Nonparametric mixture modeling for Poisson processes

Athanasios Kottas
Department of Applied Mathematics and Statistics
University of California, Santa Cruz

CRiSM workshop on Model Uncertainty
University of Warwick
June 1, 2010

Outline

1. Introduction

- Point processes are stochastic process models for events that occur separated in time or space — several applications: traffic engineering, software reliability, neurophysiology, weather modeling, forestry, ...

- Poisson processes, along with their extensions (Poisson cluster processes, marked Poisson processes, etc.), play an important role in the theory and applications of point processes (e.g., Kingman, 1993; Guttorp, 1995; Diggle, 2003; Møller & Waagepetersen, 2004)

Introduction

- Bayesian work includes methods based on:
  $\longrightarrow$ weighted gamma processes and Lévy processes(e.g., Lo & Weng, 1989; Wolpert & Ickstadt, 1998; Best et al., 2000; Gutiérrez-Peña & Nieto-Barajas, 2003; Ishwaran & James, 2004)
  $\longrightarrow$ piecewise constant functions driven by Voronoi tessellations (e.g., Heikkinen & Arjas, 1998, 1999)
  $\longrightarrow$ log-Gaussian Cox processes (e.g., Møller et al., 1998; Brix & Diggle, 2001; Liang, Carlin & Gelfand, 2009)

- **Objective:** develop a nonparametric mixture modeling framework for Poisson processes, including marked Poisson processes

2. Modeling and inference for Poisson processes

- For a non-homogeneous Poisson process (NHPP), defined on the observation window $\mathcal{R}$ with intensity $\lambda(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathcal{R}$, which is a non-negative and locally integrable function for all bounded $\mathcal{B} \subseteq \mathcal{R}$, the following hold true:
  - *i.* For any such $\mathcal{B}$, the number of points in $\mathcal{B}$, $N(\mathcal{B}) \sim \text{Poisson}(\Lambda(\mathcal{B}))$, where $\Lambda(\mathcal{B}) = \int_{\mathcal{B}} \lambda(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$ is the NHPP cumulative intensity function (mean measure)
  - *ii.* Given $N(\mathcal{B})$, the point locations within $\mathcal{B}$ are i.i.d. with density $\lambda(\boldsymbol{x}) / \int_{\mathcal{B}} \lambda(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$

- Consider first temporal NHPPs with $\mathcal{R} \subset \mathbb{R}^+$ — extensions to spatial NHPPs where $\mathcal{R} \subset \mathbb{R}^2$ (though $\mathcal{R}$ can be of arbitrary dimension) $\rightarrow$ assume bounded observation window, and without loss of generality, $\mathcal{R} = (0,1)$ ($\mathcal{R} = (0,1) \times (0,1)$ for spatial NHPPs)

Modeling and inference for Poisson processes

- Consider a NHPP observed over the time interval $\mathcal{R} = (0, 1)$ with events that occur at times $0 < t_1 < t_2 < ... < t_N < 1$
- The likelihood for the NHPP intensity function λ is proportional to

$$\exp\{-\int_0^1 \lambda(u)\mathrm{d}u\} \prod_{i=1}^{N} \lambda(t_i)$$

- **Key observation:** $f(t) = \lambda(t)/\gamma$, where $\gamma = \int_0^1 \lambda(u)\mathrm{d}u$, is a density function on $(0, 1)$
- Hence, (f, γ) provides an equivalent representation for λ, and so a nonparametric prior model for f, with a parametric prior for γ, will induce a semiparametric prior for λ — in fact, since γ only scales λ, it is f that determines the shape of the intensity function λ

- **Beta Dirichlet process mixture model** for f

$$f(t) \equiv f(t; G) = \int \mathrm{be}(t; \mu, \tau) \mathrm{d}G(\mu, \tau),$$

where $\mathrm{be}(t; \mu, \tau)$ is the Beta density on $(0, 1)$, with mean $\mu \in (0, 1)$ and scale parameter $\tau > 0$

- Dirichlet process (DP) prior for the mixing distribution, $G \sim \mathrm{DP}(\alpha, G_0)$

- Base distribution $G_0 \equiv G_0(\mu, \tau) = G_{01}(\mu) G_{02}(\tau)$
  $\rightarrow$ uniform distribution on $(0, 1)$ for $G_{01}(\mu)$
  $\rightarrow$ inverse gamma distribution (with random scale parameter) for $G_{02}(\tau)$

Modeling and inference for Poisson processes

- Full Bayesian model:

$$\exp(-\gamma)\gamma^N \left\{ \prod_{i=1}^{N} \int \mathrm{be}(t_i; \mu, \tau)\, \mathrm{d}G(\mu, \tau) \right\} p(\gamma) p(G \mid \alpha, \beta) p(\alpha) p(\beta)$$

→ DP prior structure $p(G \mid \alpha, \beta)p(\alpha)p(\beta)$ for G and its hyperparameters
→ reference prior for γ, $p(\gamma) \propto \gamma^{-1}$

- Letting $\boldsymbol{\theta} = \{(\mu_i, \tau_i) : i = 1, ..., N\}$, we have

$$p(\gamma, G, \boldsymbol{\theta}, \alpha, \beta \mid \text{data}) = p(\gamma \mid \text{data}) p(G \mid \boldsymbol{\theta}, \alpha, \beta) p(\boldsymbol{\theta}, \alpha, \beta \mid \text{data})$$

→ $p(\gamma \mid \text{data})$ is a gamma$(N, 1)$ distribution
→ MCMC with Metropolis steps to sample from $p(\boldsymbol{\theta}, \alpha, \beta \mid \text{data})$
→ $p(G \mid \boldsymbol{\theta}, \alpha, \beta)$ is a DP with updated parameters – can be sampled using the DP constructive definition

- Full posterior inference for λ, Λ, and any other functional of the NHPP

Modeling and inference for Poisson processes

- Flexible density shapes through mixing of Betas
- Alternative mixture model formulations for monotonic intensity functions
  - utilize the representation of non-increasing densities on $\mathbb{R}^+$ as scale mixtures of uniform densities
  - a prior model for non-increasing intensities:

$$\lambda(t;G) = \gamma \int \theta^{-1} \mathbf{1}_{t \in (0,\theta)} \mathrm{d}G(\theta), \;\; t \in (0,1)$$

    with $G \sim \mathrm{DP}(\alpha, G_0)$, where G_0 has support on $(0,1)$, e.g., it can be defined by a Beta distribution
  - similarly, for non-decreasing NHPP intensities:

$$\lambda(t;G) = \gamma \int \theta^{-1} \mathbf{1}_{(t-1) \in (-\theta,0)} \mathrm{d}G(\theta), \;\; t \in (0,1)$$

    with $G \sim \mathrm{DP}(\alpha, G_0)$, where again G_0 has support on $(0,1)$
- Logit-normal DP mixtures for more efficient posterior simulation

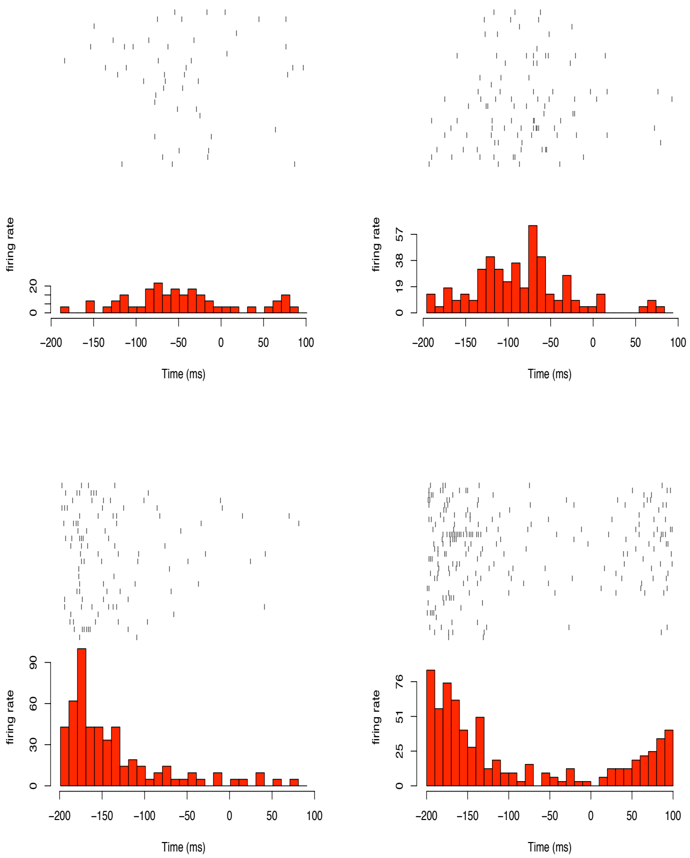
Application to neuronal data analysis

- One of the key techniques in neuroscience involves recording of electrical activity of neurons in laboratory animals
- The technique studies action potentials (spikes) generated by the neuron and measured using an electrode inserted into the animal's brain
- The firing times (times at which spikes occur) are recorded to provide the neuronal data
- In this context, the focus of statistical modeling approaches is on the temporal evolution of the neuronal firing activity
- Bayesian nonparametric modeling for neuronal data arising as firing times from a single neuron under two distinct experimental conditions (Kottas & Behseta, 2009)

Modeling and inference for Poisson processes

- Motivating neurophysiological study: neurons recorded from the primary motor cortex area (M1) of a Macaque monkey's brain while performing the sequential task of reaching a series of illuminating targets on a touch-sensitive screen (Matsuzaka, Picard and Strick, 2007)
- The animal was trained to respond to the visual stimuli under two conditions
  → *repeating* condition: a sequence of targets would appear on the screen in a repeating order
  → *random* condition: the stimuli were sent in a pseudo-random order
- Two neurons (29 and 32):
  → for **neuron 29**, 21 and 32 trials were recorded under the random and repeating mode, respectively, resulting in a total of 108 firing times for the random condition and 224 for the repeating condition
  → in both conditions, 20 trials were recorded for **neuron 32**, achieving a total of 52 and 102 firing times under the random and repeating condition

Modeling and inference for Poisson processes



Raster and PSTH plots for the firing times under neurons 32 and 29 (top and bottom row, respectively). The left panels correspond to the random condition, and the right panels to the repeating condition.

- Data from each neuron can be represented in their most general form through vectors $\{y_{ij}^{(\ell)} : i = 1, ..., M^{(\ell)}; j = 1, ..., m_i^{(\ell)}\}$, where $y_{ij}^{(\ell)}$ is the j-th firing time in the i-th trial under condition $\ell = 1, 2$

- Given our inferential objective of comparison of firing intensities under the two conditions, it suffices to consider modeling for the firing times aggregated over all trials — for condition $\ell = 1, 2$, the data vector

$$\boldsymbol{t}^{(\ell)} = \{t_i^{(\ell)} : i = 1, ..., N^{(\ell)}\}$$

  where $N^{(\ell)} = \sum_{i=1}^{M^{(\ell)}} m_i^{(\ell)}$ is the total number of firing times from all trials, and $t_i^{(\ell)}$ is the i-th spike time in the aggregated set of firing times

- NHPP a plausible model for the underlying point process of aggregated firing times (pooled point patterns across a large number of replicated trials follow approximately a NHPP model)

Modeling and inference for Poisson processes

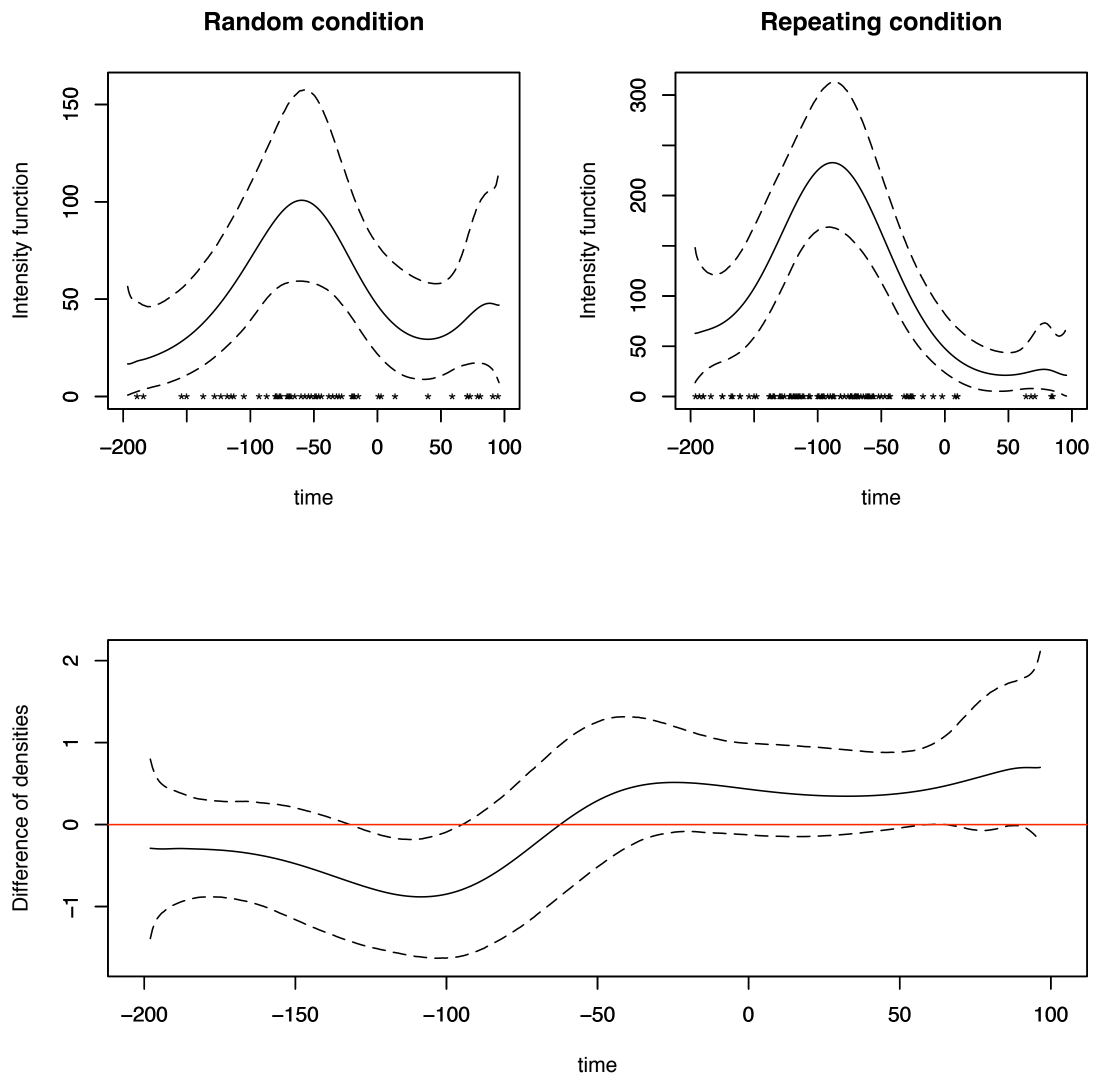
- NHPP with intensity function $\lambda^{(\ell)}(\cdot)$ under condition $\ell = 1, 2$
- Beta DP mixture model for firing intensities

$$\lambda^{(\ell)}(t) \equiv \lambda^{(\ell)}(t; \gamma^{(\ell)}, G^{(\ell)}) = \gamma^{(\ell)} f^{(\ell)}(t; G^{(\ell)}) = \int \mathrm{be}(t; \mu, \tau)\, \mathrm{d}G^{(\ell)}(\mu, \tau)$$

with (independent) DP priors for mixing distributions $G^{(1)}$ and $G^{(2)}$

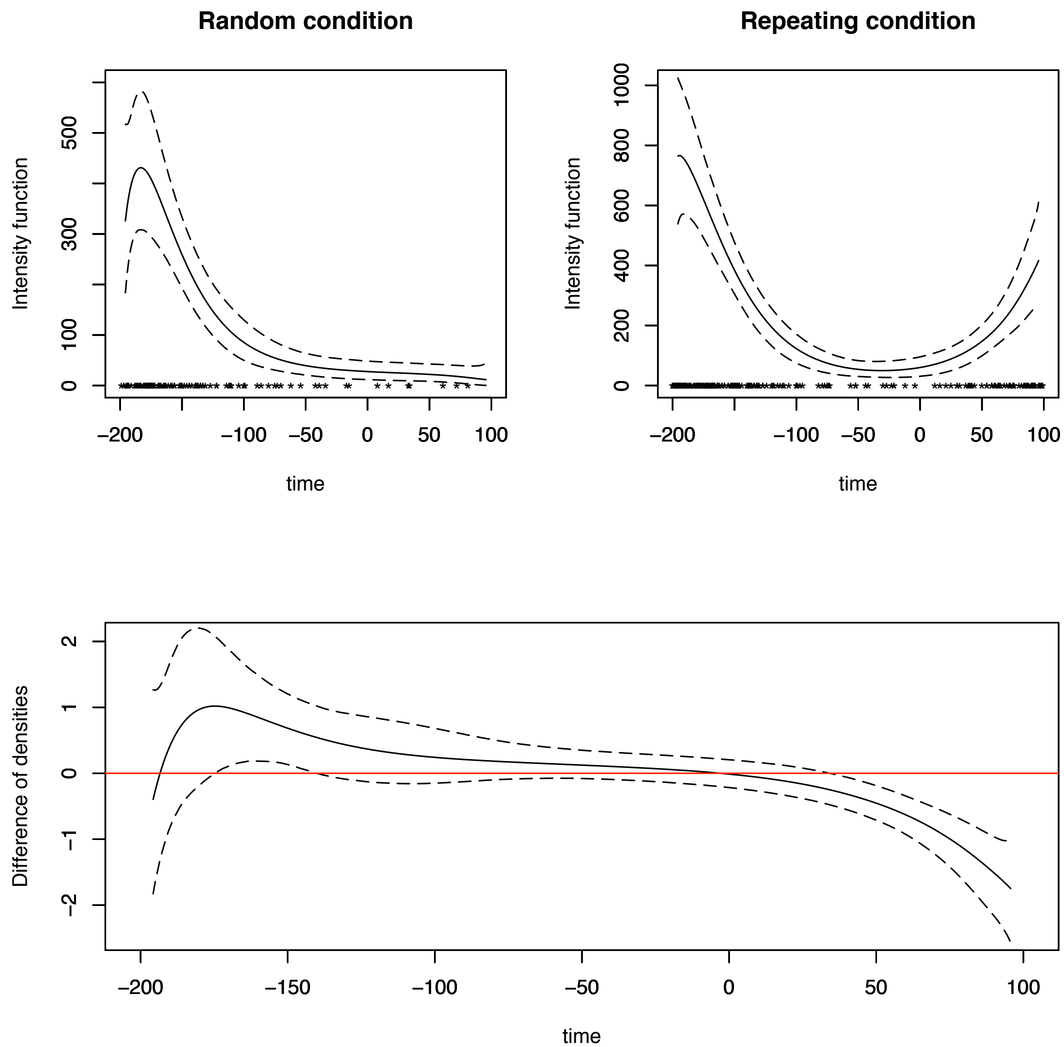
- Direct comparison of intensities $\lambda^{(1)}(\cdot)$ and $\lambda^{(2)}(\cdot)$ is hindered by their different scales, $\gamma^{(1)}$ and $\gamma^{(2)}$
- Work instead with densities $f^{(1)}(\cdot)$ and $f^{(2)}(\cdot)$
$\rightarrow$ posterior point and interval estimates for function $f^{(1)}(\cdot) - f^{(2)}(\cdot)$
$\rightarrow$ entire posterior distribution $p(f^{(1)}(t_0; G^{(1)}) - f^{(2)}(t_0; G^{(2)}) \mid \text{data})$ for specific points t_0 in the experimental time interval
- Inference for local and global differences in the neuronal firing intensities

Modeling and inference for Poisson processes



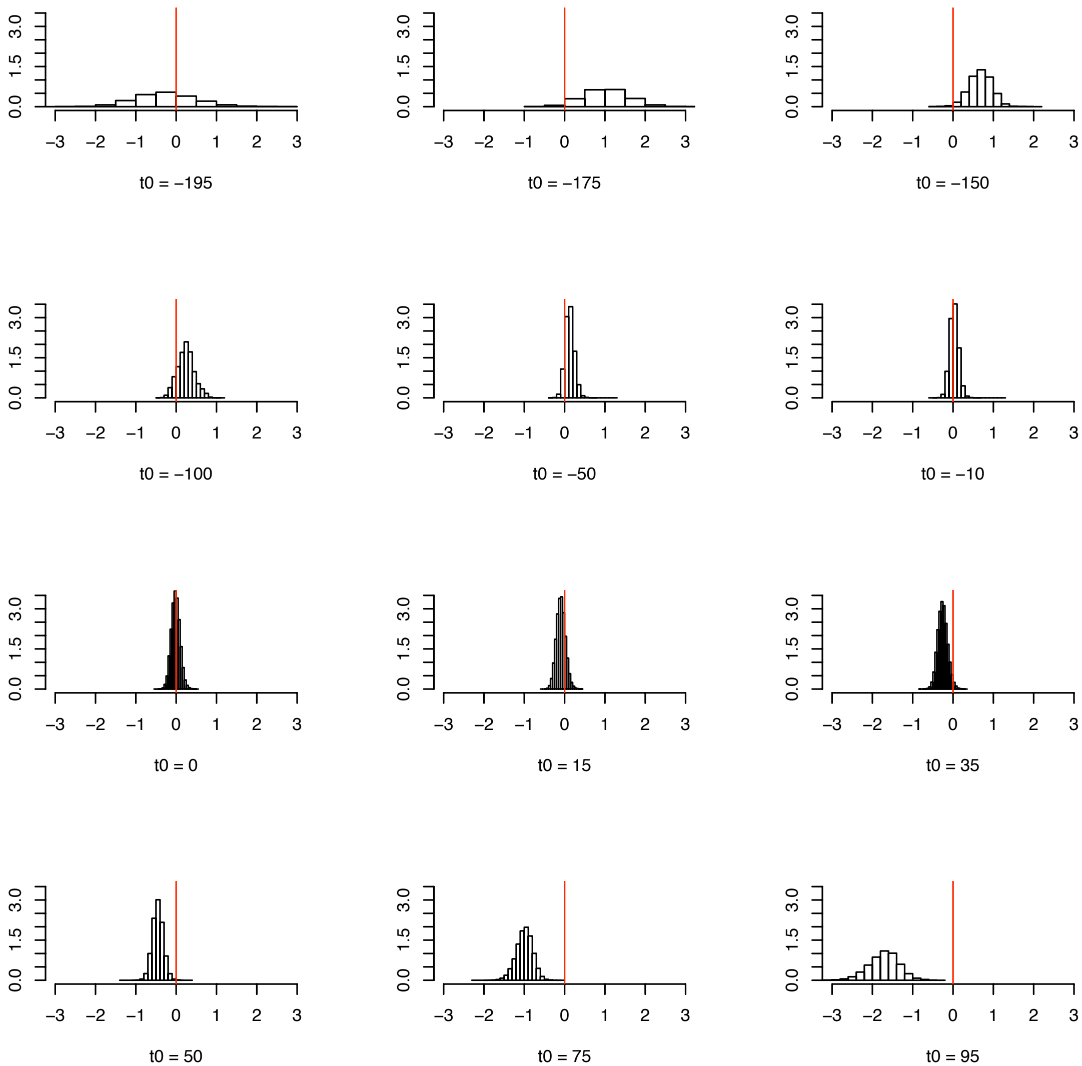
Neuron 32. Posterior mean and 95% interval estimates for the intensity functions. Posterior mean and 95% interval band for the difference of density functions between the random and repeating condition.

Modeling and inference for Poisson processes



Neuron 29. Posterior mean and 95% interval estimates for the intensity functions. Posterior mean and 95% interval band for the difference of density functions between the random and repeating condition.

Modeling and inference for Poisson processes



Neuron 29. Posterior distributions for the difference of density functions between the random and repeating condition at 12 time points.

3. Extensions to spatial point process modeling

- Again, bounded event data are rescaled such that point locations $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ all lie within the unit square, $\mathcal{R} = (0,1) \times (0,1)$
- Two-dimensional extension of the Beta DP mixture model using a Sarmanov dependence factor to induce a bounded bivariate kernel density with Beta marginals
- DP mixture model for spatial NHPP intensity $\lambda(\boldsymbol{x}; G)$:

  $$\gamma \int \text{be}(x_1; \mu_1, \tau_1)\text{be}(x_2; \mu_2, \tau_2)\,(1 + \rho(x_1 - \mu_1)(x_2 - \mu_2))\,\mathrm{d}G(\mu_1, \mu_2, \tau_1, \tau_2, \rho),$$

  where $G \sim \text{DP}(\alpha, G_0)$ and G_0 is built from independent uniform distributions for μ_1 and μ_2, and inverse gamma distributions for τ_1 and τ_2, and a conditional uniform distribution for ρ over the region such that $1 + \rho(x_1 - \mu_1)(x_2 - \mu_2) > 0$, for all $\boldsymbol{x} \in \mathcal{R}$.

Extensions to spatial point process modeling

- Models that build the mixture from bivariate beta kernels are flexible (e.g., resistant to edge effects), but inefficient in implementation (require augmented Metropolis-Hastings algorithms for posterior simulation)

- Conditionally conjugate alternative by applying individual logit transformations to each coordinate dimension and mixing over bivariate Gaussian density kernels:

$$\lambda(\boldsymbol{x}; G) = \gamma \int \mathrm{N}\left(\mathrm{logit}(\boldsymbol{x}); \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \frac{1}{\boldsymbol{x}'(1-\boldsymbol{x})} \,\mathrm{d}G(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \qquad G \sim \mathrm{DP}(\alpha, G_0),$$

  with standard conjugate normal/inverse-Wishart form for G_0

- Simpler mixture formulations based on kernels with independent components (with dependence between x_1 and x_2 induced by mixing)

Extensions to spatial point process modeling

Applications to extreme value analysis

- Analysis of extremes of a process observed over time using the *threshold approach* — based on the bivariate point pattern comprising the times of exceedances and the excess values
- Modeling for the corresponding NHPP intensity using the bivariate Beta mixture model (Kottas & Sansó, 2007)
- Current work (with Abel Rodriguez and Ph.D. student Ziwei Wang):
  - more structured mixture kernels to obtain connections with standard approaches on theoretical properties of tail behavior
  - application to inference for environmental extremes using spatial DP mixture models for temporal NHPP intensities

4. A modeling framework for marked Poisson processes

- Many applications involve *marks* – a set of random variables associated with each point event – such that the data generating mechanism is characterized as a **marked point process**

- Intensity mixture framework particularly powerful for modeling marked Poisson processes to provide a unified inference framework for point pattern data (Taddy & Kottas, 2010)
  - Semiparametric approach: combines nonparametric model for NHPP density with modeling for the conditional mark distribution
  - Fully nonparametric framework: builds inference for the marginal point process intensity **and** the conditional mark distribution through a mixture model for the joint mark-location point process

A modeling framework for marked Poisson processes

- Data structure: each point $\boldsymbol{x}_i$, $i = 1, ..., N$, in the observation window $\mathcal{R}$ has associated mark $\boldsymbol{y}_i$ with values in the mark space $\mathcal{M}$, which may be multivariate and may comprise both categorical and continuous variables

- **Semiparametric modeling approach:** joint model for the conditional mark density $h(\boldsymbol{y} \mid \boldsymbol{x})$ and the point process intensity through

$$\phi(\boldsymbol{x}, \boldsymbol{y}) = \lambda(\boldsymbol{x}) h(\boldsymbol{y} \mid \boldsymbol{x}) = \gamma f(\boldsymbol{x}) h(\boldsymbol{y} \mid \boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{R}, \, \boldsymbol{y} \in \mathcal{M}.$$

  – the conditioning in $h(\boldsymbol{y} \mid \boldsymbol{x})$ does not involve any portion of the history of the point process other than point $\boldsymbol{x}$

  – the *Marking theorem* yields that the marked point process

$$\{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x} \in \mathcal{R}, \boldsymbol{y} \in \mathcal{M}\}$$

  is a NHPP with intensity function given by $\phi(\boldsymbol{x}, \boldsymbol{y})$ for $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{R} \times \mathcal{M}$ (and by its extension to $\mathcal{B} \times \mathcal{M}$ for any bounded $\mathcal{B} \supset \mathcal{R}$)

A modeling framework for marked Poisson processes

- **Nonparametric modeling framework:** work directly with the Poisson process defined over the joint location-mark observation window, $\mathcal{R} \times \mathcal{M}$, with intensity $\phi(\boldsymbol{x}, \boldsymbol{y})$
- The (inverse of the) marking theorem yields that the joint process is also the marked Poisson process of interest, provided the marginal intensity $\int_{\mathcal{M}} \phi(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\boldsymbol{y} = \lambda(\boldsymbol{x})$ is locally integrable
- Define a process over the joint location-mark space with intensity function

$$\phi(\boldsymbol{x}, \boldsymbol{y}; G) = \gamma \int \mathrm{k}^{\boldsymbol{x}}(\boldsymbol{x}; \theta^{\boldsymbol{x}}) \mathrm{k}^{\boldsymbol{y}}(\boldsymbol{y}; \theta^{\boldsymbol{y}}) \mathrm{d}G(\theta^{\boldsymbol{x}}, \theta^{\boldsymbol{y}}) = \gamma f(\boldsymbol{x}, \boldsymbol{y}; G)$$

$\rightarrow$ where the mark kernel $\mathrm{k}^{\boldsymbol{y}}(\boldsymbol{y}; \theta^{\boldsymbol{y}})$ has support on $\mathcal{M}$, and $G \sim \mathrm{DP}(\alpha, G_0)$
$\rightarrow$ the integrated intensity can be defined in terms of either the joint or marginal process, such that $\gamma = \int_{\mathcal{R}} \lambda(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_{\mathcal{R}} \left[\int_{\mathcal{M}} \phi(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\boldsymbol{y} \right] \mathrm{d}\boldsymbol{x}$

- Note that the marginal intensity, and hence the marked point process, are properly defined with locally integrable intensity functions:

$$\begin{aligned}\int_{\mathcal{M}} \phi(\boldsymbol{x},\boldsymbol{y};G)\mathrm{d}\boldsymbol{y} &= \gamma \int_{\theta^{\boldsymbol{x}}} \mathrm{k}^{\boldsymbol{x}}(\boldsymbol{x};\theta^{\boldsymbol{x}}) \int_{\theta^{\boldsymbol{y}}} \left[\int_{\mathcal{M}} \mathrm{k}^{\boldsymbol{y}}(\boldsymbol{y};\theta^{\boldsymbol{y}})\mathrm{d}\boldsymbol{y}\right] \mathrm{d}G(\theta^{\boldsymbol{x}},\theta^{\boldsymbol{y}}) \\ &= \gamma \int \mathrm{k}^{\boldsymbol{x}}(\boldsymbol{x};\theta^{\boldsymbol{x}})\mathrm{d}G^{\boldsymbol{x}}(\theta^{\boldsymbol{x}}) = \gamma f(\boldsymbol{x};G) = \lambda(\boldsymbol{x};G)\end{aligned}$$

$\rightarrow$ $G^{\boldsymbol{x}}(\theta^{\boldsymbol{x}})$ is the marginal mixing distribution, which has an implied DP prior with base density $g_0^{\boldsymbol{x}}(\theta^{\boldsymbol{x}}) = \int g_0(\theta^{\boldsymbol{x}},\theta^{\boldsymbol{y}})\mathrm{d}\theta^{\boldsymbol{y}}$
$\rightarrow$ we have thus recovered the original DP mixture model for the marginal location NHPP over $\mathcal{R}$

A modeling framework for marked Poisson processes

- In general, both the mixture kernel and base distributions will be built from independent components corresponding to marks and to locations

- Alternatively to the generic independent kernel approach, the special case of a combination of real-valued continuous marks with the logit-normal kernel models allows for joint multivariate-normal kernels
  - e.g., a temporal point process with continuous marks is specified via bivariate normal kernels as

$$\phi(t, y; G) = \gamma \int \mathrm{N}\left([\mathrm{logit}(t), y]'; \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \frac{1}{t(1-t)} \mathrm{d}G(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

  - specification is easily adapted to spatial processes or multivariate continuous marks through the use of higher dimensional normal kernels

- Implied conditional regression (curve fitting) using nonparametric mixture models (Müller et al., 1996; Rodriguez et al., 2009; Taddy & Kottas, 2009, 2010)

A modeling framework for marked Poisson processes

- A quick note on posterior inference:

∗ posterior simulation for the finite dimensional part of the parameter vector (excluding the random mixing distribution G)

∗ here, obtaining also the posterior distribution for G is key for proper conditional inference under the DP mixture model

∗ posterior samples G_L using the DP stick-breaking definition with a truncation approximation (involving L terms):
→ to the conditional posterior distribution of G (Gelfand & Kottas, 2002)
→ or, more efficiently, to only the prior using the representation of the conditional posterior measure of G as a mixed random distribution with point masses at the distinct cluster locations and the remaining weight on the prior DP measure for G (Pitman, 1996; Ishwaran & Zarepour, 2002)

A modeling framework for marked Poisson processes

- Inference for the conditional mark density at $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ is available through

$$h(\boldsymbol{y}_0 \mid \boldsymbol{x}_0; G_L) = \frac{f(\boldsymbol{x}_0, \boldsymbol{y}_0; G_L)}{f(\boldsymbol{x}_0; G_L)} = \frac{\int \mathrm{k}^{\boldsymbol{x}}(\boldsymbol{x}_0; \theta^{\boldsymbol{x}}) \mathrm{k}^{\boldsymbol{y}}(\boldsymbol{y}_0; \theta^{\boldsymbol{y}}) \mathrm{d}G_L(\theta^{\boldsymbol{x}}, \theta^{\boldsymbol{y}})}{\int \mathrm{k}^{\boldsymbol{x}}(\boldsymbol{x}_0; \theta^{\boldsymbol{x}}) \mathrm{d}G_L^{\boldsymbol{x}}(\theta^{\boldsymbol{x}})}$$

  where the integrals are available as finite sums

- Also, inference for the conditional expectation $\mathbb{E}(\boldsymbol{y} \mid \boldsymbol{x}_0; G_L)$, and the marginal location intensity

- Key features of the joint mixture modeling framework:
  - provides flexible specifications for multivariate mark distributions comprising both categorical and continuous marks
  - particularly appealing for modeling spatially correlated marks
  - allows (relatively) easy prior specification and posterior simulation

Data illustrations

- **Simulation example:** temporal NHPP on $\mathcal{R} = (0, 1)$
  - intensity $\lambda(t) = 250\,(\text{be}(t; 1/11, 11) + \text{be}(t; 4/7, 7))$ (so $\gamma = 500$)
  - $N = 481$ events accompanied by binary marks z and continuous marks y generated from $h(y, z \mid t) = h(y \mid z, t)\text{Pr}(z \mid t)$
  - $\text{Pr}(z = 1 \mid t) = t^2$ and $h(y \mid z, t)$ built from $y = -10(1-t)^4 + \varepsilon$, with $\varepsilon \sim \text{N}(0, 1)$ if $z = 0$, and $\varepsilon \sim \text{gamma}(4, 1)$ if $z = 1$
  - non-linear heteroskedastic marginal regression function for y given t

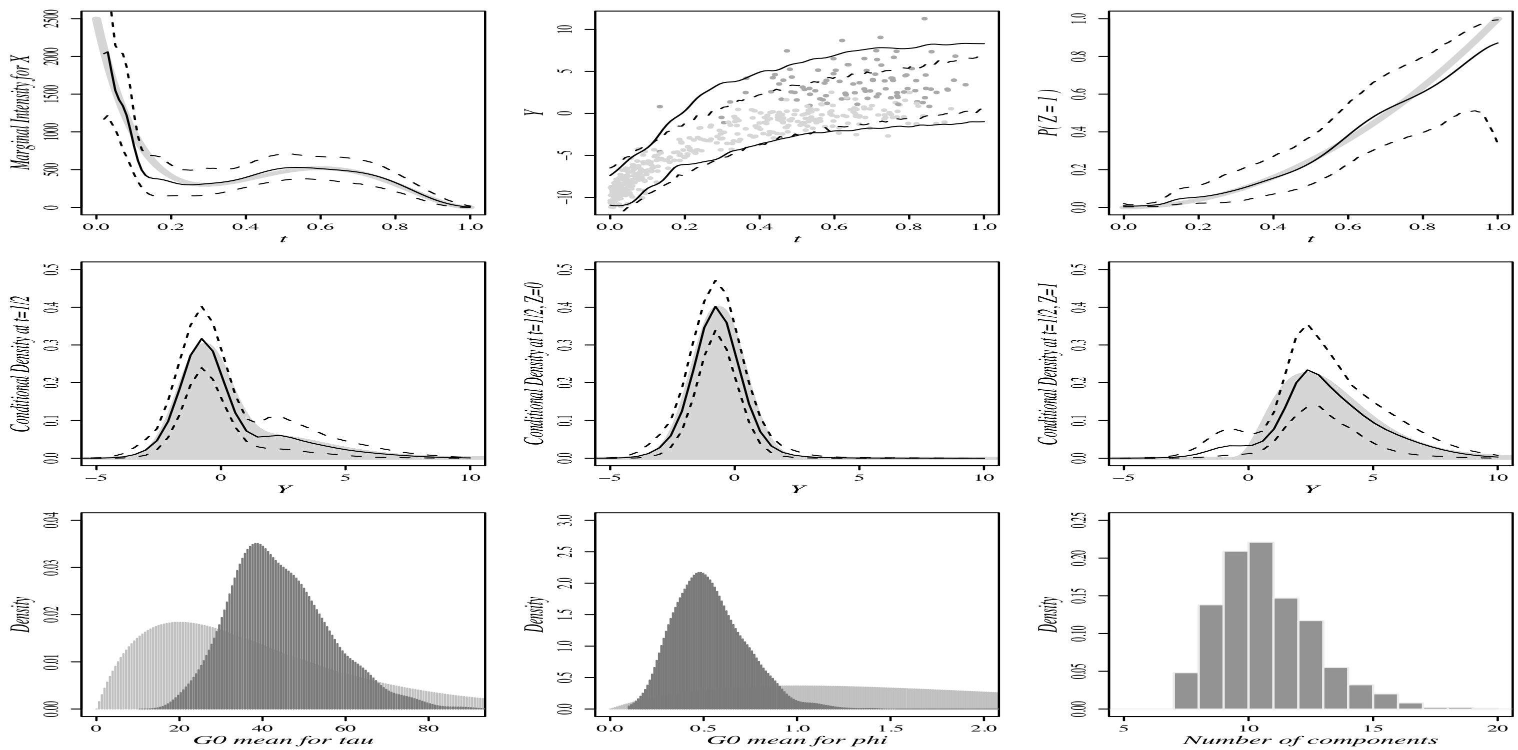
- Nonparametric DP mixture model:

$$f(t, y, z; G) = \int \text{be}(t; \mu, \tau)\text{N}(y; \eta, \phi)q^z(1-q)^{1-z}\text{d}G(\mu, \tau, \eta, \phi, q)$$

∗ $g_0 = \mathbf{1}_{\mu \in (0,1)}\text{gamma}(\tau^{-1}; 2, \beta_\tau)\text{N}(\eta; 0, 20\phi)\text{gamma}(\phi^{-1}; 2, \beta_\phi)\text{be}(q; 0.5, 1)$

∗ reference prior for γ; gamma priors for α, β_τ and β_ϕ

A modeling framework for marked Poisson processes



Top row: mean and 90% interval for the marginal intensity $\lambda(t; G)$, the data (dark grey for $z = 1$), and 90% predictive intervals based on both $h(y \mid t; G)$ (solid lines) and GP regression (dotted lines), and mean and 90% intervals for $\Pr(z = 1 \mid t; G)$. Middle row: mean and 90% intervals for conditional densities for y at $t = 1/2$, marginalized over z (left panel) and conditional on z (middle and right panels), with true densities plotted in grey. Bottom row: posterior samples for τ, ϕ and for the number of latent mixture components.

A modeling framework for marked Poisson processes

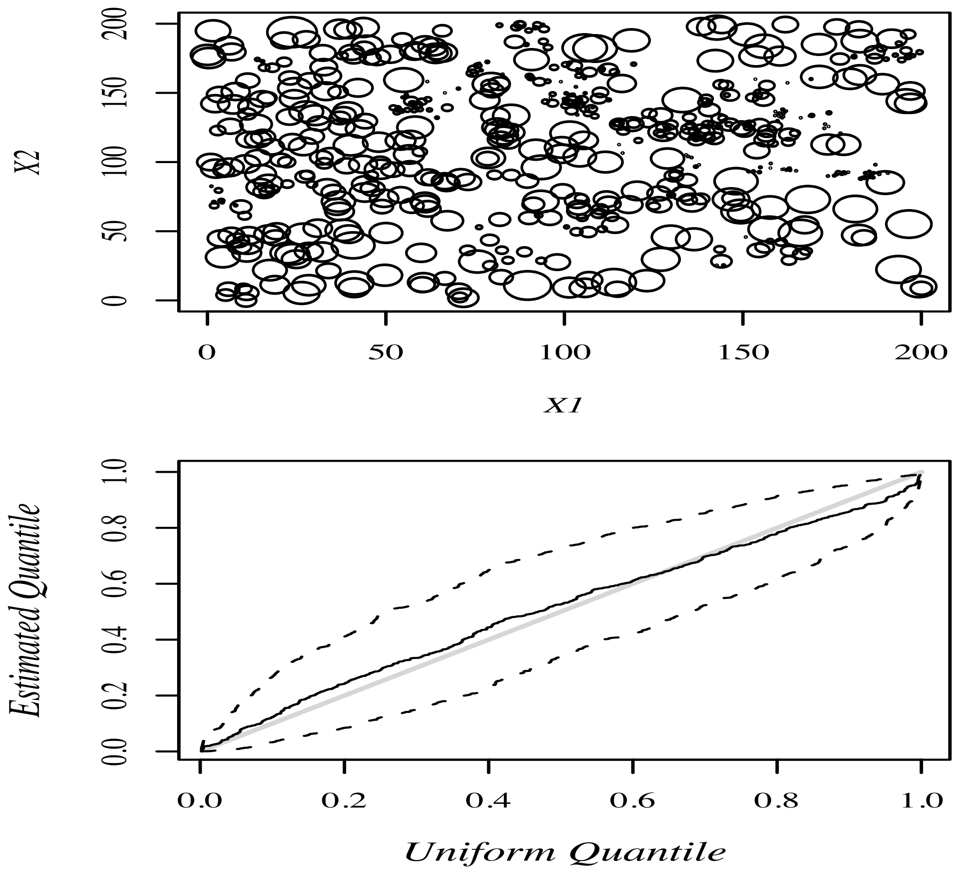
- **Longleaf pine trees:** locations and diameters of $N = 584$ Longleaf pine trees in a 200×200 meter patch of forest in Thomas County, GA
  - measured mark is diameter at breast height (1.5 m), or *dbh*, recorded only for trees with greater than 2 *cm dbh*
  - data were analyzed by Rathburn & Cressie (1994) with a space-time survival point process
  - Poisson processes are generally viewed as an inadequate model for forest patterns, due to the dependent birth process by which trees occur
  - NHPP should be flexible enough to account for variability in tree counts at a single time point and, here, we concentrate on inference for the conditional *dbh* mark distribution

A modeling framework for marked Poisson processes

- Model checking through the *time-rescaling* theorem
- Temporal point processes: if the point pattern $0 < t_1 \leq t_2 \leq ... \leq t_N < 1$ is a realization from a NHPP with cumulative intensity $\Lambda(t) = \int_0^t \lambda(s)ds$
  - the transformed point pattern $\{\Lambda(t_i) : i = 1, ..., N\}$ is a realization from a homogeneous Poisson process with unit rate
  - the $u_i = 1 - \exp\{-(\Lambda(t_i) - \Lambda(t_{i-1}))\}$, $i = 1, ..., N$, are independent uniform random variables on $(0, 1)$ (with $\Lambda(0) = 0$)
- Approach can be extended to spatial NHPPs by applying the rescaling to each margin of the observation window (Cressie, 1993)
- Similar approach for the marks: work with the $v_i = H(y_i \mid \boldsymbol{x}_i)$, for $i = 1, \ldots, N$, where $H(y \mid \boldsymbol{x}; G) = \int_{-\infty}^{y} h(s \mid \boldsymbol{x})\mathrm{d}s$ is the cdf of the conditional mark distribution
- Under any DP mixture model formulation, obtain entire posterior distributions for the u_i and v_i

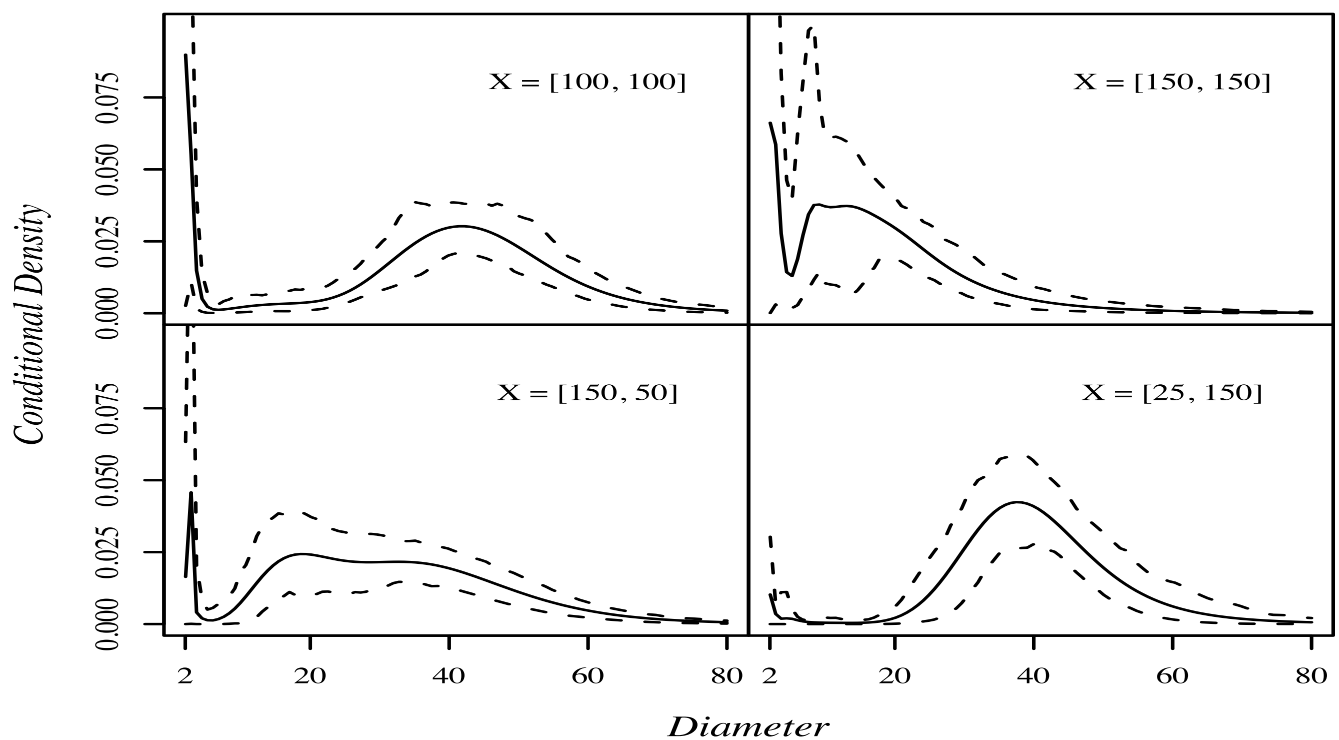
A modeling framework for marked Poisson processes

Longleaf pine trees. Data (point size proportional to tree diameter), and a Q-Q plot (mean and 90% interval) for the c.d.f. of the conditional mark distribution evaluated at data.



A modeling framework for marked Poisson processes

Posterior mean and 90% intervals for the conditional *dbh* mark density at four $\boldsymbol{x}$ values for spatial location.



5. Conclusions

- Unified inference framework for Poisson process intensities built directly from density estimation
- Modeling framework particularly powerful for marked Poisson processes
- Extensions to modeling for spatial NHPP intensities observed over discrete time (Taddy, 2009)
- A couple of things we are currently looking at:
  - spatial DDP modeling for temporal NHPP intensities (extreme value analysis applications)
  - nonparametric prior models for more general point processes (neuronal data analysis settings)

Acknowledgments

- Matt Taddy (University of Chicago Booth School of Business)
- Bruno Sansó (UCSC)
- Ziwei Wang and Abel Rodriguez (UCSC)
- Sam Behseta (California State University, Fullerton)

- Funding from *NSF, Division of Environmental Biology* and *NASA, Applied Information Systems Research* and *NASA-UARC* programs

- Contact info:
  e-mail: thanos@ams.ucsc.edu, web: http://www.ams.ucsc.edu/∼thanos

- UCSC Dept of Applied Math and Statistics website: www.ams.ucsc.edu

MANY THANKS!