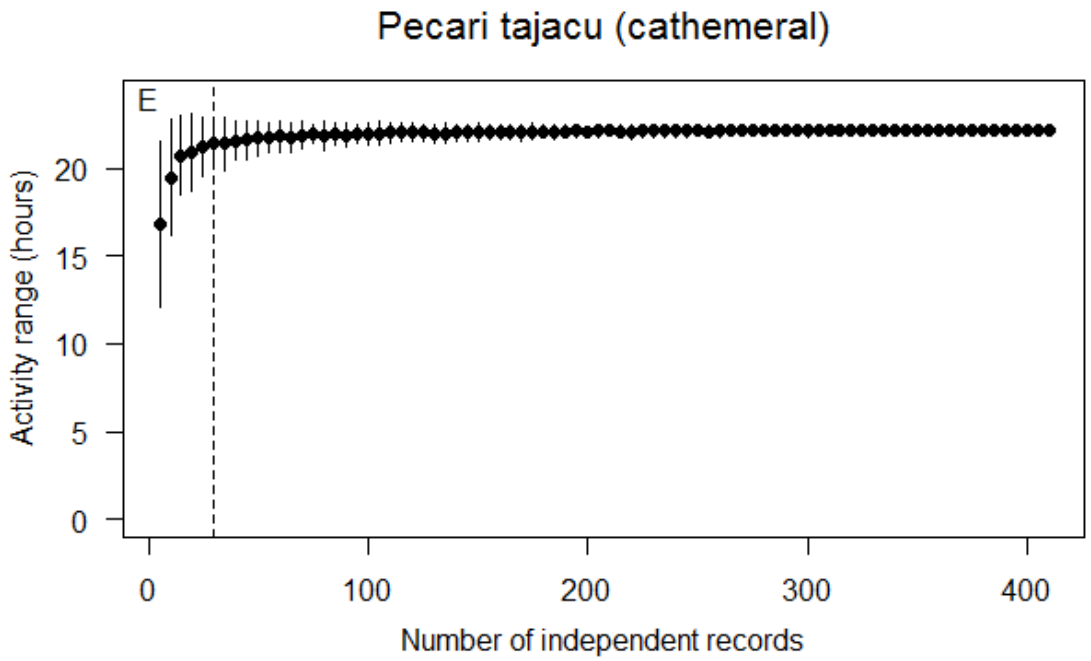


Figure A2.1. Simulation showing the sensibility of the activity range estimation to the number of independent records used to estimate it. The black dots indicate the mean value for the activity range estimation over 100 simulations for each interval of five photographic records (increasing number of records: 5, 10, 15, 20, …, total). The vertical black bars are the 95% confidence interval generated by the 100 simulation with replacement for each interval of five records. The vertical dashed lines show our established minimum (30 independent records). (A) lowland tapir (nocturnal), (B) crab eating fox (diurnal), (C) coati (diurnal with an afternoon crepuscular peak), (D) agouti (diurnal with a morning crepuscular peak) and (E) collared peccary (cathemeral).

The simulations showed that the activity range estimation was robust for sampling sizes above 30 independent records, our stablished minimum, regardless of the activity pattern (e.g., unimodal [diurnal, nocturnal], bimodal [crepuscular] or multimodal [cathemeral]). This robustness is a characteristic of the use of the kernel density to estimate the activity range, which relies on density probability of a random sampling from a biological process (Oliveira-Santos et al. 2013).

Based on this analyses and simulations, we conclude that, at least in our case, the camera traps produced reliable and unbiased estimates, allowing cross-species and cross-site analyses, provided that the minimum of 30 independent records was respected.

Finally, even if significant variance in activity range was explained by the number of records that a species has (what it is not true), we could account for this bias by including the number of records (*n*) by species as covariate at the model. This procedure would still allow us to produce reliable estimates of the effect of body mass on activity range, even in the presence of bias. As exercise, we added this covariate to the model: log(activiy)~diet * log(mass)+***n***+phylogentic control]. The covariate had no significant effect, (beta = -0.00002), which means that we need a difference of 100 thousand records to produce a spurious increase or decrease of two hours in the activity range estimation.

Appendix 2

Detailed activity range results

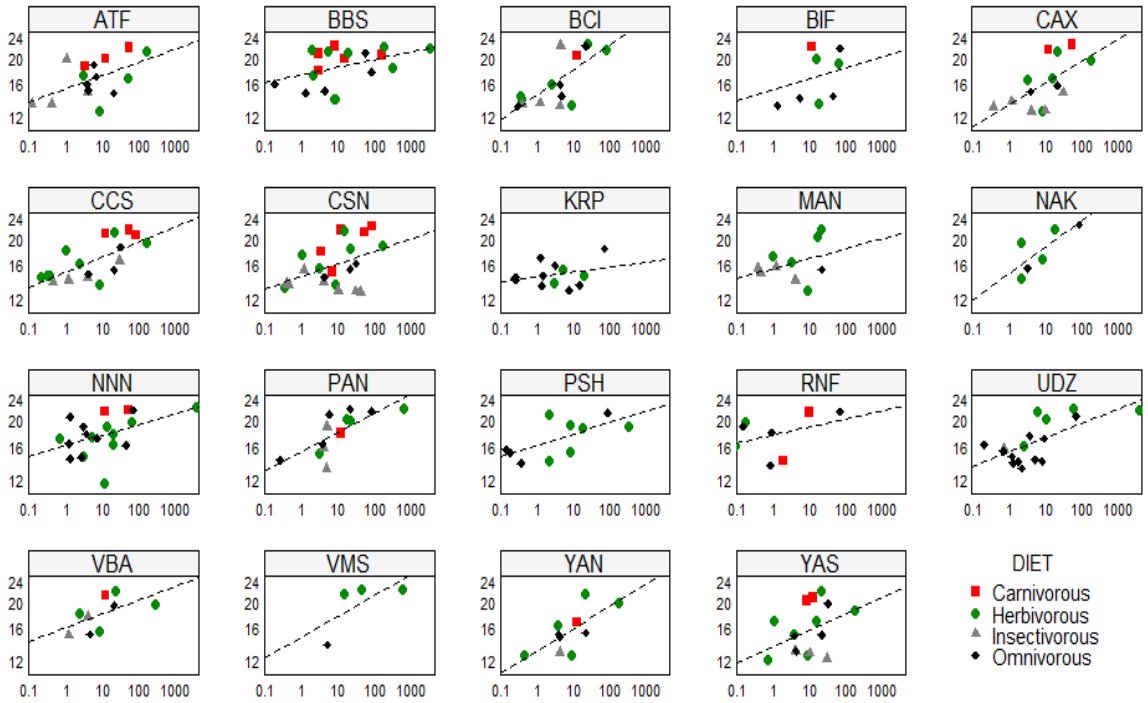


Figure A2.1. Relationship between activity range (y axis) and body mass (x axis) for 249 mammal populations by study site. The activity range is described in hours. Body mass is described in Kilograms. Dot colour is a guide to diet: carnivores, red; herbivores, green; insectivores, grey; and omnivores, black. Africa: BIF = Bwindi Impenetrable National Park (Uganda), KRP = Korup National Park (Cameroon), NNN = Nouabalé-Ndoki National Park (Democratic Republic of Congo), RNF = Ranomafana National Park (Madagascar), UDZ = Udzungwa Mountains (Tanzania), VMS = Virunga Massif (Republic of Congo); Asia: BBS = Bukit Barisan (Indonesia), NAK = Nam Kading National Protected Area (Laos), PSH = Pasoh Forest Reserve (Malaysia); Central America: BCI = Barro Colorado Nature Monument and Soberanía National Park (Panama), VBA = Volcán Barva (Costa Rica); South America: ATF = Santa Catarina Atlantic Forest (Brazil), CAX = Caxiuanã National Forest (Brazil), CCS = Cocha Cashu-Manu National Park (Peru), CSN = Central Suriname Nature Reserve (Suriname), MAN = Manaus (Brazil), PAN = Pantanal (Brazil), YAN = Yanachaga-Chemillén National Park (Peru), YAS = Yasuni National Park (Ecuador).

Appendix 3

Activity level results

The activity level showed higher amplitude than the activity range, varying five-fold across species, from 3.8 hours for the red acouchi *Myoprocta acouchi* in Central Suriname Nature Reserve (Suriname) to 19.2 hours for the Central American Red Brocket (*Mazama temama*) in Barro Colorado Nature Monument (Panama). The first 67 axes explained 95% of variation in phylogenetic distance. After sequential inclusion, one axis was enough to control phylogenetic autocorrelation of model residuals (Moran's I $p = 0.021$ before inclusion; $p = 0.319$ after inclusion).

There was a positive scaling relationship between activity level and body mass and this relationship differed among diet categories (Figure A3.1a). Carnivore activity level scaled with body mass by a factor 0.097 (CI = 0.027 – 0.167; $p = 0.007$), while in the other three diet categories it did not scale significantly with body mass, as the 95% confidence intervals of the slope estimates included zero (Figure A3.1b). Intercept estimates for each diet group did not differed significantly (Figure A3.1c). Our model accounted for 40.5% of the variation in activity level, where phylogeny (PVR axis) explained 4.3% and ecological variables explained 36.2% (body mass 27.9%, diet 4.8% and interaction body mass:diet 3.4%; Table A3.2). The variance of the random intercepts by site ($\sigma^2 = 0.001$) was about 45 times lower than the residual variance ($\sigma^2 = 0.045$), meaning that variation explained by site differences was negligible.

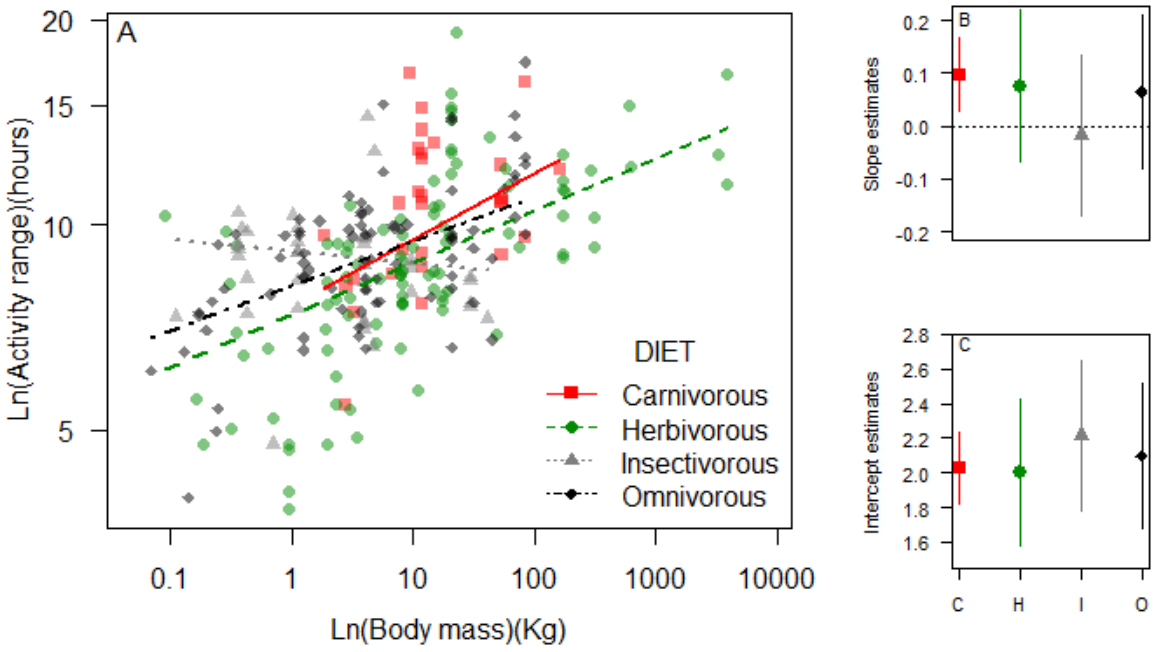


Figure A3.1. Scaling of activity level with body mass in 249 populations of terrestrial mammals across 19 tropical forests. (a) Overall log-log relationship between activity level and body mass. Dashed lines are GLMM fits to the data. Both axes are in natural log scale but real values were printed. (b) Slope estimates. Dots indicate the estimated values and lines the 95% confidence intervals. (c)Intercept estimates. Dots indicate the estimated values and lines the 95% confidence intervals. The dashed black line marks the zero. Colour is a guide to diet: carnivores, red; herbivores, green; insectivores, grey; and omnivores, black.

Table S3.1. Model (GLMM-phy) estimates. Estimate = mean estimates. Lower and upper = lower and upper limits of the 95% confidence intervals. % explained variance = activity level variance explained by each component (phylogeny [C1] and ecological variables [body mass, diet and interaction body mass:diet]).

Variable	Estimate	Lower	Upper	% explained variance
Intercept	2.029	1.823	2.234	-
Log(mass)	0.097	0.027	0.167	27.9
Diet herbivorous	-0.028	-0.247	0.190	4.8
Diet insectivorous	0.184	-0.041	0.410	
Diet omnivorous	0.067	-0.147	0.281	
C1	0.805	0.409	1.200	4.3
Diet herbivorous:log(mass)	-0.021	-0.094	0.052	3.4
Diet insectivorous:log(mass)	-0.114	-0.198	-0.030	
Diet omnivorous:log(mass)	-0.032	-0.106	0.042	

Appendix 4
R-code

```
library(nlme); library(bbmle); library(ape); library(MuMIn); library(maps); library(car);
library(plotrix); library(adephylo); library(phylobase); library(PVR); library(png); library(lme4);
library(circular); library(LMERConvenienceFunctions)

###############
## getting data ##
###############

mam1=read.table("mammals(atlforest.to.nouab).csv", T, sep=",")
mam2=read.table("mammals(pant.to.yas).csv", T, sep=",")
mam1$sp=paste(mam1$genus, ".", mam1$species, sep="")
mam2$sp=paste(mam2$genus, ".", mam2$species, sep="")
data=rbind(mam1, mam2)
data=data[data$sp!="Homo.sapiens",]
data=data[data$sp!="Domesticated.cow",] #domestic
data=data[data$sp!="Domesticated.dog",] #domestic
data=data[data$sp!="Bos.primigenius",]  #extict
data1=data[data$site!="Atlantic.Forest" & data$site!="Pantanal",]
data2=data[data$site=="Atlantic.Forest" | data$site=="Pantanal",]
datetime=as.POSIXlt(paste(data1$date,data1$hour),format="%m/%d/%Y %H:%M:%S")
nrow(data1)
block=paste(data1$site,data1$sp,data1$camera,datetime$mday,datetime$mon,datetime$year,dateti
me$hour)
data1=data1[!duplicated(block),]
nrow(data1)
data2=rbind(data1,data2)
nrow(data2)
data2$site=as.factor(data2$site)
data2$sp=as.factor(data2$sp)

#########################
## estimating activity range ##
#########################

names(table(data2$sp))[table(data2$sp)>30]
d4=as.POSIXlt(data2$hour,format="%H:%M:%S")
n=as.data.frame(table(paste(data2$sp,data2$site)))
species=as.vector(n[which(n[,2]>30),1])
ar=matrix(NA,length(species),1)
rownames(ar)=species
for (i in 1:length(species)){
time=circular(d4[which(paste(data2$sp,data2$site)==species[i])]$hour, units='hours')
res=modal.region.circular(x=time,q=0.95,bw=5)
a=numeric()
for(j in 1:nrow(res$zeros)){
a[j]=as.numeric(res$zeros[j,2])-as.numeric(res$zeros[j,1])}
ar[i]=sum(a)

###################
```

```
## managing results ##
####################

sp=unlist(strsplit(species," "))[seq(1,length(unlist(strsplit(species," "))),2)]
site=unlist(strsplit(species," "))[seq(2,length(unlist(strsplit(species," "))),2)]
order=data.frame(unique(data2[,c("sp","order")]));order=order[!is.na(order$order),]
ar1=data.frame(site,sp,n=n[n[,2]>30,2],ar=ar)
mass=read.table("mass_jones.txt",T) #PanTHERIA
ar1=merge(ar1,mass,"sp",all.x=T)
ar1=merge(ar1,order,"sp",all.x=T,all.y=F)
diet=read.table("ar.txt",T)[,c("sp","diet")] #PanTHERIA
ar1=merge(ar1,diet,"sp", all.x=T)
ar1$mass=ar1$mass/1000

## mean by species ##

msp=aggregate(ar1$ar,list(ar1$sp),mean)
colnames(msp)=c("sp","ar")
msp=merge(msp,mass,"sp", all.x=T) #PanTHERIA
msp=merge(msp, diet, "sp", all.x=T) #PanTHERIA
sdsp=aggregate(ar1$ar,list(ar1$sp),sd)
msp$mass=msp$mass/1000

## basic results ##

n=numeric()
for(a in 1:length(unique(ar1$site))){
  n[a]=length(ar1$sp[ar1$site==as.character(unique(ar1$site)[a])])
}
hist(n, breaks=20); median(n); min(n); max(n)

#################################
## extracting phylogenetic distances ##
#################################

tree1=read.nexus ("FritzTree.rs200k.100trees.tre") #Fritz et al. (2009)
bphylo=tree1[[1]]
sp.split=strsplit (as.character(ar1$sp), ".", fixed=T)
sp.tree=as.character()
for(d in 1:length(ar1$sp)){
sp.tree[d]=paste(sp.split[[d]][1], "_", sp.split[[d]][2], sep="")}
ar1$sp.tree=sp.tree

trouble=ar1$sp.tree[!ar1$sp.tree %in% bphylo$tip.label]
unique(trouble)

## changing species names to closely related species that are in the phylotree ##

sp.tree1=gsub("Cephalophus_harveyi", "Cephalophus_natalensis", ar1$sp.tree)
sp.tree1=gsub("Cercocebus_sanjei", "Cercocebus_galeritus", sp.tree1)
sp.tree1=gsub("Dendrohyrax_validus", "Dendrohyrax_dorsalis", sp.tree1)
sp.tree1=gsub("Dendrohyrax_validus", "Dendrohyrax_dorsalis", sp.tree1)
```

```
sp.tree1=gsub("Mazama_nemorivaga", "Mazama_bricenii", sp.tree1)
sp.tree1=gsub("Rhynchocyon_udzungwensis", "Rhynchocyon_cirnei", sp.tree1)
sp.tree1=gsub("Thrichomys_pachyurus", "Thrichomys_apereoides", sp.tree1)
ar1$sp.tree1=sp.tree1
trouble=ar1$sp.tree1[!ar1$sp.tree1 %in% bphylo$tip.label]
unique(trouble)

## cutting.tree ##
sp.cut=bphylo$tip.label[!bphylo$tip.label %in% ar1$sp.tree1]
tree2=drop.tip(bphylo, sp.cut)

################
## mixed model ##
################

lmePCZ=lme(log(ar)~diet*log(mass), random=~1|site, data=ar1)

## checking phylogenetic autocorrelation in the model's residuals ##
z=1/cophenetic(tree2)
diag(z)=0
z[is.infinite(z)]=0
res=residuals(lmePCZ)
res.df=data.frame(ar1$sp.tree1, as.numeric(residuals(lmePCZ)))
names(res.df)=c("sp", "res")
res1=aggregate(res.df$res, list(res.df$sp), mean)
names(res1)=c("sp", "res")
res.tree=rep(NA, nrow(res1))
for(e in 1:nrow(res1))
{
  line=which(tree2$tip.label==as.character(res1$sp[e]))
  res=res1$res[e]
  res.tree[line]=res
}
res2=data.frame(tree2$tip.label, res.tree)
Moran.I(res2$res, z)

#################
### PVR analysis ##
#################

library(bootStepAIC)
library(PVR)
msp$sp.tree1=unique(ar1$sp.tree1)
ar.tree2=msp$ar[match(tree2$tip.label, msp$sp.tree1)]
mass.tree2=msp$mass[match(tree2$tip.label, msp$sp.tree1)]
pvr=PVR(x=PVRdecomp(tree2),phy=tree2,trait=ar.tree2,envVar=mass.tree2,method="moran")
pcoapvr=PVRdecomp(tree2)
cumsum(pcoapvr@Eigen$values/sum(pcoapvr@Eigen$values))
sp.pvr=data.frame(pcoapvr@Eigen$vectors, pcoapvr@phylo$tip.label)
names(sp.pvr)[128]="sp"
sp.pvr$sp=as.character(sp.pvr$sp)
sp.tot=data.frame(ar1$sp.tree1)
```

```
names(sp.tot)="sp"
axis=merge(sp.tot, sp.pvr, "sp", all.x=T)
a=data.frame(ar1[,1:10],axis[,2:68])
pvrs=paste("c",1:67,sep="")
mod=list()
f=formula(log(ar)~+log(mass)*diet)
mod[[1]]=lme(f,random=~1|site,data=a)
df=data.frame(mod[[1]]$data$sp.tree1, as.numeric(residuals(mod[[1]])))
names(df)=c("sp", "res")
df1=aggregate(df$res, list(df$sp), mean)
names(df1)=c("sp", "res")
res.tree1=rep(NA, nrow(df1))
for(g in 1:nrow(df1)){
line=which(tree2$tip.label==as.character(df1$sp[g]))
res=df1$res[g]
res.tree1[line]=res}
df2=data.frame(tree2$tip.label, res.tree)
names(df2)=c("sp", "res")
p=Moran.I(df2$res, z)$p.value

pvrs=paste("c",1:67,sep="")
mod=list()
f=formula(log(ar)~+log(mass)*diet)
mod[[1]]=lme(f,random=~1|site,data=a)
p=numeric()
for(i in 1:length(pvrs)){
nex=paste("~ . + ", pvrs[i],sep = "")
f=update(f,nex)
mod[[i+1]]=lme(f,random=~1|site,data=a)
df=data.frame(mod[[i+1]]$data$sp.tree1, as.numeric(residuals(mod[[i+1]])))
names(df)=c("sp", "res")
df1=aggregate(df$res, list(df$sp), mean)
names(df1)=c("sp", "res")
res.tree1=rep(NA, nrow(df1))
for(g in 1:nrow(df1)){
line=which(tree2$tip.label==as.character(df1$sp[g]))
res=df1$res[g]
res.tree1[line]=res}
df2=data.frame(tree2$tip.label, res.tree1)
names(df2)=c("sp", "res")
p[i]=Moran.I(df2$res, z)$p.value}
plot(p, pch=19)
abline(h=0.05)
Moran.I(as.numeric(residuals(mod[[5]])), z)
summary(mod[[5]])
intervals(mod[[5]])
r.squaredGLMM(mod[[5]])

## table of estimated parameters values (95% confidence intervals) ##

par.df=as.data.frame(intervals(mod[[5]])$fixed)
```

```
coefs=data.frame(ic=as.numeric(par.df[1,]),ih=as.numeric(par.df[1,])+as.numeric(par.df[3,]),ii=as.n
umeric(par.df[1,])+as.numeric(par.df[4,]),io=as.numeric(par.df[1,])+as.numeric(par.df[5,]),sc=as.nu
meric(par.df[2,]),sh=as.numeric(par.df[2,])+as.numeric(par.df[10,]),si=as.numeric(par.df[2,])+as.nu
meric(par.df[11,]),so=as.numeric(par.df[2,])+as.numeric(par.df[12,]))
coefs=data.frame(t(coefs))
colnames(coefs)=c("low","est","upp")
coefs

## the same but using lme4 that allows to estimate the variance partitioning ##
model=lmer(log(ar)~diet*log(mass)+c1+c2+c3+c4+(1|site),data=a)
summary(model)
par=pamer.fnc(model)
confint(model)
r2=sum(par[ncol(par)])
r2phy=sum(par[ncol(par)][paste("c",1:4,sep=""),])
r2eco=r2-r2phy
```

REREERENCES

Burton, A. C. et al. 2015. Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. – J. Appl. Ecol. 52: 675–685.

Gómez, H. et al. 2005. Dry season activity periods of some Amazonian mammals. – Stud. Neotrop. Fauna E. 40: 91–95.

Oliveira-Santos, L. G. R. et al. 2013. Using conditional circular kernel density functions to test hypotheses on animal circadian activity. – Anim. Behav. 85: 269–280.

Rowcliffe, M. J. et al. 2014. Quantifying levels of animal activity using camera trap data. – Methods Ecol. Evol. 5: 1170-1179.

Rowcliffe, M. J. et al. 2016. Wildlife speed cameras: Measuring animal travel speed and day range using camera traps. – Remote Sens. Ecol. Conserv. 2: 84–94.