

# Estudo da relação entre os parâmetros de um modelo neutro de campo médio e um modelo neutro de espaço explícito

Mori, Danilo Pereira; Coutinho, Renato Mendes; Lima, Renato A.F. ; Prado, Paulo Inácio K.

3 de abril de 2018

## Introdução

Os modelos neutros de campo médio apresentam dois parâmetros para descrever a distribuição de abundância de espécies (SAD): theta e m (Hubbell 2001, Etienne 2005). Já o modelo neutro de espaço explícito coalescente depende de dois parâmetros: U (taxa de especiação na paisagem) e d (distância média de dispersão dos indivíduos na metacomunidade) (Rosindell et al 2008).

### Relação entre $\theta$ e $U$

Theta é o chamado número fundamental da biodiversidade, número adimensional que pode ser descrita como a relação entre o número de indivíduos( $J_M$ ) e da taxa de especiação da metacomunidade (U):  $\theta = 2J_M U$  (Hubbell 2001). Assim, se  $\theta = 50$ , então, “o produto de duas vezes o tamanho da metacommunidade e a taxa de especiação (novas espécies por evento de nascimento na metacomunidade) é igual 50 novas espécies por geração” (Ricklefs 2012). Considerando a metacomunidade como uma paisagem finita composta de habitat e não habitat e saturada (todas as unidades de habitat estão colonizadas antes do evento de morte - *zero-sum assumption*), podemos escrever  $\theta$  como:

$$eqn1 : \theta = 2ApDA_{media}U$$

Onde  $DA_{media}$  = densidade media de indivíduos na paisagem (ind./ha), p = porcentagem de cobertura vegetal da paisagem (# unidades de habitat/ # total de unidades na paisagem [REESCREVER]), A = área total da paisagem (500 ha) e U = taxa de especiação na paisagem. Essa igualdade pressupõem que a densidade de unidades de habitat (ou de não habitat) é constante na paisagem. Considerando  $DA_{media} = DA_{obs}$ , podemos utilizar essa igualdade para obter U a partir do modelo neutro de campo médio (EI) e  $\theta$  a partir do modelo neutro de espaço explícito (EE).

### Relação entre m e função de dispersão

$m$  é a probabilidade de uma morte na ‘comunidade local’ ser sucedida de um evento de imigração (de movimentação sucedida de colonização e amadurecimento até a idade reprodutiva) de um indivíduos da metacomunidade. Assim, se  $m = 0.1$ , então 10% dos indivíduos na amostra (i.e. a ‘comunidade local’) são imigrantes que chegaram ao acaso da paisagem ao redor (adaptado de Ricklefs 2012). Esse parâmetro descreve a limitação à dispersão no nível da amostra, imaginando uma comunidade local que se relaciona por dispersão com um o conjunto de comunidades locais onde todas podem potencialmente trocar indivíduos (a metacomunidade/paisagem) (Etienne et al. 2007). Chisholm & Lichstein (2009) (doravante C&L09) estabeleceram uma relação entre  $m$  e  $A$  (a área da comunidade local):

When an individual at location (x, y) in the local community dies, the replacement individual may, by virtue of the random dispersal and recruitment processes, be from within the local community (i.e., within the quadrat) or from outside the local community (i.e., from outside the quadrat). Define  $m_{x,y}$  as the probability that the replacement individual at location (x, y) is drawn from outside the local community. This parameter will be highest for individuals on the edges of the quadrat and smallest for individuals at the centre of the quadrat, where  $m_{x,y} \gg 0$  for large A. We define  $m$  as the average value of  $m_{x,y}$  across the whole of the local community as follows:

$$m = \frac{1}{A} \int \int_A m_{x,y} dx dy$$

Essa equação pressupõem que a comunidade local está inserida em uma paisagem homogênea e que a dispersão ocorre de maneira radial, ou seja, com igual probabilidade para todas as direções. Utilizar essa equação no contexto da simulação coalescente (Rosindell et al. 2008) requer duas mudanças: i) na simulação a dispersão não ocorre de maneira radial, ao invés, ela é descrita como a resultante de dois sorteio independente em eixos ortogonais; ii) explicitar a função de dispersão, que descreve a probabilidade de um indivíduo colonizar uma unidade de habitat em função da distância entre está unidade de habitat e a planta progenitora [REVISÃO]. Especificamente para nosso conjunto de simulações selecionamos apenas amostras contíguas que aproximamos como quadrados o que permite reescrever a equação de C&L09 a partir do lado da amostra, L:

$$m = \left( \frac{1}{L} \int_{-L/2}^{L/2} m_x(x) dx \right)^2$$

$$m_x = 1 - \int_{-L/2}^{L/2} K(x-y) dy$$

Onde K é a função de dispersão. Como aproximamos todas as amostras por quadrados podemos escrever o lado da amostra como função de J e DA:  $L = 100\sqrt{J/DA}$ . Como utilizamos a mesma função de dispersão para realizar o sorteio nos dois eixos  $m_x = m_y$  e podemos apenas calcular  $m_x$  e elevar o resultado ao quadrado (primeira equação). Note que na 2a eq. a integral de K no espaço inteiro é 1, e limitamos o espaço à região dentro do quadrado, então a integral < 1, e  $m_x < 1$ ; na 1a eq. tomamos uma média de  $m_x$  no espaço, então isso também é sempre menor que 1 (Coutinho, e-mail “Reunião 29mar” de 8abril18). Em outras palavras Coutinho descreve o conjunto de equações como:

a gente pega um ponto de coordenadas x,y no plot, e vê qual a chance de ele ter vindo de fora (isso é o m). Pra isso, usamos o kernel, mas na prática sorteamos o deslocamento no eixo x e no y separadamente. Na direção x, a chance de cair no patch (que fica na região entre  $-L/2$  e  $L/2$ ) é a integral do kernel de  $-L/2$  a  $L/2$ , a chance de cair fora é o complementar, a integral de  $-\infty$  a  $-L/2$  mais a integral de  $L/2$  a  $\infty$ . Como a integral do kernel todo tem que ser 1, é mais fácil fazer  $1 - \text{integral de } -L/2 \text{ a } L/2$  na eq. 2, do que fazer as 2 integrais até  $(+/-)$  infinito. Essa integral te dá o  $m_x$ , e você calcula a média disso sobre todos os x, na eq. 1. Daí você faz o mesmo na direção y, mas dá o mesmo resultado, daí a expressão fica ao quadrado na eq. 1.

A distribuição de probabilidade que utilizamos na função de dispersão foi a distribuição Laplace. A distribuição Laplace, que pode ser entendida como uma distribuição exponencial negativa espelhada; a distribuição exponencial é comumente utilizada para estudar o efeito de processos que operam em larga escala, como aqueles em populações em paisagens fragmentadas, na chuva de propágulos (Clark et al. 1999). Assim, podemos reescrever o conjunto de equações considerando uma distribuição Laplace. Considerando a distribuição de Laplace com a seguinte parametrização de sua função de densidade de probabilidade  $K(x) = \frac{\alpha}{2} e^{-|\alpha x|}$ , temos:

$$m = \left( \frac{1 - e^{-\alpha L}}{\alpha L} \right)^2$$

Onde  $\alpha = 1/b$ , b é o parâmetro escalar da distribuição de Laplace que pode ser escrito em função do desvio-padrão da distância de valores (sd) como  $b = sd/\sqrt{2}$ . Assim podemos reescrever a equação de m em função de sd:

$$eq1 : m = sd \frac{1 - e^{-\frac{\sqrt{2}L}{sd}}}{\sqrt{2}L}$$

Com a equação 1 podemos calcular m a partir do desvio padrão da função de dispersão. Para calcular sd a partir de m, utilizamos o ramo principal da função W de Lambert ( $W_0$ ):

$$eq2 : sd = \frac{\sqrt{2}L}{mW_0\left(-\frac{e^{-1/m}}{m}\right) + 1}$$

### Função de dispersão e chuva de propagulos

A chuva de propágulos pode ser entendida como o produto da fecundidade e função de dispersão (Clark et al. 1999). Por conta do pressuposto da equivalência funcional todos os indivíduos produzem o mesmo número de propágulos por unidade de tempo (Hubbell 2001), assim, podemos simular cenários de limitação à dispersão em função da porcentagem de propágulos que permanece até determinada distância da planta progenitora. Dessa maneira, não precisamos definir a dispersão em termos de distância *per se* mas em termos de porcentagem de indivíduos que permanecem na área imediata da planta progenitora. Para isso é necessário estabelecer uma distância padrão da planta progenitora e estimar ou definir a porcentagem de indivíduos que se mantêm até esta distância padrão. Como distância padronizamos  $l_{cel}$ , assim, cada paisagem possui uma distância padrão que depende da densidade observada de indivíduos naquela paisagem [REESCREVER]. Podemos estimar qual a porcentagem de propágulos até a distância padrão que um determinado  $sd$  gera, partindo de um  $m$  (eqn 2); ou podemos informar *a priori* quais as porcentagens de interesse e estimar o  $sd$  necessário para gerar tais porcentagens. Na simulação coalescente, utilizamos 12 valores de porcentagem para simular os cenários de limitação à dispersão: 99%, seq(95,50,by=-5)% e 25%. Apesar da simulação coalescente ser bem eficiente e permitir simular paisagens infinitas, funções de dispersão que apresentam dispersão muito elevadas são computacionalmente muito onerosas (Rosindell et al. 2008) e apresentariam pouco realismo biológico (REFERÊNCIA), logo, não utilizamos porcentagens muito baixas (e.g.<1%).

Para estimar o  $sd$  necessário para gerar uma determinada porcentagem de propágulos até a distância padronizada, desenvolvemos uma função no ambiente de programação R (R language team). A seguir o código utilizado nessa função, note que a função permite utilizar 3 distribuições de probabilidade (uniforme, normal e Laplace), contudo utilizamos apenas a distribuição Laplace.

```
library(rmutil)

qkernel <- function(sigma, kernel, p, density = 20852/50, npoints = 100000) {
  kernel <- match.arg(kernel, choices = c("normal", "gaussian", "laplace",
    "uniform"))
  d_ind_MA <- 100/sqrt(density)
  if (kernel == "laplace") {
    b_laplace <- sigma/sqrt(2)
    X_laplace <- d_ind_MA * round(rlaplace(npoints, s = b_laplace)/d_ind_MA)
    Y_laplace <- d_ind_MA * round(rlaplace(npoints, s = b_laplace)/d_ind_MA)
    dist_laplace <- sqrt(X_laplace^2 + Y_laplace^2)
    result <- quantile(dist_laplace, p)
  }
  if (kernel == "normal" | kernel == "gaussian") {
    b_norm <- sigma
    X_norm <- d_ind_MA * round(rnorm(npoints, sd = b_norm)/d_ind_MA)
    Y_norm <- d_ind_MA * round(rnorm(npoints, sd = b_norm)/d_ind_MA)
    dist_norm <- sqrt(X_norm^2 + Y_norm^2)
    result <- quantile(dist_norm, p)
  }
  if (kernel == "uniform") {
    b_unif <- sigma/2
    X_unif <- d_ind_MA * round(runif(npoints, min = -b_unif, max = b_unif)/d_ind_MA)
    Y_unif <- d_ind_MA * round(runif(npoints, min = -b_unif, max = b_unif)/d_ind_MA)
  }
}
```

```

    dist_unif <- sqrt(X_unif^2 + Y_unif^2)
    result <- quantile(dist_unif, p)
}
return(unname(result))
}

sigkernel <- function(kernel, p, distance, density = 20852/50, npoints = 100000,
  sigma.min = 1, sigma.max = 100) {
  f1 <- function(x) distance - qkernel(x, kernel, p, density, npoints)
  uniroot(f1, lower = sigma.min, upper = sigma.max)
}

```

Nossa função para estimar o  $sd$  necessário para gerar determinada porcentagem até a distância padrão consiste de duas funções:  $qkernel$  e  $sigkernel$ .  $sigkernel$  retorna o parâmetro de dispersão necessário para que a função de dispersão tenha a distância padrão como quantil da porcentagem de interesse. Essa função depende da função  $qkernel$ , que por sua vez retorna o valor quantil que acumula determinado percentil de propágulos. A função  $qkernel$  pressupõem o processo de dispersão aleatório como está implementado na simulação coalescente descrito por Rosindell et al. (2008), ou seja, como duas amostras independentes da distribuição de probabilidade em eixos ortogonais, a distância de dispersão é descrito como o módulo do vetor resultado da amostra nos eixos ortogonais (figura 1).

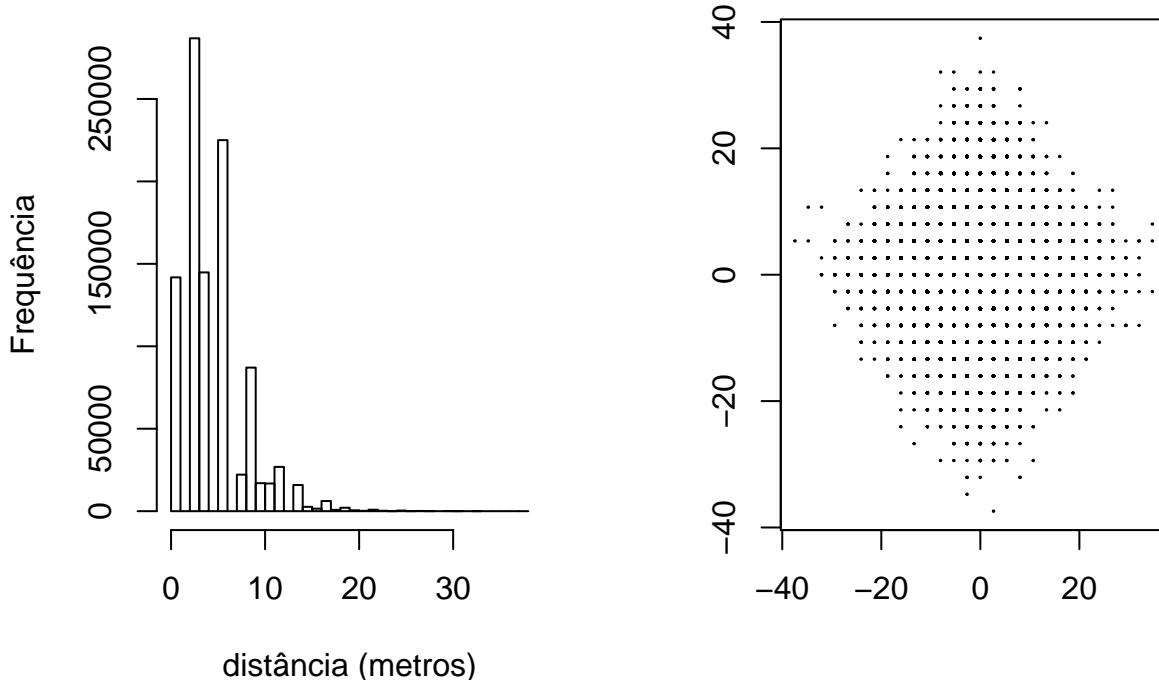


Figura 1. Simulação do processo de dispersão usado na simulação coalescente (Rosindell et al. 2008), consideramos DA = 1400 indivíduos/ha, desvio-padrão da função de dispersão = 4 metros e 1e6 pontos. No primeiro gráfico temos a distribuição de distâncias à planta progenitora (m). No segundo gráfico temos a distribuição da chuva de propagulos ao redor da planta progenitora que se encontra na origem do gráfico ( ponto (0,0)). Notem que diferente de uma dispersão que assume que a distribuição de probabilidade é bivariada (que aproximaria de um círculo), a maneira como a simulação coalesceente (Rosindell et al. 2008) aproxima a dispersão faz com que a forma da chuva de propagulos se aproxime de um losango.

Podemos calcular qual a correspondente chuva de propagulo (%k) de um  $m$  conhecido, considerando a densidade observada e o tamanho da amostra. Considerando  $m=0.11$  (valor estimado para BCI REFERÊNCIA),  $J=20852$  (número de indivíduos em BCI),  $A=50\text{ha}$  (área amostrada em BCI) e  $DA=417.04$ , podemos calcular  $sd$  correspondente considerando uma função de dispersão com distribuição de Laplace:

```

J = 20852
DA = 20852/50
(sd = 0.11 * 100 * pi * sqrt((2 * J)/DA)/4) #eqn 2

## [1] 86.3938

```

O  $sd$  estimado é bem elevado, acredito que o principal motivo para isso é que a área amostrada é grande e a densidade dos indivíduos é baixa, gerando um baixo valor de indivíduos (J) para uma respectiva área. Assim, a função de dispersão descreve eventos de colonização em unidades de habitat que possuem arrestra grande. Vamos avaliar qual é a correspondente chuva de propágulos considerando uma distribuição Laplace:

```

library(rmutil)
# distância entre os indivíduos e a distância padrão
d_ind_MA <- 100/sqrt(DA) #tanto a metrica da paisagem quanto a distancia de referencia
b_laplace <- sd/sqrt(2) ## relação entre sigma e b, a variância da laplaciana
## gerando valores de uma distribuição laplaciana e convertendo para a
## distância em metros
X_laplace <- d_ind_MA * round(rlaplace(100000, s = b_laplace)/d_ind_MA)
Y_laplace <- d_ind_MA * round(rlaplace(100000, s = b_laplace)/d_ind_MA) # idem para Y
dist_laplace <- sqrt(X_laplace^2 + Y_laplace^2) #distâncias dos pontos até a origem
(percentil <- length(dist_laplace[dist_laplace <= d_ind_MA])/length(dist_laplace)) # percentil

## [1] 0.00742

```

Uma maneira de avaliar se o percentil está corretamente é utilizar esse valor na função quantil e ver se o valor é próximo de ‘d\_ind\_MA’. O quantil correspondente à esse percentil é igual 6.9100514, valor próximo em escala de 4.8967862. Uma segunda avaliação pode ser utilizar o valor do quantil para calcular o percentil

```

(percentil2 <- length(dist_laplace[dist_laplace <= unname(quantile(dist_laplace,
percentil))])/length(dist_laplace)) # percentil do quantil

## [1] 0.00742

```

O valor de percentil calculado pela distância referência, a distância entre indivíduos, e o quantil correspondente ao percentil possuem o mesmo valor. O erro da simulação deve ocorrer pelo uso da função ‘round’ na simulação de valores de uma distribuição Laplace em dois eixos ortogonais. Portanto, existe um problema de aproximação da função que provavelmente está relacionado com a necessidade de tornar discreta uma distribuição de probabilidade que é contínua, assim, esse erro independente da  $DA_{obs}$ .

Isso quer dizer que apenas 0.742% dos propágulos se mantêm até cerca de 4.8967862 da planta progenitora. Considerando um modelo coalescente, então isso representaria que menos de 1% dos propagulos permanece até  $l_{cel}$  da planta progenitora, ou seja, até duas vezes a área da sua copa (assumindo uma árvore que ocupa toda a área da unidade de habitat). Esse é um cenário de ausência de limitação à dispersão no nível da comunidade local. Acredito que esse valor ocorre pois no modelo neutro, o evento de dispersão é resultado do processo de dispersão *per se* e do de colonização seguido do amadurecimento até a idade reprodutiva. Os dados de BCI são de DBH  $\geq 10$  cm, considerando que existe considerável correlação entre o diâmetro da árvore e sua idade, podemos considerar que estamos amostrando em BCI indivíduos que já existem há muitas gerações. E como a produção de propagulos é muitas vezes mais rápido que o processo de dispersão (*sensu* teoria neutra) podemos considerar que já ocorreram diversos eventos de dispersão *per se* de modo que todas as classes de distância já foram apresentados, permitindo o surgimento de indivíduos em longas distâncias de suas respectivas árvores progenitoras. Pelo percentil apresentado, esse efeito deve ser forte o suficiente para produzir populações com índices de agregamento que tendem a zero. Uma vez que o modelo de campo médio, desconsidera o efeito da limitação à dispersão no nível de suas escalas (comunidade local e metacomunidade) e considera esse efeito apenas entre escalas (Etienne 2005), podemos considerar razoável que o percentil tenha sido tão baixo, isso mostra coerência interna do modelo, apesar do baixo realismo biológico. (Seria interessante realizar uma pesquisa bibliográfica para avaliar as questões levantadas nesse parágrafo - MAS NÃO AGORA).

Uma possível pergunta nesse contexto é qual a equivalente chuva de propágulo para que  $m = 1$ , ou seja, sem qualquer tipo de limitação à dispersão. Como esse valor provavelmente é muito baixo, vou aumentar o número de pontos utilizado para estimar o percentil:

```
# sd máximo teórico
sd_max <- 100 * pi * sqrt((2 * J/DA))/4 #eqn 2
b_laplace <- sd_max/sqrt(2)
# aumentei o número de pontos para maximizar a resolução do percentil
X_laplace <- d_ind_MA * round(rlaplace(1000000, s = b_laplace)/d_ind_MA)
Y_laplace <- d_ind_MA * round(rlaplace(1000000, s = b_laplace)/d_ind_MA)
dist_laplace <- sqrt(X_laplace^2 + Y_laplace^2)
(percentil_max <- length(dist_laplace[dist_laplace <= d_ind_MA])/length(dist_laplace)) # percentil

## [1] 0.000079
```

O percentil máximo teórico 0.0079% é o esperado para quando não existe limitação à dispersão entre as escala comunidade local e a metacomunidade, ou seja, todos os indivíduos tem capacidade de produzir propágulos que podem se dispersar para qualquer unidade de habitat da metacomunidade.

Alternativamente podemos estimar  $sd$  que gera um determinado percentil de propágulos até a distância referência. Essa abordagem é principalmente importante na simulação coalescente, onde não possuímos ferramentário estatístico para estimar os parâmetros do modelo U e a chuva de propagulos (%k) a partir do ajuste da  $SAD_{obs}$  por verossimilhança . Na simulação coalescente podemos estimar U a partir da riqueza observada utilizando uma modificação da formula semi-analítica apresenta no apêndice A de Rosindel et al. (2008). A formula estabelece relação entre a probabilidade de colonização de um singleton por evento de morte (ou seja, a probabilidade de uma nova espécie aparecer na paisagem - U) e a riqueza observada partindo de uma árvore genealogica da comunidade quando  $U > 0$ , ou seja, quando há monodominância. No modelo coalescente, reconstruímos a arvore genealógica da comunidade no espaço, assim a função de dispersão altera a arvore genealógica da comunidade e por consequência a relação entre U e a riqueza observada. Dessa maneira, para trabalhar com o modelo neutro espacialmente explícito informamos *a priori* qual a chuva de propágulos de interesse e estimamos  $sd$  correspondentes. Utilizamos 12 valores de porcentagem de propágulos que permanecem até  $l_{cel}$  metros da planta progenitora: 99%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 25%. Vamos calcular os respectivos valores de  $sd$  que produzem os percentis de interesse utilizando a função sigkernel:

```
# data frame com os percentis de interesse
df_ <- data.frame(l_cel = 100/sqrt(DA),
                  percentil = c(99, seq(95,50,by=-5), 25)/100,
                  kernel_type = "laplace",
                  DA_obs = DA,
                  J_obs = J,
                  stringsAsFactors = FALSE)

f_sapply <- function(i){
  sigkernel(kernel = df_[i,"kernel_type"],
            p = df_[i,"percentil"],
            density = df_[i,"DA_obs"],
            distance = df_[i,"l_cel"])$root
}
df_$sd <- sapply(1:nrow(df_),f_sapply)
```

Com os valores de  $sd$  podemos estimar o correspondente 'm' utilizando a eqn2:

```
library(car)
df_ %>% mutate(m = sd * 4/(100 * pi * sqrt((2 * J)/DA)))
scatterplotMatrix(~percentil + sd + m, data = df_, main = "")
```

```

## Warning in smoother(x[subsx], y[subsy], col = smoother.args$col[i], log.x =
## FALSE, : could not fit positive part of the spread

```

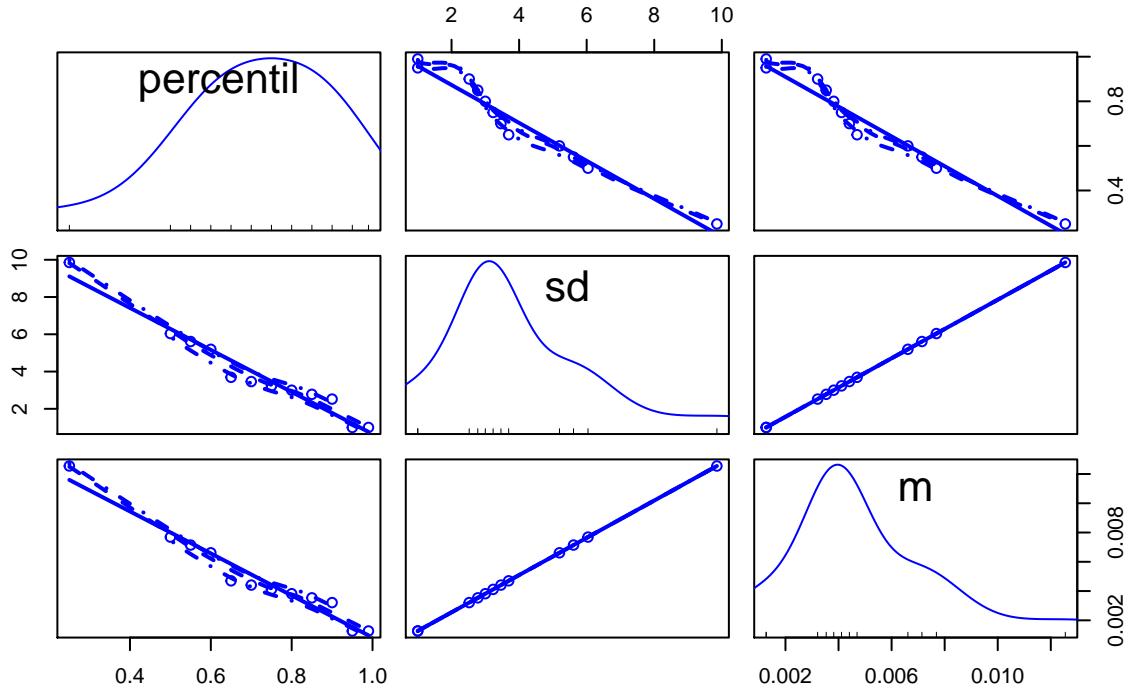


figura 2. Relação entre o percentil de propágulos que permanecem até a distância padrão de  $100/\sqrt{DA}$  que é a distância dos indivíduos na simulação coalescente (painel[1,1]), o desvio padrão correspondente  $sd$  (painel[2,2]) e a taxa de imigração calculada a partir de  $sd$  e da equação 2. A relação entre  $m$  e  $sd$  é proporcional para valores de  $J$  e  $DA$  iguais (eqn 2); já a relação do percentil e  $sd$  é descrita por uma distribuição Laplace que foi transformada para valores discretos (detalhes no texto).

Os valores utilizados até aqui são observados em BCI, as características da amostra de BCI difere daquelas presentes no TreeCo em: i) paisagem aproximada como homogênea, ou seja, existem apenas unidades de habitat na paisagem (não há fragmentação); ii) a área amostrada de BCI é cerca de 50 vezes maior que a moda das áreas amostradas em nosso filtro de trabalhos do TreeCo ( $\sim 1\text{ha}$ ); iii) BCI conta com o censu de árvores com  $DBH \geq 10\text{ cm}$ , os sítios selecionados no TreeCo utilizaram  $DBH \geq 5\text{ cm}$ , espera-se que isso tenha efeito em  $DA_{obs}$ . Assim, na próxima sessão vou fazer a conversão de parâmetros do modelo EI e EE das paisagens selecionadas no TreeCo

## Convertendo Parâmetros de modelos neutros para as paisagens de TreeCo

Foram selecionados 80 trabalhos fitossociológicos que amostram de maneira contígua pelo menos 1 ha de indivíduos arbóreos na Mata Atlântica, Brasil, e com critério de inclusão  $DBH \geq 5\text{ cm}$ , além disso os trabalhos foram realizados em paisagens que mudaram pouco até o ano de 2000, ano em que a foto área do recorte de paisagem foi tomada (detalhes em outros trabalhos do TreeCo REFERÊNCIAS). Anteriormente ajustei o modelo de Etienne (2005) às SADs amostradas nos trabalhos fitossociológicos, obtendo  $\theta$  e  $m$ . Concomitantemente, trabalhamos com o modelo coalescente nos mesmos trabalhos, utilizamos recortes de paisagens de 500 ha concêntricos à coordenada central dos trabalhos fitossociológicos, esses recortes de paisagem consideram a estrutura de habitat e não habitat na paisagem, assim, transcrevemos a estrutura de paisagem para a simulação; ao redor da coordenada central geramos um quadrado com  $J$  unidades de habitat, que representa a comunidade local (detalhes no métodos). Cada matriz de paisagem apresenta o número de elementos (unidades de habitat e de não habitat) tal que a densidade de elementos é igual a  $DA_{obs}$ . A simulação coalescente ocorre nas matrizes de paisagens geradas e conta com duas etapas: primeiro para estimar a taxa de especiação necessária ( $U$ ) para se obter a riqueza observada no equilíbrio dado uma função

de dispersão; então utilizando U e a mesma função de dispersão simulamos 100 SADs neutras. Aqui vamos focar apenas na conversão de parâmetros entre o modelo de campo médio de Etienne (2005) (modelo EI) e o modelo de espaço explícito coalescente (modelo EE).

### Convertendo parâmetros do modelo EI

O objetivo aqui é utilizar as eqn 1 e 2 para converter  $\theta$  e  $m$  para U e %k (porcentagem da chuva de propagulos que se mantêm até o dobro da copa da planta progenitora). Primeiro vamos olhar os dados para o modelo EI:

```
# data frame com todos os ddos e as conversões de parâmetros antigas
# (erradas)
df_resultados <- read.table(file = "df_resultados.txt", header = T, as.is = T)
# filtrando os dados para o modelo EI e selecionando as colunas de interesse
df_EI <- df_resultados %>% filter(modelo == "EI") %>% select(SiteCode, p, S,
  DA, J, theta, m)
summary(df_EI[, c(2, 4:7)])
```

	p	DA	J	theta
## Min.	:0.007424	Min. : 595	Min. : 595	Min. : 4.119
## 1st Qu.	:0.291633	1st Qu.:1138	1st Qu.: 1163	1st Qu.: 18.634
## Median	:0.672751	Median :1474	Median : 1488	Median : 25.716
## Mean	:0.602527	Mean :1497	Mean : 1792	Mean : 30.416
## 3rd Qu.	:0.921663	3rd Qu.:1745	3rd Qu.: 1803	3rd Qu.: 39.461
## Max.	:1.000000	Max. :3148	Max. :12105	Max. :121.666
## m				
## Min.	:0.04292			
## 1st Qu.	:0.46322			
## Median	:0.78243			
## Mean	:0.69719			
## 3rd Qu.	:0.99864			
## Max.	:0.99997			

Com esses dados podemos aplicar as eqn 1 e 2:

```
df_EI %>% mutate(U = theta / (2 * 500 * DA * p), #eqn1
                     sd_k = m * 100 * pi * sqrt((2*J)/DA)/4) #eqn2
scatterplotMatrix(~ J + DA + p + theta + U + m + sd_k,
                  data= df_EI)
```

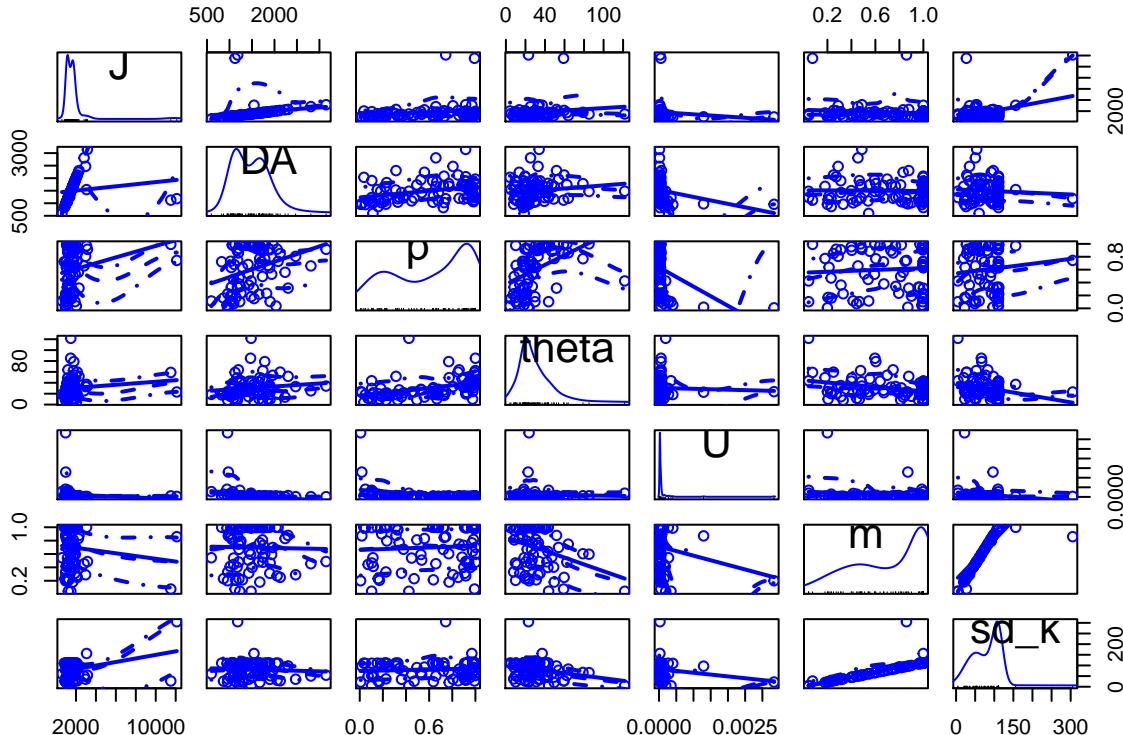


Figura 3.

Valores utilizados na conversão de parâmetros do modelo EI ( $J$ ,  $DA$  e  $p$ ) e os parâmetros do modelo EI ( $\theta$ ,  $U$ ,  $m$  e  $sd$ ) e seus respectivos parâmetros do modelo EE ( $U$  e  $sd$ ).

As amostras descrevem uma linha entre  $J$  e  $DA$ , com duas notórias exceções: há dois valores de  $J$  que são muitas vezes superiores aos demais valores; esses dois valores ocorrem em paisagens com cobertura vegetal elevada e distoam da nuvem de pontos dos parâmetros  $\theta$ ,  $U$ ,  $m$  e  $sd$ . A distribuição de  $p$  é levemente bimodal, apresentando a maioria dos valores em paisagens com alta cobertura vegetal, paisagens com valores intermediários são os menos representados - idealmente gostaríamos de uma distribuição uniforme, contudo, grande parte dos trabalhos fitossociológicos ocorrem em paisagens pouco modificadas, o que explica a alta frequência de coberturas vegetais (REFERÊNCIAS - treeco).  $U$  apresenta dois valores distoantes da nuvem de pontos, esses valores ocorrem em baixos valores de  $p$ , assim, uma possibilidade é que esses valores distoantes sejam resultado de uma interação de valores baixos de  $p$  com altos valores de  $DA$ , contudo, como ocorrem em valores baixos de  $p$  não acho recomendado retira-los dada a escassez de pontos em baixos  $p$ . Existe coerência entre  $m$  e  $sd$ , onde esperamos uma relação proporcional entre  $m$  e  $sd$  (eqn 2), contudo, parece que os valores distoantes de  $J$  estão influenciando na conversão entre as variáveis.

Parece que existe coerência entre as conversões de parâmetros dos modelos, vou seguir estimando qual a correspondente chuva de propágulos para os  $sd$  calculados. Para isso vou desenvolver uma formula adaptada de qkernel, que ao invés de retornar o quantil, retorna o percentil de interesse:

```
f_percentil.kernel <- function(i, df_ = df_EI) {
  percentilkernel <- function(sigma, density, npoints = 100000) {
    # metrica da simulacao e distancia de referencia
    d_ind_MA <- 100/sqrt(density)
    # relacao entre sd e o parametro escalar da distribuicao Laplace
    b_laplace <- sigma/sqrt(2)
    # sorteios de duas distribuicoes identicas Laplace em eixos ortogonais
    X_laplace <- d_ind_MA * round(rlaplace(npoints, s = b_laplace)/d_ind_MA)
    Y_laplace <- d_ind_MA * round(rlaplace(npoints, s = b_laplace)/d_ind_MA)
    # calculando a distancia dos pontos ate a origem
    dist_laplace <- sqrt(X_laplace^2 + Y_laplace^2)
    # Percentil
    quantile(dist_laplace, i)
  }
  percentilkernel
}
```

```

percentil <- length(dist_laplace[dist_laplace <= d_ind_MA])/length(dist_laplace)
return(percentil)
}
kernel_percentil <- percentilkernel(sigma = df_[i, "sd_k"], density = df_[i,
    "DA"])
return(kernel_percentil)
}
df_EI$kernel_percentil <- sapply(1:nrow(df_EI), f_percentil.kernel) #armazenando

scatterplotMatrix(~J + DA + m + sd_k + kernel_percentil, data = df_EI)

```

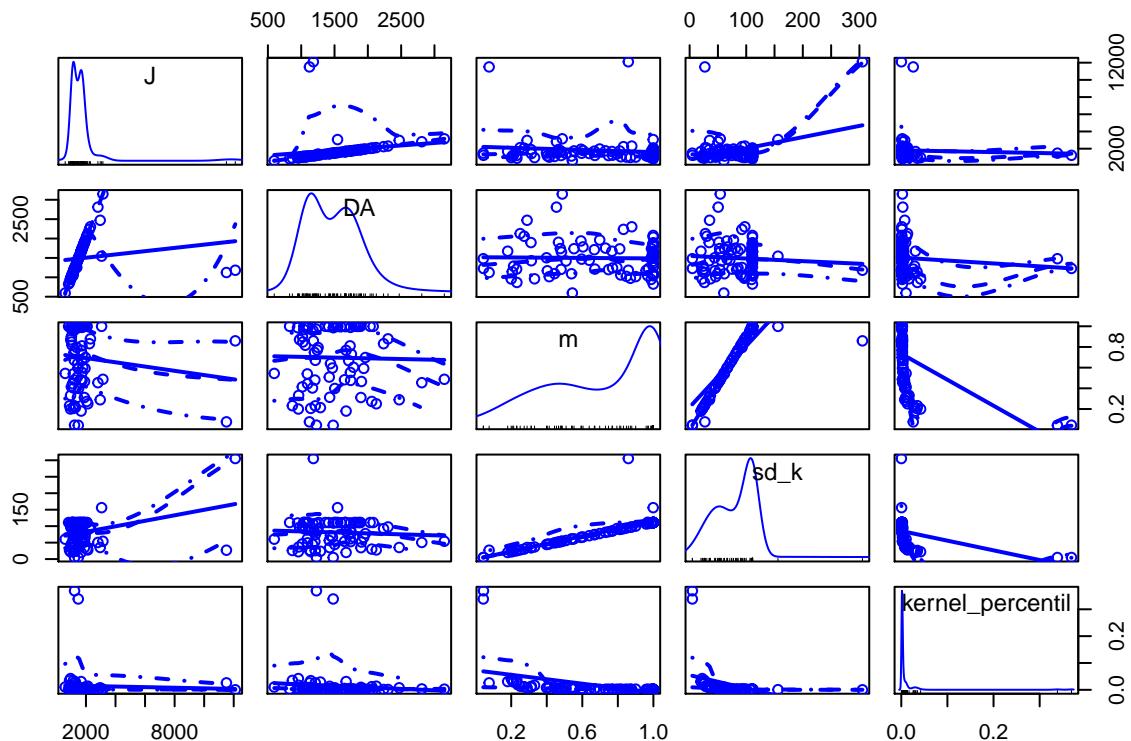


figura 4.

Os valores utilizados na conversão (J e DA) e as variáveis convertidas (m e sd) e a chuva de proágulos correspondente (kernel\_percentil).

Os valores de %k variam de ~ 0.028% até cerca de 37.07%, sendo a mediana ~ 2.56 % e o terceiro quartil ~ 6.57% o quê indica que o modelo EI pressupõem em geral indivíduos com alta capacidade de dispersão - o quê indica pouca realismo biológico.

### Convertendo parâmetros do modelo EE

No modelo coalescente informamos *a priori* os valores de %k, então convertemos para o respectivo *sd*, e estimamos o U necessário para se obter a riqueza observada no equilíbrio. Vamos calcular utilizando as eqn 1 e 2 os correspondentes valores de  $\theta$  e  $m$  para cada respectivo cenário de %k e U:

```

# selecionando os dados
df_EE <- df_resultados %>% filter(modelo == "EE") %>% select(SiteCode, kernel_percentil, p, S, DA, J, U)

#convertendo os parâmetros
df_EE %>% mutate(theta = 2 * 500 * p * DA * U, #eqn1
                     m = sd * 4 / (100 * pi * sqrt( (2*J)/DA ) ) ) #eqn2

```

