

大気環境変動の統計解析システムの使用法

■更新記録

オリジナルドキュメント作成日：2013 年 5 月 作成者：京都大学生存圏研究所 新堀淳樹

統計解析システムの構成

本統計解析システムは以下の 5 つの機能から構成されている。

- (1) 平均値検定：2 データの分布についての平均値差異
- (2) 相関係数解析：データ間の相互相関係数と無相関検定
- (3) コヒーレンス解析：周波数スペクトルの各成分に対するコヒーレンスと位相
- (4) S 変換解析：各周波数成分の卓越周波数とスペクトル密度の時間変化
- (5) トレンド検定：線形回帰直線の変化傾向の有無

なお、一般に観測データは必ずしも時間・空間に関するサンプル間隔が一定でなく、また欠測を含む場合もあるので、解析に先立って線形補間を行う機能もある。

統計解析システムを使う

本統計解析システムは単なる IDL プロシージャ群なので、IDL ver6.3 以降が動く Windows PC、Mac、Linux、Sun なら、どの環境でも動かすことができる。生存圏研究所の大学関連サーバー(renkei1)にある zip ファイルをダウンロードして解凍すればよい。

<http://mdb01.rish.kyoto-u.ac.jp/~shinbori/udas/>

ここにある statistical_package.zip というファイルがそれに該当する。これを自分の PC 上で展開し、それを /tdas_7_00/iugonet/tool/以下に移動させる。この作業を行う前に、IUGONET の解析ソフトウェア(UDAS)がインストールされている必要がある。UDAS のインストール方法は、以下のサイトが参考になる。

http://www.iugonet.org/software/documents/udas_installation_j20120216.pdf

このドキュメントには、UDAS の親になっている TDAS のインストール方法や IDL のパスの設定方法などが記載されている。なお、以下のサイトには、過去に行われた IUGONET データ解析講習会で配布された資料なども閲覧できる。

<http://www.iugonet.org/software/install.html#install>

統計解析システム procedures

■ udata_interpolation

一般に観測データは、必ずしも時間と空間に関するデータ間隔が一定でなく、欠測を含む場合が多い。本統計解析システムはフーリエ変換を用いたスペクトル解析等、時間に対して等間隔のデータを要求する解析を含んでいるため、解析対象のデータを整形する時刻間隔均等化やデータ補間を事前に行う機能が必要である。時刻間隔の均等化は読み込まれたデータ内の最小の時間間隔を基準に行われる。例えば、最小サンプリング間隔が1時間のデータを読み込む場合、2時間以上の欠測が存在する箇所すべてについて1時間毎のデータ補間を行う。具体的には、以下のように動かせばよい。

udata_interpolation, vname1, vname2

vname1, vname2 には、それぞれ 1 次元時系列データの tplot 変数名が入る。2 種類の時系列データを同時に補間するこのプロシージャを動作させると、IDL コンソール上に以下のメッセージが出力され、***_interpol という tplot 変数名ができる。今後の統計解析システムを構成しているプロシージャにかませる tplot 変数名は、***_interpol というものを使用する。

下のメッセージは、太陽活動度指数(黒点数と電波フラックス)のデータについて行った結果である。

```
> udata_interpolation, 'ssn', 'f10_7_solar_flux'
```

```
Start Time : 1947-02-14/00:00:00
```

```
End Time : 2013-03-31/00:00:00
```

```
86400.000 [sec] : Time Interval
```

```
71343 : number of ssn data
```

```
24197 : number of f10_7_solar_flux data
```

```
24153 : number of ssn_interpol data (interpolated ssn data)
```

```
24153 : number of f10_7_solar_flux_interpol data (interpolated f10_7_solar_flux data)
```

```
0 : interpolation number of ssn_interpol data
```

```
642 : interpolation number of f10_7_solar_flux_interpol data
```

なお、キーワードで補間する開始時刻、終了時刻の設定や 2 データのうち、どちらを基準にするかを定めることができる。

■ ucross_cor

2 種類の 1 次元時系列データ間の相互相関係数を求め、その値に対して無相関検定を行う。具体的には、以下のように動かせばよい。

ucross_cor, vname1, vname2

vname1, vname2 には、それぞれ 1 次元時系列データの tplot 変数名が入る。ただし、データ補間プロシージャ udata_interpolation で生成された***_interpol という tplot 変数名を指定する。これを実行すると、以下のようにコンソール上に 2 データ間の相互相関係数と無相関検定結果が出力される。

>ucross_cor, 'ssn_interpol', 'f10_7_solar_flux_interpol', sl=0.01

```
-----cross correlation status-----
|          maximun correlation coefficient          =      0.94658822
|          lag of max correlation                  =           0
| statistically significant (significance Lv         =      0.0100000)
|          minimun correlation coefficient          =     -0.53807378
|          lag of min correlation                  =          1908
| statistically significant (significance Lv         =      0.0100000)
-----
```

なお、上記のコマンドのようにキーワード sl を用いて、統計検定に用いる有意水準を指定することができる。なにもキーワードを指定しない場合は、有意水準 5% で検定を行う。

■ udifference_test

本統計解析システムに含まれる平均値検定は、与えられた 2 標本が同一母集団からの標本とみなせるかどうかを検定する。言い換えれば、2 標本の平均値が同一とみなせるかを検定する機能である。

具体的な処理工程は、初めに 2 標本の分布が正規分布であるかを検定し、その結果 2 標本とも正規分布とみなされれば、平均値が同一かを判定する Welch 検定を行い、そうでない場合は Mann-Whitney の U 検定を用い、平均値が同一かを判定する。注意すべき点として、データがあらかじめ正規分布することが推測される場合には、Welch 検定を使用する方が検定の精度は高い。そのような場合には、オプションで Welch 検定を指定して、平均値検定を使用することが望ましい。具体的には、以下のように動かせばよい。

udifference_test, vname1, vname2

vname1, vname2 には、それぞれ 1 次元時系列データの tplot 変数名が入る。ただし、データ補間プロシージャ udata_interpolation で生成された***_interpol という tplot 変数名を指定する。これを実行すると、以下のようにコンソール上に 2 データ間の平均値検定結果が出力される。

```
> udifference_test, 'ssn_interpol', 'f10_7_solar_flux_interpol', test_sel=1
-----Welch test result-----
t      105.31927      t0      1.9600146
There is a significant difference between these data with significance level =
0.0500000
-----
```

なお、上記のコマンドのようにキーワードs1を用いて、統計検定に用いる有意水準を指定することができる。なにもキーワードを指定しない場合は、有意水準5%で検定を行う。また、キーワードtest_selを用いてWelch検定とMann-WhitneyのU検定を選択できる。

test_sel=1 : Welch検定

test_sel=2 : Mann-WhitneyのU検定

キーワードを指定しない場合は、2データの正規分布判定を行い、自動的にWelch検定、またはMann-WhitneyのU検定が選択される。

■ utrend_test

これは、データ値の大きさの順位によって傾向性を検定し、トレンド検定に付加的に回帰直線の傾きおよびその信頼区間の情報を与えるプロシージャである。具体的には、以下のよう動かせばよい。

utrend_test, vname1

vname1には、1次元時系列データのtplot変数名が入る。ただし、データ補間プロシージャ udata_interpolationで生成された***_interpolというtplot変数名を指定する。これを実行すると、以下のようにコンソール上に1データのトレンド検定結果が出力され、図1に示すようなグラフが出現する。

```
> utrend_test, 'ssn_interpol'
-----trend test result-----
Z=      -34.2137
Max  |Z| =      155.409
```

Threshold (at S_Level = 0.0500000) : 1.64500

There is a negative trend.

slope -0.00205963

error of slope 0.000107970

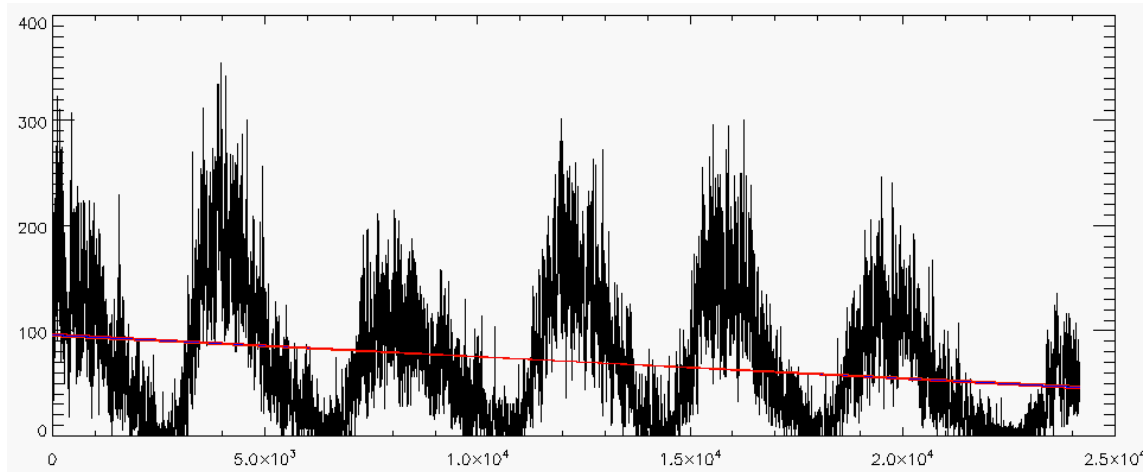


図1. 太陽黒点データに回帰直線を引いたグラフ。青線が回帰直線を表し、赤線が傾きの誤差を含めた回帰直線を示す。

なお、上記のコマンドのようにキーワードs1を用いて、統計検定に用いる有意水準を指定することができる。なにもキーワードを指定しない場合は、有意水準5%で検定を行う。

■uspec_coh

これは、2種類の時系列データに対してそれぞれの周波数スペクトル、クロススペクトルを求め、周波数スペクトルの各成分に対するコヒーレンスと位相を計算するプロシージャである。具体的には、以下のように動かせばよい。

uspec_coh, vname1, vname2

vname1, vname2には、それぞれ1次元時系列データのtplot変数名が入る。ただし、データ補間プロシージャudata_interpolationで生成された***_interpolというtplot変数名を指定する。これを実行すると、以下のようにコンソール上にコヒーレンス解析結果と各データの周波数スペクトル、コヒーレンスと位相のプロットが現れる。

```
> uspec_coh,'ssn_interpol','f10_7_solar_flux_interpol', deltat=86400
```

```
-----Coherence analysis result-----
```

```
coherence confidence interval = 6.20484e-005
```

max coherence = 0.99761842

main_period = 2.98117e+008 (data intervals)

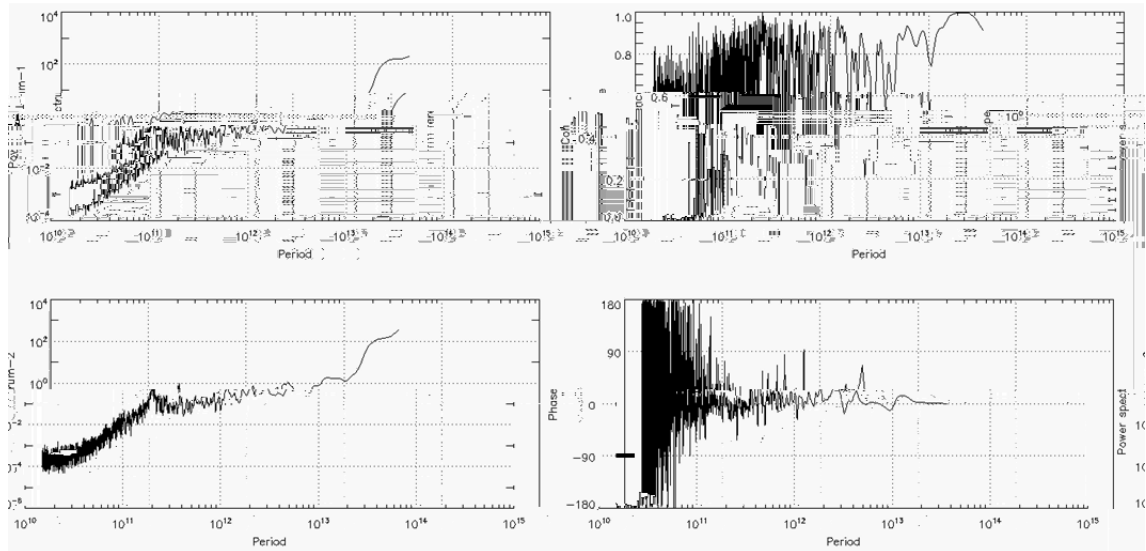


図2. 各データの周波数スペクトル、コヒーレンスと位相のプロット。

なお、`uspec_coh`には様々なキーワードを入れることができ、お好みに応じてフィルターをかけるデータ点の個数などを変更できる。主なキーワードは、以下である。

`deltat` : データの時間間隔

`sl` : 有意水準(デフォルト値=0.05)

`width` : フィルターをかけるデータ点の個数(デフォルト値= 10.0)

`wd` : プロットのウインドウ番号(デフォルト値= 1)

`window` : 窓関数の種類。Hanning、boxcar、gaussian、triangleの4つが選択可能
(デフォルトの窓関数 : hanning)

`xsize` : プロットウインドウのXサイズ(デフォルト値=1400)

`ysize` : プロットウインドウのYサイズ(デフォルト値=700)

■ `ustrans_pwrspc`

これは、1種類の時系列データに対してS(Stockwell)変換を行い、周波数成分の時間的に周波数が変化する様子や振幅変調を計算するプロシージャである。具体的には、以下のよう
に動かせばよい。

`ustrans_pwrspc, vname1, /sampling, /abs`

vname1には、1次元時系列データのtplot変数名が入る。ただし、データ補間プロシージャ udata_interpolationで生成された***_interpolというtplot変数名を指定する。これを実行すると、***_interpol_stpwrspecというtplot変数が作成される。このtplot変数に対してtplot コマンドを施してあげれば、図3のようにS変換したスペクトルが現れる。

> ustrans_pwrspc,'f10_7_solar_flux_interpol',/sampling,/abs

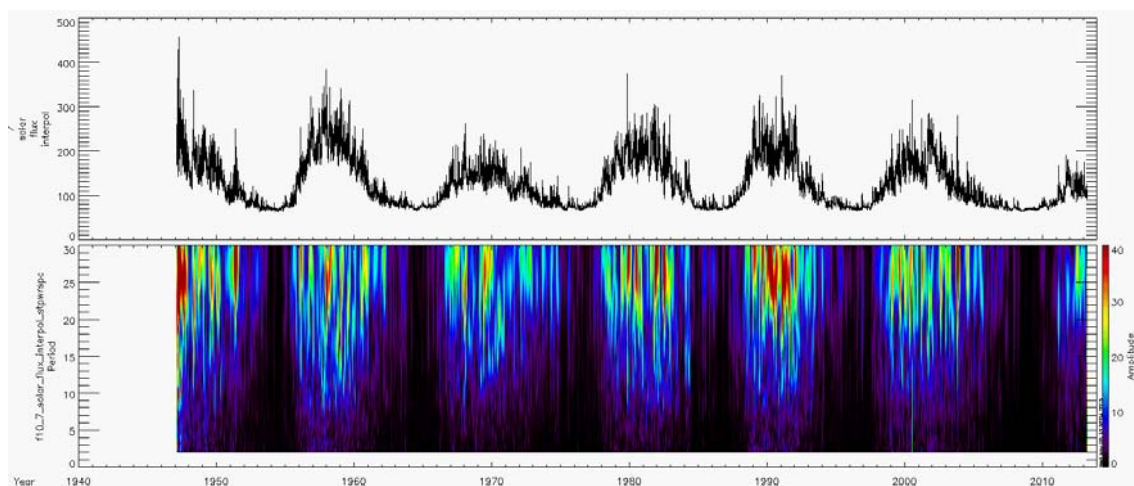


図3. 太陽電波フラックス(F10.7)の時系列プロットとそれをS変換したスペクトル図。

なお、ustrans_pwrspcには様々なキーワードを入れることができ、お好みに応じてフィルターをかけるデータ点の個数などを変更できる。主なキーワードは、以下である。

Help : 全てのキーワードとパラメータの説明を表示

verbose : エラーフラグとそのサイズ

samplingrate : 必須キーワード。これをいれるとS変換を行う。

入れるときは/samplingでよい。

maxfreq : S変換を行う最大周波数

minfreq : S変換を行う最小周波数

freqsmplngrate : 周波数間隔

power : パワー出力 (選択必須キーワード)

abs : 振幅の出力(選択必須キーワード)

Rremoveedge : 5%のテッパを用いて両サイドのエッジを除去。