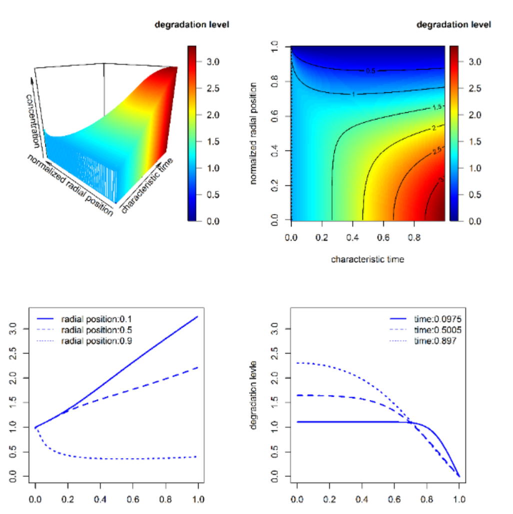
Statistical Modeling for Spatio-Temporal Degradation Data

Xiao Liu
IBM Thomas J. Watson Research Center

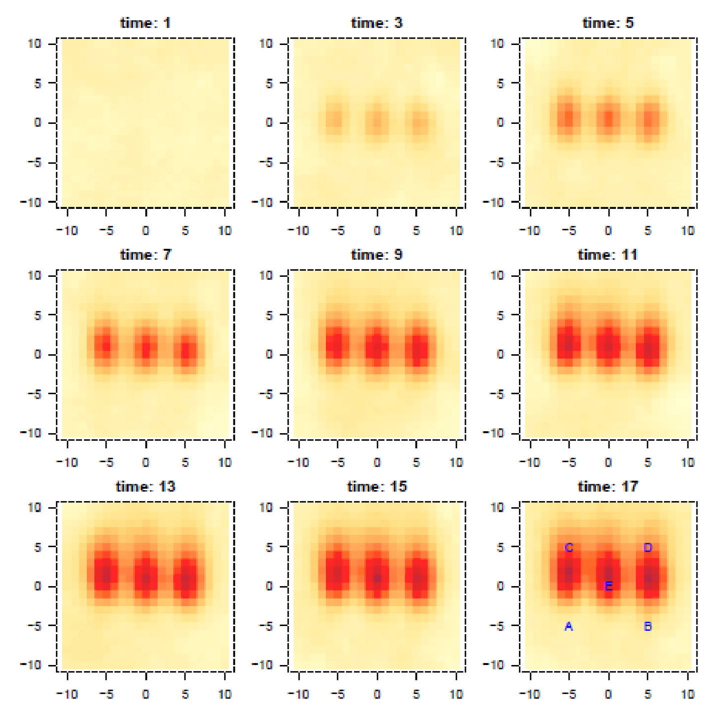
Spatio-Temporal Degradation Data

- Spatio-temporal models arise when data are collected across time and space, and one must take account of both temporal and spatial correlations.
- Data from spatio-temporal processes are common in the real-world, representing variety of interactions across processes and scales of variabilities.
- *Motivating Example 1:*
  - Autocatalytic Degradation and Erosion in Polymer Microspheres.



- *Motivating Example 2:*
  - Degradation measured on a two-dimensional surface (a 21×21 pixel array) over 9 equally-spaced time intervals.

  - Some unique challenges:
  - ➢ Spatial **dependence and propagation**
  - ➢ Spatial **heterogeneity**
  - ➢ Complex **spatio-temporal correlation structures** (anisotropic, space-time non-separable, etc.)



Basic Framework and Key Results

- Consider a discrete-in-time and continuous-in-space spatio-temporal random field,

$$\{Y(s,t); s \in R^d, t \in N^+\}$$

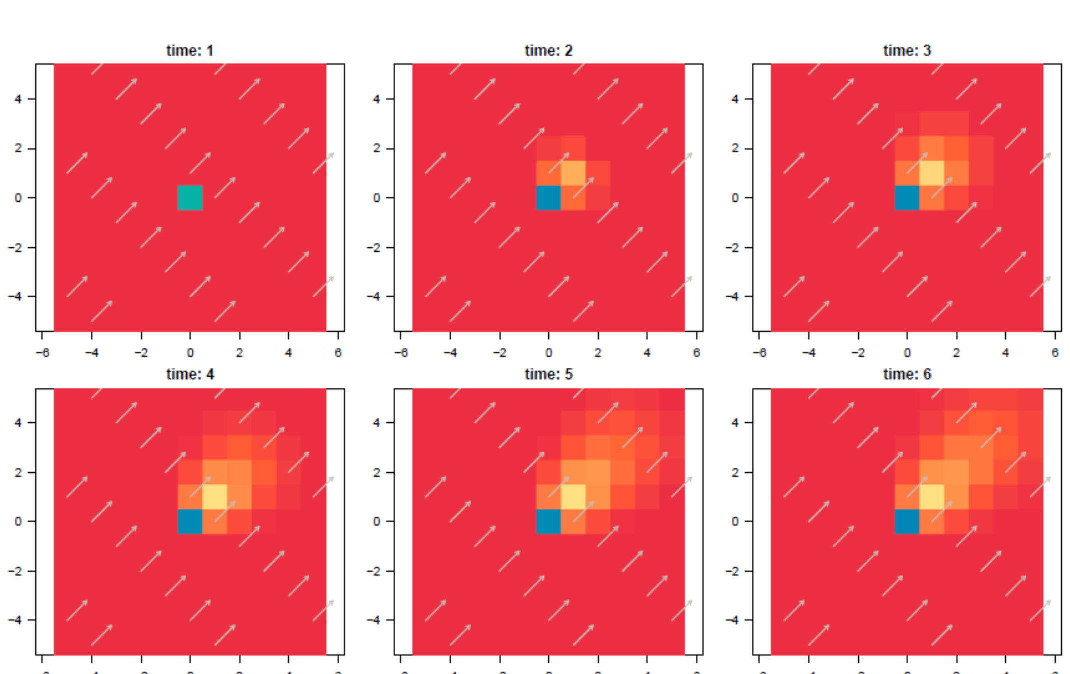
where $Y(s,t)$ represents the degradation at time t and location s in a d-dimensional space. Without loss of generality, we let $d=2$ which corresponds to the degradation over a two-dimensional surface.

- Based on the assumption of an additive accumulation of degradation (the fundamental idea behind existing stochastic degradation models), $Y(s,t)$ takes a general additive form:

$$Y(s,t) = \underbrace{G_\Delta(s,t)}_{\text{A generation process}} + \underbrace{Z(s,t)}_{\text{A propagation process}}$$

  - ➢ ***Generation process***: A spatial process that represents the amount of degradation generated at location s over the time interval $(t-\Delta, t]$:
$$G_\Delta(s,t) = g_\Delta(s,t) + \xi(s,t)$$
where $g_\Delta(s,t)$ is the mean, and $\xi(s,t)$ is a white-in-time Gaussian spatial process with covariance function $c_\Delta(\cdot) = \Delta \cdot c(\cdot)$.
  - ➢ ***Propagation process***: amount of degradation propagated to a certain location at time t can be expressed as a linear combination of the degradation at neighboring locations at the previous time stamp weighted by some spatial Kernel function:

- A **convolution model** with a Gaussian Kernel is used to model the propagation process, i.e., the influence of the degradation at neighboring areas: a linear combination of the degradation at neighboring locations at time $t-\Delta$, weighted by some spatial Kernel function.
$$Z(s,t) = \zeta_\Delta\{\omega_\Delta * Y(s,t-\Delta)\}$$
where $*$ denotes the convolution operation, ω_Δ is a Gaussian convolution kernel, and ζ_Δ is a scaling factor.
  - ➢ An illustration



Basic Framework and Key Results (cont.)

- A discrete-in-time and continuous-in-time spatio-temporal degradation process:

$$Y(s,t) = g_\delta(s,t) + \varepsilon_\delta(s,t) + A + B + C$$

where ϕ is a Gaussian density function, and

$$A = \sum_{i=1}^{n}\{e^{-i\lambda\delta}\phi_{i,t}(s) * g_\delta(s,t-i\delta)\}$$
$$B = e^{-n\lambda\delta}\phi_{i,t}(s) * Z(s,t-\Delta)$$
$$C = \sum_{i=1}^{n}\{e^{-i\lambda\delta}\phi_{i,t}(s) * \varepsilon_\delta(s,t-i\delta)\}$$

- For the stochastic degradation process defined above, the covariance of the degradation between (s_1,t_1) and (s_2,t_2) (assuming t_2-$t_1 = \iota\delta$ for ι=0,1,2,…) is given by

$$\text{cov}(Y(s_1,t_1),Y(s_2,t_2)) = \sum_{i=0}^{n}\left(\tilde{\Psi}_{i,t_1} * \Psi_{\iota+i,t_2} * c_\delta\right)(d) + I_{\iota=0}c_\delta(d)$$

where

$$\Psi_{\iota+i,t_2} = \begin{cases} e^{-i\lambda\delta}\phi_{i,t}(s), & i > 0 \\ 1, & i = 0 \end{cases}$$

and $d = s_2 - s_1$, $\tilde{\Psi}_{i,t_1}(s) = \Psi_{i,t_1}(-s)$, and $I_{\iota=0} = 1$ if $\iota = 0$.

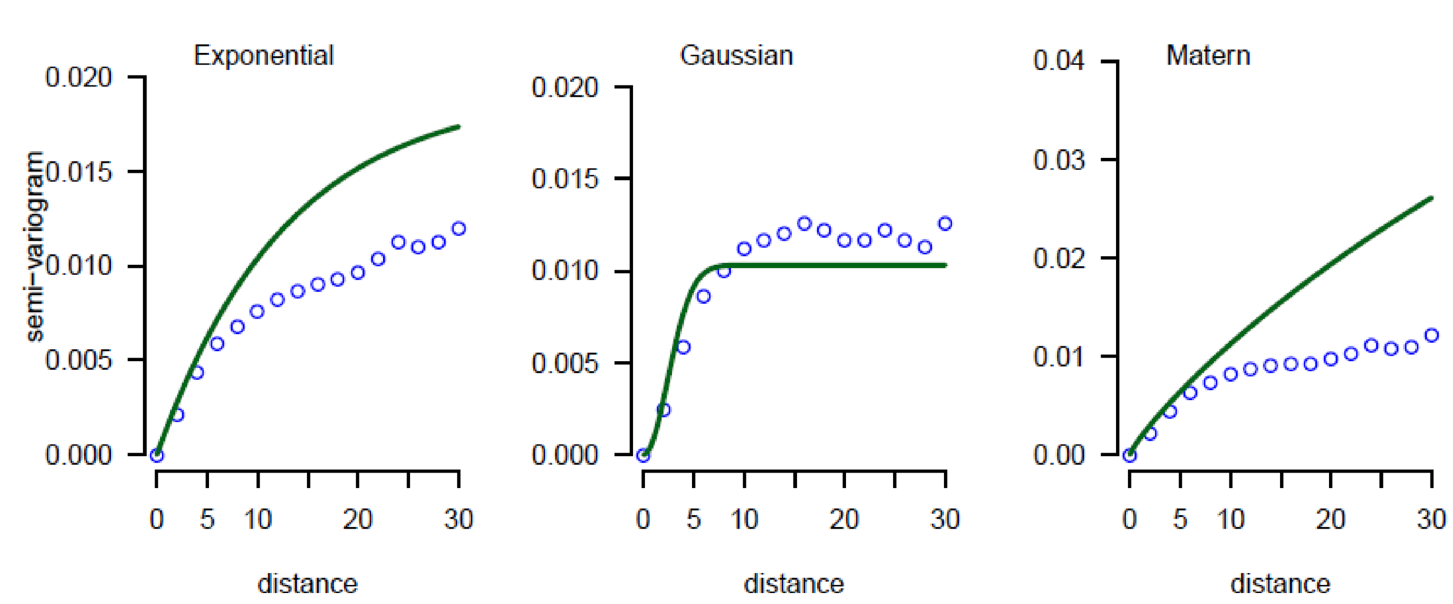
- The covariance is anisotropic and space-time non-separable, if $v \neq 0$.

Numerical Illustration

- Maximum Likelihood Estimation for parameter estimation

		covariance function $c(\cdot)$		
		Exponential	Gaussian	Matérn
parameters	λ	0.127	0.09	0.149
	$\boldsymbol{v}$	(-0.040,0.499)	(-0.004,0.793)	(0.006,0.598)
	ρ_1	1.119	2.247	0.802
	ρ_2	0.192	0.301	0.216
	θ_1	0.019	0.010	0.071
	θ_2	12.883	11.564	81.560
	θ_3	N/A	N/A	0.434
	β	0.977	1.251	1.108

- Graphical model validation:
  - ➢ Comparison between the empirical semi-variogram estimated from the Cressie-Hawkins estimator and the theoretical semi-variogram computed based on the estimated model parameters, respectively assuming Exponential, Gaussian, and Matern covariance functions for $c(\cdot)$.



References

Details can be found in the manuscript, which is available from the author.

Contact

Dr. Xiao Liu
Office 29-252, 1101 Kitchawan Road, Yorktown Heights, New York, 10598
E-mail: liuxiao@us.ibm.com