1) flow data

Flow data pre- and post-LYX, LJX dams are measured at a roughly 10-day basis (旬Xun in Chinese, one month = 3 Xuns -> one Xun ~= 10 days) and are collected from June 2000 to June 2010 (10 water years over June to June). As the simulation period is set to a calendar year Jan-Dec, data of the first half year and the last half year are dropped for consistency. The 10-day flow data over Jan 2001 to Dec 2009 are then used to sample synthetic streamflow data. Time series of flow data are shown in Figure 1.

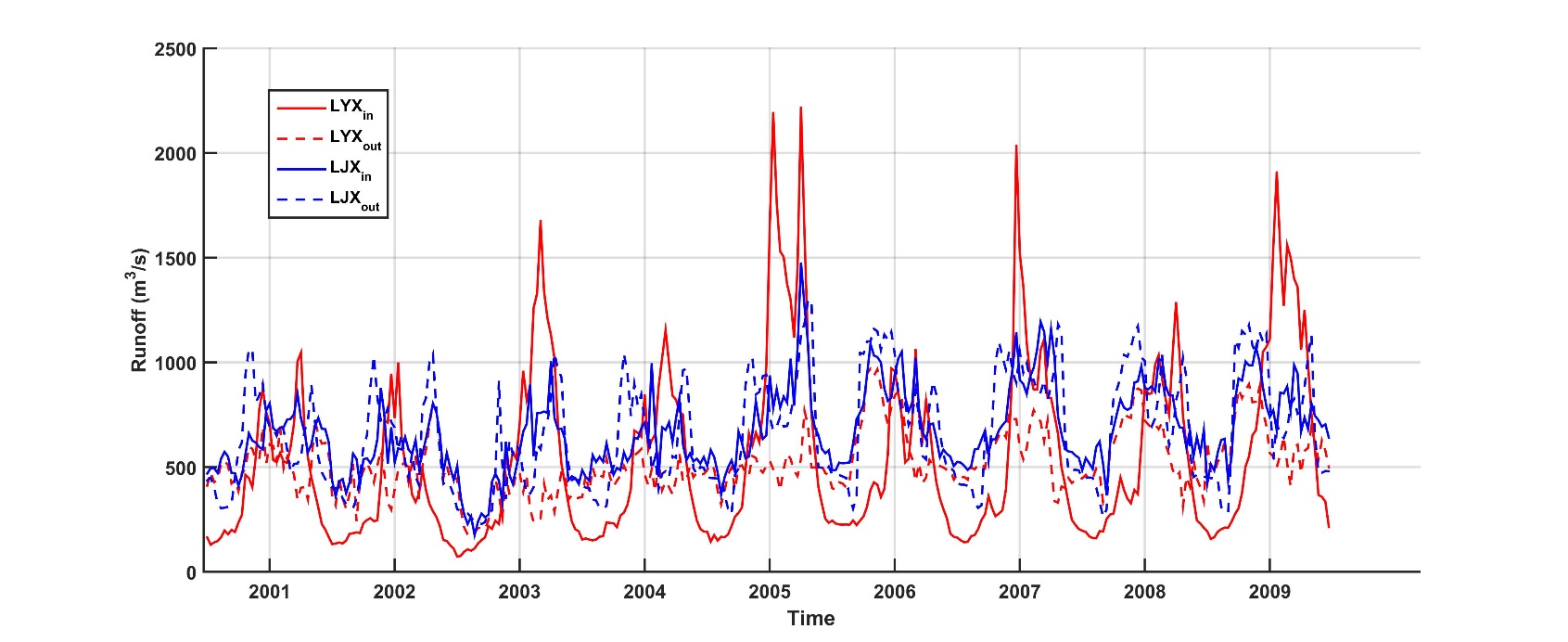


Figure 1. Flow over 2001-2019 for pre-LYX (red solid), post-LYX (red dashed), pre-LJX (blue solid), and post-LJX (blue dashed)

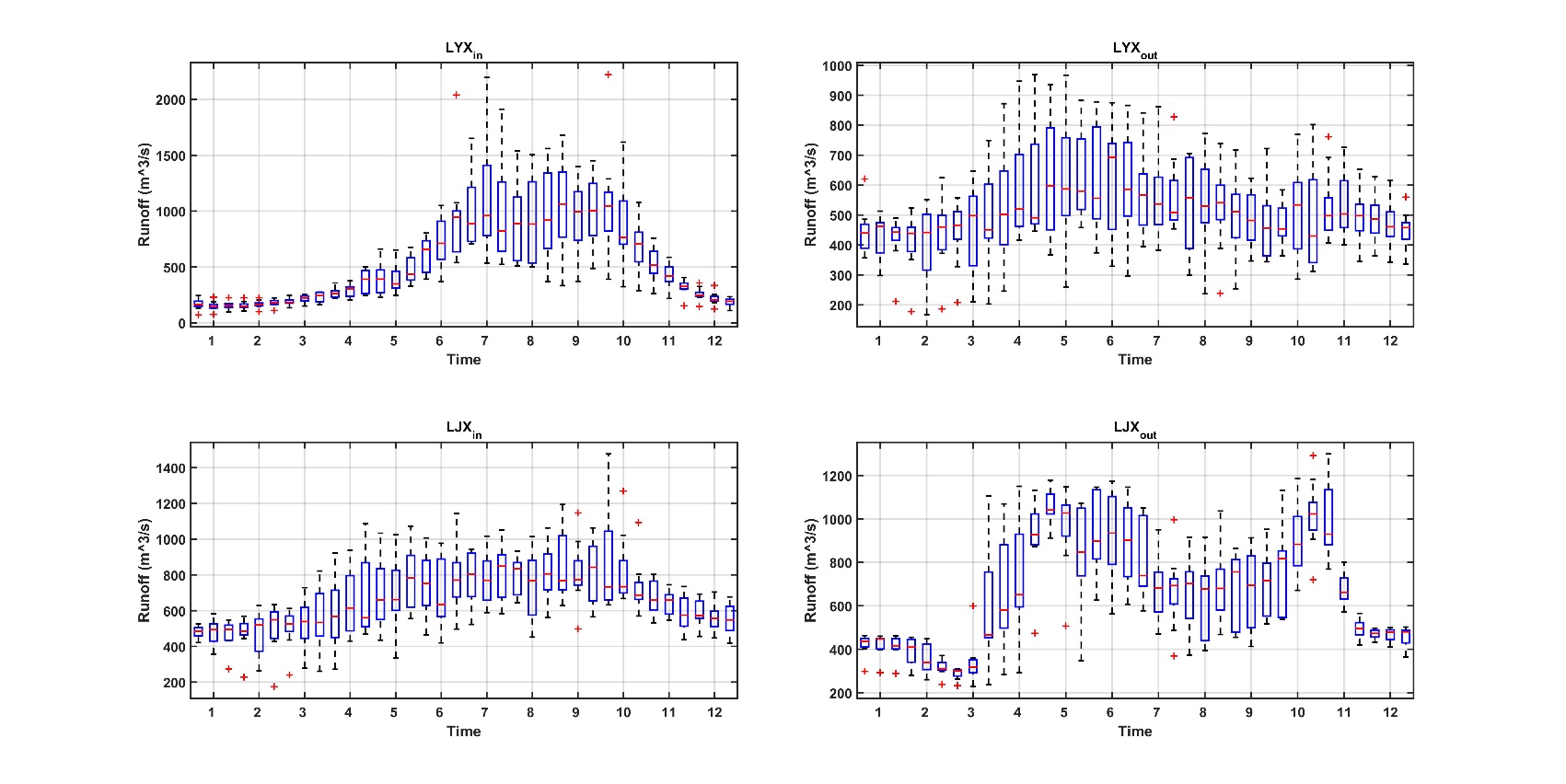
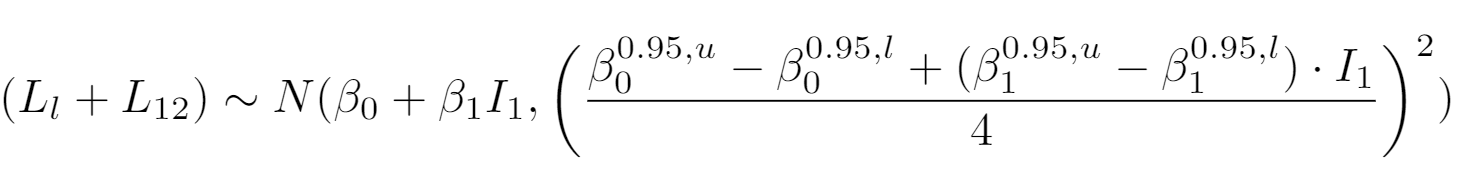


Figure 2. Seasonality of flow for pre-LYX (top left), post-LYX (top right), pre-LJX (bottom left), and post-LJX (bottom right)

Seasonality of flow is shown in Figure 2. Pre-LYX, post-LYX, and pre-LJX flows show similar monomodal patterns while the peak flow is more smoothed for post-LYX and pre-LJX flows. Interestingly, post-LJX flow shows a somewhat bimodal pattern.

Still, we observed good linear relationship between the inflow (pre-LYX) and the lateral flow (pre-LJX – post-LYX) as shown in Figure 3. A two-step algorithm is used to resample the flow data (i.e. inflow and lateral flow). The inflow is generated using a synthetic streamflow generator that preserve the temporal autocorrelation. The lateral flow is then conditioned on the sampled inflow data and is approximated using a Gaussian random variable to preserve the positive correlations between inflow and lateral flow. The standard deviation linearly increases with the amplitude of inflow and is approximated as ¼ of difference between the upper and lower 95% fitted lines.



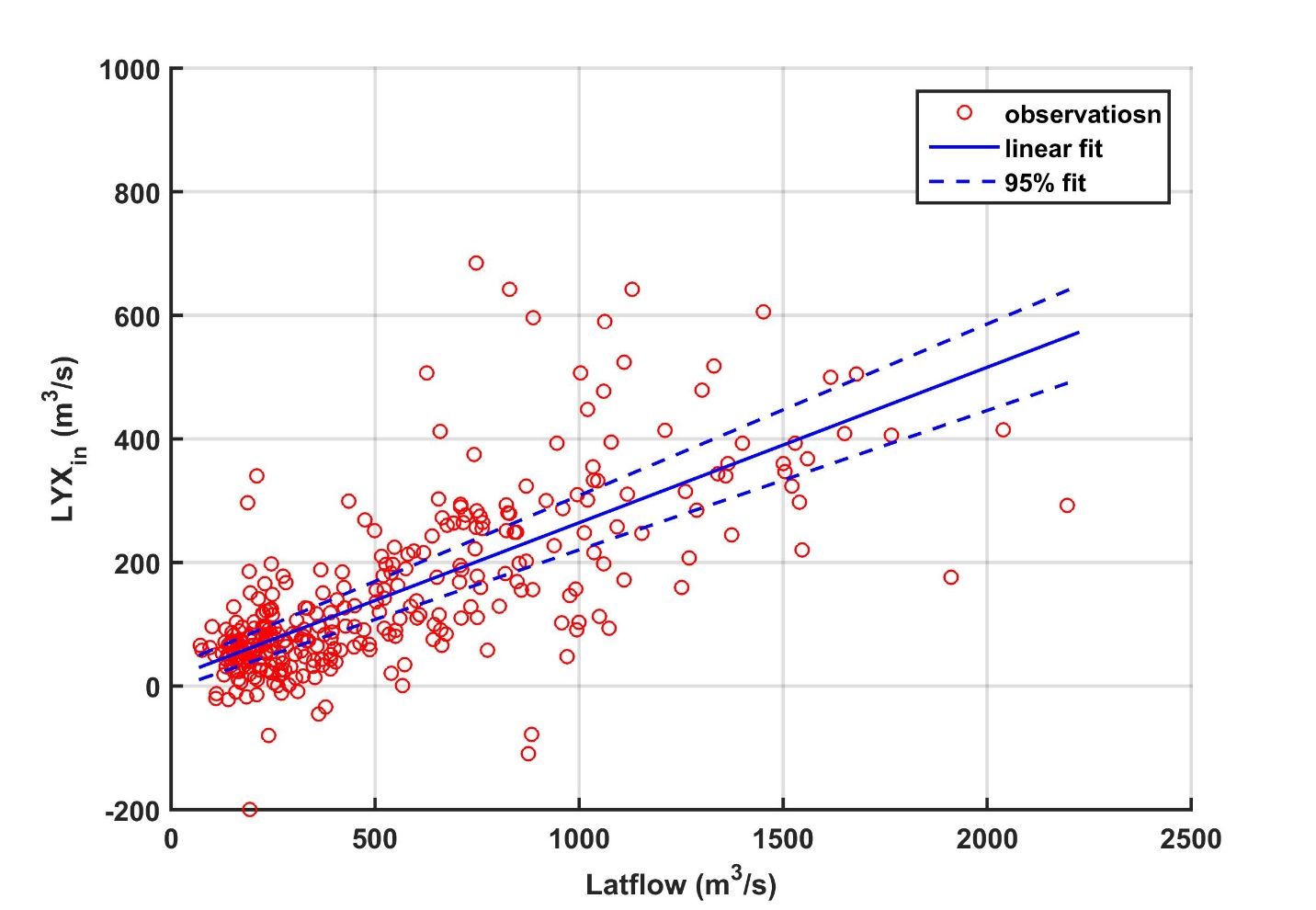


Figure 3. Scatter plot between the inflow and lateral flow and the linear fits

2) SOW generator

An auto-regressive model is approximated based on the auto-correlation structure of the observational flow as follows Kirsch et al., 2013 to generate synthetic streamflow at pre-LYX. The observational flow data is denoted by y or Y. For flow in the matrix form Y, the matrix is organized such that each row is flow of one year and each column is flow data of the given 10-day period across years. Therefore, Y is a 9x36 matrix here. The algorithm consists of 5 steps:

i) Y1 = log(Y0) to reduced skewness;

ii) Ya = (Y1 – mean(Y1))/std(Y1) to reshape flow data to roughly normal;

iii) to sample uncorrelated normal random variables X(i,j)~N(0,1) ();

iv) to decompose the autocorrelation structure of Ya using the Cholesky factorization corr(Ya) = Y’\*Y = Q’\*Q;

v) to impose the autocorrelations onto X to generate Z such that Z = X\*Q

It is easy to prove that Z, the generated samples, preserve the autocorrelations. corr(Z) = Z’\*Z = Q’\*X’\*X\*Q. X’\*X ~= I since X are uncorrelated normal rv. corr(Z) = Z’\*Z = Q’ \*Q = corr(Ya).

The assumption of normality is well assumed after step (i) and (ii) as shown in Figure 4.

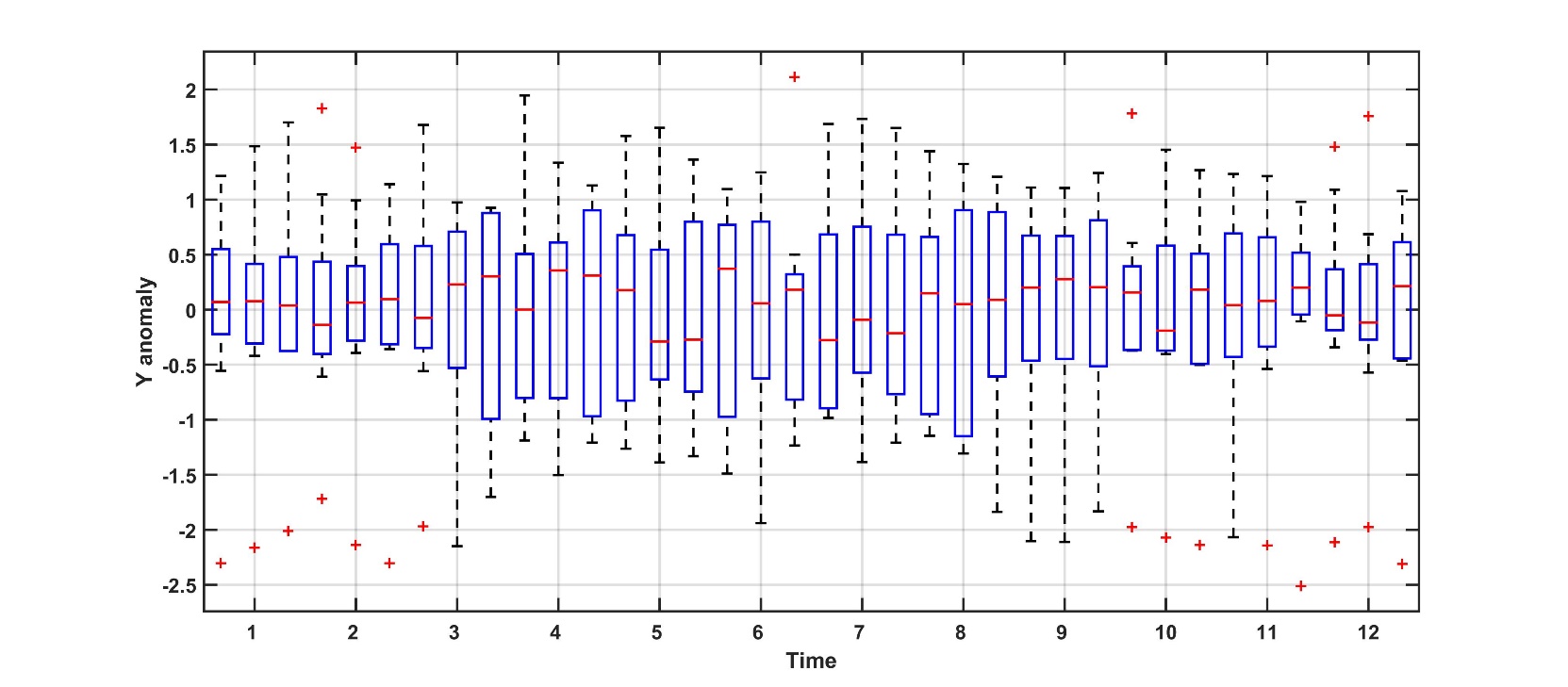


Figure 4. Boxplot of Ya

N is set to 900 and autocorrelations as calculated given by corr(Ya) and corr(Z) are compared in Figure 5, which shows very good consistency. The N = 900 samples for each 10-day period are then divided into 100 equally-sized subsets to estimate the time-varying lag-1 and lag-2 correlations as shown in Figure 6. Consistent inner-annual variations are observed for the lag-1 and lag-2 correlations between Ya and Z. It should be noted that we are only imposing the inner-annual autocorrelation to the sampled flow data given how we organize the data matrix Y (lag <= 36).

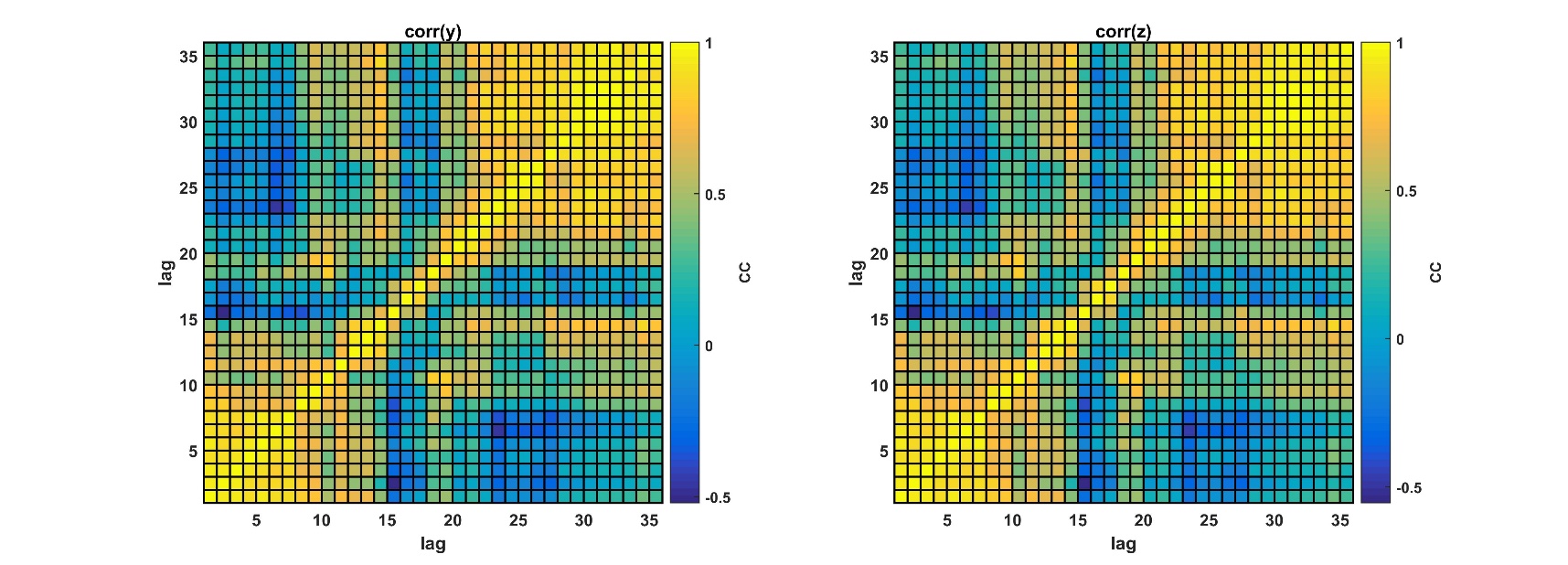


Figure 5. Autocorrelation maps for Ya (left) and Z (right)

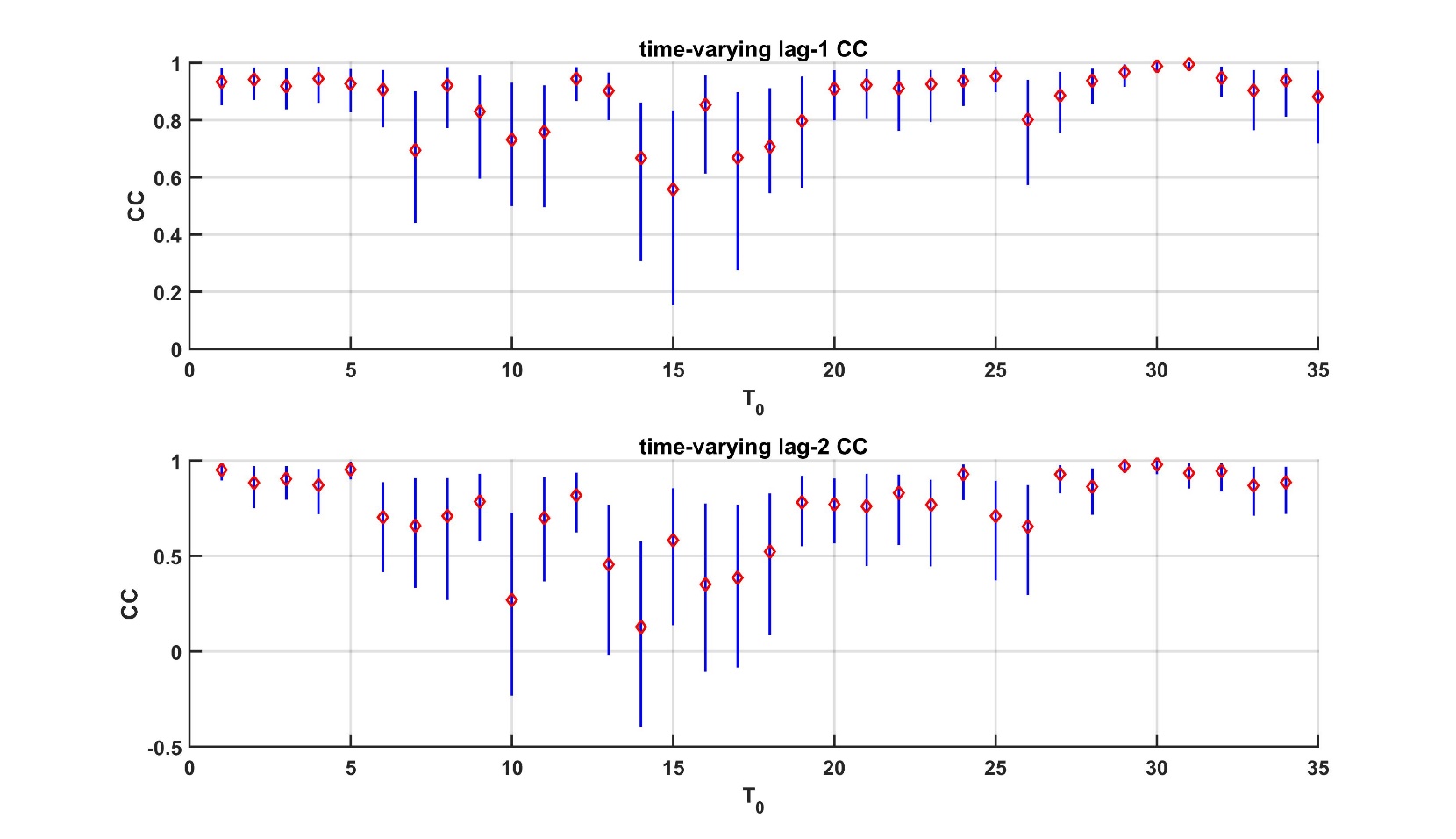


Figure 6. Lag-1 (top) and lag-2 (bottom) correlations for Ya (red diamonds) and Z (blue line for the 90th-10th percentile range)