Piecewise-deterministic Markov processes for spatio-temporal population dynamics

Samuel Soubeyrand and Rachid Senoussi
Biostatistics and Spatial Processes research unit

Workshop “Statistique pour les PDMP”
Nancy – February 2-3, 2017
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

- Examples:
  - Dynamics of bluefin tuna in the Mediterranean see
  - Invasion of Europe by the Asian predatory wasp
  - Recurrence of the avian flu in Europe

- Huge diversity of modeling approaches, e.g.:
  - Diffusion
  - Trajectory
  - Branching process
  - Point process
  - Areal process
  - Regression
  - etc.
(Quasi-)mechanistic models for population dynamics

- Models based on reaction-diffusion equations
  - Aggregated model

- Models based on spatio-temporal point processes
  - Individual-based model

- 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
Contents of the presentation

- Coalescing Colony Model
- Metapopulation epidemic model
- Trajectory models from auto-regressive processes
A precursory example of PDMP in population dynamics: the Coalescing Colony Model (Shigesada et al., 1995)
Modeling stratified diffusion in biological invasions

- Biological invasions may be driven by various modes of dispersal
  - Ex.: Stratified dispersal process
    - neighborhood diffusion
    - long-distance dispersal

- Impact of long-distance dispersal: acceleration of range expansion
Range expansion by neighborhood diffusion

- Diffusion equation with a Malthusian growth term (Skellam, 1951)

\[
\frac{\partial n}{\partial t} = D \left( \frac{\partial^2 n}{\partial x^2} + \frac{\partial^2 n}{\partial y^2} \right) + \epsilon n
\]

- \( n((x, y), t) \): local population density at location \((x, y)\) and time \(t\)
- \( D \): diffusion coefficient
- \( \epsilon \): intrinsic growth rate of the population

**Property**

The rate of spread at the front of the population range asymptotically approaches \(2\sqrt{\epsilon D}\) when a small population is initially introduced at the origin.
Change with time in the population density:

Establishing phase followed by a constant rate spread
The property above is robust to some modifications of the growth term

Ex.: Diffusion equation with a logistic growth term (Fisher-KPP)

\[
\frac{\partial n}{\partial t} = D \left( \frac{\partial^2 n}{\partial x^2} + \frac{\partial^2 n}{\partial y^2} \right) + \epsilon (1 - n)n 
\]
Invasion by stratified diffusion

- Homogeneous environment
- Invading species expanding its range by both neighborhood diffusion and long-distance dispersal
- Simple approximation of Skellam or Fisher-KPP equations augmented by long-distance dispersal:

\[ c = 2\sqrt{\epsilon D} \]: constant rate expansion
\[ \lambda(r) \]: rate of generation of new colonies by a colony with radius \( r \)

- the establishing phase is neglected
Coalescing Colony Model

- **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

⇒ One obtains a spatio-temporal PDMP
Shigesada et al. (1995) characterized the variation in the range expansion $\tilde{r}(t)$ of the total population

- $\lambda(r) = \lambda_0 \Rightarrow \tilde{r}(t)$ constant
- $\lambda(r) = \lambda_0 r \Rightarrow \tilde{r}(t)$ bi-phasic
- $\lambda(r) = \lambda_0 r^2 \Rightarrow \tilde{r}(t)$ continually accelerates

Fig. 4.—Three types of range-versus-time curves. Range expansion patterns commonly have an establishment phase (arrow), expansion phase (solid line), and saturation phase (dashed line), successively. The expansion phase is classified into three types. Type 1 shows linear expansion. Type 2 exhibits biphasic expansion, with an initial slow slope followed by a steep linear slope. In type 3, the rate of expansion continually increases with time.

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 $\sigma = 40$, $c = 15$. 
Bayesian inference for a PDMP modeling the dynamics of a metapopulation

(Soubeyrand, Laine, Hanski and Penttinen, 2009)
A metapopulation epidemic model viewed as a PDMP

- Disks: host populations labelled by $i$
- Colored disks: infected host populations
- Points: contaminating particles released by infected hosts and dispersed with kernel $h$ (cluster point process)

- Flow: deterministic growth $t \mapsto g_i(t) = g(t - T_i)$ of the disease in infected populations ($T_i$: infection time for $i$)
- Jump: particles deposited in healthy populations may generate new infections ($g_i(T_i^-) = 0$, $g_i(T_i) > 0$)
Infections of populations (jumps) depend on a spatio-temporal point process governed by the inhomogeneous intensity:

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

where

- $t \mapsto c_j g(t - T_j)$ gives the evolution of the infection strength of $j$, which is deterministic after $T_j$ (flow),
- $h$ 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 ($Y_{n,i}^{obs}$: healthy / infected / NA) of populations at the end of successive epidemic seasons
  - Covariates $Z_i$
Application: inference of the dynamics of *Podosphaera plantaginis* in Åland archipelago

- **Data:**
  - Observation of sanitary states ($Y_{n,i}^{obs}$: healthy / infected / NA) 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:

$$p(\theta, T | Y^{obs}_n, Y^{obs}_{n-1}, Z) \propto p(Y^{obs}_n | T, \theta, Y^{obs}_{n-1}) p(T | \theta, Y^{obs}_{n-1}, Z) \pi(\theta)$$

$$= p(Y^{obs}_n | T) \pi(\theta) b(\theta, Y^{obs}_{n-1}, Z) \prod_i \exp\{ -a_i \Lambda(t_{\text{end}}, x_i) \}$$

$$\times \prod_i \exp\{ -a_i \Lambda(T_i, x_i) \} \lambda(T_i, x_i)$$

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

- MCMC
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:

$$p(\theta, T \mid Y_n^{obs}, Y_{n-1}^{obs}, Z) \propto p(Y_n^{obs} \mid T, \theta, Y_{n-1}^{obs})p(T \mid \theta, Y_{n-1}^{obs}, Z)\pi(\theta)$$

$$= p(Y_n^{obs} \mid T)\pi(\theta)b(\theta, Y_{n-1}^{obs}, Z) \prod_{i \text{ healthy at } t_{end}} \exp\{-a_i\Lambda(t_{end}, x_i)\}$$

$$\times \prod_{i \text{ infected}} \exp\{-a_i\Lambda(T_i, x_i)\}\lambda(T_i, x_i)$$

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

- MCMC
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:

$$p(\theta, T \mid Y_{n}^{obs}, Y_{n-1}^{obs}, Z) \propto p(Y_{n}^{obs} \mid T, \theta, Y_{n-1}^{obs})p(T \mid \theta, Y_{n-1}^{obs}, Z)\pi(\theta)$$

$$= p(Y_{n}^{obs} \mid T)\pi(\theta)b(\theta, Y_{n-1}^{obs}, Z) \prod_{i \text{ healthy at } t_{end}} \exp\{-a_i\Lambda(t_{end}, x_i)\}$$

$$\times \prod_{i \text{ infected}} \exp\{-a_i\Lambda(T_i, x_i)\}\lambda(T_i, x_i)$$

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

- MCMC
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:

$$p(\theta, T \mid Y_{n}^{obs}, Y_{n-1}^{obs}, Z) \propto p(Y_{n}^{obs} \mid T, \theta, Y_{n-1}^{obs}) p(T \mid \theta, Y_{n-1}^{obs}, Z) \pi(\theta)$$

$$= p(Y_{n}^{obs} \mid T) \pi(\theta) b(\theta, Y_{n-1}^{obs}, Z) \prod_{i \text{ healthy at } t_{end}} \exp\{-a_i \Lambda(t_{end}, x_i)\}$$

$$\times \prod_{i \text{ infected}} \exp\{-a_i \Lambda(T_i, x_i)\} \lambda(T_i, x_i)$$

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.
Perspective 1: Towards random jumps with spatial extents

- Random jumps in the epidemic model results from independent population-to-population dispersal events.

- However, the random jumps could be correlated in space and time.
Incorporating area–to–area dispersal

- Random jump: dispersal from a set of aggregated patches $I$ to a set of aggregated patches $J$

\[ g_i(t) = g_i(t^-) + \Delta_{ji}(\{c_jg_j(t) : j \in J\}), \quad \forall i \in I \]

- Challenge: defining such a jump process yielding to a tractable posterior
Perspective 2: Fitting a PDE-based PDMP to epidemiological surveillance data

- Estimation of the introduction time and location of an invasive species

- Handling multiple introductions modeled as a Markov process
  - Example of objective: characterizing the inter-jump duration, and its eventual non-stationarity