Cosmic Ray Flux Correlation between McMurdo and Jang Bogo Stations

E. Kittiya,[a,*] W. Nuntiyakul,[a] P. Evenson,[b] A. Saiz,[c] D. Ruffolo,[c] S. Oh[d] and A. Seripienlert[e]

[a]Chiang Mai University, THAILAND

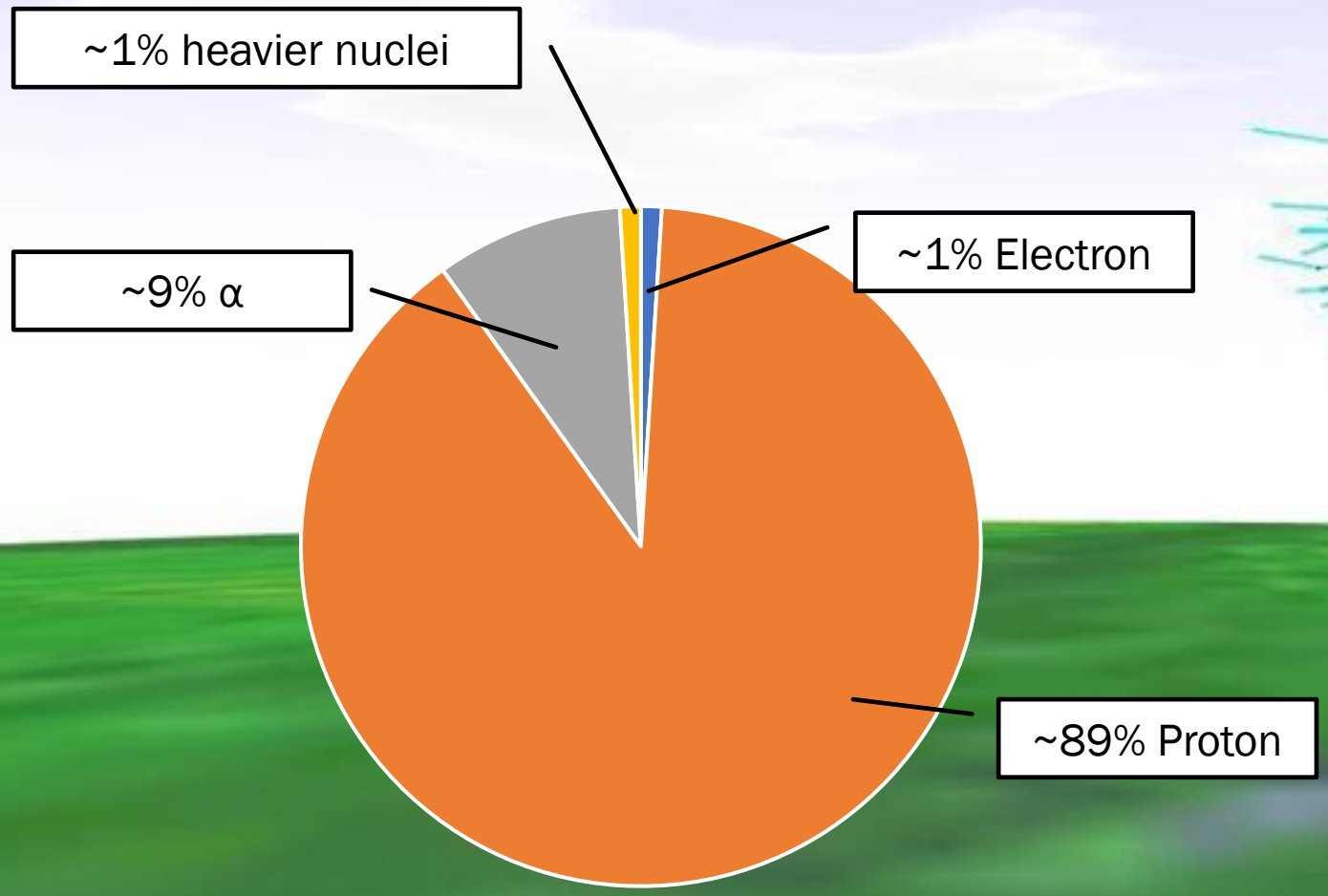
[b]University of Delaware, USA

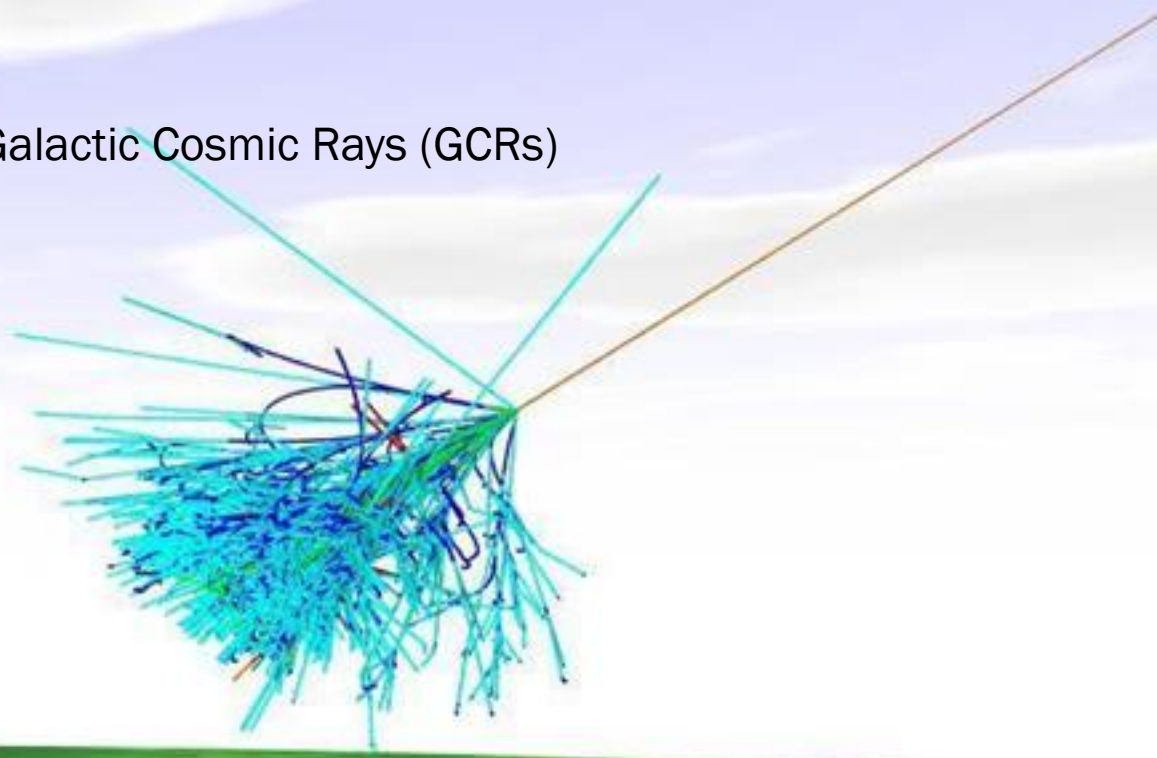
[c]Mahidol University, THAILAND

[d]Chonnam National University, SOUTH KOREA

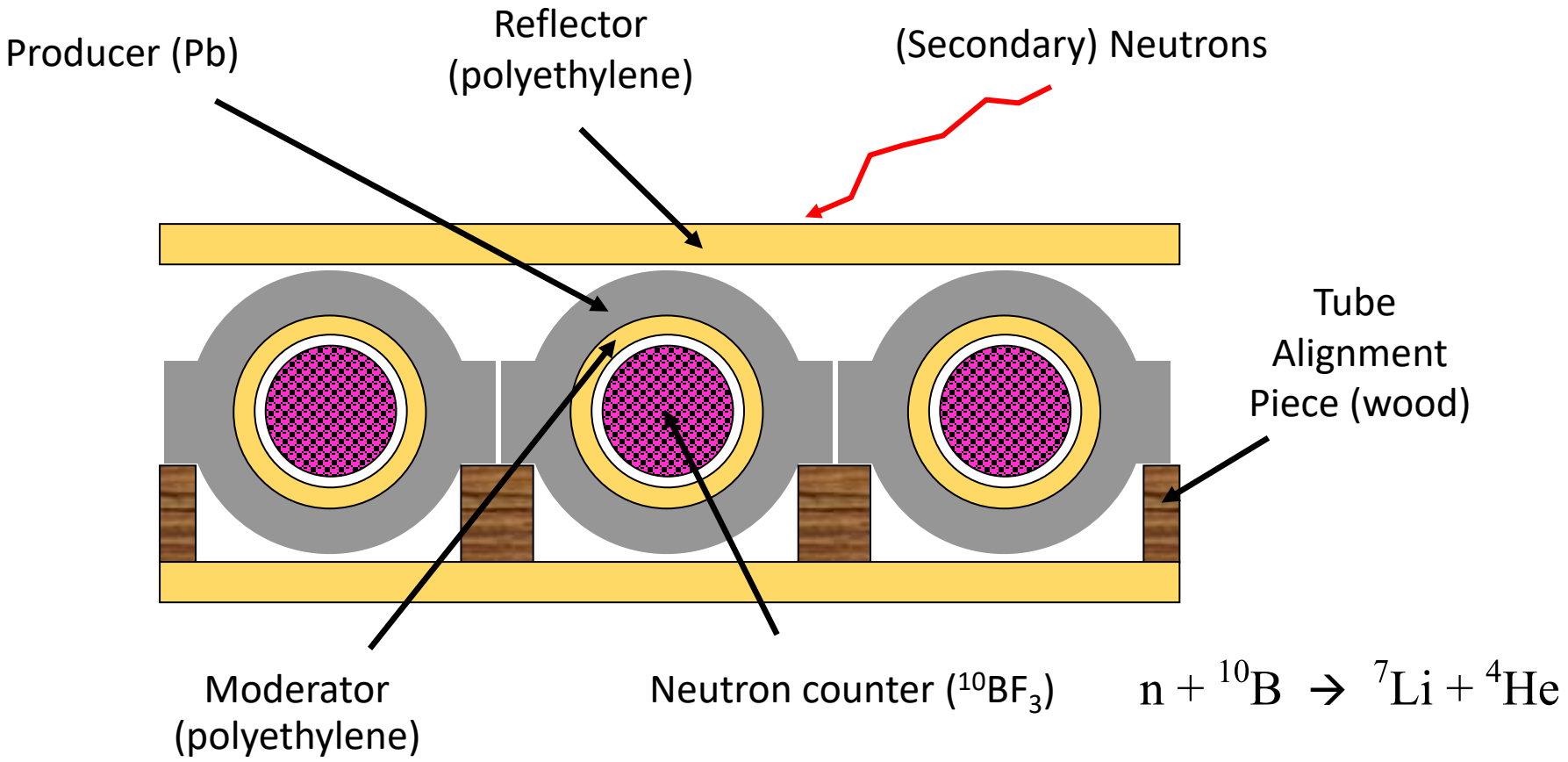
[e]National Astroniomical Research Institute of Thailand (NARIT), THAILAND

Solar Energetic Particles (SEPs) and Galactic Cosmic Rays (GCRs)





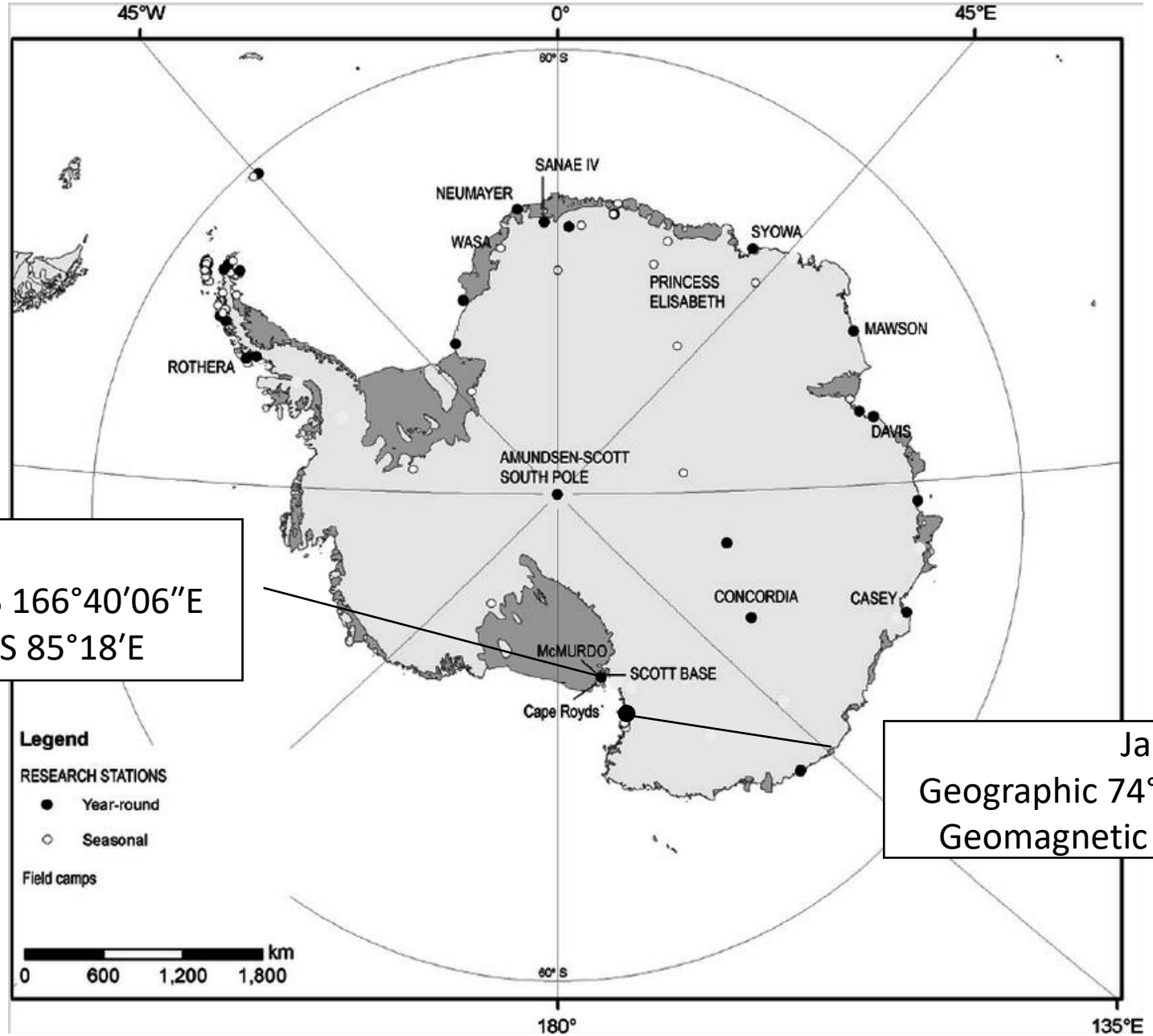
Hajo Drescher, Frankfurt U.



courtesy: Nuntiyakul et al.

Figure 3 Standard Neutron Monitor

McMurdo
Geographic 77°50’47”S 166°40’06”E
Geomagnetic 77°3’S 85°18’E



Jang Bogo
Geographic 74°37’26”S 164°13’44”E
Geomagnetic 78°58.8’S 72°22.8’E

McMurdo and Jang Bogo stations

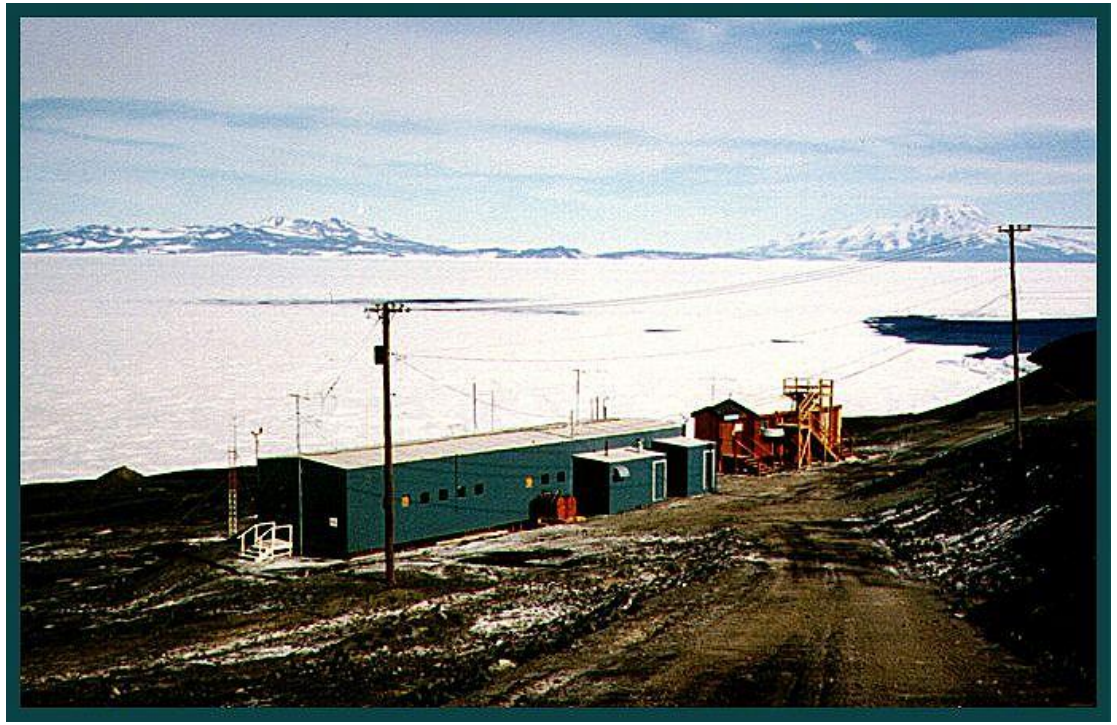
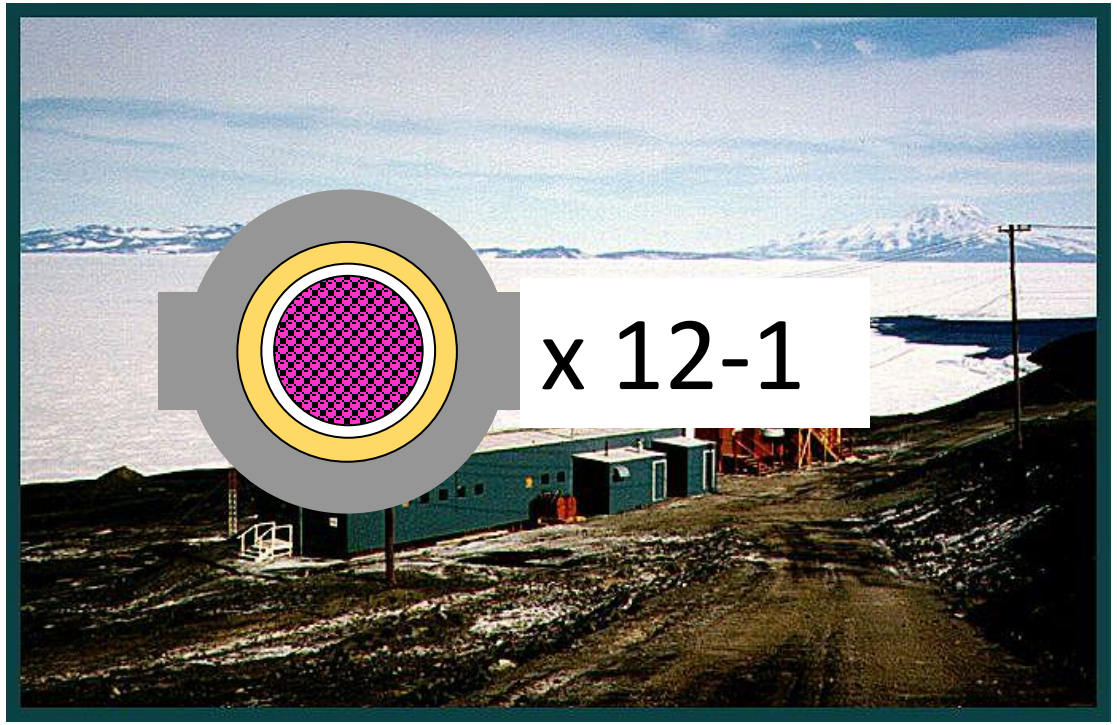


Figure 1 Bird-eye view of McMurdo Station (courtesy: nmdb database)

Figure 2 Bird-eye view of Jang Bogo Station (courtesy: KOPRI)

McMurdo and Jang Bogo stations



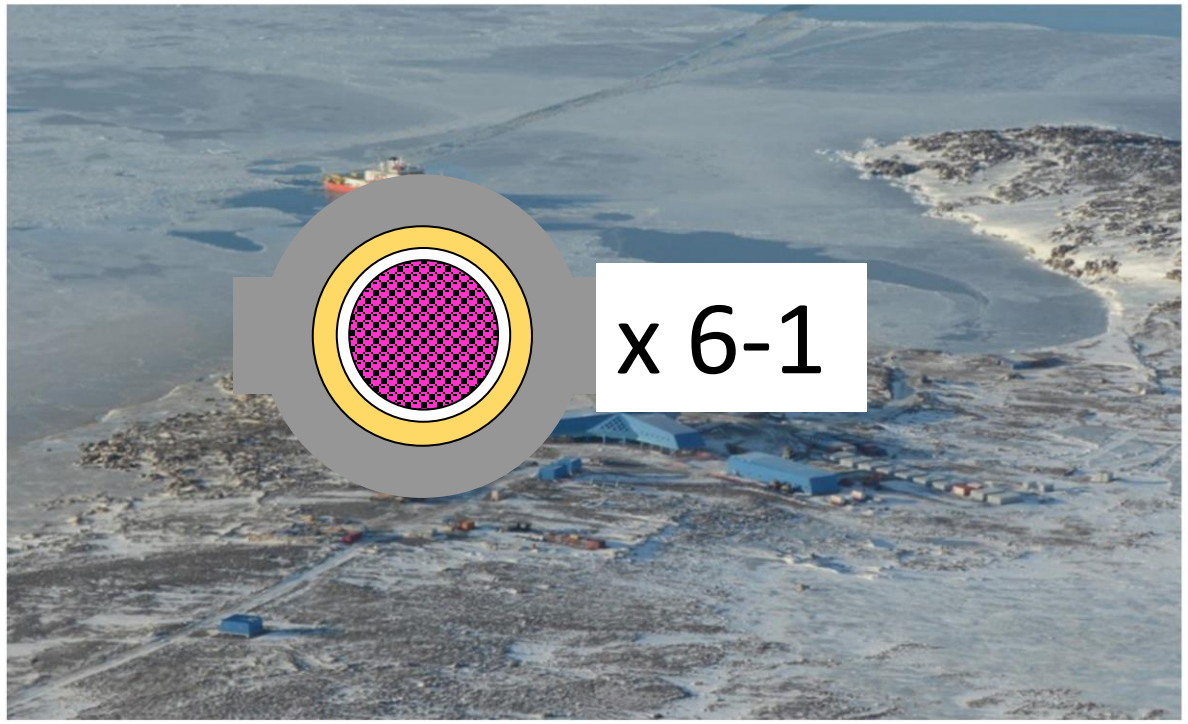
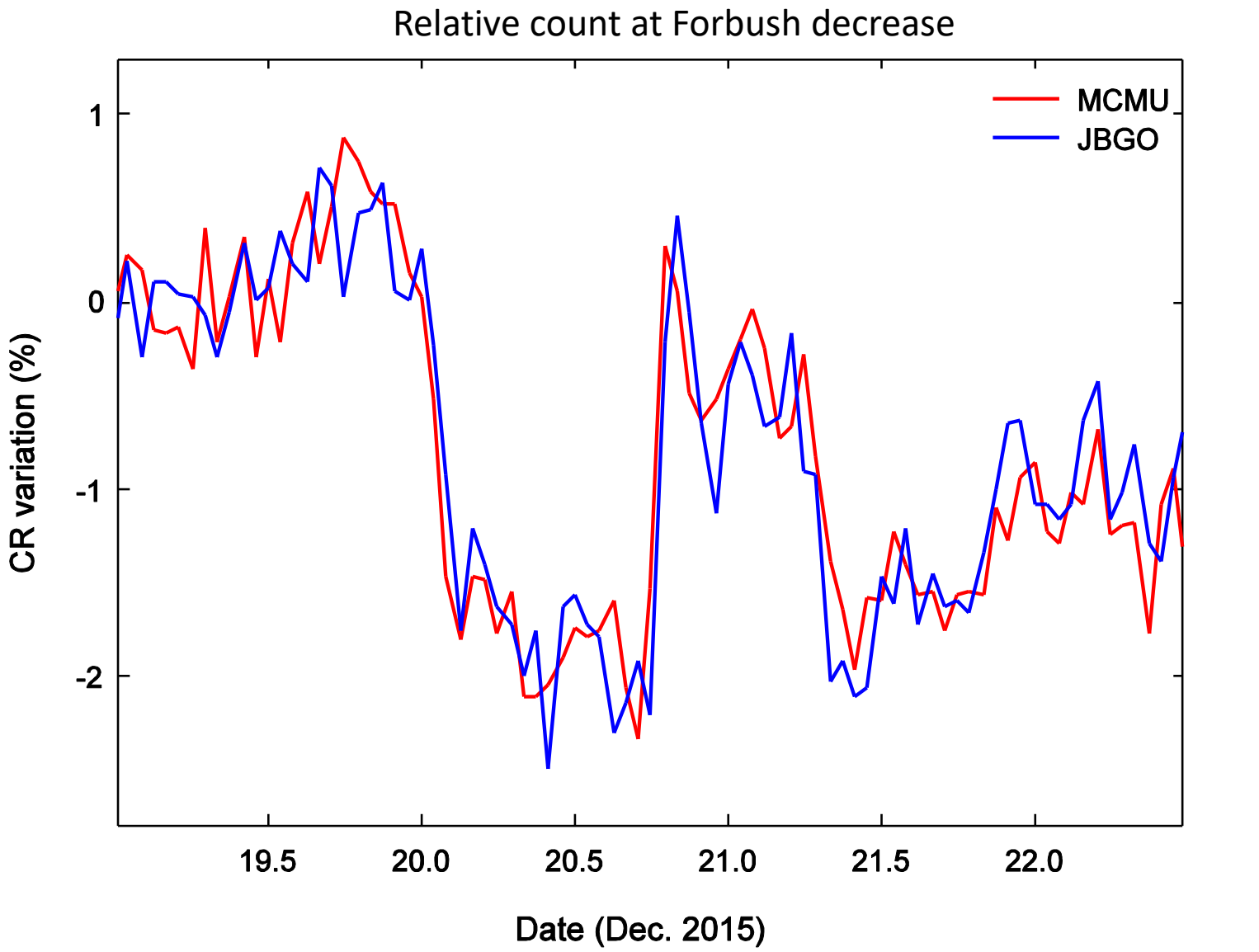


Figure 4 12-1 counter tubes located at McMurdo during December 2015 to October 2016

Figure 5 6-1 counter tubes located at Jang Bogo during December 2015 to October 2016

Data during December 2015 to January 2017 has been analyzed in this work!



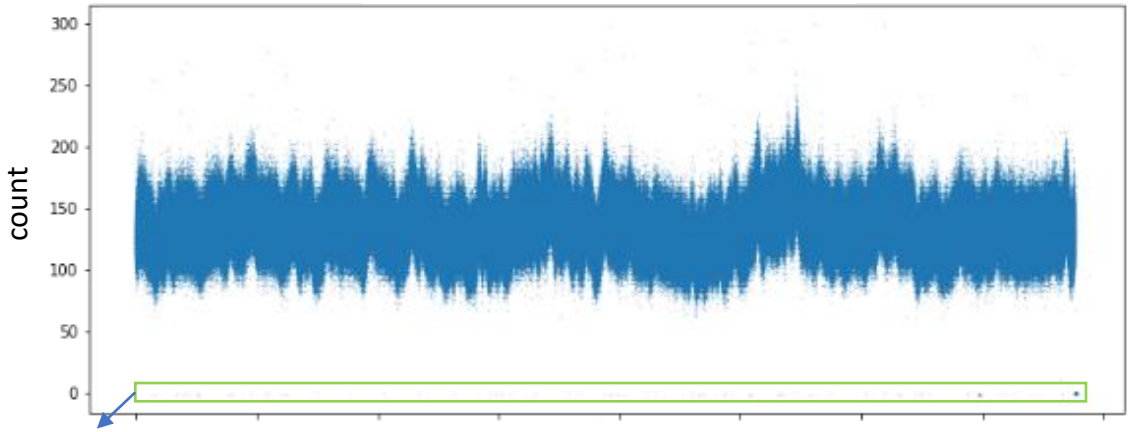
Nuntiyakul et al.

Data correction and cleanup

P_0 is defined for each stations

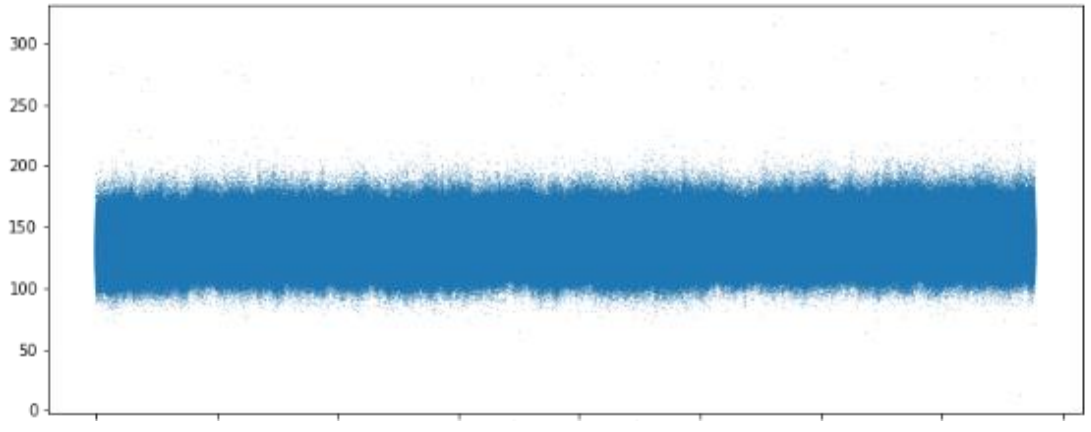
$$c_i = c_i^* e^{-\beta(P-P_0)}$$

Uncorrected count (c_i^*)

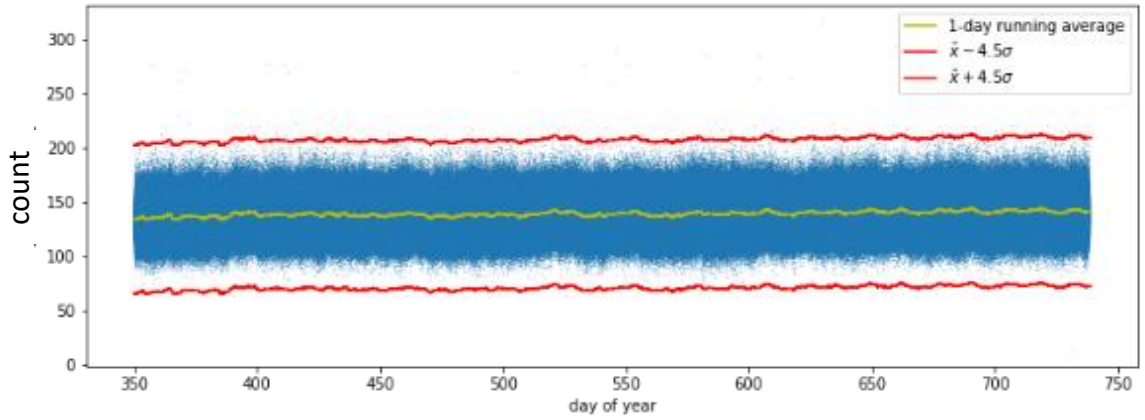


Missing data marked as 0 or -1

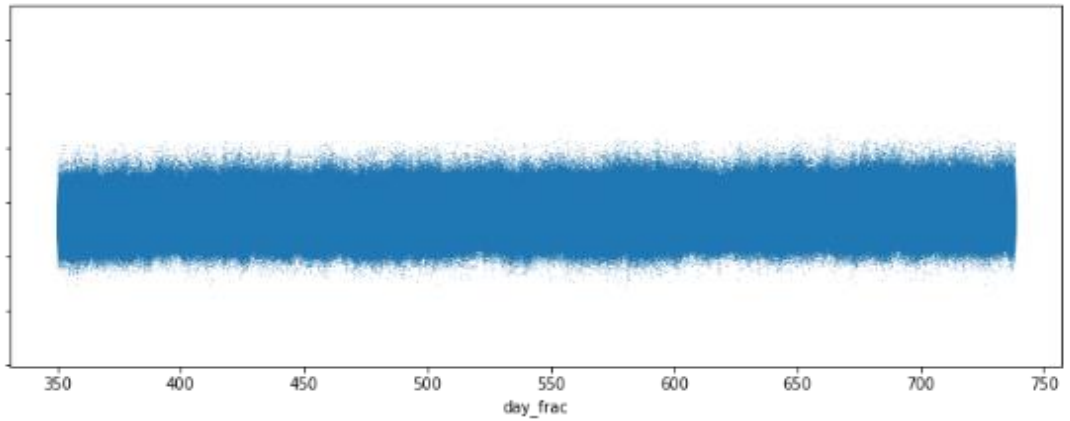
Corrected count c_i



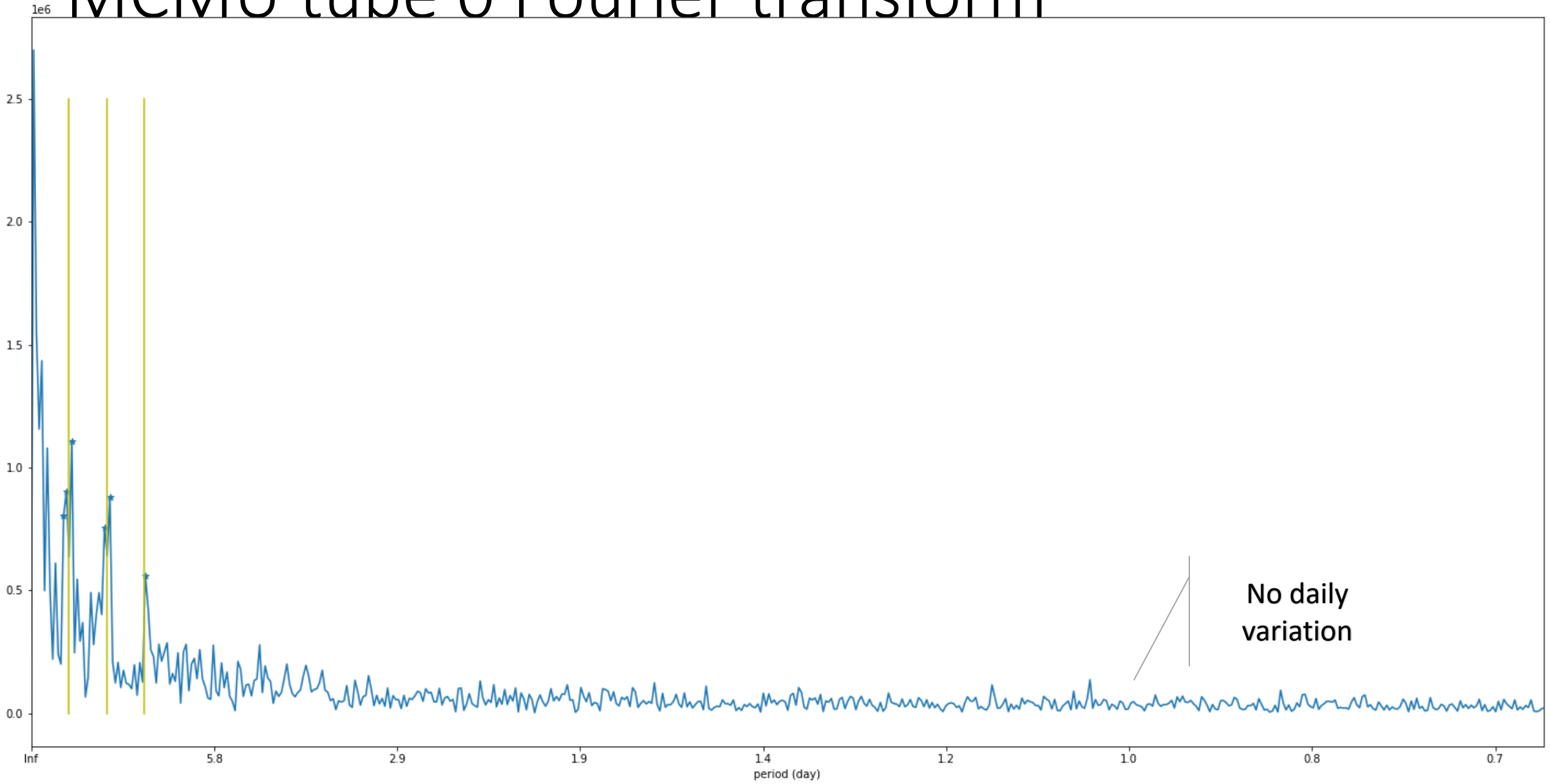
Corrected count c_i



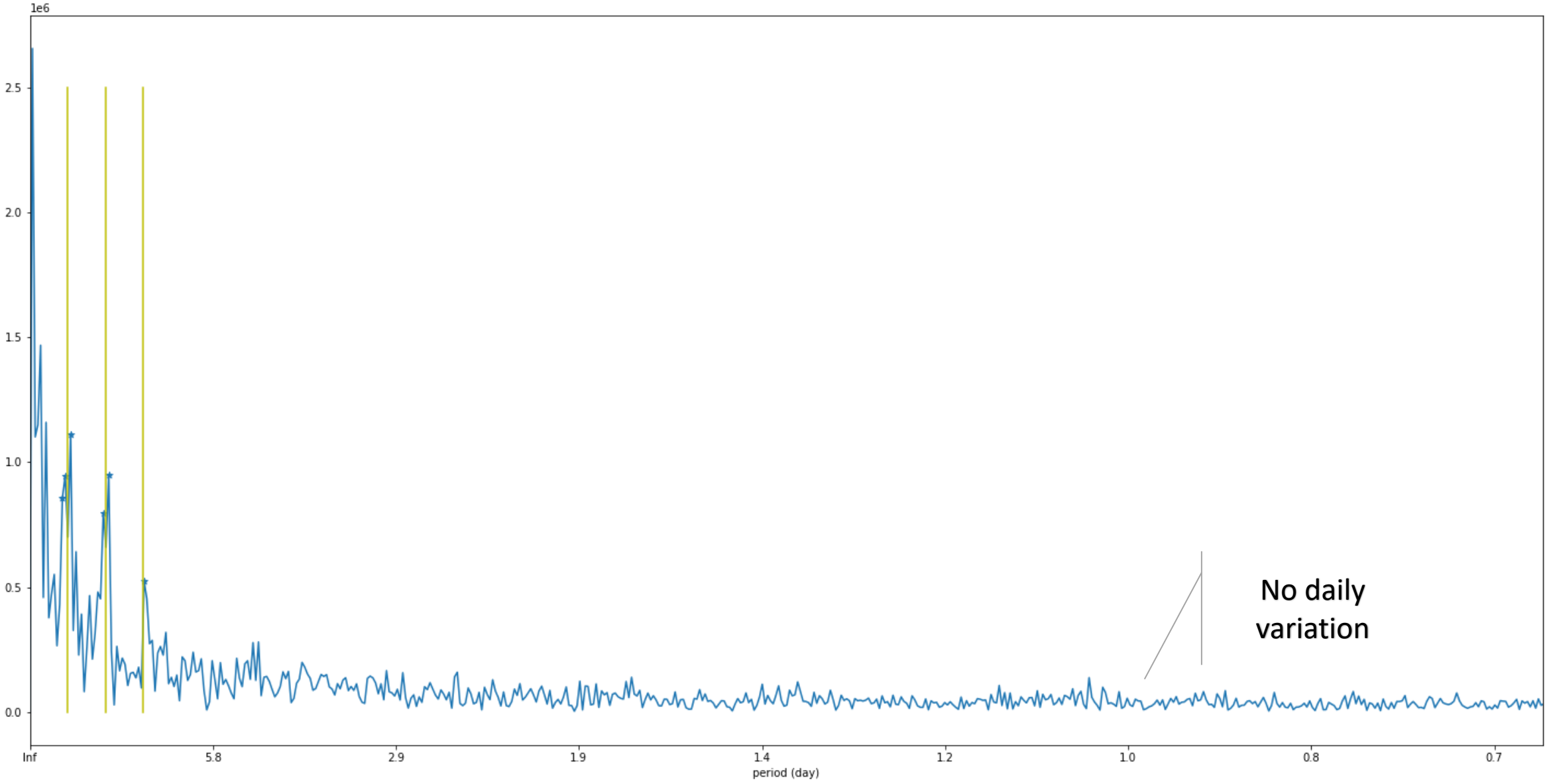
Corrected count c_i w/o outlier



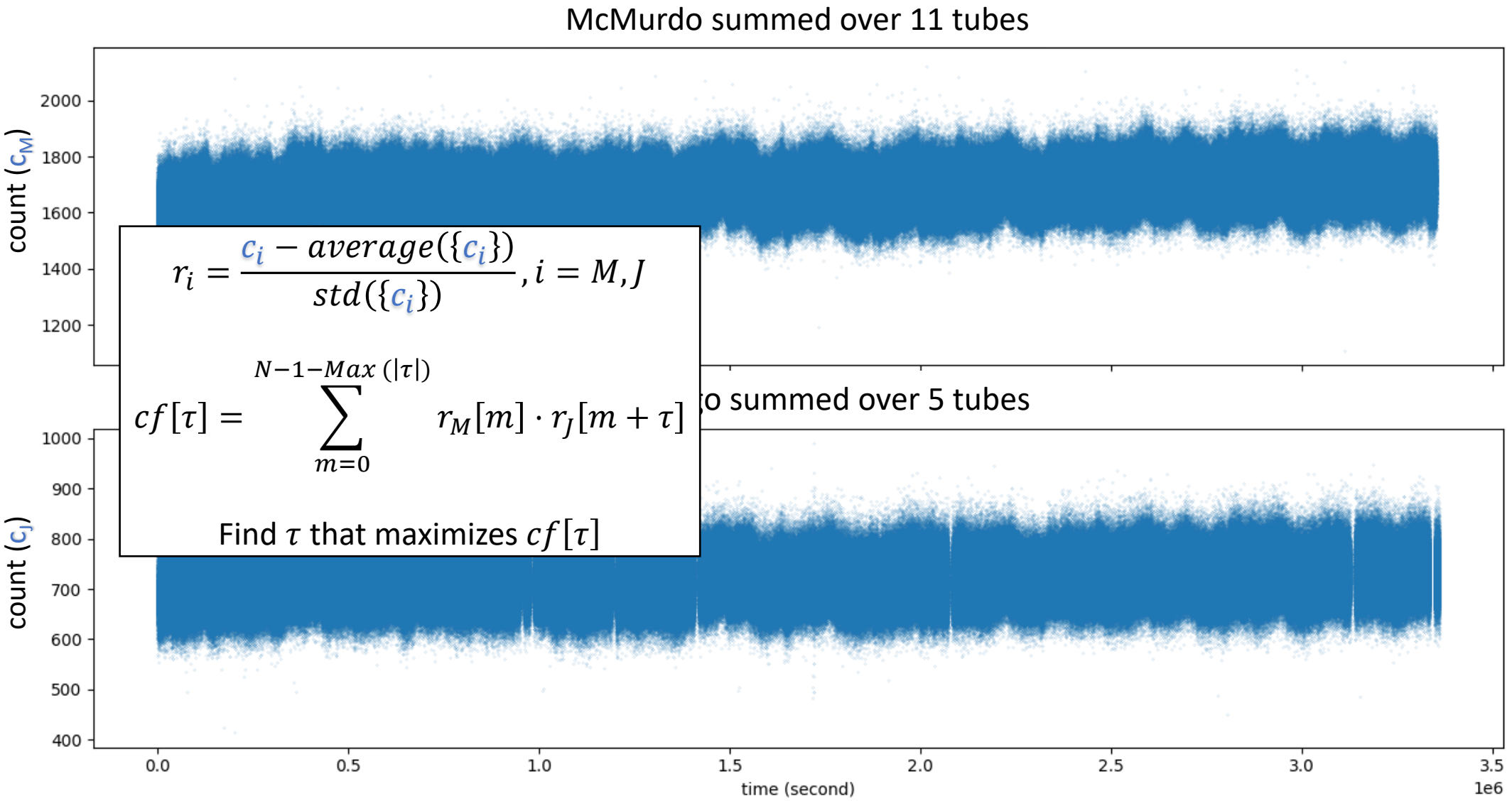
MCMU tube 0 Fourier transform

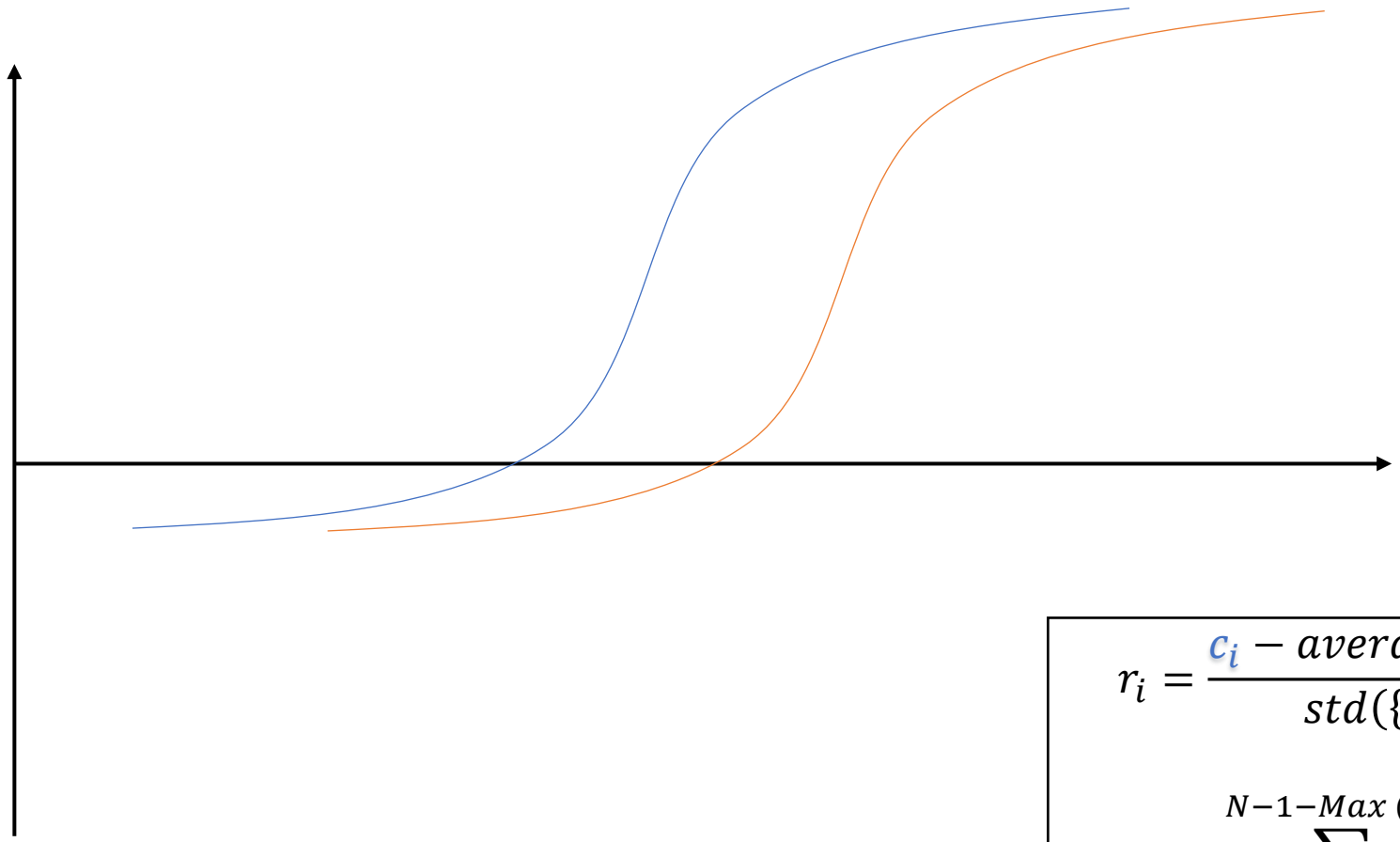


JBGO tube 0 Fourier transform



Whole station count

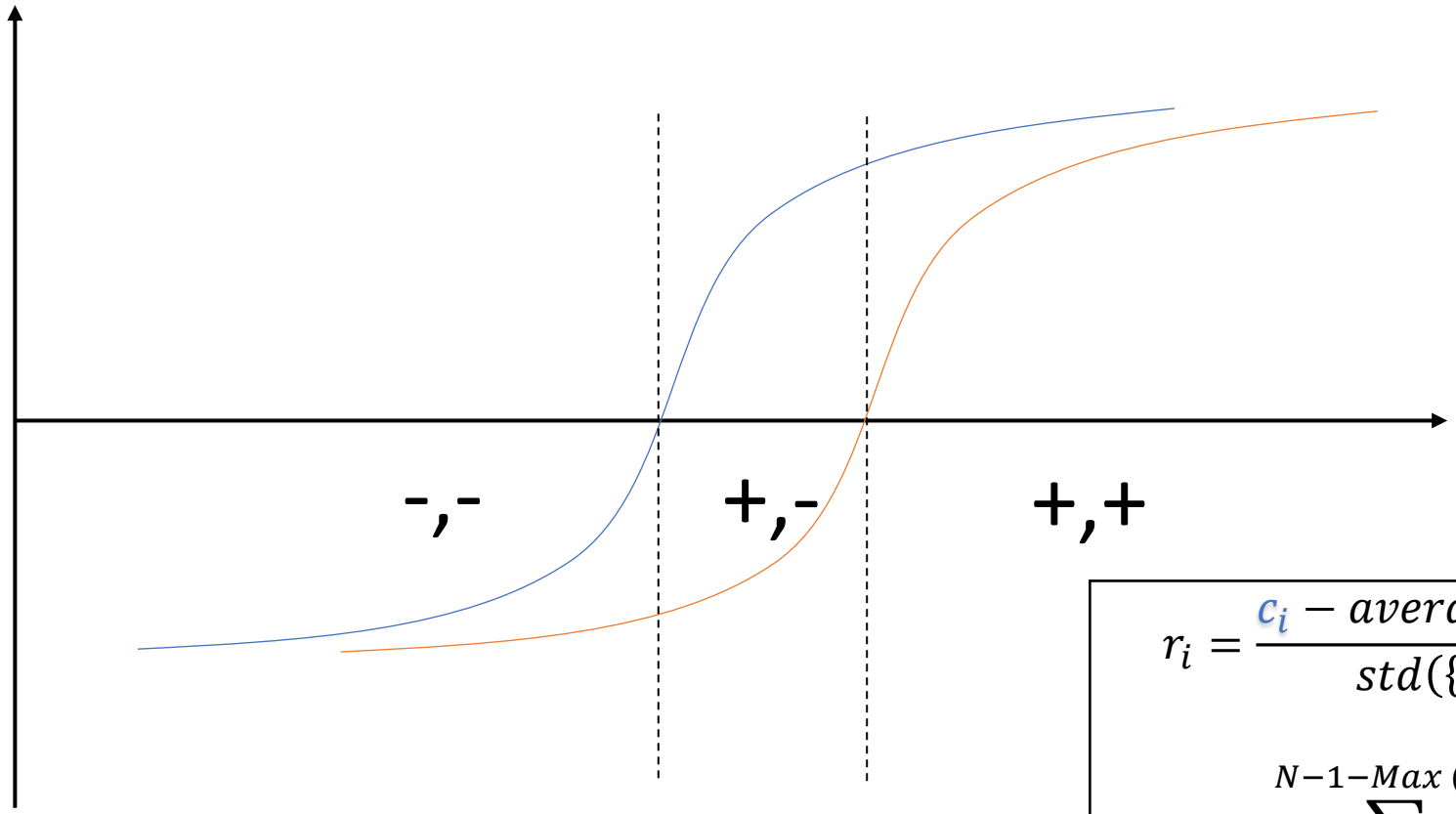




$$r_i = \frac{c_i - average(\{c_i\})}{std(\{c_i\})}, i = M, J$$

$$cf[\tau] = \sum_{m=0}^{N-1-Max\,(|\tau|)} r_M[m] \cdot r_J[m+\tau]$$

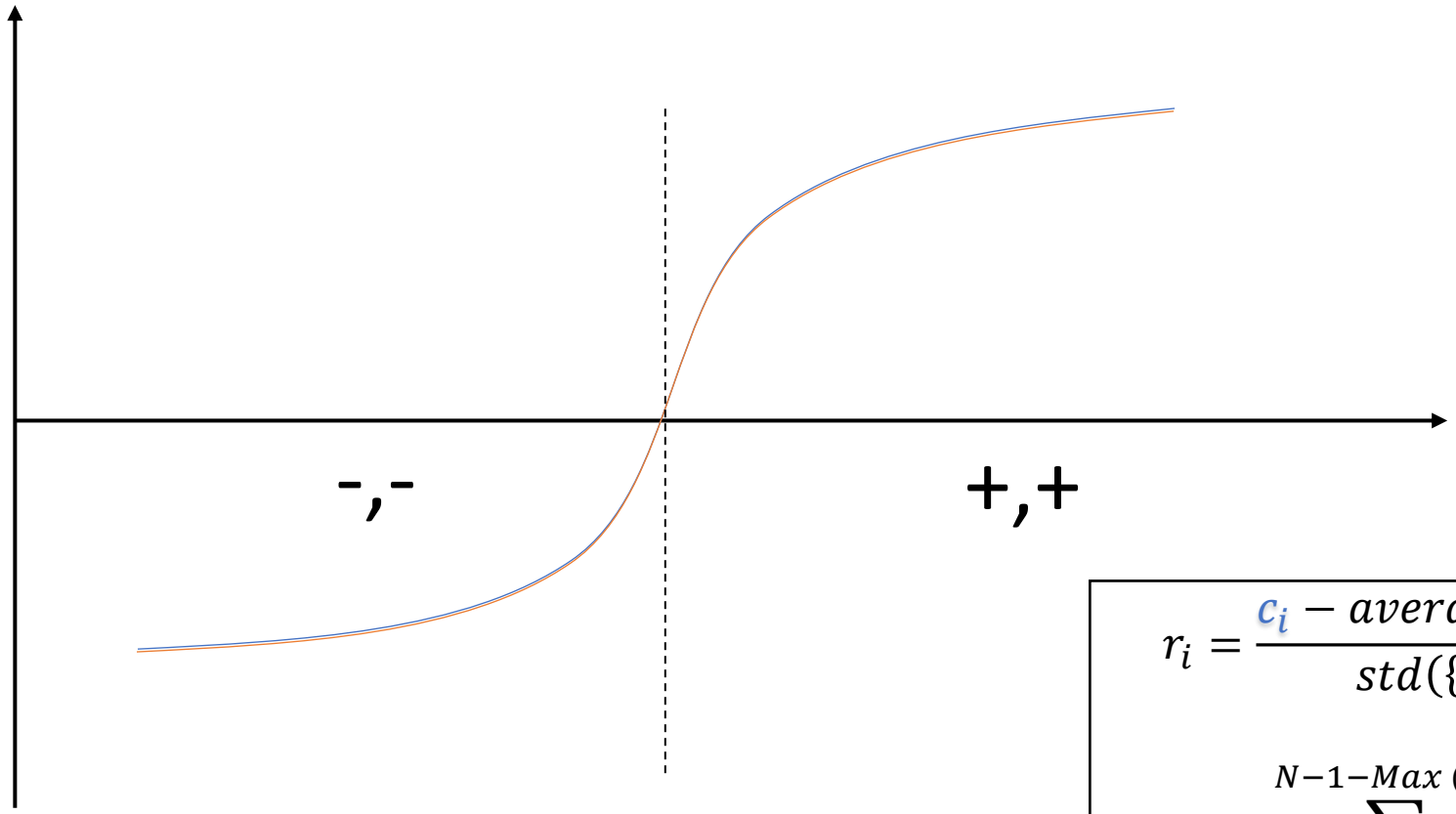
Find τ that maximizes $cf[\tau]$



$$r_i = \frac{c_i - average(\{c_i\})}{std(\{c_i\})}, i = M, J$$

$$cf[\tau] = \sum_{m=0}^{N-1-Max\,(|\tau|)} r_M[m] \cdot r_J[m+\tau]$$

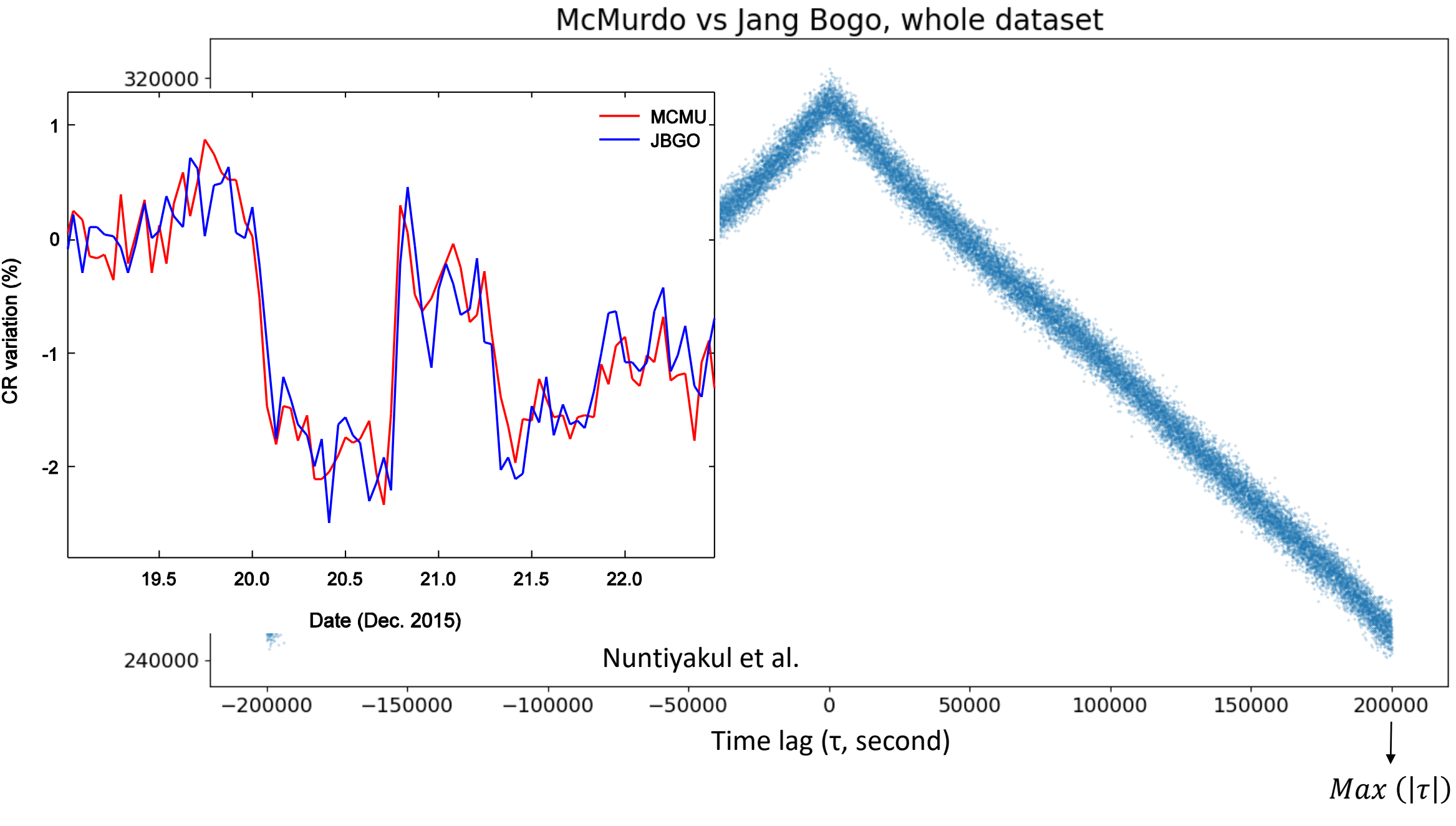
Find τ that maximizes $cf[\tau]$



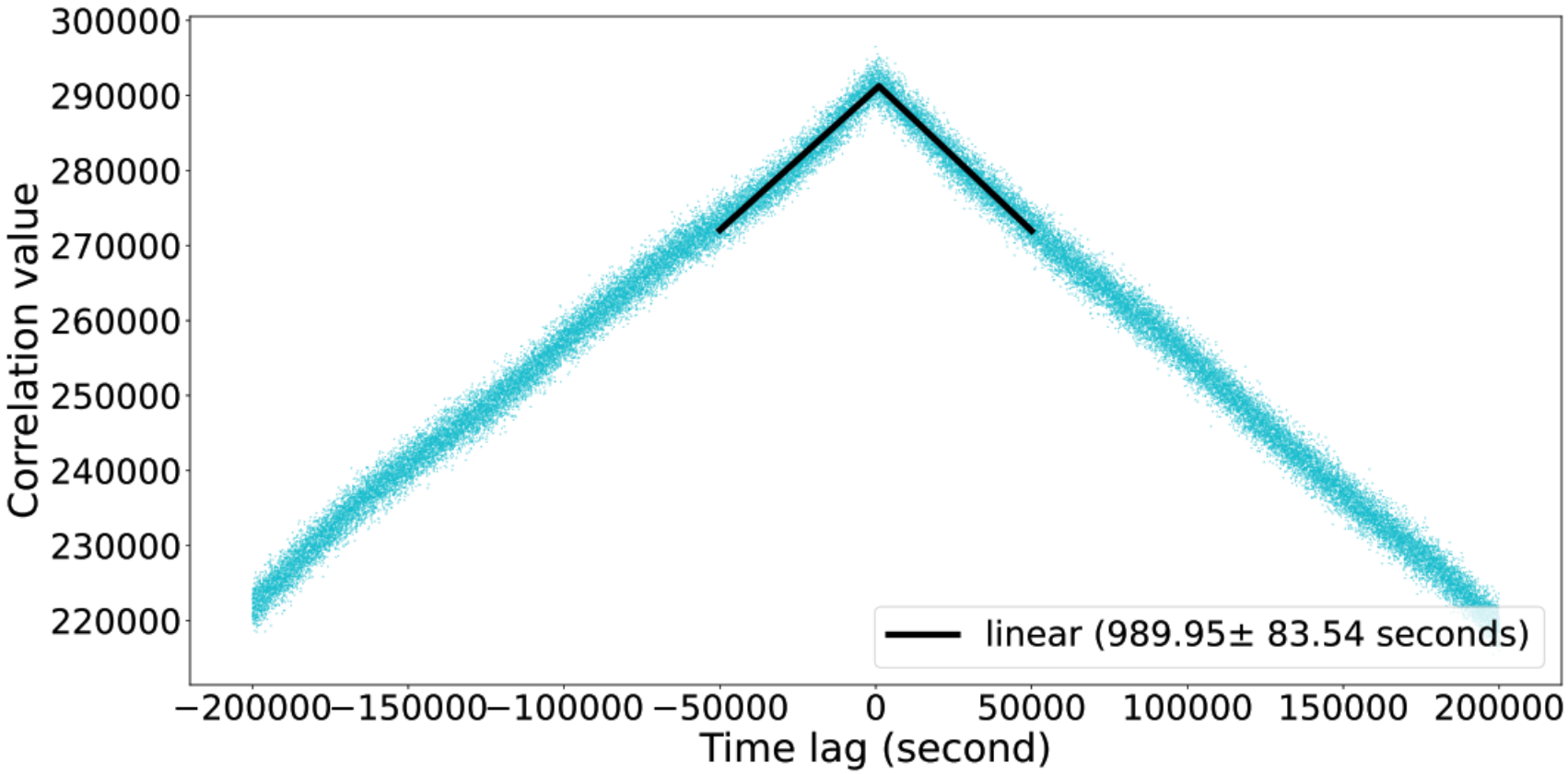
$$r_i = \frac{c_i - average(\{c_i\})}{std(\{c_i\})}, i = M, J$$

$$cf[\tau] = \sum_{m=0}^{N-1-Max\,(|\tau|)} r_M[m] \cdot r_J[m+\tau]$$

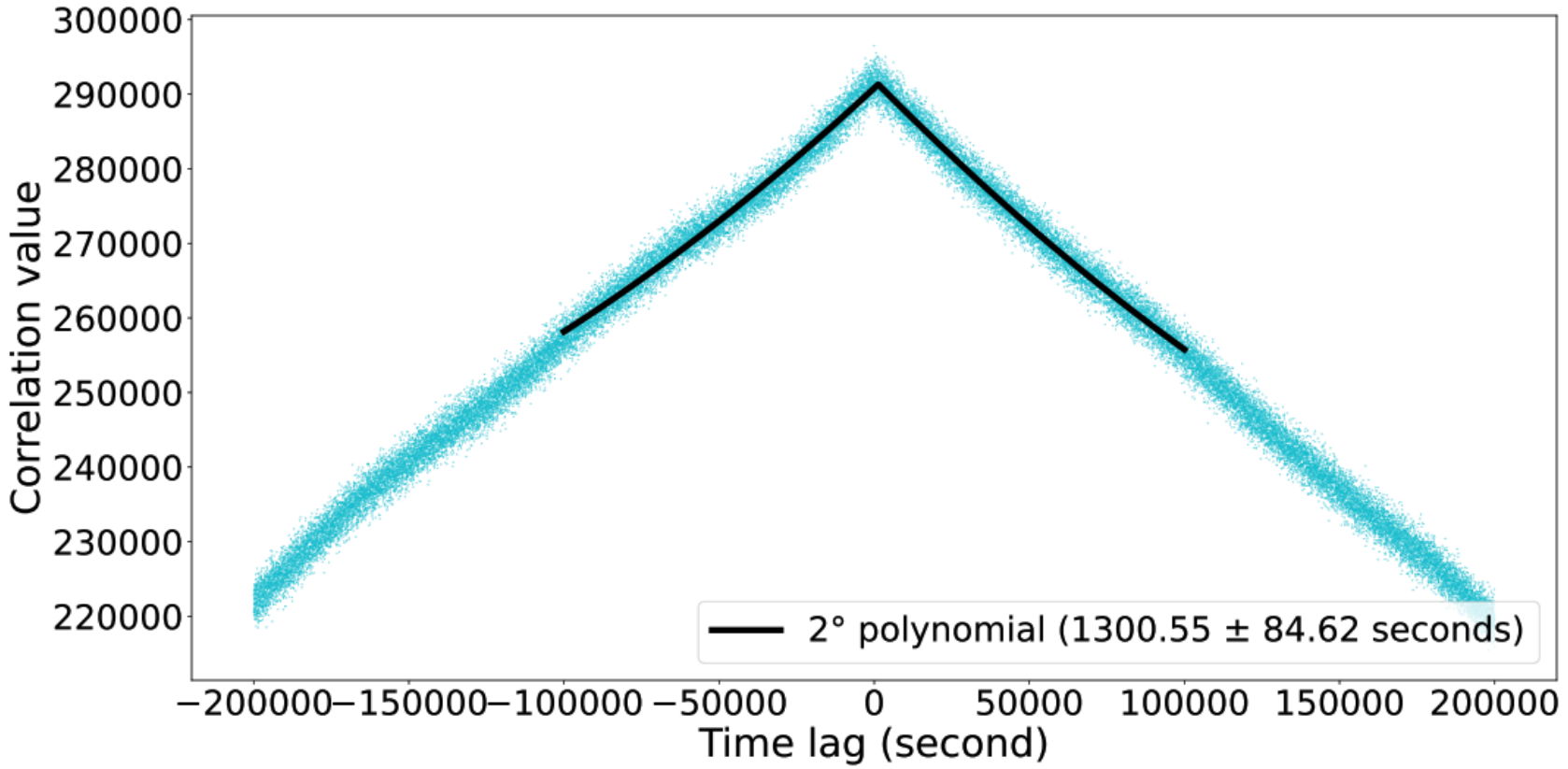
Find τ that maximizes $cf[\tau]$


McMurdo vs Jang Bogo, whole dataset
MCMU
JBGO
CR variation (%)
Date (Dec. 2015)
Nuntiyakul et al.
Time lag (τ, second)
Max (|τ|)

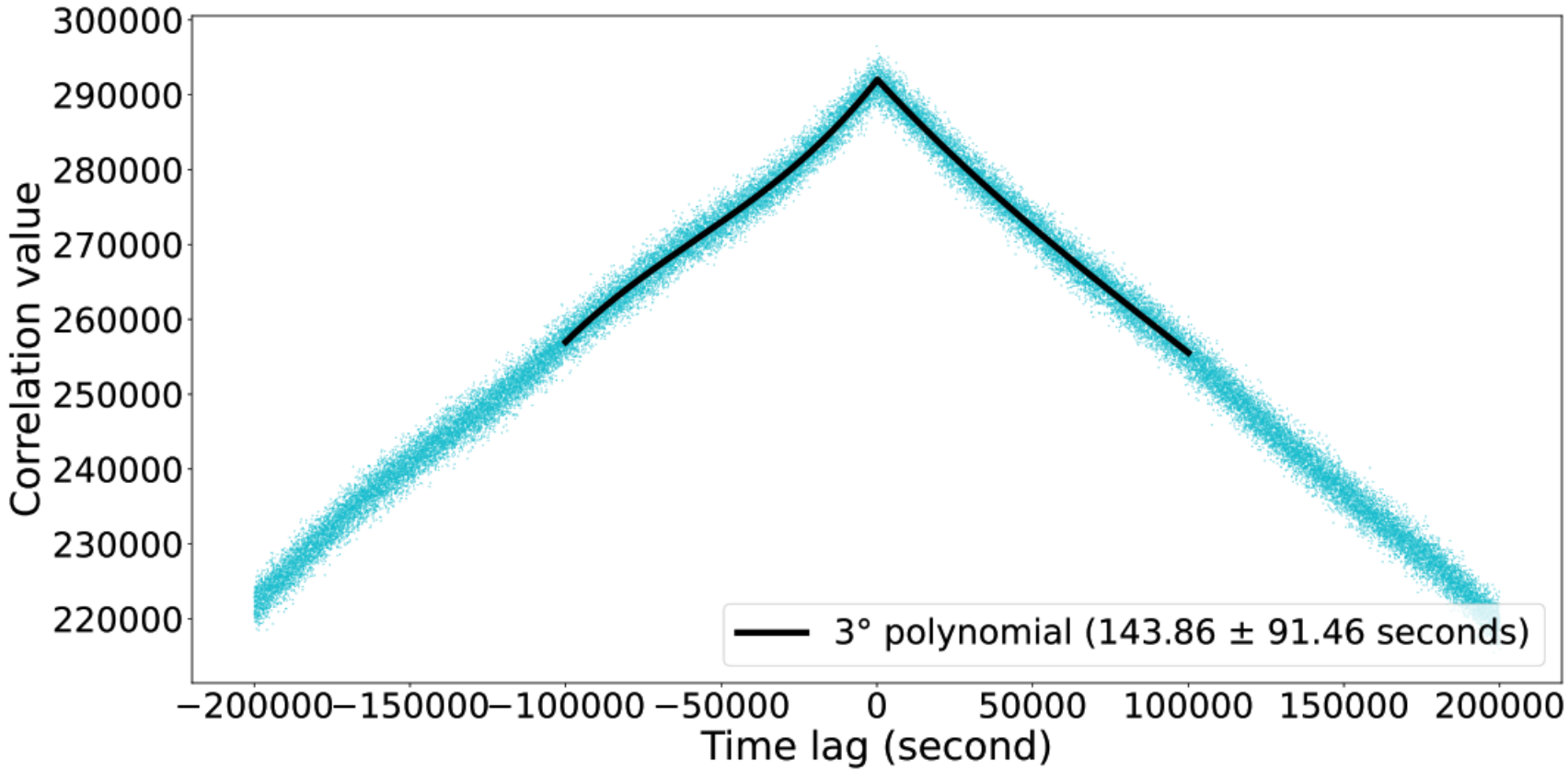
Linear, whole station



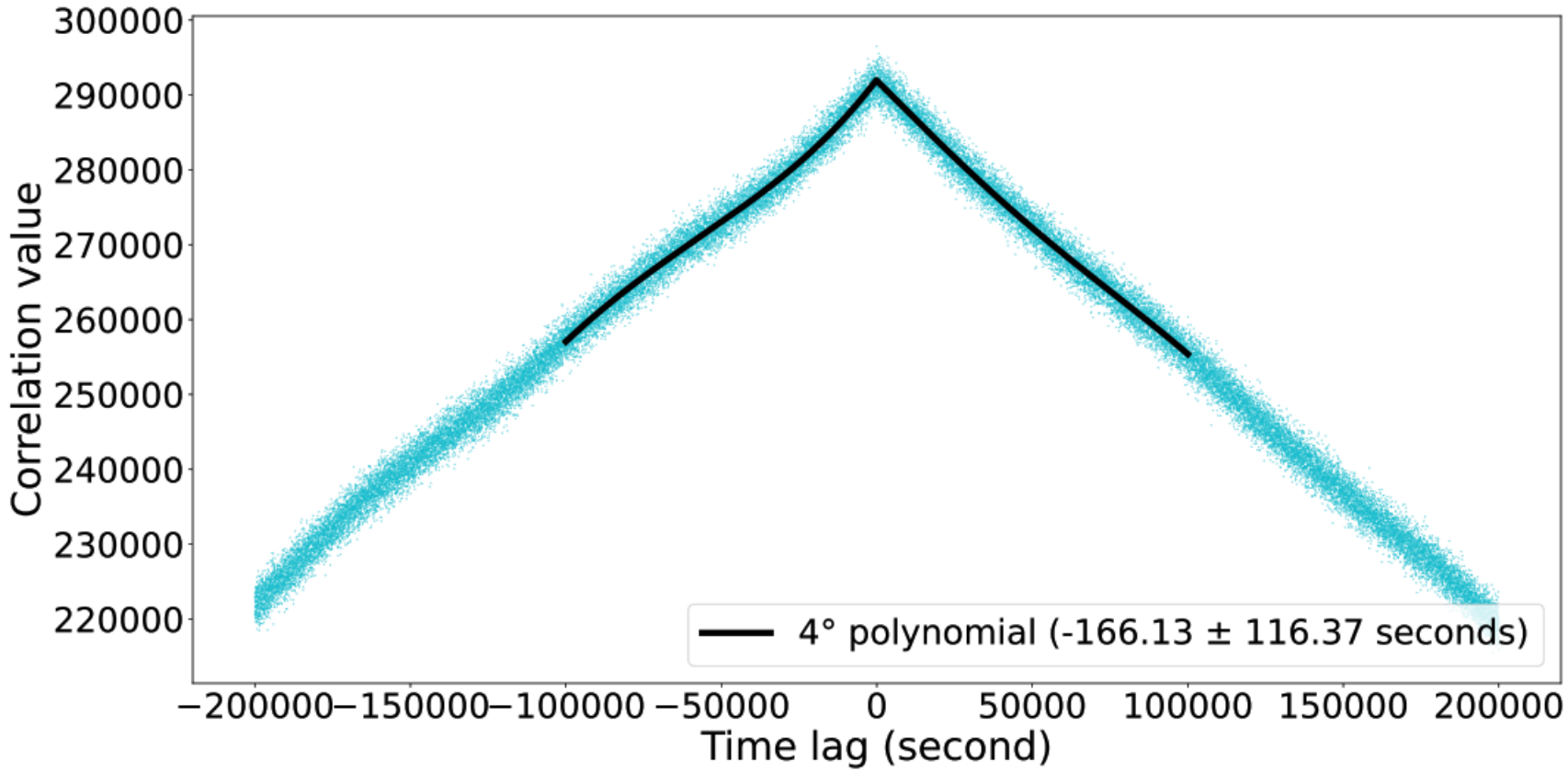
2° polynomial, whole station



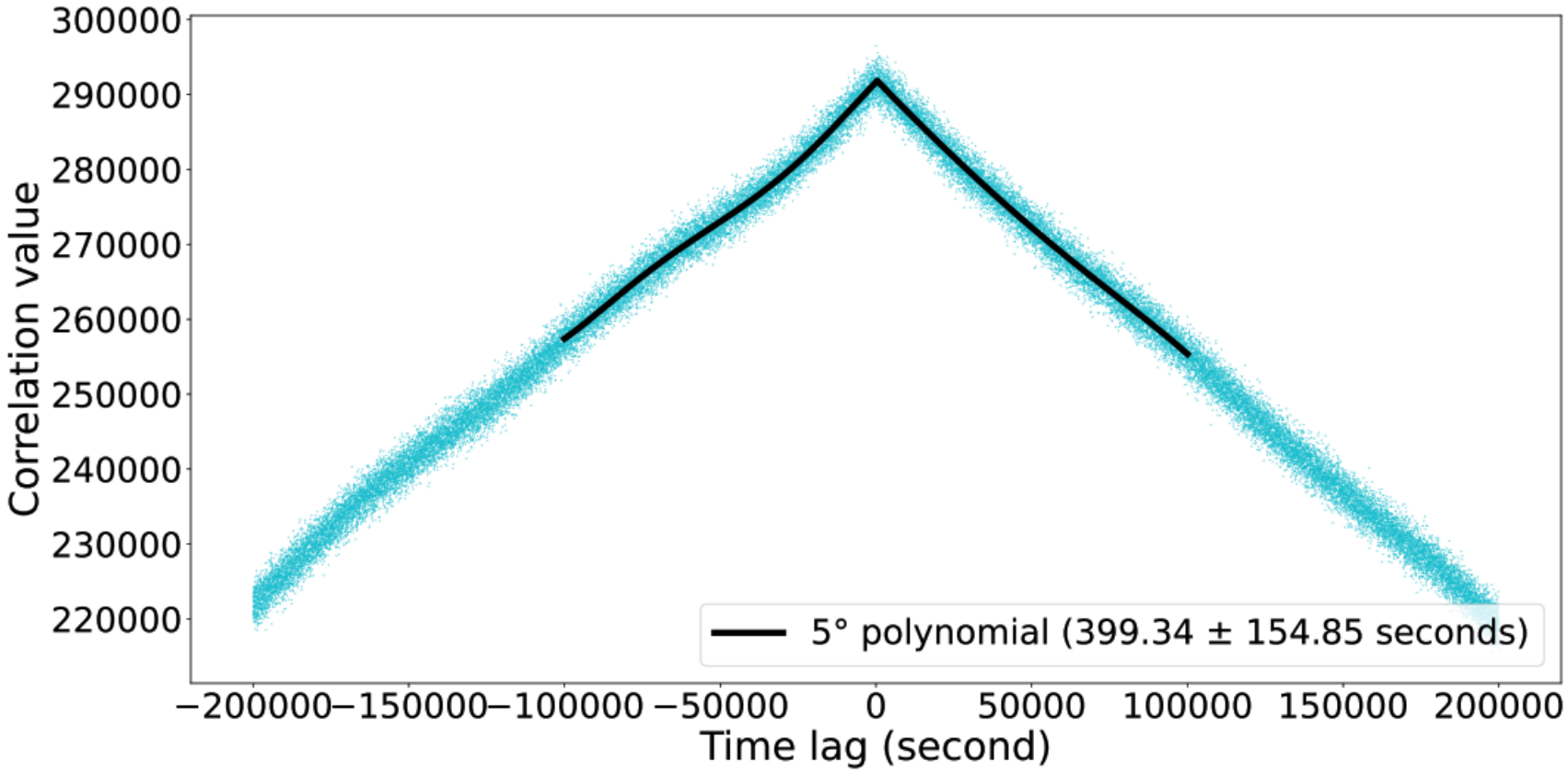
3º polynomial, whole station



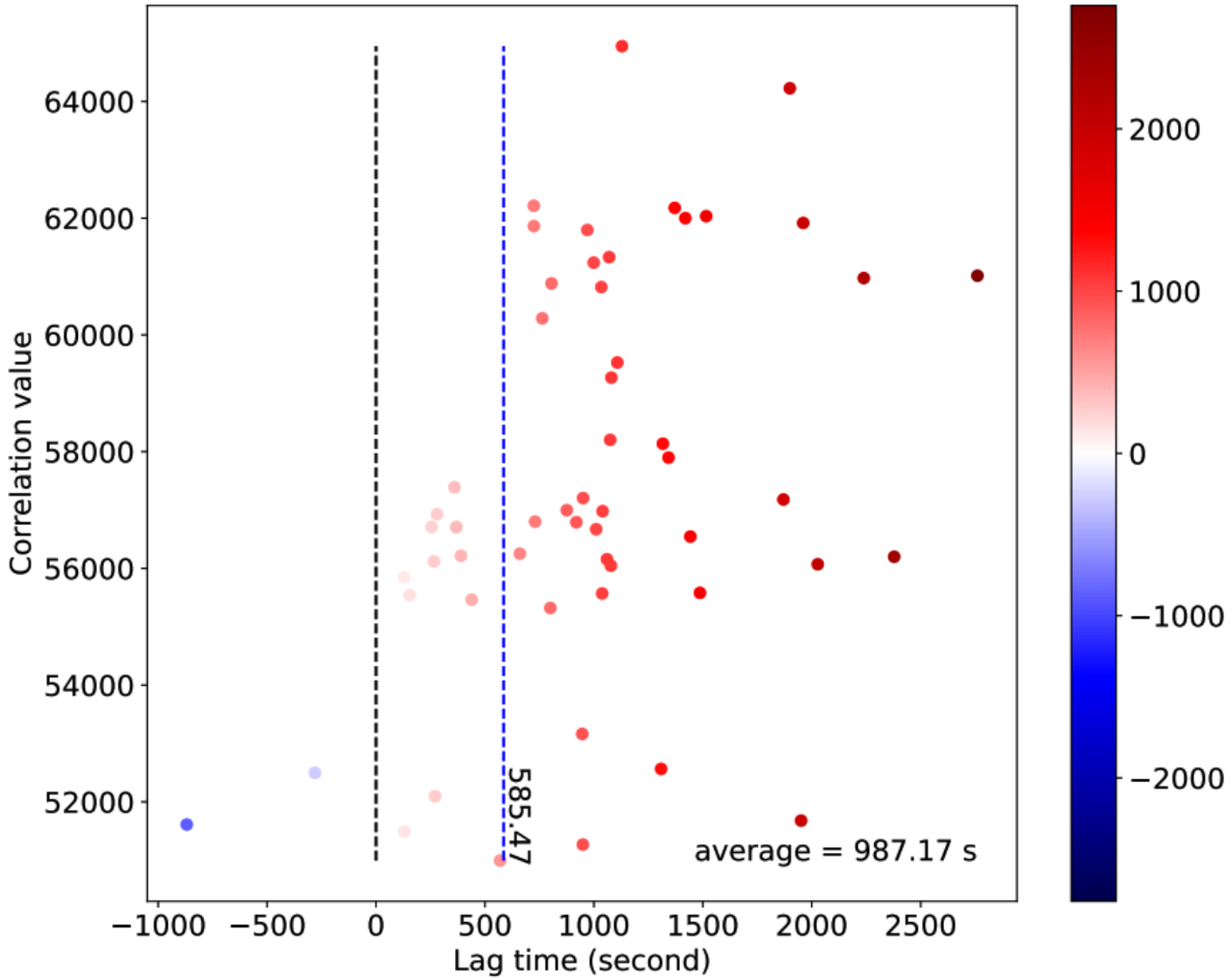
4º polynomial, whole station



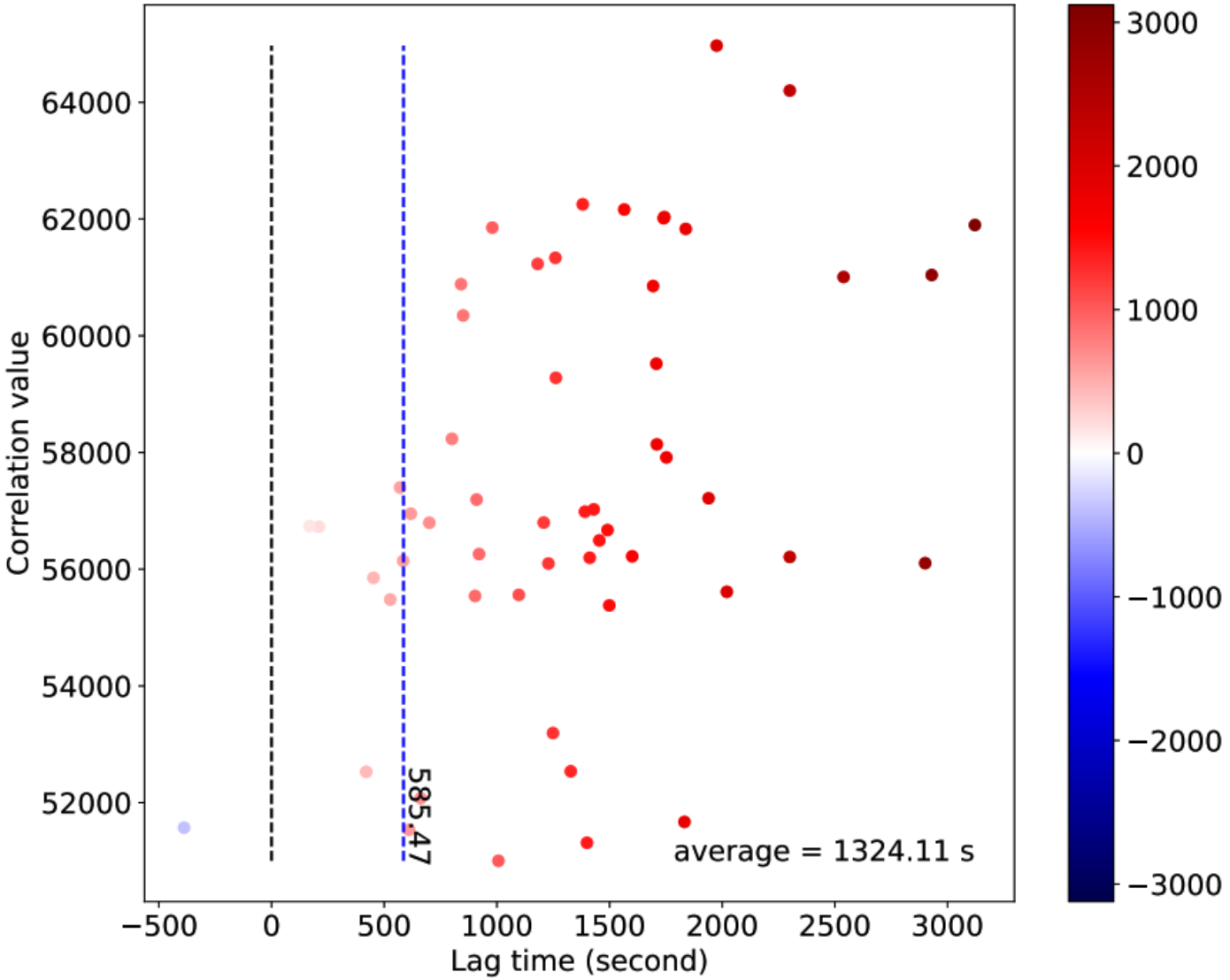
5° polynomial, whole station



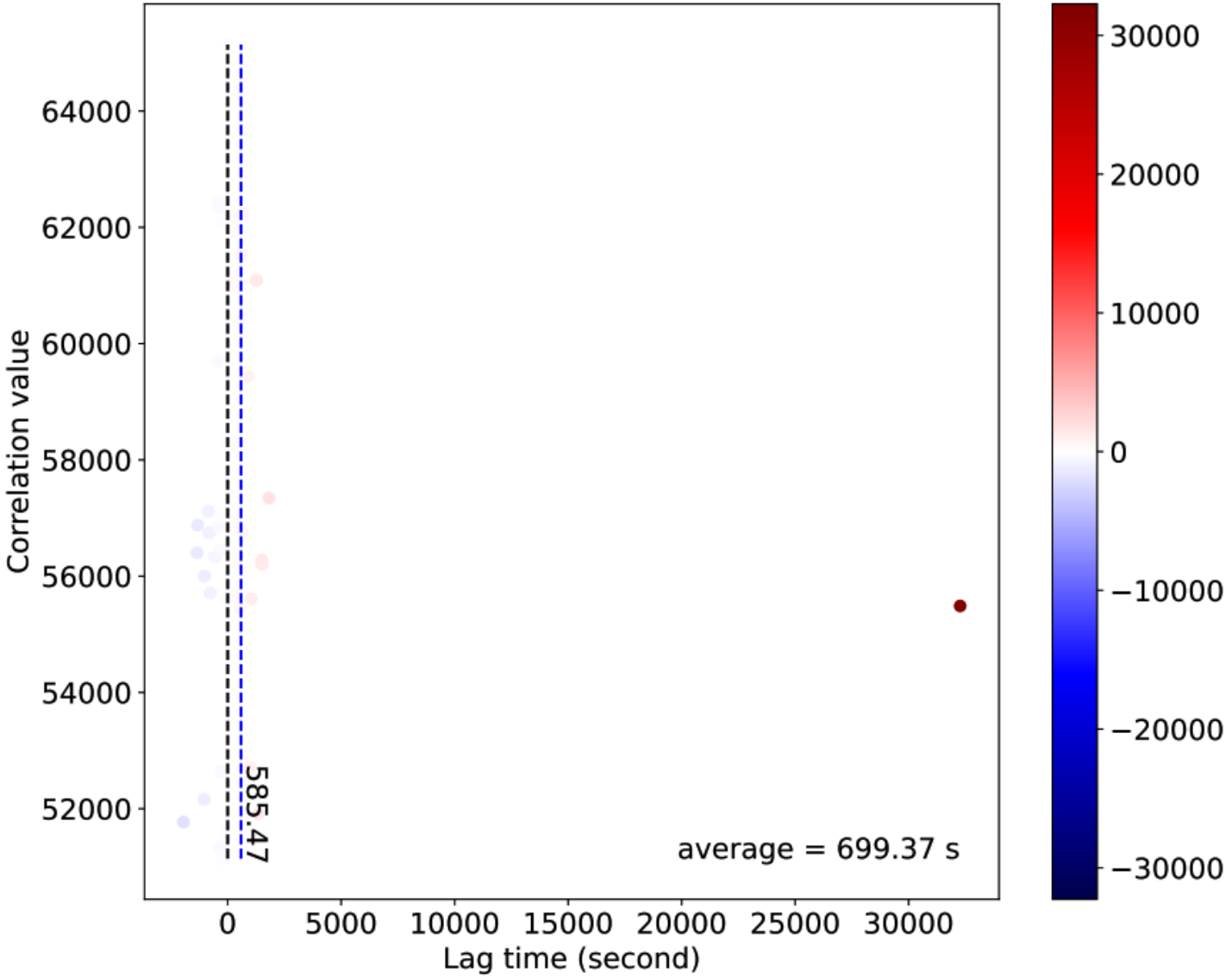
Linear, tube by tube, peak only



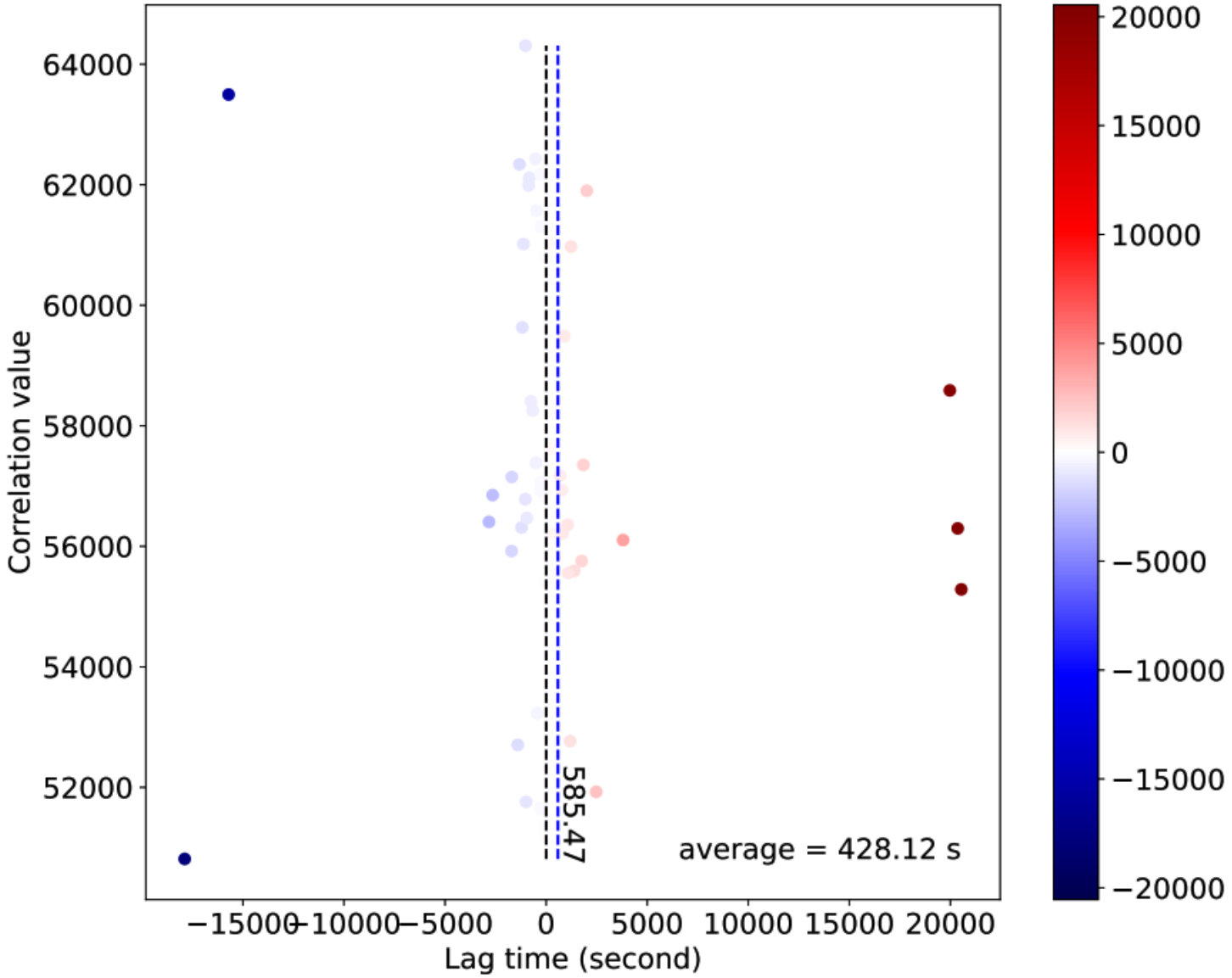
2° polynomial, tube by tube, peak only



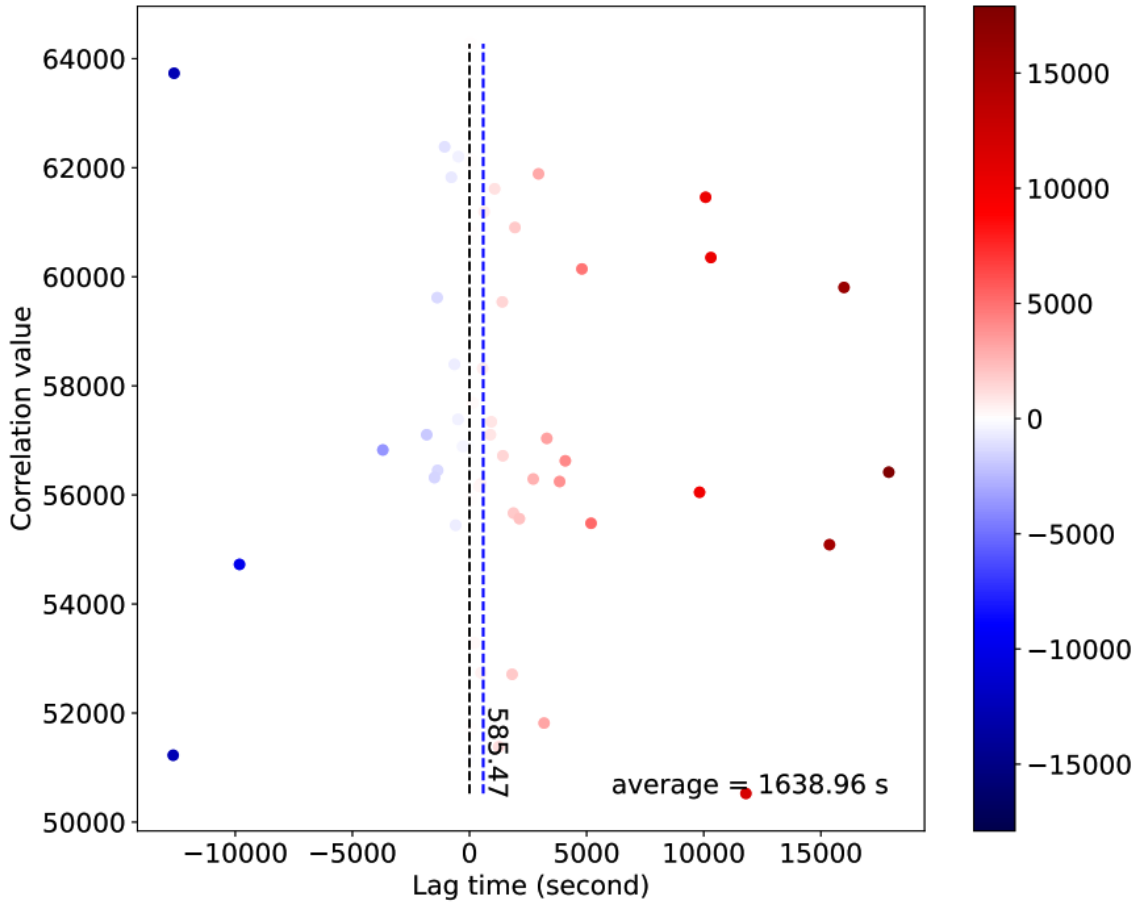
3º polynomial, tube by tube, peak only



4º polynomial, tube by tube, peak only



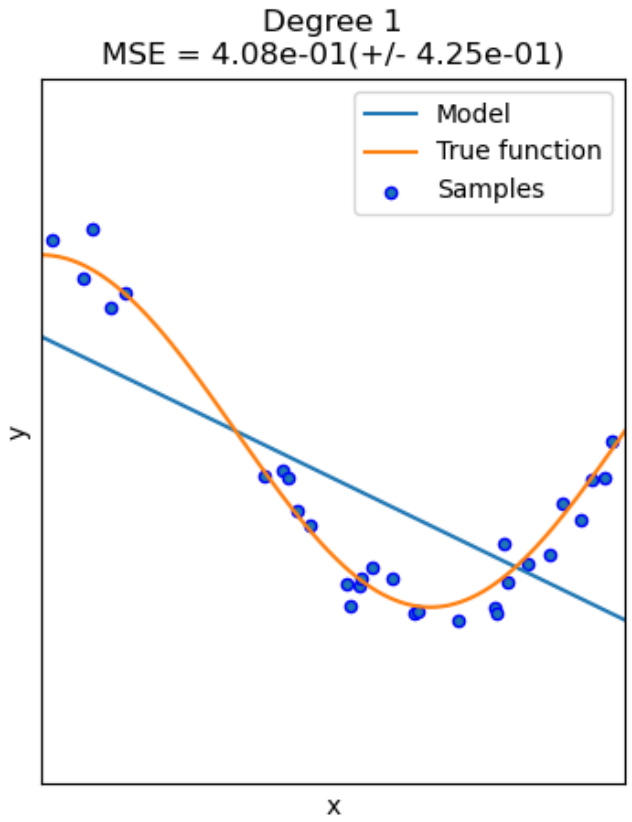
5° polynomial, tube by tube, peak only



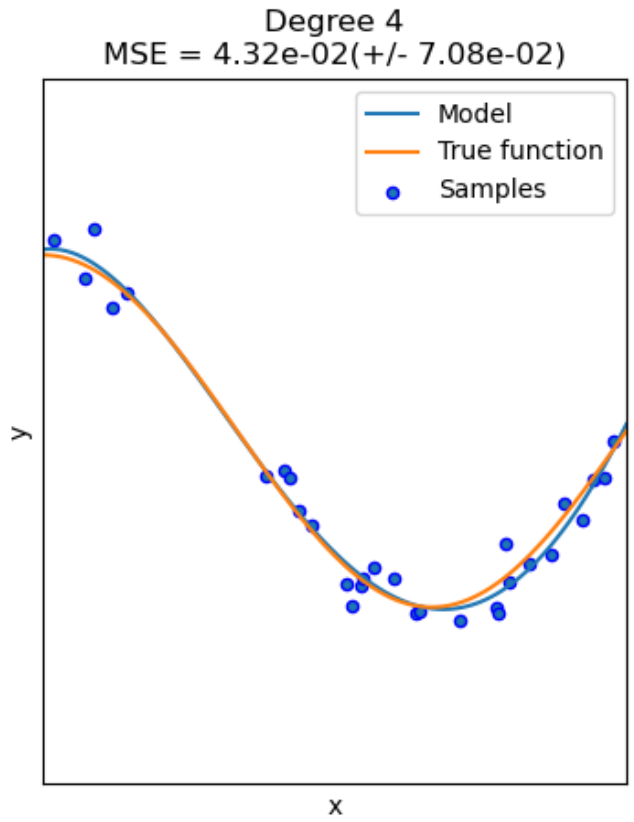
Which one?

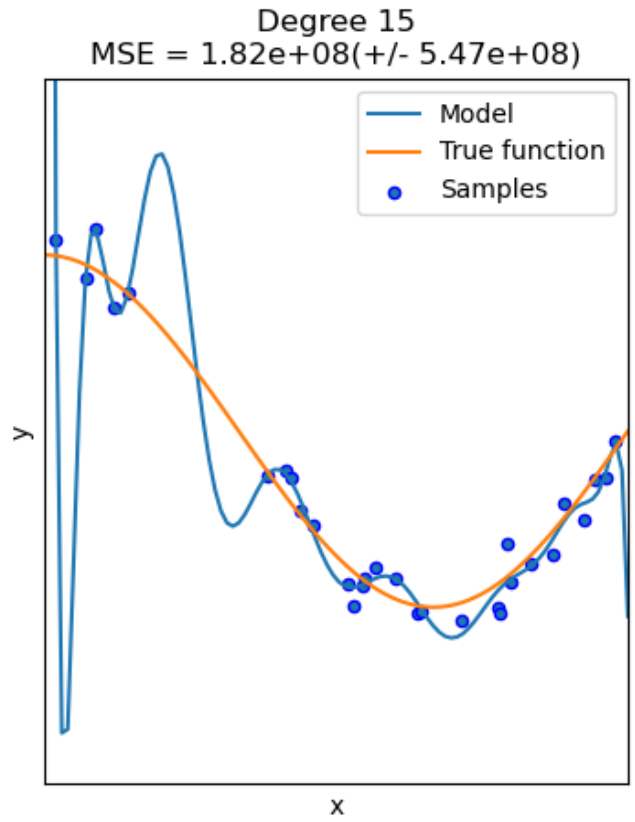
- Introduce variation while keeping the middle value
  - By randomly removing data
- Inspired by k-fold cross-validation (k=5 in this work)

Underfitting and overfitting



underfitting

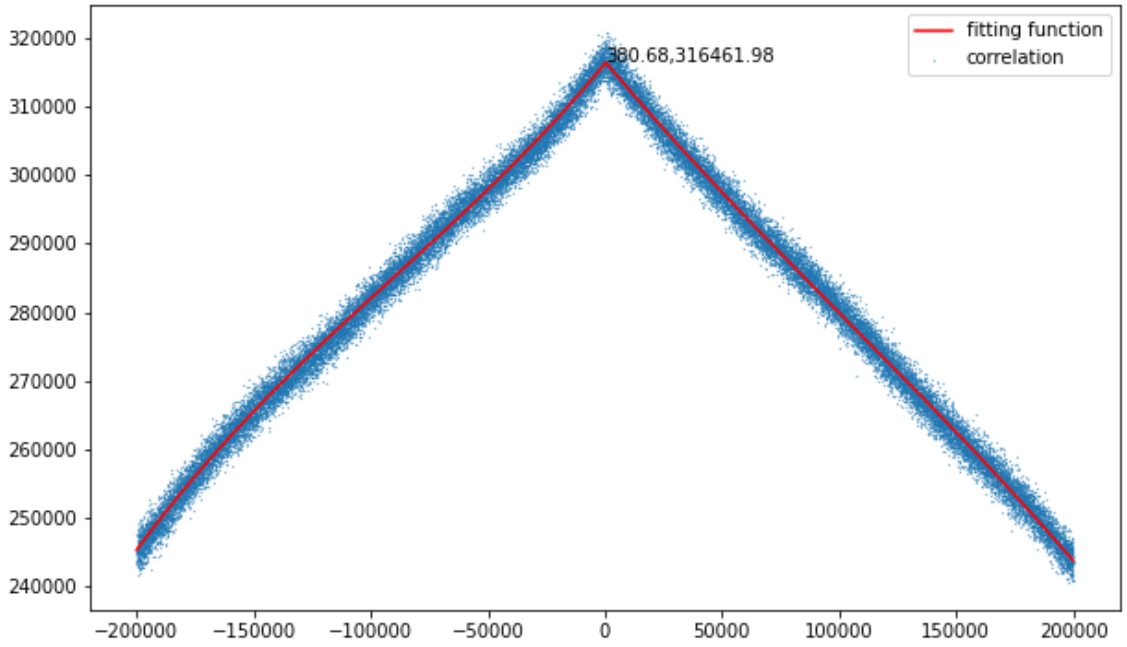




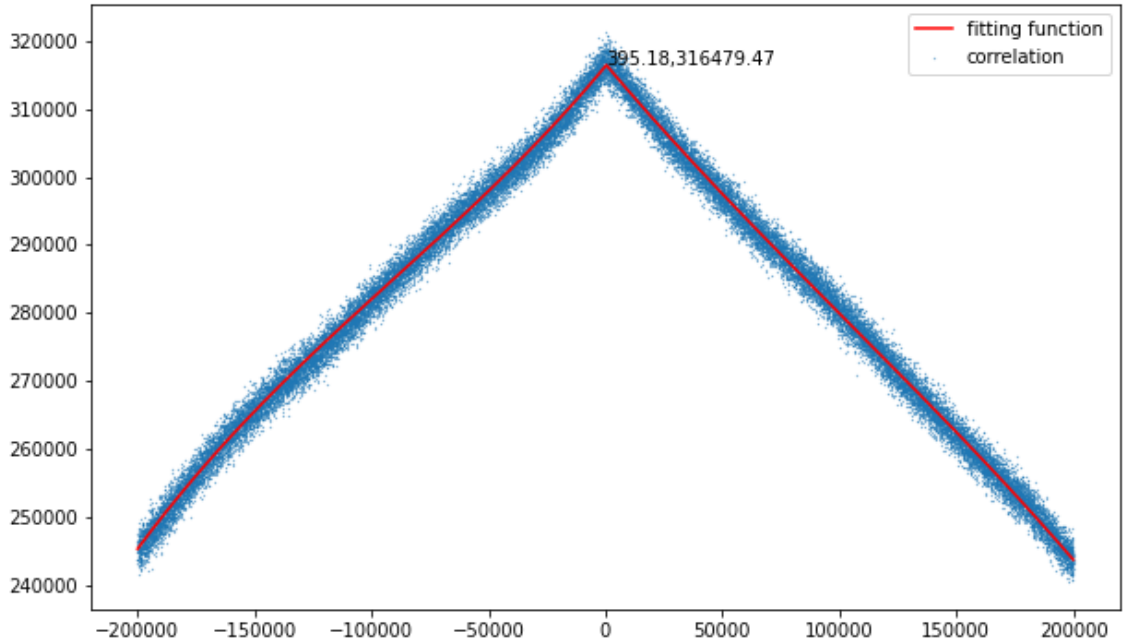
overfitting

5-fold cross-validation

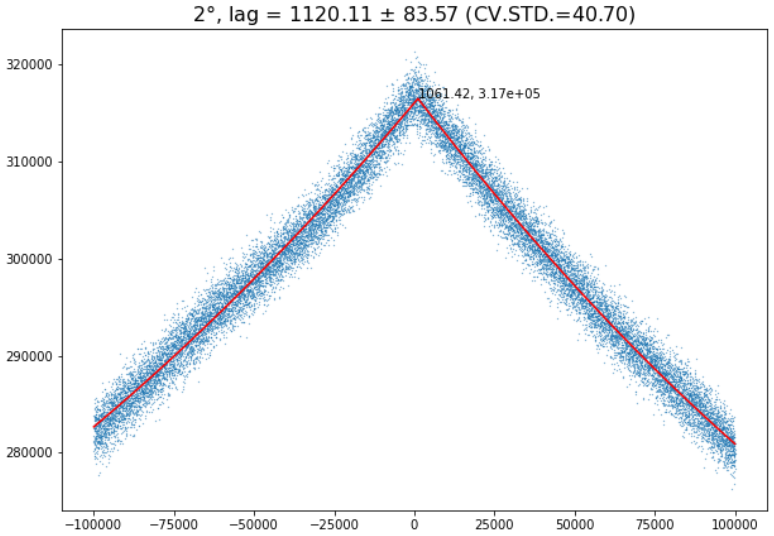
1/5 removed

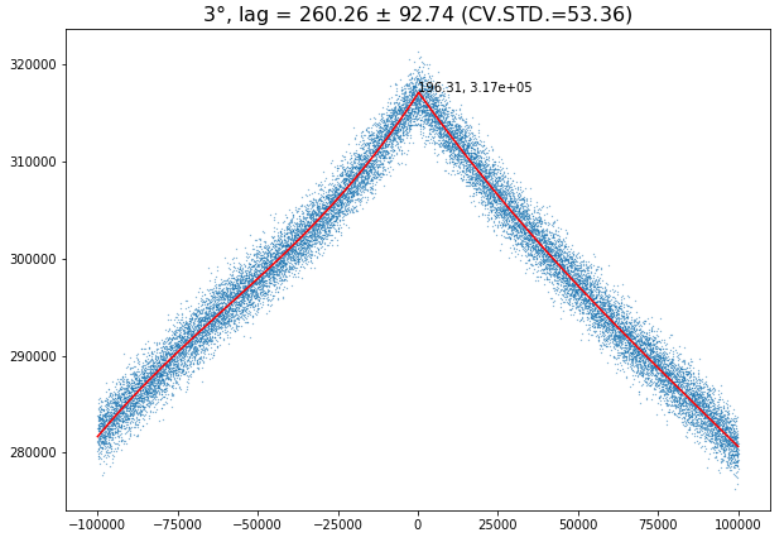


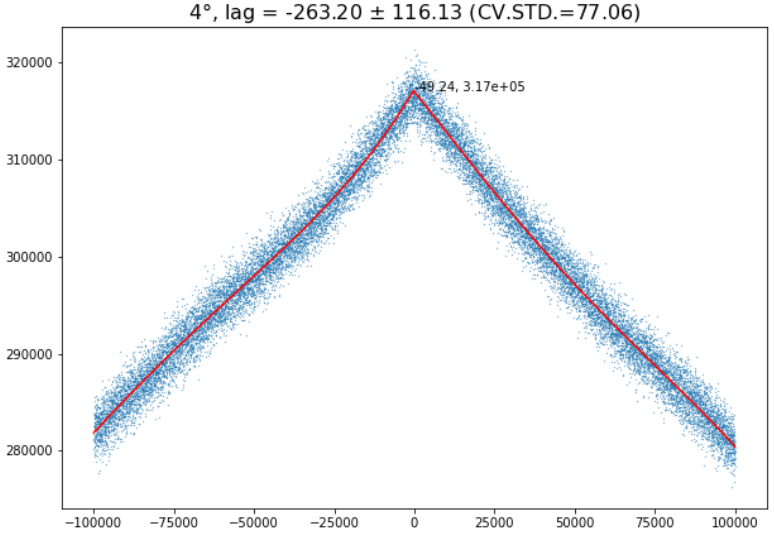
other 1/5 removed

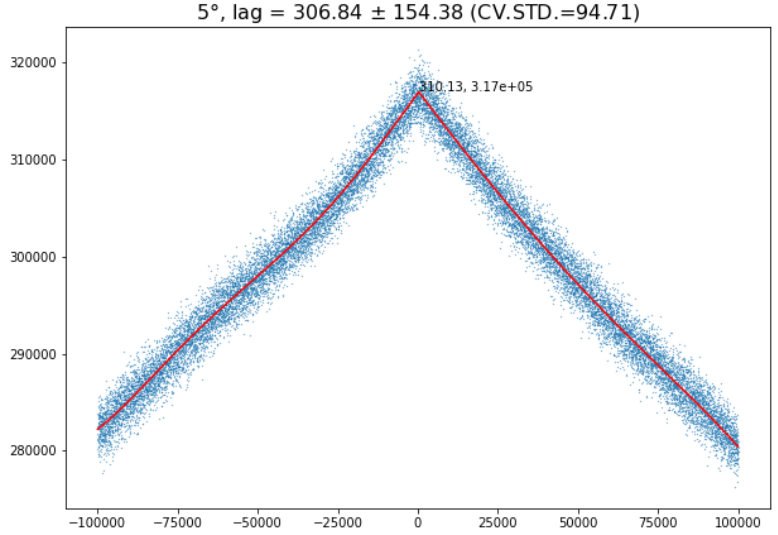


5-fold cross-validation, from -100k to 100k seconds, 1000 iterations





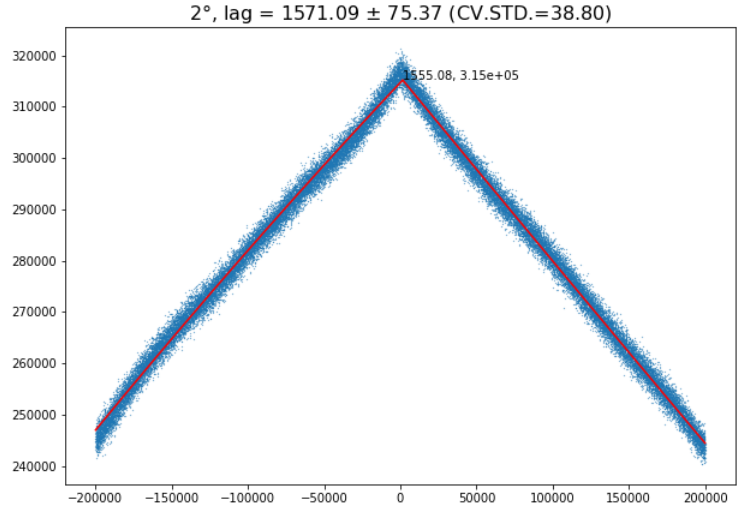


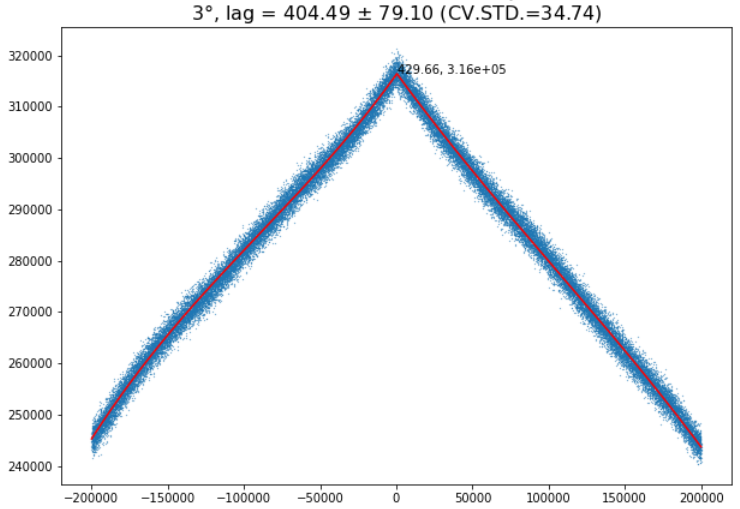


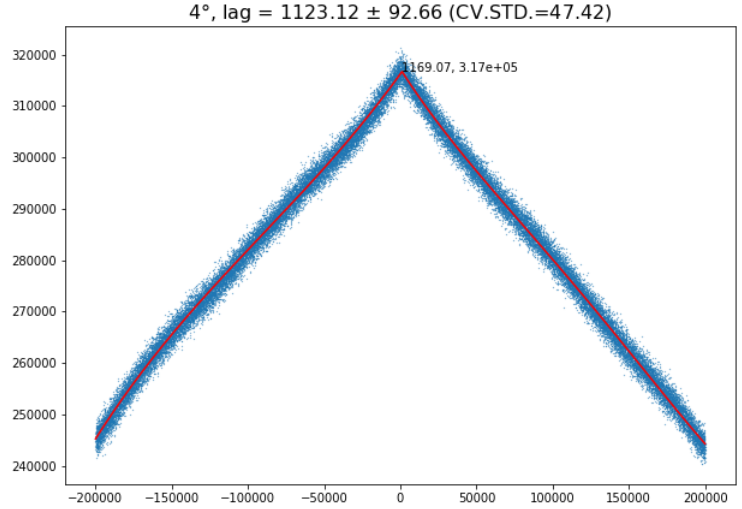
CV.STD = cross-validation STD calculated from 5*iteration values of τ at peak

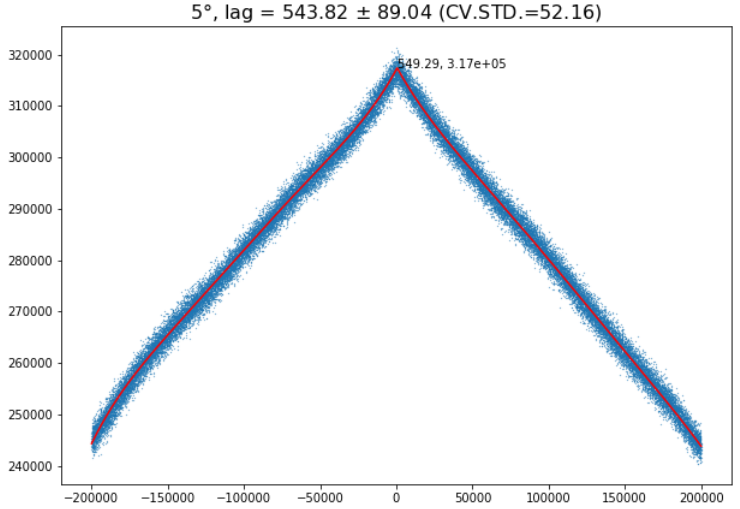
iteration = 1000

5-fold cross-validation, 1000 iterations from -200k to 200k seconds, 1000 iterations





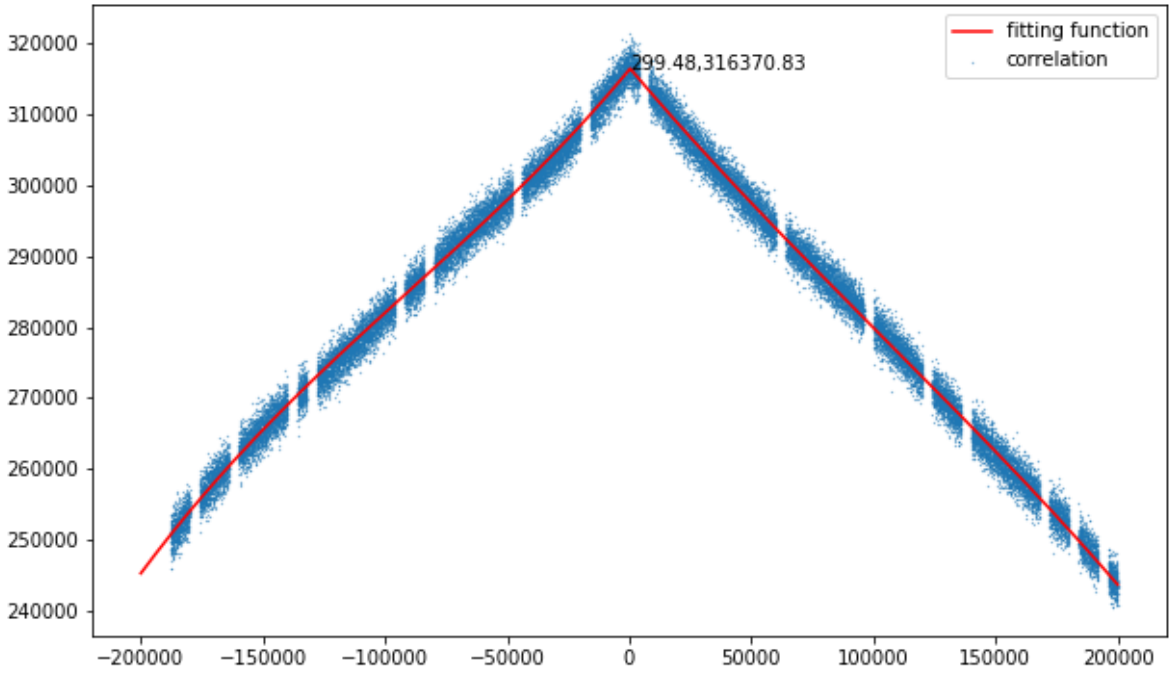




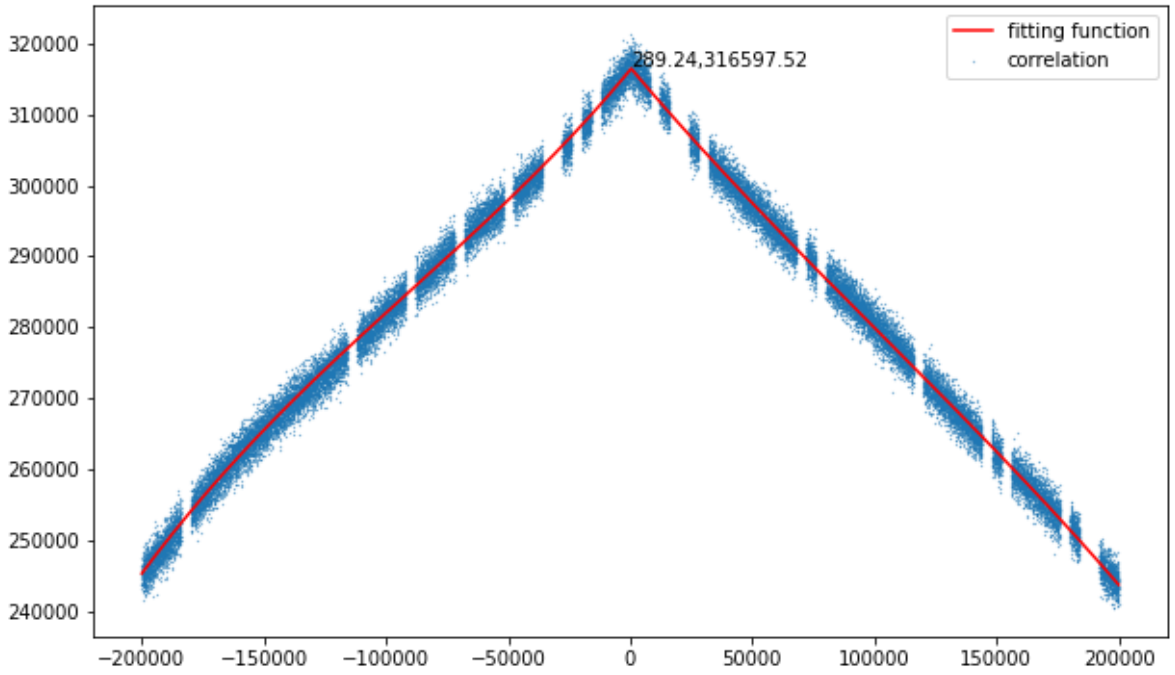
CV.STD = cross-validation STD

5-fold cross validation, 100 chunks

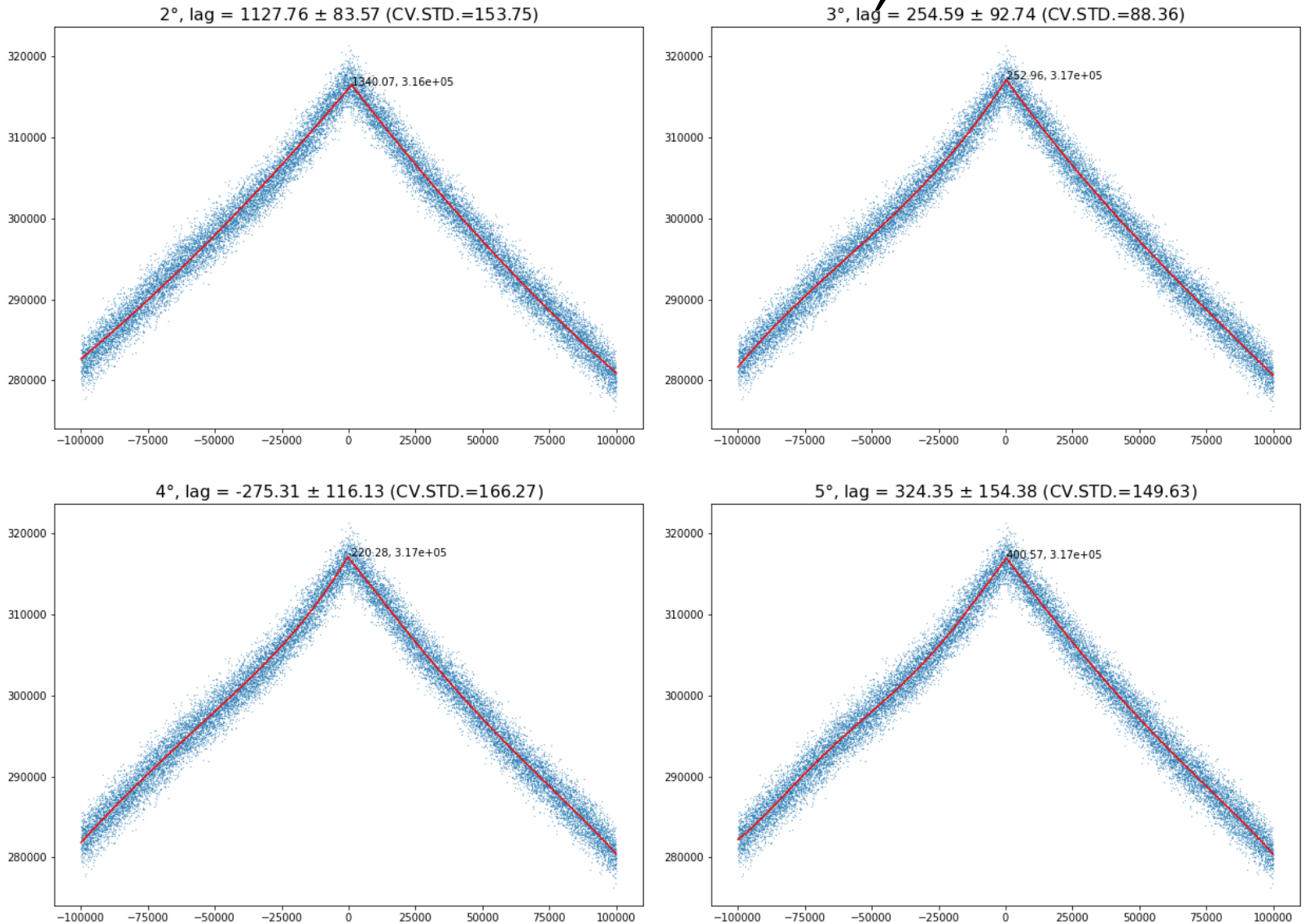
1/5 of chunks removed



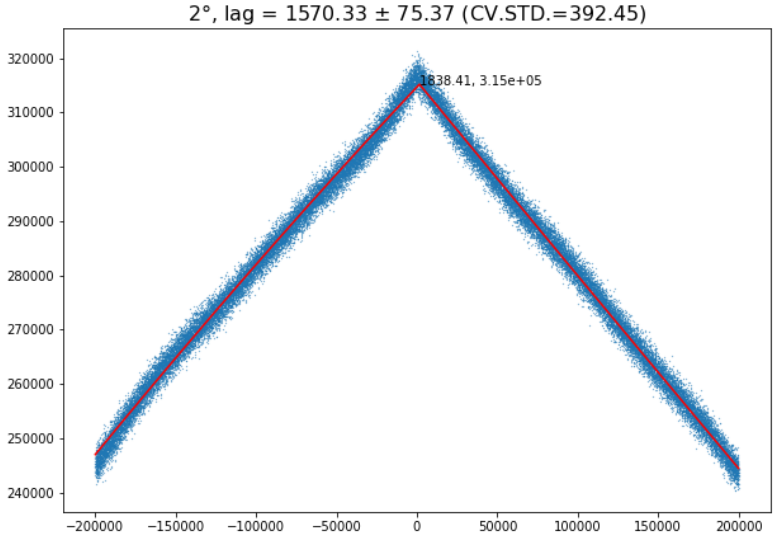
other 1/5 of chunks removed

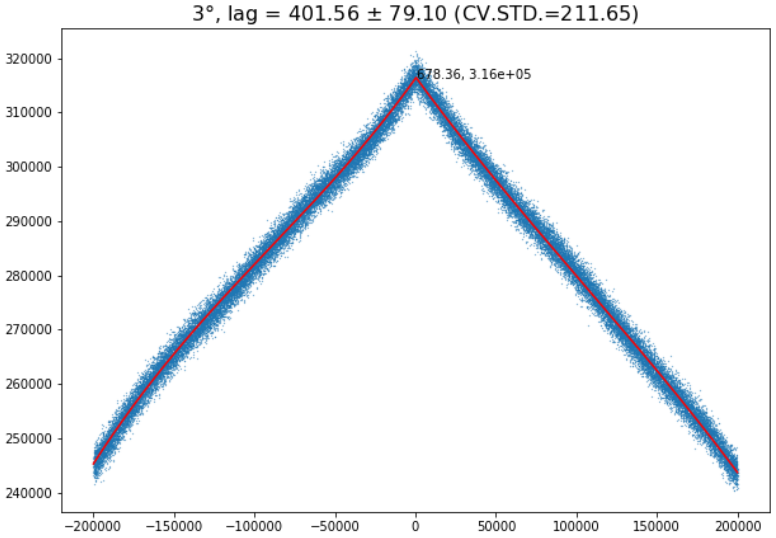


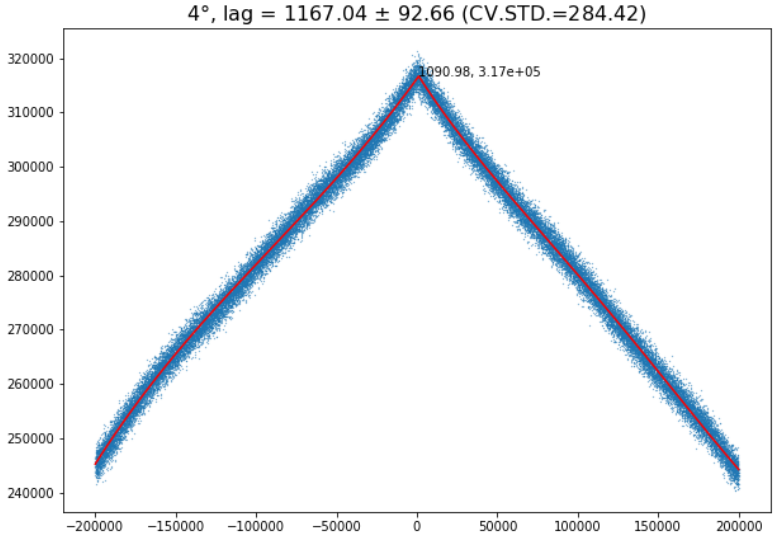
5-fold cross 5-fold cross validation, 100 chunks from -100k to 100k seconds, 1000 iterations

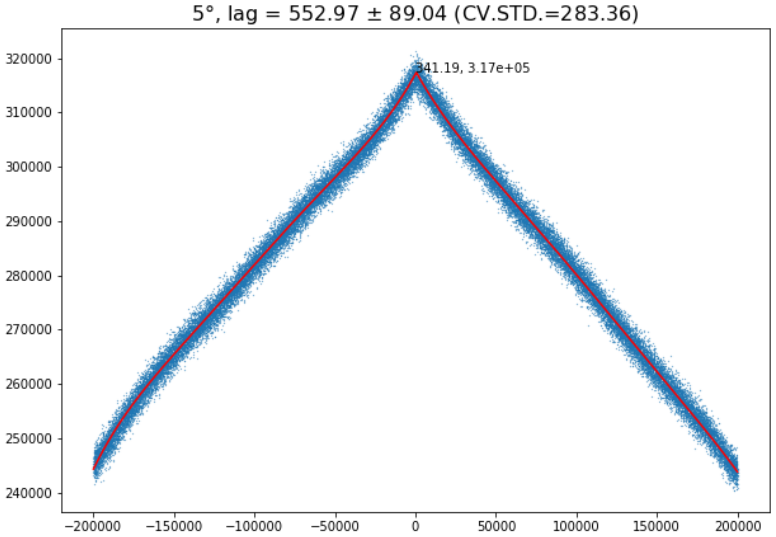


5-fold cross validation, 100 chunks from -200k to 200k seconds, 1000 iterations









Works on progress

- Segment correlation i.e., correlation of data at MCMU at 03:00-06:00
- temporal correlation i.e., $corr(n, \tau) = \sum_i r_M[i+n] \cdot r_J[i+n+\tau]$
- Fewer sample, recalculate average and SD for every τ while taking missing data into account
  - for each τ if $r_M[m]$ or $r_J[m+\tau]$ is missing, both are treated as missing
  - Performance problem

$$r_i = \frac{c_i - average(\{c_i\})}{std(\{c_i\})}, i = M, J$$

$$cf[\tau] = \sum_{m=0}^{N-1-Max\,(|\tau|)} r_M[m] \cdot r_J[m+\tau]$$

Find τ that maximizes $cf[\tau]$

Thank you