Piecewise-deterministic Markov processes for spatio-temporal population dynamics

Samuel Soubeyrand
Biostatistics and Spatial Processes research unit

11th SSIAB workshop – Rennes – May 27, 2016

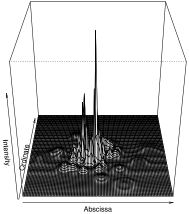
Spatio-temporal population dynamics

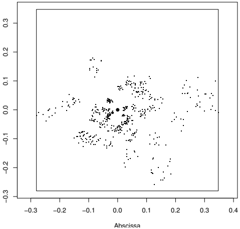
- Population dynamics: vast topic
  - of particular interest in ecology and epidemiology
  - studied at various scales, from the microscopic scale to the global scale

- Huge diversity of modeling approaches, e.g.:
  - Diffusion
  - Trajectory
  - Branching process
  - Point process
  - Areal process
  - Regression
  - etc.

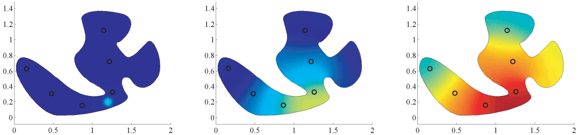
Classical approaches in the BioSP research group

- At BioSP: frequent use/construction of (quasi-)mechanistic models for population dynamics
- Typical examples of models:
  - Models based on non-stationary Neyman-Scott point processes





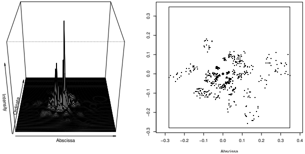
  - Models based on reaction-diffusion equations



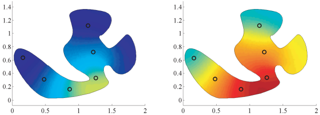
- Trade-off b/n model realism and estimation complexity

Spatio-temporal PDMP: the missing link for modeling population dynamics

- Extreme 1: models with stochastic behavior and lots of degrees of freedom



- Extreme 2: models with deterministic behavior and a few degrees of freedom

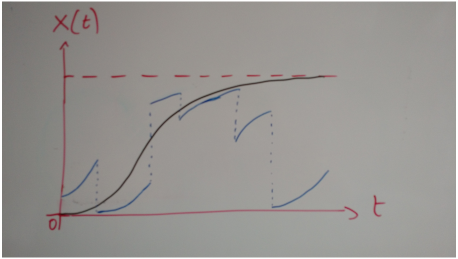


- **Need for intermediate models** to achieve rapid, realistic and consistent inference

→ **Spatio-temporal piecewise-deterministic Markov processes can play this role**

Piecewise-deterministic Markov process (PDMP)

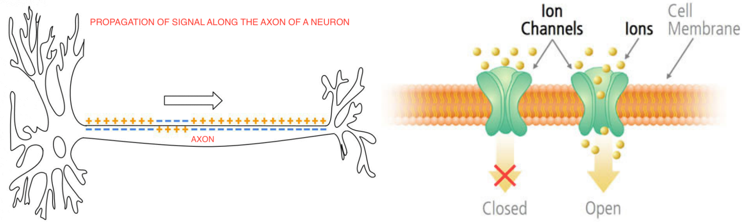
- PDMP: cadlag Markov process made of deterministic motion punctuated by random jumps
- A PDMP $\{X(t)\}_{t \geq 0}$ depends on:
  - a flow function $\phi : (x, t) \mapsto \phi(x, t)$
  - a jump rate function λ
  - a transition measure Q for jump amplitudes



- Challenges (for probabilists and statisticians):
  - Probabilistic properties, including asymptotic behavior
  - Estimation
  - Simulation

Examples of PDMP

- Numerous examples of temporal models
  - Insurance risk models
  - Queueing models
  - ...
- A spatio-temporal example: transmission by the axon of an electrical signal (see Azaïs et al., 2014)



  - Flow: local potential along the axon governed by a PDE
  - Jumps: opening and closing of ion channels
  - Joint modeling of flow and jumps (i.e. of local potential and channel states)

Contents of the presentation: short review of three spatio-temporal PDMPs in population dynamics

- Coalescing Colony Model
- Midge flight model
- Metapopulation epidemic model

A precursory example of PDMP in population dynamics: the Coalescing Colony Model (Shigesada et al., 1995)

- Flow: a colony forms a disk of radius r expanding in space at constant speed c → deterministic range expansion of colonies
- Jumps: new colonies are generated by an existing colony with rate $\lambda(r)$ and are located at distance L from the mother colony → stochastic generation of new colonies

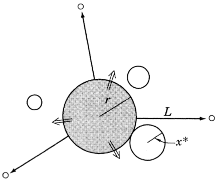


Fig. 5.5 Range expansion in the coalescing colony model. The primary colony (shaded circle) expands its range by short-distance dispersal and at the same time emits long-distance migrants at rate $\lambda(r)$, which settle at distance L ahead of the front of the primary population. Each offspring colony which expands at rate c coalesces with the primary population when their ranges overlap. Upon coalescence, the range of the primary colony including the offspring colony is immediately reshaped in a circular pattern, keeping the total area unchanged.

⇒ One obtains a spatio-temporal PDMP

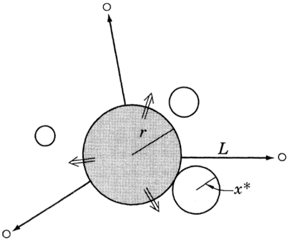


Fig. 5.5 Range expansion in the coalescing colony model. The primary colony (shaded circle) expands its range by short-distance dispersal and at the same time emits long-distance migrants at rate $\lambda(r)$, which settle at distance L ahead of the front of the primary population. Each offspring colony which expands at rate c coalesces with the primary population when their ranges overlap. Upon coalescence, the range of the primary colony including the offspring colony is immediately reshaped in a circular pattern, keeping the total area unchanged.

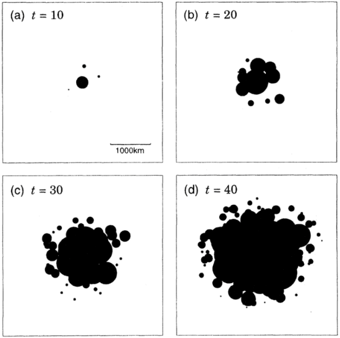
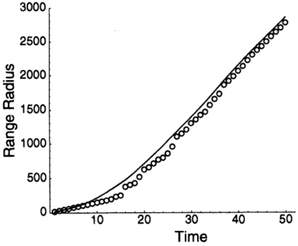


Fig. 5.9 Snapshots of range expansion from computer simulation of the coalescing colony model. Long-distance dispersal occurs stochastically at rate $\lambda(r) = 0.005r$, and its leap distance follows a truncated Gaussian distribution with mean $L = 300$ and standard deviation $\sigma = 40$, $c = 15$.

- The Coalescing Colony Model can also be viewed as a spatio-temporal Boolean process
- Shigesada et al. (1995) characterized, for instance, the temporal evolution of the range expansion of the total population



An example of spatio-temporal trajectory model (Girardin and Senoussi, in prep.)

- Girardin and Senoussi are interested in linking continuous- and discrete-time autoregressive processes with random coefficients and random jumps (CRARE and RARE, resp.)
    - How to discretize a CRARE process by conserving the autoregressive property?
    - How to embed a RARE process into a CRARE process?
- The links are made with so-called *aller-retour* formulas

Embedding theorem

Theorem

Let $\mathbf{y}$ *be a multivariate discrete time* $RARE_d(1)$ *process satisfying*

$$\mathbf{y}(n+1) = \mathbf{\Phi}(n)\mathbf{y}(n) + \mathbf{u}(n+1), \quad n \in \mathbb{N},$$

such that $\log \mathbf{\Phi}(n)$ *is well-defined for all* n.
Then $\mathbf{y}$ *can be embedded into a* $CRARE_d(1)$ *process* $\mathbf{x}$ *satisfying*

$$d\mathbf{x}(t) = \mathbf{A}(t)\mathbf{x}(t-)dt + d\mathbf{z}(t), \quad t \in \mathbb{R}_+,$$

where the flow is governed by the matrix

$$\mathbf{A}(t) = \sum_{n \geq 0} \log(\mathbf{\Phi}(n-1))\mathbf{1}_{[n,n+1[}(t)$$

and $\mathbf{z}$ *is a jump process with jumps* $\Delta\mathbf{z}_n = \mathbf{u}(n)$ *at times* $T_n^{\mathbf{z}} = n$.

⇒ One obtains a spatio-temporal PDMP

Application

- Consider the discrete-time equation:

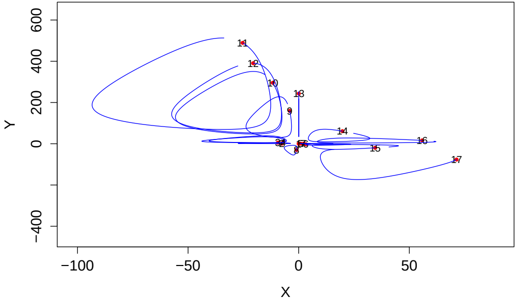
$$\mathbf{y}(n+1) = \mathbf{\Phi}(n)\mathbf{y}(n) + \mathbf{u}(n+1),$$

with

$\mathbf{\Phi}$ = exp(rotation matrix) in the plane

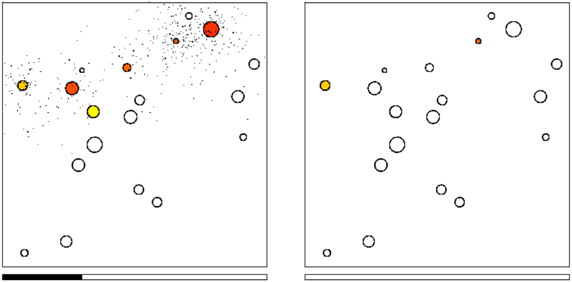
$\mathbf{u}$ = difference between two gamma-distributed 2D-vectors

- Simulation of the embedding continuous process, which is a PDMP (*midge flight model*):



An example of spatio-temporal metapopulation epidemic model (Soubeyrand, Laine, Hanski and Penttinen, 2009)

- Disks: host populations labelled by i
- Colored disks: infected host populations
- Flow: deterministic growth $t \mapsto g_i(t) = g(t - T_i)$ of the disease in infected populations (T_i: infection time for i)
- Infected host populations release contaminating particles (cluster point process) dispersed with the kernel h
- Jump: any particle deposited in a healthy population may generate an infection ($g_i(T_i^-) = 0$, $g_i(T_i) > 0$)



- Infections of populations (jumps) depend on a spatio-temporal point process governed by the inhomogeneous intensity:

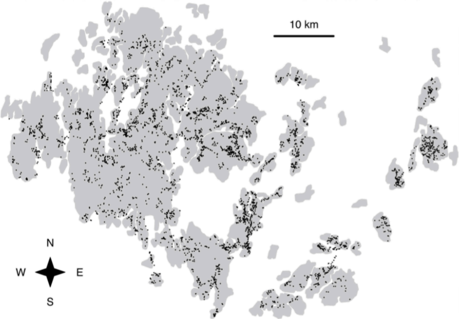
$$\lambda(t,x) = \sum_{j \in \mathcal{I}_t} c_j g(t - T_j) h(x - x_j)$$

where $t \mapsto g(t - T_j)$ gives the evolution of the infection strength of j, which is deterministic after T_j (flow), and h is the spatial dispersal kernel

$\Rightarrow$ One obtains a spatio-temporal PDMP

Application: inference of the dynamics of *Podosphaera plantaginis* in Åland archipelago

- Data:
  - Observation of sanitary states ($Y^{obs}_{n,i}$: healthy / infected) of populations at the end of successive epidemic seasons
  - Covariates Z_i



Bayesian estimation

- Estimation of model parameters and latent variables, e.g.:
  - parameters of the growth functions g_i
  - parameters of the dispersal kernel h
  - infection times T_i
- Joint posterior distribution:

$$
\begin{aligned}
p(\theta, \mathbf{T} \mid \mathbf{Y}_n^{obs}, \mathbf{Y}_{n-1}^{obs}, \mathbf{Z}) &\propto p(\mathbf{Y}_n^{obs} \mid \mathbf{T}, \theta, \mathbf{Y}_{n-1}^{obs}) p(\mathbf{T} \mid \theta, \mathbf{Y}_{n-1}^{obs}, \mathbf{Z}) \pi(\theta) \\
&= p(\mathbf{Y}_n^{obs} \mid \mathbf{T}) \pi(\theta) b(\theta, \mathbf{Y}_{n-1}^{obs}, \mathbf{Z}) \prod_{i \text{ healthy at } t_{end}} \exp\{-a_i \Lambda(t_{end}, x_i)\} \\
&\quad \times \prod_{i \text{ infected}} \exp\{-a_i \Lambda(T_i, x_i)\} \lambda(T_i, x_i)
\end{aligned}
$$

with $\Lambda(t, x) = \int_{t_0}^{t} \lambda(t, x) dt$

- MCMC

Posterior distributions of infection times (i.e. jump times) for a few populations

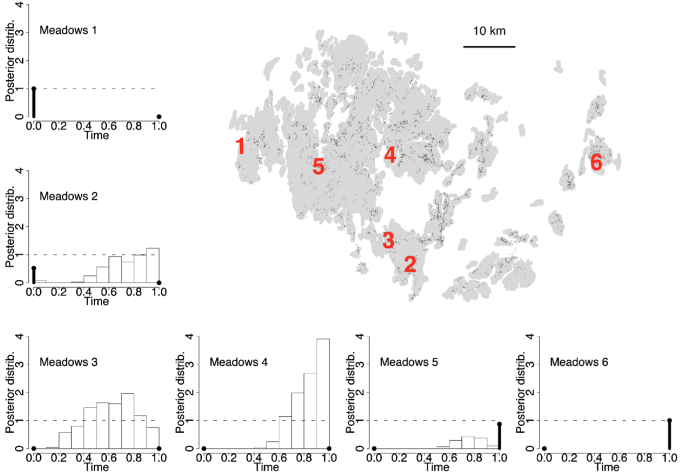
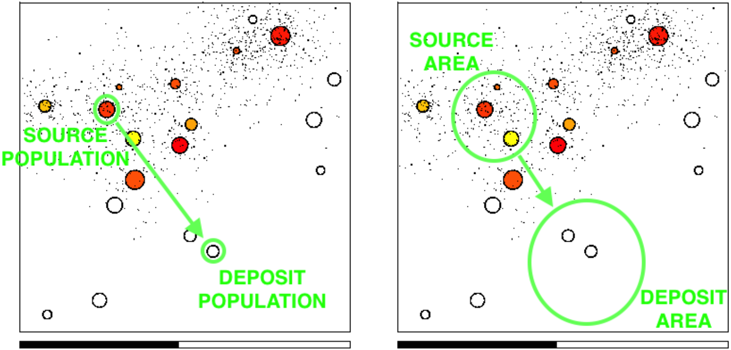


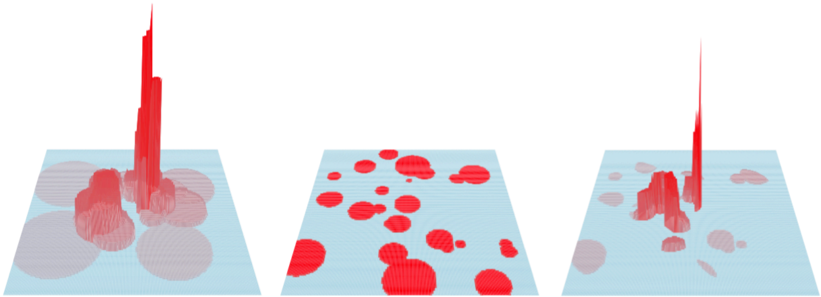
Figure C1: Posterior distributions of the infection times of six different meadows in the year 2004. In each plot, the dots at times 0 and 1 give the posterior probabilities that the infection time is 0 and 1, respectively. The horizontal dashed line corresponds to the uniform prior distribution that was used here. The map of the Åland Islands shows the locations of the six meadows.

Towards random jumps with spatial extents

- Random jumps in the epidemic model results from population–to–population dispersal
- Group dispersal leads us to consider PDMPs incorporating random jumps, which are correlated in space and time



Spatio-temporal correlation of infections

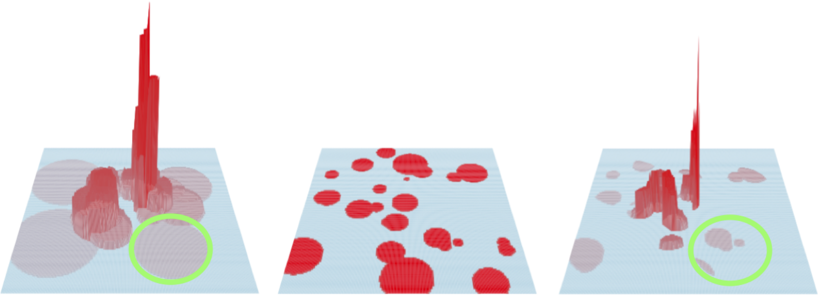


Dispersal

Populations

Dispersal × populations

Figure from Soubeyrand, Mrkvička and Penttinen (2014)



Dispersal Populations Dispersal × populations

Figure from Soubeyrand, Mrkvička and Penttinen (2014)

Incorporating area–to–area dispersal

- Random jump: dispersal from a set of aggregated patches $\mathcal{J}$ to a set of aggregated patches $\mathcal{I}$

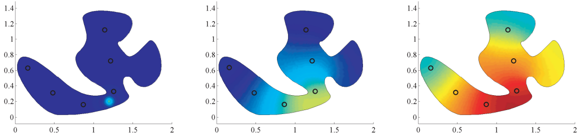
$$g_i(t) = g_i(t^-) + \Delta_{\mathcal{J}i}(\{c_j g_j(t) : j \in \mathcal{J}\}), \qquad \forall i \in \mathcal{I}$$

- Challenge: defining such a jump process yielding to a tractable posterior

Perspectives

Enriching the class of spatio-temporal PDMPs
for population dynamics

Building spatio-temporal PDMPs based on PDE
which are more “continuous” than the Coalescing Colony Model
(properly defining the jump process)



Developing simulation and estimation for these models