Group Dispersal

Probabilistic Characterization and Ecological Implications

Samuel Soubeyrand
INRA – Biostatistics and Spatial Processes

Paris, December 15, 2017

Group dispersal has been reported for many types of organisms and biological propagules:

- terrestrial and marine animals
- plant seeds and pollen
- fungal spores, peat moss spores, bacteria, viruses

Group dispersal

- can occur as rare extreme events (e.g. rains of organic matters caused by extreme storms; McAtee, 1917)
- or can be the rule (e.g. seed dispersal by frugivores; Pizo and Simao, 2001)

Living things raining down from the sky

SHOWERS OF ORGANIC MATTER.

By Waldo L. McAtee, Assistant Biologist.

[Address: U. S. Bureau of Biological Survey, Washington, D. C.]

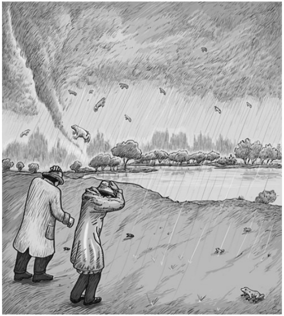
[Paper presented to the Biological Society of Washington, Jan. 27, 1917.]

CONTENTS.

Introduction 217
Spurious showers—
Insect larvæ 218
Ants 218
Honey; sugar 218
Grain 218
Black rain 218
"Blood" rain 218
Manna 218
True showers—
Red rains; dust 218
Plants and invertebrates—
Pollen; "sulphur" rain 219
Hay 219
Wheat 220
Meteoric "paper" 220
Jelly; "flesh" 220
Insects 220
Molluscs 20
Vertebrates—
Frogs; toads 221
Fish 221
Other vertebrates 223
Wind as a distributing agent—
Vertebrates 223
Plant seeds 223
Spores, statoblasts, eggs 223
Conclusion 224

(McAtee, 1917, Monthly Weather Review)

(Gravure de Magnus, 1555)



(Le Magasin Pittoresque, 1836)

Long-distance seed dispersal by frugivores may increase seed survival and yield plant aggregates

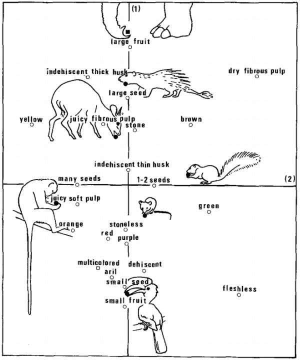


Fig. 2. Factorial plane 1–2 of the multifactorial analysis showing the interrelationships among the six group of consumers and the fruit characters (*white circle*: active variable for fruit; *black circle*: active variable for consumers; *white square*: supplementary variable for fruit; *black square*: supplementary variable for consumer

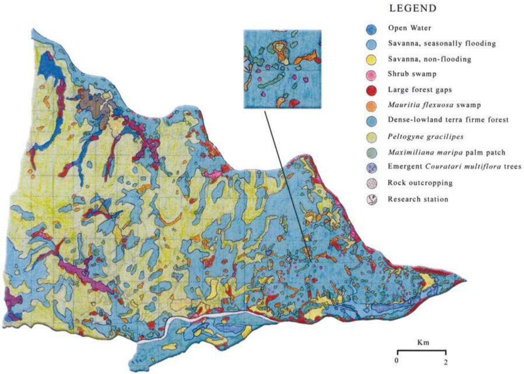
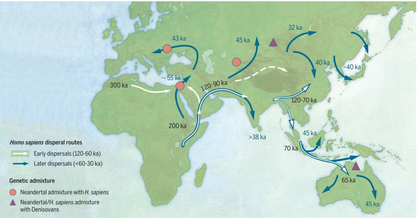


FIG. 1. Meso-scale map of plant associations and some plant species on the eastern third of Maracá Island Ecological Reserve, Roraima, Brazil. Nonflooding savanna supports a mixed vegetation of grasses and shrubs (*Curatella americana* and *Byrsonima* sp.). Forest gaps include open areas near the river dominated by *Triplaris* sp. and *Cecropia* sp. trees. *Peltogyne gracilipes* forest is dominated by this leguminous species but also supports palms and species in the Sapotaceae at low densities. The inset shows a close-up of a section of forest with intermingled *Couratari multiflora* emergents (pink) and *M. maripa* palm patches (dark green); the two species co-occur throughout the eastern end of the study site.

(Gautier-Hion et al., 1985) (Fragaso et al., 2003)

Group dispersal for *H. sapiens*

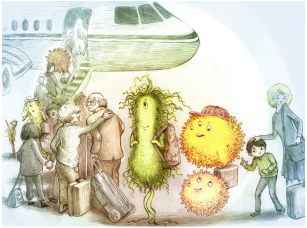
"The clearest demonstration of the effectiveness of the social networks and technological efficiency of *H. sapiens* lies in the evidence for our species' dispersal into new habitats. [...] Three of this dispersals (Sahul, the Philippines and Paleo-Honshu) could only have been accomplished by using boats or rafts that could be steered and that probably needed sails or oars for propulsion..."
(Dennell, chap. 3, In Boivin et al., 2017)



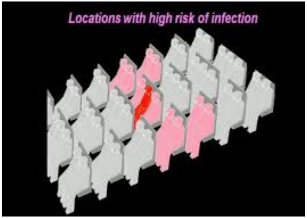
(Bae et al., 2017)

(Homo sapiens, J. Malaterre, 2005)

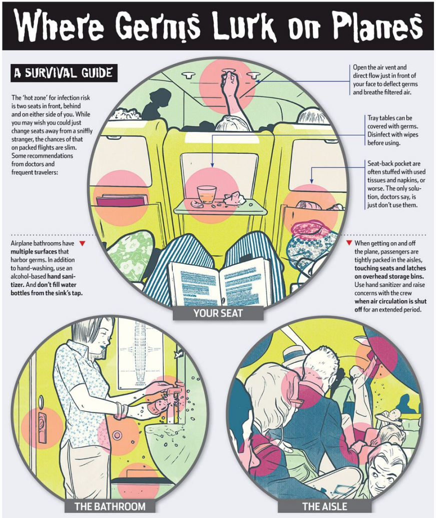
The airplane: a semi-closed setting facilitating transmission and synchronous movements of pathogens



(Benjamin Arthur for NPR)



(ANSYS)

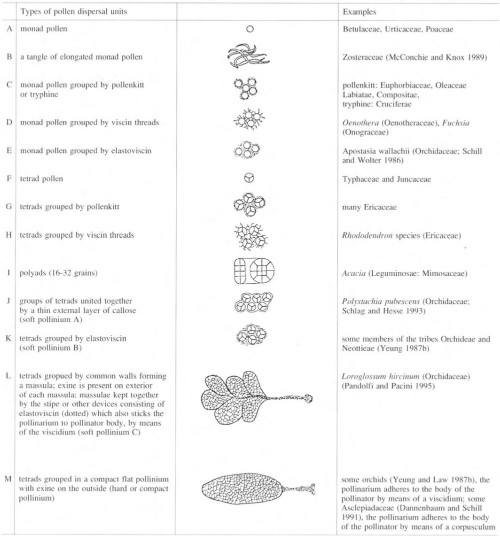


(The Wall Street Journal, 2011)

Clumps of pollen grains

Examples of pollen dispersal units:
(Pacini and Franchi, 1999)

Rk: Clumping also classically occurs for fungal spores
(Ingold, 1971; Rapilly, 1979)

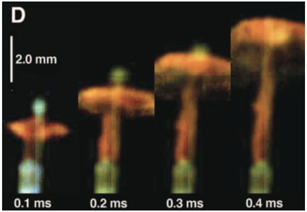


	Types of pollen dispersal units		Examples
A	monad pollen		Betulaceae, Urticaceae, Poaceae
B	a tangle of elongated monad pollen		Zosteraceae (McConchie and Knox 1989)
C	monad pollen grouped by pollenkitt or tryphine		pollenkitt: Euphorbiaceae, Oleaceae Labiatae, Compositae, tryphine: Cruciferae
D	monad pollen grouped by viscin threads		*Oenothera* (Oenotheraceae), *Fuchsia* (Onograceae)
E	monad pollen grouped by elastoviscin		Apostasia wallachii (Orchidaceae; Schill and Wolter 1986)
F	tetrad pollen		Typhaceae and Juncaceae
G	tetrads grouped by pollenkitt		many Ericaceae
H	tetrads grouped by viscin threads		*Rhododendron* species (Ericaceae)
I	polyads (16-32 grains)		*Acacia* (Leguminosae: Mimosaceae)
J	groups of tetrads united together by a thin external layer of callose (soft pollinium A)		*Polystachia pubescens* (Orchidaceae; Schlag and Hesse 1993)
K	tetrads grouped by elastoviscin (soft pollinium B)		some members of the tribes Orchideae and Neottieae (Yeung 1987b)
L	tetrads gropued by common walls forming a massula; exine is present on exterior of each massula: massulae kept together by the stipe or other devices consisting of elastoviscin (dotted) which also sticks the pollinarium to pollinator body, by means of the viscidium (soft pollinium C)		*Loroglossum hircinum* (Orchidaceae) (Pandolfi and Pacini 1995)
M	tetrads grouped in a compact flat pollinium with exine on the outside (hard or compact pollinium)		some orchids (Yeung and Law 1987b), the pollinarium adheres to the body of the pollinator by means of a viscidium; some Asclepiadaceae (Dannenbaum and Schill 1991), the pollinarium adheres to the body of the pollinator by means of a corpusculum

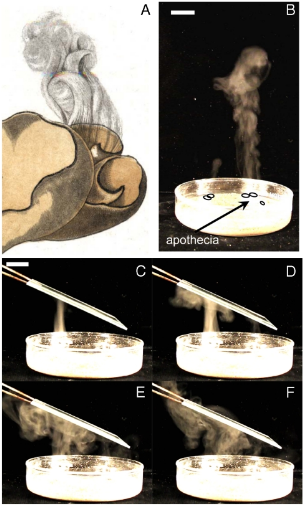
Fig. 1. Different types of pollen dispersal units with examples. Pollen grains are not drawn to scale. Types of pollen dispersal units as A. and F. are dispersed by air currents; B. by sea water currents; D., E., H., I., J., K., L., M., by animals, mainly insects; C. and G. by animals and/or air currents.

Release of groups of spores

Ejection of spores from a *Sphagnum* moss capsule

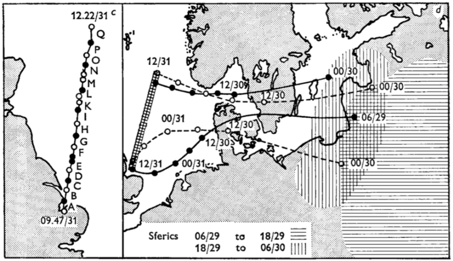


(Whitaker and Edwards, 2010)

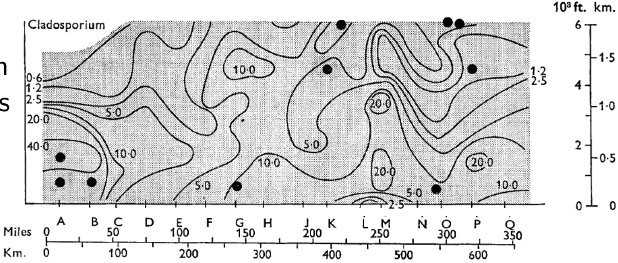


Ejection of spores of *Sclerotinia sclerotiorum*: (Roper et al., 2010)

Aggregates of spores in the atmosphere



Large-scale distribution of Cladosporium spores in the air:
(Hirst et al., 1967)



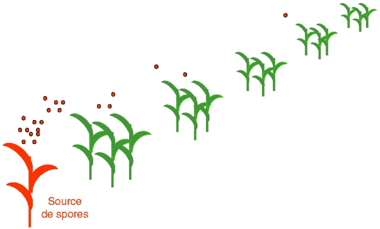
Let's mark out group dispersal for this talk

Definition (Soubeyrand et al., 2015)

Group dispersal occurs when groups of individuals/propagules start movement from the same place and time, travel following correlated paths and then settle at positively correlated locations (i.e. at nearby locations)

- We will only consider the dispersal of windborne propagules
- Propagules forming a group have the same time and point of origin
- A group is not only a cluster of propagules observed on the ground as the result of several dispersal events from several parental sources

Dispersal of windborne propagules



Sources of particles generate a spatially structured *rain* of particles

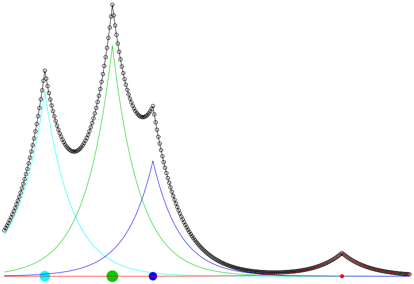
- rain of particles $\rightarrow$ spatial point process
- spatial structure $\rightarrow$ inhomogeneous intensity of the process

Examples: fungal spores, pollen grains, seeds

Intensity of the spatial point process formed by the deposit locations of the particles

The intensity is often obtained by a convolution between

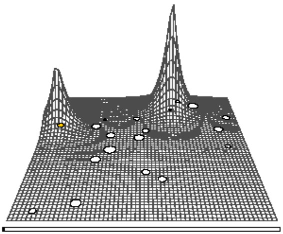
- the source process (spatial pattern and strengths) and
- a parametric dispersal kernel

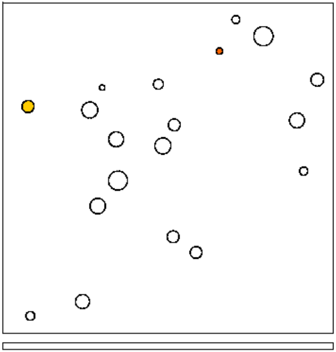


Intensity of the spatial point process formed by the deposit locations of the particles

The intensity is often obtained by a convolution between

- the source process (spatial pattern and strengths) and
- a parametric dispersal kernel





Dispersal kernel

Definition

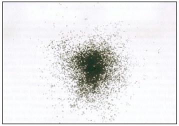
The *dispersal kernel* is the probability density function of the deposit location of a particle released at the origin

The shape of the kernel is a major topic in dispersal studies: it partly determines

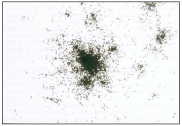
- the propagation speed
- the spatial structure of the population
- the genetic structure of the population

Main characteristics of dispersal kernels:

- short- or long-distance dispersal (Minogue, 1989)

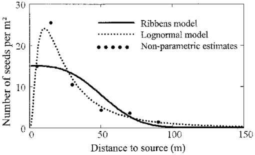


Noyau exponentiel

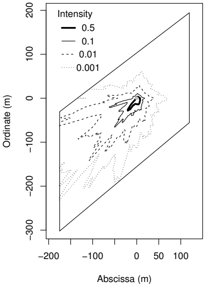


Noyau de Pareto

- possible non-monotonicity (Stoyan and Wagner, 2001)



- anisotropy (Soubeyrand et al., 2007-2008)



Beyond models with i.i.d. deposit locations

- In independent dispersal models (IDM), deposit locations of particles released by a given source are i.i.d. under the dispersal kernel

  Therefore, the set of deposit locations $\mathcal{X}$ satisfies:

  $$\mathcal{X} \mid \{x_i, \lambda_i : i \in \mathcal{I}\} \sim \text{Poisson Point Process}\left(x \mapsto \sum_{i \in \mathcal{I}} \lambda_i f(x - x_i)\right)$$

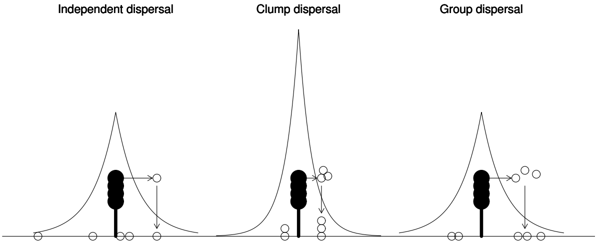
  $\mathcal{I}$ = set of particle sources $\qquad$ x_i = location of source i

  f = dispersal kernel $\qquad$ λ_i = strength of source i

- In contrast, group dispersal models (GDM) incorporate dependencies in deposit locations of particles

Contents of the talk

- Construction of group dispersal models and resulting dependencies between deposit locations of particles



- Implications of group dispersal on population dynamics and evolution

The Group Dispersal Model (GDM): A Neyman-Scott point process with double inhomogeneity

Soubeyrand, Roques, Coville and Fayard (2011)

Deposit equation for particles

- Single point source of particles located at the origin of $\mathbb{R}^2$
- J: number of groups of particles released by the source
- N_j: number of particles in group $j \in \{1, \dots, J\}$
- X_{jn}: deposit location of the n-th particle of group j satisfying

$$X_{jn} = X_j + B_{jn}(\nu||X_j||),$$

where X_j: final location of the "center" of group j,
B_{jn}: isotropic Brownian motion describing the relative movement of the n-th particle in group j with respect to X_j
ν: positive parameter governing the dislocation of groups

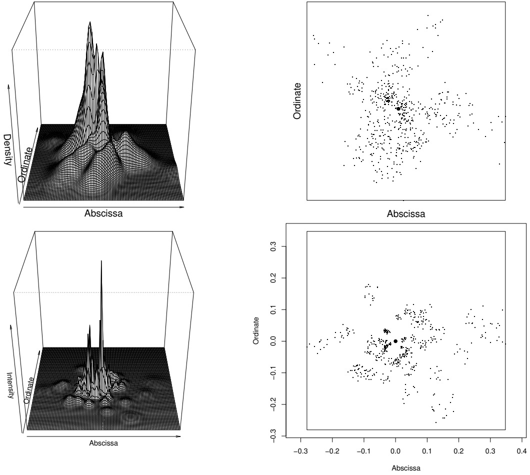
Assumptions about the deposit equation

- The random variables J, N_j, X_j and the random processes $\{B_{jn} : n = 1, \dots, N_j\}$ are mutually independent
- Number of groups: $J \sim \text{Poisson}(\lambda)$
- Number of particles in group j: $N_j \overset{\text{indep.}}{\sim} p_{\mu,\sigma^2}(\cdot)$
- Group center location: $X_j \overset{\text{indep.}}{\sim} f_{X_j}(\cdot)$
  (features of f_{X_j}: decrease at the origin is more or less steep, tail more or less heavy, shape more or less anisotropic...)
- The Brownian motions B_{jn} are centered, independent and with independent components
  They are stopped at time $t = \nu||X_j||$
  Then,

$$B_{jn}(\nu||X_j||) \overset{\text{indep.}}{\sim} \text{Normal}(0, \nu||X_j||I)$$

Dispersal from a single source under the GDM

- Simulations:



- This model can be viewed as a Neyman-Scott point process with double inhomogeneity (in the center pattern and in the offspring diffusion) or a non-stationary Cox point process

Discrepancies from independent dispersal

- Marginal probability density function (dispersal kernel):

$$f_{X_{jn}}(x) = \int_{\mathbb{R}^2} f_{X_{jn}|X_j}(x \mid y) f_{X_j}(y) dy = \int_{\mathbb{R}^2} \phi_{\nu,y}(x) f_{X_j}(y) dy.$$

- The particles are n.i.i.d. (not independently but identically distributed) from this p.d.f. while in the classical dispersal models the particles are i.i.d. from a dispersal kernel
- The Group Dispersal Model (GDM) is compared with the independent dispersal model (IDM1) having the same marginal dispersal kernel
- IDM1: the number of particles in each group is assumed to be one $\Rightarrow$ particles are independently drawn under the p.d.f. $f_{X_{jn}}$

Moments analysis and spatial structure of the population

X: Deposit location of a particle

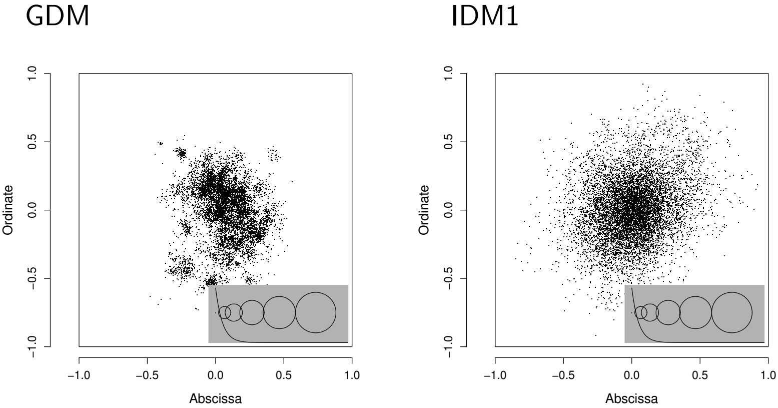
$Q(x+dx)$: Count of points in $x+dx$

Criterion	Model	Value
$E(X)$	GDM	$\binom{0}{0}$
	IDM1	$\binom{0}{0}$
$V(X)$	GDM	$V(X_j)+\nu E(\|\|X_j\|\|)I$
	IDM1	$V(X_j)+\nu E(\|\|X_j\|\|)I$
$E(\|\|X\|\|^2)$	GDM	$E(\|\|X_j\|\|^2)+2\nu E(\|\|X_j\|\|)$
	IDM1	$E(\|\|X_j\|\|^2)+2\nu E(\|\|X_j\|\|)$
$E\{Q(x+dx)\}$	GDM	$\lambda\mu f_{X_{jn}}(x)dx$
	IDM1	$\lambda f_{X_{jn}}(x)dx$
$V\{Q(x+dx)\}$	GDM	$\lambda[\mu f_{X_{jn}}(x)dx+(\sigma^2+\mu^2-\mu)E\{\phi_{\nu,X_j}(x)^2\}(dx)^2]$
	IDM1	$\lambda f_{X_{jn}}(x)dx$
$\mathrm{cov}\{Q(x_1+dx)$	GDM	$\lambda(\sigma^2+\mu^2-\mu)E\{\phi_{\nu,X_j}(x_1)\phi_{\nu,X_j}(x_2)\}(dx)^2$
$,Q(x_2+dx)\}$	IDM1	0

$\Rightarrow$ GDM: larger variance of $Q(x+dx)$ and positive covariance (decreasing with distance)

$\Rightarrow$ A characterization of clusters of particles under the GDM

Spatio-temporal simulations



⇒ Observation of multiple foci under the GDM

Farthest particle

Definition

The maximum dispersal distance in one generation is

$$R^{max} = \max\{R_{jn} : j \in \mathcal{J}, n \in \mathcal{N}_j\}$$

where $R_{jn} = ||X_{jn}||$
$\mathcal{J} = \{1, \ldots, J\}$ if $J > 0$ and the empty set otherwise
$\mathcal{N}_j = \{1, \ldots, N_j\}$ if $N_j > 0$ and the empty set otherwise

By convention, if no particle is released ($J = 0$ or $N_j = 0$ for all j), then $R^{max} = 0$

Distribution of R^{max}

$$R^{max} = \max\{R_{jn} : j \in \mathcal{J}, n \in \mathcal{N}_j\}$$

Under the GDM and IDMs, the distribution of the distance between the origin and the furthest deposited propagule is zero-inflated and satisfies:

$$P(R^{max} = 0) = \exp\left[\lambda\{p_{\mu,\sigma^2}(0) - 1\}\right]$$
$$f_{R^{max}}(r) = \lambda f_{R_j^{max}}(r)\exp\{\lambda(F_{R_j^{max}}(r) - 1)\}, \qquad \forall r > 0,$$

where $f_{R_j^{max}}$ is the p.d.f. of the distance $R_j^{max} = \max\{R_{jn} : n \in \mathcal{N}_j\}$ between the origin and the furthest deposited propagule of group j, and $F_{R_j^{max}}$ is the corresponding cumulative distribution function ($F_{R_j^{max}}(r) = P(R_j^{max} = 0) + \int_0^r f_{R_j^{max}}(u)du$).

$\rightarrow$ Distribution of R_j^{max} ?

Distribution of R_j^{max} under the IDM1

Under the IDM1, $N_j = 1$ for all $j \in \mathcal{J}$ and, consequently, $p_{\mu,\sigma^2}(0) = 0$ and

$$
\begin{aligned}
f_{R_j^{max}}(r) &= f_{R_{jn}}(r) \\
&= \int_0^{2\pi} r f_{X_{jn}}((r\cos\theta, r\sin\theta))d\theta
\end{aligned}
$$

Distribution of R_j^{max} under the GDM

Under the GDM, the distribution of R_j^{max} is zero-inflated and satisfies:

$$
\begin{aligned}
P(R_j^{max} = 0) &= p_{\mu,\sigma^2}(0) \\
f_{R_j^{max}}(r) &= \int_{\mathbb{R}^2} f_{R_j^{max}|X_j}(r \mid x) f_{X_j}(x) dx \\
&= \sum_{q=1}^{+\infty} q p_{\mu,\sigma^2}(q) \int_{\mathbb{R}^2} f_{R_{jn}|X_j}(r \mid x) F_{R_{jn}|X_j}(r \mid x)^{q-1} f_{X_j}(x) dx,
\end{aligned}
$$

where $f_{R_{jn}|X_j}$ is the conditional distribution of R_{jn} given X_j satisfying:

$$
\begin{aligned}
f_{R_{jn}|X_j}(r \mid x) &= 2r \int_0^{r^2} h_1(u,x) h_2(r^2 - u, x) du, \\
h_i(u,x) &= \frac{f_i(\sqrt{u}, x) + f_i(-\sqrt{u}, x)}{2\sqrt{u}}, \quad \forall i \in \{1,2\}, \\
f_i(v,x) &= \frac{1}{\sqrt{2\pi\nu||x||}} \exp\left(-\frac{(v - x^{(i)})^2}{2\nu||x||}\right), \quad \forall i \in \{1,2\},
\end{aligned}
$$

$x = (x^{(1)}, x^{(2)})$ and $F_{R_{jn}|X_j}(r \mid x) = \int_0^r f_{R_{jn}|X_j}(s \mid x) ds$.

Farthest particle and spatial structure of the population

Expressions of the distribution of R^{max} for the GDM and the IDM1 lead to:

Theorem

Consider a GDM and an IDM1 characterized by the same parameter values except that:
- for the GDM: $E(J) = \tilde{\lambda}$, $E(N_j) = \tilde{\mu}$ and $V(N_j) = \sigma^2$,
- for the IDM1: $E(J) = \tilde{\lambda}\tilde{\mu}$, $E(N_j) = 1$ and $V(N_j) = 0$,
($\Rightarrow$ same marginal dispersal kernel).
Then, for all $r > 0$ the probability $P(R^{max} \geq r)$ is lower for the GDM than for the IDM1.

Interpretation: The population of particles is less concentrated in probability for the IDM1 than for the GDM

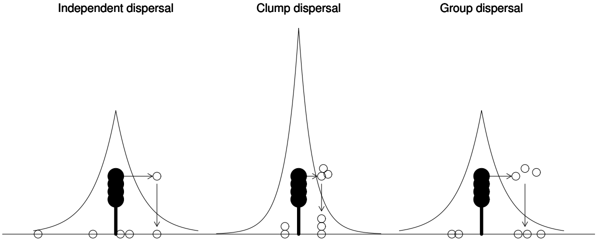
Consequence: With group dispersal, one can generate multiple foci whereas the particles are more concentrated

Evolution between independent, clump and group dispersal

Soubeyrand, Sache, Hamelin and Klein (2015)

Three dispersal strategies:

- **I** variants: **independent** movements of all propagules
- **C** variants: **clumps** of propagules sticked together and settling at the same location
- **G** variants: **groups** of propagules simultaneously released and settling at different but correlated locations



Question: how limits and fragmentation of the habitat shape the frequencies of **I**, **C** and **G** variants?

Model

Approximately the same spatio-temporal model than above

Except:

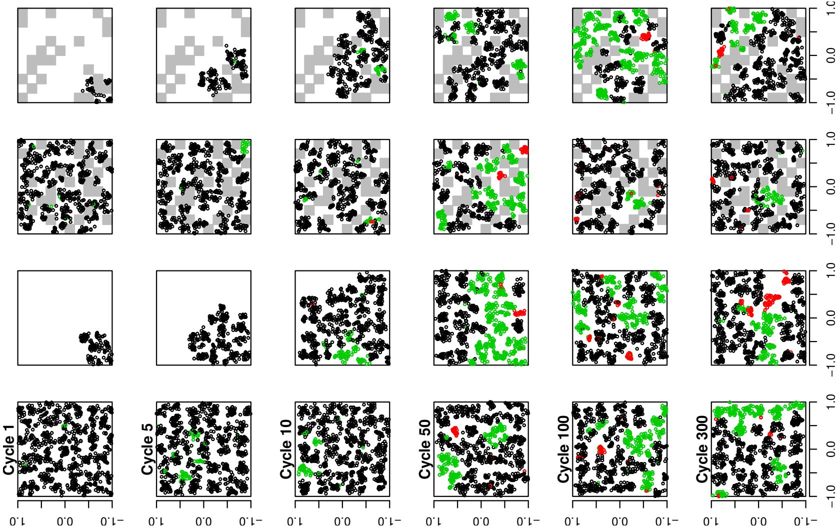
- f_{X_j} is an isotropic case of the normal inverse Gaussian (NIG) dispersal kernel (Klein et al., 2003): for $x \in \mathbb{R}^2$,

$$f_{X_j}(x) = \frac{\delta^2 e^\tau}{2\pi} \frac{\{(1+\delta^2||x||^2)^{-1/2} + \tau\}}{1+\delta^2||x||^2} \exp\{-\tau(1+\delta^2||x||^2)^{1/2}\},$$

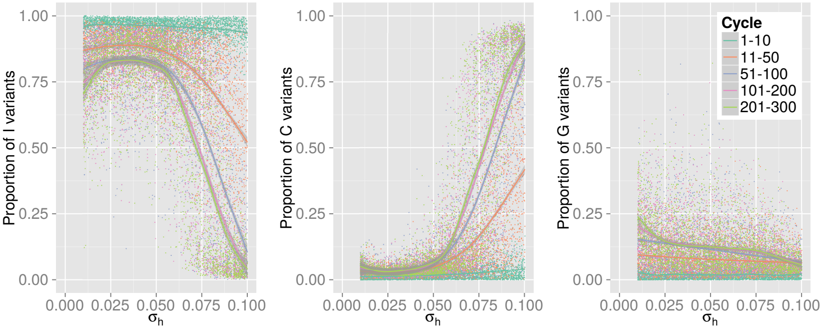
  f_{X_j} includes a settling velocity parameter that depends on the mass and the volume of the dispersed entities
  $\Rightarrow$ clumps will disperse at shorter distances
- Incorporation of an evolutionary process for the dispersal strategy
  - Evolution between **I**, **C** and **G** variants
  - Evolution of the distribution of the clump/group size
- Incorporation of a density-dependence

Simulations of the demo-genetic model

4 settings: invasive/endemic dynamic × fragmented/continuous habitat (**I** variants; **C** variants; **G** variants)

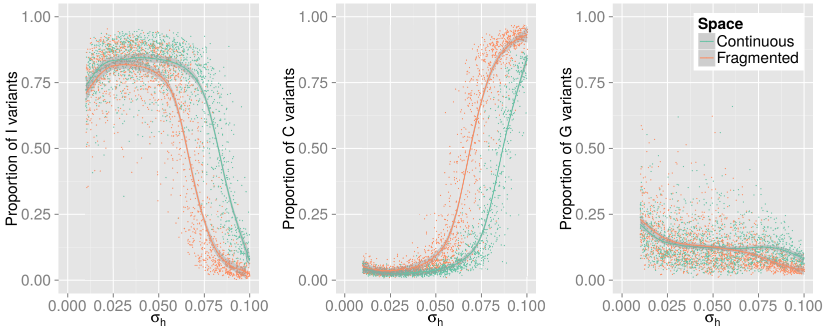


Co-existence of the three dispersal strategies



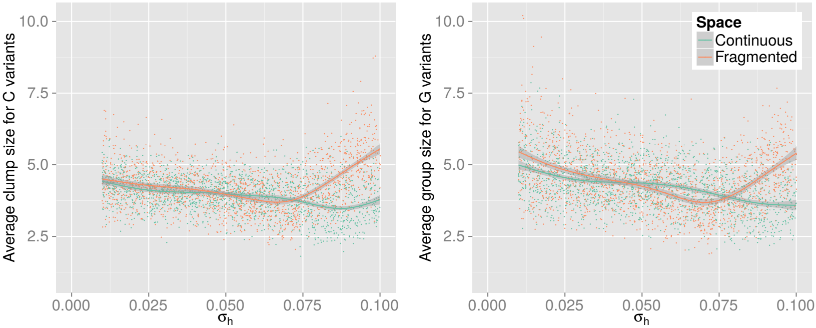
- Increasing the horizontal turbulence parameter σ_h (i.e. increasing the mean dispersal distance) advantages **C** variants at the expense of **I** variants
- **G** variants are never dominant but less affected than **I** variants by the increase of **C** variants with σ_h
- Spatial heterogeneity in dispersal strategies: **C** and **G** variants tend to be located closer to the habitat borders than **I** variants

Effect of fragmentation on the frequencies of variants



- Fragmentation is in favor of **C** variants

Effect of fragmentation on the clump and group sizes



- Similar group and clump sizes in each setting "$\sigma_h \times$ fragmentation"
- For high values of σ_h, larger sizes in fragmented habitats
- Optimal clump and group sizes: trade-off determined by the density-dependence and the probability for propagules released by **C** and **G** variants to be deposited outside the habitat

Group dispersal and metapopulation dynamics

Soubeyrand and Laine (2017)

- Markovian random walk for the size of a metapopulation whose dynamics includes group dispersal and Allee effect

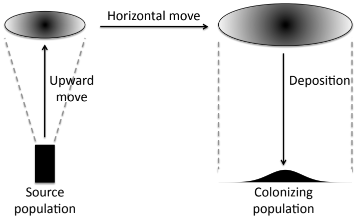
$\Rightarrow$ Discrete-time Stochastic Patch Occupancy Model (SPOM) in the metapop. terminology

- Question: How the interaction between group dispersal and Allee effect shape the metapopulation dynamics?

Model skeleton

- Migration/destruction event: a source population is simultaneously affected by:
  - the migration of a fraction λ_m of the population
  - the destruction of a fraction λ_d of the population

  (because of an extreme weather event for example)



- Random dispersal distance $\Rightarrow$ Random diffusion of the colonization population
- Allee effect: threshold for survival of source pop. and for emergence of colonizing pop.

Dynamic of the metapopulation size

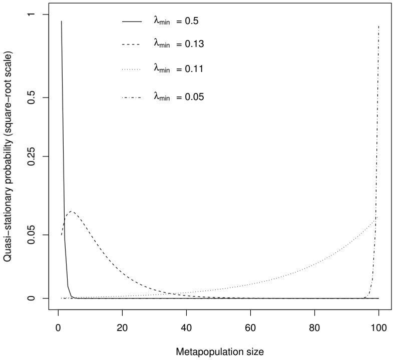
- N_i: metapopulation size after the migration/destruction event occurring at time t_i
- The process $\{N_i\}_i$ is a Markovian random walk
- Specifying the model allows us to provide explicit forms for transition probabilities:

$$P(N_{i+1} = N_i + 1)$$
$$P(N_{i+1} = N_i - 1)$$
$$P(N_{i+1} = N_i)$$

Quasi-stationary distribution and extinction time

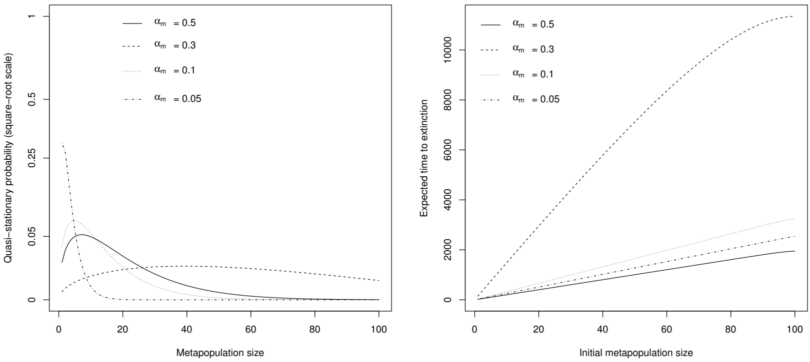
- In general, after a sufficiently large time, any metapopulation governed by the transition probabilities provided above vanishes
- Quasi-stationary distribution of the metapopulation size: conditional probability distribution for the size of the metapopulation given that the metapopulation has not reached extinction
- Expected time to extinction: average number of migration/destruction events that leads to the extinction of the metapopulation, given the initial size of the metapopulation

Effect of the Allee threshold on the quasi-distribution of the metapopulation size



- Smaller the Allee threshold, larger the population size in the quasi-stationary state

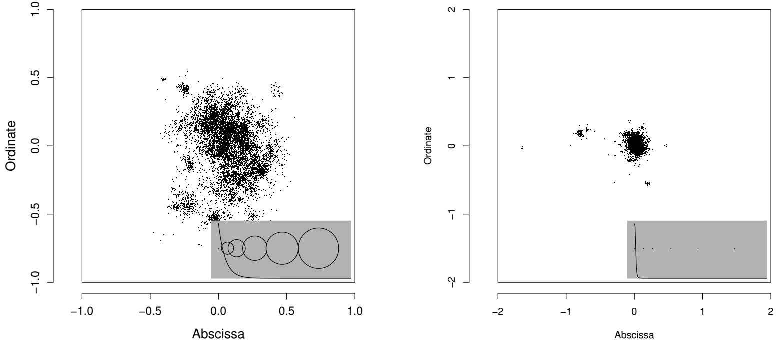
Effect of the maximum migrating fraction α_m on the quasi-distribution of the metapopulation size and the expected time to extinction



- Small α_m: $\approx$ metapopulation dynamic without group dispersal
- Non-monotonous effect of α_m: intermediate values of α_m lead to larger and more sustainable metapopulations
- Heuristically, large values of α_m plays against the survival of SP; small values of α_m plays against the emergence of CP

Conclusions

- Group dispersal is encountered in many cases ... but dispersal models generally assume independent transports of particles
- Group dispersal generates patterns with multiple foci
  $\rightarrow$ this is obvious!
- Group dispersal generates patterns with multiple foci whereas the population is more concentrated
  $\rightarrow$ remarkable difference b/n group and long-distance dispersal



Conclusions

- Clump dispersal (a special case of group dispersal) is favored by habitat fragmentation and limits
- A theoretical study of group dispersal in fragmented habitat: Soubeyrand, Mrkvička and Penttinen (2014)
- Intermediate group sizes may lead to larger and more sustainable metapopulations in the presence of an Allee effect

- Perspective: Fitting group dispersal model to data and testing independent *vs* group dispersal
  $\rightarrow$ MCMC for Neyman-Scott point processes with double inhomogeneity (Mrkvička and Soubeyrand, 2017)

Example of dynamic with eventual group dispersal

Wheat yellow rust epidemic in an experimental field
(Sache and Schermesser)