

# EOF 解析入門

高橋芳幸 (神戸大学大学院理学研究科)

# はじめに

- 「EOF 解析」は, 気象学・海洋学における主成分分析の呼称.
  - EOF = Empirical Orthogonal Function ; 経験的直交関数
- EOF 解析は, 気象学・海洋学において, 変動構造の抽出に用いられている.

# 大気の解析における例：北極振動

- 地球の北半球における冬季の高緯度大気の年々変動において卓越する構造
  - 「北極域と中緯度の圧力がシーズンのように振動」のようにはしは表現される。



北半球の冬季の海面更生気圧（を 1000 hPa 圧力面のジオポテンシャル高度の変動に変換したもの）における経験的直交関数 (EOF) 第一モード。  
(Thompson and Wallace, 1998)

# はじめに

- 今回は、[2021年12月10日の itpass セミナー](#)での主成分分析の解説を踏まえて EOF 解析の基礎について紹介し、試しに使ってみた結果を紹介する。
  - ~~地球の再解析データの EOF 解析~~
  - ~~地球のモデル計算結果の EOF 解析~~
  - 金星のモデル計算結果の EOF 解析
- 当初の目的は EOF 解析について勉強することだったが…
  - 惑星大気に対しても EOF 解析は適用されているようだが、まだ知られていない変動構造があるだろう。それらに繋がる何か、もしついでに見つかるなら嬉しい。

# 目次

- はじめに
- [復習] 主成分分析の説明
- EOF 解析の説明
- EOF 解析の実装
- EOF 解析例
  - 練習
  - DCPAM 金星実験結果を用いた大気変動構造の解析
- まとめ
  
- おまけ
  - EOF 解析例
    - NCEP/NCAR 再解析データを用いた北極振動の解析
    - DCPAM AMIP 実験結果を用いた北極振動の解析

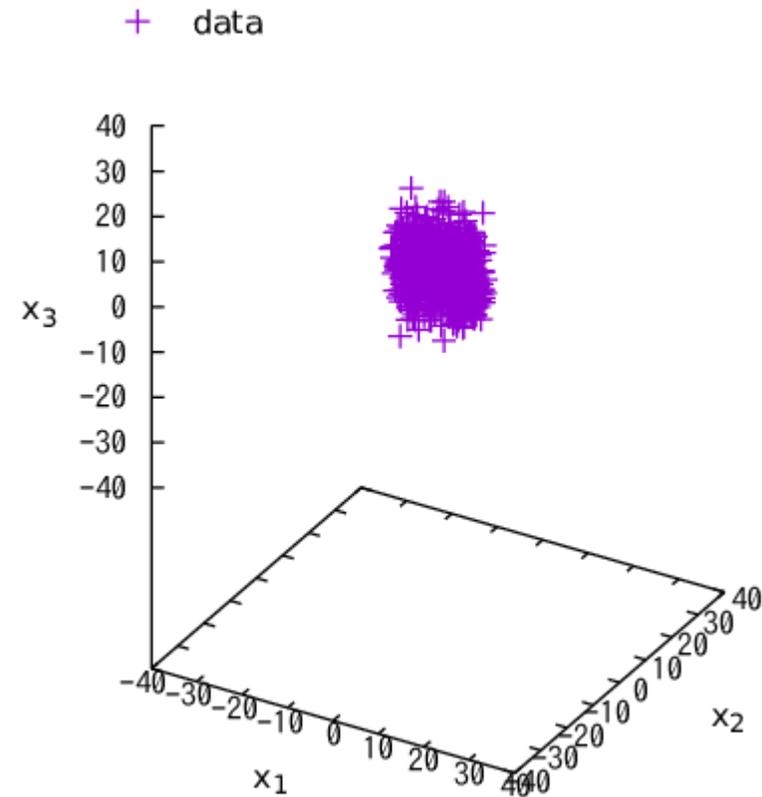
# [復習] 主成分分析の説明

# [復習] 主成分分析 – 定性的な説明

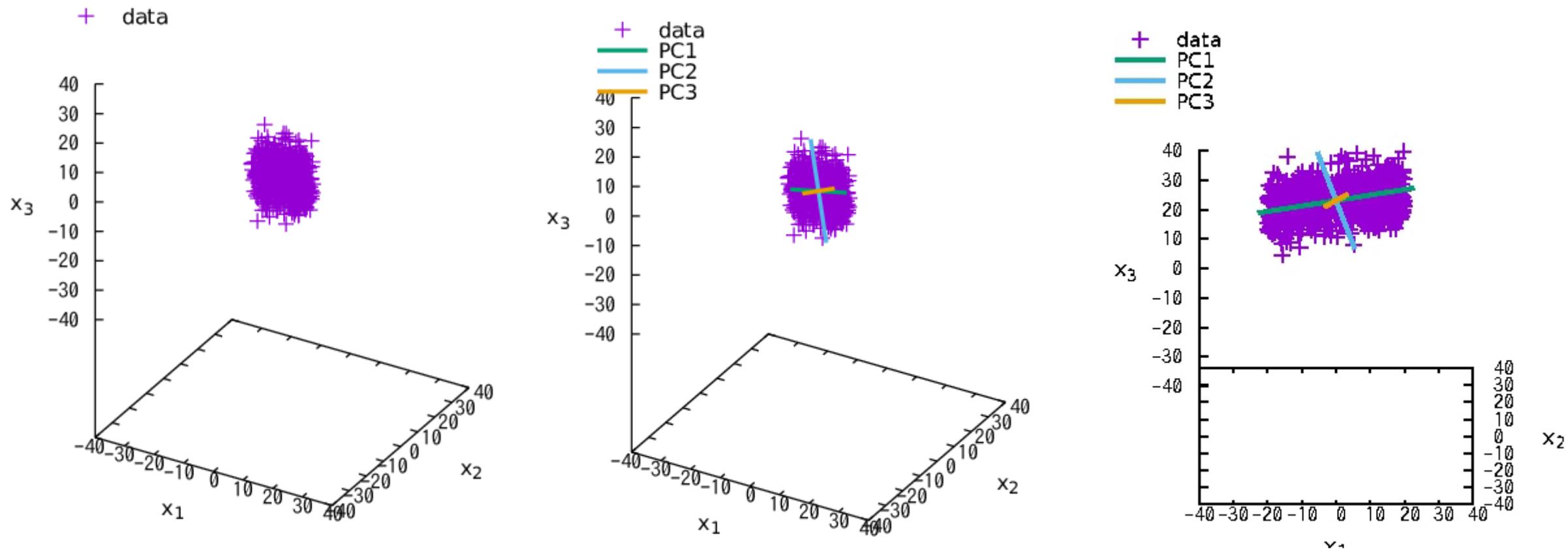
- 3 変数の値の組  $N$  個  
 $(x_{1,1}, x_{2,1}, x_{3,1}), (x_{1,2}, x_{2,2}, x_{3,2}),$   
 $\dots, (x_{1,N}, x_{2,N}, x_{3,N})$

を考える.

- ただし, 平均を引いてある.
- 例えば,
  - あるクラスの生徒の算数と理科と英語の点数
  - 圧力, 温度, 比湿
  - ...



# [復習] 主成分分析 - 定性的な説明



データがばらついている方向を探す.

# [復習] 主成分分析 解説 (3次元)

結局,  $N$  個のデータの組

$$(x'_{1,1}, x'_{1,2}, x'_{1,3}), (x'_{2,1}, x'_{2,2}, x'_{2,3}), \dots, (x'_{N,1}, x'_{N,2}, x'_{N,3})$$

のばらつく方向の単位ベクトル  $\mathbf{a} = (a_1, a_2, a_3)$  は,

$$\mathbf{V}\mathbf{a} = \lambda\mathbf{a}$$

$$\mathbf{V} = \begin{pmatrix} \sigma_{x'_1}^2 & C_{x'_1, x'_2} & C_{x'_1, x'_3} \\ C_{x'_1, x'_2} & \sigma_{x'_2}^2 & C_{x'_2, x'_3} \\ C_{x'_1, x'_3} & C_{x'_2, x'_3} & \sigma_{x'_3}^2 \end{pmatrix}$$

$$a_1^2 + a_2^2 + a_3^2 = 1$$

と書ける.  $\mathbf{V}$  は分散と共分散を要素とする行列 (分散共分散行列) で, 問題は  $\mathbf{V}$  の固有値問題に帰着する.

# [復習] 主成分分析 解説 (2次元)

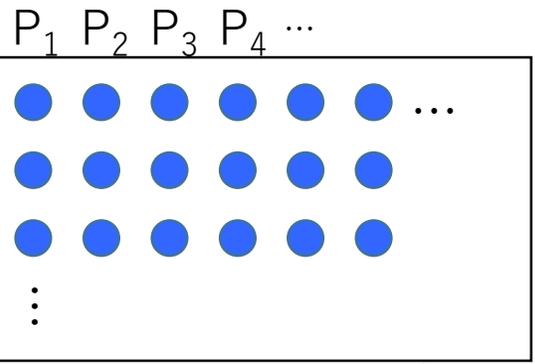
- 前回のセミナー (2021年12月10日 itpass セミナー) での2次元の説明は [こちら](#) (下の URL) を参照のこと.
  - [https://itpass.scitec.kobe-u.ac.jp/seminar/lecture/fy2021/211210/pub/slide\\_pca\\_2021-12-10.pdf](https://itpass.scitec.kobe-u.ac.jp/seminar/lecture/fy2021/211210/pub/slide_pca_2021-12-10.pdf)

# EOF 解析の説明

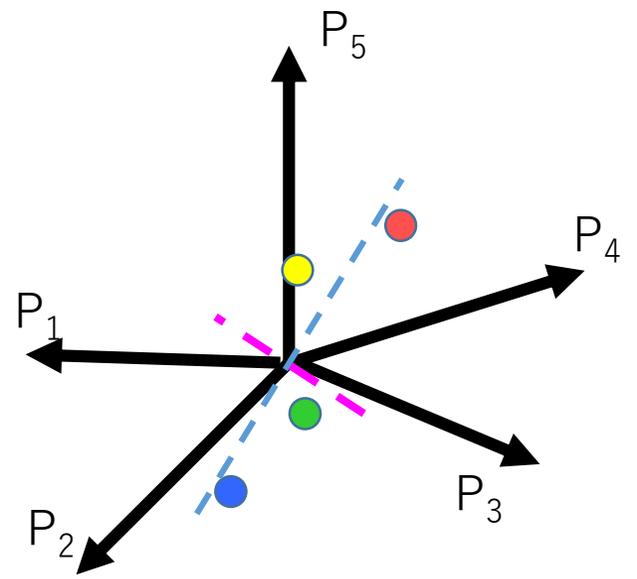
# EOF 解析 - 定性的な説明

物理量 A の  
時空間分布

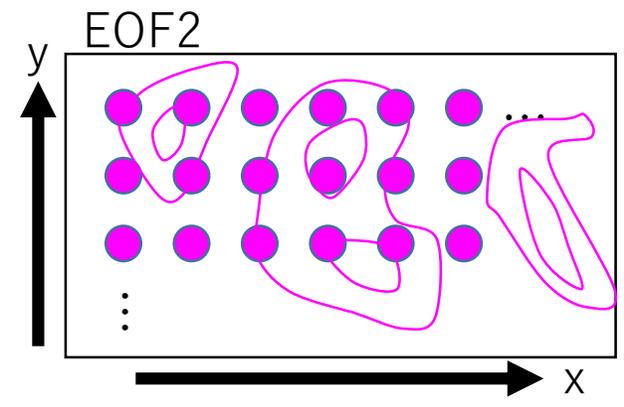
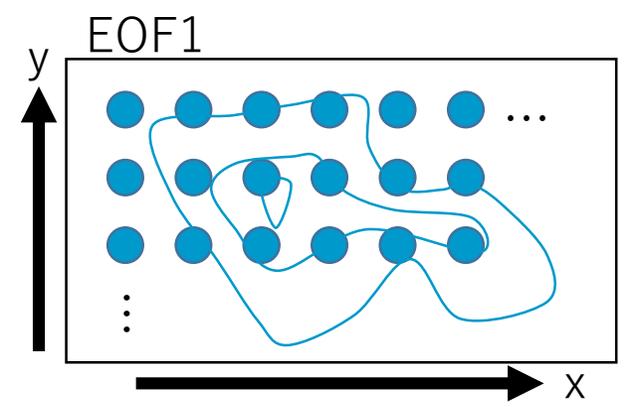
● は空間格子点 (M 点)



空間分布の時系列を  
各空間格子点の値を  
軸とする M 次元空間に  
プロットして  
ばらつく方向を求める  
(主成分分析)



ばらつく方向は、  
時間変動の大きな  
空間構造を表す



⋮

# EOF 解析概要

- $M$  個の格子点における空間データを  $M$  次元のデータと見なして主成分分析.
  - データ例
    - 海面更生気圧
      - 空間格子点  $M$  個 (空間 1 次元でも 2 次元でも ... 良い)
      - 時間格子点  $N$  個
- 結果として,  $M \times M$  の時間に対する分散共分散行列の固有値問題を解くことになる.
- 補足
  - 各格子点に重みがある場合には重みをかけて解析する.
    - 以下の説明では重みについては触れていないことに注意.
  - 時間方向のデータを含めて考える場合もあるようである.

# 主成分分析 解説 (M 次元)

$M$  個の空間格子点上の値を一次元に並べたデータの組  $N$  個を考える (時間方向のデータ数が  $N$  個).

$$(x'_{1,1}, x'_{1,2}, \dots, x'_{1,M}), (x'_{2,1}, x'_{2,2}, \dots, x'_{2,M}), \dots, (x'_{N,1}, x'_{N,2}, \dots, x'_{N,M})$$

つまり,

$$\mathbf{x}'_n = (x'_{n,1}, x'_{n,2}, \dots, x'_{n,M}) \quad (n = 1, 2, \dots, N)$$

を考える.

( $x'_{n,m}$  は  $n$  番目の時刻における  $m$  番目の空間点の値. ただし平均を引いてある.)

# 主成分分析 解説 (M 次元)

$x'_n$  の, 単位ベクトル  $a$  への射影の分散が最も大きくなるとき,  $a$  は

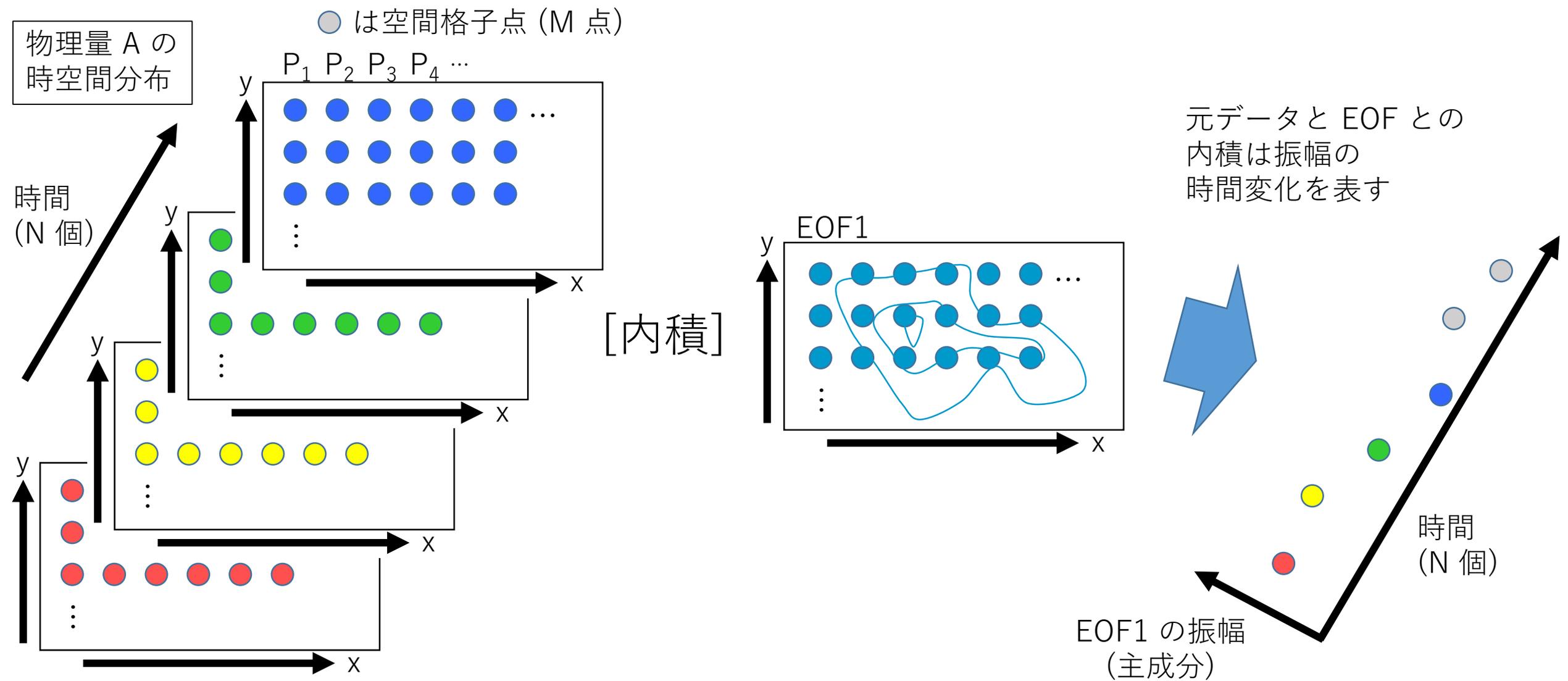
$$V a = \lambda a$$

$$V = \begin{pmatrix} \sigma_{x'_1}^2 & C_{x'_1, x'_2} & \cdots & C_{x'_1, x'_M} \\ C_{x'_2, x'_1} & \sigma_{x'_2}^2 & \cdots & C_{x'_2, x'_M} \\ \vdots & \vdots & \ddots & \vdots \\ C_{x'_M, x'_1} & C_{x'_M, x'_2} & \cdots & \sigma_{x'_M}^2 \end{pmatrix}$$

$$a_1^2 + a_2^2 + \cdots + a_M^2 = 1$$

を満たす. ただし,  $V$  は分散と共分散を要素とする行列 (分散共分散行列) で,  $\sigma_{x'_i}$  は  $x'_i$  の分散,  $C_{x'_i, x'_j} = C_{x'_j, x'_i}$  は  $x'_i, x'_j$  の共分散である ( $V$  は対称行列). つまり, 問題は  $V$  の固有値問題に帰着する.

# EOF 解析 - 定性的な説明



# 主成分分析 解説 (M 次元)

固有問題を解いて求めた  $k$  番目の固有ベクトル  $\mathbf{a}_k = (a_{k,1}, a_{k,2}, \dots, a_{k,M})$  への元の偏差

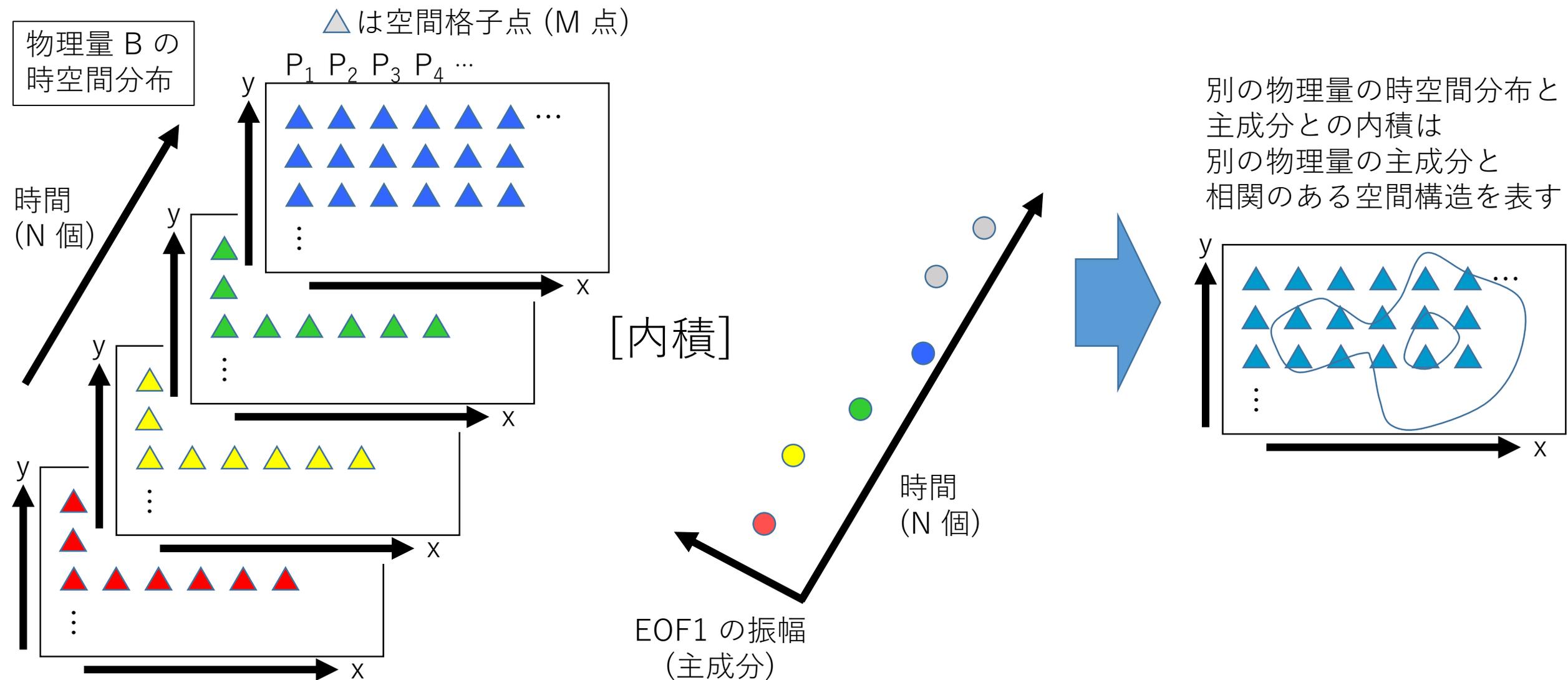
$$\begin{aligned} \mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_N \\ = (x'_{1,1}, x'_{1,2}, \dots, x'_{1,M}), (x'_{2,1}, x'_{2,2}, \dots, x'_{2,M}), \dots, (x'_{N,1}, x'_{N,2}, \dots, x'_{N,M}) \end{aligned}$$

( $x'_{n,m}$  は  $n$  番目の時刻における  $m$  番目の空間格子点の偏差) の射影

$$z_{n,k} = \mathbf{x}'_n \cdot \mathbf{a}_k = \sum_{m=1}^M x'_{n,m} a_{k,m}$$

は、その固有ベクトルが表す構造の振幅に対応し、主成分 (Principal Component; PC) と呼ばれる ( $z_{n,k}$  は  $n$  番目の時刻の  $k$  番目の EOF の振幅).

# EOF 解析 - 定性的な説明



# 主成分分析 解説 (M 次元)

EOF を求めた物理量とは別の物理量の偏差

$$\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_N \\ = (y'_{1,1}, y'_{1,2}, \dots, y'_{1,M}), (y'_{2,1}, y'_{2,2}, \dots, y'_{2,M}), \dots, (y'_{N,1}, y'_{N,2}, \dots, y'_{N,M})$$

( $y'_{n,m}$  は  $n$  番目の時刻における  $m$  番目の空間格子点の偏差) の主成分 (PC) への射影

$$\mathbf{y}_k^* = \sum_{n=1}^N \mathbf{y}'_n z_{n,k}, \quad y_{m,k}^* = \sum_{n=1}^N y'_{n,m} z_{n,k}$$

は, 物理量,  $k$  番目の EOF と相関のある  $y'$  の構造を表す. ( $y_{m,k}^*$  は  $k$  番目の EOF と相関のある  $m$  番目の空間格子点の偏差.)

# EOF 解析で得られるもの

- EOF 解析からは, 主に時間変化する定在性の空間構造が得られる
  - 定常な構造は得られない
  - 伝播性の構造も得られるが, 等しい振幅を持つ二つの構造として検出される
  - 得られる構造は互いに直交
- 他の物理量の EOF の主成分への回帰により, 相関のある他の物理量の変動構造も得られる
- ただし, 構造が得られてもそれを生じる物理機構が EOF 解析でわかるわけではない.

# EOF 解析の亜種

- EOF 解析には, 方法を発展させた/利用した亜種がある
  - Rotated EOF (REOF)
  - Extended EOF (EEOF)
  - Complex/Hilbert EOF
  - Frequency Domain EOF (FDEOF)
- 今回はそれら亜種については触れない.
  - 興味があれば, Hannachi et al. (2007) を参照すると良いかもしれない.
    - Hannachi, A., Jolliffe, I.T. and Stephenson, D.B. (2007), Empirical orthogonal functions and related techniques in atmospheric science: A review. Int. J. Climatol., 27: 1119-1152. <https://doi.org/10.1002/joc.1499>

# EOF 解析の実装

# EOF 解析 実装

- 今回は, 分散共分散行列の固有値問題を LAPACK の DSYEV サブルーチン (倍精度実対象行列の固有値, 固有ベクトルを計算) を用いて解く.
  - LAPACK - Linear Algebra PACKage
    - 数値線形代数計算のための Fortran ライブラリ
    - debian ならば liblapack-dev パッケージで利用可能
  - コンパイル例

```
$ gfortran -I /usr/lib/x86_64-linux-gnu/lapack ¥  
-L /usr/lib/x86_64-linux-gnu/lapack ¥  
-o pca pca.f90 -llapack
```
- なお, 分散共分散行列の固有値問題を解く代わりに, データを収めた行列 (サンプル数 × 変数の数) を特異値分解して特異値と特異ベクトルを求めることも等価である.
- 補足
  - 「主成分分析」「ワイン」などの単語を使って検索すると, Python や R を使った分析の解説をたくさん見つけることができる.

EOF 解析 — 確認

# EOF 解析 — 確認

- 時空間構造 (空間 2 次元) を用意し, その EOF を求めることで EOF 解析から得られるものを確認する.
  - 二つの例を紹介

EOF 解析 — 確認 1

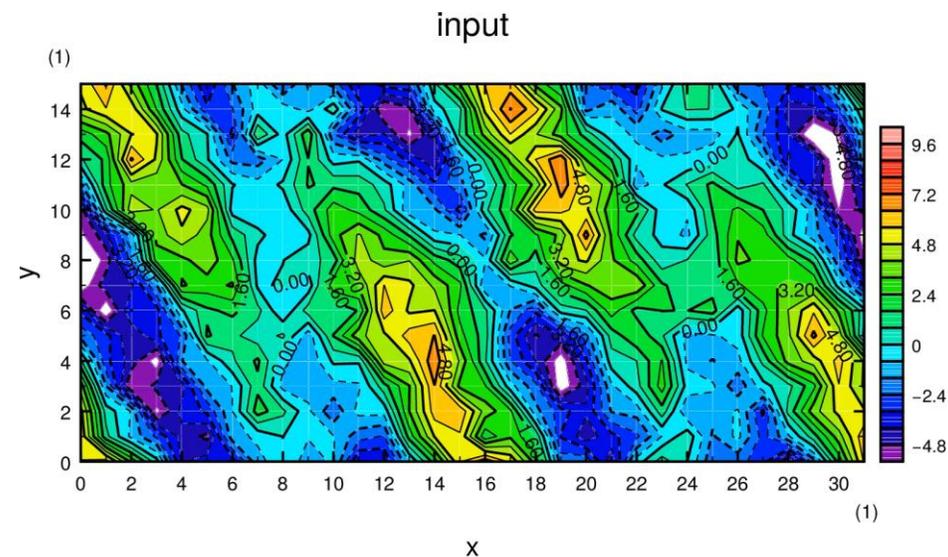
# 与えるデータ

- 擾乱構造

$$\begin{aligned}\phi(x, y, t) = & 3 \cos \left( 2 \frac{2\pi}{L_x} x + \frac{2\pi}{L_y} y \right) \cos \left( 2 \frac{2\pi}{T} t \right) \\ & + 3 \cos \left( 4 \frac{2\pi}{L_y} x + \frac{2\pi}{L_y} y \right) \cos \left( 5 \frac{2\pi}{T} t \right) \\ & + 6 \exp \left\{ - \left( \frac{x - \frac{L_x}{2}}{\frac{L_x}{9}} \right)^2 - \left( \frac{y - \frac{L_y}{2}}{\frac{L_y}{5}} \right)^2 \right\} \cos \left( 3 \frac{2\pi}{T} t \right) \\ & + \varepsilon(x, t)\end{aligned}$$

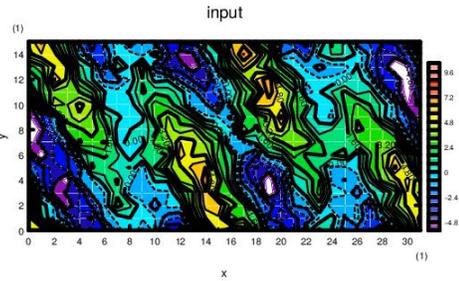
- 格子点数

- x 方向 :  $L_x$  32 点
- y 方向 :  $L_y$  16 点

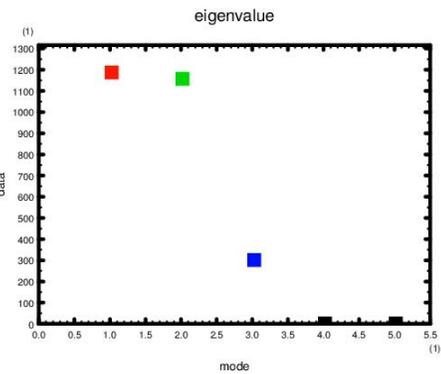


# EOF 解析結果

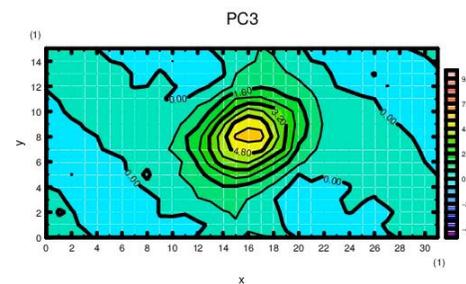
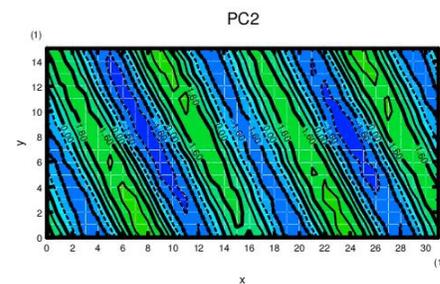
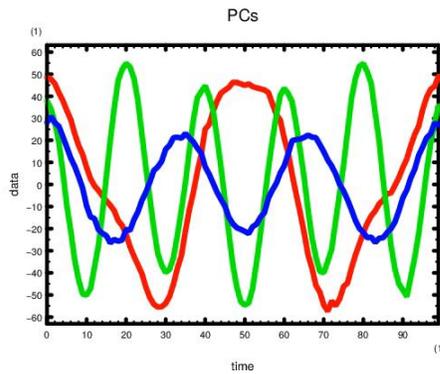
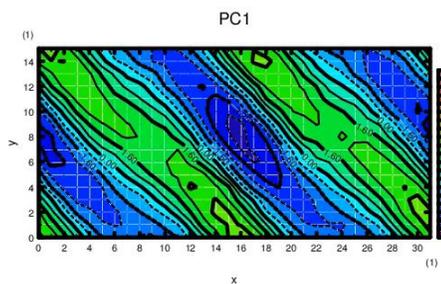
入力



EOF 振幅



EOF 構造



三つの擾乱が検出  
しかし, 与えた項が綺麗には分離して検出  
されない

$$\begin{aligned} \phi(x, y, t) = & 3 \cos\left(2\frac{2\pi}{L_x}x + \frac{2\pi}{L_y}y\right) \cos\left(2\frac{2\pi}{T}t\right) \\ & + 3 \cos\left(4\frac{2\pi}{L_y}x + \frac{2\pi}{L_y}y\right) \cos\left(5\frac{2\pi}{T}t\right) \\ & + 6 \exp\left\{-\left(\frac{x - \frac{L_x}{2}}{\frac{L_x}{9}}\right)^2 - \left(\frac{y - \frac{L_y}{2}}{\frac{L_y}{5}}\right)^2\right\} \cos\left(3\frac{2\pi}{T}t\right) \\ & + \varepsilon(x, t) \end{aligned}$$

EOF 解析—確認 2

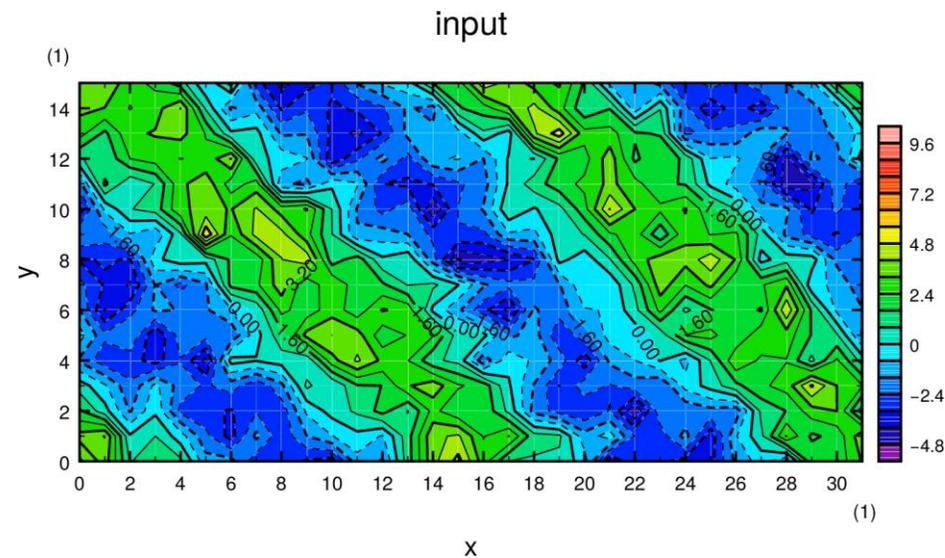
# 与えるデータ

- 擾乱構造

$$\phi(x, y, t) = 3 \cos \left( 2 \frac{2\pi}{L_x} x + \frac{2\pi}{L_y} y - 5 \frac{2\pi}{T} t \right) + \varepsilon(x, t)$$

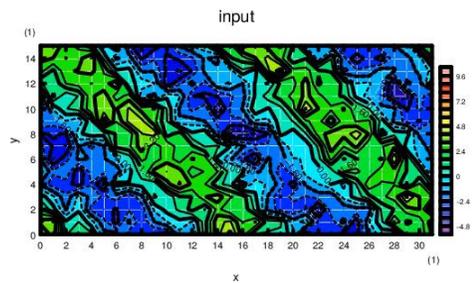
- 格子点数

- x 方向 :  $L_x$  32 点
- y 方向 :  $L_y$  16 点

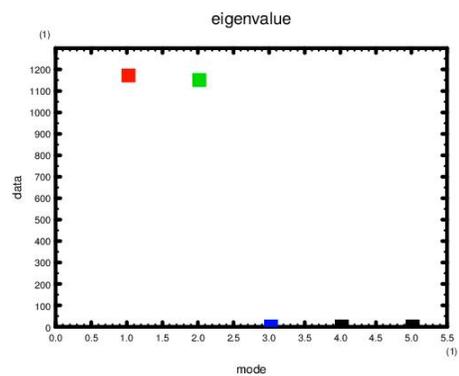


# EOF 解析結果

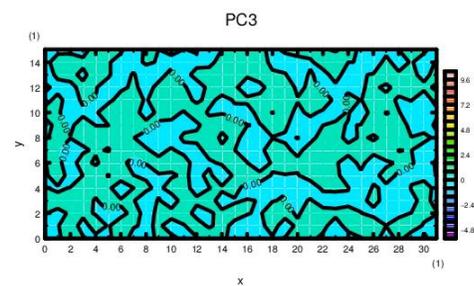
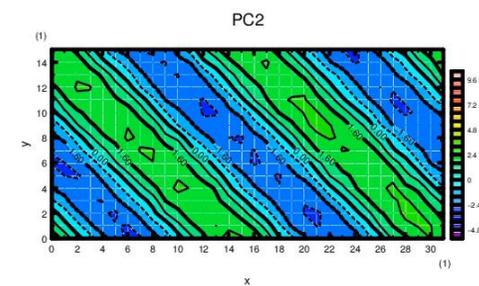
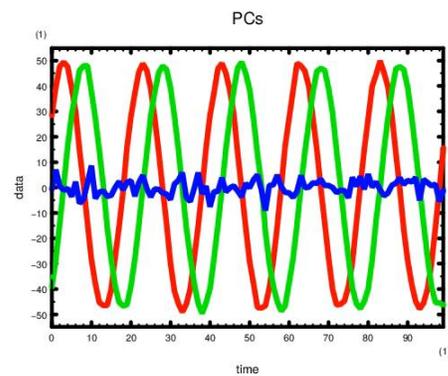
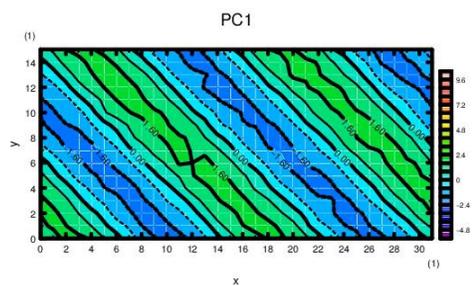
入力



EOF 振幅



EOF 構造

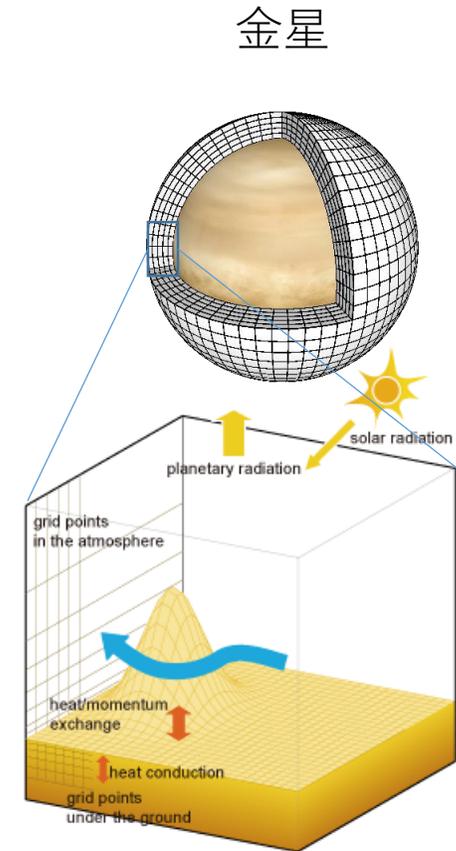


二つの定在性擾乱として検出

# 金星大気の EOF 解析 DCPAM 金星大気実験

# DCPAM 金星大気実験 — モデル

- DCPAM
  - 地球流体電脳倶楽部有志で開発している惑星大気大循環モデル
- DCPAM 開発版
  - 力学過程
    - プリミティブ方程式
      - 水平スペクトル変換法
      - 鉛直差分法 (Arakawa and Suarez, 1983)
  - 使用物理過程
    - 金星大気放射モデル
      - 大気気体組成分布, 雲分布固定
    - Mellor and Yamada (1982) level 2.5
    - 乾燥対流調節
    - 惑星表面・土壌熱拡散モデル



← 本当は雲を描きたいが描けていない

# DCPAM —モデル

## 力学過程の方程式系

方程式系は, 静水圧平衡を仮定し, 鉛直方向に  $\sigma = p/p_s$  座標を用いたプリミティブ方程式系.

運動方程式 
$$\frac{\partial \vec{u}}{\partial t} = -(\vec{u} \cdot \nabla) \vec{u} - \dot{\sigma} \frac{\partial \vec{u}}{\partial \sigma} - f \vec{k} \times \vec{u} - \nabla \Phi + \frac{RT}{p_s} \nabla p_s + \vec{F}$$

熱力学の式 
$$\frac{\partial \Phi}{\partial \sigma} = -\frac{RT}{\sigma}$$

$$\frac{\partial T}{\partial t} = -(\vec{u} \cdot \nabla) T - \dot{\sigma} \frac{\partial T}{\partial \sigma} + \frac{\kappa T}{p} \frac{dp}{dt} + Q$$

連続の式 
$$\frac{\partial p_s}{\partial t} = -\nabla \cdot (p_s \vec{u}) - \frac{\partial}{\partial \sigma} (p_s \dot{\sigma})$$

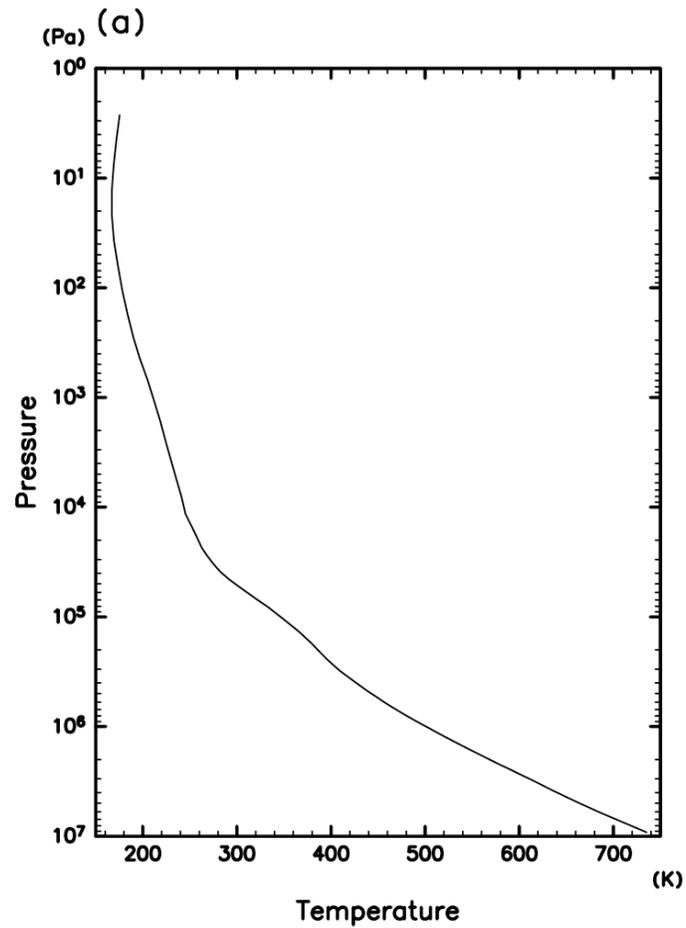
$$\frac{\partial q}{\partial t} = -(\vec{u} \cdot \nabla) q - \dot{\sigma} \frac{\partial q}{\partial \sigma} + S_q$$

境界条件:

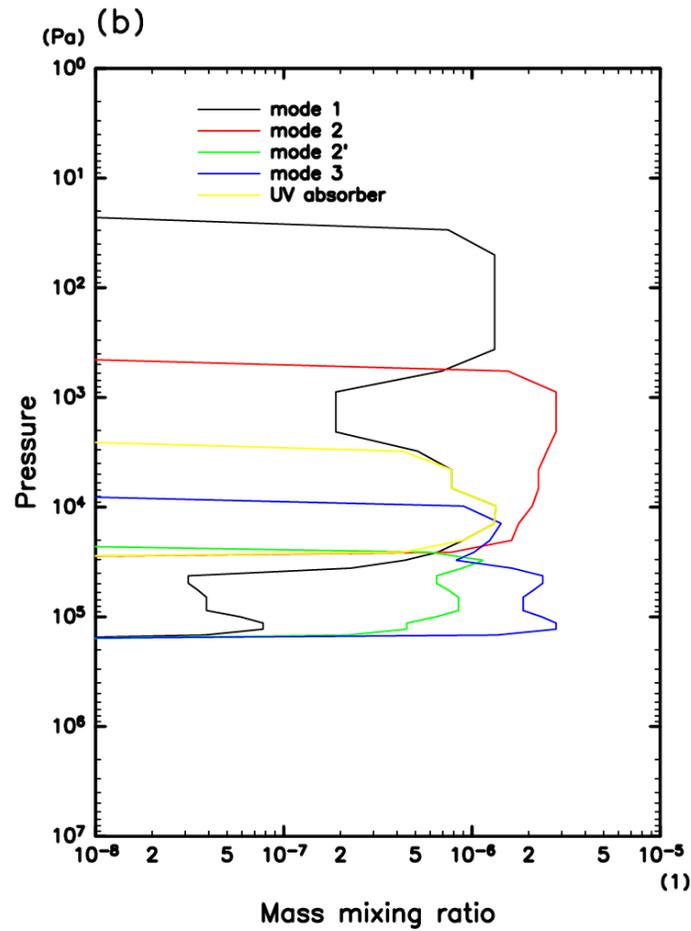
$$\dot{\sigma} = \frac{d\sigma}{dt} = 0 \quad \text{at } \sigma = 0, 1$$

$\vec{u}$  は水平風速,  $p$  は気圧,  $p_s$  は惑星表面気圧,  $T$  は温度,  $q$  は物質の混合比,  $\Phi$  はジオポテンシャル,  $\kappa = R/C_p$ ,  $R$  は気体定数,  $C_p$  は定圧比熱,  $\vec{F}$ ,  $Q$ ,  $S_q$  はそれぞれ摩擦, 非断熱加熱, 物質の生成・消滅.

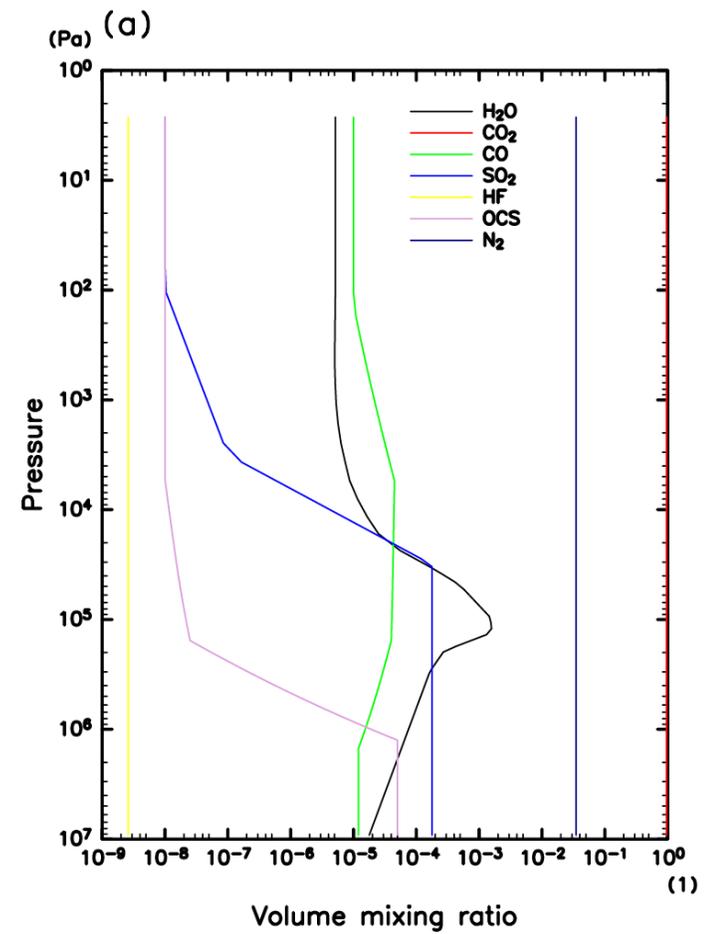
初期条件温度分布



放射計算に用いている  
雲分布



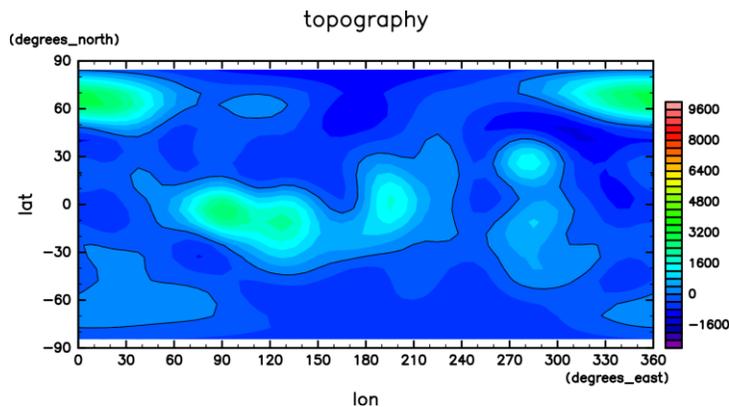
放射計算に用いている  
気体組成分布



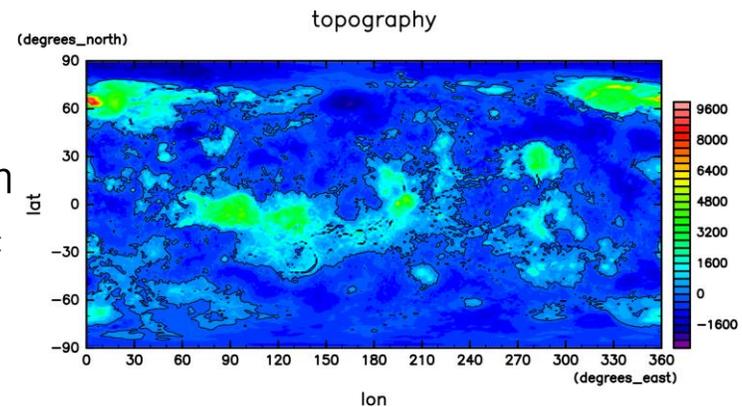
# DCPAM 金星大気実験 — 実験条件

- 惑星パラメータ : 金星
  - 自転角速度 :  $-2.993 \times 10^{-7} \text{ s}^{-1}$  (周期: 242.97 地球日)
  - 太陽日 : 117 地球日
- 地形 : Magellan 観測結果を平滑化
  - モデル解像度に合わせて平滑化 & 球面調和関数で全波数 15 で切断
- 解像度 : T15L52
- 初期条件 : 静止, 水平一様な VIRA の鉛直分布を持つ温度
- 積分期間 : 200000 地球日

実験に用いた  
地形

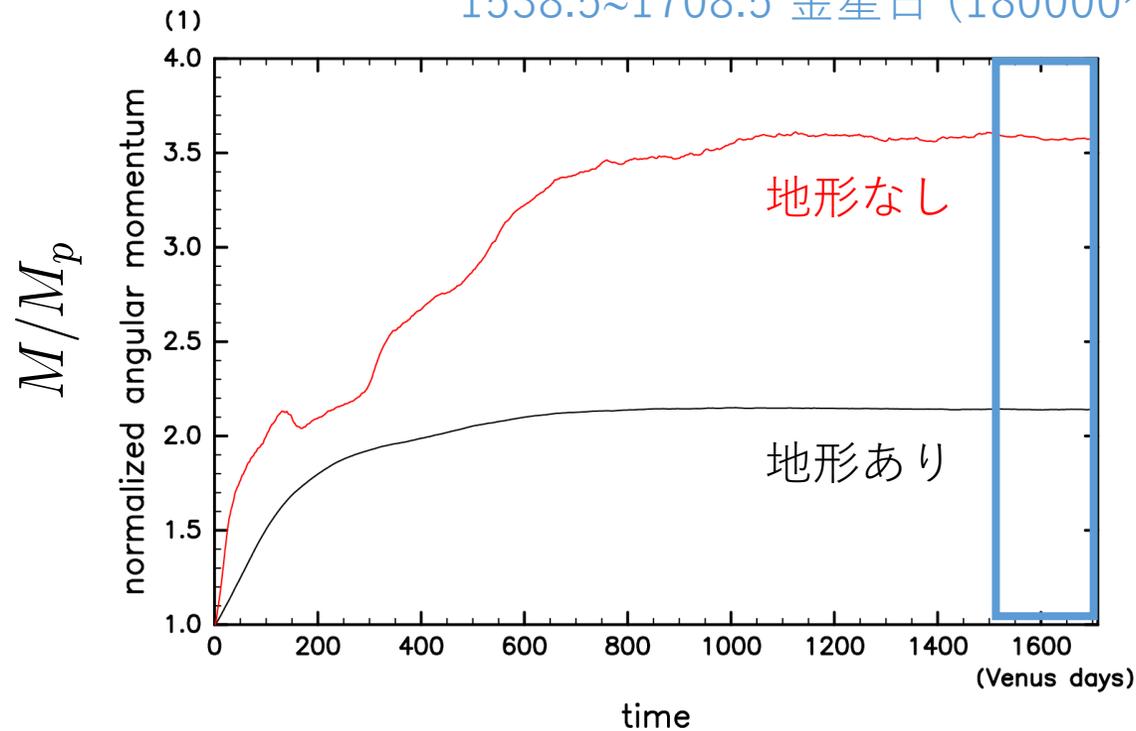


Magellan  
観測結果



# スピニアップの様子

1538.5~1708.5 金星日 (180000~199890 地球日) の結果を解析

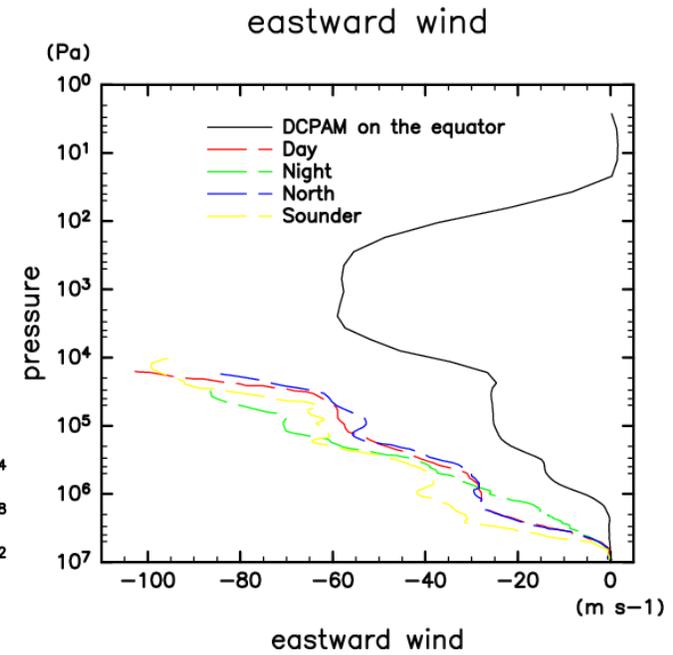
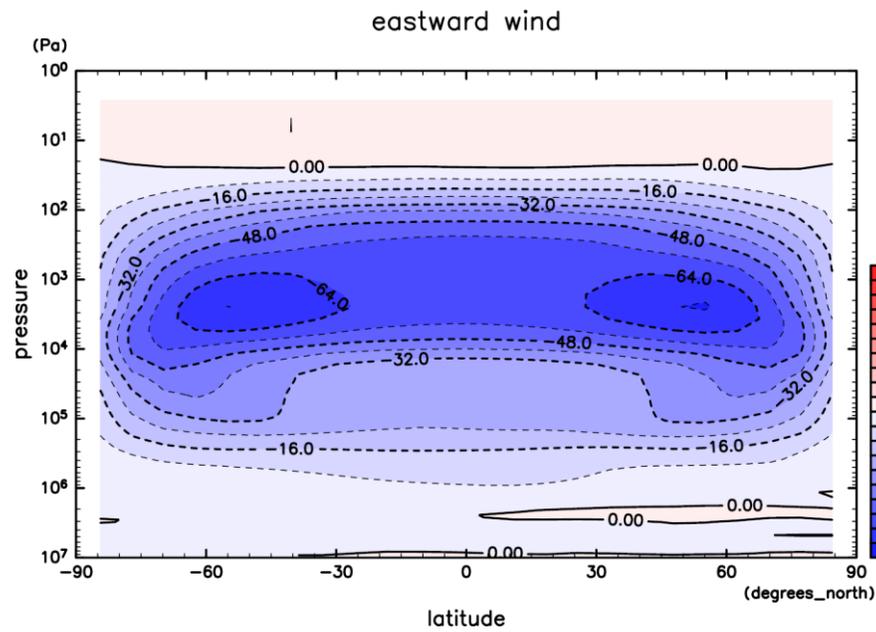
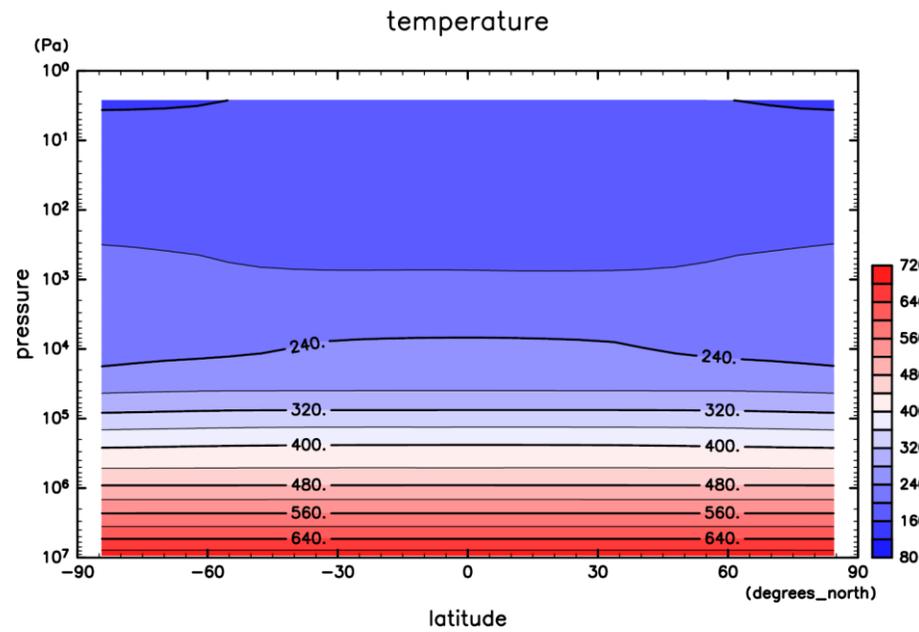


$$M = M_r + M_p$$

$$M_r = \frac{1}{4\pi} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{z_s}^{\infty} \rho u a \cos \phi \cos \phi dz d\phi d\lambda,$$

$$M_p = \frac{1}{4\pi} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{z_s}^{\infty} \rho \Omega a^2 \cos^2 \phi \cos \phi dz d\phi d\lambda$$

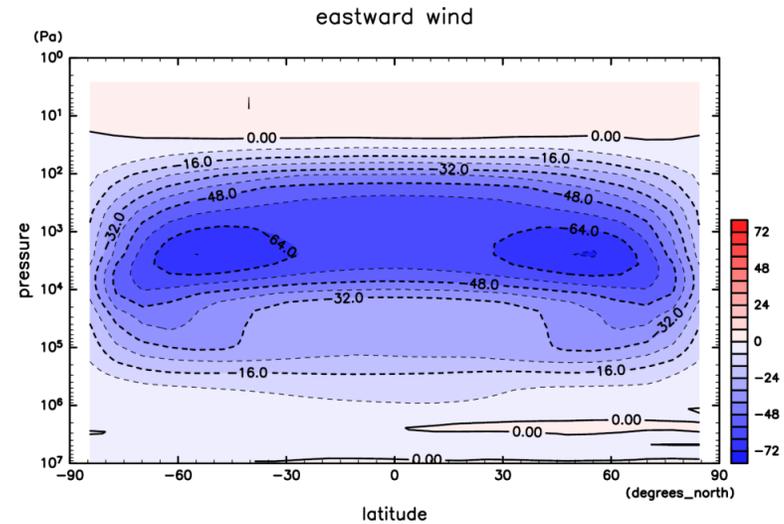
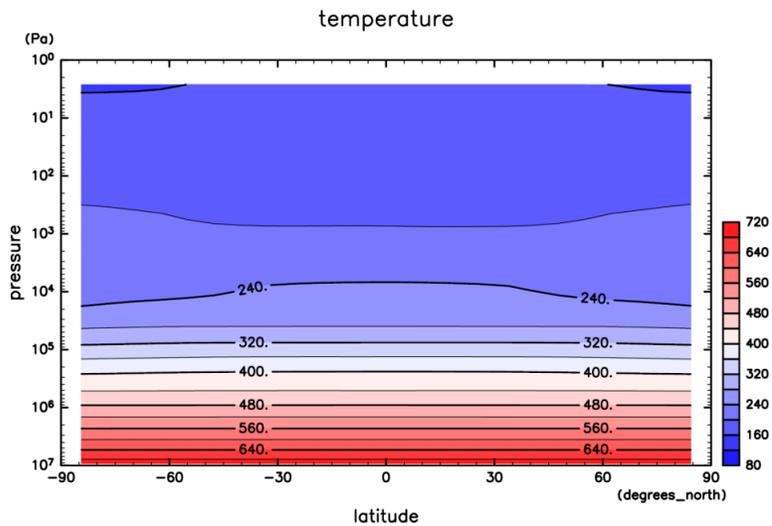
# 東西平均場



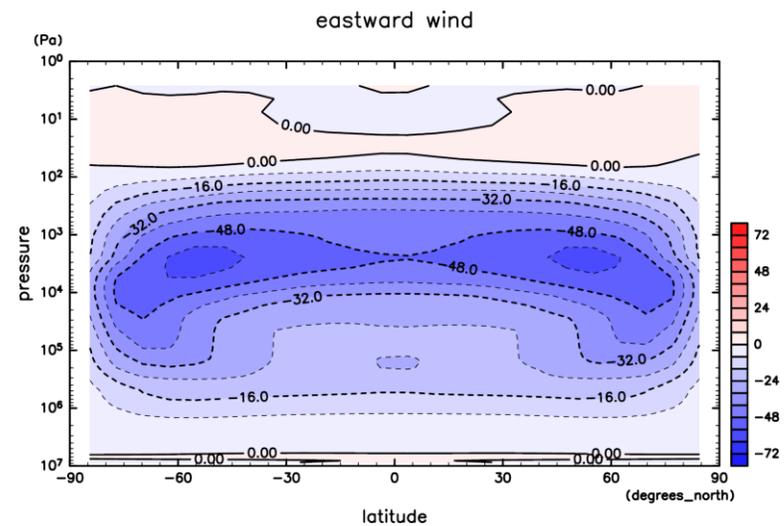
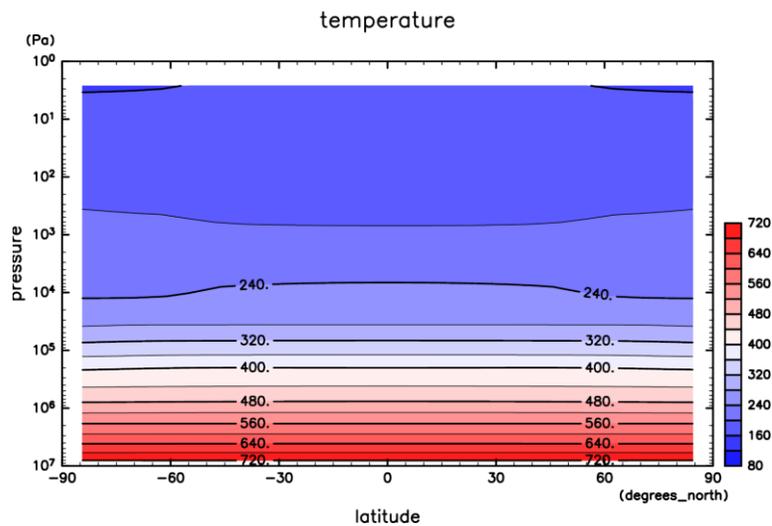
- 温度はほぼ水平一様
- 東西風は実際よりも弱い (~1/2)

# 東西平均場 地形の有無の比較

地形あり



地形なし



# DCPAM 金星大気実験結果の EOF 解析

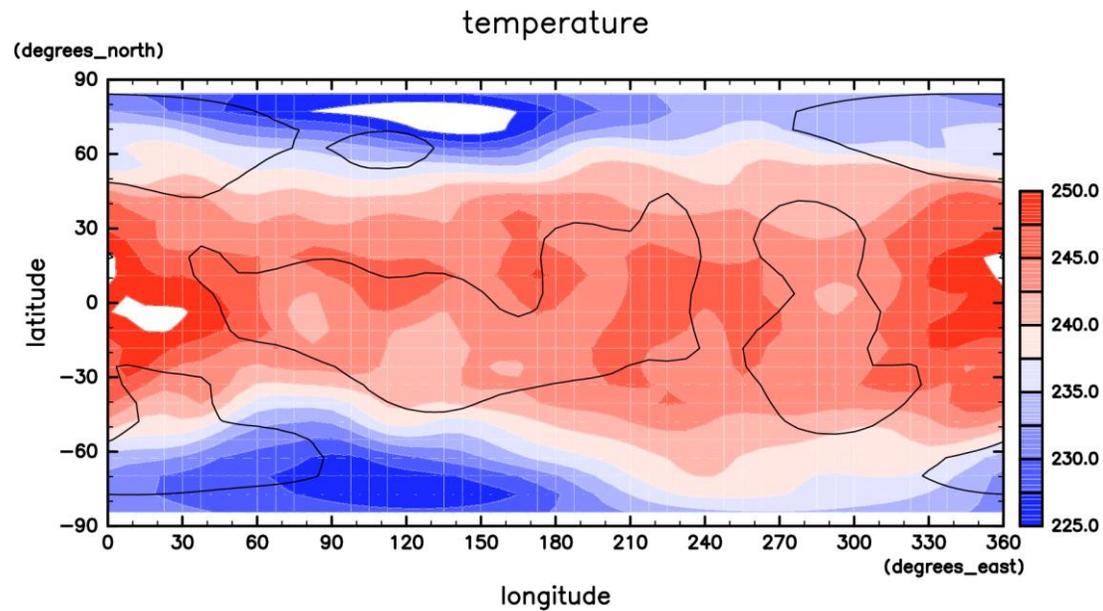
- 解析方針

- データ期間：19890 地球日 ( $85 \times 2$  金星日 =  $85 \times (117 \times 2)$  地球日)
- 物理量
  - 0.1 bar 圧力面の温度
    - 2 金星日 ( $117$  地球日  $\times 2$ ) 間の平均値
    - 0.1 bar 圧力面はおよそ雲層上端に位置し, およそあかつき LIR が温度を観測する高度
- 地理的範囲：全球

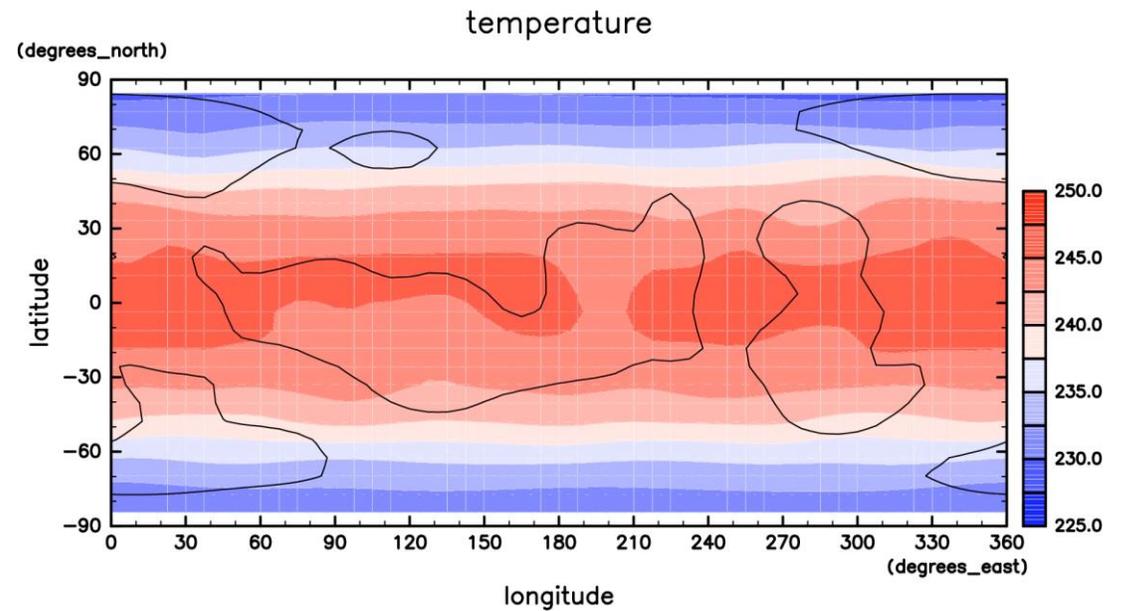
# 0.1 bar の温度の時間変化

瞬間値

(1 地球日毎, 動画期間 1000 地球日間 (8.5 金星日間))

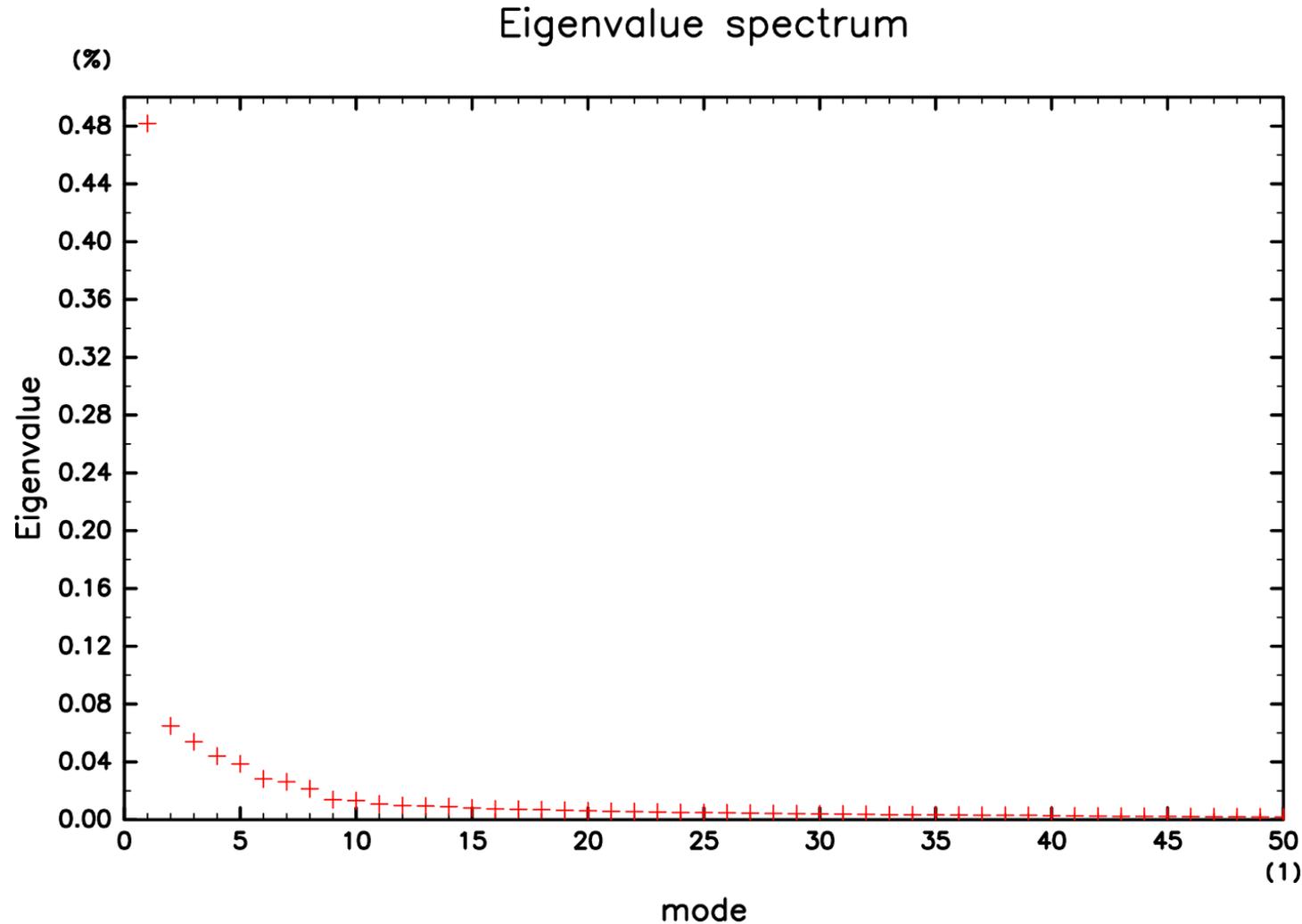


2 金星日平均値 (117 × 2 地球日平均値)



(おそらく高緯度の青い領域が一周するのが 1 金星日)

# 固有値スペクトル

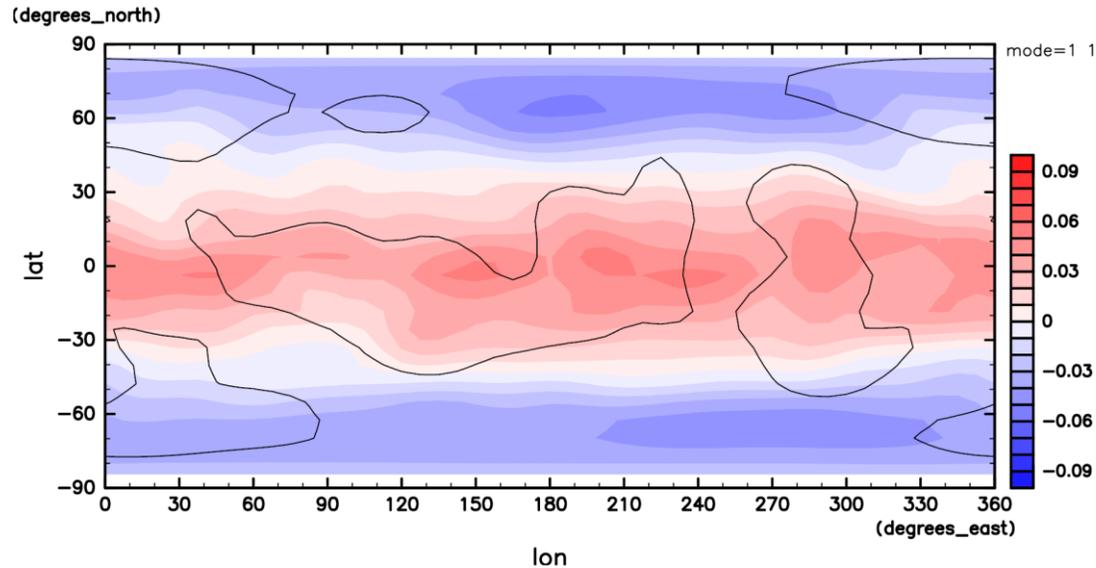


EOF1 が卓越.

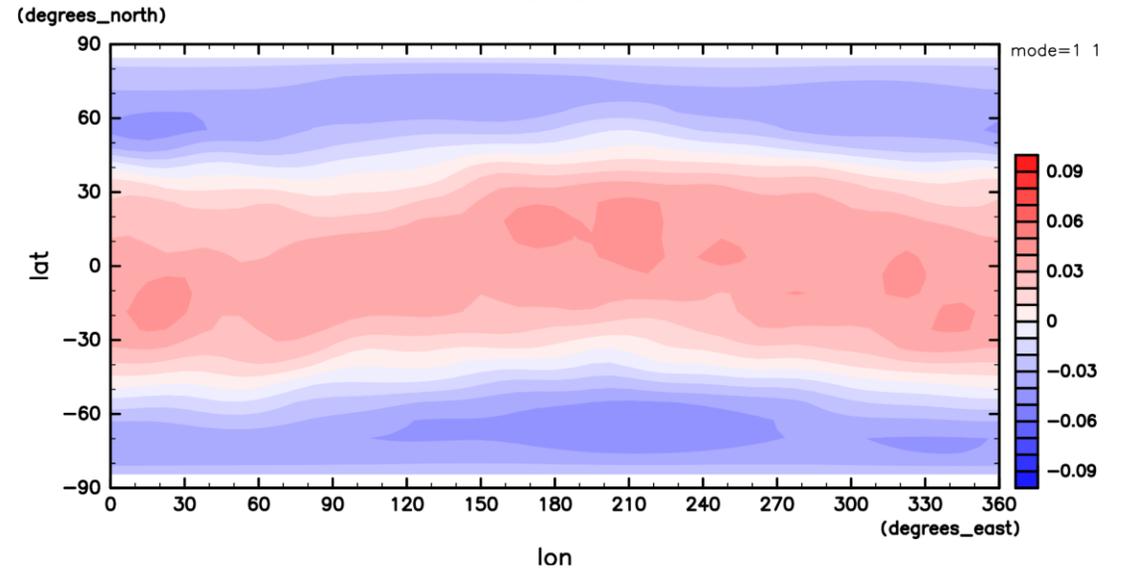
本来は信頼性を  
評価する必要あり.  
(しかし今回は EOF1  
が卓越することに  
変わりはないだろう.)

# EOF1 の構造

地形あり  
data



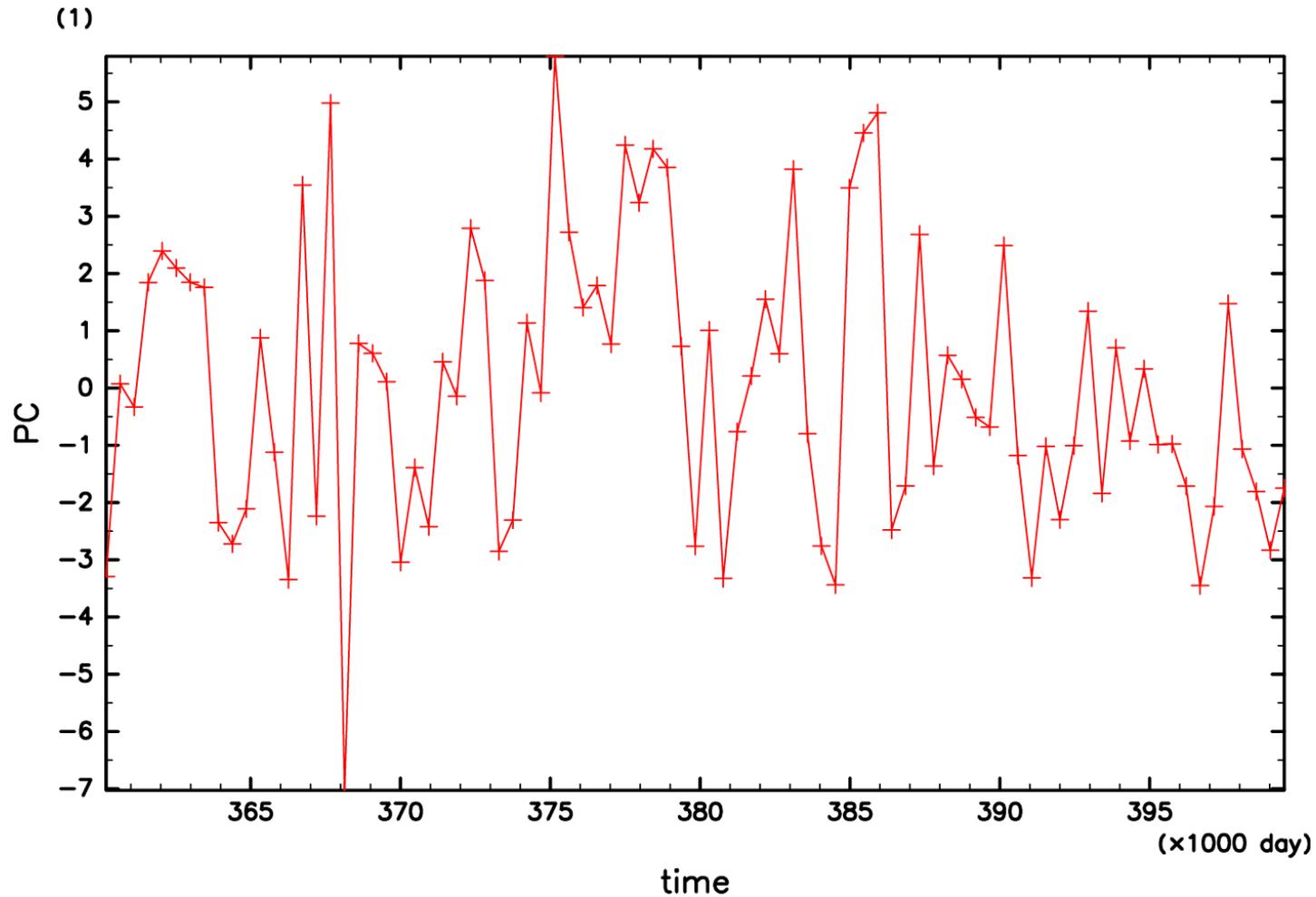
地形なし  
data



- ほぼ東西一様
  - 経度依存する構造が全くないわけではないが有意かどうかわからない

# EOF1 主成分

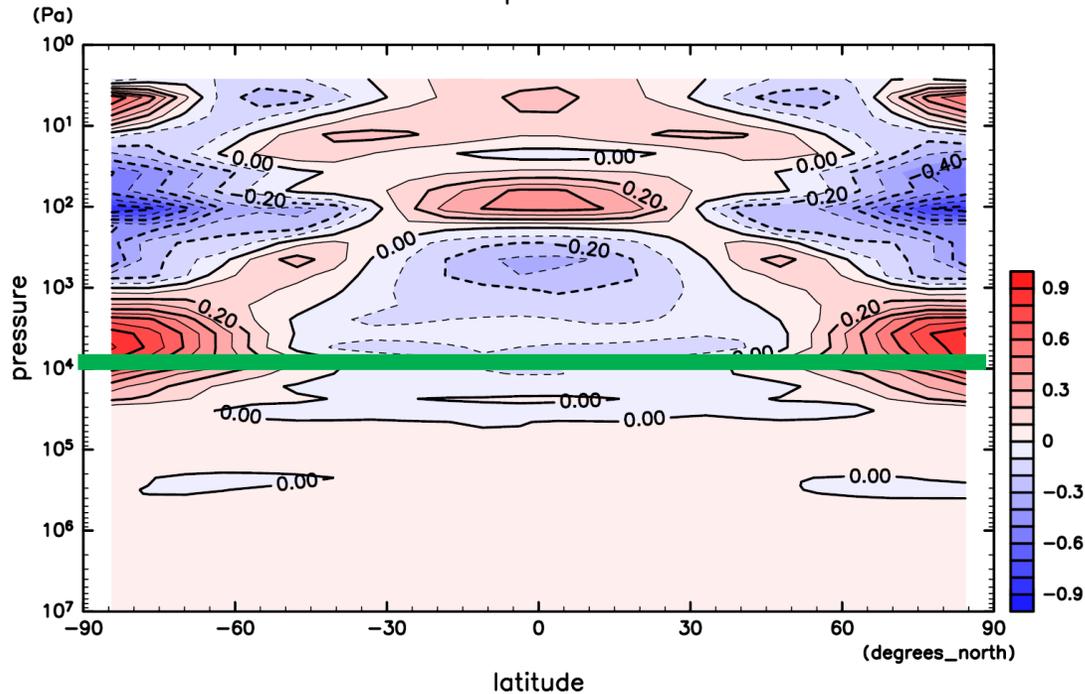
EOF1



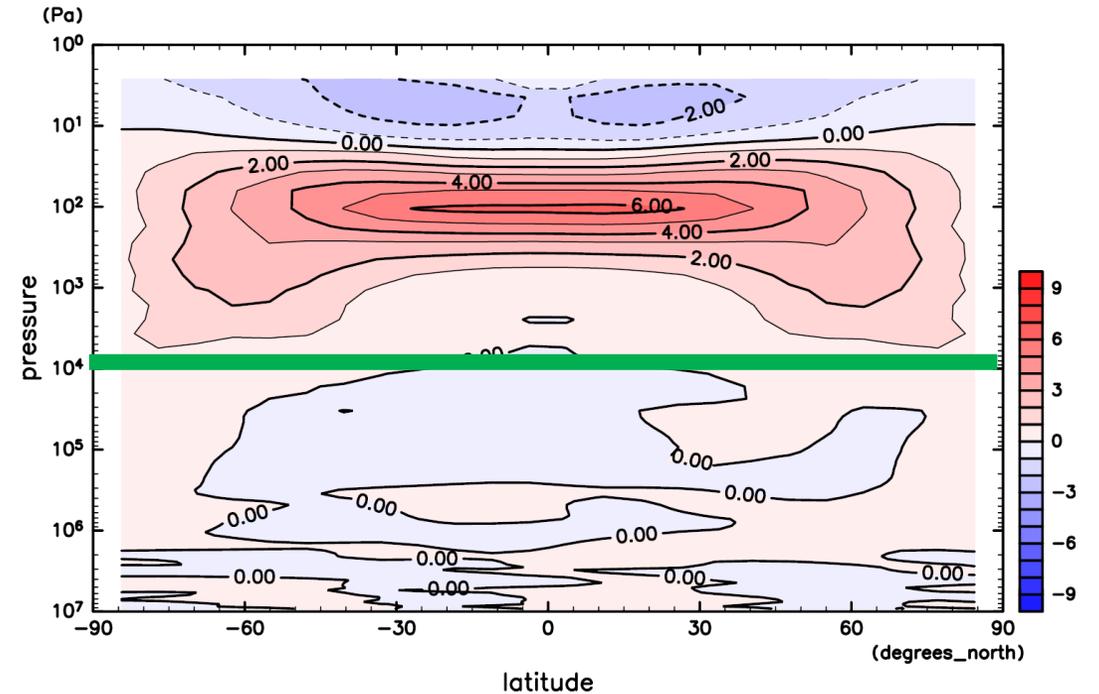
数千地球日（数十金星日）  
程度の時間スケールで  
変動？

# 0.1 bar 温度 EOF1 主成分への射影成分の 東西平均の子午面分布

温度  
temperature



東西風  
eastward wind



- 0.01 bar 付近では位相がほぼ反転
- 下層にはほぼ影響しない

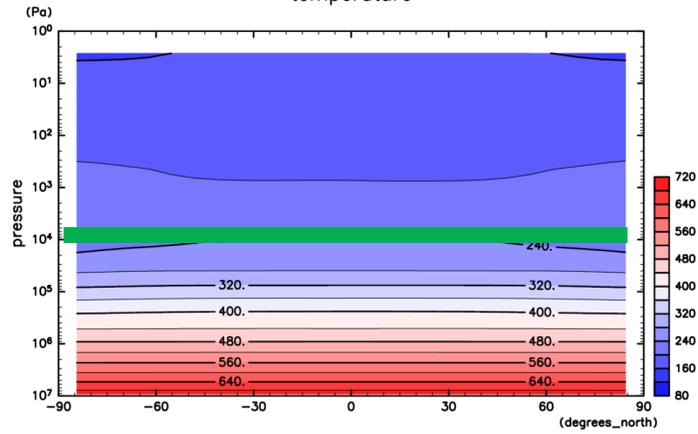
- 0.001 bar 付近の東西風(スーパーローテーションの上部)と関係

# 何が起こっているのか

東西平均

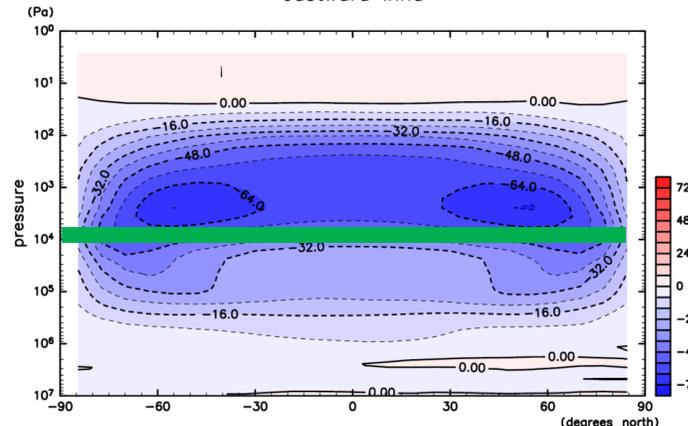
温度

temperature



東西風

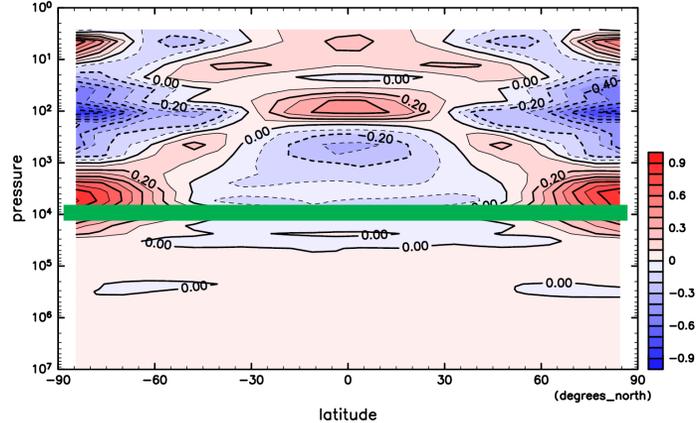
eastward wind



EOF1

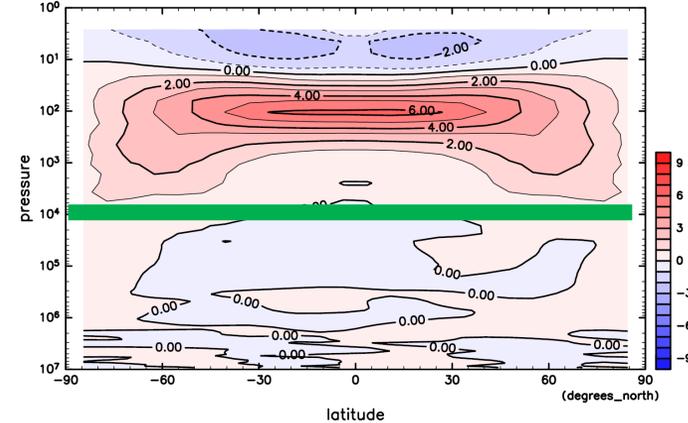
温度

temperature



東西風

eastward wind



- EOF1 の温度と東西風の関係 (の向き) は, 温度風と整合的.

- 温度変動と東西風変動のどちらに起因?

- 上層低緯度の東西風変動に起因?

- 例えば雲層での熱潮汐波励起が時間変動すれば説明できる?

- 計算では熱潮汐波の振幅が(大きく)時間変化するとは考えにくい.

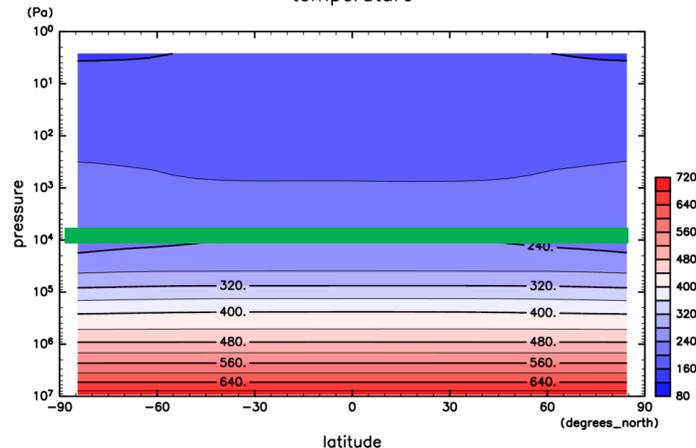
- 温度の変動に起因?

- 次ページへ.

# 何が起こっているのか

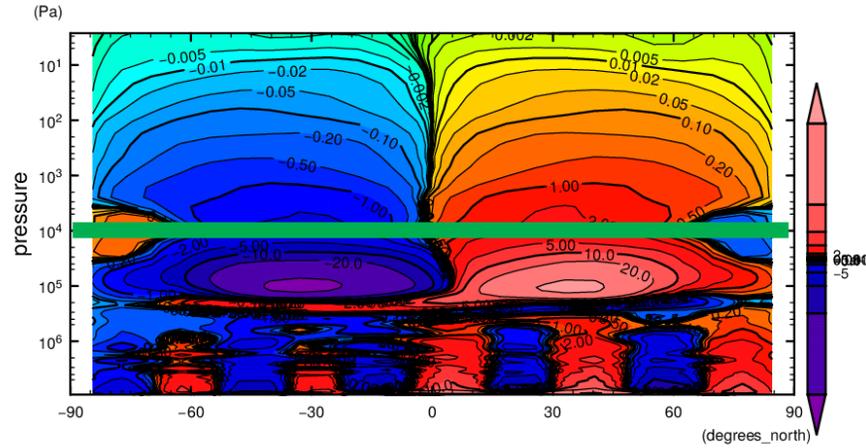
温度

temperature



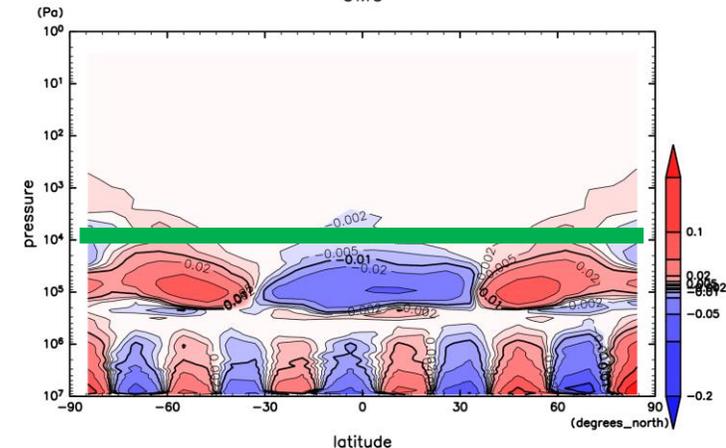
子午面循環質量流線関数 (オイラー平均)

MSF



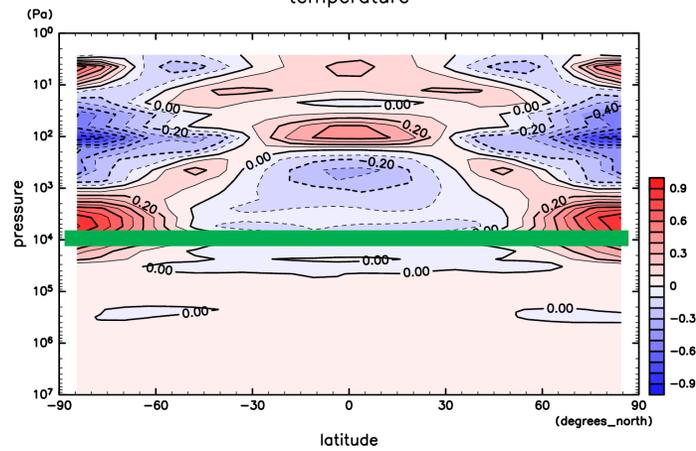
鉛直風 ( $\omega = dp/dt$ )

OMG



東西  
平均

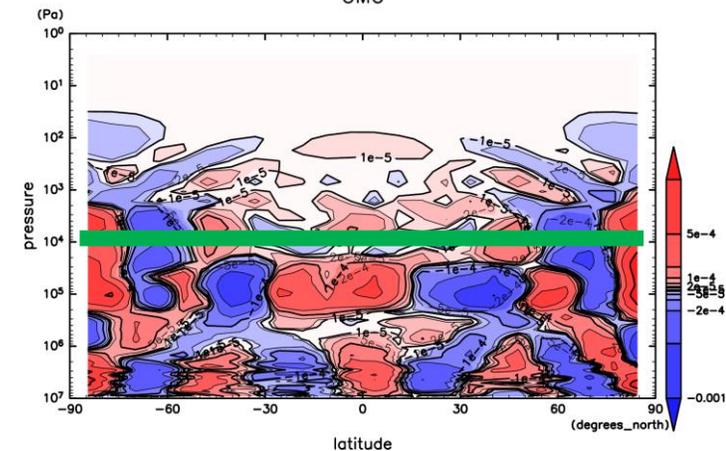
temperature



latitude

latitude

OMG

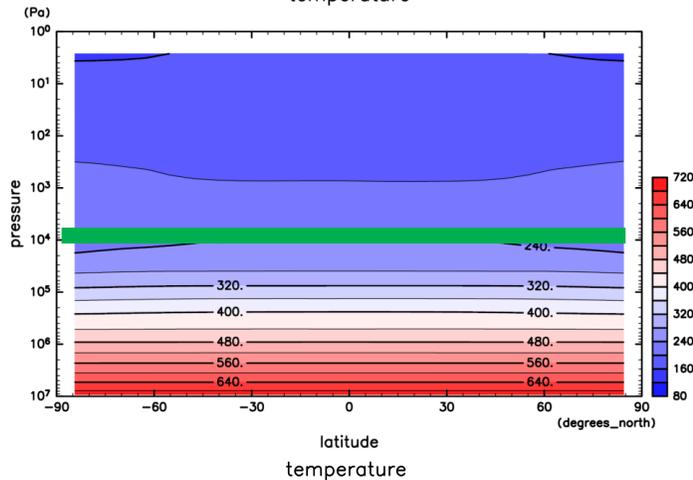


EOF1

# 何が起こっているのか

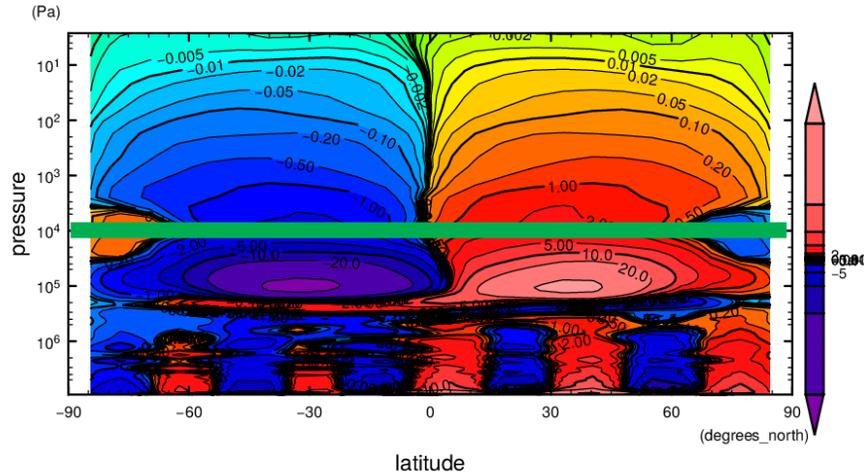
温度

temperature



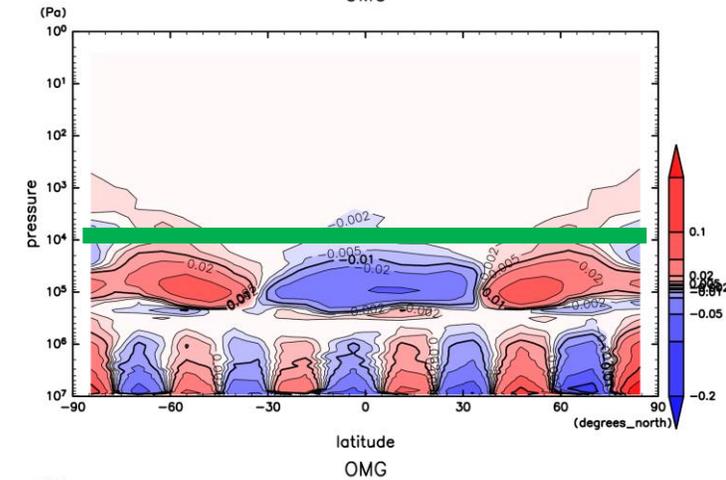
子午面循環質量流線関数 (オイラー平均)

MSF



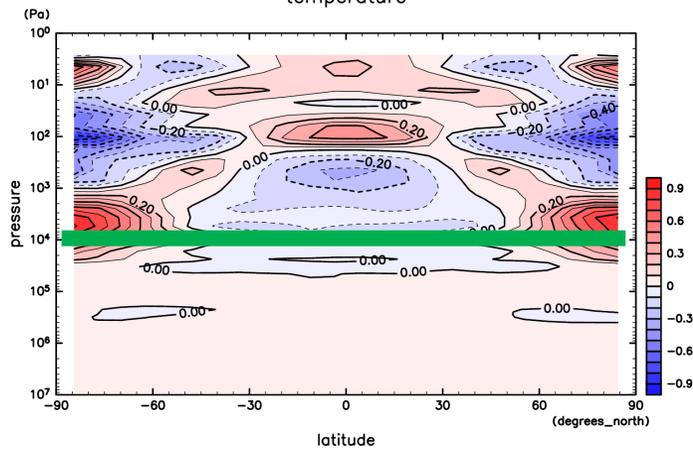
鉛直風 ( $\omega = dp/dt$ )

OMG



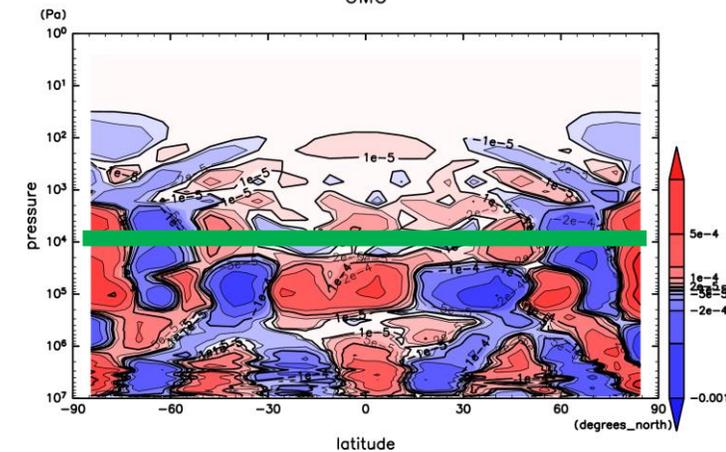
東西  
平均

temperature



- 0.1 bar 付近の高緯度には間接循環が存在.
- その領域の鉛直風の EOF1 主成分の大きさは平均値の ~10%.
- 間接循環を駆動する波の活動の時間変化が原因だろう.
  - なぜ変動する?
  - 時間スケールは何が決めている?
  - EP フラックスの変動を調べれば良いのだろうけど...

EOF1



# まとめ

- 大気海洋分野では EOF 解析と呼ばれる主成分分析によって時間変動構造の抽出が行われている。
- 今回は、複数のデータに対して EOF 解析を適用した。
  - 練習として関数を組み合わせて生成したデータを EOF 解析したところ、伝播性の擾乱は二つの主成分に分解されることを確認した。
  - NCEP/NCAR 再解析データの北半球高緯度域の海面更生気圧等の EOF 解析により、先行研究で示されていた北極振動の信号を検出できることを確認した。
  - DCPAM AMIP 実験の北半球高緯度域の海面更生気圧等の EOF 解析により、再解析データと整合的な北極振動の信号を検出できた。
  - DCPAM 金星大気実験の雲層上端付近の温度等の EOF 解析により、最も卓越する変動は低緯度と高緯度が逆位相でほぼ東西一様な構造であった。この変動はスーパーローテーション上端の低緯度に最大値を持つ東西風の変動と関係。
    - この変動が実際にあるかどうかはわからない。
- EOF 解析から (直接的に) 大気変動の物理機構は理解できないが、大気中の (時間) 変動構造の抽出手法としてはおもしろい。

おまけ

# EOF 解析の説明

# 主成分分析 解説 (M 次元)

$M$  個の空間格子点上の値を一次元に並べたデータの組  $N$  個を考える (時間方向のデータ数が  $N$  個).

$$(x_{1,1}, x_{1,2}, \dots, x_{1,M}), (x_{2,1}, x_{2,2}, \dots, x_{2,M}), \dots, (x_{N,1}, x_{N,2}, \dots, x_{N,M})$$

つまり,

$$\mathbf{x}_n = (x_{n,1}, x_{n,2}, \dots, x_{n,M}) \quad (n = 1, 2, \dots, N)$$

を考える.

( $x_{n,m}$  は  $n$  番目の時刻における  $m$  番目の空間点の値)

# 主成分分析 解説 (M 次元)

このデータを, それぞれの格子点における平均  $\left( \overline{x_i} = \frac{1}{N} \sum_{n=1}^N x_{n,i} \right)$  からの偏差で表すと

$$\begin{aligned} \mathbf{x}'_n &= (x'_{n,1}, x'_{n,2}, \dots, x'_{n,M}) \\ &= (x_{n,1} - \overline{x_1}, x_{n,2} - \overline{x_2}, \dots, x_{n,M} - \overline{x_M}) \quad (n = 1, 2, \dots, N) \end{aligned}$$

となる.  $x'_{n,m}$  は  $n$  番目の時刻における  $m$  番目の空間点の平均からの偏差である.  
この  $M$  次元空間内のデータの, 原点を通る直線への射影を考える.  
この原点を通る直線は,  $M$  個の空間格子点を作る空間構造を表す.

# 主成分分析 解説 (M 次元)

このデータの, 原点を通る直線の単位接線ベクトルを  $\mathbf{a} = (a_1, a_2, \dots, a_M)$  ( $|\mathbf{a}|^2 = a_1^2 + a_2^2 + \dots + a_M^2 = 1$ ) とすると, 射影は

$$z_n = \mathbf{x}'_n \cdot \mathbf{a} = x'_{n,1}a_1 + x'_{n,2}a_2 + \dots + x'_{n,M}a_M$$

となる. ここで  $z_n$  は, 直線に射影された点 (直線への垂線と直線の交点) と原点との距離 (符号付き) であり, すべての点を射影すると, 元の  $M$  次元分布を 1 次元分布に変換したことになる.

# 主成分分析 解説 (M 次元)

ここで, 射影された点  $z_n$  の分散が最も大きくなる直線 (の接線ベクトル) を求める. (このとき, 最も情報の損失が小さい.)

$z_n$  の分散は,

$$\sigma_z^2 = \frac{1}{N-1} \sum_{n=1}^N (z_n - \bar{z})^2 = \frac{1}{N-1} \sum_{n=1}^N \{(\mathbf{x}'_n - \bar{\mathbf{x}}') \cdot \mathbf{a}\}^2 = \frac{1}{N-1} \sum_{n=1}^N (\mathbf{x}'_n \cdot \mathbf{a})^2$$

であり, 求めたいのは, 分散  $\sigma_z^2$  が最大となる  $\mathbf{a}$  である.

# 主成分分析 解説 (M 次元)

言い換えると,  $|\mathbf{a}|^2 = a_1^2 + a_2^2 + \cdots + a_M^2 = 1$  の条件の下で  $\frac{\partial \sigma_z^2}{\partial a_1} = \frac{\partial \sigma_z^2}{\partial a_2} = \cdots = \frac{\partial \sigma_z^2}{\partial a_M} = 0$  となる  $\mathbf{a}$  を求めることである.

ラグランジュの未定乗数法を用いると, 定数  $\lambda$  を用いて

$$F(\mathbf{a}, \lambda) = \sigma_z^2 - \lambda(a_1^2 + a_2^2 + \cdots + a_M^2 - 1)$$

を定義して,

$$\frac{\partial F}{\partial a_1} = 0, \quad \frac{\partial F}{\partial a_2} = 0, \cdots, \quad \frac{\partial F}{\partial a_M} = 0, \quad \frac{\partial F}{\partial \lambda} = 0$$

を満たす  $\mathbf{a}$  を求めれば良い.

# 主成分分析 解説 (M 次元)

計算してみると,

$$\begin{aligned}\frac{\partial F}{\partial a_1} &= 2 \frac{1}{N-1} \sum_{n=1}^N (\mathbf{x}'_n \cdot \mathbf{a}) x'_{n,1} - 2\lambda a_1 \\ &= 2 \frac{1}{N-1} \sum_{n=1}^N x'_{n,1} a_1 + x'_{n,2} a_2 + \cdots + x'_{n,M} a_M x'_{n,1} - 2\lambda a_1 \\ &= 2 \frac{1}{N-1} \sum_{n=1}^N (x'_{n,1})^2 a_1 + 2 \frac{1}{N-1} \sum_{n=1}^N x'_{n,1} x'_{n,2} a_2 \\ &\quad + \cdots + 2 \frac{1}{N-1} \sum_{n=1}^N x'_{n,1} x'_{n,M} a_M - 2\lambda a_1 \\ &= 2\sigma_{x'_1}^2 a_1 + 2C_{x'_1, x'_2} a_2 + \cdots + 2C_{x'_1, x'_M} a_M - 2\lambda a_1 = 0,\end{aligned}$$

# 主成分分析 解説 (M 次元)

同様にして計算すると,

$$\frac{\partial F}{\partial a_1} = 2\sigma_{x'_1}^2 a_1 + 2C_{x'_1, x'_2} a_2 + \cdots + 2C_{x'_1, x'_M} a_M - 2\lambda a_1 = 0,$$

$$\frac{\partial F}{\partial a_2} = 2C_{x'_1, x'_2} a_1 + 2\sigma_{x'_2}^2 a_2 + \cdots + 2C_{x'_1, x'_M} a_M - 2\lambda a_2 = 0,$$

⋮

$$\frac{\partial F}{\partial a_M} = 2C_{x'_1, x'_M} a_1 + 2C_{x'_1, x'_M} a_2 + \cdots + 2\sigma_{x'_M}^2 a_M - 2\lambda a_M = 0,$$

$$\frac{\partial F}{\partial \lambda} = -(a_1^2 + a_2^2 + \cdots + a_M^2 - 1) = 0$$

# 主成分分析 解説 (M 次元)

整理すると,

$$\mathbf{V}\mathbf{a} = \lambda\mathbf{a}$$

$$\mathbf{V} = \begin{pmatrix} \sigma_{x'_1}^2 & C_{x'_1, x'_2} & \cdots & C_{x'_1, x'_M} \\ C_{x'_2, x'_1} & \sigma_{x'_2}^2 & \cdots & C_{x'_2, x'_M} \\ \vdots & \vdots & \ddots & \vdots \\ C_{x'_M, x'_1} & C_{x'_M, x'_2} & \cdots & \sigma_{x'_M}^2 \end{pmatrix}$$

$$a_1^2 + a_2^2 + \cdots + a_M^2 = 1$$

と書ける. ただし,  $C_{x'_i, x'_j} = C_{x'_j, x'_i}$  である ( $\mathbf{V}$  は対称行列).  $\mathbf{V}$  は分散と共分散を要素とする行列 (分散共分散行列) で, 問題は  $\mathbf{V}$  の固有値問題に帰着する.

# 主成分分析 解説 (M 次元)

なお, 元データ  $x_{n,m}$  ( $n$  番目の時刻における  $m$  番目の空間格子点の値) をまとめ

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,M} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,M} \end{pmatrix}$$

と表す.

# 主成分分析 解説 (M 次元)

各格子点の時間平均 (気候値)

$$\overline{x}_m = \frac{1}{N} \sum_{n=1}^N x_{n,m}$$

を成分に持つベクトル

$$\overline{\boldsymbol{x}} = (\overline{x}_1, \overline{x}_2, \dots, \overline{x}_M)$$

を定義する.

# 主成分分析 解説 (M 次元)

平均からの偏差は,

$$\mathbf{X}' = \mathbf{X} - \mathbf{1}\bar{x} = \left( \mathbf{I} - \frac{1}{N} \mathbf{1}\mathbf{1}^T \right) \mathbf{X}$$

と書ける. ただし, ここで,  $\mathbf{I}$  は単位行列であり,  $\mathbf{1} = (1, 1, \dots, 1)^T$  である.  
これを用いると, 分散共分散行列  $\mathbf{V}$  は

$$\mathbf{V} = \mathbf{X}'^T \mathbf{X}'$$

と書ける.

# 北極振動の EOF 解析 NCEP/NCAR 再解析データ

# 北極振動の EOF 解析

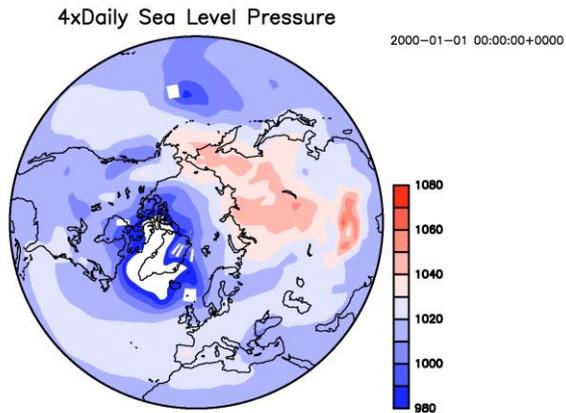
- EOF 解析によって性質を既によく調べられている北極振動について, EOF 解析を自分で解析してみる.
  - これまでに示されてきた解析結果と比べることで解析方法が正しいことを確認する.
  - 北極振動の性質を自分で確認してみる.

# NCEP/NCAR 再解析データの EOF 解析

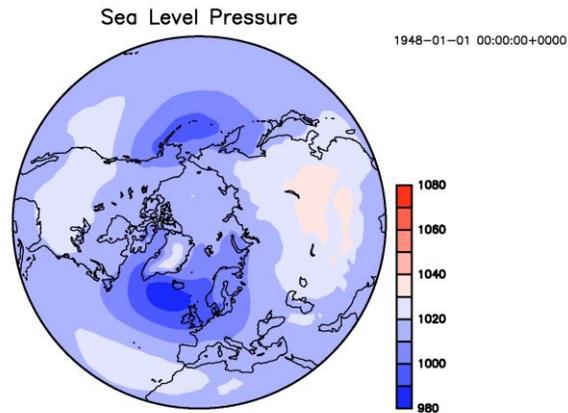
- 使用データ
  - NCEP/NCAR 再解析データ (Kalnay et al., 1996)
- 解析方針
  - Hannachi (2004) に従う.
  - データ期間：1948-2000 年 (の 12 月, 1 月, 2 月)
  - 物理量
    - 海面更生氣圧 (SLP; Sea Level Pressure) の月平均値の偏差
      - 1945-2000 年の SLP の月平均値を用いて平均的な年変化を求めることで, そこからの偏差を求める.
  - 地理的範囲：北半球 20° 以北

# EOF 解析の処理：海面更生氣圧

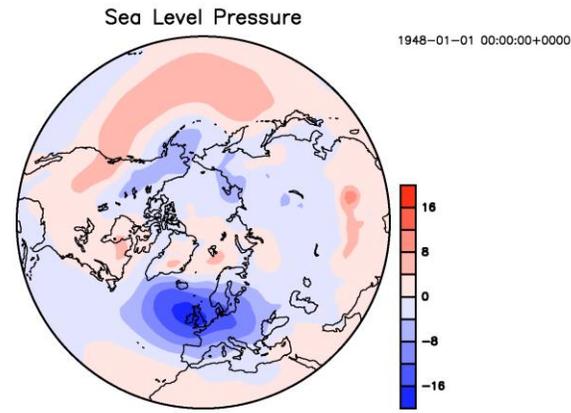
瞬間値  
(2000 年)



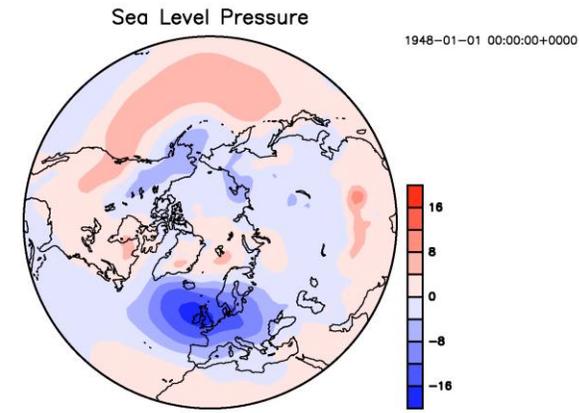
月平均値  
(1948-2000 年)



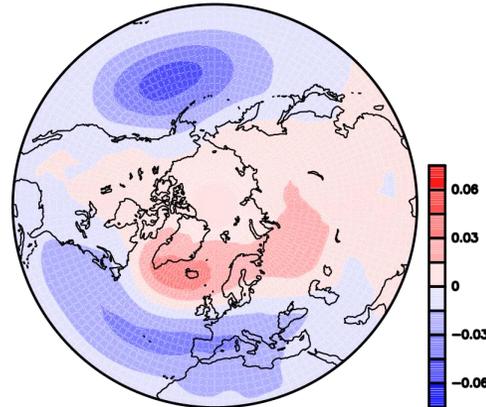
気候値からの偏差  
(1948-2000 年)



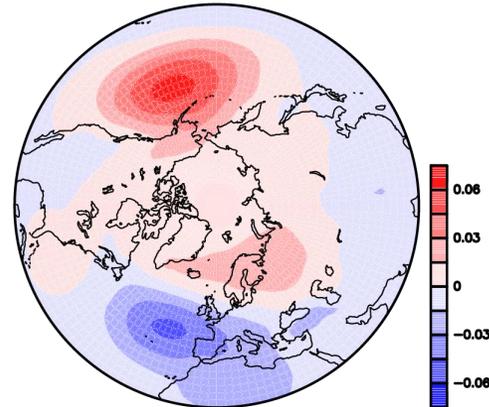
気候値からの偏差  
(1948-2000 年)  
12, 1, 2 月のみ



EOF1  
EOF1

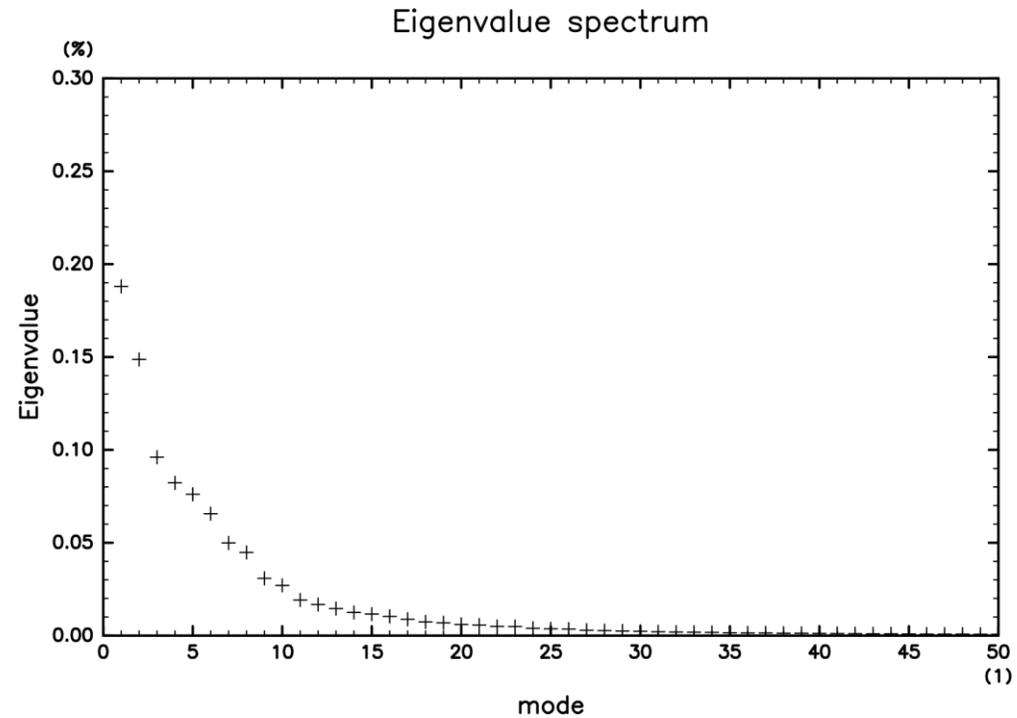


EOF2  
EOF2

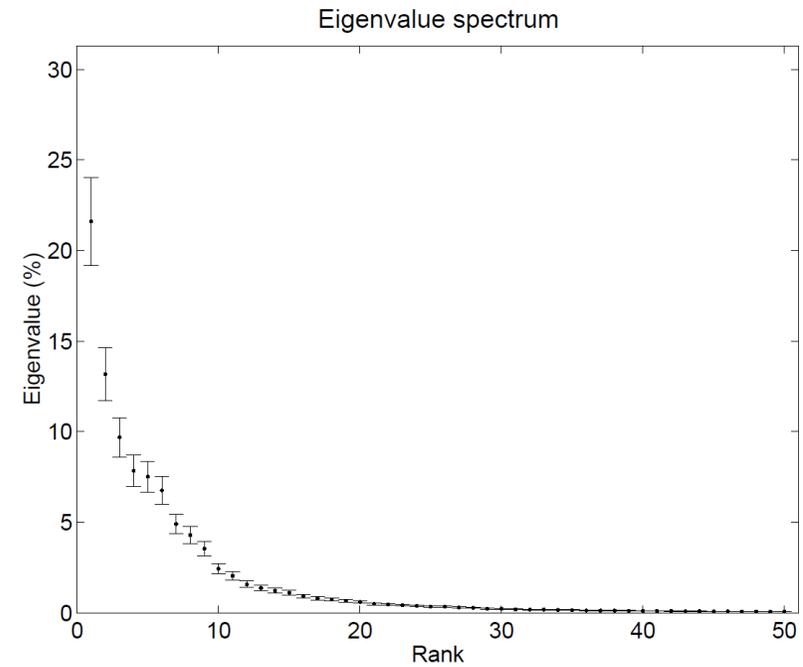


# 固有値スペクトル

今回の解析

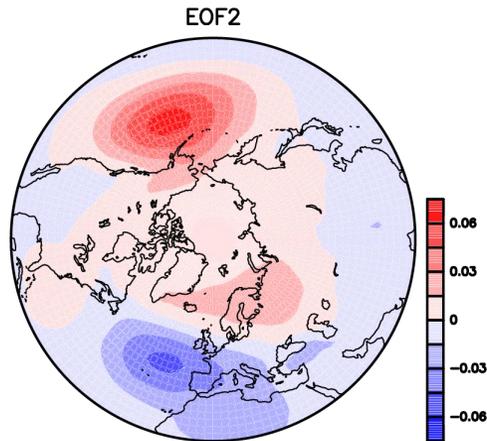
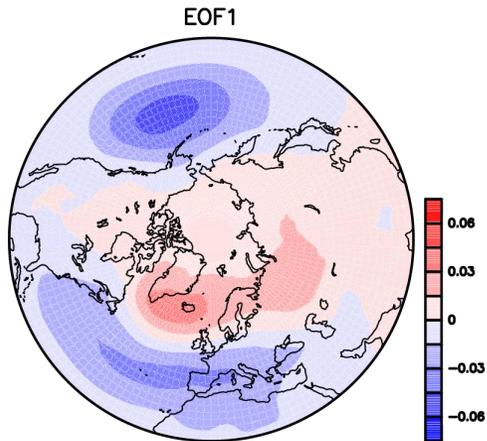


Hannachi (2004)



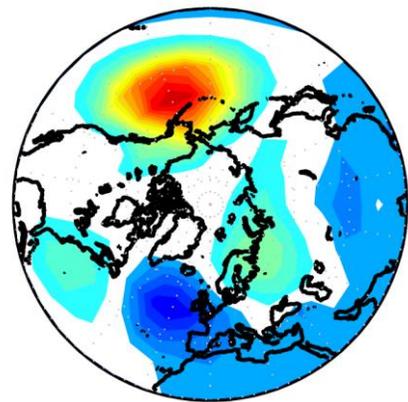
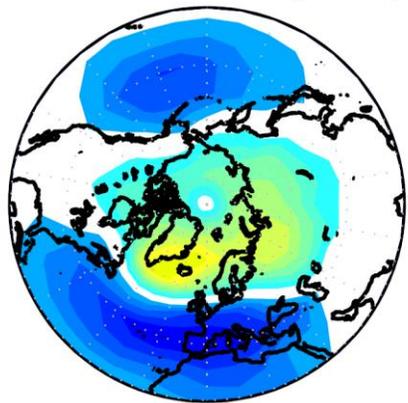
# EOF 構造

今回の解析



DJFSLP EOF1 (21%)

DJFSLP EOF2 (13%)

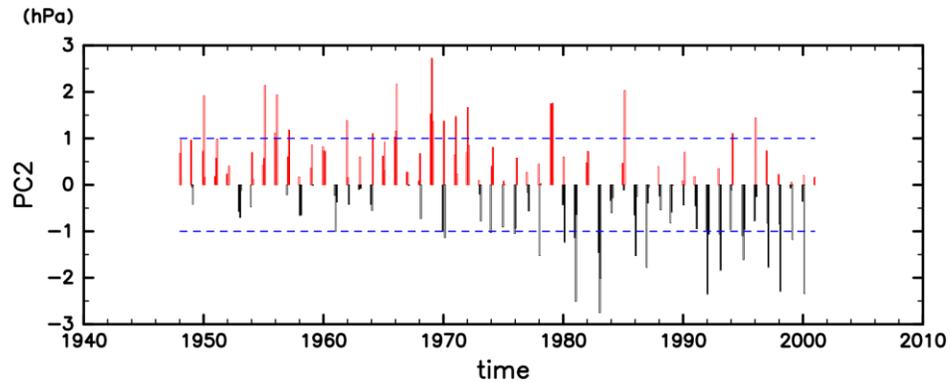
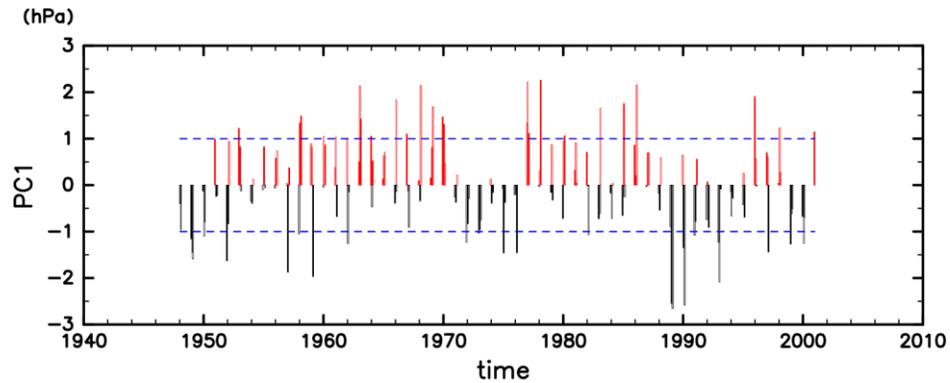


Hannachi (2004)

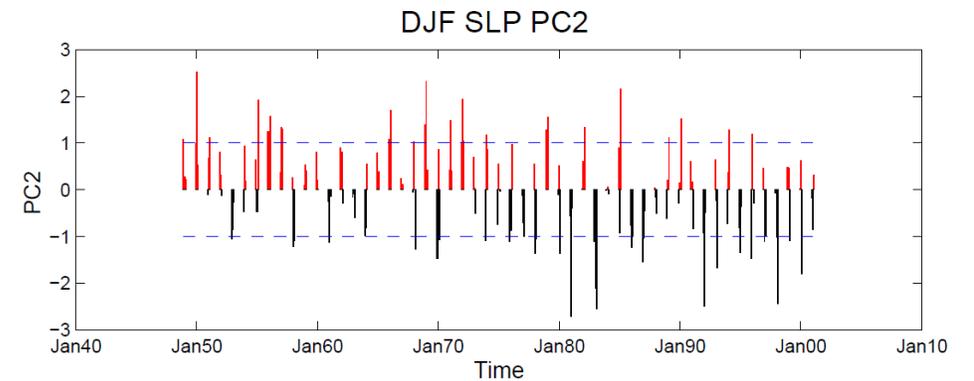
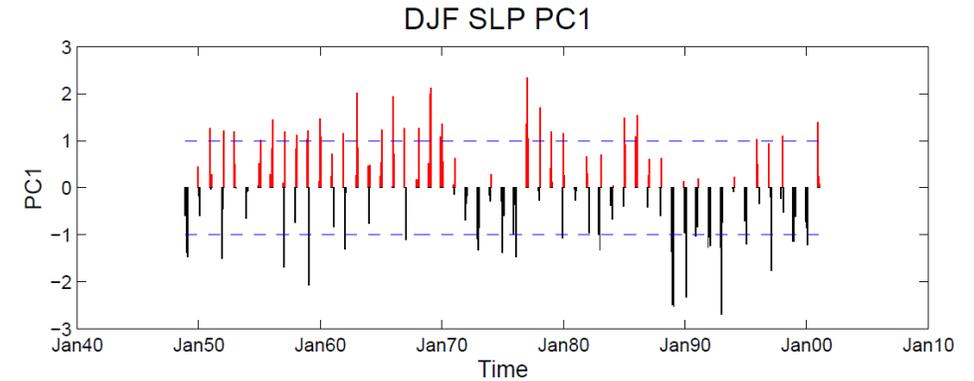
- EOF1, 2 ともに Hannachi (2004) で示されていたのとよく似た構造が検出された.
  - 特徴
    - 高緯度の圧力偏差の中心はグリーンランド付近.
    - 中緯度の圧力偏差の中心はアラスカ沖と大西洋.
  - 注意
    - 今回の解析で得られる EOF は無次元

# 主成分 (PC)

今回の解析

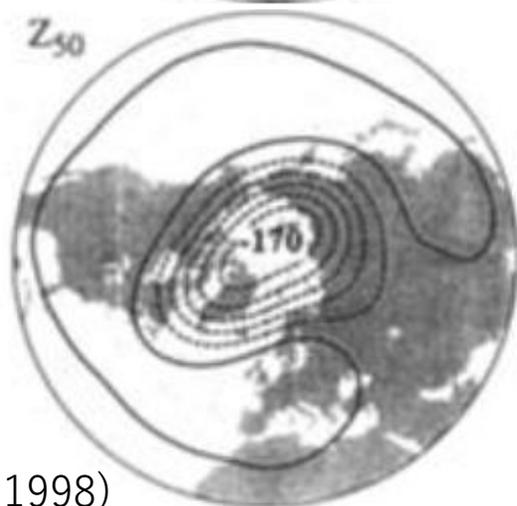
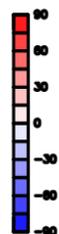
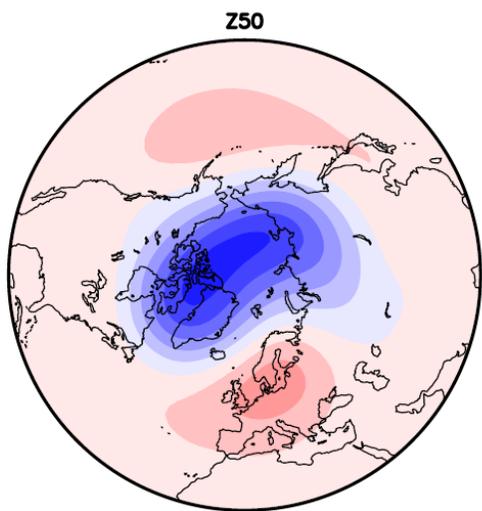
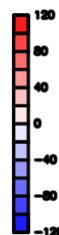
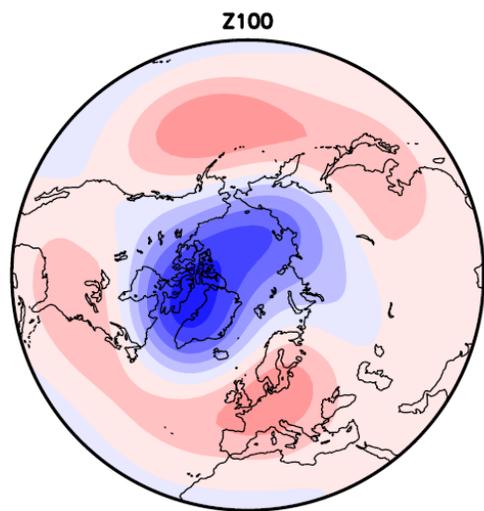
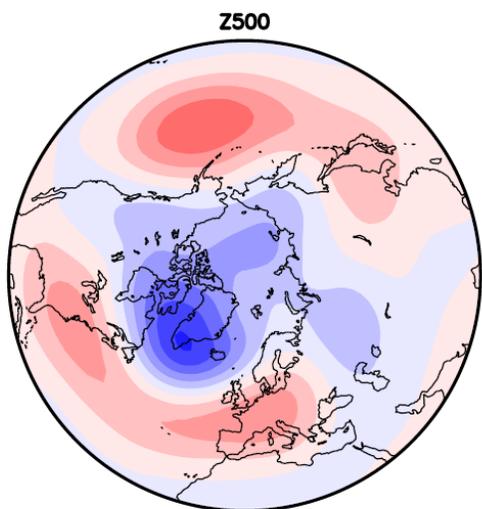
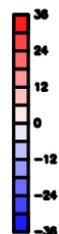
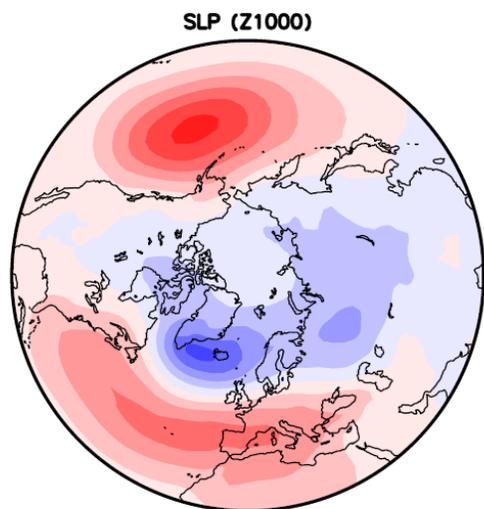


Hannachi (2004)



Hannachi (2004) とよく似た主成分が検出された.

# 鉛直構造 (EOF1)



(Thompson and Wallace, 1998)

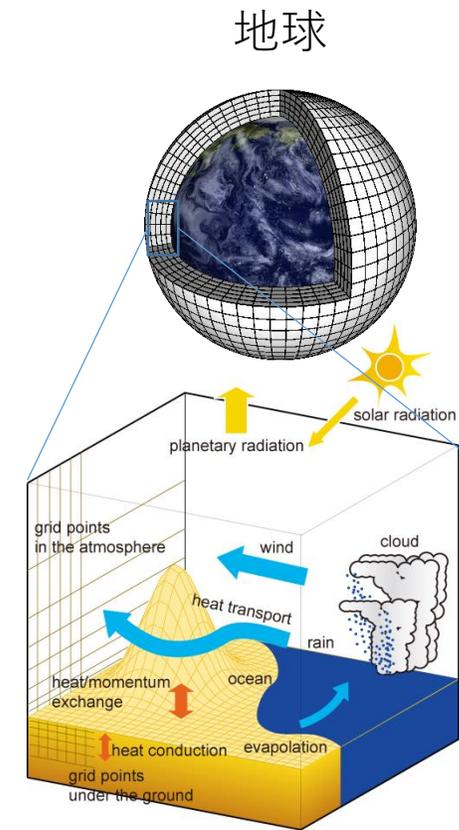
# 北極振動の EOF 解析 DCPAM AMIP 実験

# AMIP

- Atmospheric Model Intercomparison Project
  - 大気モデル相互比較プロジェクト
- MIP
  - 各国の研究グループの大気モデルを同じ条件で用いて結果を比較することを通して、モデル全体およびその部品のパフォーマンスを把握し、その改良の指針を得ることを目指す
  - これまでに様々な種類の相互比較プロジェクト (MIP) が行われてきた
    - OMIP - Ocean Model Intercomparison Project
    - CMIP - Coupled Model Intercomparison Project
    - PMIP - Palaeoclimate Modelling Intercomparison Project
    - GeoMIP - Geoengineering Model Intercomparison Project
  - など

# DCPAM AMIP 実験 — モデル

- DCPAM
  - 地球流体電脳倶楽部有志で開発している惑星大気大循環モデル
- DCPAM 開発版
  - 力学過程
    - プリミティブ方程式
      - 水平スペクトル変換法
      - 鉛直差分法 (Arakawa and Suarez, 1983)
  - 使用物理過程
    - DCPAM 地球大気放射モデル
    - Mellor and Yamada (1982) level 2.5
    - 乾燥対流調節
    - Relaxed Arakawa-Schubert (Moorthi and Suarez, 1992) の拡張版
    - 非対流性凝結 (Le Treut and Li, 1991)
    - 簡単雲モデル
    - 惑星表面・土壌熱拡散モデル
    - バケツモデル (Manabe, 1969)



# DCPAM AMIP 実験 — モデル

## 力学過程の方程式系

方程式系は, 静水圧平衡を仮定し, 鉛直方向に  $\sigma = p/p_s$  座標を用いたプリミティブ方程式系.

運動方程式 
$$\frac{\partial \vec{u}}{\partial t} = -(\vec{u} \cdot \nabla) \vec{u} - \dot{\sigma} \frac{\partial \vec{u}}{\partial \sigma} - f \vec{k} \times \vec{u} - \nabla \Phi + \frac{RT}{p_s} \nabla p_s + \vec{F}$$

熱力学の式 
$$\frac{\partial \Phi}{\partial \sigma} = -\frac{RT}{\sigma}$$

$$\frac{\partial T}{\partial t} = -(\vec{u} \cdot \nabla) T - \dot{\sigma} \frac{\partial T}{\partial \sigma} + \frac{\kappa T}{p} \frac{dp}{dt} + Q$$

連続の式 
$$\frac{\partial p_s}{\partial t} = -\nabla \cdot (p_s \vec{u}) - \frac{\partial}{\partial \sigma} (p_s \dot{\sigma})$$

$$\frac{\partial q}{\partial t} = -(\vec{u} \cdot \nabla) q - \dot{\sigma} \frac{\partial q}{\partial \sigma} + S_q$$

境界条件:

$$\dot{\sigma} = \frac{d\sigma}{dt} = 0 \quad \text{at } \sigma = 0, 1$$

$\vec{u}$  は水平風速,  $p$  は気圧,  $p_s$  は惑星表面気圧,  $T$  は温度,  $q$  は物質の混合比,  $\Phi$  はジオポテンシャル,  $\kappa = R/C_p$ ,  $R$  は気体定数,  $C_p$  は定圧比熱,  $\vec{F}$ ,  $Q$ ,  $S_q$  はそれぞれ摩擦, 非断熱加熱, 物質の生成・消滅.

# DCPAM AMIP 実験 — 実験条件

- 惑星パラメータ : 地球の値
- 地形 : ETOPO1
- 土地利用 : Matthews (1983, 1984, 1985)
- 海表面温度, 海氷面密度 : AMIPII 設定値
  - データ期間
    - 1870年1月~2010年3月
- O<sub>3</sub> 分布, CO<sub>2</sub>/CH<sub>4</sub>/N<sub>2</sub>O 混合比(全球一様), 太陽放射フラックス, 耕作率 : CMIP 設定値
  - データ期間
    - O<sub>3</sub> 分布 : 1850年~2009年
    - CO<sub>2</sub>/CH<sub>4</sub>/N<sub>2</sub>O 混合比 : 1765年~2500年
    - 太陽放射フラックス : 1882年~2008年
    - 耕作率 : 1500年~2100年
- 解像度 : T42L26
- 初期条件 : 等温静止大気
- 積分期間 : 1870年1月1日~2009年12月31日
  
- 注意 : 1882年以前, 2009年は入力データが揃っていない

# 東西平均温度 DJF 気候値 (1948-2000 年)

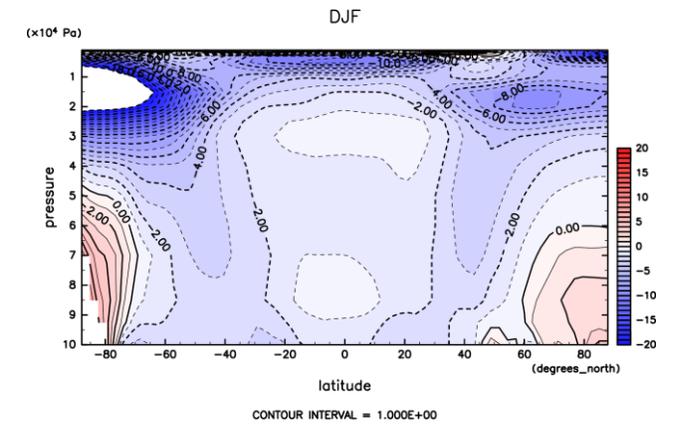
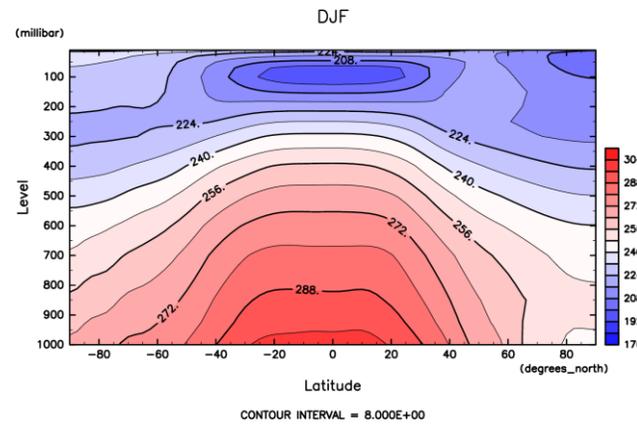
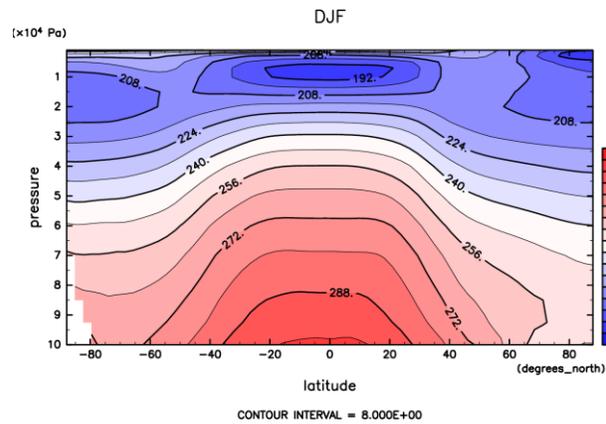
DCPAM

NCEP/NCAR 再解析

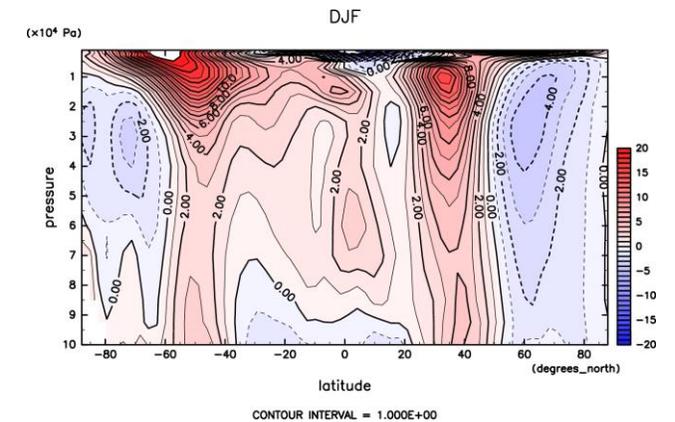
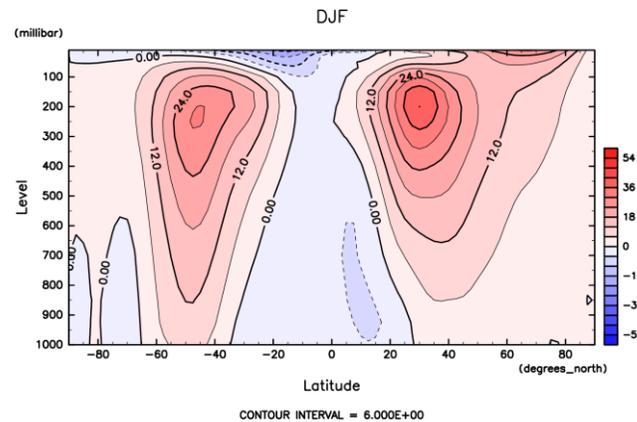
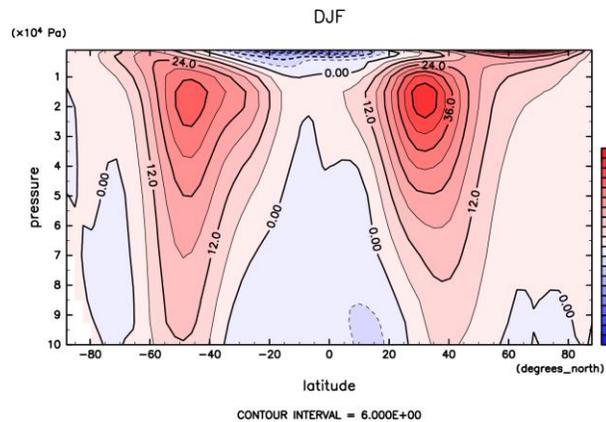
差

([DCPAM] - [NCEP/NCAR 再解析])

温度

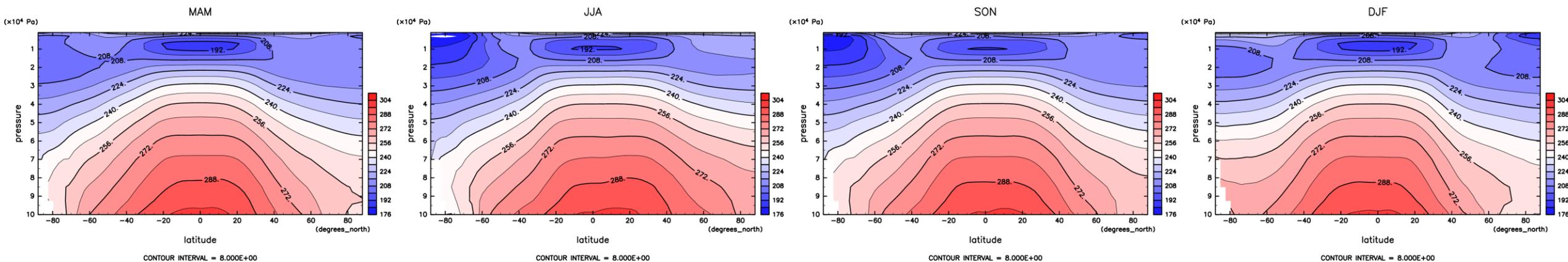


東西風

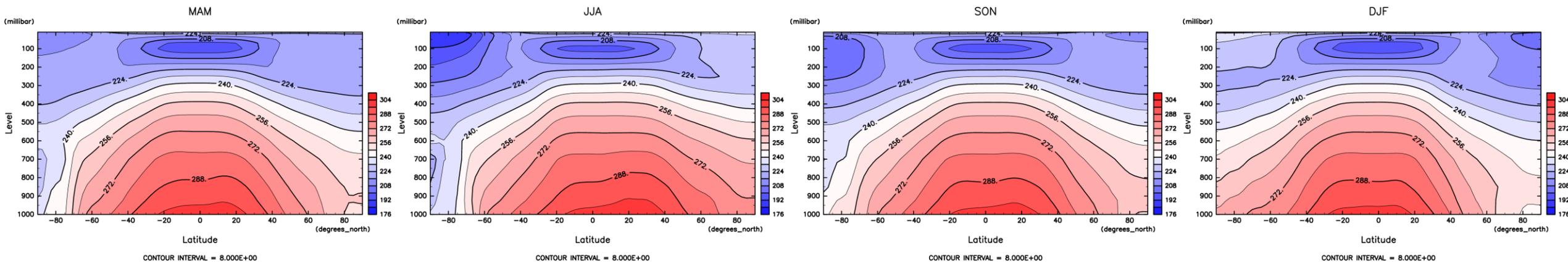


# 東西平均温度 気候値 (1948-2000 年)

## DCPAM

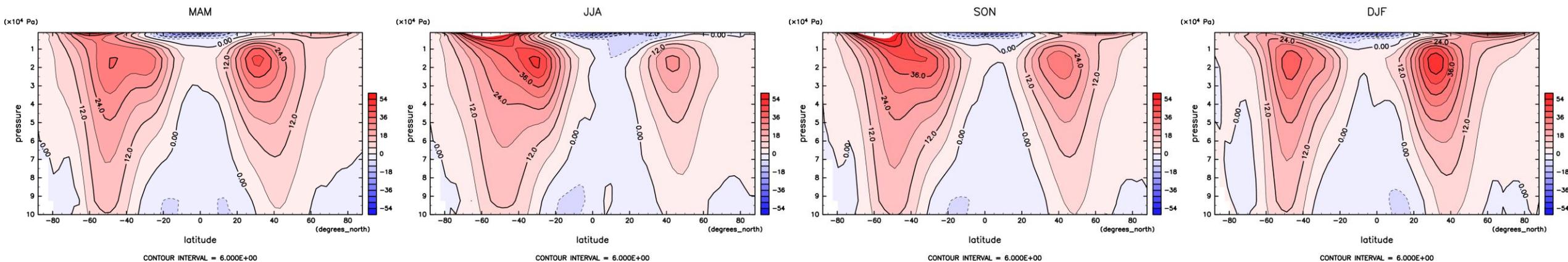


## NCEP/NCAR 再解析

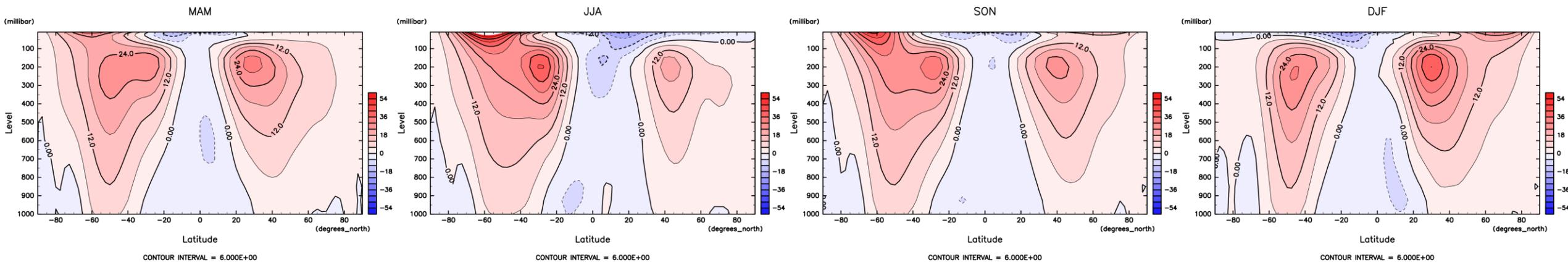


# 東西平均東西風 氣候值 (1948-2000 年)

## DCPAM

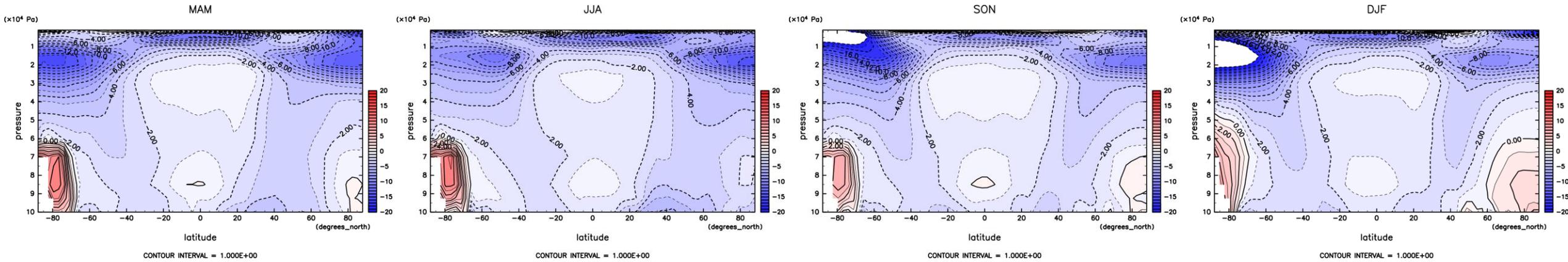


## NCEP/NCAR 再解析

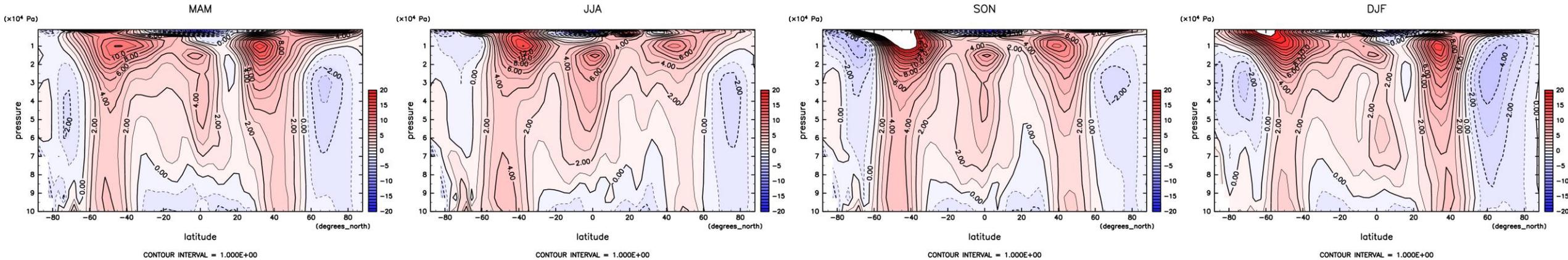


# 東西平均の差 気候値 (1948-2000 年) [DCPAM] - [NCEP/NCAR 再解析]

温度



東西風

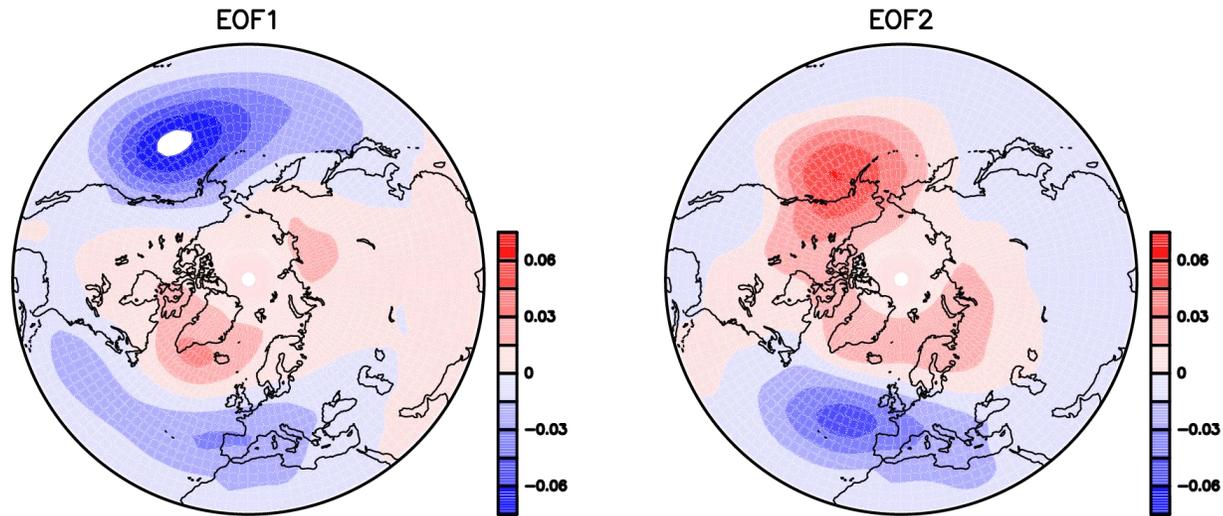


# DCPAM AMIP 実験結果の EOF 解析

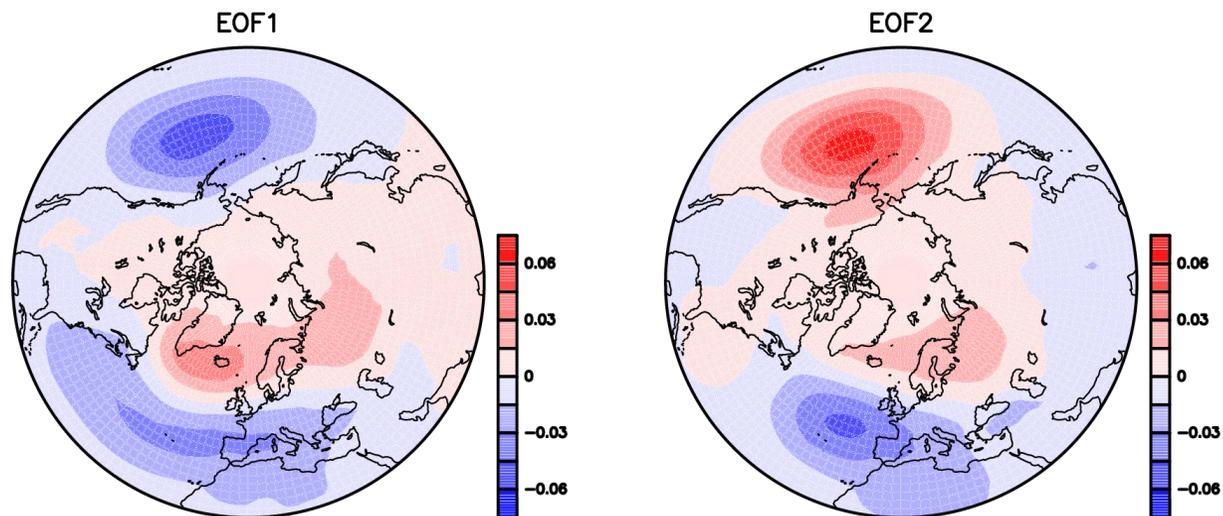
- 使用データ
  - DCPAM AMIP 実験
- 解析方針
  - NCEP/NCAR 再解析データの解析と同じ (Hannachi (2004) に従う).
  - データ期間 : 1948-2000 年 (の 12 月, 1 月, 2 月)
  - 物理量
    - 海面更生気圧 (SLP; Sea Level Pressure) の月平均値の偏差
      - 1945-2000 年の SLP の月平均値を用いて平均的な年変化を求めることで, そこからの偏差を求める.
  - 地理的範囲 : 北半球 20° 以北

# EOF 構造

DCPAM



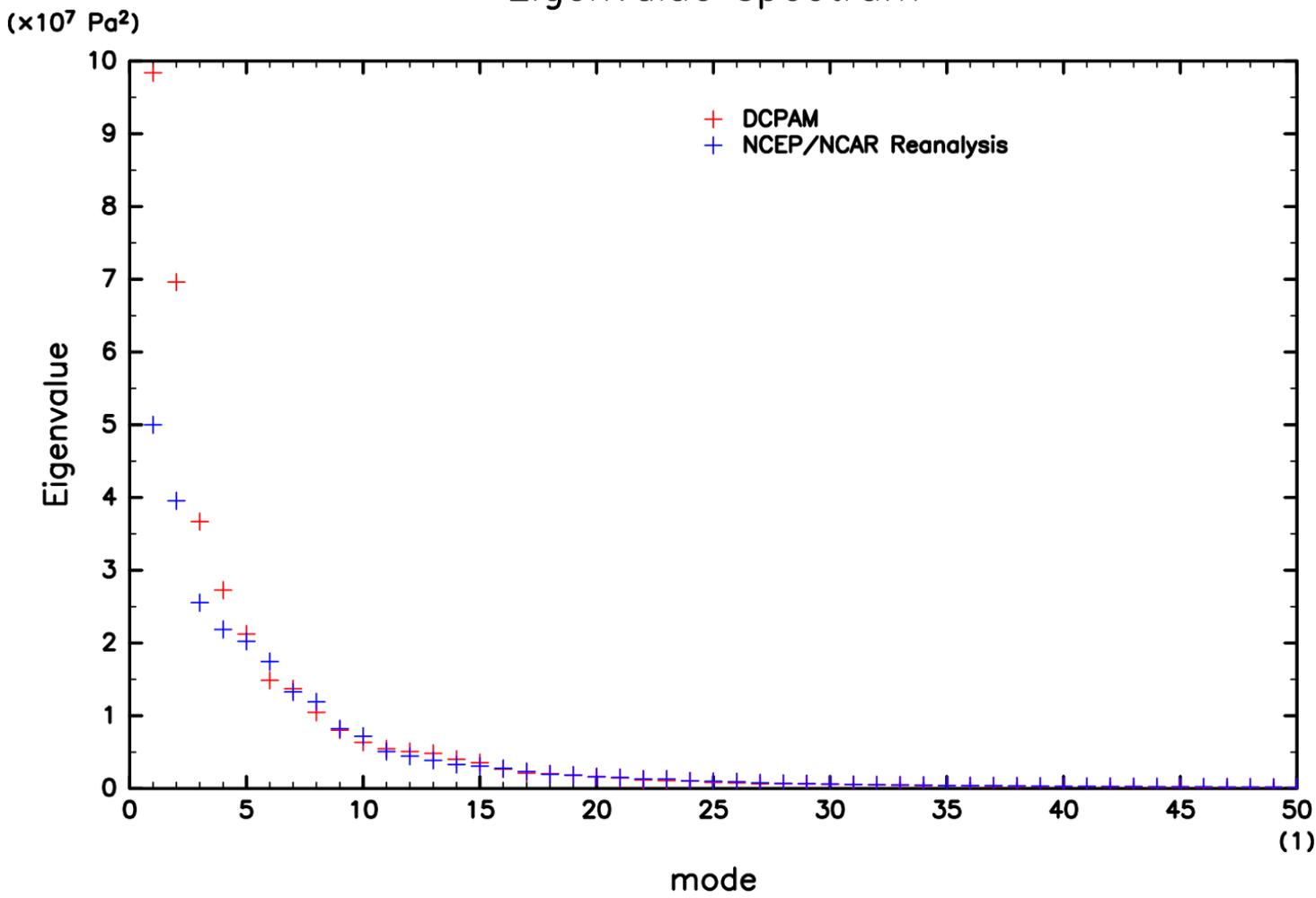
NCEP/NCAR  
再解析



- DCPAM は再解析データに見られる EOF1, EOF2 の構造を概ね表現

# 固有値スペクトル

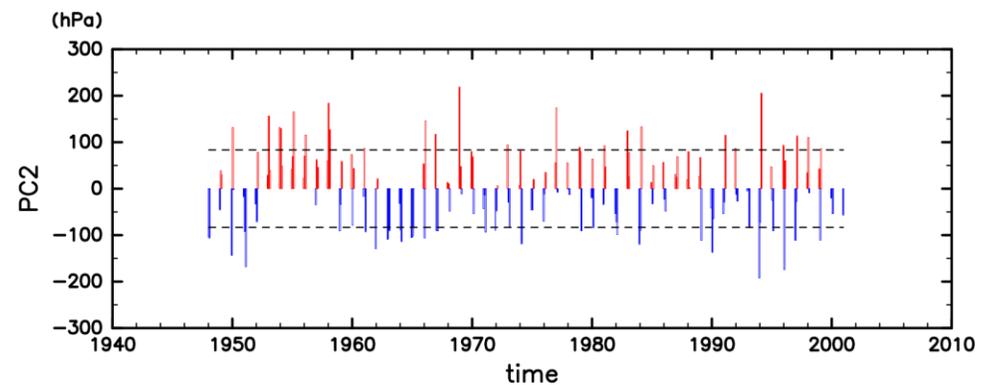
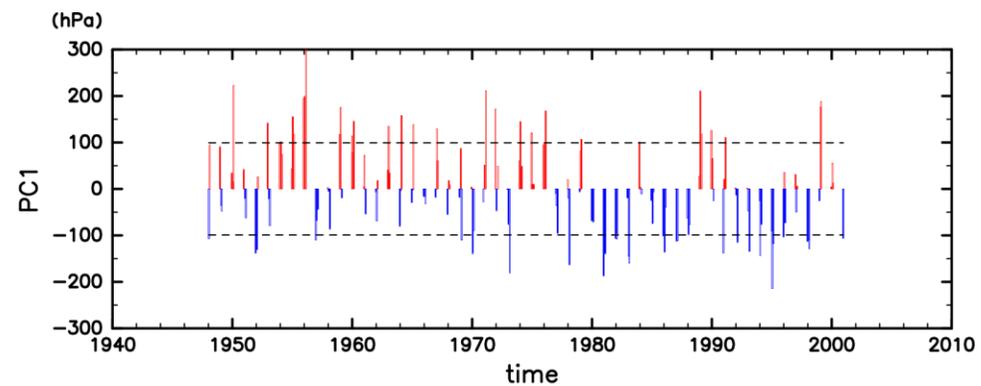
Eigenvalue spectrum



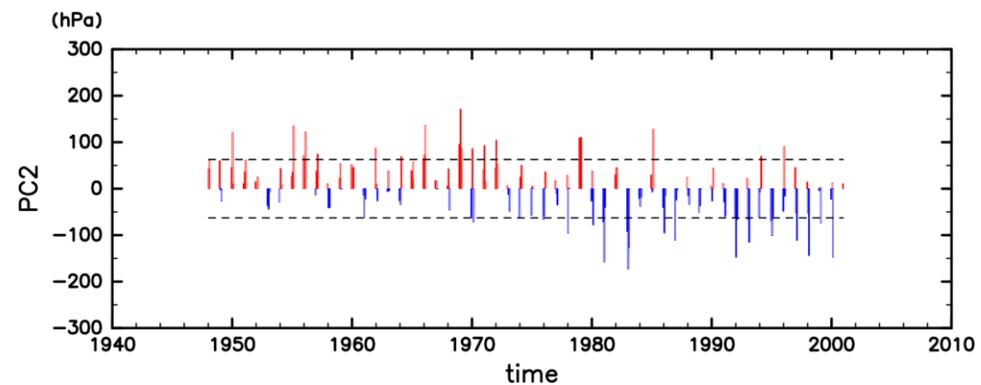
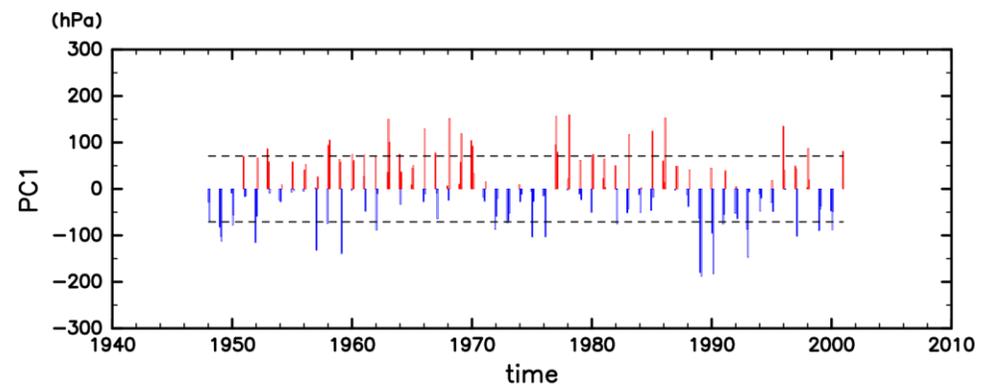
- DCPAM の EOF1, EOF2 の固有値は再解析データに比べて ~2 倍

# 主成分 (PC)

DCPAM

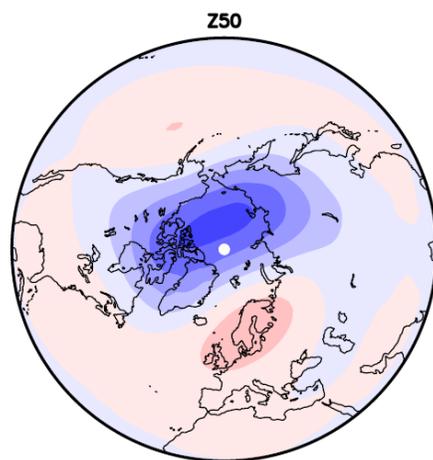
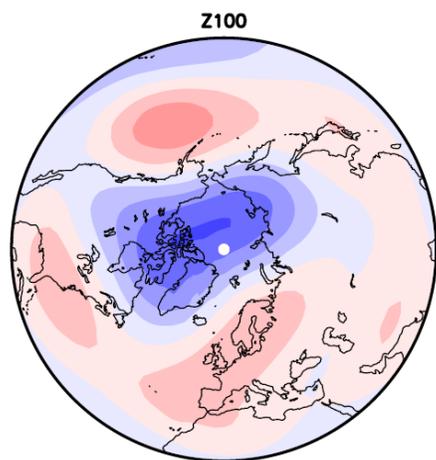
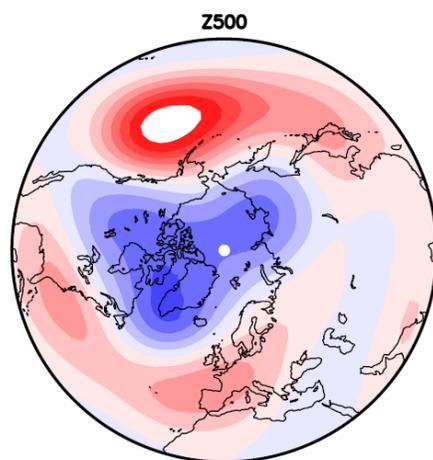
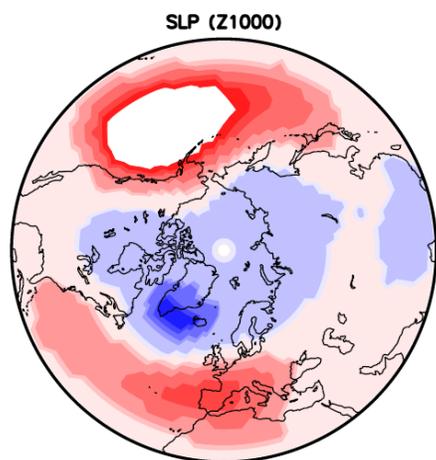


NCEP/NCAR 再解析



# 鉛直構造 (EOF1)

DCPAM



NCEP/NCAR 再解析

