RESEARCH EXPERIENCES IN SOLID EARTH SCIENCE FOR STUDENTS
RESESS

A Non-Parametric Approach to Infer Slopes from Short Time-Series with Seasonal Fluctuation: Application to GPS Time-Series

Investigators: Brian Chung (Boston College) and Dr. Corné Kreemer (University of Nevada, Reno)

Introduction

A common element of time-series data in fields such as geophysics, hydrology, and astronomy is seasonal is seasonal signal. This variation indubitably obscures the trends that exist in the time-series. In cases with seasonal signals, the slope estimates can have low bias as long as the data span is long enough to outweigh the effects of the variation of the seasonal signal. The traditional method of calculating slopes of time-series is least-squares regression. *Blewitt and Lavallée* [2002] conclude that least-squares slope estimates that account for seasonal signal produce reasonable biases for data spans above 2.5 years long. They add that least squares slope estimates not accounting for seasonal signal produce acceptable biases for data spans beyond 4.5 years. The Theil-Sen method is a nonparametric approach to time-series regression that does not depend on the bold assumptions of parametric methods like least-squares. This method is additionally robust to outliers, whereas least-squares can be heavily influenced by the presence of outliers.

Figure 1A (left). MAGNET GPS monument and antenna, named DIRT, located in Glen Canyon National Recreation Area

Figure 1B (right). Common winter problem of snow interference causing GPS stations to cease producing position measurements due to a lack of solar power to charge the battery

Geodesists often encounter data spans that are too short for least-squares estimation. They calculate plate velocities based on the position data from fixed monuments around the world. Technological malfunctions as well as natural, anthropogenic, and economic factors often limit time spans over which the sensitive stations can produce useable data. *Van Dam et al.* [2001], note the effects of seasonal atmospheric and hydrospheric loading on the lithosphere, which alters GPS position measurements. A challenge at the forefront of geodesy is generating accurate and meaningful plate velocity estimates from short timespan time series. The seasonal Theil-Sen method is henceforth examined as a viable slope estimator for GPS time-series shorter than 2.5 years.

Methods

Theil-Sen Estimator

There are $\binom{n}{2}$ possible pairs of points in a time-series of *n* data points and each pair has a characteristic, but not always unique, slope. *Theil* [1950] proposes using the median of these slopes as an estimator for the regression coefficient. The seasonal estimator matches only data points within a range of exactly an integer number of years plus/minus a time interval, dt. For time-series based on the GPS position data of interest, this is an estimate of tectonic plate velocity. As this estimator is insensitive to outliers and does not rely on the strong assumptions of the least-squares estimator, it may produce smaller biases compared to the default least-squares approach.

Time-Series Simulations

(1) Generate set of 250 artificial time-series for each combination of: true velocity (mm/yr), white noise amplitude, seasonal signal amplitude, time length (yr); and starting phase (radians) for least-squares, and dt (days) for Theil-Sen
(2) Calculate seasonal and non-seasonal least-squares velocity estimates and a seasonal Theil-Sen velocity estimate
(3) Average velocity estimates over the 250 for each type of estimator
(4) Calculate the absolute bias for each method from the known true velocity

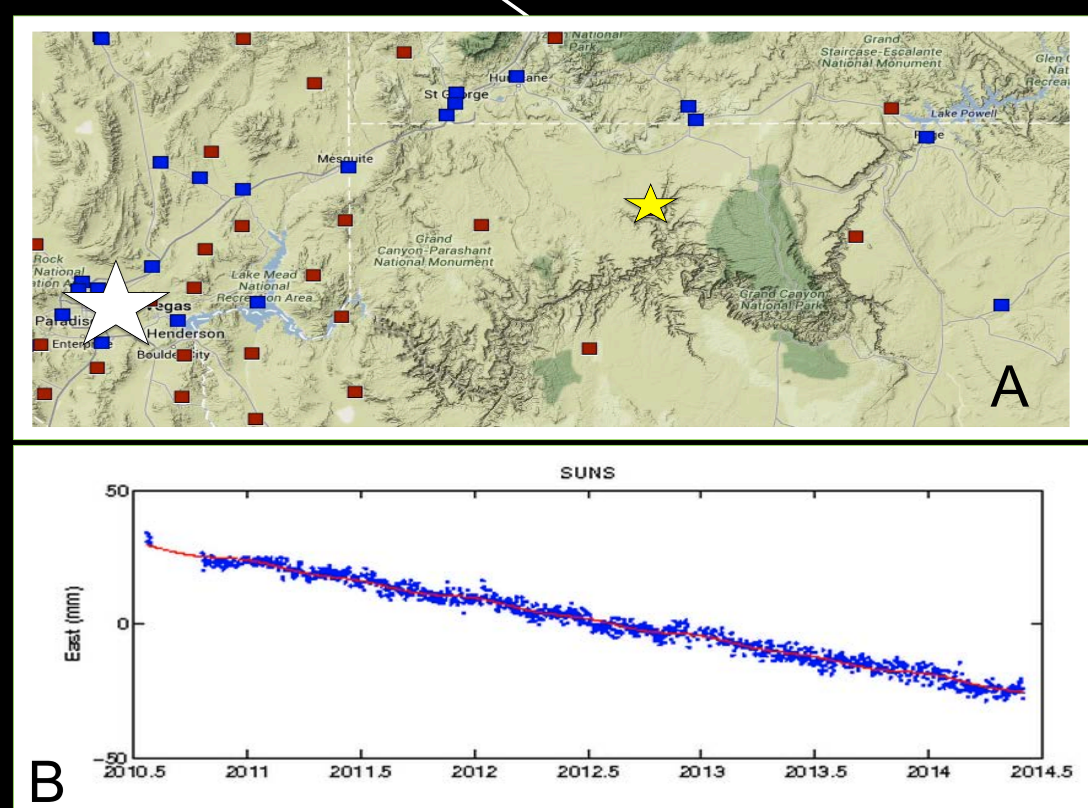


Figure 2. (A) Map view of the Nevada-Utah-Arizona intersection in the neighborhood of Grand Canyon National Park; Yellow Star: Site of MAGNET GPS station, SUNS; White Star: Las Vegas (for reference); (B) Horizontal-East/West component of SUNS position data

Application to Real Data: SUNS (Figure 2) is chosen as a reliable MAGNET station consistently (≥ 2.5 years) producing position measurements from which a seasonal least-squares velocity estimate can be used as truth.

(1) Calculate Theil-Sen velocity estimates for varying time lengths (spanning from 1.1 years to 2.5 years) and chosen dt (spanning from 0 to 30 days) over varying sections in the time-series

(2) Calculate the absolute bias for each estimate from the velocity from the seasonal least-squares estimate over the entire available data (≥ 2.5 years), which is considered the true velocity

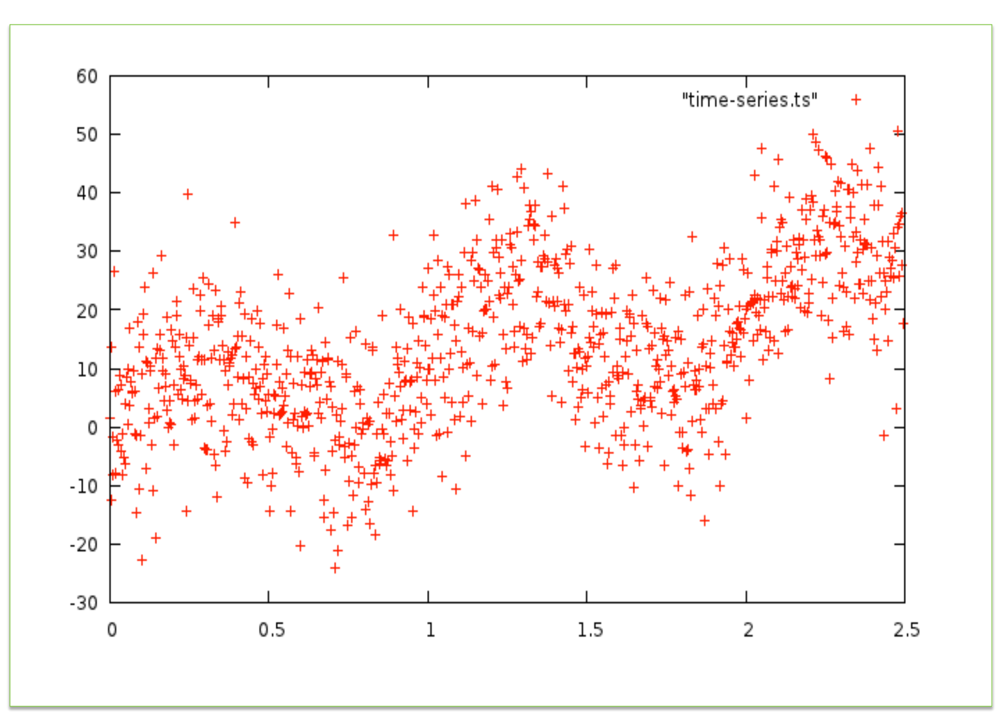


Figure 3. Sample artificially generated time-series with data span length 2.5 years, white noise amplitude 10, seasonal signal amplitude 10/phase 0 radians, true velocity 10 mm/yr, and no flicker noise; two hundred and fifty randomized realizations of each given set of aforementioned parameters contribute to each velocity estimate; X-axis: net time in years; Y-axis: net position in millimeters

Results

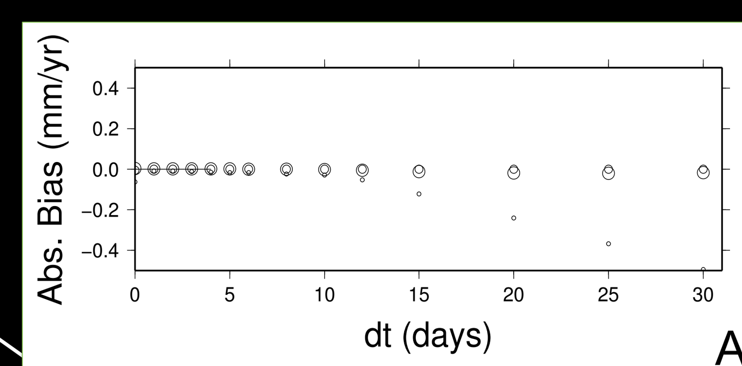
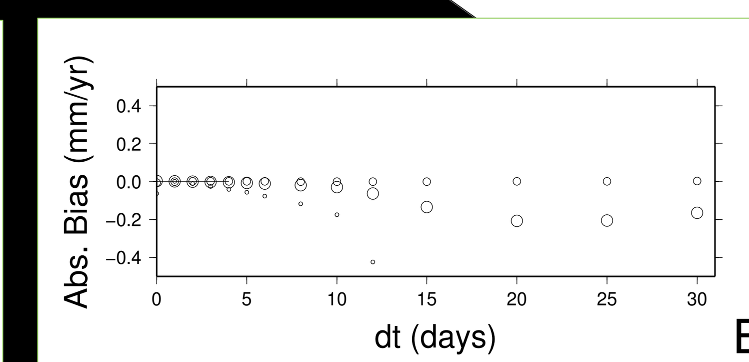


Figure 4. Theil-Sen Seasonal increment, dt (in days) versus absolute velocity bias (in mm/yr) with seasonal signal amplitudes [A] 1 [B] 10; ; circle size is proportional to length (in years) of data span, large: 2.1, medium: 1.5, small: 1.1



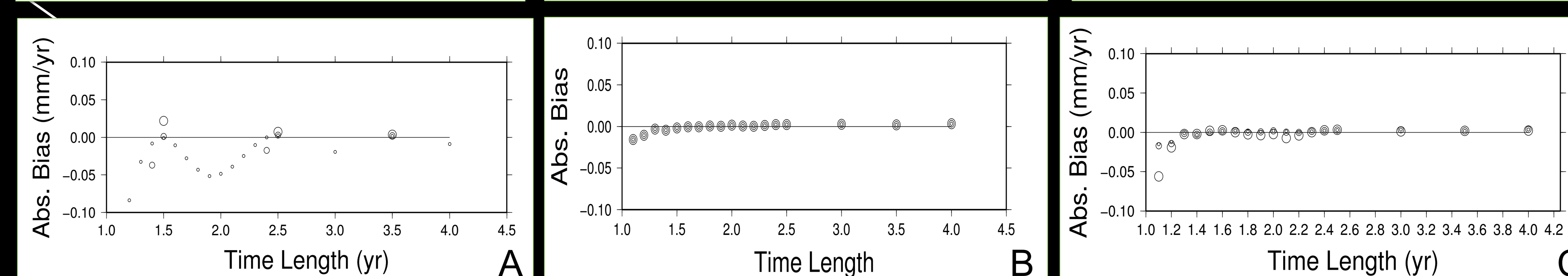


Figure 5. Data span time length (in years) versus absolute velocity bias (in mr/yr) using the [A] non-seasonal least-squares estimator, [B] seasonal least-squares estimator, [C] seasonal Theil-Sen estimator with dt = 5 days; circle size is proportional to amplitude of seasonal signal large: 10, medium: 1, small: 0.1

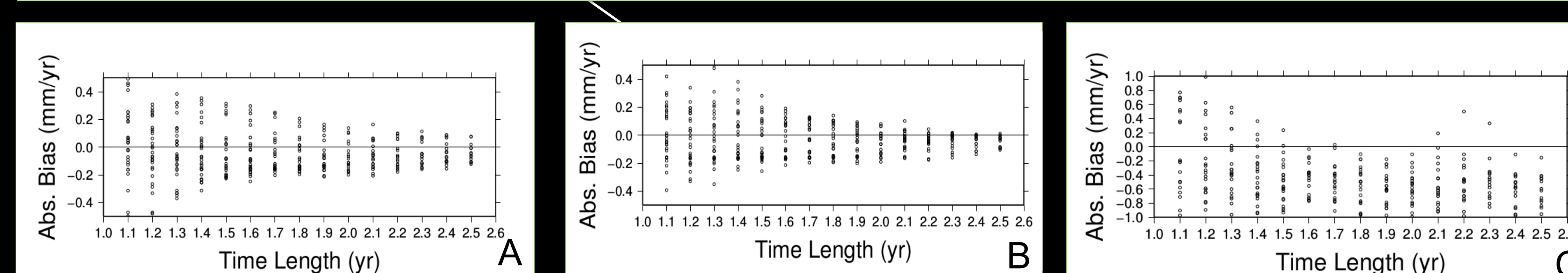


Figure 6. Spread of absolute velocity biases (in mm/yr) for Theil-Sen velocity estimates (using dt = 5 days) for given data span time lengths within each time series (in years); Application to [A] SUNS station, [B] artificial time series mimicking SUNS without flicker noise, [C] artificial time series mimicking SUNS with flicker noise (with α = 1); circle size is proportional to amplitude of seasonal signal large: 10, medium: 1, small: 0.1

Conclusions

(1) Figure 4 reveals the that the ideal dt for a Theil-Sen estimate depends on the amplitude of the seasonal signal for the time-series. Shorter time spans and seasonal signal amplitudes require a smaller dt, whereas the estimator performs similarly for small and large dt on time-series with small seasonal signal amplitudes

(2) Figure 5A confirms the results of *Blewitt and Lavallée* [2002] in that biases become unreasonably large and erratic when applying the least-squares estimator to time series data with seasonal signals. However, Figure 5B reveals that the least squares estimator produces low biases that are comparable to the biases generated by the seasonal Theil-Sen estimator under this methodology. Therefore, the Theil-Sen estimator is proven to be comparable to the seasonal least-squares estimator in terms of bias.

(3) The absolute velocity biases for the Theil-Sen estimator on SUNS (shown in Figure 6A) are discouragingly large. However, the bias spreads produced by varying dt (not shown here) were consistently similar to Figure 6A; therefore, there was the amplitude of the seasonal signal in the chosen time-series was miniscule.

(4) Flicker noise is also concluded to be negligible because Figure 6B resembles Figure 6A much more closely than Figure 6C. The time-series for Figures 6B and 6C mimicked SUNS in most parameters, with presence of flicker noise as the main difference between them.

Discussion

While this study is rigorous in realizations of white noise, seasonal signal, phase, and relevant data span lengths, it lacks considerations of other phenomena in GPS time-series. Such phenomena include flicker noise, additional colored noise, and steps. Therefore, the current findings that the seasonal Theil-Sen estimator is, at best, equally as effective as the seasonal least-squares estimator may be premature. The additional garnered information about the seasonal Theil-Sen estimator will provide a foundation for future studies comparing these estimators. These future studies might seek to increase the rigor in simulating time-series with more detailed noise. The comparison of these estimators is also important to view with more rigor in the context of real data from stations like SUNS.

Acknowledgements

This project would not have been possible without the collaboration and insight of Dr. Geoff Blewitt (University of Nevada, Reno). Thank you to Bill Hammond for his field work organization and to Bret Pecoraro for his steadfast and expert maintenance and data downloading on the ever-growing networks of GPS stations. Thank you to Yang Zhang for his programming assistance and to Cheri Briggs for her administrative support. Lastly, thank you to the UNAVCO/RESESS, NSF, ExxonMobil, the University of Nevada Reno, the Nevada Bureau of Mines and Geology, and the Nevada Geodetic Lab for their support in this adventure.

UNAVCO NSF ExxonMobil nbmg Nevada Geodetic Laboratory