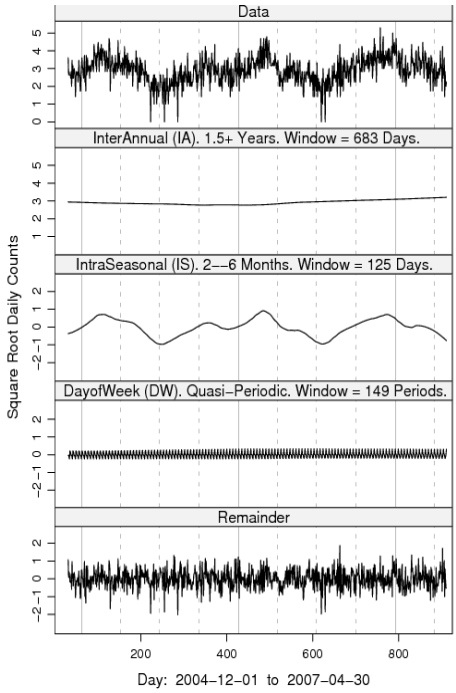
STL and Local Regression for Modeling Disease Surveillance Counts

David Anderson, Ryan Hafen, Ross Maciejewski, William S. Cleveland, David S. Ebert, Mourad Ouzzani, Shaun J. Grannis, Ahmad Abusalah
Purdue University, RVAC COE

Project Scope: Seasonal decomposition of time series by Loess (STL) is a well-proven, flexible method for modeling time series with periodic components. We are using STL to model daily disease counts for syndromic surveillance. Loess decomposes the square root daily counts into inter-annual, intra-seasonal, periodic day-of-the week, and remainder components. The inter-annual component models long term trends such as a hospital expanding capacity. The intra-seasonal component models yearly seasonal effects, such as the flu. The day-of-the-week component models the change in the propensity for patients to show up more on certain days of the week than others; the typical pattern is a high count on Monday, a decline during the rest of the work days, and a rise over the weekend. STL decomposition helps the analyst better visualize the trends in the data from each component with the noise removed.



Our work is part of the Indiana state syndromic disease surveillance program in collaboration with the Indiana State Department of Health (ISDH). The state monitors chief-complaints at emergency rooms across Indiana and we are receiving their data feed.

Recent Progress: We have completed the modeling of daily counts related to several syndrome categories for the 72 emergency rooms that have so far come on line in the Indiana surveillance program. The figure at the left shows the STL decomposition for respiratory-related emergency room visits for one hospital for 29 months.

Future Plans: We will develop STL-based control charts to assist with automated outbreak detection. The STL visualization and detection algorithms will be integrated into a visual analytics system being developed at PURVAC.

Relevance to Listed Research Areas: STL modeling is relevant to the Advanced Data Analysis research area. This particular application of STL statistical modeling is related to early detection of disease events, and is therefore also relevant to the Biological Threats and Countermeasures research area.

Publications: A paper based on this project is in preparation.