Subject: Re: Seasonal Variation in Trend Analysis
Posted by Kenneth P. Bowman on Thu, 27 Sep 2007 03:01:44 GMT
View Forum Message <> Reply to Message

In article <MPG.2164877cbc85365998a09c@news.frii.com>,
David Fanning <news@dfanning.com> wrote:

> Folks,
>
> Does anyone happen to have an IDL example of some
> code that might remove seasonal variation in a time
> series? Or some suggestions for how to proceed in IDL?
> I can see that I might want to use a model that has
> sin and cosine terms, but I can't see how to find the
> coefficients of such a model in IDL.

There are several ways to do this, David, depending on exactly what
you are trying to do.

Let's assume that you have evenly-spaced data, say monthly,
and you have complete years of data (12*nyears data points).
Those assumptions are not essential, but make the analysis
simpler.

One approach is to remove the climatological annual cycle. This is
done by computing the average of all the Januaries, all the Februaries,
etc. Then subtract the average January from each of the individual
Januaries. The residual is called the "anomaly".

Another approach is to remove the annual harmonic (a perfect
sinusoid). This can be done with an FFT, but is usually easier to do
with simple regression. (FFT and regression are equivalent for
evenly-spaced data with no missing values.)

The difference between these two methods is that the second
method removes only the first harmonic of the annual cycle.
The first method removes the first *six* harmonics. In many
cases the annual cycle is dominated by the first harmonic,
so the results of the two methods are similar. In some cases, the
semi-annual harmonic might be important, in which case the second method
would not work as well. The second method, however, guarantees
a smooth estimate of the annual cycle. (Note that you could also compute
the higher harmonics, with FFT or REGRESS, and then remove as many
harmonics as you want.)

Here is a sample program that illustrates both approaches using some
synthetic monthly data. The values are assumed to be located at the
centers of the months.

Cheers, Ken

```
;   Removing the climatological mean

nyears  = 4                             ;Number of years
nmonths = 12*nyears                     ;Number of months
t       = (0.5 + FINDGEN(nmonths))/12   ;Time in years

z = COS(2.0*!PI*t) + 0.2*RANDOMN(seed, n)       ;Data with an annual cycle
z = REFORM(z, 12, nyears)                       ;Rearrange for convenience
zclim = REBIN(TOTAL(z, 2)/nyears, 12, nyears)   ;Compute climatological mean

PLOT, t, z, PSYM = 1                    ;Plot original data
OPLOT, t, zclim                         ;Plot climatological mean
OPLOT, t, z - zclim, LINESTYLE = 1      ;Plot anomaly

cr = ''
READ, cr, PROMPT = 'Enter <cr> to continue : '



;   Removing the annual harmonic

z = COS(2.0*!PI*t - 0.75) + 0.2*RANDOMN(seed, n)    ;Data with an annual cycle
x = TRANSPOSE([[SIN(2.0*!PI*t)], [COS(2.0*!PI*t)]]) ;Create predictor variables

coeffs = REGRESS(x, z, CONST = zmean, YFIT = yfit)  ;Fit the sinusoids
PRINT, SQRT(TOTAL(coeffs^2)), ATAN(coeffs[0], coeffs[1]) ;Amplitude and phase

PLOT, t, z, PSYM = 1                    ;Plot original data
OPLOT, t, yfit                          ;Plot climatological mean
OPLOT, t, z - yfit, LINESTYLE = 1       ;Plot anomaly
```