

GPUを用いた気候変動影響評価における全球規模 トレンド分析の高速化

山本 昭夫†

喜連川 優‡

†東京大学地球観測データ
統融合連携研究機構‡国立情報学研究所
東京大学生産技術研究所

1. はじめに

近年、地球温暖化に起因するとみられる気候変動とそれがもたらす地球環境への影響が大きな問題として社会的に関心を集めている。気候変動に関する議論では、将来の変化予測が不可欠である。気候変動の特徴を把握し、その影響評価を行うには、様々な物理量の長期変化傾向（トレンド）を見積もる必要があるが、長期間のデータを取得することができ、小さな変化であっても統計的に有意なトレンドを検出可能な気候モデルによる数値シミュレーションが有効な手段となる。

将来の気候変化特性の研究において、従来は全球平均されたデータの変動に関する議論が多く行われてきたが、最近の研究によると、全球平均気温は温暖化傾向にある場合でも、ある地域では寒冷化しているなど、地域・季節間の違いが小さくないことが明らかになってきている。また、気候特性を考える上で重要な降水量においては、地域ごとにトレンドに増加、減少が顕著に異なり、地域ごとの評価が不可欠である[1]。こうした地域や季節による違いを検出するには、気候モデルの全ての格子点ごとに特定の時間範囲においてトレンド分析を行う必要があるが、全球平均と比較して対象となるデータ量が増加するため、実行時間が長くなってしまいうという問題がある。

本稿では、全球規模トレンド分析の高速化を目的として、GPU (Graphics Processing Unit) を用いたトレンド分析の実行時性能評価を行い、その有効性を検証した結果について述べる。

2. トレンド分析

気候変動による影響評価では、数十年から百年程度の期間において、気温などの物理量の経年変化にトレンドがあるかどうかを調べるのが重要である[2]。ここではトレンドを見積もる方法およびその統計的有意性検定の方法について簡潔に説明する。

トレンドの見積もりに用いることが多い線形回帰分析は、時系列の開始または終了付近にある周囲のデータから大きく外れた値の影響を受けやすいため、よりロバストな手法を用いるべきである[3]。また、分析対象の変動が非正規である場合、変動の分布関数に関する仮定が必要ないノンパラメトリックな方法が有効である。パラメトリックな方法に比べ検出力が劣るという欠点があるが、気候データ等の変動の分布関数が未知の場合には有効な方法である。

本研究ではノンパラメトリック法の一つである Mann-Kendall 検定を用いて、ある時系列の各要素が独立で同一の確率分布に従う、すなわち「時系列にトレンドが存在しない」という帰無仮説が成立するか否かを検定する。帰無仮説が棄却されると、その系列は傾向変動を持つと判定される。具体的な手順は以下の通りである[3]。

サンプル数 N の時系列に対して、次の式により検定統計量 τ を定義する。

$$\tau = 4 \left(\sum_{i=1}^{N-1} n_i / [N(N-1)] \right) - 1 \quad (1)$$

ここで、 n_i は対象となる時系列において、 i 番目の要素値よりも大きい値が i 番目より後ろにいくつあるかを示したものである。 τ が正（負）の時には、その時系列は増加（減少）傾向にある。

τ の統計的有意性を調べるためには、 τ の値を次式で定義される τ_g と比較する。

$$\tau_g = \pm t_g \sqrt{(4N+10)/9N(N-1)} \quad (2)$$

t_g は与えられた有意水準における標準正規分布表の値である。検定すべき仮説（帰無仮説）は「時系列のトレンドはない」であり、 $|\tau| > |\tau_g|$ の時に帰無仮説は棄却されて「時系列のトレンドは統計的に有意である」との結果が得られる。

また、検定の結果、トレンドが統計的に有意となった時系列に対して、任意のデータ対 (x_j, x_k) ($1 \leq k < j \leq N$) から次式(3)により定義される傾斜推定量 β (Sen's slope estimator) [4] を算出する。

GPU acceleration of global trend analysis in climate change impact assessment

† Akio Yamamoto, Earth Observation Data Integration and Fusion Research Initiative, The University of Tokyo

‡ Masaru Kitsuregawa, National Institute of Informatics / Institute of Industrial Science, The University of Tokyo

$$\beta = \text{median} \left(\frac{x_j - x_k}{j - k} \right), \quad \forall k < j \quad (3)$$

これは時系列の増減の大きさを表す推定値として用いられ、 $\beta < 0$ の場合、時系列には減少トレンドがあることを示し、逆に $\beta > 0$ となった時は増加トレンドがあることを示す[5].

3. GPUによる高速化

気候モデルによる数値シミュレーション結果の一つの格子点に着目すると、その格子点の時系列に対するトレンド分析は、他の格子点における時系列の分析処理とは独立である。すなわち、全球規模トレンド分析は並列度が極めて高く、並列処理による高速化が大いに見込まれる。

本研究では、GPUを用いた超並列処理による実行時間の短縮を目標とする。筆者らが全球規模トレンド分析の OpenMP による並列化[6]で検討したように、気候モデル出力のデータ構造である 3次元配列を、トレンド分析で扱う格子点の時系列に連続的かつ効率的にアクセス可能となるよう再配置した配列データに対し、次の2つの方式で実装を行う。

3.1. 空間並列化

式(1)から(3)の計算を単一のカーネル関数として作成する。ホストからカーネル関数を起動する際、ブロック数を緯度方向の格子点数、スレッド数を経度方向の格子点数として空間方向、すなわち緯度・経度面における処理の並列化を行う。

3.2. 時空間並列化

式(1)、(2)で定義される検定統計量に基づく統計的有意性検定と、式(3)による傾斜推定量の算出を行う計算モジュールをそれぞれ個別のカーネル関数として作成する。前者のカーネル関数 Stats は、ブロック数を全格子点数、スレッド数を時系列の長さ $N-1$ で起動し、式(1)に示した時系列における i 番目の要素に対する検定統計量の算出を各スレッドに割り当て並列実行する。

また、後者のカーネル関数 Slope では中央値を求めるためのソートに GPU 上に実装した Bitonic sort を利用することにより、時間方向処理の並列化をさらに進める。

4. 性能評価

本研究の性能評価では、WCRP/PCMDI の CMIP データセット[7]のうち、格子点数 320×160 の気候モデル MIROC3.2-HIRES の出力データを用いる。100年間(2001~2100年)にわたる地表気温の年平均値時系列を対象として、Intel Core i7-3770K および NVIDIA GeForce GTX690 (Kepler GK104) から成るシステム上で CUDA[8]により実装

した全球規模 Mann-Kendall トレンド分析プログラムの実行時間を計測し、その性能評価を行う。

4.1. 結果

再配置前の 3次元配列データを対象とした CPU 逐次処理による全球規模トレンド分析の実行時間は 756.98 秒であった。この時間を基準値として、GPU による高速化の効果を検証した。

空間並列化の実行時間は 48.14 秒であるのに対し、時間方向処理の並列化を加えると実行時間は 3.709 秒まで短縮され、CPU 逐次処理と比較して約 204 倍の大幅な速度向上が得られた。実行時間のうちカーネル関数については Stats が 0.930 秒、Slope が 1.307 秒となり、ホスト・デバイス間のデータ転送時間は 0.565 秒であった。

5. おわりに

本稿では、気候変動予測結果を対象とした Mann-Kendall 検定に基づく全球規模トレンド分析の高速化を目的として、GPU を用いた超並列トレンド分析の実行時性能評価について述べた。

主要な計算モジュールを GPU/CUDA により個別のカーネル関数として実装することで極めて高速な分析処理が可能となり、対話的に操作できる気候変動影響評価システムを開発および構築するためのコア技術を確立したものと見える。

謝辞 本研究は文部科学省委託事業「地球環境情報プラットフォーム構築推進プログラム」の支援を受けたものである。

参考文献

- [1] 小林, 宝, 中北: 全球気候モデル出力を用いた日本域の 100 年確率日降水量の将来予測, 水工学論文集, 54, pp.223-228, 2010
- [2] 西澤: 気象観測データの長期トレンドの統計解析, 応用数理, 18 (3), pp.27-38, 2008
- [3] 松山, 谷本: 実践! 気候データ解析【第二版】, 古今書院, 2005
- [4] Sen, P. K. : Estimates of the Regression Coefficient Based on Kendall's Tau, Journal of American Statistical Association, 63 (324), pp.1379-1389, 1968
- [5] 徐, 竹内, 石平: 日本の平均気温・降水量時系列におけるジャンプ及びトレンドに関する研究, 水工学論文集, 46, pp.121-126, 2002
- [6] 山本, 喜連川: 気候変動予測における全球規模トレンド分析の OpenMP による並列化に関する検討, 情報処理学会第 77 回全国大会講演論文集, 1A-01, 2015
- [7] CMIP - Coupled Model Intercomparison Project, <http://cmip-pcmdi.llnl.gov/>
- [8] CUDA C Programming Guide, NVIDIA, 2016