On Sunday, 26 October 2014 01:00:08 UTC+2, siumt...@gmail.com wrote:

> I do not use FFT because I have missing data . I provided you with monthly timeseries which does not have missing data. But generally, I use monthly datasets that have missing values.

Ok. You should have told so then...

However, you can still repeat my harmonic analysis using a Fourier transform for irregularly sampled data (which is really a kind of multiple regression since the fourier transform is linear).

> That is why I used multiple regression.
>
> Assumption about regression is though that the dataset follow gaussian distribution. Maybe I am not getting the correct coefficients of the sine and cosine terms because my data is skewed (I can not think any thing else)

This is true in some sense, since there residual signal after subtracting your seasonal signal still is significant and not really noiselike (nonzero skewness and/or kurtosis, colored). One could say that you have a very low SNR, if you interpret the residual signal as noise.

> Second, I believe my model is correct because Suppose you have monthly temperature datasets. You can represent your seasonal cycle using the harmonics as I have .

You haven't told whether your goal is to extract the seasonal signal or to fit the full signal to some mode.

Indeed you can respresent a lowpass seasonal signal with period 12 moths with your model, using 2x4 coeffcents (plus a constant term you should include or subtract from the original signal). What I meant is that at least for the signal you posted, this seasonal component does not contain much energy. Thus, your model yields a very bad fit to your original signal. I showed in my previous post, with a simplified harmonic analysis, that most of your signal energy is contained in harmonic components with longer periods than 12 months.

However, this is now diverging away from IDL and into signal processing theory. Your problem is not IDL. From the style of your code, I assume you inherited it from someone who wrote it in the 90's (or didn't change style since then...). The code probably works correctly for the problem it was written for. REGRESS has a few keywords that will provide goodness-of-fit statistics. I suggest you try them, and you will see that your model is not good for the full signal.

--
Yngvar