Towards a unified descriptive theory for spatial ecology: *Mathematica* Code Implementation

Upload Data

Import your data as a matrix where rows give the species position, the column is the species type, and the entry *(i,j)* of the matrix gives the abundance of the species *j* in the coordinate {gx(*i*),gy(*i*)}. In other words, you need to have three lists: one with all the species positions, one with the corresponding species label, and the last one with the corresponding abundance.

```
{{gx, gy}, sp, ab} = data;

(*Note that pos={gx,gy} gives the x-y coordinates of the species)

Uspecie = Union[sp];(*List of the different species in the ecosystem*)
NSpecies = Length[Uspecie];(*Total number of different species*)
TotAbundnace = Total[ab](*Total number of individuals in the ecosystem*);
```

Empirical PCF

These lines create the pcf empirical correlation function of the ecosystem, as described in the methods of the paper. If you already have a list giving the correlations avaraged among species for different distances, you can skip this part. Call the correlation list **empPCF.**

```
{Lx, Ly} = {1000, 500}
(*Dimensions of the whole ecosystem area. These numbers referes to the
  rectangular sampled area of 1000 m x 500 m of the BCI forest;*)
Asub = 100;(*Area of the plots →
 This number gives the level of resolution of our analyis. You need to choose
    a resolution so to have enough statistics and at the same time reduces
    the noise. The output of the PCF gives you an idea of this trade-
   off: you need to have a n≥30 points and not too noisy*)
allpos = Tuples[{Range[Lx / Asub], Range[Ly / Asub]}];
(*this line creates all the plots that cover the ecosystem area*)
xyCpos = Flatten[Table[{(2 i + 1) / 2, (2 j + 1) / 2} // N,
     {i, 0, (Lx / Asub) - 1, 1}, {j, 0, (Ly / Asub) - 1, 1}], 1];
  (*Gives the center plot coordinates*)
allpair = Tuples[{allpos, allpos}];
(*all possible combinations of plot positions*)
nplot = Range[Length[allpos]];(*total number of plots*)

(*These lines are needed to count the presence
 of species and their abundance in each of the plots*)
Xpos[p_] := Flatten[Position[gx, x_ ? (Asub * (p - 1) ≤ # < Asub * p &)], 1]
Ypos[q_] := Flatten[Position[gy, x_ ? (Asub * (q - 1) ≤ # < Asub * q &)], 1]
rule1 = Flatten[Table[{allpos〚i〛 → nplot〚i〛}, {i, 1, Length[allpos]}], 1];
rule2 = Flatten[Table[{nplot〚i〛 → allpos〚i〛}, {i, 1, Length[allpos]}], 1];
XYpos[{p_, q_}] := Intersection[Xpos[p], Ypos[q]];
possub = Map[XYpos, allpos];
(*It might take a while. If you have the last version of Mathematica,
use ParallelMap instead of Map*)
XYS[i_] := Transpose[{gx〚possub〚i〛〛, gy〚possub〚i〛〛, sp〚possub〚i〛〛}];
SabA[i_] := Count[sp〚possub〚i〛〛, #] & /@ Uspecies
SA[i_] := Boole /@ (MemberQ[sp〚possub〚i〛〛, #] & /@ Uspecies)
Spos = Table[SA[i], {i, 1, Length[possub]}];

(*These lines calculate the PCF for all
 different species and averaged over all species*)

dist[{i_, j_}] := Sqrt[Total[(Abs[xyCpos〚i〛 - xyCpos〚j〛])^2]]
index[{i_, j_}] := {i, j}
alldist = Map[dist, allpair /. rule1];
rvec = Union[alldist];
indexpair = Map[index, allpair /. rule1];
posdist = Position[alldist, #] & /@ rvec;
Cov[{i_, j_}] := Mean[SabA[i] SabA[j]] / (Mean[SabA[i]] Mean[SabA[j]]) // N

allcov = Table[Map[Cov, indexpair〚Flatten[posdist〚i〛, 1]〛], {i, 1, Length[rvec]}]
(*It might take a while. If you have the last version of Mathematica,
use ParallelTable instead of Table*)


empPCF = Transpose[{rvec * Asub, Map[Mean, allcov]}];
```

Fit of the empirical PCF

Important assumption: g(r) is supposed to fit well the PCF data.

```
g[r_] := 1 + ────── (─)² BesselK[0, ─]
            2 π   λ               λ
```

$$g[r\_] := 1 + \frac{1}{2\pi}\left(\frac{\rho}{\lambda}\right)^2 \text{BesselK}\left[0, \frac{r}{\lambda}\right]$$

```
λ = .;
ρ = .;
S0 = .;
(*bets fit of the empirical pcf -
 you may want to remove some points from the tail if they are too noisy*)
fitcor = NonlinearModelFit[Rest[empPCF], g[r], {{λ, 10 000}, {ρ, 50 000}}, r];
```

Goodness of fit statistics

```
Grid[Transpose[{#, fitcor[#]} &[{"AdjustedRSquared", "AIC", "BIC", "RSquared"}]],
 Alignment → Left]
```

```
AdjustedRSquared  0.999099
AIC               34.2898
BIC               38.1773
RSquared          0.999166
```

```
fitcor["ParameterTable"]
```

	Estimate	Standard Error	t-Statistic	P-Value
λ	2.30845×10^7	3.37374×10^7	0.684242	0.500117
ρ	6.32245×10^7	8.82087×10^7	0.71676	0.480165

```
σλ1 = fitcor["ParameterTableEntries"]〚1, 2〛;
σρ1 = fitcor["ParameterTableEntries"]〚2, 2〛;
```

```
fitcor["ParameterConfidenceIntervalTable"]
```
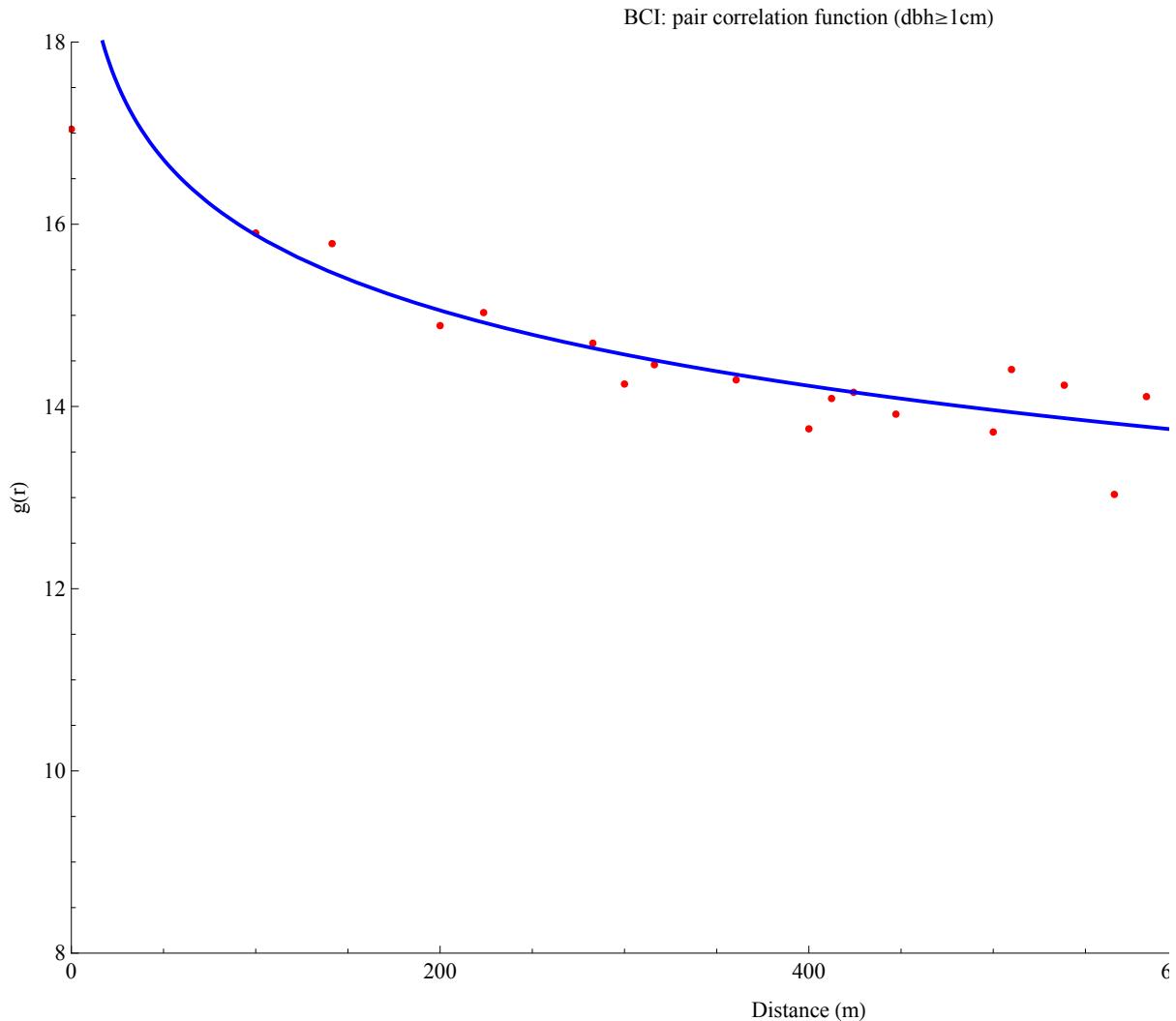
	Estimate	Standard Error	Confidence Interval
λ	2.30845×10^7	3.37374×10^7	$\{-4.63989 \times 10^7, 9.25679 \times 10^7\}$
ρ	6.32245×10^7	8.82087×10^7	$\{-1.18445 \times 10^8, 2.44894 \times 10^8\}$

```
Plot[fitcor[r], {r, Min[rvec * Asub], Max[rvec * Asub]}, PlotStyle → {Blue, Thick},
 PlotRange → {{0, 800}, {8, 18}}, Frame → {{True, False}, {True, False}},
 FrameLabel → {"Distance (m)", "g(r)"}, PlotLabel → "Pair correlation function",
 LabelStyle → {FontFamily → "Times", FontSize → 12},
 Prolog → {Red, PointSize[0.005], Point /@ empPCF}]
```


BCI: pair correlation function (dbh≥1cm)
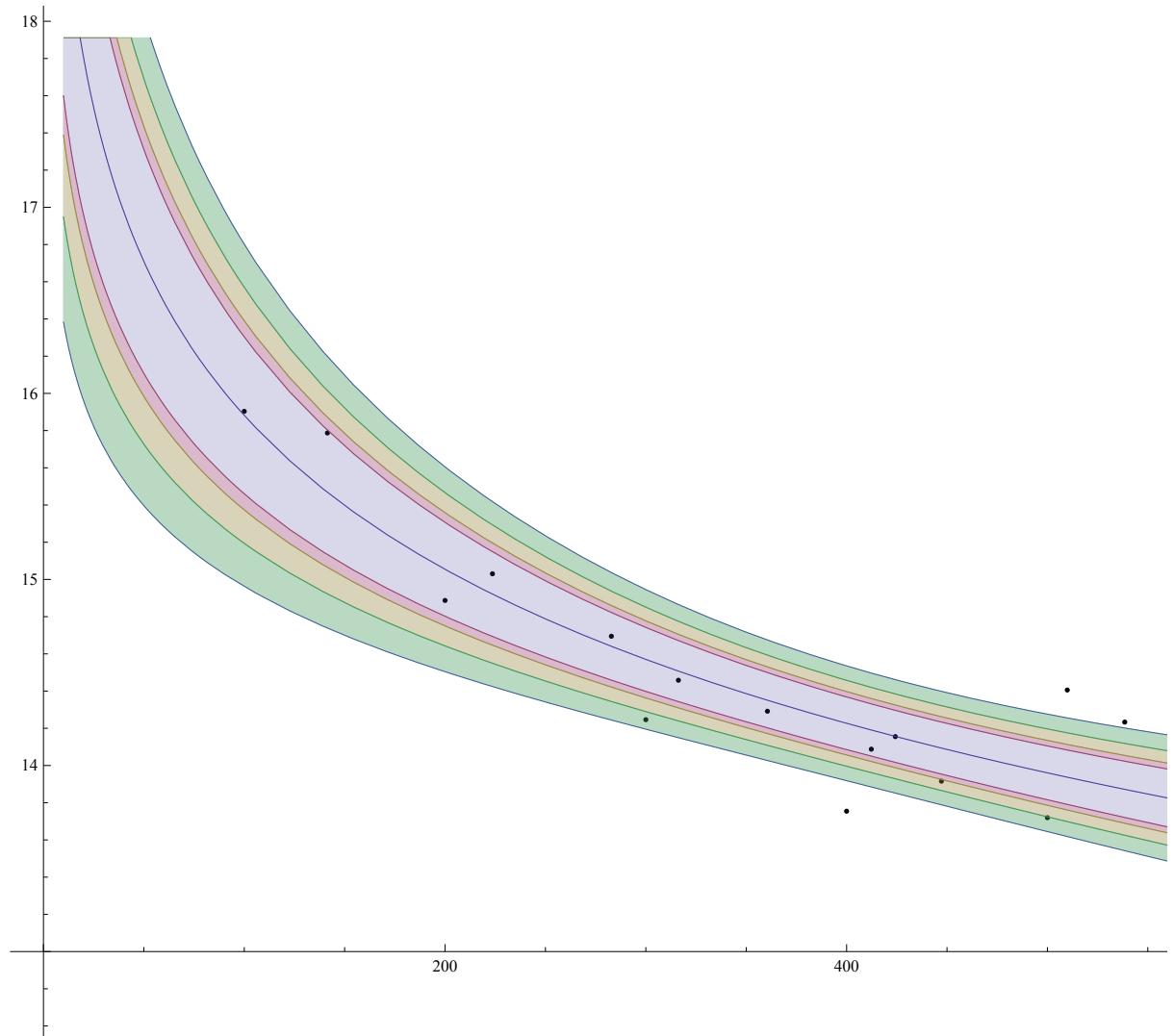
```
{bands90[x_], bands95[x_], bands99[x_], bands999[x_]} =
  Table[fitcor["MeanPredictionBands", ConfidenceLevel → cl],
   {cl, {.9, .95, .99, .999}}];
 (*Visualize the confidence bands with the model and data:*)
```

```
Show[ListPlot[Rest[empPCF], PlotStyle -> {Black, PointSize[0.003]}],
 Plot[{fitcorBCI[r], bands90[r], bands95[r], bands99[r], bands999[r]},
  {r, Min[rvec * Asub], Max[rvec * Asub]},
  Filling → {2 → {1}, 3 → {2}, 4 → {3}, 5 → {4}}]]
```


RSA and SAR prediction

```
ind = Table[Total[SabA[i]], {i, 1, Max[nplot]}];
(*number of individuals in each plot*)
```

α2p e **β2** are the α e β as described in the Supp. Methods. They represent the key result of the model through which we can make a prediction on the RSA and SAR at different scales (of radius L)

Important assumption here: the Gamma distribution is supposed to fit well the RSA data (at the whole scale).

```
λ = . ;
ρ = . ;
ν = . ;
α = . ;
β = . ;
```

$$\alpha 2p[L\_] := \pi \left(\frac{L}{\rho}\right)^2 \left(1 - \frac{2\lambda}{L} \left(\text{BesselK}\left[1, \frac{L}{\lambda}\right] \text{BesselI}\left[1, \frac{L}{\lambda}\right]\right)\right) \Big/$$
$$\left(\text{BesselK}\left[1, \frac{L}{\lambda}\right] \text{BesselI}\left[0, \frac{L}{\lambda}\right] + \text{BesselK}\left[0, \frac{L}{\lambda}\right] \text{BesselI}\left[1, \frac{L}{\lambda}\right]\right)\right)^{-1} ;$$

```
(*α2 is the adimensional shape parameter of the Gamma distribution. ρ
 e λ are the characteristics distances of the model,
L is the radius of a circular area*)
```

$$\beta 2[L\_] :=$$
$$(\text{Total}[\text{ind}] / (\text{Nspecies} * \text{Lx} * \text{Ly})) * \rho^2 \left(1 - \frac{2\lambda}{L} \left(\text{BesselK}\left[1, \frac{L}{\lambda}\right] \text{BesselI}\left[1, \frac{L}{\lambda}\right]\right)\right) \Big/$$
$$\left(\text{BesselK}\left[1, \frac{L}{\lambda}\right] \text{BesselI}\left[0, \frac{L}{\lambda}\right] + \text{BesselK}\left[0, \frac{L}{\lambda}\right] \text{BesselI}\left[1, \frac{L}{\lambda}\right]\right) ;$$

```
(*β2 has the dimensions of  ind  . Here we have fixed b/μ, i.e.  b = N    *)
                            specie                             μ   S A_tot
```

Prediction given by the model

$$\text{rsapres2medp}[n\_, L\_] := \text{CDF}\left[\text{GammaDistribution}[\alpha 2p[L], \beta 2[L]], 2^{n+1}\right] -$$
$$\text{CDF}[\text{GammaDistribution}[\alpha 2p[L], \beta 2[L]], 2^n] ;$$
$$(*\int_{2^n}^{2^{n+1}} p[x] dx, \text{ where } p[x] \text{ is the RSA with } n=0,1,2... *)$$

```
sarmedp[L_] := 1 - CDF[GammaDistribution[α2p[L], β2[L]], 1];
```
$$(*\text{SAR with } \frac{b}{\mu} = \frac{N}{S \, A_{tot}} \text{ fixed using the data}*)$$

$$\text{SAR}[r\_] := \text{NSpecies} * \frac{\text{sarmedp}[r]}{\text{sarmedp}\left[\sqrt{\frac{Lx*Ly}{\pi}}\right]} \text{ /. fitcor}["\text{BestFitParameters}"] ;$$

```
(*dowscaled SAR at radius r when all the information
   in the study region is available. The largest scale is Lx*
```
$$\text{Ly. The parameters are those from the PCF. Here } r < \sqrt{\frac{Lx*Ly}{\pi}} \text{ *)}$$

$$\text{SAD}[n\_, r\_] := \text{NSpecies} * \frac{\text{rsapres2medp}[n, r]}{\text{sarmedp}\left[\sqrt{\frac{Lx*Ly}{\pi}}\right]} \text{ /. fitcor}["\text{BestFitParameters}"] ;$$

```
(*dowscaled SAD at radius r when all the information
   in the study region is available. The largest scale is Lx*
```
$$\text{Ly. The parameters are those from the PCF. Here } r < \sqrt{\frac{Lx*Ly}{\pi}} \text{ *)}$$

$$
\texttt{upRichness = NsppSamp} * \frac{\texttt{sarmedp}\left[\sqrt{\frac{\texttt{Lx*Ly}}{\pi}}\ ,\ \texttt{dens},\ \lambda,\ \rho\right]}{\texttt{sarmedp}\left[\sqrt{\frac{\texttt{samp*Asp}}{\pi}}\ ,\ \texttt{dens},\ \lambda,\ \rho\right]}\ ;
$$

```
(*upscaled species richness at the whole scale from
  the information availbale from the samples. NsppSamp is the
  total number of species found in the aggregated samples;
samp is the total number of samples whose area is Asp;
dens is the mean population density per
  species as obtained from the aggregated samples;
λ and ρ are the parameters found from the PCF calculated from the samples. *)
```

```
UPdens = NSpecies * dens / upRichness;
(*mean population density per species at the scale of the study region*)
```

$$
\texttt{downSAR[r\_] := upRichness} * \frac{\texttt{sarmedp[r, UPdens, }\lambda,\ \rho]}{\texttt{sarmedp}\left[\sqrt{\frac{\texttt{Lx*Ly}}{\pi}}\ ,\ \texttt{UPdens},\ \lambda,\ \rho\right]}\ ;
$$

```
(*Downscaled Sar from the information

  available from the samples. Here r<√(Lx*Ly/π) . *)
```