Nonparametric Time Series Analysis: A review of Peter Lewis’ contributions to the field

Bonnie Ray
IBM T. J. Watson Research Center

*Joint Statistical Meetings 2012*
Outline

- Background
  - My connection to Peter

- Joint work: Nonparametric time series analysis
  - Motivation
  - Methodological foundation
  - Applications and Extensions

- Current related work

- Impact of Peter’s work
How I knew Peter

IBM T. J. Watson Research Center
Yorktown Heights, NY

Carmel Beach
Carmel, CA

Naval Postgraduate School
Monterey, CA
Underlying motivation for much of our joint work

20 Years of Daily Sea Surface Temps at Granite Canyon

(a)

Log Temperature

(b)

Residuals

Time in Days from 3/1/1971
Key idea

- Peter recognized that non-parametric regression techniques, under development in the late 80’s and early 90’s, could be applied in the time series context to model non-linear time series phenomena.

- Focused on Multivariate Adaptive Regression Splines technique (MARS)
  - Fits truncated linear splines functions to the data with optimal knot points selected automatically.

- Extended MARS to TS-MARS
  - Used MARS algorithm to modeled nonlinear univariate time series using lagged values of the series itself and possible exogenous covariates.
  - Results in nonlinear threshold models that are continuous in the domain of the predictor variables (ASTAR, SMASTAR).
  - More general than self-exciting threshold-type models (SETAR, TARSO), which identify piecewise linear functions over disjoint subregions and are discontinuous at the boundaries of the domain of interest.

- Main publications
MARS Applied to Granite Canyon Data

- Model for 5 Years of SSTs using Wind Direction and Wind Speed

\[ X_t = \]
\[ 2.192(0.0036) + 0.878(0.0079)(X_{t-1} - 2.13) + \\
+1.616(0.2770)(2.22 - X_{t-34}) + \\
+0.013(0.0018)(WS_{t-1} - 1.10) + I(WD_{t-1} \in \{1, 2\}) \\
-0.035(0.0018)(WS_{t-1} - 1.10) + I(WD_{t-1} \in \{2, 3\}) \\
-.499(0.0060)(X_{t-1} - 2.13) + (2.75 - X_{t-8}) + (2.68 - X_{t-17}) + \\
-0.584(0.0999)(2.27 - X_{t-34}) + (WS_{t-1} - 1.10) + I(WD_{t-1} \in \{2, 3\}) \\
-0.517(0.1174)(X_{t-49} - 2.510) + (WS_{t-1} - 3.00) + I(WD_{t-1} \in \{1, 4, 5\}) \\
+4.665(1.0344)(2.51 - X_{t-49}) + (2.26 - X_{t-24}) + I(WD_{t-1} \in \{2, 3\}) \]

Suggests that when the wind blows from the Northwest on the previous day, the SST tends to decrease.

Reflects the fact that the average time between storm fronts in the vicinity of Granite Canyon in the winter is about 8 days.

Suggests a coupling of SSTs with SSTs approximately 49 days previous, dependent on the wind direction and speed.
Periodic Autoregressive Models: Characterizing River Flows

- **Periodic time series**
  - Correlation structure does not change from cycle to cycle, but differs from period to period within a cycle
  - For example, monthly data may have a yearly cycle, but the correlation between observations in Jan and Feb is different from the correlation structure between observations in Feb and Mar

*Scatter plots of the logarithm of mean monthly flow of the Fraser River over the time period March 1913–December 1991*
Innovations in Modeling Periodic Time Series: P-CASTAR

- Adapted nonlinearity tests for threshold-type behavior to the case of periodic time series
- Applied MARS algorithm to time series exhibiting periodic behavior to capture non-linear relationships
  - Initially modeled each subseries separately using MARS algorithm
  - Introduced the use of categorical predictors representing each period within a cycle to simultaneously model nonlinear behavior for each period
  - Each response weighted to adjust for heteroskedasticity of the residuals in different periods and weights updated iteratively

\[ X_t = 7.15 - 0.782(7.307 - X_{t-1})_+ + 0.757(X_{t-1} - 7.307)_+ - 0.125I(t \text{ mod } 12 \in \{1, 4, 5, 12\}) + 0.293I(t \text{ mod } 12 \in \{5, 7\}) 
+ 0.815I(t \text{ mod } 12 \in \{4, 5, 6\}) 
+ 0.614(7.828 - X_{t-1})_+I(t \text{ mod } 12 \in \{5, 7\}) \]

Change the mean level of the model only

Impact of Peter’s work using MARS to model nonlinear structure in time series

- TSMARS methodology has been used to model energy price series, mobile communication channels, foreign exchange rates, brain dynamics, ozone extremes, nuclear safeguards and non-proliferation, ….. For example,

- Ideas extended to multivariate time series modeling
  - Kooperberg Bose and Stone (*JASA*, 1997) developed PolyMARS (PMARS) algorithm to extend the advantages of the MARS algorithm over simple recursive partitioning to the multiple classification problem
  - DeGooijer and Ray (*CSDA*, 2003) applied PMARS algorithm to model vector threshold-type nonlinearity in multivariate time series
Using PMARS to model Electricity Load Data

Three weeks of half-hourly electricity load data (n=1008) from the Australian states of New South Wales (NSW) and Victoria (VIC)

- Used temperature data, available only for NSW, as an additional predictor, along with Time of Day (TOD) and Time of Week (TOW) indicator variables

- Results
  - Interactions between the $TOD_t$ and lagged loads, suggesting that prior electricity usage acts to modulate $TOD_t$ effects
  - Several terms involving lagged loads contained thresholds, indicating that electricity loads exhibit different behavior when usage is above or below certain levels
  - Model showed a feedback relationship between the electricity loads of the two states

Related work from IBM Research: Scalable Matrix-valued Kernel Learning and High-dimensional Nonlinear Causal Inference

**Innovations**

- Propose a general matrix-valued **multiple kernel learning framework** to fit non-parametric models to multivariate time series, i.e. kernels are selected dynamically from a library of kernels based on the local structure of the data
- Allow a broad class of mixed norm regularizers, including those that induce **sparsity**, to be imposed on a dictionary of vector-valued Reproducing Kernel HilbertSpaces (RKHS)

**Resulting models**

- Non-parametric nonlinear, sparse temporal-causal models
- May be viewed as non-parametric multivariate extension of Group Lasso and related sparse learning models

**Applications**

- Weekly log returns of multiple related stocks
- Time-course gene expression microarray data
  - Modeled the expression levels of 2397 unique genes simultaneously measured at 66 time points corresponding to various developmental stages and grouped into 35 functional groups based on their gene to infer causal interactions between functional groups, as well obtain insight on within group relationships between genes.
Continuing impact of Peter’s work using MARS to model nonlinear structure in time series

- Motivated theoretical work on **limiting properties, boosting, improved partitioning algorithms**, etc., e.g.

- Ideas extended to **nonlinear transfer function-type models**

- Most recent work found directly linked to work of Lewis and Ray
Conclusion

Operations Research  Distinguished Professor
Emeritus Peter Lewis,
1932—2011

A leader in the fields of computer simulation, applied statistics and probability, and operations research. …..A common theme running through the comments about Peter Lewis by many of his colleagues and former students is his extraordinary influence on their professional careers and his steadfast encouragement and support of their work.