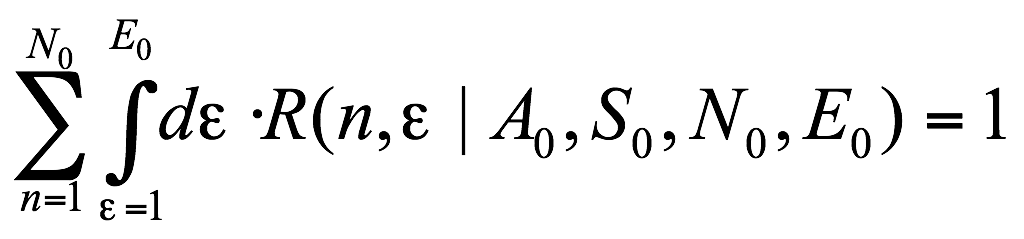
**SUPPORTING INFORMATION**

**Appendix S1.** Derivation of the distribution of metabolic rates across individuals.

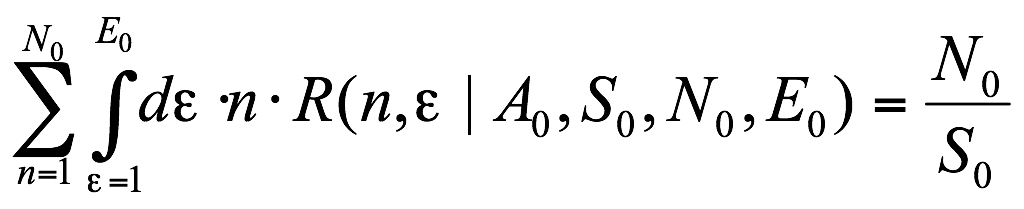
“MaxEnt” is an inference procedure that yields least-biased probability distributions that are constrained by prior knowledge. The core theoretical construct in the Maximum Entropy Theory of Ecology (METE) from which the distribution of metabolic rates is derived is the ecological structure function, *R*(*n*,*ε*|*S*0, *N*0, *E*0), a joint, conditional probability distribution. *R* describes how individuals are distributed over species and how metabolism is distributed over individuals. The meaning of *n* and *ε* is evident from the definition of *R*. In particular, *R∙dε* is the probability that if a species is picked at random from the species pool, then it has abundance *n*, and if an individual is picked at random from that selected species, then its metabolic rate is in the interval (*ε*, *ε*+d*ε*).

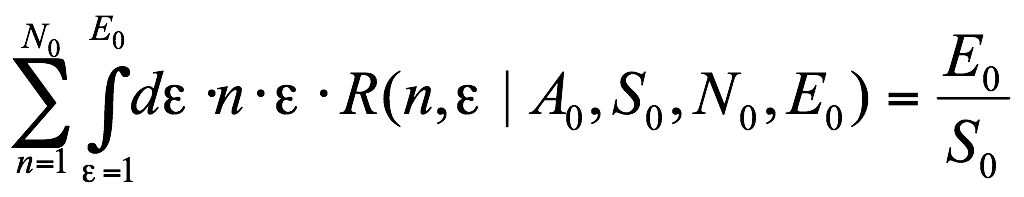
*S*0, *N*0, *E*0 are the state variables whose ratios provide the constraints on *R*. *S*0 is the total number of species (from within any chosen taxonomic category) in a community, *N*0 is the total number of individuals in all those species, and *E*0 is the total metabolic rate of all those individuals. Note that *ε* and *E*0 are rates of energy use, and are therefore measures of power, not energy. By choice of units, we define the minimum metabolic rate of all the individuals to be 1. Note, further, that *n* is a discrete variable, while *ε* is continuous.

The normalization condition on *R* is:

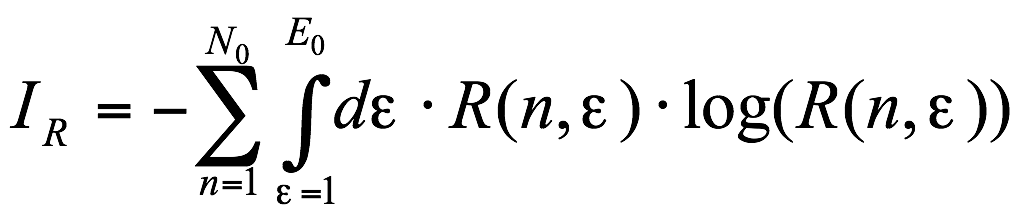
 (S1)

Two other constraints on the ecosystem structure function result from the two independent ratios formed from the state variables: *N*0/*S*0 is the average abundance per species and *E*0/*S*0 is the average over species of the total metabolic rates of the individuals within the species. This gives us the following two constraint equations on *R*:

 (S2)

 (S3)

The distribution *R*(*n*,*ε*) can now be obtained by maximizing the Shannon information entropy of that distribution:

 (S4)

subject to the constraints of Eqs. S1 - S3.

Using the method of Lagrange multipliers, the MaxEnt solution (Harte et al., 2008) for the structure function *R*(*n*,|*S*0, *N*0, *E*0) is

 (S5)

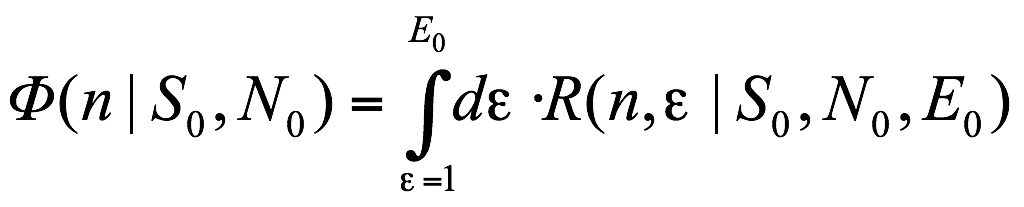
and

 (S6)

where = 1 + 2 and the ’s are Lagrange multipliers that are determined from the values of the state variables, *S*, *N*, *E*. In particular:

 (S7)

From the distribution *R*, various metrics of macroecology can be derived. For example, the species-abundance distribution (the probability that a randomly selected species has abundance *n*) is given by integrating out the energy variable:

 (S8)

Eq. S8 follows immediately from the definition of *R*(*n*, *ε*) .

Another metric that can be derived from *R* is the intraspecific metabolic rate distribution *Θ*(*ε*|*n*). *Θ*(*ε*|*n*)d*ε* is the probability that an individual selected from a species with abundance *n* has a metabolic rate in the interval *ε*, *ε* +d*ε*. From the fundamental rule *P*(*x*|*y*)*P*(*y*) = *P*(*x*,*y*) it follows that:

 (S9)

Of interest here, the community energy distribution is given by:

 (S10)

The factors of *S*0/*N*0 and *n* in the right hand side of Eq. (S10) require explanation. The fraction of species with abundance *n* is estimated by *Φ*(*n*), and so the expected total abundance of species with abundance *n* is *nS*0*Φ*(*n*). If an individual is picked at random from the individuals pool (*N*0), as opposed to the species pool, then the probability that the individual belongs to a species with abundance *n* is given by *nS*0*Φ*(*n*)/*N*0. Similarly, *Ψ*(*ε*)*dε* is the probability that an individual picked from the entire individuals pool (*N*0) has a metabolic energy requirement between *ε* and *ε* + *dε*. Recall that *Θ*(*ε*|*n*) equals the conditional probability that the metabolic rate of an individual is between *ε* and *ε* + *dε if* the individual is from a species with abundance *n*. Then from the general relationship that

 (S11)

we have:

 (S12)

Using Eq. S9 with Eq. S12, we immediately obtain Eq. S10. The summation in Eq. S10 is readily evaluated to yield equation (1) in the text; terms of order exp(-*N*0(1 + 2) are set to zero because they are less than or equal to exp(-*S*0) << 1.

**Appendix S2.** Code used to generate model and data comparisons in the R computing language; note that working directories will need to be updated to work on a local machine. Data needed to run the script are available online (github.com/ajrominger/psi\_mete) along with R scripts, reproduced below.

library(meteR)  
devtools::install\_github('ajrominger/socorro')  
library(socorro)  
  
setwd('~/Dropbox/Research/psi\_mete')  
  
# load data  
  
rmbl <- read.csv('RMBL\_PSI\_test.csv', as.is = TRUE)  
  
bci <- read.csv('BCIS.csv', as.is = TRUE)  
bci <- bci[bci$year == 1995, ]  
bcitest <- read.csv('BCI\_PSI\_test.csv', as.is = TRUE)  
  
arth <- read.csv('gruner\_kohala.csv', as.is = TRUE)  
  
# make METE objects  
  
rmblIPD <- ipd(meteESF(S0 = 31, N0 = 877, E0 = sum(rmbl$observed)))  
# need to custom add data to IPD for rmbl  
rmblIPD$data <- sort(rmbl$observed, decreasing = TRUE)  
  
bciIPD <- ipd(meteESF(spp = bci$spp, abund = bci$count,   
 power = bci$dbh^2))  
  
# metabolic scaling from Mori et al. PNAS paper  
bciIPDmori <- ipd(meteESF(spp = bci$spp, abund = bci$count,   
 power = (bci$dbh^(8/3))^0.805))  
  
  
pdf('fig\_bciPsiMori.pdf', width = 5, height = 5)  
plot(bciIPD, log = 'xy', axes = FALSE)  
logAxis(1)  
logAxis(2)  
box()  
dev.off()  
  
arthIPD <- ipd(meteESF(spp = arth$SpeciesCode, abund = arth$Abundance,   
 power = arth$IND\_BIOM^0.75))  
  
  
# plotting theory for BCI  
e <- exp(seq(log(1), log(500000), length = 1000))  
  
jpeg('ms/fig\_PsiThr.jpg', width = 3, height = 3, units = 'in', res = 380)  
  
par(mar = c(3, 3, 0, 0) + 0.5, mgp = c(2, 0.75, 0))  
plot(e, bciIPD$d(e), log = 'xy', type = 'l', col = 'red',  
 xaxt = 'n', yaxt = 'n', ylim = 10^c(-12, 0),  
 xlab = 'Metabolic rate', ylab = 'Probability density',  
 panel.first = {  
 rect(xleft = 10, xright = 10000, ybottom = 10^par('usr')[3], ytop = 10^par('usr')[4],  
 col = 'gray60', border = NA)  
 rect(xleft = 10000, xright = 10^par('usr')[2],   
 ybottom = 10^par('usr')[3], ytop = 10^par('usr')[4],  
 col = 'gray80', border = NA)  
 })  
logAxis(1, expLab = TRUE)  
axis(2, at = 10^seq(-12, 0, by = 3),   
 labels = sapply(seq(-12, 0, by = 3),   
 function(p)   
 eval(substitute(expression(10^p), list(p = p)))))  
  
dev.off()  
  
  
# plotting data and theory  
  
jpeg('ms/fig\_PsiData.jpg', width = 8, height = 3, units = 'in', res = 380)  
  
par(mfrow = c(1, 3), oma = c(3, 2, 0, 0) + 0.5, mar = c(0, 2, 1, 0) + 0.2,   
 cex = 1, mgp = c(2, 0.75, 0))  
  
plot(exp(bcitest$ln.rank.), exp(bcitest$ln.dbh2.), log = 'xy',  
 ylim = c(1, 300000), xaxt = 'n', yaxt = 'n',  
 xlab = '', ylab = '')  
points(exp(bcitest$ln.rank.), exp(bcitest$ln.PRED\_METE.), type = 'l',   
 col = 'red')  
logAxis(1, expLab = TRUE)  
logAxis(2, expLab = TRUE)  
legend('topright', legend = c('Data', 'METE'), pch = c(1, NA),   
 lwd = c(NA, 1), col = c('black', 'red'), bty = 'n', cex = 0.9)  
mtext('A', side = 3, at = 10^(par('usr')[1] + 0.05 \*   
 diff(par('usr')[1:2])),   
 line = 0.2)  
  
plot(rmblIPD, ptype = 'rad', log = 'xy', add.legend = FALSE,   
 ylim = c(1, 100000), xaxt = 'n', yaxt = 'n',  
 xlab = '', ylab = '')  
logAxis(1, expLab = TRUE)  
logAxis(2, expLab = TRUE)  
mtext('B', side = 3, at = 10^(par('usr')[1] + 0.05 \*   
 diff(par('usr')[1:2])),   
 line = 0.2)  
  
plot(arthIPD, ptype = 'rad', log = 'xy', add.legend = FALSE,  
 xaxt = 'n', yaxt = 'n', xlab = '', ylab = '')  
logAxis(1, expLab = TRUE)  
logAxis(2, expLab = TRUE)  
mtext('C', side = 3, at = 10^(par('usr')[1] + 0.05 \*   
 diff(par('usr')[1:2])),   
 line = 0.2)  
  
mtext('Metabolic rate', side = 2, line = 0.5, outer = TRUE)  
mtext('Rank', side = 1, line = 2, outer = TRUE)  
  
dev.off()