多様な太陽地球環境データの相関解析及び その統計検定パッケージの開発

濵口 良太[1], 新堀 淳樹[2], 林 寛生[2], 津田 敏隆[2]

[1] 京都大学情報学研究科
[2] 京都大学生存圏研究所

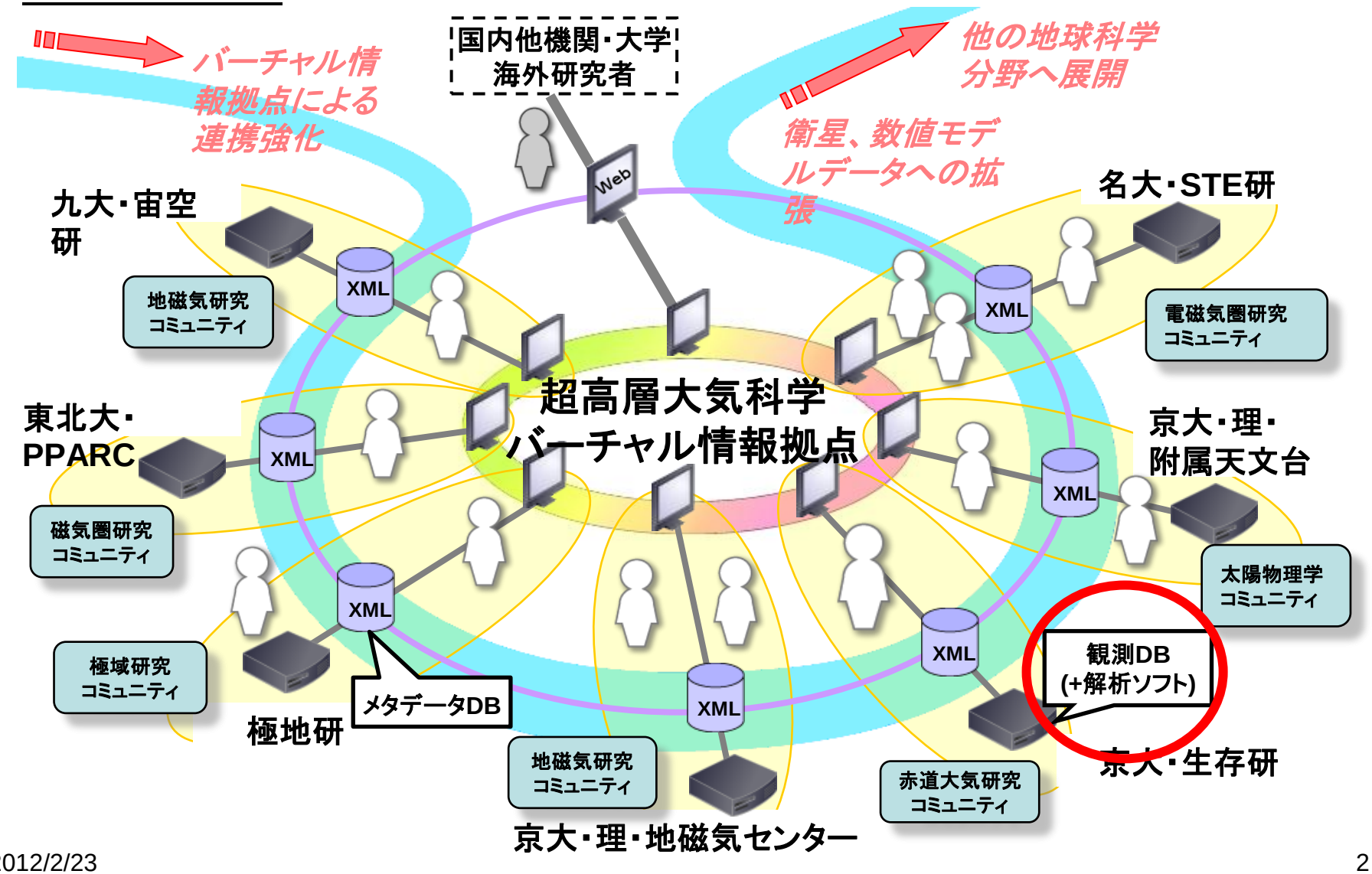
2012/2/23

発表の流れ

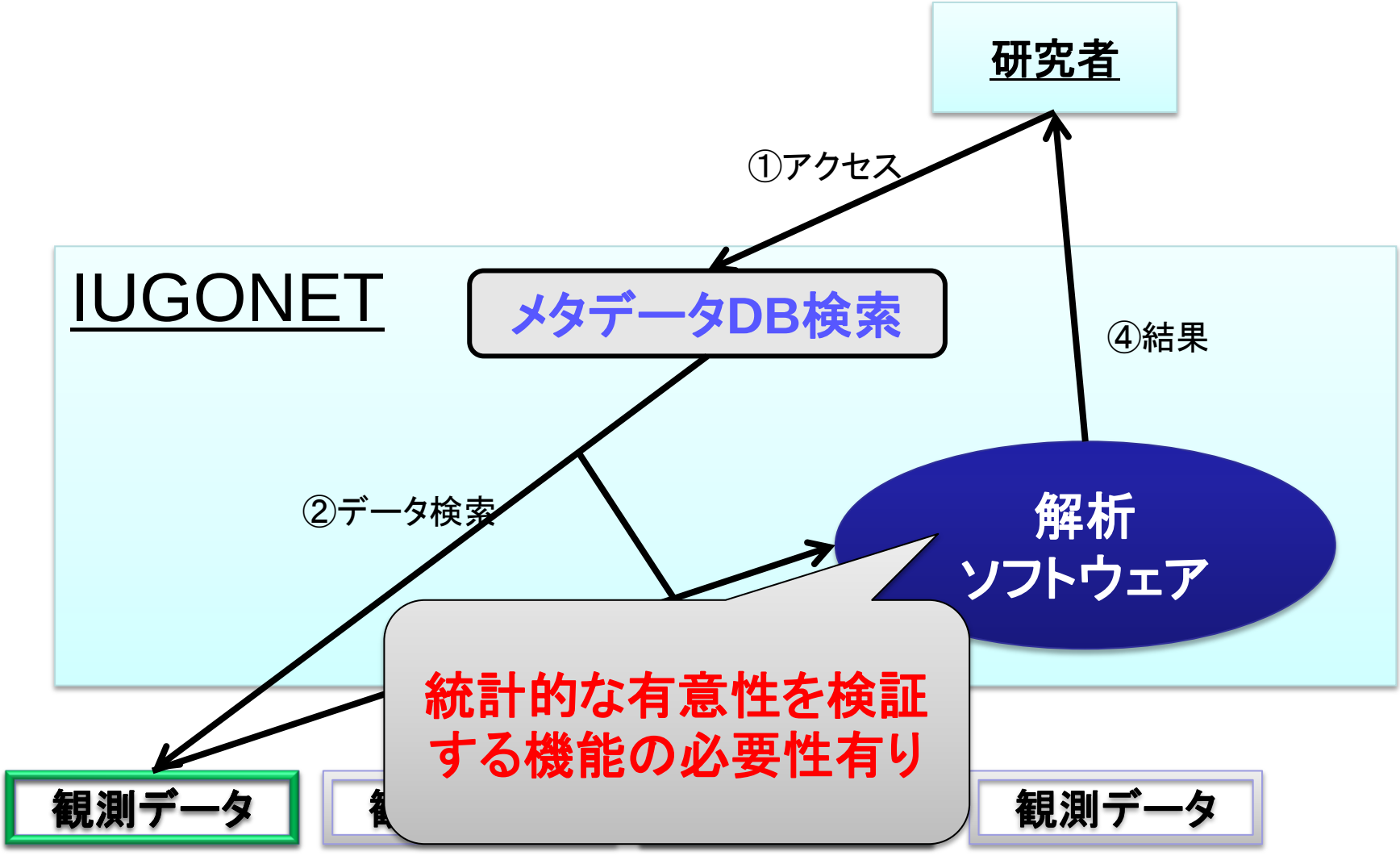
- 開発背景
- パッケージの機能説明
- パッケージ適用例

開発背景

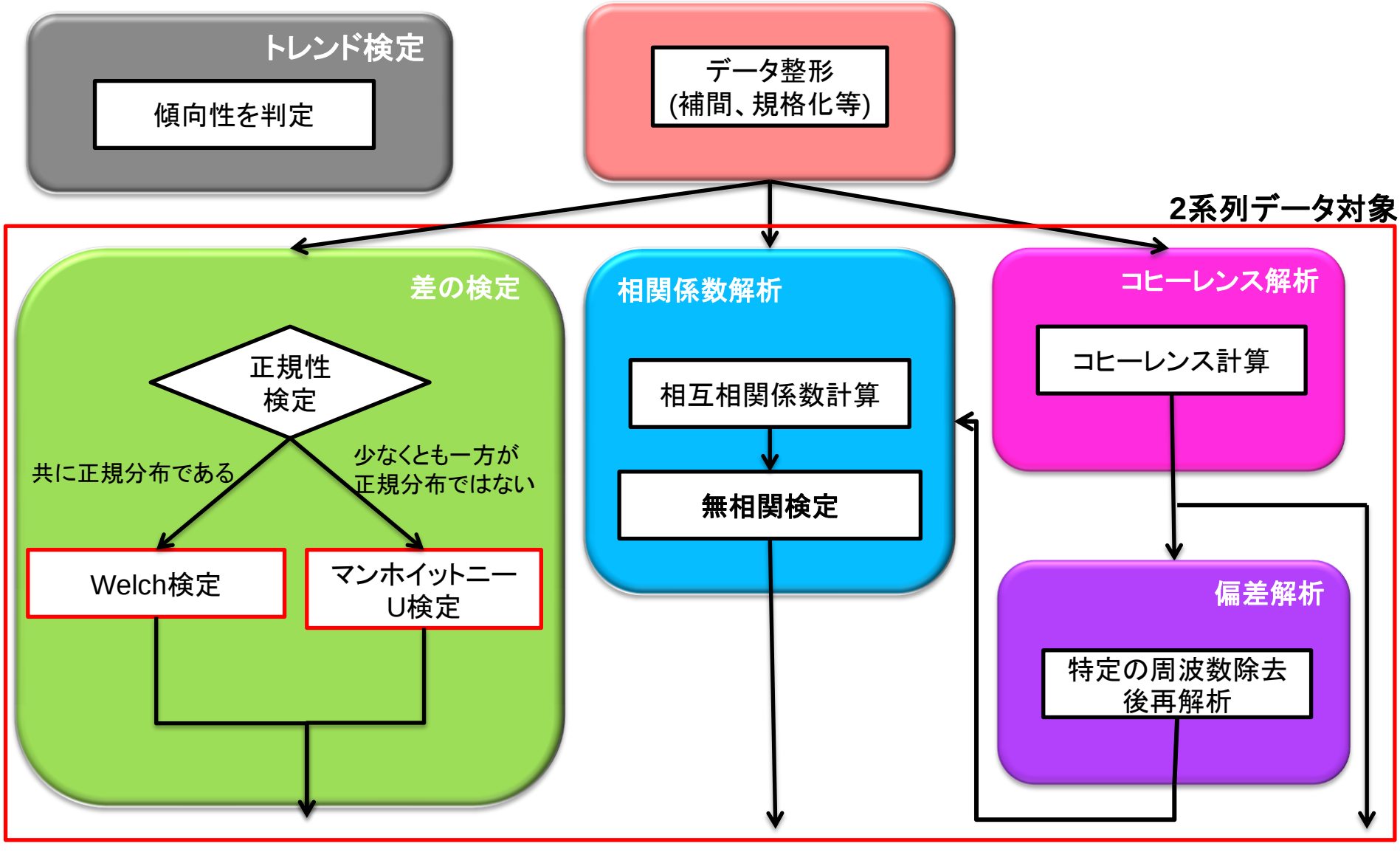
IUGONET



開発背景


研究者
①アクセス
IUGONET
メタデータDB検索
④結果
②データ検索
解析
ソフトウェア
統計的な有意性を検証
する機能の必要性有り
観測データ
観測データ

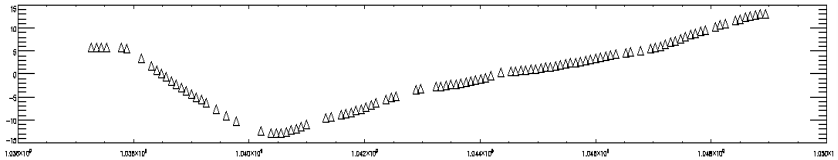
統計検定を含む解析パッケージ



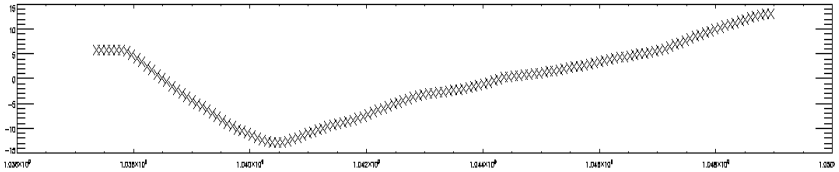
データ補間方法

・データ1内の最小時間間隔を1目盛として、等間隔に時間軸を決定
・その時間軸に対して、データ1及び2のデータを宛てていく
・時間に対応した点がない場合は線形補間によって補う

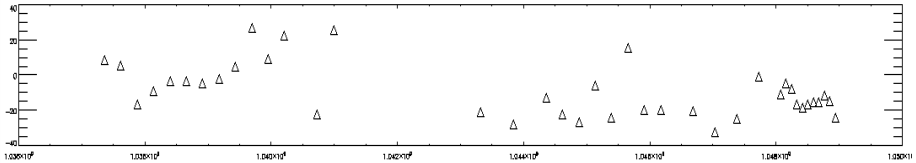
データ1



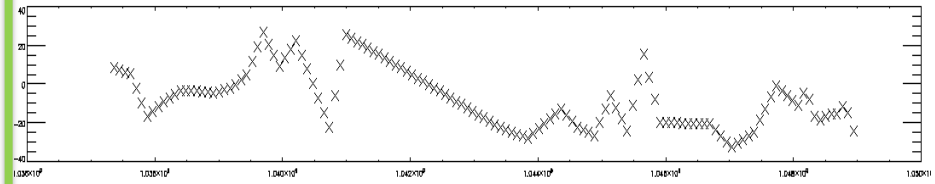
等間隔に補間



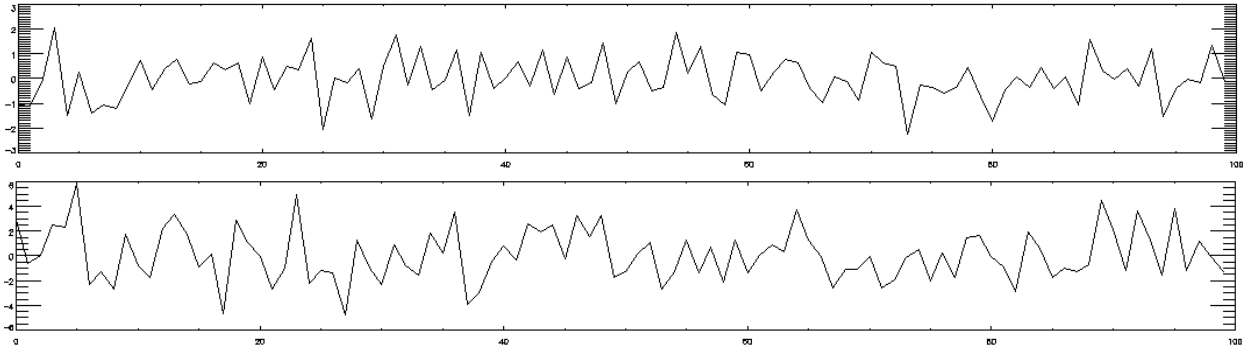
データ2



データ1の補間間隔
に合わせて補間



差の検定



振幅1倍

乱数

振幅2倍

```
-----------------cross correlation status---------------------------
|        maximun correlation coefficient        =      0.26153263
|                lag of max correlation        =           8
|NOT statistically significant (significance Lv    =    0.0100000)
|        minimun correlation coefficient        =     -0.22538732
|                lag of min correlation        =          16
|NOT statistically significant (significance Lv    =    0.0100000)
--------------------------------------------------------------------
```

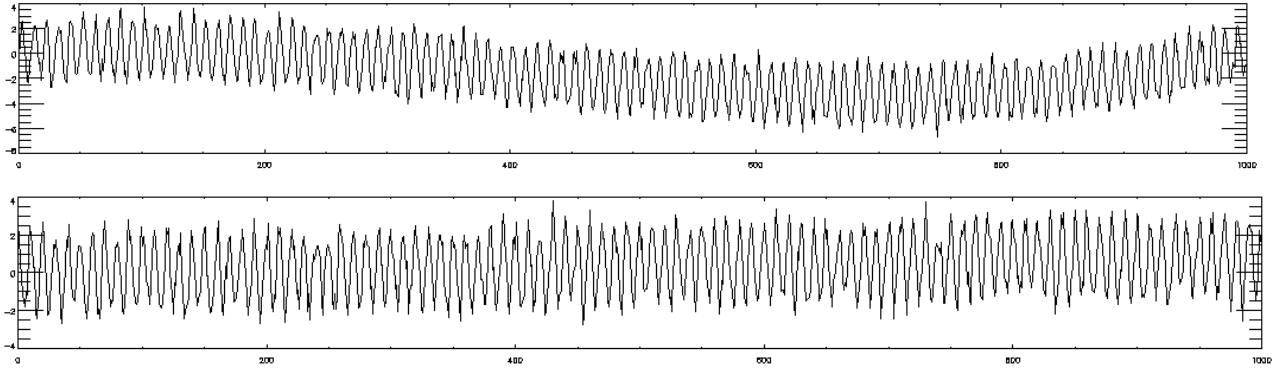
相関：低

```
comment:    NORMAL DISTRIBUTION with significance level = 0.05
      8.41067
      14.0671
comment:    NORMAL DISTRIBUTION with significance level = 0.05
Welch test
      1.03994       1.97207
 There is no difference between these data with significance level =     0.0500000
```

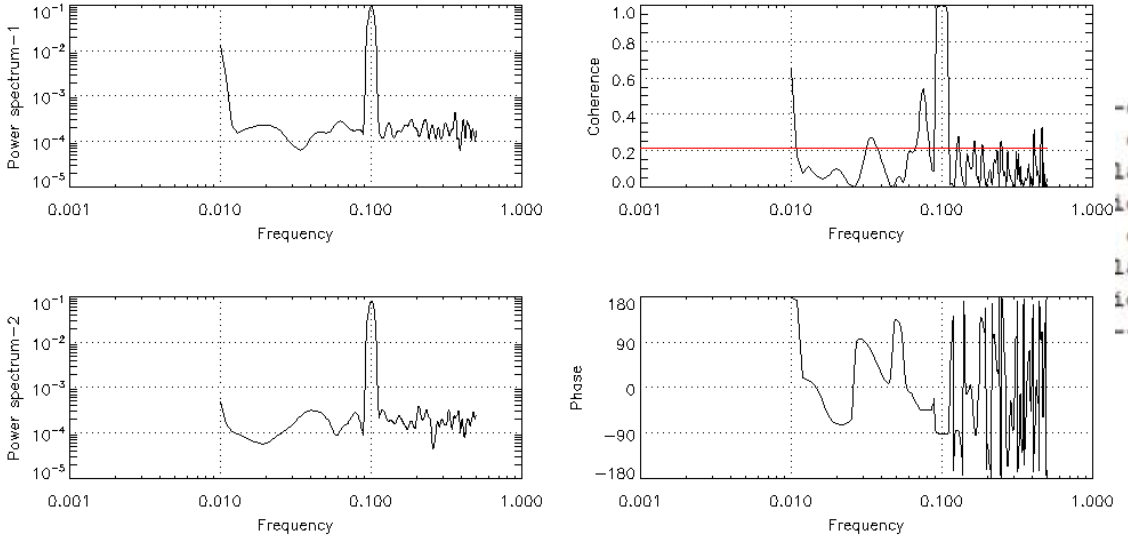
平均値は同じとみなせる

特定周波数除去処理

ある周波数（例えば日変化）が強く出ているデータの背後にある長周期変化を調べる



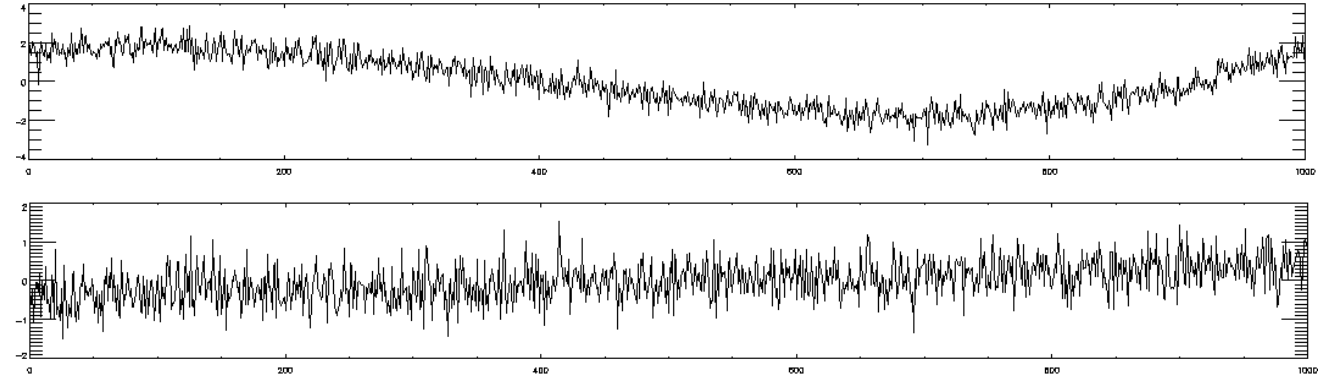
人工データ



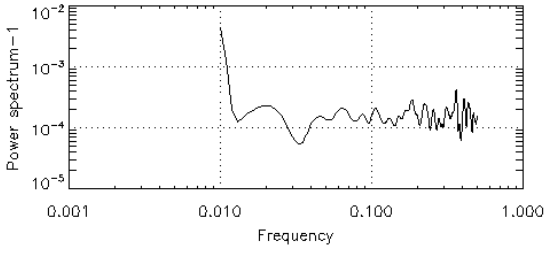

```
-cross correlation status-------------------------
 correlation coefficient        =      0.60589838
lag of max correlation        =          -3
ignificant (significance Lv    =     0.0100000)
 correlation coefficient        =     -0.76361740
lag of min correlation        =           1
ignificant (significance Lv    =     0.0100000)
--------------------------------------------------
```

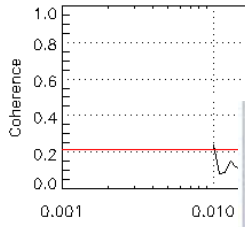
特定周波数除去処理

ある周波数（例えば日変化）が強く出ているデータの背後にある長周期変化を調べる

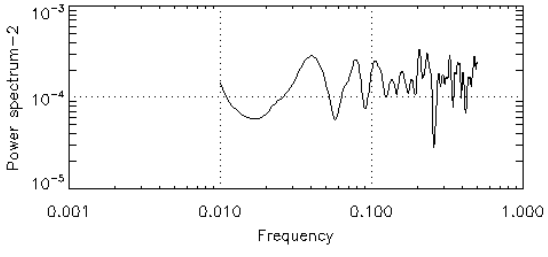


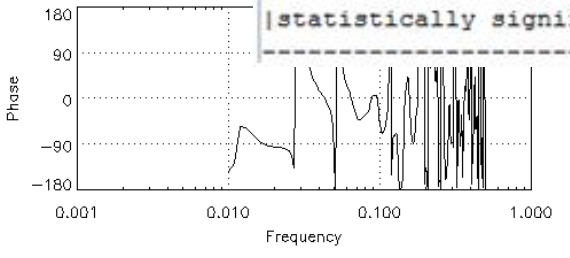
人工データ






```
-----------------cross correlation status-----------------------------
|        maximun correlation coefficient        =      0.19458576
|                lag of max correlation        =        -471
|statistically significant (significance Lv     =    0.0100000)
|        minimun correlation coefficient        =     -0.33576277
|                lag of min correlation        =        -501
|statistically significant (significance Lv     =    0.0100000)
----------------------------------------------------------------------
```

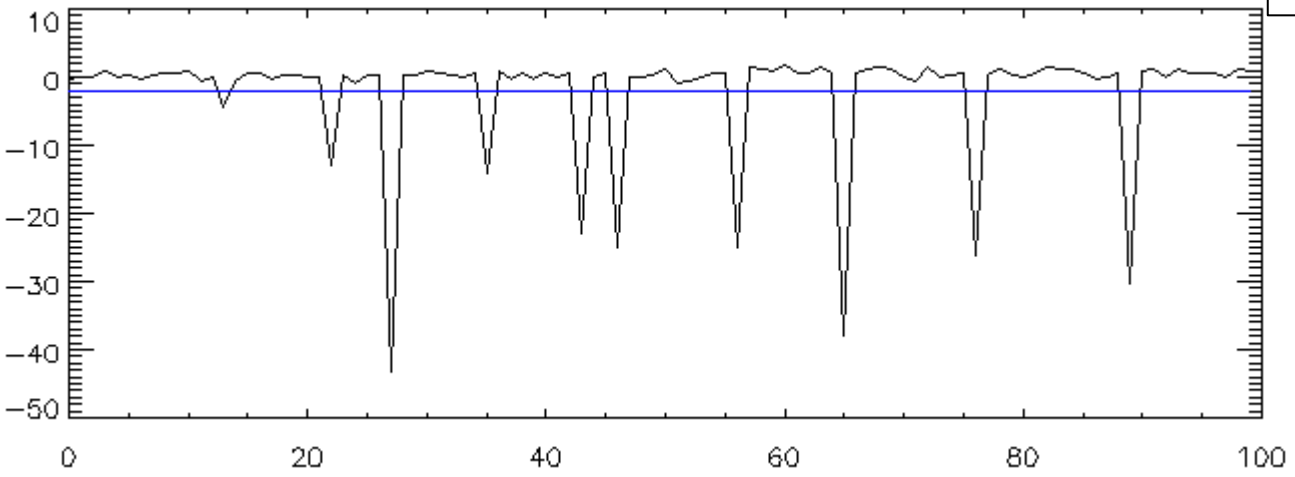




トレンド検定

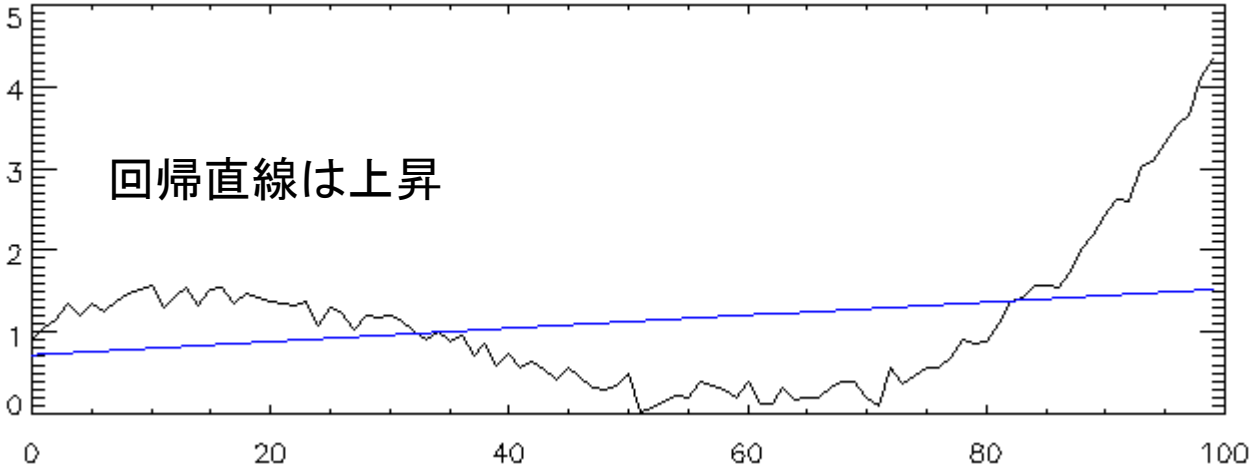
統計検定量 t がある有意水準の閾値よりも絶対値が小さい
⇒ 傾向があるとは言えない

微小増加データ



上昇トレンド有り

```
t=        3.11184
Max =        9.94987
S_Level@5%=1.645
S_Level@1%=2.325
```



トレンドなし

```
t=        0.125026
Max =        9.94987
S_Level@5%=1.645
S_Level@1%=2.325
```

パッケージオプション一覧

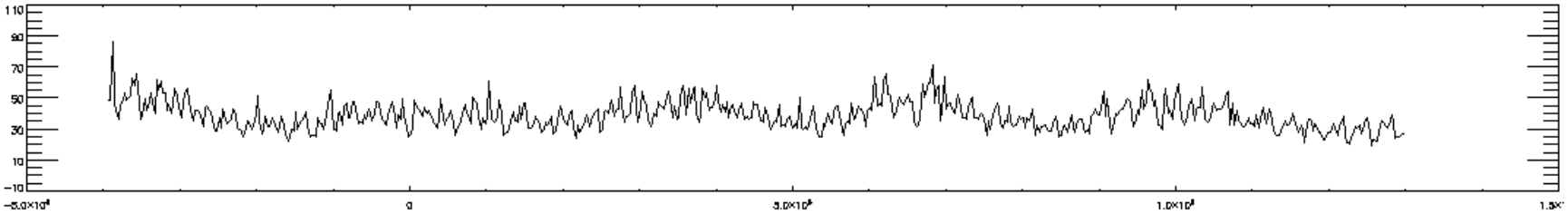
- 解析時間の指定
- 補間間隔の指定
- 相互相関の最大最小ラグ指定
- 各検定における有意水準指定
- 欠損値指定
- コヒーレンス計算時の平滑化幅
- 偏差解析を行うかどうか

パッケージ適用

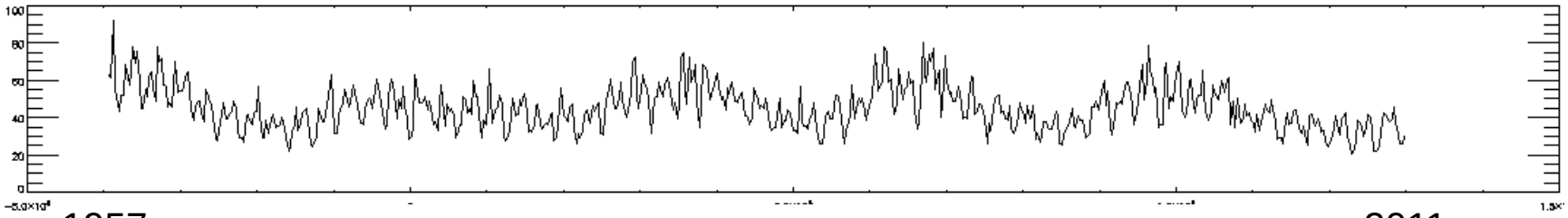
適用対象データ観測点
女満別　地磁気
北緯:43.903
東経:144.189
柿岡　地磁気
北緯:36.232
東経:140.186
流星個数
Meteor rader
Kototabang
南緯:0.203
東経:100.318
モンゴル
中国
韓国
日本
東シナ海
ネパール
ミャンマー
タイ
南シナ海
ベンガル湾
ベトナム
フィリピン
アンダマン海
タイランド湾
マレーシア
インドネシア
ジャワ海
バンダ海
アラフラ海
パプア
ニューギニア
ソロモン海

パッケージ適用

相関が高いとされるもの　⇒　柿岡、女満別の地磁気日変化振幅データ



柿岡



女満別

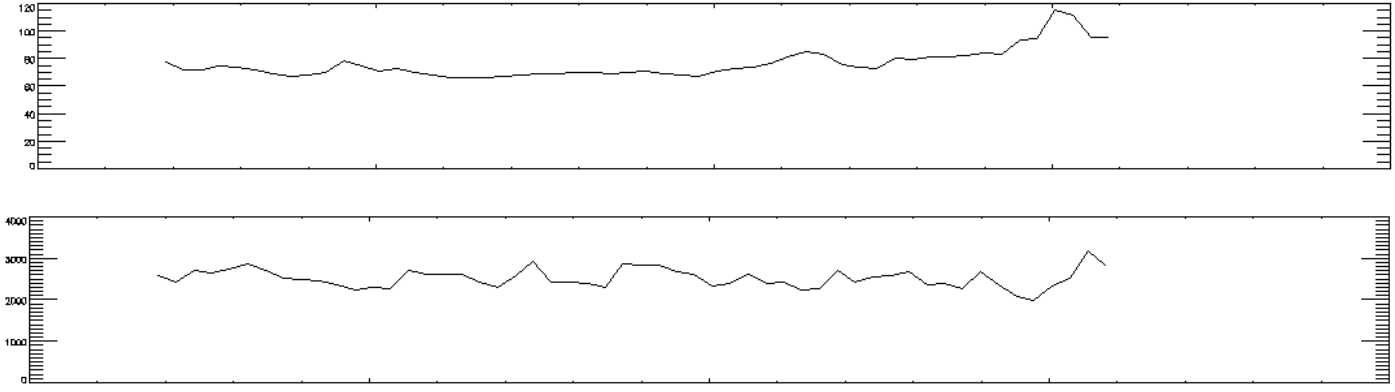
1957　　2011

```
-----------------cross correlation status--------------------------
|        maximun correlation coefficient        =     0.94283086
|               lag of max correlation        =           0
|statistically significant (significance Lv    =    0.0500000)
|        minimun correlation coefficient        =     -0.38612247
|               lag of min correlation        =        -152
|statistically significant (significance Lv    =    0.0500000)
-------------------------------------------------------------------
```

高い正相関

パッケージ適用

相関が未確認の（おそらくない）もの　⇒　F10.7 indexと日本の自殺者数



F10.7

自殺者数

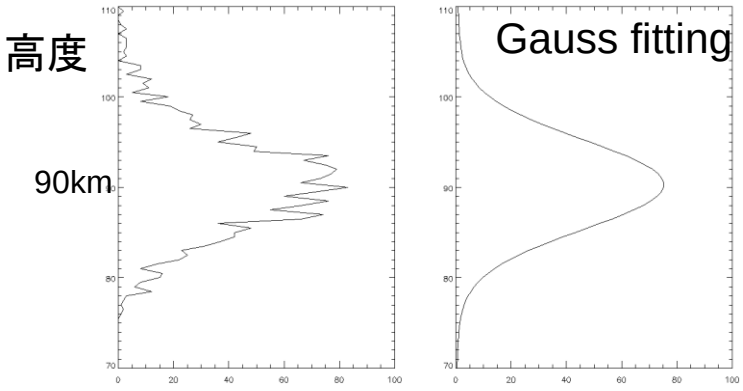
```
-----------------cross correlation status-------------------------
|        maximun correlation coefficient        =      0.22121306
|               lag of max correlation        =          16
|NOT statistically significant (significance Lv    =    0.0500000)
|        minimun correlation coefficient        =     -0.31267747
|               lag of min correlation        =          20
|NOT statistically significant (significance Lv    =    0.0500000)
------------------------------------------------------------------
```

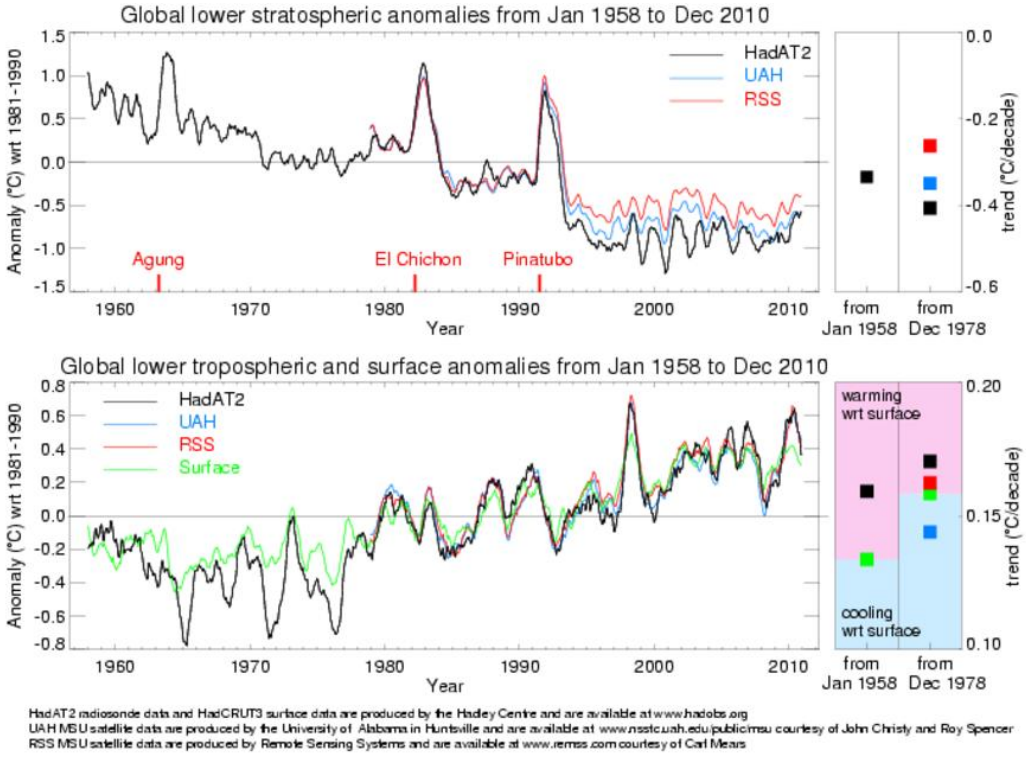
相関：低

検定：無相関である

トレンド検定適用例　流星高度トレンド

流星高度分布



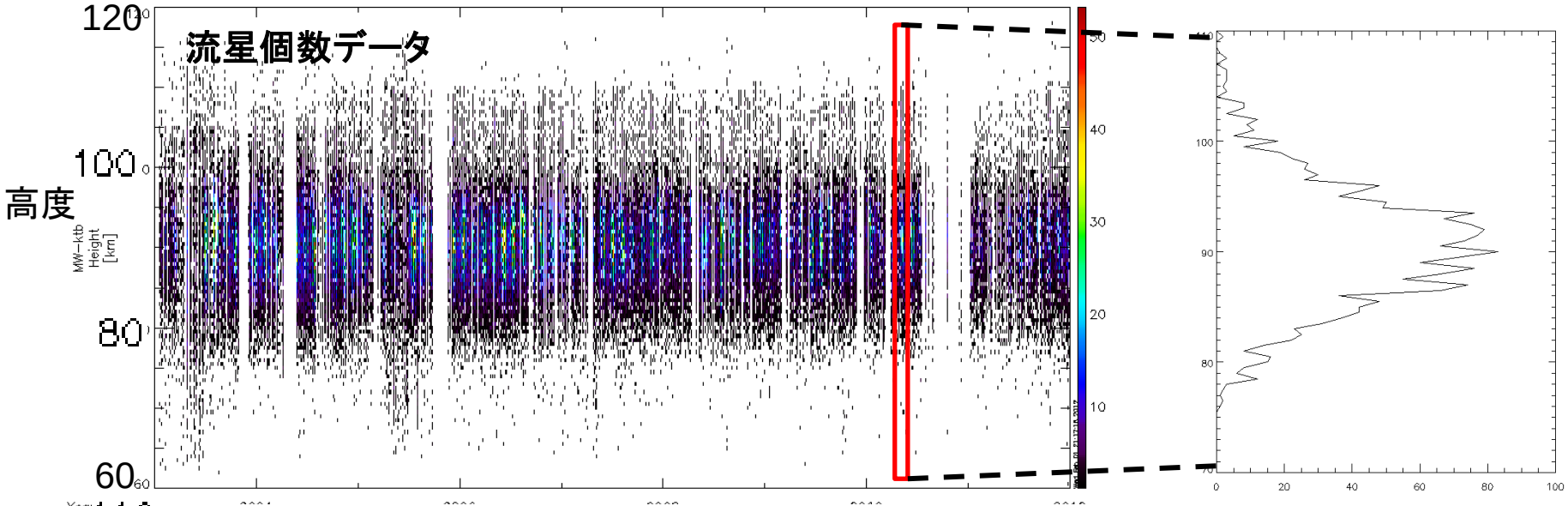


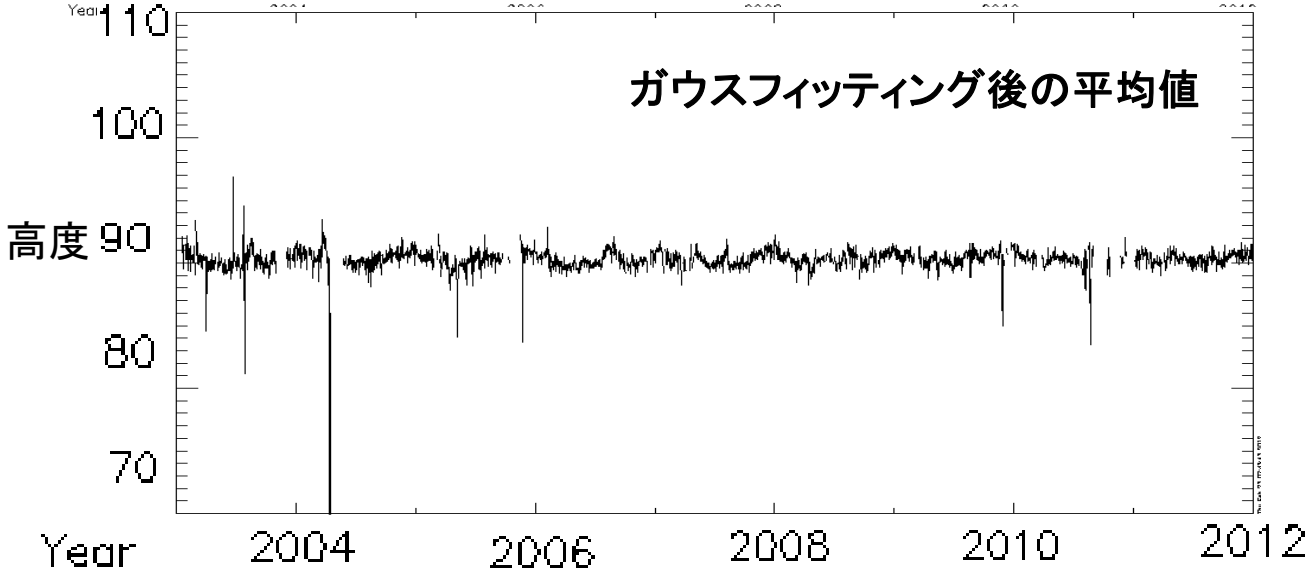
流星高度は大気密度に依存

地球温暖化により流星高度分布の変動が観測されるのではないか

大気の温度構造の変化は大気密度の変化をもたらす

トレンド検定適用例　流星高度トレンド





t= 1.69373
Max = 53.6470
S_Level@1%=2.325

有意水準1%でトレンドがあるとは言えない

まとめ

- 解析ソフトウェア上に統計的有意性を検証する機能が必要

実装

⇒　差の検定　［分布が同一とみなせるか］

⇒　相関解析　［無相関検定を含む相互相関計算］

⇒　周波数解析　［周波数領域での相関］

⇒　偏差解析　［特定周波数を差し引いたのち再度相関解析］

⇒　トレンド検定　［上昇(下降)傾向の有無を検定］

- 解析・検定パッケージを各種データに適用

⇒　現状問題なく処理が行われていることを確認