**Autocorrelation of Astronomical Time-Series Data**

Thomas Cannon - MATH550

The Autocorrelation Function (ACF) is a correlation of a signal – normally time series – onto itself, within the family of time-series analysis. The self-correlation is offset by a delay called a *lag.* In other words, it will correlate a value given in a current point in time to a point in the future relative to the current point, for all of the signal. It is commonly used to help find patterns that may repeat themselves in signal data, and can be especially helpful in data containing noise. Jefferey Scargle presents a practical overview of the ACF in the context of time-series astronomical data in his 1989 paper “Studies in astronomical time series analysis. III - Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data” (Scargle 1989). We will attempt to understand the ACF with a discussion of the section on Autocorrelation in Scargle’s paper as well as an implementation of ACF onto time-series astronomical data of Gamma-Ray Bursts (GRB’s) collected by the Burst and Transient Source Experiment (BATSE) onboard NASA’s Compton Gamma-Ray Burst Observatory (CGRO). Images of the GBR light curves and their ACF lag plots (Figures 1 to 6) are available in this paper’s appendix and will be referenced throughout.

Scargle begins by defining the ACF, as we have done above, and then he quickly draws attention to the fact that nearby bins of data – what he calls “sidelobes” – may influence the frequencies of lag plots. This happens because data that is improperly binned will tend to bleed into adjacent bins. Scargle argues that this problem can be helped a little by introducing even sampling, which we fortunately have by the data that we will be representing.

Scargle then starts to write about an example of using the ACF in an example Poisson process. Our research in the field of GRB’s also contains Poisson processes, hence the motivation behind the following analysis. GRB’s are transient astronomical events discovered in the 70’s, observed randomly about the sky, and occurring about once daily. The data from these events present themselves as increases in gamma-ray radiation on timescales ranging from milliseconds to minutes. It is widely accepted that GRB’s are composed of pulses – distinct burps of this gamma radiation (clearly seen as two pulses in Fig 1 and 3). What is not understood is how these events are created. A current avenue of research is to perform analysis on the characteristics of the GRB pulses relative to each other as they exist in the bursts. ACF may be an interesting technique to apply to GRB’s in hopes of identifying relationships in these systems.

We randomly select three bursts from a sample of over 8000 event triggers from CGRO’s BATSE and perform the ACF technique defined in the R standard library. The first burst selected – 0143 (Fig 1) – is incredibly bright by GRB standards. It also has the obvious feature of two very distinct pulses within the burst. Having two pulses as such is not a common event. Because of clarity of these two distinct pulses, we expect the ACF to be able to pick out this feature, which it does (Fig 2). At a lag of about 750, there is a very clear peak. While this is interesting, The ACF also discovered something else that is potentially fundamental about GRB pulses.

It has been proposed that GRB pulses contain signatures of shocks (Hakkila 2014). These features present themselves as periodic oscillations, overlaid on top of the intrinsic pulse shape found in the light curve data. These oscillations are often difficult to discern and obscured by a lot of noise. Evidence of these oscillations can be seen in the lag plot of 0143 (Fig 2). As the ACF correlation initially drops in roughly the first 150 lag increments on the plot, an overlaid oscillation is visually evident. More analysis will need to be done to verify that it is truly periodic, but such a visual turnout is promising. Furthermore, the shock signature time scales have been proposed to exist independent of the burst and pulse timescale. In Burst 3307 (Fig 3) it is possible to interpret oscillatory patterns in the lag plot at a different scale relative to the burst (Fig 4) and at a different time scale relative to burst 0143.

Some GRBs like 0143 have a very high signal relative to the background noise, making the structure easy to identify. However, there are numerous events that have a very low signal relative to the background noise. One in particular found its way into our sample – Burst 3370 (Fig 5). Burst 3370 is probably the most common type of GRB, commonly referred to as a “single pulse”. For GRB’s of a single pulse, we would not expect to see much in the way of ACF lag past the initial spike in the beginning, as can be seen in the lag plot for 3370 (Fig 6). However, like the two bursts before, there seems to be a very slight pattern emerging through all of the noise. As to determining if it is random or structural, it is out of the scope of this paper.

The process to implement ACF, inspired by the explanations and examples from Scargle concerning Poisson time-series processes, was executed in the R programming language (version 3.6.1), available in the appendix. While noting definitive was proven within the scope of this paper, there is potential for ACF to be used in the analysis of GRB light curves, as structure of significance seems to become evident in the ACF lag plots.

**References**

Hakkila, J., Preece, R. 2014. “Gamma-Ray Burst Pulse Shapes: Evidence for Embedded Shock Signatures?”. arXiv:1401.4047.

Scargle, J. 1989, “Studies in astronomical time series analysis. III - Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data”. ApJ, 343, 874S.

**BATSE Burst Trigger 0143**

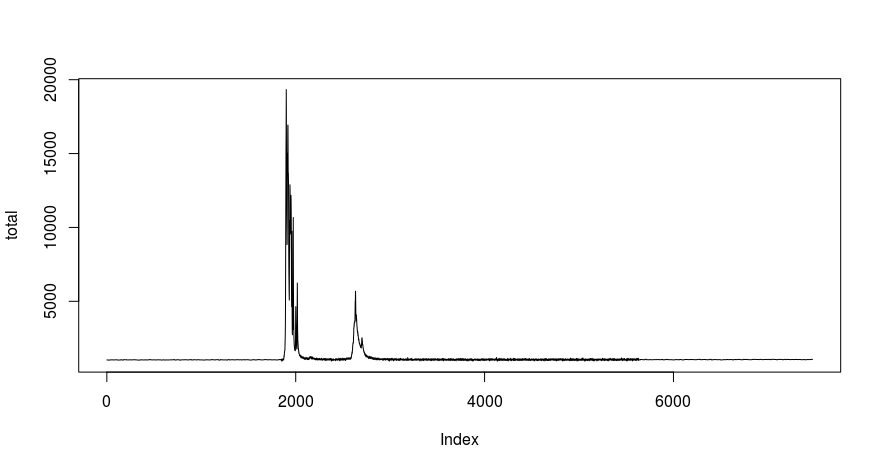


Figure 1: 0143 Light Curve

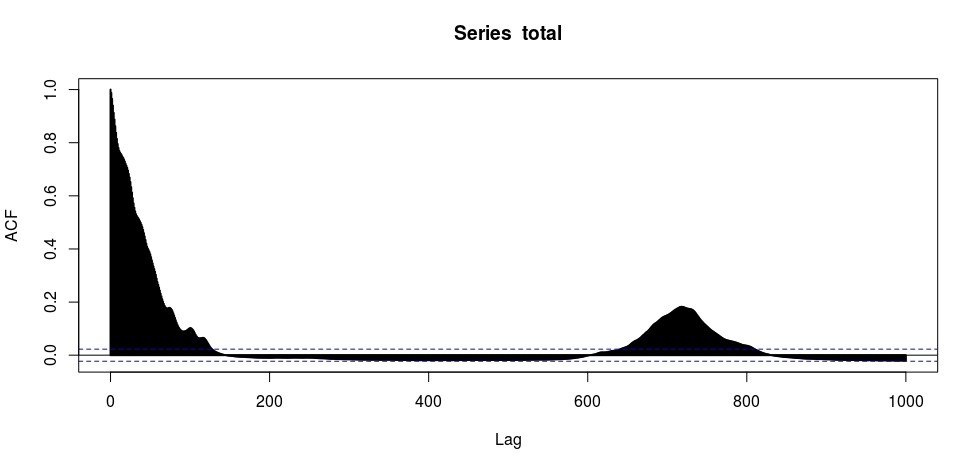


Figure 2: 0143 ACF Lag Plot

**BATSE Burst Trigger 3307**

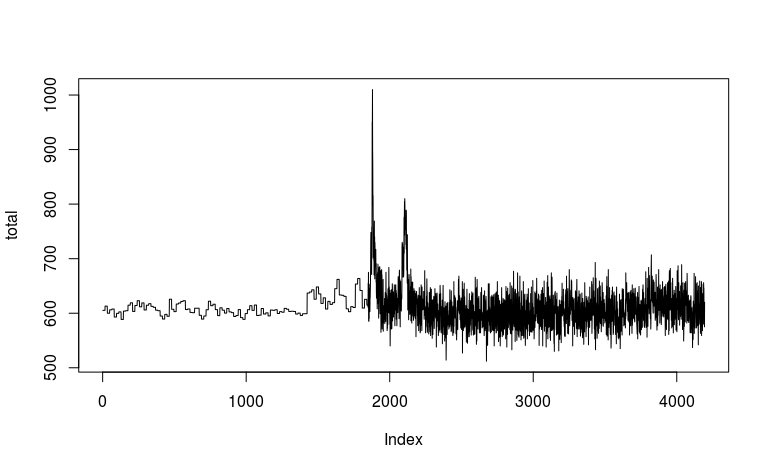


Figure 3: 3307 Light Curve

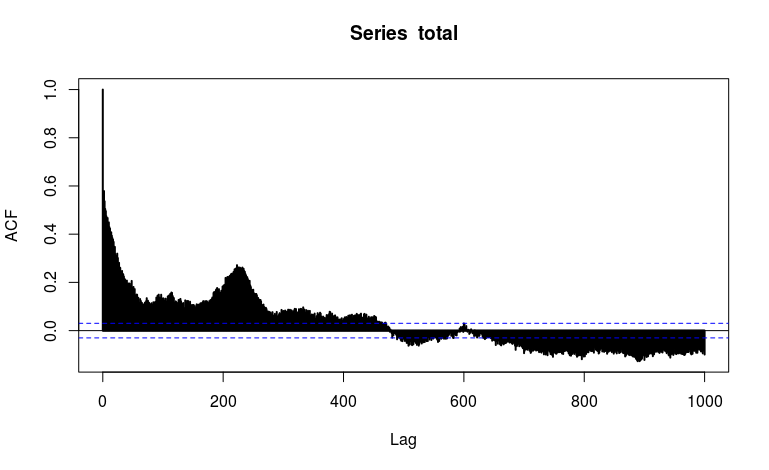


Figure 4: 3307 ACF Lag plot

**BATSE Burst Trigger 3370**

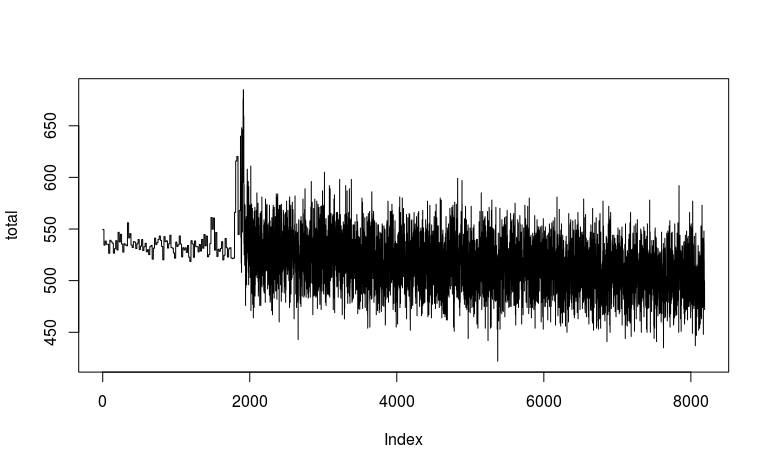


Figure 5: 3370 Light Curve

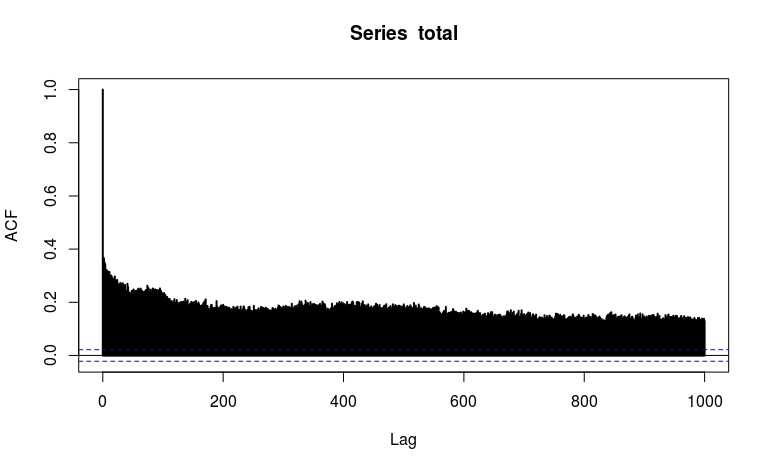


Figure 6: 3370 ACF Lag Plot

**R Code**

burst <- read.csv("/path/to/burst/cat64ms.00143",header=TRUE,sep=",")

attach(burst)

total <- CHAN1+CHAN2+CHAN3+CHAN4

total

plot(total, type="l")

acf <- acf(total,plot=TRUE,lwd=2,lag.max=1000)