

# The use of differencing to remove spurious correlations in models of geostationary 2 MeV electron flux: comparison to spectral subtraction

Laura E. Simms, Augsburg University and University of Michigan, *contact: simmsl@augsborg.edu*

ML-Helio 2022

Mark Engebretson, Augsburg University

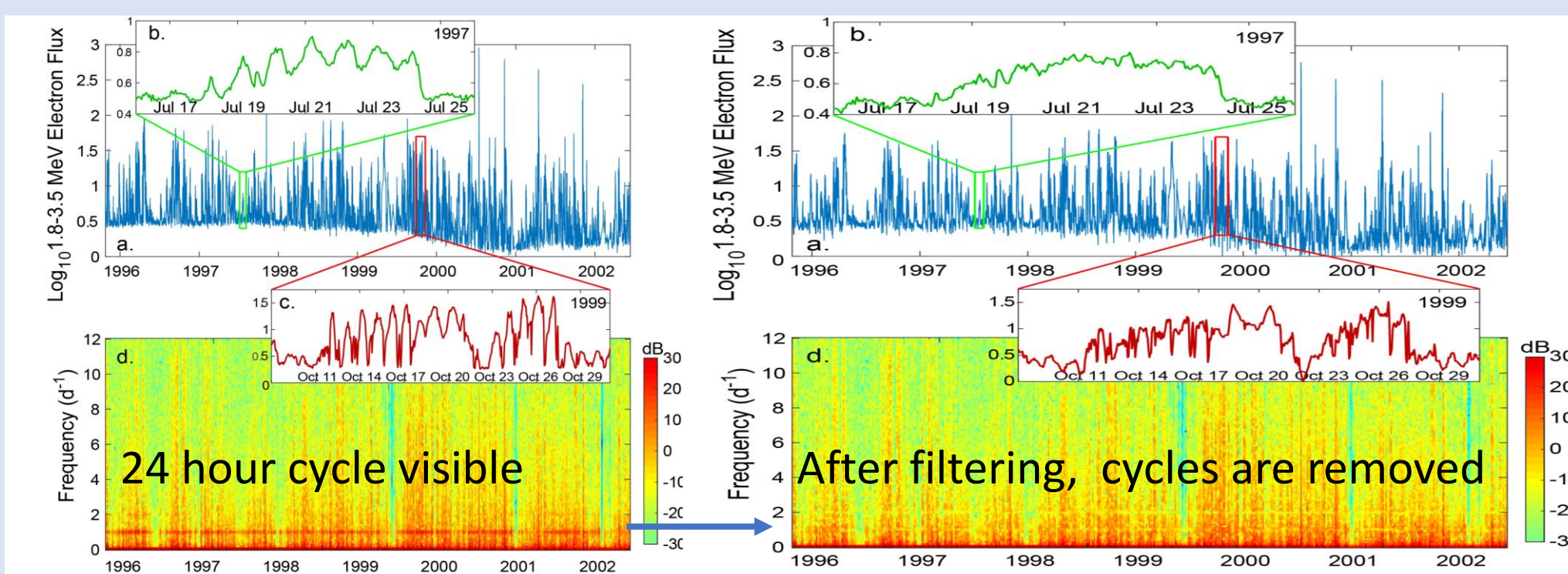
JGR: <https://doi.org/10.1029/2021JA030021>

Geoff Reeves, Los Alamos National Laboratory

- Periodicity is common in space weather data. -- Cycles range from diurnal to the 27 day and 11 year solar cycles
- Removing common cycles reduces spurious correlations
- Diurnal cycles can be removed from electron flux and a ULF index using either a bandstop filter or a differencing transformation ( $y_t - y_{t-24}$ )
- Differencing is more useful as it removes both the 24 hour noise signal and the 27 day solar cycle
- With common cycles removed, correlations between these parameters are more valid, but lower ( $\leq 0.11$ ) than previously published values

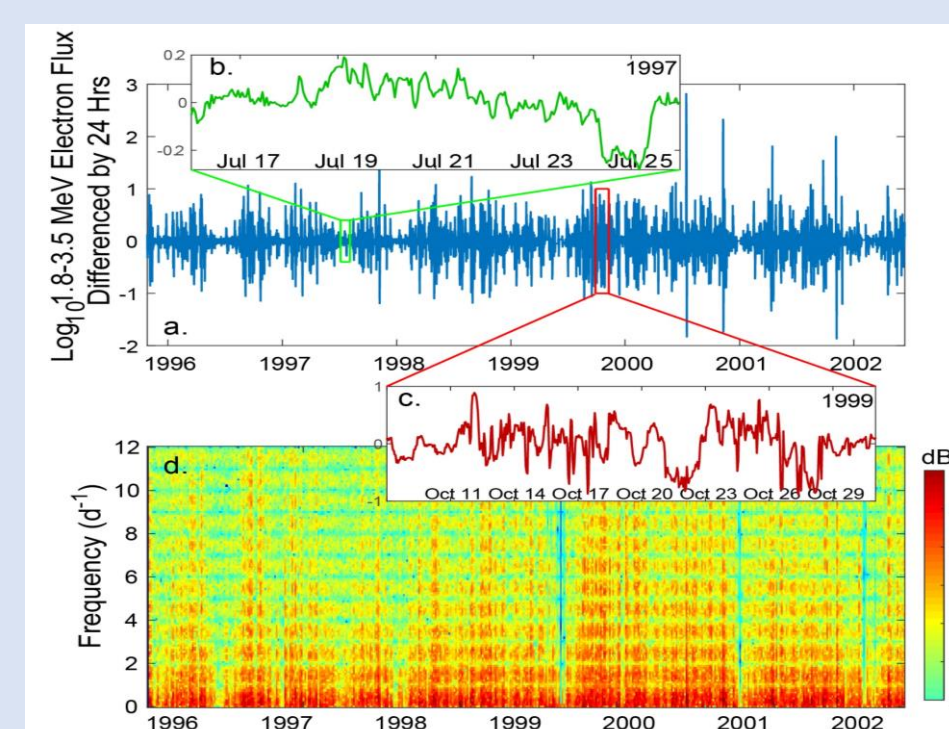
## One Approach: Bandstop filter to remove 24 hour signal

Using an FIR equiripple bandstop filter, we successfully remove both the 24 hour and 12 hour noise signals from electron flux data (LANL satellite at geosynchronous orbit)

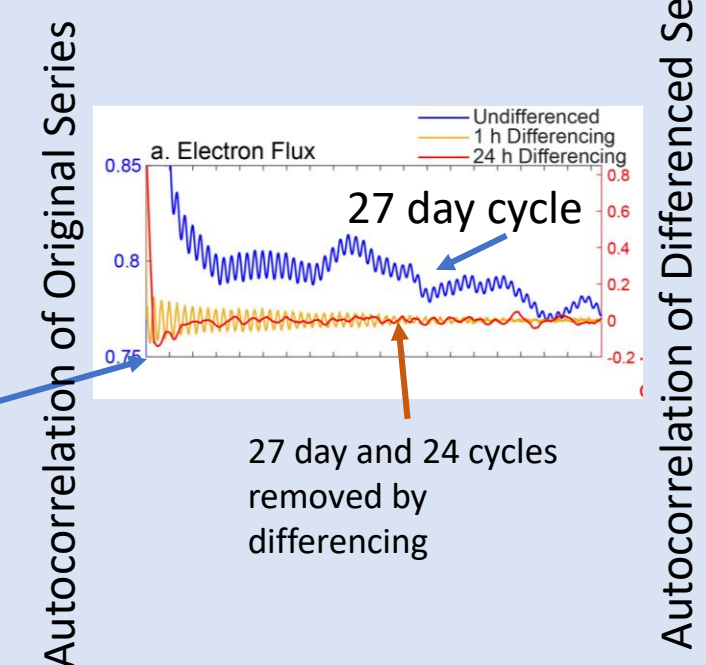


## Second Approach: Difference transformation ( $y_t - y_{t-24}$ ) removes both the 24 hour signal and the 27 day solar cycle

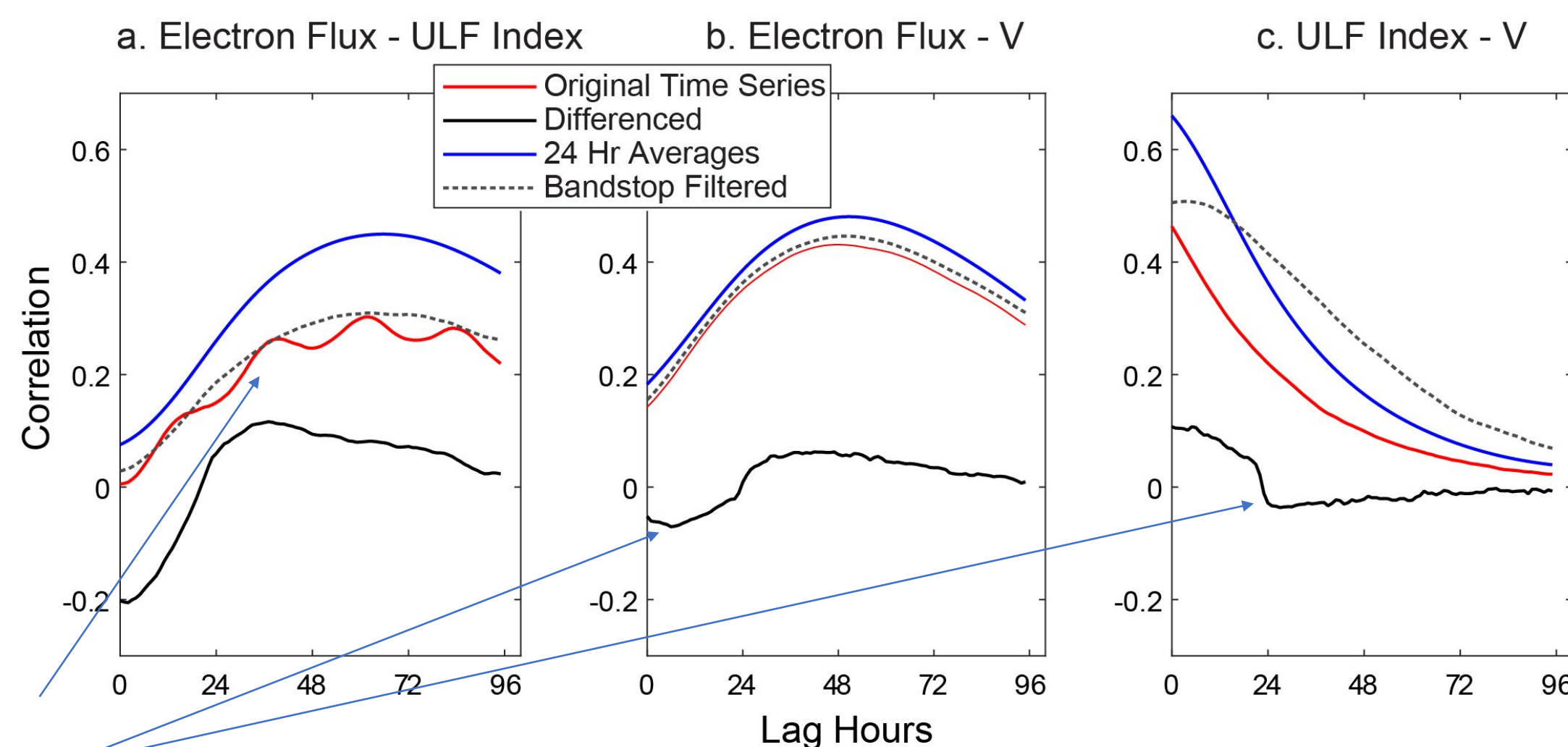
Subtracting the flux 24 hours previous from every observation removes the 24 h frequency as well as higher frequencies



Differencing also removes longer cycles, including the 27 day solar cycle



- Correlations between electron flux, a ULF index, and solar wind velocity ( $V$ ) are all lower if common cycles are removed using the difference transformation ( $y_t - y_{t-24}$ )
- Spectral subtraction removes 24 cycle but not 27 day cycle: higher correlations are result of this longer cycle not being removed



- Differenced correlation is the most valid as cycles are removed
- Differenced correlation is also the lowest
- Most of the correlation between electron flux and the ULF index is therefore due to common cycles
- Previously published high correlations suggested a strong physical relationship between these variables over these long time periods. With cycles removed, these associations are no longer as strong

## Correlations of Differenced Series Are Lower – Because Common Cycles Are Removed

### Data and Methods

- Hourly averaged  $\log_{10}$  electron fluxes ( $\log(\text{electrons}/(\text{cm}^2/\text{s}/\text{sr}/\text{keV}))$ ) in the 1.8-3.5 MeV range from the Los Alamos National Laboratory (LANL) Energetic Spectrometer for Particles (ESP) instrument located at geosynchronous orbit on the 1994-084 satellite.
- Period limited to 25 Oct 1995 – 13 Jun 2002 when there was a minimum of missing data. For spectral analysis and bandstop filtering, short periods of missing values were interpolated from surrounding values.
- Kozyreva et al. (2007) ULF Pc5 index provides an hourly measure of Pc5 (2–7 mHz) ULF power (in  $\text{nT}^2/\text{Hz}$ ) observed in the local time range from 0500 to 1500 hours by ground-based magnetometers stationed between 60 and 70°N corrected geomagnetic latitude.
- Hourly averages of the solar wind velocity (km/s) from the OMNIweb database.
- Analyses and filtering performed in MATLAB. Spectrograms were obtained from the MATLAB specgram procedure. This computes the (Hanning) windowed discrete-time Fourier transform (FFT) of a signal using a sliding window of 256 points and a 50% overlap (128 points). We used a sampling frequency of 24 hours/day.
- Work at Augsburg University was supported by NSF grants AGS-1651263 and AGS-2013648.