Dynamical Systems as Gaussian Process Prior Distributions

Nathan Wycoff

Virginia Tech

December 6, 2016

Table of Contents

Table of Contents

Mathematical Spatial Models

- Separate suite of methods for spatial analysis.
- Makes use of subject matter theory based on partial differential equation systems.
- Ex: Murray, *Mathematical Biology II: Spatial and Biomedial Applications*

How can different kinds of spatial analysis merge?

Dynamical Systems

Description of something over time.

Examples:

- Standard Equation:

$$y(t) = f(t)$$

- Differential Equation:

$$\frac{\partial^i y(t)}{\partial t^i} = f(\frac{\partial^i y(t)}{\partial t^i}, \frac{\partial^i y(t)}{\partial x_j^i})$$

- Difference Equation:

$$y_{t+1,j} = f(y_{t-i,j+k})$$

Theme

- GP kernels as priors on functions
- Informative kernels

Central Question
Can we use a dynamical system as prior information for a GP through a kernel?
Importance:

- Include information about system dynamics
- Separate signal from noise

Table of Contents

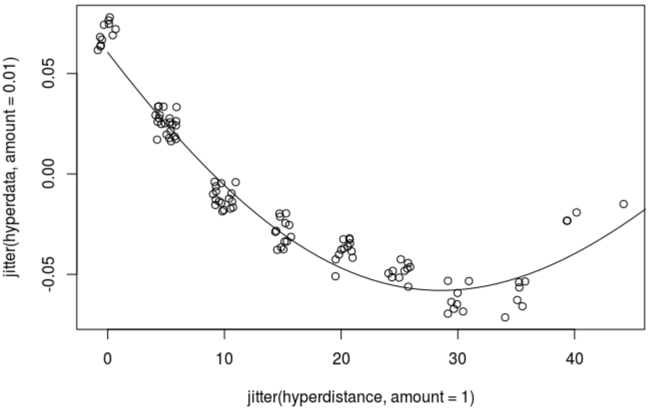
Existing Ideas

- Make predictions using the mean (Null Model)
- Make predictions using the dynamical system only.
- Make predictions using a Gaussian Process with radial basis function.
- Make predictions using a dynamical system, doing residual inference with a GP.

Proposition 1: Nonparametric Matching of Kernel to DS

Idea:

- Calculate empirical covariance of DS at train and test points.
- Nonparametric estimate as kernel for GP.
- Possible "hyperGP"



Proposition 2: Use DS to Develop Priors for Parametric Kernel

Idea:

- Fix some parametric form
- With loose priors, develop a posterior on DS with no nugget over train and test points.
- Use posterior as prior for real data analysis.

For this project: Moment-match IG prior

Table of Contents

Problem setup

- Release gas in a room.
- Observe spread along 1 dimension at some points.
- Predict concentration at unobserved points.

The scientist's model:

$$\frac{\partial S}{\partial t} = D\Delta$$

$$\frac{\partial S}{\partial t} = D\frac{\partial^2 S}{\partial x^2}$$

Observation:

$$f_{obs} = f_{true} + \epsilon$$

$$\epsilon \sim N(0, 0.1)$$

Sim 1: Setup

- The scientists model is perfect.

Scientist's Model:

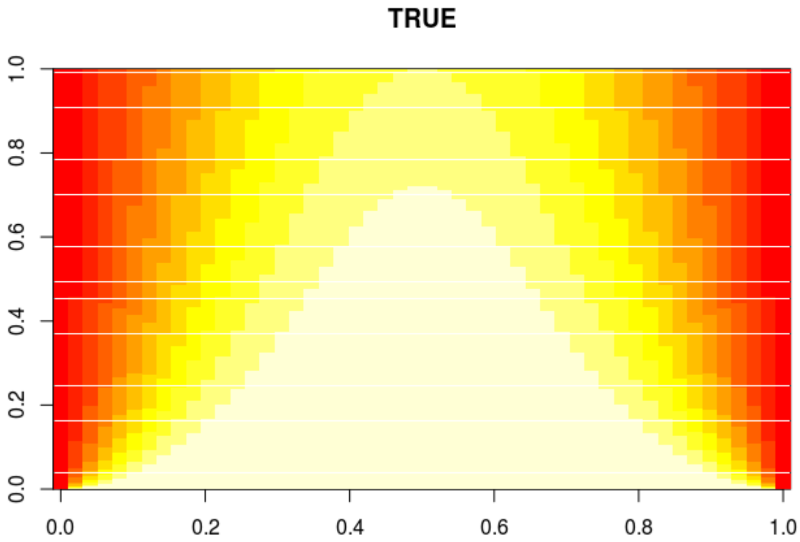
$$\frac{\partial S}{\partial t} = 10\frac{\partial^2 S}{\partial x^2}$$

Boundary conditions of zero, initial condition of point mass halfway through the region.

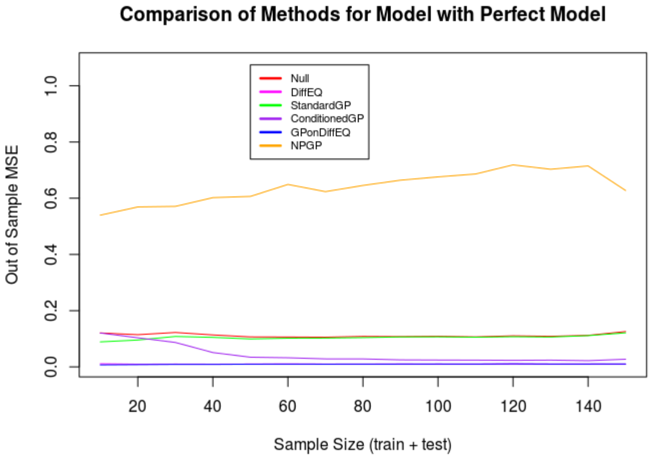
Truth:

$$\frac{\partial S}{\partial t} = 10\frac{\partial^2 S}{\partial x^2}$$

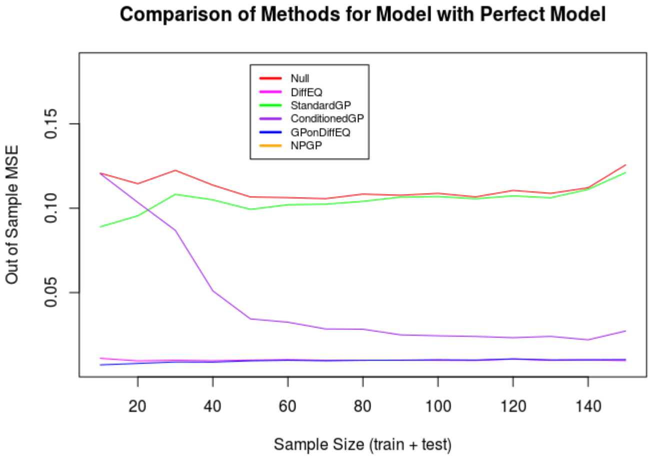
Sim 1: DiffEQ Solution



Sim 1: Results



Sim 1: Results



Sim 3: Setup

- The scientists model has the right parameter but the wrong structure.

Scientist's Model:

$$\frac{\partial S}{\partial t} = 10\frac{\partial^2 S}{\partial x^2}$$

Same boundary conditions

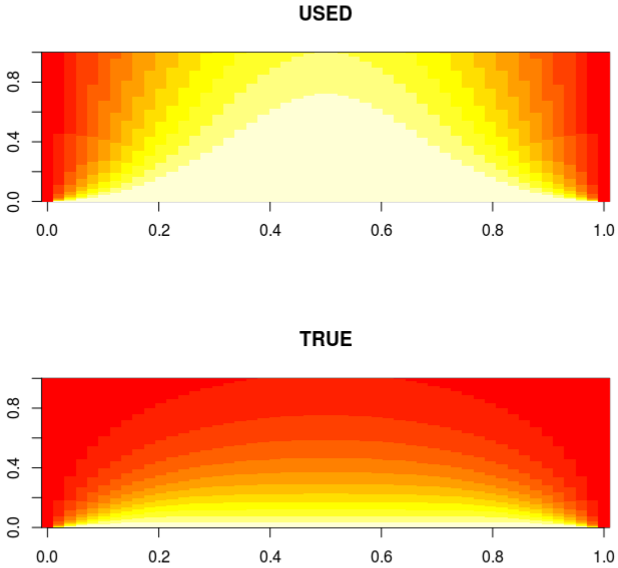
Truth:

$$\frac{\partial S}{\partial t} = 10\frac{\partial^2 S}{\partial x^2} - 0.05SR$$

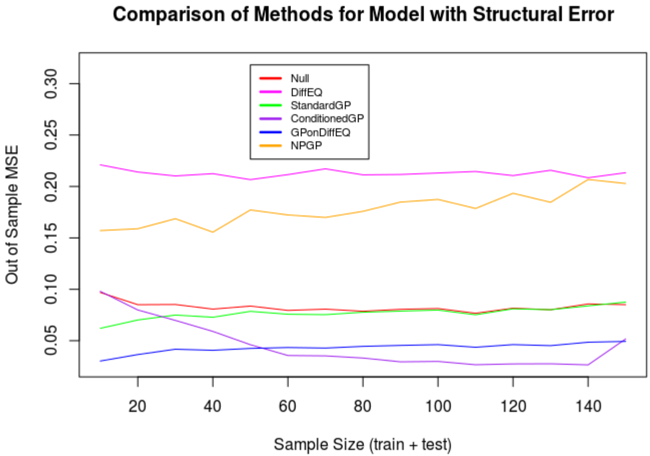
$$\frac{\partial R}{\partial t} = 10\frac{\partial^2 R}{\partial x^2} - 0.05SR$$

Correct boundary conditions for S, R has initial mass 5 everywhere.

Sim 2: DiffEQ Solution



Sim 2: Results



Future Work

Kernels in between:

- The NP kernel seems to adapt too much to the hyperdata, not enough to real data.
- Conditioned kernel may not bond to hyperdata enough.
- Can they be merged somehow?

Fit GP without integration:

$$cov(f_i, \frac{\partial f_i}{\partial x_{dj}}) = \frac{\partial k(x_i, x_j)}{\partial x_{dj}}$$

Bayes-Hermite Integration.