

クラスタ解析を用いた 気候モデルの気候場再現性能評価の試み

日本気象学会 2009年度秋季大会 スペシャルセッション

「地球温暖化に伴う大気海洋諸現象の変化を考える
—CMIP3マルチ気候モデル比較解析研究から—」

(2009年11月27日@アクロス福岡)

横井 覚¹, 高薮 縁^{1,2}, 西井 和晃³, 宮坂 貴文³,

小坂 優⁴, 中村 尚^{3,2}, アジアメトリックWG

(1:東大CCSR, 2:JAMSTEC, 3:東大院理, 4:ハワイ大IPRC)

謝辞 本研究は、環境省地球環境研究総合推進費S-5-2の支援により実施された。

モデル性能評価指標(メトリック)

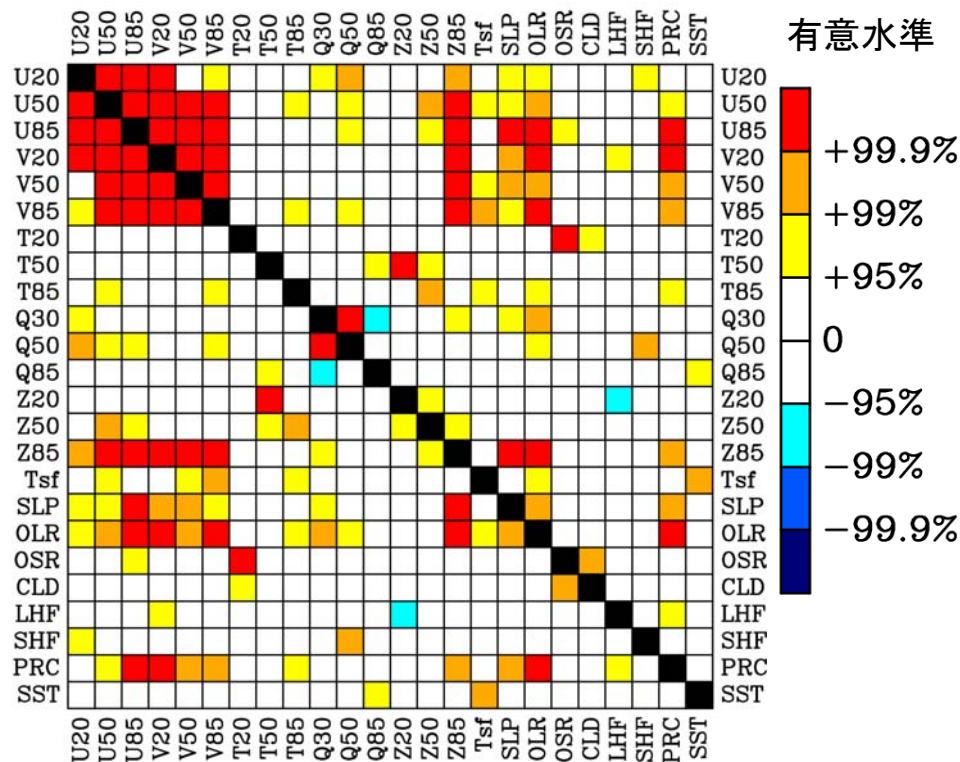
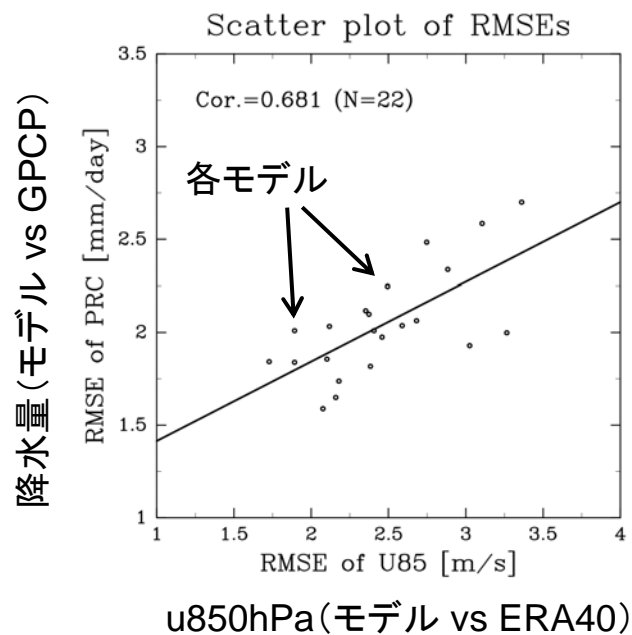
- 種々の物理量の気候値の誤差統計量の総合点で評価。
- 使い方:
 - 将来気候予測の重みづけ (e.g. Murphy et al. 2004)
 - 温暖化シグナル検出 (D&A study) (e.g. Santer et al. 2009)
 - 力学的ダウンスケーリング, 影響評価モデルの親モデル選定
 - 気候モデルの世代間比較 (e.g. Reichler and Kim 2008)
- 問題ごとにメトリックを定義する必要あり。
 - 適切な物理量, 領域, 季節, 誤差統計量を設定。
 - 適切に重みをつけて誤差をブレンド。
- S-5テーマ2での目標:

アジアの様々な気象の総合的再現性の評価指標の作成
(季節進行, 経年変動, 擾乱活動度, ...)

情報の冗長性

水平風・気温・OLR等の物理量間のRMSEの相関(JJA)

例) RMSE (u850hPa) vs RMSE (降水量)



物理量間で誤差が有意な正相関を持つペアが存在。

冗長性の軽減にむけて

冗長性があると・・・

- 将来予測で再現性能を二重に加味して重みづけしてしまう
(Over-constraint)
- メトリック定義の際の重みの統計的推定で問題
(重回帰分析での共線形性等)



冗長性を軽減する系統的方法をいろいろ試してみよう。

- クラスタ解析
 - ハードクラスタリング
 - ソフト(ファジー)クラスタリング
- 経験的直交関数展開(EOF)
- 特異値分解(SVD)
- ...

- 階層的的手法 **Ward法**
(近い因子同士順番に結合)
- 非階層的的手法 **K-mean法**
(予めクラスタ数を決め, 分類)

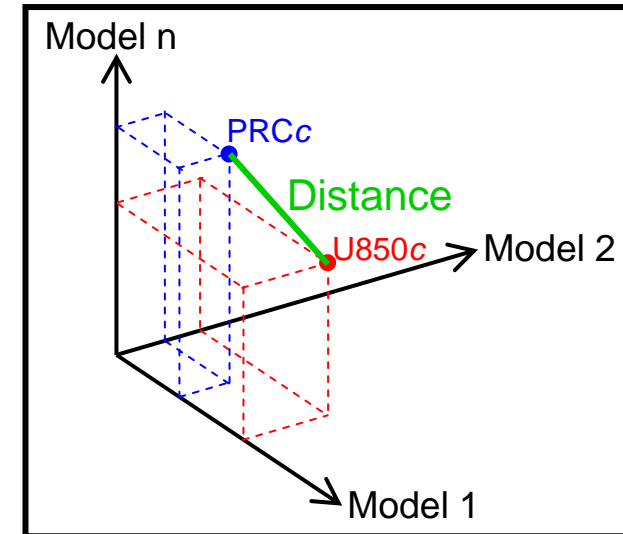
クラスタ解析詳細

- **6-8月, 月別気候値**
- 誤差統計量(因子):
 - 全球平均バイアスの絶対値
 - 全球平均偏差の二乗誤差平均(centered RMSE)
- 物理量(24種) → **因子数は43**

東西風(200, 500, 850hPa), 南北風(200, 500, 850hPa),
 気温(200, 500, 850hPa, 地表#), 比湿(300, 500, 850hPa),
 ジオポテンシャル高度(200, 500, 850hPa), 海面更正気圧+,
 潜熱&#, 顕熱&#, 外向き長波放射, 外向き短波反射, 降水量, 雲量, 海面水温#

(& centered RMSEのみ評価, + 標高1,000m以下の地点のみ評価, # 海上でのみ評価)
- 検証用データ

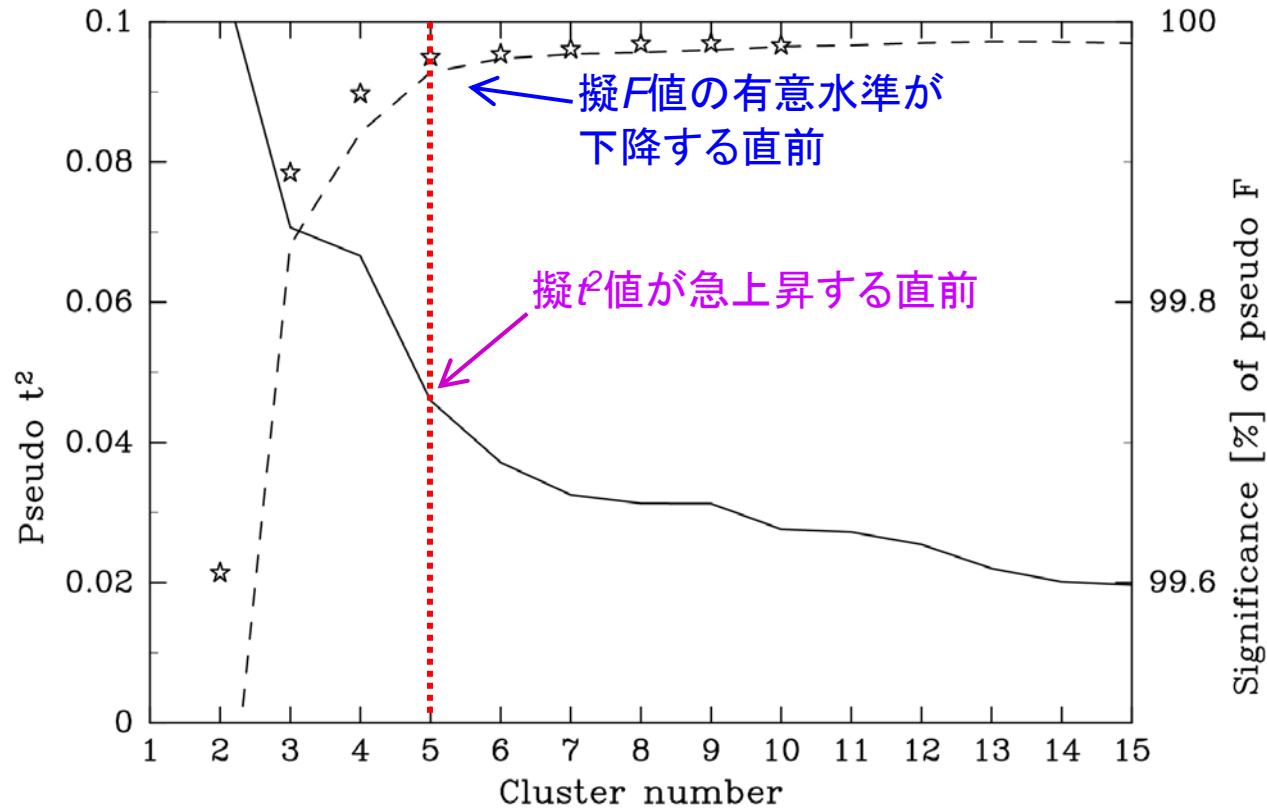
ERA40 (Uppala et al. 2005), SOC (Josey et al. 1999),
 CERES (Wielicki et al. 1996), GPCP (Adler et al. 2003),
 ISCCP-D2 (Rossow and Schiffer 1999), HadISST (Rayner et al. 2006)
- CMIP3の24モデルのうち, ECHO-G, UKMO-HadGEM1を除く**22モデル**。
- 類似度: 標準化した因子間のユークリッド距離。



クラスタ数

クラスタの分かれ方の妥当性の統計的評価指標

実線: 擬 t^2 値, 破線: 擬 F 値 (Ward法), 星印: 擬 F 値 (K-mean法)



- 両手法とも, クラスタ数5が最も妥当。

クラスタ数5

"b": global mean bias
 "c": centered RMSE

Ward法

U200c U500c U850c V200c V500c
 V850c T850c Q500c Z500c Z850c
 SLPc OLRc 降水c

T500b T850b Q850b Z200b Z500b
 T地表b SSTb

U200b U500b T500c Q300b Q300c
 Q500b Z200c T地表c 顕熱c SSTc

T200b T200c Q850c OSRc 雲量b
 雲量c 潜熱c

U850b Z850b SLPb OLRb OSRb
 降水b

K-mean法

U200c U500c U850c V200c V500c
 V850c T500c T850c Q500c Z500c
 Z850c T地表c SLPc OLRc 雲量c
 降水c

T500b T850b Q850b Z200b Z500b
 T地表b SSTb SSTc

U200b U500b Q300b Q300c Q500b
 Z200c 顕熱c

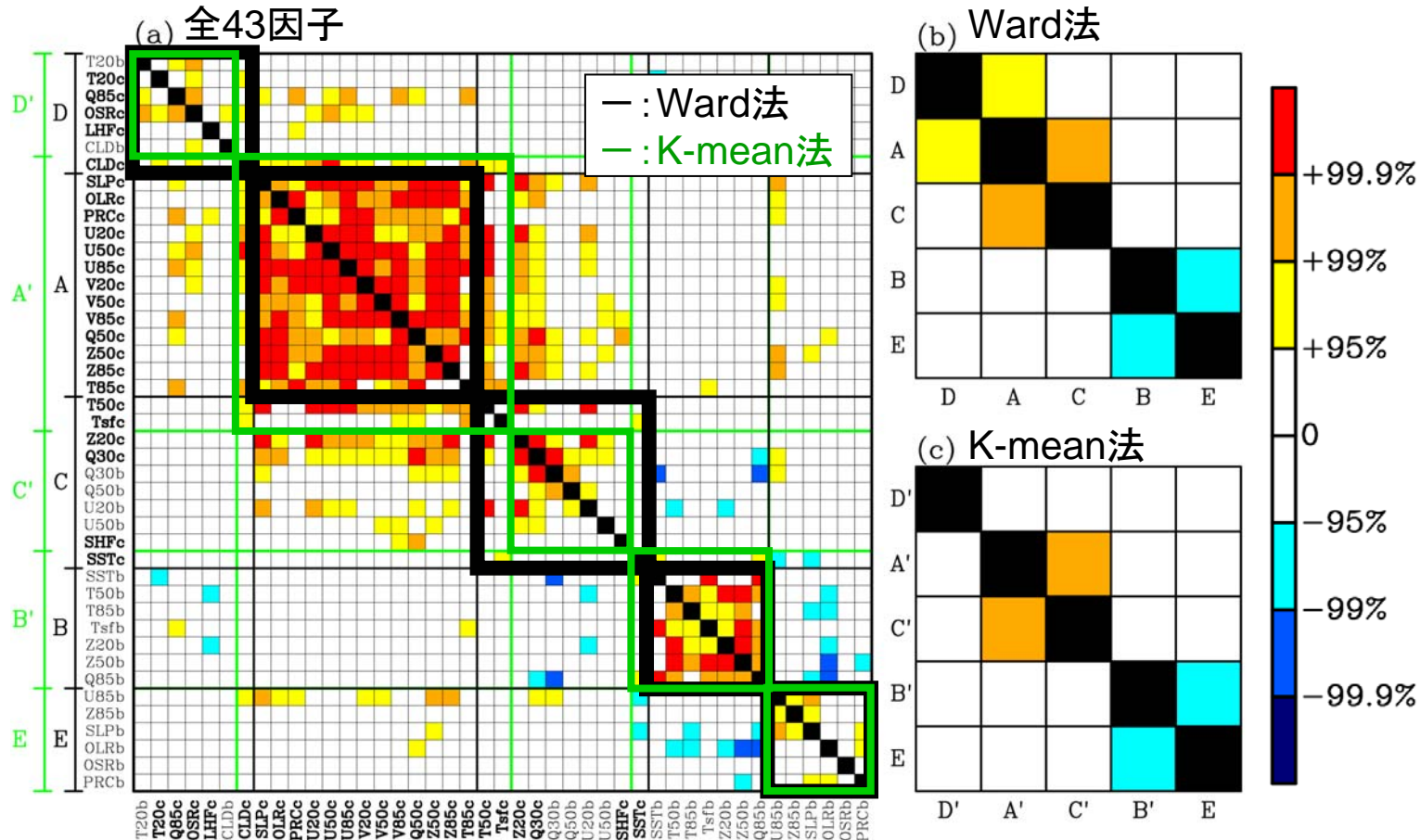
T200b T200c Q850c OSRc 雲量b
 潜熱c

U850b Z850b SLPb OLRb OSRb
 降水b

- 両手法で、結果の差はわずか(違いは4因子のみ)。
- Centered RMSEの多くがひとつのクラスタにまとまっている。
- 物理的・力学的に解釈可能な組み合わせも多い。

冗長性の軽減

因子間, 又はクラスタ間の相関係数



- 相関係数が有意な因子ほど同じクラスタにまとまりやすい傾向にある。
- クラスタ間の相関はほとんど有意でない。
 - クラスタAとCは有意な相関。Ward法では, 次のステップでこれらが結合。

物理量をいくつか減らしてみると・・・

- U500, V500, T500, Q500, Z200, Z850を除外(32因子)
- クラスタ結果はほとんど同じ(Q850cが移動したのみ)。

32因子

U200c U850c V200c V850c T850c
 Q850c Z500c SLPc OLRc 降水c

T850b Q850b Z500b T地表b SSTb

U200b Q300b Q300c T地表c 顕熱c
 SSTc

T200b T200c OSRc 雲量b 雲量c
 潜熱c

U850b SLPb OLRb OSRb 降水b

43因子

U200c U500c U850c V200c V500c
 V850c T850c Q500c Z500c Z850c
 SLPc OLRc 降水c

T500b T850b Q850b Z200b Z500b
 T地表b SSTb

U200b U500b T500c Q300b Q300c
 Q500b Z200c T地表c 顕熱c SSTc

T200b T200c Q850c OSRc 雲量b
 雲量c 潜熱c

U850b Z850b SLPb OLRb OSRb
 降水b

他の季節 (K-mean法)

DJF (n=5)

U200c U500c U850c V200c
 V500c V850c T500c Z500c
 Z850c SLPc Q500c Z200c
 U200b Q300b Q300c
 Q500b

T500b T850b Q850b Z200b
 Z500b

T850c T地表b SSTb T地表c
 顕熱c SSTc

T200b T200c OSRb 雲量b

U850b SLPb OLRb 降水b
 U500b Q850b Z850b OLRc
 OSRc 雲量c 潜熱c 降水c

MAM (n=6)

U200c U500c U850c V200c
 V500c V850c T500c Z500c
 Z850c SLPc Z200c

U200b Q300b Q300c
 Q500b 顕熱c U500b T200c

T500b T850b Q850b Z200b
 Z500b

T200b T850c T地表b SSTb
 T地表c OSRb OSRc SSTc

Q500c Q850c OLRc 雲量c
 潜熱c 降水c

U850b SLPb OLRb 降水b
 Z850b 雲量b

SON (n=6)

U200c U500c U850c V200c
 V500c V850c T500c Z500c
 Z850c SLPc OLRc Q50c
 潜熱c 顕熱c 降水c

U200b Q300b Q300c
 Q500b U500b T200c Z200c

T500b T850b Q850b Z200b
 Z500b T地表b SSTb

T200b Q850c OSRc 雲量b

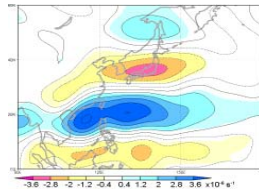
T850c Z850b T地表c 雲量c
 SSTc

U850b SLPb OLRb 降水b
 OSRb

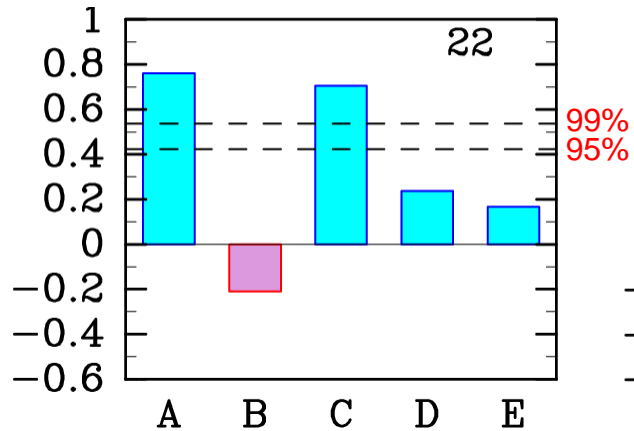
- 最適クラスタ数や分かれ方は異なってくるものの、類似点も多い。

経年変動再現性指標と気候値再現性クラスタの相関

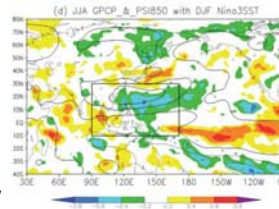
PJパターン



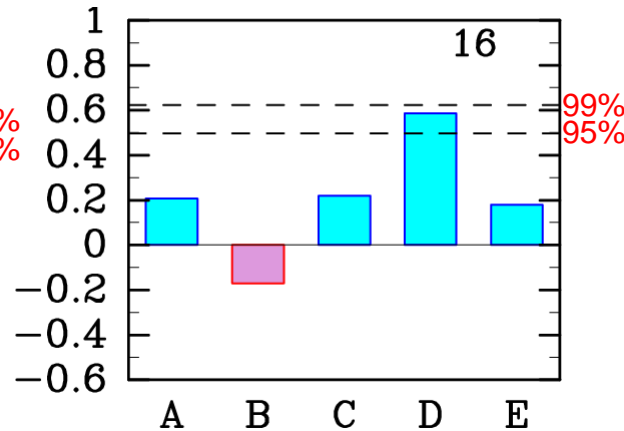
(Kosaka et al.)



Nino3SSTに
対する降水応答



(Ose et al. 2009, JMSJ)



注) 予稿集は熱帯気候値を用いたクラスタ

A U200c U500c U850c V200c
V500c V850c T850c Q500c
Z500c Z850c SLPc OLRc
降水c

B T500b T850b Q850b Z200b
Z500b T地表b SSTb

C U200b U500b T500c Q300b
Q300c Q500b Z200c SSTc
T地表c 顕熱c

D T200b T200c Q850c OSRc
雲量b 雲量c 潜熱c

E U850b Z850b SLPb OLRb
OSRb 降水b

- 経年変動パターンの再現性指標(S-5-2で開発)とクラスタ平均との相関。
- 現象により、有意な正相関を示すクラスタが異なる。
- 注)今回は全球気候場。これから、領域をきちんと設定して開発する予定。

まとめ

- 物理量間で気候モデルの誤差統計量に見られる有意な正相関(≡情報の冗長性)を軽減する試みを行った。
- クラスタ解析の結果, 物理的・力学的に解釈可能なクラスタに分けることができた。
 - 22という高次元なため, クラスタリングが困難という懸念もあった(≡「次元の呪い」)
- 2種類のクラスタ解析手法では結果がほぼ一致し, 因子を減らした実験でも結果はほぼ一致。
 - これだけではrobustnessは言えないとはいえ, 手法の妥当性を示唆している。