

気候モデルのバイアス検出と補正

田中賢治・萩澤佑樹*・佐久間良一**・小尻利治

* 京都大学工学部

** 京都大学大学院工学研究科

要 旨

本研究では各機関から提供される地域気候モデル(RCM)出力値に含まれるバイアス情報を効率的に検出・補正する手法の開発を行う。個々のグリッド毎にバイアス情報を算定しても情報が多くなりすぎる上に、検証情報の関係で評価不可能な格子が出てくるため、水系別あるいは都道府県別といった領域単位でモデル評価を行う。バイアス検出では各気象要素の月平均値のみならず頻度分布を算出し、各階級別にモデルバイアスを評価する。バイアス補正では、領域別の月平均値をできるだけ観測値と一致させることに留意しつつ、極値(頻度分布の両側の裾野部)の再現性を高めるべく、各階級別に補正係数を調整するアルゴリズムを考案した。

キーワード: 地域気候モデル, バイアス検出, バイアス補正, 頻度分布, 極値

1. はじめに

地球温暖化に対する危機感が実感のものとして感じられるようになった昨今、将来気候の具体的な変動予測についての社会的ニーズはより高まってきている。2007年2月に発表されたIPCC(気候変動に関する政府間パネル)第4次評価報告書は、国内の報道でも大きく取り上げられ、地球温暖化に対する危機意識はわが国でも次第に高まりつつある。

地球温暖化の影響評価および対策検討のため、地域レベルでの信頼性の高い気候変化予測情報が必要となったのをきっかけに、2007年度から気候変動予測に関する国家的な2つのプロジェクトが発足した。文部科学省による「21世紀気候変動予測革新プログラム」と、環境省による「地球温暖化に係る政策支援と普及啓発のための気候変動シナリオに関する総合的研究」である。前者は過去5年間に地球シミュレータを用いた地球温暖化予測で実績を上げた「人・自然・地球共生プロジェクト」の後継として予測そのものの改良・高度化を図ること、後者は気候モデルの温暖化将来予測結果の総合的な解析、地域気候モデルなどの利用による空間詳細化、また社会経済シナリオの空間詳細化等を利用して、気候モデルの予測結果を社会に「実感」可能な情報として翻訳す

ることを目的としている(江守, 2007)。

これまで、モデル開発者側からデータ利用者側へほとんど一方的な情報の流れで、必ずしも情報交換が十分ではなく、例えば日射量など影響評価に必要な要素が出力されていないということがあった。これまでのプロジェクトに見られた問題点を克服すべく、モデル側と利用者側の情報交換を促進し橋渡しをすることが、著者らが参画している上記環境省プロジェクトのサブ課題S-5-3「温暖化影響評価のためのマルチモデルアンサンブルとダウンスケーリングの研究」の重要な役割となる。

近年の気候モデルの発展は目覚ましいものがあるが、それにも増して、気候予測情報の利用者側の求める解像度や精度も高まり、モデルが実現できる精度と利用者側の求める精度の間には依然ギャップが存在する。本研究では各機関から提供されるRCM出力値に含まれるバイアス情報を効率的に検出・補正する手法の開発を行う。モデル格子数や格子位置が変更された場合にも迅速にモデルの性能評価を行えるように、できるだけ汎用的なシステムを構築することを目指す。こういったバイアス検出・補正システムにより、利用者にとって有用な情報が得られるだけではなく、モデル開発者がモデルの改良・調整を効率的に進めることも可能となる。

2. 使用したデータおよびデータ処理

2.1 モデル出力値

環境省地球環境研究総合推進費 S-5-3 では気象研、防災科研、筑波大の各機関が 20km 解像度 RCM による力学的ダウンスケーリングを実施しており (Table 1 参照), 本研究では, サンプルデータとして 2002 年 1 月から 2004 年 12 月 (計 36 ヶ月) の再現計算結果の提供を受けた。サンプルデータには日射, 風速等様々な地上気象要素が含まれるが, 本研究では降水量と気温のデータのみを使用した。

Table 1 Name of each RCM in this study

institution	Name of RCM
MRI (Meteorological Research Institute)	NHM-MRI
NIED (National Research Institute for Earth Science and Disaster Prevention)	RAMS-DP
University of Tsukuba	RAMS-TU

2.2 検証データ (観測値)

モデル検証用データとして AMeDAS 毎時データ (気象データベース・アメダス, 制作: ウェザートイ, 発売元: (財) 気象業務支援センター) の気温, 降水量を使用した。観測地点数は年により異なる。全ての観測地点番号を分析した結果, 2044 地点 (ただし, 観測サイトの移動や観測要素の変更に伴う番号の付け替えも含む) が抽出された。データ解析を容易にするために, 一旦 2044 地点全ての時刻に欠測値を与え, データを順次読み込み, 有効データ (欠測ではない値) が見つければ置き換えるという作業を通じて, 全ての年に共通なデータセット (観測年に依存せず統一的に扱えるデータセット) を作成した。

2.3 地理データ (マスクデータ)

流域・非集水域メッシュデータ W07_52M (国土数値情報 <http://nlftp.mlit.go.jp/ksj/>) の県別の水系コード情報, 北海道の支庁コード情報 (気象研・内山博士より提供) から, 各 RCM のモデルグリッドで表現した水系別マスク, 都道府県別 (北海道は支庁別) マスクを計算する。水系コードは一級河川を含む単一水系域 (001~199), 主要二級河川を含む単

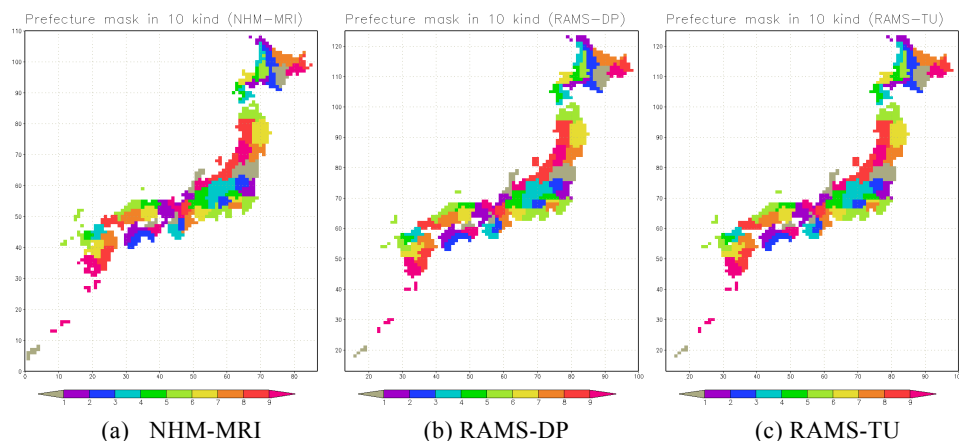


Fig.1 Prefecture unit mask for each RCM grids (60 categories)

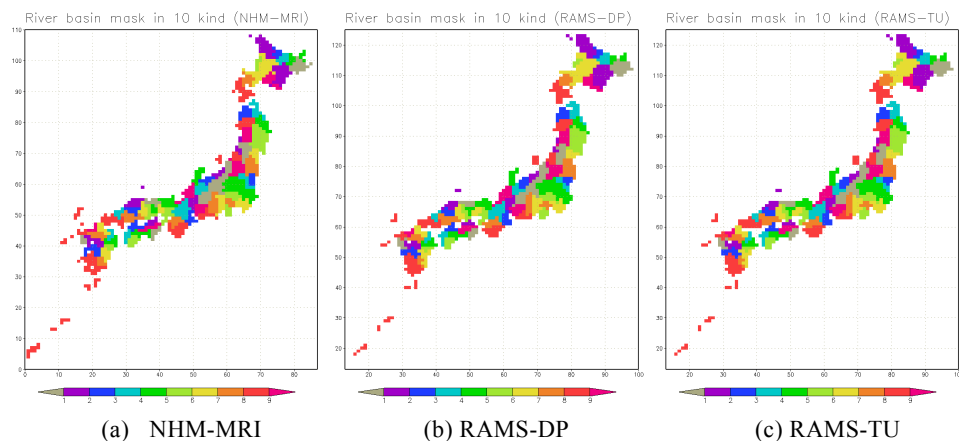


Fig.2 River basin unit mask for each RCM grids (78 categories)

一水系域（201～499）、複合水系域（501～949）、流出口のない水系域（951～999）で区分されているが、201 以降のコードについては隣接する（あるいは再近隣の）一級水系に集約する。ただし、小さな水系は最低でも 5 メッシュが含まれるようにさらに集約した結果、78 水系に区分された。都道府県マスクについては 46 都府県に北海道の 14 支庁を加えて 60 区分に分類される。各 RCM の解像度は同じであるが、図法の違い、基準経度の違いにより、格子中心位置が少しずつずれ、その結果、各モデル格子がどの水系（県）に分類されるかが微妙に異なる (Fig. 1, Fig. 2)。できるだけ汎用性をもたせるため、基本的にモデル格子中心座標の情報のみからこれらマスクデータを作成できるシステムを構築した。

2.3 モデル格子と観測点情報との対応付け並びに観測データの並べ替え

アメダス観測点の位置情報、各 RCM のモデル格子中心座標情報から、検証データの存在する各観測点がどのモデル格子に最も近いかを判定し、観測点とモデル格子を対応づける。また 1 つのモデル格子に対応する観測点が複数存在する場合には (Fig. 3 参照)、有効データに関して平均値を算出し、モデル値と対応させるべき観測値を 1 つ用意する。各モデル格子に最も近い観測点を 1 つ探索するのではなく、各観測点に対応するモデル格子を探索するため、観測点の配置状況によってはどの観測点とも対応しない格子が出てくることになる。本研究では県別あるいは水系別といったある領域単位でモデルバイアスは評価されるため、評価対象にならなかった格子のバイアス情報は同一領域内の他の格子から算定されたバイアス情報と同様であると仮定することでバイアス補正を行う。なお、気温に関しては、モデル標高と観測点標高の差の分を補正する（観測値をモデル標高気温に補正する）。最後に、モデルバイアスを評価する単位（都道府県別、水系別）毎に各観測値を並べ替え、1 次元配列に格納する。

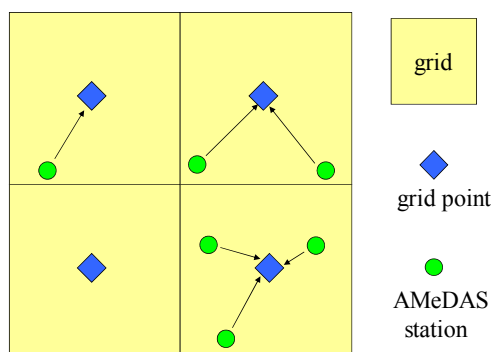


Fig.3 Corresponding grid for each observation station

2.4 RCM 出力の並べ替え

モデルバイアスを評価する単位（都道府県別、水系別）毎に各RCM出力（地上2次元場）各格子点の値を並べ替え、1次元配列に格納する。これにより、各モデル出力のフォーマットが統一されるとともに、海上の値など、評価対象外の情報を落とせる。

3. モデルバイアスの検出

3.1 月平均降水量と降水強度頻度分布

各RCMグリッドにおけるモデル値と観測値が、共に欠測でない時の1時間降水量のみを扱い、これらを各月ごと各エリア別に、サンプルランである2002年～2004年の3年間分の平均値を算定する。降水量の場合、モデル値の平均月降水量を観測値の平均月降水量で割った比を月降水量のバイアスとする。

頻度分布の算出においても平均月降水量の比較と同様に、モデル値と観測値が共に欠測でない時の1時間降水量を 1 個の標本として数える。各月ごとに、1エリア内の全てのグリッドがもつ標本（3年分）を集めた。1ヶ月が30日のとき、各月ごとの標本の大きさは、1エリア内に1つのグリッドしか標本をもっていない場合であれば2160(24x30x3)、9つのグリッドが標本をもっている場合だと19440(24x30x3x9)になる。これら1時間降水量を、降水強度別に階級分けして頻度分布に表す。階級ごとの境界値は、1.0mm/hr, 2.0mm/hr, 5.0mm/hr, 10mm/hr, 15mm/hr, 20mm/hr, 30mm/hrの7つを設定し、合計8つの階級に分けることを基本とするが、月別、エリア別の最大降水強度に応じて、頻度分布の階級を調整する。頻度分布の階級の基本の境界値を $clev_i(i=1\sim7)$ とし、最大降水強度 (P_{max}) に応じて調整された境界値を $clev_i(i=1\sim7)$ とすると、

$$\begin{aligned} P_{max} < 6 & \quad clev_i = 0.2clev_i \\ 6 < P_{max} < 15 & \quad clev_i = 0.5clev_i \\ 15 < P_{max} < 60 & \quad clev_i = 1.0clev_i \\ 60 < P_{max} & \quad clev_i = 2.0clev_i \end{aligned}$$

以上のように設定された降水強度階級を用いて、観測値とモデル出力のそれぞれについて頻度分布を計算する。さらに、極端現象の再現性を評価するために、90%、99%、99.9%の3つの非超過確率降水量（上位10%、1%、0.1%降水量）を算定する。

3.2 月平均気温と気温頻度分布

降水量の場合と同様に、各 RCM グリッドにおけるモデル値と観測値が、共に欠測でない時刻の気温のみを扱い、これらを各月ごと各エリア別に、サン

ランである 2002 年～2004 年の 3 年間分の平均値を算定する。気温の場合、モデル値の月平均気温から観測値の月平均気温を差し引いた値を月平均気温のバイアスとする。気温の頻度分布を求める際の階級については、月によりまた緯度帯により、適切なレンジが異なるため、モデル値と観測値の全標本の中の最大値 (Tmax) と最小値 (Tmin) を 10 階級に分割することにした。頻度分布の階級の基本の境界値を $clev_i (i=1\sim9)$ とすると、

$$clev_i = \text{int}(Tmin) + \text{real}(i) \times \frac{\text{real}(\text{int}(Tmax) - \text{int}(Tmin))}{9}.$$

以上のように設定された気温階級を用いて、観測値とモデル出力のそれぞれについて頻度分布を計算する。さらに、両側（低温側と高温側）の極端現象の再現性を評価するために、0.1%, 1%, 10%, 90%, 99%, 99.9% の 6 つの非超過確率気温（上位 10%, 1%, 0.1% の低温と高温）を算定する。

3.3 モデルバイアス検出結果

Fig.4 は各 RCM の 2 月と 7 月の平均降水量のバイアス(比)を都道府県別に示したものであり、モデル降水量が過大なエリアには青色が、過少なエリアには赤色がつけられている。なお、黄色いエリアはモデル出力が観測値に対して $\pm 10\%$ 以内に収まっている場所である。Fig.5 は各 RCM の 2 月と 7 月の平均気温のバイアス(差)を都道府県別に示したものであり、高温バイアスのエリアには赤色が、低温バイアスのエリアには青色がつけられている。なお、黄色いエリアはモデル出力が観測値に対して $\pm 0.5K$ 以内に収まっている場所である。

これらの図からわかる通り、モデルバイアスは季節的にも空間的にも大きくばらつく。ただし、各 RCM はまだ調整段階であり、使用したデータはサンプルデータであることに注意されたい。本研究ではこのように様々なモデル出力についてモデルバイアスを効率よく検出するシステムを構築することを目的としているため、ここでは個々のモデルについてバイアスの特徴を詳しく述べることはしない。

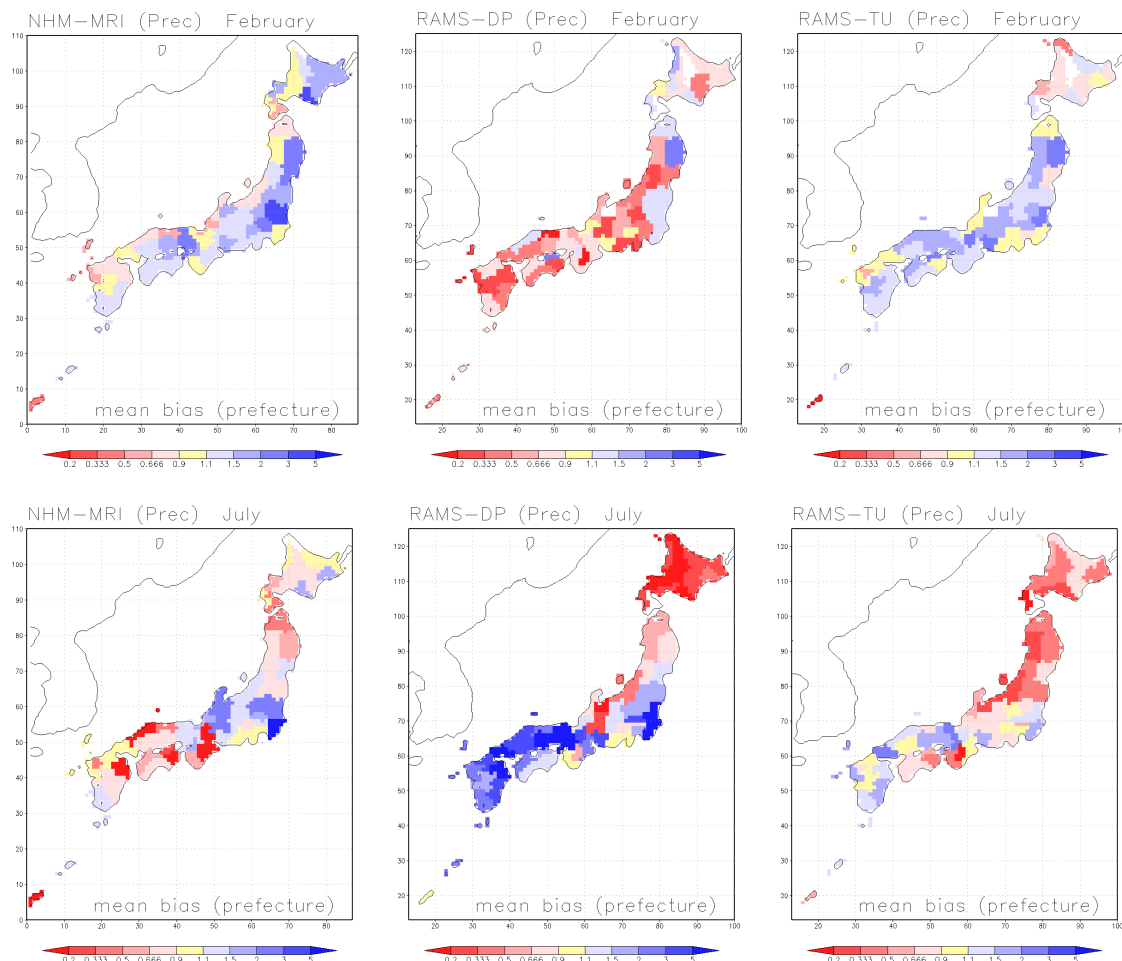


Fig.4 Monthly mean precipitation bias by prefecture unit for each RCM (upper: February, lower: July).

Too much (wet) area is blue, dry area is red. Small bias (ratio is close to 1) is yellow.

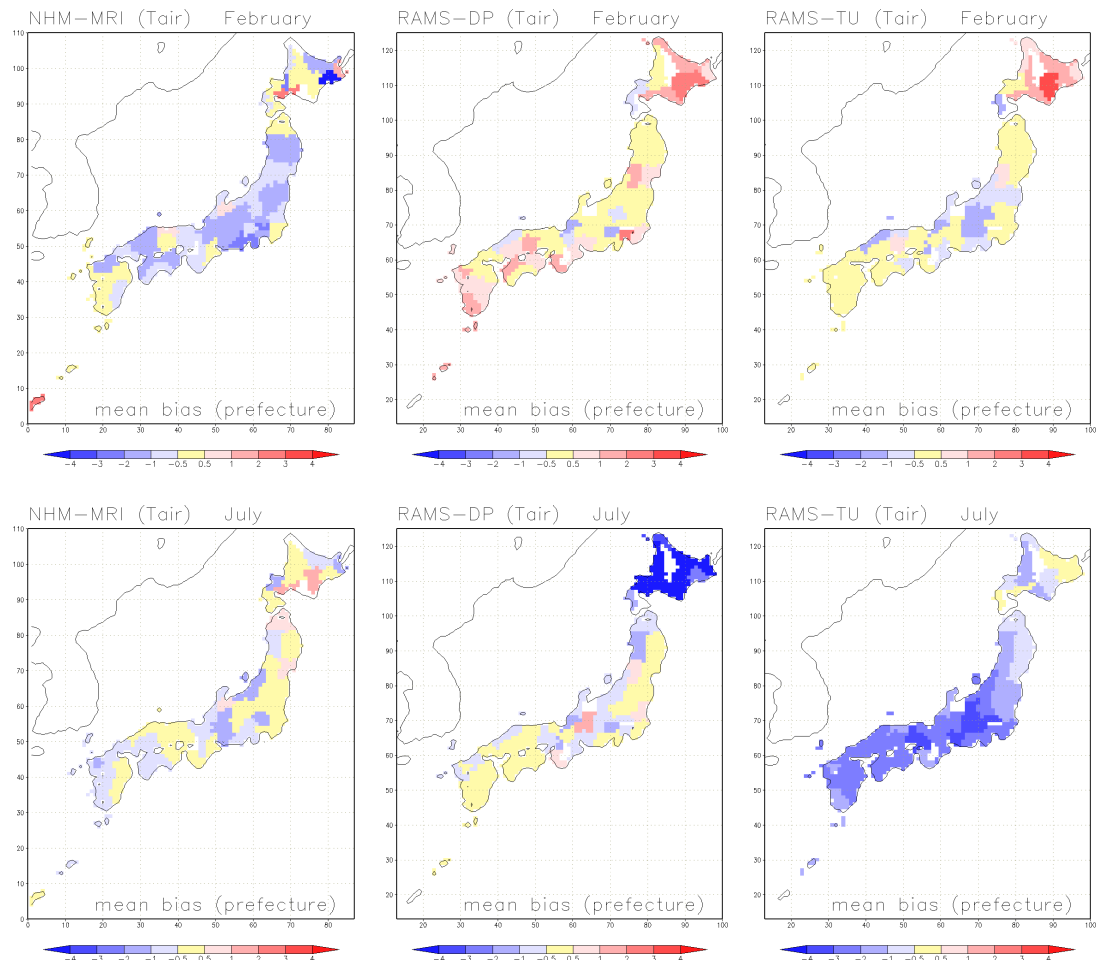


Fig.5 Monthly mean temperature bias by prefecture unit for each RCM (upper: February, lower: July).
Too high (warm) area is red, cold area is blue. Small bias (difference is $\pm 0.1K$) is yellow.

3.4 エリア別の頻度分布の評価

降水強度と気温の頻度分布の例をそれぞれ Fig. 6, Fig. 7 に示す。黒点は観測値、青線はモデル出力であり、簡易な補正として月平均バイアスで補正をかけた結果も緑線で示されている。Fig. 6 の左図は NHM-MRI の 2 月の秋田県の場合であるが、観測値では 10mm/hr 以上の降水強度が含まれているが、モデル出力では 5~7.5mm/hr の階級までしか頻度分布が存在しない。Fig. 4 の左上図でも示されている通り、2 月の秋田県の降水量バイアスはほとんどなく、月平均値だけを見ていると問題がないように思えるが、このように頻度分布をとることで、モデル出力の問題点が明確になる。また、このように月平均値が合っている場合には、簡易補正をかけてもほとんど補正がかからず、頻度分布もほとんど改善されない。Fig. 6 の中図は NHM-MRI の 7 月の福島県の場合であり、全体的に頻度が低いものの、30mm/hr 以上の階級まで分布が広がっており頻度分布の形状は観測値に近い。また、この場合、月平均値の過少バイアスの情

報で補正をかけることで、頻度分布もある程度改善されている。Fig. 6 の右図は NHM-MRI の 7 月の宮崎県の場合であり、30mm/hr 以上の階級まで頻度分布が広がるべきところモデル出力では 5~10mm/hr の階級で終わっている。この場合も簡易補正で少しは頻度分布が改善されるが、実際の頻度分布には程遠い。

Fig. 7 の左図は NHM-MRI の 7 月の静岡県の場合であり、モデル出力値の頻度分布は両側の裾野の広がりも含めてかなり精度が良いが、簡易補正をかけることで、高温側が改善され、低温側はやや悪くなっている。Fig. 7 の中図は NHM-MRI の 7 月の大分県の場合であり、Fig. 5 の左下図でも示されている通り、月平均気温のバイアスはほとんどないが、頻度分布の裾野の広がりとは全く異なっており、簡易補正をかけても頻度分布は全く改善されていない。Fig. 7 の右図は RAMS-TU の 7 月の石川県の場合であり、大きな低温バイアスが見られる。簡易補正をかけることで頻度分布が全体に高温側にシフトするが、その結果低温の極値の再現性がかえって悪くなっている。

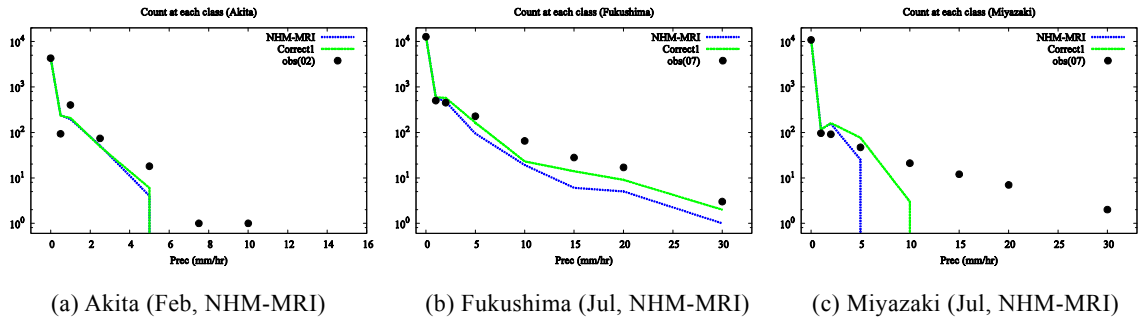


Fig.6 Example of frequency distribution of precipitation intensity (blue line: model output, green line: corrected value by monthly mean bias, dot: observation)

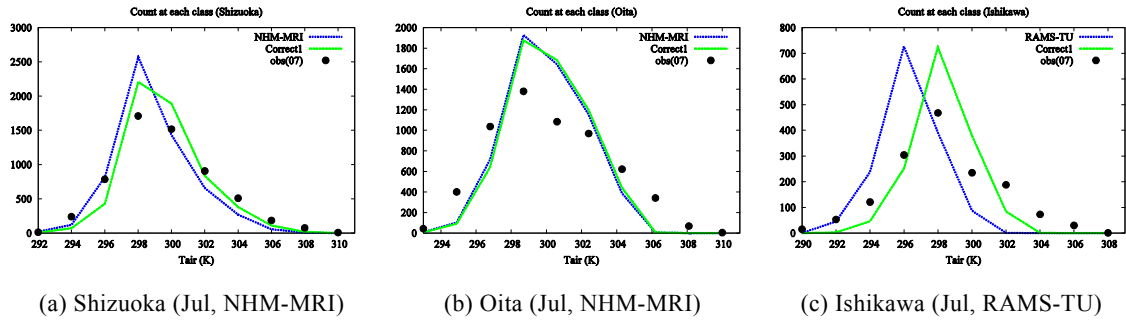


Fig.7 Example of frequency distribution of precipitation intensity (blue line: model output, green line: corrected value by monthly mean bias, dot: observation)

以上，いくつかの例で示した通り，頻度分布も考慮した形でモデル出力値を補正する手法を開発する必要がある。

4. 頻度分布を考慮したモデルバイアス補正

簡易な補正法として，まずは月平均値を合わせることがよくされる。本研究ではさらに極値（頻度分布の両側の裾野部）の再現性を高めるべく，各階級別に補正係数を調整する手法の開発に着手した。

頻度分布を考慮した補正法では，月平均値バイアスを解消する平均の補正係数をまず各階級に一律に適用し，頻度分布がどのように修正されるかをあらかじめ求めておく（仮補正）。次に，各階級の補正係数をある割合で変化させた場合，その前後の階級の頻度分布がどのように変化するかをあらかじめ算定しておき，月平均値を保存させるという制約条件のもと各階級の補正係数を様々変化させ，頻度分布のピーク位置や極値の再現性の高い組み合わせを順次探索していくというものである。

4.1 頻度分布調整アルゴリズム

(a) 最適比率の設定

まず，本補正のアルゴリズムで用いる比率について説明する。最適比率 (optimalratio) は，仮最適

比率 (provisional optimal ratio) を微調整したものであり，この仮最適比率とは，全階級に対して一律な値である固定比 (fixed ratio) と，各階級によって異なる値である変動比 (variable ratio) により構成されている。これらの関係を図で表したものが Fig.8 である。それぞれの役割を簡単に述べるとすれば，固定比は月平均値を観測値に近づけるための比率であり，変動比は階級頻度分布を観測値に近づけるための比率である。固定比は，月平均値バイアスを解消するものであるため，降水量（比）の場合はその逆数（観測値の平均月降水量／モデル値の平均月降水量），気温（差）の場合は符号を逆にしたもの（観測値の月平均気温－モデル値の月平均気温）となる。

続いて，本補正のアルゴリズムに変動比を導入した動機について触れる。簡易補正によってモデル値が向上した点は，平均月降水量が観測値と等しくなった点である。つまり，全体的な補正としての効力はあるが，階級ごとにおける部分的な補正としては効果が無い。この簡易補正の欠点を補うものとして，本補正では変動比を用いるのである。

降水量（比）の場合の変動比 (α) は，固定比 ($w0$) に対して 0.1 倍から 0.9 倍まで変化させるように設定する。

$$\alpha = 0.1w0 \times i \quad (i=-9 \sim 9)$$

気温（差）の場合は 0.5K 単位で-4K から 4K までの範囲で設定する。

$$\alpha = 0.5 \times i \quad (i = -8 \sim 8)$$

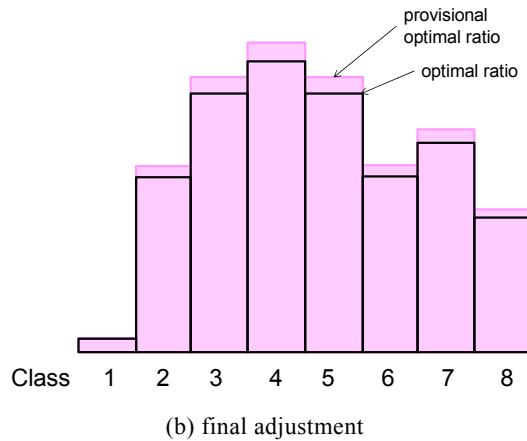
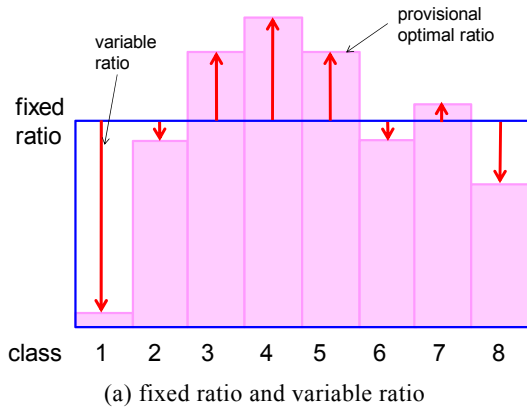


Fig.8 Schematic image of optimal ratio

(b) 本補正の判断基準

本補正によるモデルバイアス改善度の指標として、以下の値を用いる。

$$dvar = \sum_{i=1}^8 (mvar_i - ovar_i) a_i \quad (1)$$

以降では、この値 $dvar$ を補正誤差と呼ぶこととする。この補正誤差は、各階級別の降水量（気温）に重み a_i をつけた上での、モデル値と観測値の階級別合計量の差であり、0 に近づくほど、その補正の精度が高まっていることを示す。

頻度分布の裾野部分（降雨強度が強い階級）の階級の補正に重点を置き、補正誤差算定における重み (a_i) を次のように設定することで、極値の再現回数の差が小さくなるようにした。

$$\begin{aligned} a_i &= clev_i & (clev_i > 1) \\ a_i &= 1 & (clev_i \leq 1) \end{aligned}$$

気温についても同様に設定すると、かえって頻度分布のピーク階級がずれてしまう場合が出てくるので、気温については全ての階級で $a_i=1$ とした。

(c) 各変動比による変化量

本項では、各変動比によるモデルバイアス変化量の検討方法について述べる。本補正は後述する単位操作を繰り返し、各変動比がモデルバイアスに及ぼす変化量を予め調べておくことで、次項で記述する補正誤差最小化のための処理を助けている。記述を簡便にするために、降水量のみにして記す。

モデル値に関して、変動比を変化させることにより、本補正後のモデル値がどの階級に移り、元の階級の合計降水量と移動先の階級の合計降水量にどのような変化をもたらすかを1つの単位操作とする。

Fig. 9 は、この単位操作が行われる判断過程をフローチャートにしたものである。欠測でないモデル値に、簡易補正と本補正をそれぞれ行うことにより、本補正後の階級が簡易補正後の階級と等しければ単位操作 1 を、等しくなければ単位操作 2 を行う。変動比を変化させながら、この図で示されている単位操作までの一連の流れを、各月・各エリアにて、欠測でない全ての降水量で行う。単位操作の式は、以下に示す。

単位操作 1

$$store(j, ic0, ic1) = \sum \{c(j) \cdot prc\} \quad (2)$$

単位操作 2

$$store(j, ic0, ic1) = \sum \{(-f) \cdot prc\} \quad (3)$$

$$store(j, ic0, ic2) = \sum \{(c(j) + f) \cdot prc\} \quad (4)$$

ただし、

j : 変動比の数

$ic0$: 元のモデル値が属する階級

$ic1$: 簡易補正後に属する階級

$ic2$: 本補正後に属する階級

$c(j)$: j 番目の変動比

f : 固定比

prc : モデル値の1時間降水量

$store(j, ic0, ic1)$: j 番目の変動比によって、階級 $ic0$ から階級 $ic1$ へ移動した降水量

$store(j, ic0, ic2)$: j 番目の変動比によって、階級 $ic0$ から階級 $ic2$ へ移動した降水量

これら単位操作によって出された、変動比 $c(j)$ ごとの階級 k から階級 l への移動した降水量 $store(j, k, l)$ を、次項で述べるアルゴリズムにて駆使する。

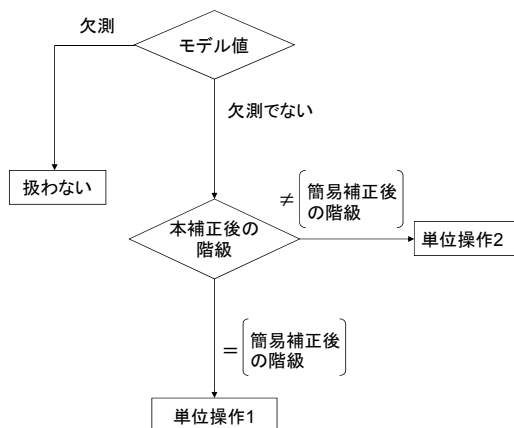


Fig.9 Flowchart of unit correction

移動先の階級 1

	1	2	3	4	5	6	7	8	
元の階級 k									
1									
2									
3									
4									
5									
6									
7									
8									
和	dif(1)	dif(2)	dif(3)	dif(4)	dif(5)	dif(6)	dif(7)	dif(8)	

Fig.10 calculation of dif

移動先の階級 1

	1	2	3	4	5	6	7	8	和
元の階級 k									
1									del(1)
2									del(2)
3									del(3)
4									del(4)
5									del(5)
6									del(6)
7									del(7)
8									del(8)

Fig.11 calculation of del

移動先の階級 1

	1	2	3	4	5	6	7	8	和
元の階級 k									
1									
2									
3									
4									
5									
6									
7									
8									
和									ds

Fig.12 calculation of ds

1	-0.14	-80.68	0	0	0	0	0	0	0	-80.68
2	-0.14	55.43	-118.42	0	0	0	0	0	0	-62.99
3	-0.14	0	141.69	-302.68	0	0	0	0	0	-161
4	-0.14	0	0	-17.17	-27.84	0	0	0	0	-45
5	-0.14	0	0	0	-18.38	0	0	0	0	-18.38
6	-0.14	0	0	0	16.61	-20.76	0	0	0	-4.15
7	-0.14	0	0	0	0	0	0	0	0	0
8	-0.14	0	0	0	0	0	0	0	0	0

1	0.14	80.68	0	0	0	0	0	0	0	80.68
2	0.14	-76.28	139.27	0	0	0	0	0	0	62.99
3	0.14	0	-100.84	261.84	0	0	0	0	0	161
4	0.14	0	0	-63.31	108.31	0	0	0	0	45
5	0.14	0	0	0	-36.19	54.58	0	0	0	18.38
6	0.14	0	0	0	0	4.15	0	0	0	4.15
7	0.14	0	0	0	0	0	0	0	0	0
8	0.14	0	0	0	0	0	0	0	0	0

Fig.13 Example of the table of store(j,k,l) and del(k) (upper: $\alpha = -0.14$, lower: $\alpha = +0.14$)

(d) 補正誤差最小化のためのアルゴリズム

前項までに設定した値を使って、本研究のメインテーマである、モデルバイアス最小化のためのアルゴリズムを紹介する。このアルゴリズムは、4.1(a)で設定した複数の変動比を使って、それぞれの階級でどのような変動比を固定比に加えれば、4.1(b)に

示した補正誤差を最小化できるのかを特定する。

(1) 最初に、簡易補正後のモデル値について各階級の合計降水量を求め、それらを式(1)の $mvar$ に代入したものを、補正誤差の初期値と定める。元のモデル値でなく、簡易補正後のモデル値に関するモデルバイアスを初期値として求めるのは、簡易補正によ

って平均月降水量が観測値とほぼ等しい値になっているからという理由と、一律の固定比をまず全階級にかけておくことによって、 n 階級に属する降水量が $(n-2)$ 階級に移動しているのに、 $(n-1)$ 階級に属する降水量が $(n+1)$ 階級に移動するという好ましくないケースが起こるのを防ぐという理由がある。

(2) 4.1(c) で予め導き出していた $store(j, k, l)$ を、Fig. 10 のようにして移動先の階級 l ごとに足し合わせて、これを dif とする。この dif (各変動比による移動先の階級ごとの合計降水量変化) を、式 (1) で用いた $mvar$ (各階級の合計降水量) に各々加えていき、式 (1) によって再び補正誤差 ($dvar$) を算出する。また、補正誤差を算出するときには、両隣の階級とその階級の変動比の差の絶対値和が 1.0 を越えないような変動比をとることで、(1) で挙げた降水量の好ましくない階級移動を防ぐと共に、計算時間の短縮を図っている。

(3) さらに $store(j, k, l)$ を用いて、 ds (各変動比による元の階級ごとの合計降水量変化の和) を導出する。まず、Fig. 11 のように $store(j, k, l)$ を元の階級 k ごとに足し合わせて、その和を $del(k)$ (各変動比による元の階級ごとの合計降水量変化) とする。さらに、この $del(k)$ を Fig. 12 で表されるように全ての階級で足し合わせたものを ds とする。Fig. 13 は $store(j, k, l)$ と $del(k)$ の例である。

(4) (2) で算出した補正誤差が初期値よりも小さくなっており、かつ、(3) で導出した ds が全階級合計降水量の 3 パーセント未満であるときに、この補正誤差を新たな補正誤差として、それ以降の他の補正誤差と比較する対象にする。

(5) (2) から (4) の繰り返しによって、最も補正誤差が小さくなったときに、その補正誤差を導き出した各階級の変動比を最適な変動比の組み合わせとし、これらの最適な変動比と固定比を足し合わせたものを、各階級の仮最適比率とする。

(6) 最後に、この比率をかけたモデル値の平均月降水量と観測値の平均月降水量の比をとり、微調整として、その逆比をアルゴリズムで導き出した仮最適比率にかけることによって本補正の最適比率とする。なお、(3) で「 ds が全体量の 3 パーセント未満」という条件を設けたのは、この最後の微調整が「微調整」の範囲を超えてしまわないようにするためである。なぜなら、仮に ds が全階級合計降水量の中でかなりの重みを占めてしまったら、最後の微調整によって、アルゴリズムで改善された頻度分布が悪化してしまう可能性があるからである。

以上の本補正におけるアルゴリズムの全体的な流れをフローチャートとして、Fig. 14 に示す。

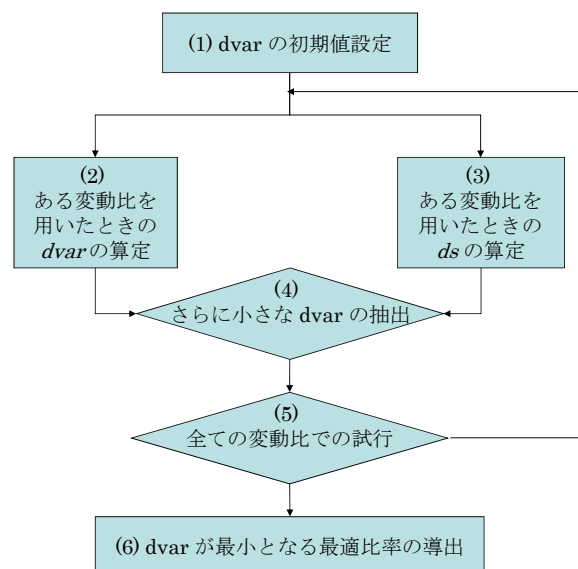


Fig.14 Flowchart of bias correction system

4.2 頻度分布を考慮した補正法の適用結果

3.4 で示した例について、頻度分布を考慮した補正法を適用した結果をそれぞれ Fig. 15, Fig. 16 に示す。いずれの場合も簡易補正に比べて格段に頻度分布形状が改善されていることが見てとれる。ただし、図 6 の右図のように元のモデル値の分布があまりにもかけ離れている場合には、実際の頻度分布に近づけることは困難である。補正係数の探索範囲を広げればもう少し頻度分布の裾野を右側に引っ張ることはできるが、その場合莫大な計算時間を要することになり、また、あまりに過剰な補正を施すことは「補正」の範疇を越える作業でもある。モデル出力自身がどの程度まで頻度分布を再現できれば極値の再現性も含めて補正が可能になるのかといったこと（モデル改良のためのガイドラインのようなもの）も今後検討していく予定である。

4.3 極値の再現性について

Fig. 17 は、各都道府県別の 99% 非超過確率（上位 1%）降水強度の散布図であり、上図は 2 月に関するもの、下図は 7 月に関するものである。図中に相関係数と RMSE の値も示されている。

2 月については、頻度分布を考慮した補正により CC と RMSE が大幅に改善されていることが分かる。7 月に関しても改善はされているものの、まだ十分な値とは言えない。現在のアルゴリズムでは元のモデル出力値の精度がもう少し上がらないと、たとえ補正をかけたとしても、影響評価研究に耐える精度で極値を再現することはできない。興味深いことに、RAMS-TU の 2 月の例（右上図）では、月平均値による簡易補正をかけると、元のモデル出力値よりも極値の再現性が悪くなる場合があるということである。

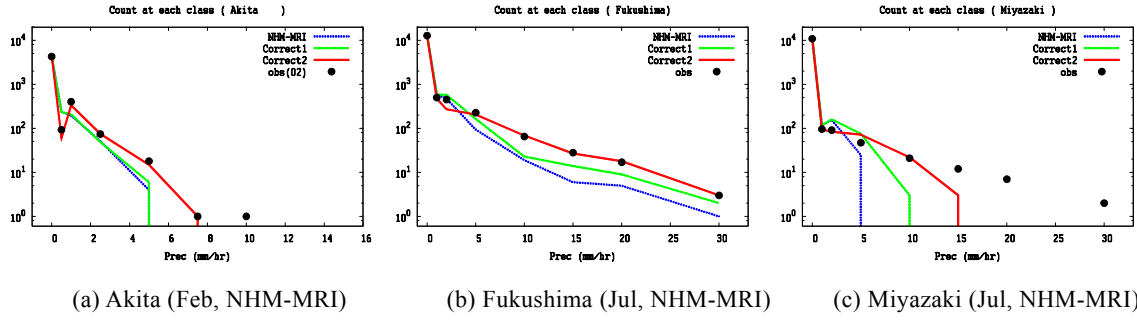


Fig.15 Example of frequency distribution of precipitation intensity (blue line: model output, green line: corrected value by monthly mean bias, red line: corrected value considering frequency distribution, dot: observation)

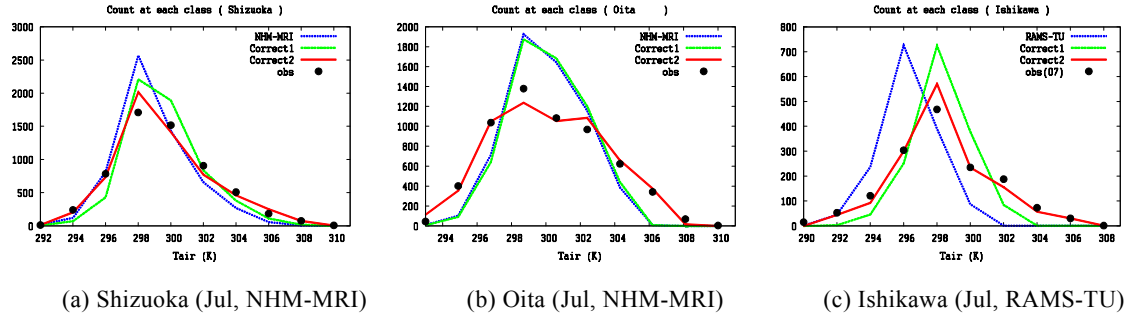


Fig.16 Example of frequency distribution of precipitation intensity (blue line: model output, green line: corrected value by monthly mean bias, red line: corrected value considering frequency distribution, dot: observation)

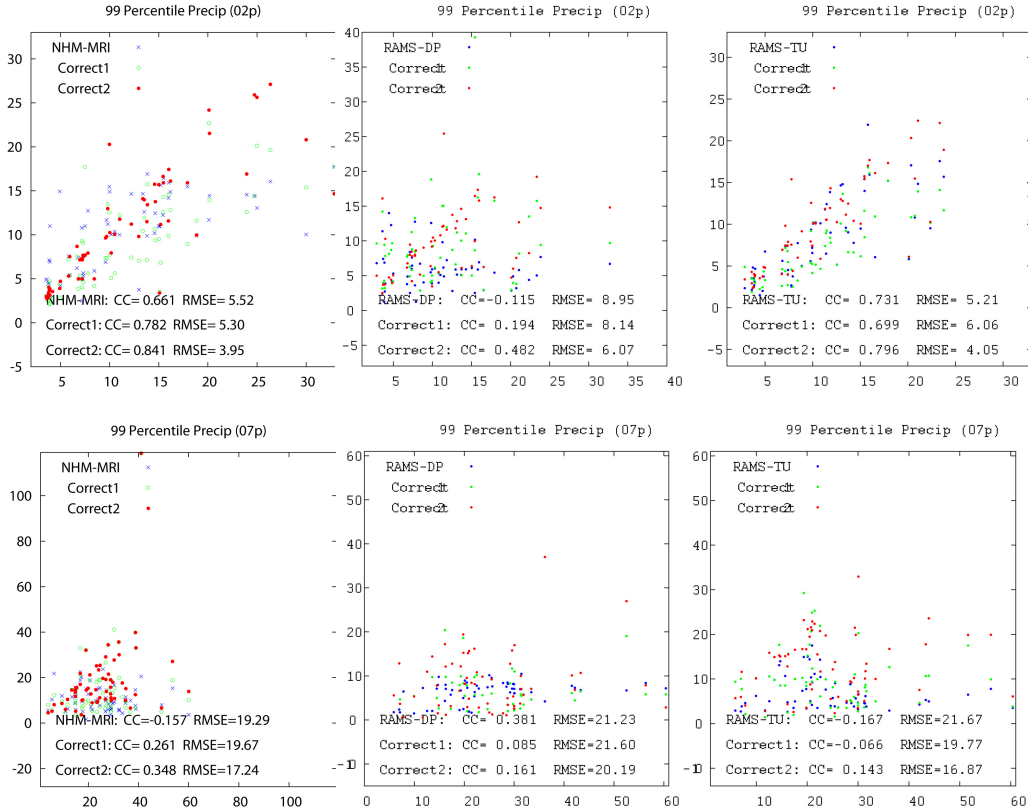


Fig.17 Scatter diagram of extreme precipitation (99percentile) for each RCM output and bias corrected value (blue: model output, green: corrected value by monthly mean bias, red: corrected value considering frequency distribution, upper:February, lower: July)

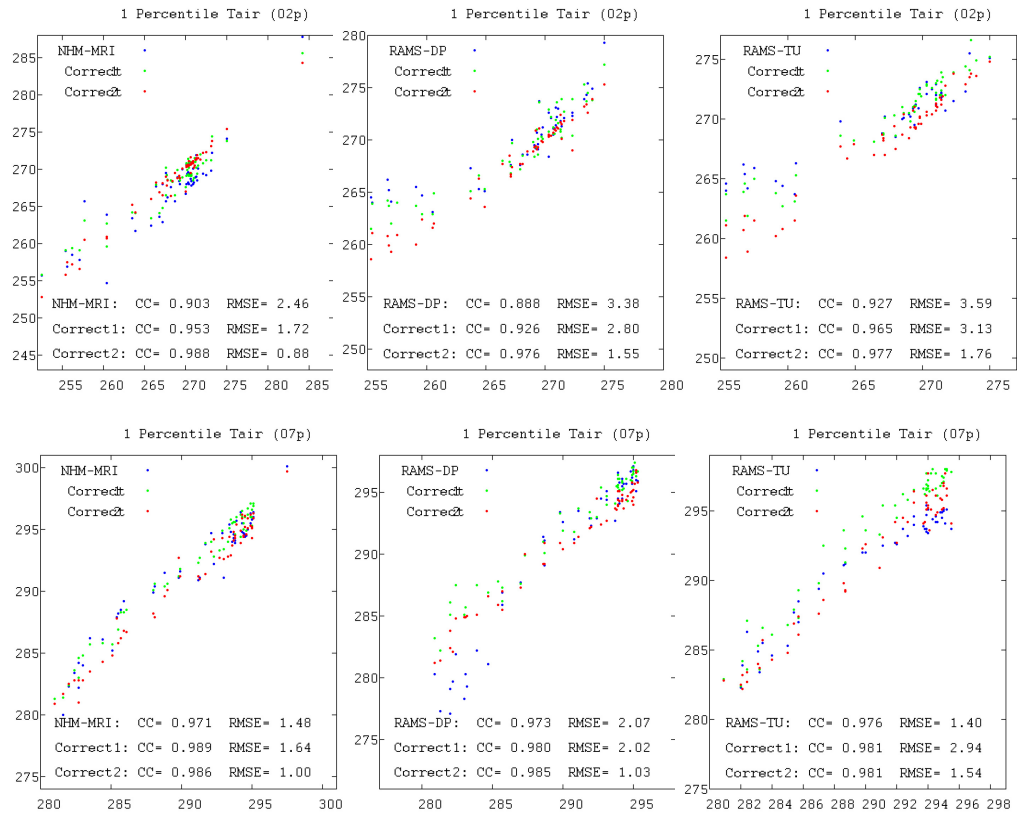


Fig.18 Same figure as Fig.17 but for low temperature extreme (1 percentile)

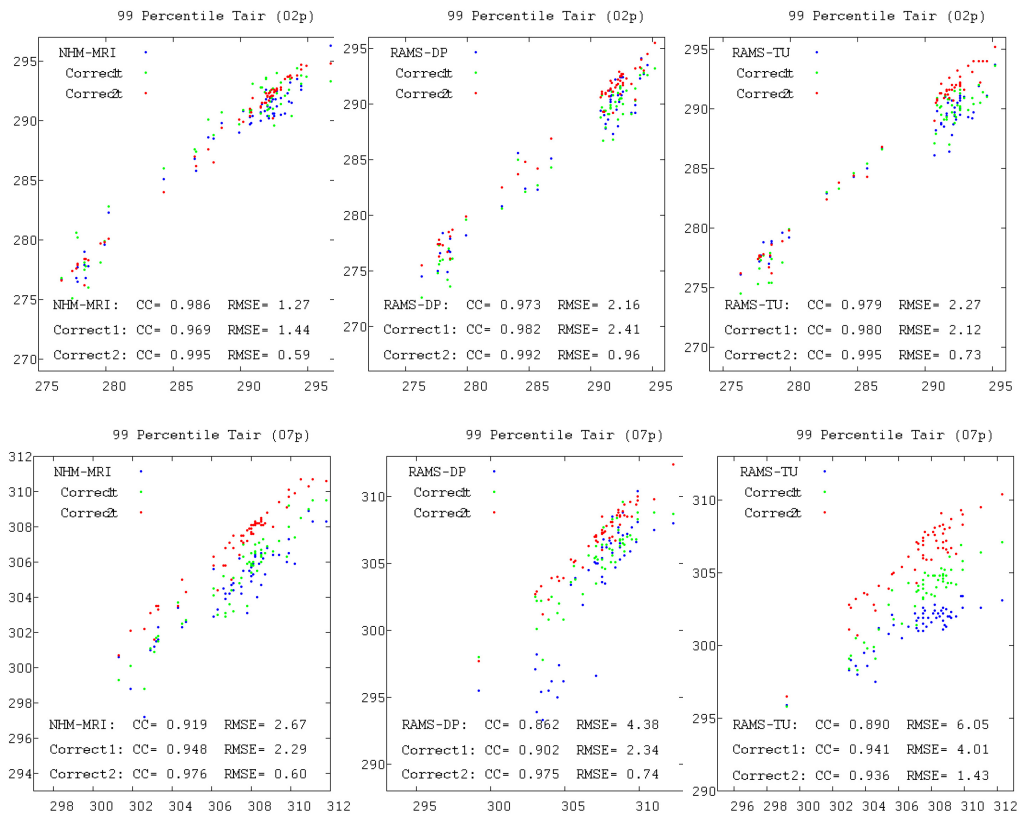


Fig.19 Same figure as Fig.17 but for high temperature extreme (99 percentile)

Table 2 Input and Output of bias detection/correction system

プログラム : GSI/basmask.f

入力データ

W07_52M_*.txt (**=01~47) 流域・非集水域メッシュ
code_Lon1-Lon2E_Lat1-Lat2N.dat (北海道の支庁別マスクデータ)
../RCM/comp/orgMODEL.txt (MODEL=NHM-MRI, RAMS-DP, RAMS-TU) モデル格子中心座標

出力データ

prefecarea.txt 各都道府県に属する RCM グリッド数 (NHM-MRI, RAMS-DP, RAMS-TU)
basinarea.txt 各水系に属する RCM グリッド数 (NHM-MRI, RAMS-DP, RAMS-TU)
gridpMODEL.gad 各 RCM 格子の都道府県別マスク
gridrMODEL.gad 各 RCM 格子の水系別マスク

プログラム : RCM/comp/stnmatch.f

入力データ

addramd.txt アメダス観測点の位置, 標高情報
../GSI/prefecarea.txt 各都道府県に属する RCM グリッド数
../GSI/basinarea.txt 各水系に属する RCM グリッド数
../GSI/gridpMODEL.gad 各 RCM 格子の都道府県別マスク
../GSI/gridrMODEL.gad 各 RCM 格子の水系別マスク
orgMODEL.txt モデル格子中心座標
../AMEDAS/amedas1hrYYYY.gad (YYYY=2002~2004) アメダスデータ

出力データ

gridmatchMODEL.gad 各 RCM 格子に属する観測点数
maskorderMODEL.gad 各 RCM 格子を 1 次元配列に並べ替える場合の順位
amedas/topoMODEL.gad (A=r,p) 評価するエリア別に並べ替えた一次元配列の観測点標高データ
amedas/VARMODELYYYYA.gad (VAR=Prec, Tair) 評価するエリア別に並べ替えた一次元配列の観測データ

プログラム : RCM/comp/readmodel.f

入力データ

../GSI/prefecarea.txt 各都道府県に属する RCM グリッド数
../GSI/basinarea.txt 各水系に属する RCM グリッド数
../GSI/gridpMODEL.gad 各 RCM 格子の都道府県別マスク
../GSI/gridrMODEL.gad 各 RCM 格子の水系別マスク
gridmatchMODEL.gad 各 RCM 格子に属する観測点数
maskorderMODEL.gad 各 RCM 格子を 1 次元配列に並べ替える場合の順位
降水量, 気温のモデル出力値(MM=01~12)
../NHM_Ver0/YYYYMM/VAR_105x115.dat (VAR=smqr, smqs, smqg, tsfs)
../RAMSD/VER1/YYYY_ps/MM/1hr/VAR (VAR=precipr, tempf2m)
../RAMST/1mon/S5V0MM.YYYYMM.grd

出力データ

model/VARMODELYYYYA.gad (VAR=Prec, Tair) 評価するエリア別に並べ替えた一次元配列のモデル出力値

プログラム : RCM/comp/biaspdfprec.f, biaspdfair.f

入力データ

../GSI/prefecarea.txt 各都道府県に属する RCM グリッド数
../GSI/basinarea.txt 各水系に属する RCM グリッド数
../GSI/table-prefec.txt 各都道府県区分の名前
../GSI/table-river.txt 各水系の名前
model/VARMODELYYYYA.gad (VAR=Prec, Tair) 評価するエリア別に並べ替えた一次元配列のモデル出力値
amedas/VARMODELYYYYA.gad (VAR=Prec, Tair) 評価するエリア別に並べ替えた一次元配列の観測データ

出力データ

gnu/MODEL/pct-statA.txt (A=r,p) エリア別の非超過確率降水強度の統計値
gnu/MODEL/percentileMMA.txt (MM=01~12, A=r,p) 月別エリア別の非超過確率降水強度
gnu/MODEL/weightMMA.txt (MM=01~12, A=r,p) 月別エリア別の平均バイアスと階級別補正係数
gnu/MODEL/biaspdfMM-NNA.txt (MM=01~12, NN=01-60) 頻度分布, 階級別総降水量, 積算相対頻度

プログラム : meanbias.f

入力データ

../GSI/gridpMODEL.gad 各 RCM 格子の都道府県別マスク
../GSI/gridrMODEL.gad 各 RCM 格子の水系別マスク
gnu/MODEL/weightMMA.txt (MM=01~12, A=r,p) 月別エリア別の降水量の平均バイアスと階級別補正係数
gnuT/MODEL/weightMMA.txt (MM=01~12, A=r,p) 月別エリア別の気温の平均バイアスと階級別補正係数

出力データ

meanbiasMODEL-1dA.gad (A=r,p) 各 RCM の月平均値のバイアス (1 次元)
meanbiasMODEL-2dA.gad (A=r,p) 各 RCM の月平均値のバイアス (2 次元)

Fig. 18, Fig. 19 は、それぞれ各都道府県別の 1%, 99% 非超過確率気温（上位 1% の低温，上位 1% の高温）の散布図であり，上図は 2 月に関するもの，下図は 7 月に関するものである。いずれの図においても，頻度分布を考慮した補正により CC と RMSE が大幅に改善されていることが分かる。例えば RAMS-TU の 7 月の 99% 非超過確率気温 (Fig. 19 の右下) では，比較的大きな低温バイアスが存在するものの，気温に関しては元のモデル出力の頻度分布形状も比較的滑らかで観測値から極端に外れていないため，本研究で提案したアルゴリズムがうまく機能していると言える。

5. おわりに

本研究では気象研，防災科研，筑波大の各機関からサンプルデータとして提供を受けた 2002 年から 2004 年の再現計算結果をモデル出力値，AMeDAS 毎時データを検証データとして，各 RCM 出力値に含まれるバイアス情報を効率的に検出・補正するシステムの開発に着手した。

モデル格子数や格子位置が変更された場合にも迅速にモデルの性能評価を行えるように，できるだけ汎用的なシステムとなるように心掛けた。個々のグリッド毎にバイアス情報を出しても情報が多くなりすぎる上に，検証情報の関係で評価不可能な格子も出てくるため，水系別あるいは都道府県別といった領域単位でモデル評価を行うこととした。

バイアス検出では各気象要素の月平均値のみならず頻度分布を算出し，各階級別にモデルバイアスを算定するとともに，様々な例を通じて，月平均値のみに基づく補正では，特に極値の再現性という意味で不十分であることを指摘し，頻度分布も考慮した形でモデル出力値を補正する手法を開発する必要があることを示した。

バイアス補正では，領域別の月平均値をできるだけ観測値と一致させることに留意しつつ，極値(頻度分布の両側の裾野部)の再現性を高めるべく，各階級別に補正係数を調整するアルゴリズムを考案し，試験的にバイアス補正を試みた。降水量においては必ずしも十分に補正アルゴリズムが機能していなかったが，用いたデータがわずか 3 年と短く，サンプル数が少ない関係で頻度分布形状が不自然になっていた

ことも考えられるので，今後はさらに長期間のデータを用いてアルゴリズムの改良を検討していきたい。

謝 辞

本研究は，環境省の地球環境研究総合推進費 (S-5-3) の支援により実施されました。

参考文献

- 江守正多 (2007) : 地球環境研究センターニュース Vol. 18, No. 3.
- 田中賢治，渡辺裕司，相馬一義，池淵周一 (2005) : 領域気象モデルによる降水量プロダクトの定量的な利用にむけて，水文・水資源学会 2005 年研究発表会要旨集，pp. 314-315.
- Tanaka, K., T. Kojiri, Y. Fujihara (2006) : Bias correction of the meteorological variables from RCM for hydrological application, The Advance Report of ICCAP, pp.43-46.
- Tanaka, K., T. Kojiri, Y. Fujihara (2007) : Bias correction of the meteorological variables from RCM for hydrological application, XXIV General Assembly of IUGG, MS016-60.
- Kenji Tanaka, Kazuaki Yorozu (2007) : Selecting the best meteorological forcing dataset for each river basin, Eos Trans. AGU, 88(52), Fall Meet. Suppl., Abstract H33C-1455.

付 録

本研究で提案したバイアス検出補正システムは以下のプログラムで構成される。各プログラムの入出力リストを Table 2 にまとめた。

1. basmask.f : 水系別マスク，都道府県別マスクデータの作成
2. stnmatch.f : モデル格子と観測点情報との対応付け並びに観測データの並べ替え
3. readmodel.f : RCM 出力の並べ替え
4. biaspdfprec.f, biaspdfair.f :
モデルバイアスの検出，頻度分布計算
頻度分布を考慮したモデルバイアスの補正

Bias Detection and Correction of Climate Model Output

Kenji TANAKA, Yuuki HAGIZAWA*, Ryoichi SAKUMA** and Toshiharu KOJIRI

* Faculty of Engineering, Kyoto University

** Graduate School of Engineering, Kyoto University

Synopsis

This study aims to develop the method for detecting and correcting the bias information of RCMs. This system is designed to be general so that it can follow the change in model setting. Bias information for each individual grid is too much for handling, and some grids may have no evaluation data. Therefore, model bias is evaluated in each river basin or prefecture unit. Not only the monthly mean value, but also frequency distribution is evaluated for each unit area. A new bias correction method that can reproduce the extreme values while keeping the monthly mean value is proposed. In most cases, this new method works well in getting better frequency distribution and extreme values.

Keywords: Regional Climate Model, bias detection, bias correction, frequency distribution, extreme value