

## Research Statement

My research falls in three areas: nonparametric techniques for point process data, environmental applications, and portfolio optimization in financial statistics. My dissertation focuses on two methods in repeated independent realizations of a point process. The first, spike-time distance, is used to characterize the distance between two point patterns. A second method for repeated realizations of point processes is prototype analysis, which is used to identify the typical pattern, called the prototype. In separate research, I have collaborated on environmental applications and modeling techniques in financial statistics.

### *Dissertation*

Computing spike time distance or determining the prototype for large point pattern data sets in multidimensional spaces was infeasible in previous work. In (1), we presented an algorithm to compute spike time distance using dynamic programming and proved its convergence under general conditions. This algorithm provides a generalized framework that can be applied to similar distance measures. We also proposed a kernel smoothing technique for prototype estimation and applied our methods to neuronal spike trains in cats. The prototype may be thought of as a sort of median, generalized so as to apply to point patterns. Each of these techniques readily extends to multiple dimensions.

My current dissertation research focuses on classification and model checking. Given a training set of patterns with identified classes, I am investigating the use of spike-time distance, multidimensional scaling, and linear classification techniques to classify patterns based on the training data (2). I also am researching the properties of prototypes under limiting conditions for applications in evaluating point process models (3).

Additional ongoing work includes increasing accessibility to and applying those methods developed in my research. I am currently writing an R package for generalized spike-time distance and multidimensional prototypes (4), which will be discussed in a companion paper (5). I am also working on a proposed multidimensional application of the package to characterize the annual distribution of wildfire locations, times, and sizes in southern California.

### *Additional Research*

My dissertation research has been supplemented by collaborations on three projects, two focused on environmental science and one on stock portfolio optimization. The first was a time series analysis to investigate Pacific Decadal Oscillation, a measure of large-scale

climate variability (6). During 2008-9, I reviewed the accuracy of the current methodology used at the US Environmental Protection Agency to assess public exposure to pesticide residues in food (7). Most recently, I collaborated with Professor Nicolas Christou of UCLA Statistics in the production of an R package for financial statistics. The package focuses on stock portfolio optimization, and we will present the package in an upcoming paper (8, 9).

### *References*

- (1) Diez DM, Schoenberg FP, and Woody CD. *Analysis of neuronal responses to stimuli in cats using point process prototypes*. In review.
- (2) Diez DM. *Applying distance measures in point pattern classification*. In progress.
- (3) Diez DM. *Model testing for point process data using spike time distance*. In progress.
- (4) Diez DM, Tranbarger KE, and Schoenberg FP. *ppDistMeasures*. R package. In progress.
- (5) Diez DM and Schoenberg FP. *ppDistMeasures: Point Pattern Prototypes and Spike-Time Distance*. Planned paper.
- (6) Nairn-Birch N, Diez DM, Eslami E, Macias Fauria M, Johnson E, and Schoenberg FP. *Simulation and Estimation of Probabilities of Phases of the Pacific Decadal Oscillation*. Environmetrics. In press.
- (7) Diez DM and Villanueva P. *Decomposition Techniques in the Application of Pesticide Exposure*. In progress.
- (8) Diez DM and Christou N. *stockPortfolio*. R package. October, 2009.
- (9) Diez DM and Christou N. *stockPortfolio: An R Package for Stock Portfolio Modeling and Management*. In progress.