

The number of lags should depend upon what correlation time you think might be in the data. For one year of data, you would look for times less than a year. That might be seasonal (90 days) or shorter. So 150 lags might work. If you look at multiple years, then you might look for annual or semi-annual variations so you might look at 2 or 3 times 365. On the other hand, if you did the daily data points over say, 4 years, you would have all the correlation times with periods shorter than 4 years, including seasonal or shorter. The same is true if you do the whole data set to look for solar-cycle correlations. Then the lagged correlation function will have variations on all the sub-time scales that might make it hard to separate.

You can get around this by perhaps starting with individual years and determining any correlations less than or equal to seasonal times. Then you can average the data over season (90 days) and do multiple years to see there are annual or semi-annual relationships. Finally, you can average over a year to see if there are relationships with correlation times longer than a year (e.g. solar cycle).

The correlation routine will limit you so that you have at least 2 points of overlap between the two data sets at each end of the lagged correlation. This is so that it can calculate the uncertainty values. All the statistical estimates will have a term like $1/(N-1)$, and when $N=1$ this blows up!