

Jh130037-DA02

地殻変動連続データを用いた スロースリップイベント高精度自動検出アルゴリズムの研究

木村 武志 (防災科学技術研究所)

概要

西南日本下のプレート境界上の大地震発生域周辺では、ゆっくりすべりという現象が発生している。ゆっくりすべりの効果的検出手法の開発は、プレート境界全体のすべり過程の理解に重要である。本研究では、地殻変動連続データからゆっくりすべりを高精度に検出するために、独立成分分析の適用について検討した。九州大学情報基盤研究開発センターの Fujitsu PRIMERGY CX-400 上で MATLAB および Fortran それぞれによるシステム開発を行った。データのノイズ特性・適用するデータ長等に対する独立成分分析の性能評価を行うため、大量の人工傾斜変動データの解析をした。その結果、ランダムウォークノイズの影響により抽出した独立成分のみによるゆっくりすべり検出は困難であるが、同時に推定した応答行列を評価に組み込むことで安定して検出出来ることが分かった。また、適用するデータ長が短いほど精度良く検出出来ることが分かった。

1. 研究の目的と意義

研究の目的

日本列島下のプレート境界付近で発生するゆっくりすべり (Slow Slip Event, SSE) を、広域・高密度に展開する高感度地殻変動観測網データから高精度で自動検出する解析基盤を開発する。これまでに申請者らは、複数の SSE の震源モデルを予め仮定し、多観測点の傾斜変動データと整合するモデルを探索することによって、SSE を自動的に検出する手法を開発している (Kimura et al., 2011) 。本研究では独立成分分析 (Independent Component Analysis, ICA) などの信号処理・データマイニング手法を適用することにより、この手法を高度化し、SSE の検出能力の向上および推定される震源モデルの高精度化を目的とする。それにより、これまで大地震を引き起こしてきたプレート境界におけるすべり特性の時空間分布の特徴を明らかにする。

研究の意義

日本列島の地下では陸側プレートの下に海洋プレートが沈み込んでいる。これらのプレート境界では、2011年東北地方太平洋沖地震を始め、マグニチュード 8 を越える大地震が繰り返し発生している。近年の高感度地殻変動観測網の展開やデータ解析手法の発達により、西南日本のプレート境界では、過去に大地震が発生した領

域の深部延長で通常地震とは異なる SSE という現象が発生していることが分かってきた。SSE は通常地震と比べ震源でのすべりがゆっくりと進行するため、通常地震計では検知が困難であり、地殻変動データを用いた検出・解析が行われている。SSE の発生を多数の地殻変動データから効果的に検出することにより、その発生状況の正確な把握が可能になり、これまで大地震を起こしてきたプレート境界におけるすべり特性などの状態に対する理解が進むことが期待できる。

防災科学技術研究所が展開する高感度地震観測網 (Hi-net) に併設されている高感度加速度計によって得られる地殻の傾斜変動データは、(1) カバーしている空間範囲が広いこと、(2) 観測点密度が高いこと (約 20km 間隔)、(3) SSE などの地殻活動による変動を高感度に捉えられること、等の利点がある。SSE が数ヶ月～半年間隔で繰り返し発生している四国地方から紀伊半島・東海地域にかけては、約 100 観測点の傾斜データ (各観測点では南北・東西方向の 2 成分のデータ) が 20Hz サンプルングで取得されている。これまでに、Kimura et al.[2011] では、SSE の震源モデルを予め仮定した自動検出アルゴリズムを提案し、四国地域の傾斜データを対象とした SSE の準リアルタイム自動検出システムを試験的に運用し

ている。しかしながら、これらの傾斜データには潮汐や気圧変化、降雨などの影響により、SSEなどの地殻活動とは関係のない変動成分が重畳しているため、上述のアルゴリズムでは誤検出や取り逃しが発生してしまっている。FA（因子分析）やICAは、異なる時空間分布特性を持つ現象をアダプティブに分離・抽出することが出来るため、これらの手法を傾斜データに適用することにより、SSEによる信号の検知能力の向上や震源モデル推定の高精度化が期待できる。

このようにして蓄積されたSSEの発生履歴や高精度SSE震源モデルは、プレート境界における大地震の発生機構の解明にとって重要な基礎データとなる。また、日本国内には傾斜データ以外にもGPSやひずみ計などの地殻変動観測体制も整備されており、そのような大規模多様なデータ（ビッグデータ）を並列計算機上で高速かつ適切に解析可能な基盤技術の研究は、地球科学の諸分野、大規模工学システムの監視、社会経済データ解析からのリスク評価、システムログデータ解析による電気機器の最適制御・コミュニティのスマート化など、広範な学問領域に展開可能なものである。

2. 当拠点公募型共同研究として実施した意義

(1) 共同研究を実施した拠点名および役割分担

共同研究を実施した拠点名

国立大学法人九州大学

共同研究の役割分担

- ・木村武志（防災科研）
地殻変動データ準備・地殻活動のモデル推定
 - ・徳永旭将（高度情報科学技術研究機構→統数研）
データ解析アルゴリズムの開発・実装
 - ・中村和幸（明治大学）
時空間データ解析・モデリングに関する情報提供
 - ・深沢圭一郎（九州大学）
九州大学・情報基盤研究開発センターの計算機資源を利用する上での支援
- ### (2) 共同研究分野
- 超大規模データ処理系応用分野

(3) 当公募型共同研究ならではの事項など

ICAなどのブラインド信号分離手法は、多点観測されたデータから潜在変数を見出す手法として脳波解析などの分野で応用されてきた。一方、昨今では様々な分野で高次元の観測データが蓄積されつつあり、それらの中からより本質的な潜在構造を帰納的に見出し、背後にある支配法則の理解やリスクマネジメントに活かす高次元データ駆動型科学の発展が期待されている。このようなデータドリブンの研究では、データ計測やその生成過程に関する専門知見のみならず、情報科学、コンピュータサイエンス、統計科学、計算機科学など、分野の垣根を超えた共同研究実施体制が必要である。本研究では、地震・測地学分野における高次元データ駆動型科学の実問題を解決するために、地震学分野、測地学分野、統計科学分野および計算機科学分野の専門的知見を備えた研究者による分野横断的な共同研究を実施した。特に、高次元データ駆動型の研究では、データの背後にあるダイナミクスを直接的に仮定しないことから、信頼性のある（あるいは解釈可能な）結果を得るためには、非常に多数のパラメータにおいて解析を実施する必要がある、大型計算機の活用が必要である。九州大学情報基盤研究開発センターの提供するFujitsu PRIMERGY CX-400では、1,000を超えるコアが利用可能であり、実験計画に応じた高並列計算が可能である。生命科学分野において100TB規模の大規模データ解析の使用実績もある。また、豊富な数値演算ライブラリを備えたMATLABが、並列計算が可能な状態で利用可能であり、トライアンドエラーのプロセスが不可避なデータ駆動型研究の実施に理想的な環境が提供されているといえる。

3. 研究成果の詳細と当初計画の達成状況

(1) 研究成果の詳細について

本研究は、ICAをHi-net傾斜変動データなどの大規模地殻変動データに適用することによる、SSEなどの地殻活動の検出やモデル推定の高精度化を目的とする。

(1-1) ICA (独立成分分析)

ICA は、観測データに含まれる原信号の独立性を仮定するだけで、原信号そのものはもちろんのこと信号の応答特性に関する情報も未知として、推定する (図 1). 従って、傾斜変動などの測地連続データから地殻活動の信号を抽出する場合には、地殻活動に伴う応答特性、即ち地殻活動のモデルを予め仮定する必要が無い. 観測データに含まれる原信号の統計的独立性を仮定するだけで、原信号を抽出し、それぞれの応答特性をも推定する.

いま、 N 次元の観測信号ベクトル $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T \in \mathbb{R}^n$ と M 次元の原信号ベクトル $\mathbf{s}(t) = (s_1(t), \dots, s_m(t))^T \in \mathbb{R}^m$ について、

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad \dots (1)$$

の関係が成り立つとする. \mathbf{A} は $m \times n$ の定数行列であり、その要素は $\mathbf{A} = (a_{ji}) \in \mathbb{R}$ と記述できるとする. 式(1)は、観測信号は原信号の線形な足し合わせによって表現できるというモデルであり、ICA の観測モデルと呼ばれる. ここで、推定された原信号を $\mathbf{u}(t) = (u_1(t), \dots, u_n(t))^T \in \mathbb{R}^n$ とし、簡単のために \mathbf{A} を正則行列とすると、 \mathbf{u} は次式によって得られる:

$$\mathbf{u}(t) = \mathbf{A}^{-1}\mathbf{x}(t) = \mathbf{W}\mathbf{x}(t) \quad \dots (2)$$

ここで、 \mathbf{A}^{-1} は \mathbf{A} の逆行列を意味する. \mathbf{W} は分離行列と呼ばれ、その要素は $\mathbf{W} = [w_1, \dots, w_n]^T$ と表現できるとする. 式(2)より、ICA は観測信号 $\mathbf{x}(t)$ から、 $\mathbf{u}(t)$ ができるだけ統計的独立となるように、分離行列 \mathbf{W} を逐次更新する手法と捉えることができる.

ICA には数多くの実装があるが、本研究では Hyvärinen and Oja [1999] らによる FastICA を用いる. FastICA では、 $\|w_i\| = 1$ という制約のもとで分離

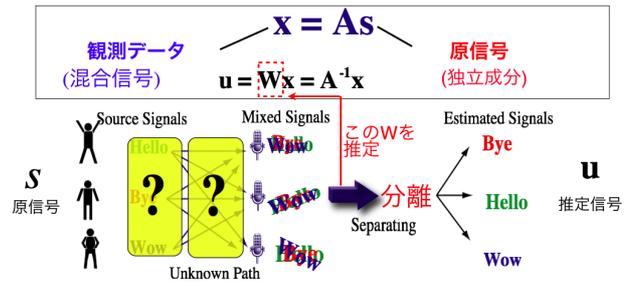


図 1. ICA の観測モデル.

信号の非ガウス性を最大化させる次のような目的関数を導入する;

$$L(w_i) = \{E[G(u_i)] - E[G(v)]\}^2 - \beta \{\|w_i\|^2 - 1\} \quad \dots (3)$$

ここで、 $G(*)$ はコントラスト関数と呼ばれる非 2 次関数であり、 β はラグランジュ係数である. (3) 式から不動点 ($\partial L(w_i)/w_i = 0$) における関係を求めると、次式が導かれる:

$$E[\tilde{\mathbf{x}}g(w_i^T \tilde{\mathbf{x}})] - \beta w_i = 0 \quad \dots (4)$$

ここで、 g は G の導関数を表す. (4) 式にニュートン法を適用することによって、最終的に次のような反復アルゴリズムが導出される;

$$w_i^+ = E[\tilde{\mathbf{x}}g(w_i^T \tilde{\mathbf{x}})] - E[g'(w_i^T \tilde{\mathbf{x}})]w_i \quad \dots (5)$$

$$w_i = w_i^+ / \|w_i^+\| \quad \dots (6)$$

ここで、 w_i^+ は w_i の更新された値を表している.

更新前後の w_i の変化量が十分小さな値となったとき、収束したと判断される.

Hyvärinen and Oja [1999] らの実装では、全ての独立成分を同時に推定する Symmetric 法と、非ガウス性の高い独立成分から順に推定する Deflation 法が提案されている. 本研究では、傾斜データに含まれる SSE 成分が、他の成分 (ノイズや気象擾乱な

ど) よりも非ガウス性が高いという想定のもと、非ガウス性が高い成分から順に推定可能な deflation法を採用することとした。Deflation法においては、次式で表現されるGram-Schmidtの直交化の繰り返し適用により、それまでに推定された分離ベクトルと直交する解空間から新しい分離ベクトルを得る;

$$w_i = w_i - \sum_{k=1}^{i-1} w_k^T w_i w_k \quad \dots (7).$$

(1-2) ICA を用いた SSE 成分抽出の性能評価

Kimura et al. [2011] で使用した四国地域の Hi-net27 観測点について、各南北・東西の 2 成分の人工傾斜変動データを作成した (図 2)。単純な問題について評価を行うため、通常傾斜変動データに含まれている潮汐や気圧応答、背景トレンドは既に十分補正されていると仮定し、①SSE に対する応答、②ランダムウォークノイズ、③ホワイトノイズによりデータが構成されているとした。

SSE 成分は、これまで最も頻繁に SSE が検出されている四国西部下のプレート境界で SSE が発生したと仮定して計算した。深さ約 40 km で 30 km × 30 km の断層面上で Mw 5.6 相当の SSE が発生した時の半無限均質媒質での傾斜変動を応答成分とする。すべり量は断層面上で一様、すべり速度は一定とし継続時間は 100 時間とした。

表 1 に、抽出能力評価において可変としたパラメータとその評価数を示す。傾斜変動などの地殻変動データには、ホワイトノイズ成分だけでなく長周期帯域で卓越するランダムウォークノイズ成分が含まれており、SSE など地殻活動の検出・解析を困難にしている。本研究では、ホワイトノイズ成分 $w(t)$ ($w(t) \sim N(0, \sigma^2)$) ・ランダムウォークノイズ成分 $r(t)$ ($r(t) - r(t-1) \sim N(0, \rho^2)$) について、それぞれの強度 (σ , ρ) (即ち SSE 応答成分に対する S/N 比) を変えた多数のデータセットに対して ICA を適用することにより、SSE 成分を抽出する指標の検討や抽出能力の評価を行った。 σ , ρ の値の検討範囲はそれぞれ SSE 成分の振幅と継続時

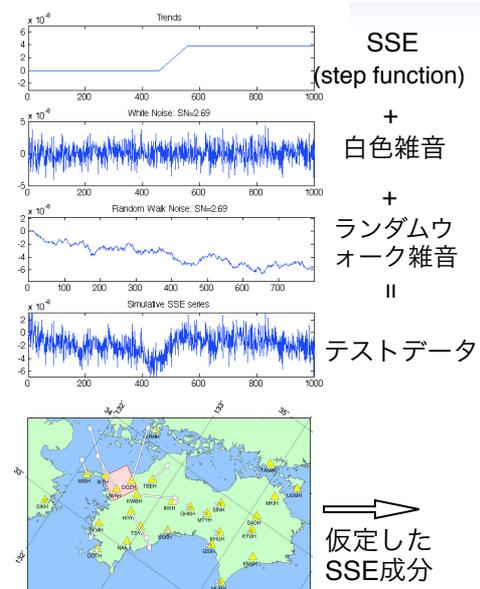


図 2. 人工傾斜データの作成例。SSE 応答 (1 段目), 白色ノイズ (2 段目) ならびにランダムウォークノイズ (3 段目) の線形和とする。各観測点における SSE 応答は、四国西部での SSE による理論傾斜変動とした。

間に対する S/N 比で 0.2-7.7 倍とし、全観測点で共通の値を持つとした。ICA を適用するデータ長 L は 200-1000 時間の 9 通りとし、それぞれ SSE はタイムウィンドウの中心で発生するとした。これらのテストデータを 100 通りの乱数のシードを用いて生成し、最終的な抽出能力評価はその 100 通りの結果を集約したものとした。

(1-3) 網羅的な SSE 抽出能力評価

【MATLAB によるプロトタイプの実装】

本研究では、高並列計算機を活用し可能な限り多様な条件下で抽出能力評価を行う。一方で、実験計画法を援用しながら効果的なパラメータの探

表 1. 抽出能力の評価を行うパラメータ数。

ホワイトノイズ強度 (σ)	16 通り
ランダムウォークノイズ強度 (ρ)	16 通り
テストデータ長 (L)	9 通り
独立成分数 (N)	4 通り
正規乱数シード (s)	100 通り
計	921, 600 通り

素範囲を絞り込んで行くプロセスには、トライアンドエラーを伴う。その過程におけるコード改変のコストを低減するため、本研究では対話型インタプリタ言語である MATLAB で、プロトタイプの実装を行った。

九州大学情報基盤研究開発センターの Fujitsu PRIMERGY CX-400 には、MATLAB での並列計算をサポートする Parallel Computing Toolbox および MATLAB Distributed Computing Server が搭載されている。これらの機能により、並列領域を指定することで容易にスレッド並列を実現することができる。また、MATLAB は通常 GUI によるインタラクティブな操作を想定しているが、バッチモードでの起動もサポートしている。本研究では、大型計算機との親和性から、GUI を用いずに Fujitsu PRIMERGY CX-400 の Portable Batch System を介してジョブを投入するスタイルを採用した。

図 3 に、プロトタイプ実装の概念図を示す。Parallel Computing Toolbox を用いて複数の MATLAB worker を生成し、コア毎に異なるパラメータを割り当てて計算を行う。ここでは、worker 毎に個別に設定ファイルを読み込み、固有のテストデータを生成後、抽出能力評価を行った。Worker 当たりの主要な作業配列は最大で 1.83MB である（パラメータにより作業配列の容量は異なる）。ファイル出力については、各 worker において独立に行うと頻度が高くなるため、並列領域終了後に一括して行うこととした。本研究では、ICA による SSE の抽出能力評価に関する詳細な基礎データを収集するため、最終的な検出精度のみではなくテストデータや独立成分の波形、応答行列の値なども含め、主要な全作業データ出力することとした。従って、一回の検出テストにおけるファイル出力は最大で 1.83MB である。

【Fortran + MPI による実装】

実験計画とプロトタイプ実装の構築後、計算効率の向上と今後の拡張性を考慮し、システム全体を Fortran コードに移植し、フラット MPI (Message-Passing Interface) によって並列化を行った。図 4 にその概念図を示す。ここでは、プ

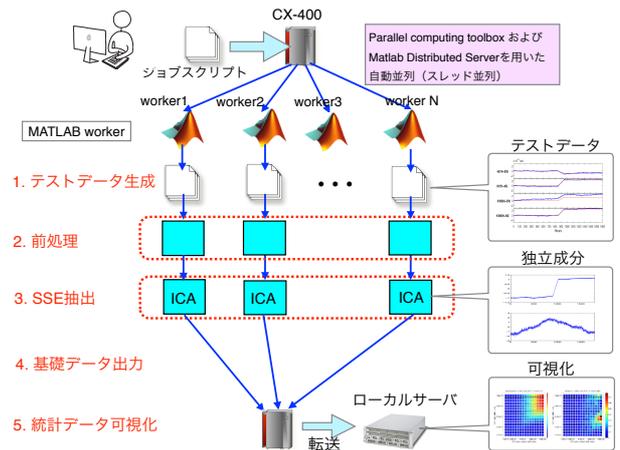


図 3. MATLAB でのプロトタイプの実装。

ロセスとパラメータの対応関係を記述したテーブルを事前に作成し、プロセス毎に異なるパラメータでテストデータを生成し、ICA の検出試験を行った。プロセスあたりの作業配列や出力ファイルの容量はプロトタイプと同様である。図 5 に、16通りのホワイトノイズ、16通りのランダムウォークノイズならびに 9 通りのテストデータ長の 2,304 通りで計算を行った際の、プロセス数と CPU 時間との関係を示す。ここでは、プロセス数は最大で 256 とした。前述の通り、今回の実装では各プロセスの計算結果を集約した検出精度の値ではなく、計算の途中で生成される作業データを各プロセスで独立に出力する。そのため、各プロセスの計算は基本的に独立であり、プロセス数に応じて計算時間の短縮が期待できる。

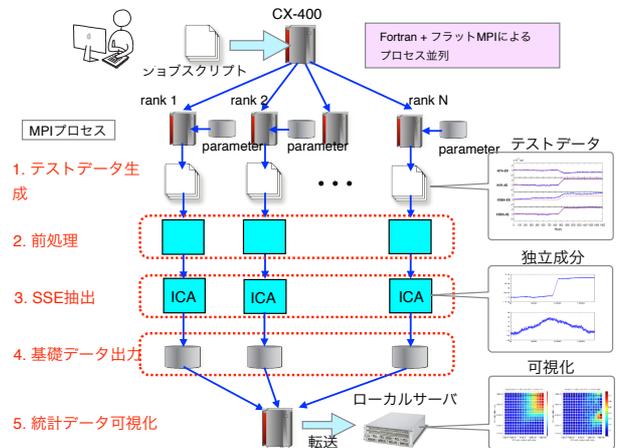


図 4. フラット MPI での実装。

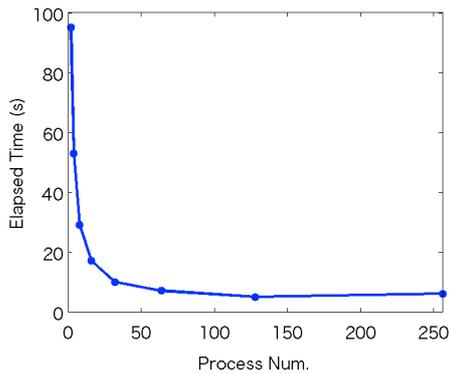


図 5. プロセス数と実行演算時間の関係.

(1-4) ノイズ特性・データ長・成分数の影響

図 6 に、与えた SSE によるシグナルとランダムウォークノイズ、ホワイトノイズそれぞれの S/N 比に対する検出結果の例を示す. (1) 抽出された独立成分と仮定した SSE のすべり時間関数の相関係数, (2) 抽出された独立成分に対する応答行列と仮定した SSE による傾斜応答の相関係数を示している. ICA によって抽出された独立成分は、仮定した SSE のすべり時間関数と次に挙げる 3 つの場合に 0.9 以上の高い相関係数を示している. ① ホワイトノイズ成分に対する S/N が高くランダムウォークノイズ成分に対する S/N 比が低い場合, ② ホワイトノイズ成分・ランダムウォークノイズ成分ともに S/N 比が高い場合, ③ ホワイトノイズ成分・ランダムウォークノイズ成分ともに S/N 比が低い場合. 一方で, ICA によって推定された独立成分に対する応答行列と仮定した SSE に対する傾斜応答を比較すると, 上記②の場合のみ 0.9 以上の高い相関係数を示し, ①③では 0.3 以下の相関係数になった. さらに抽出結果を検証するために, ①③と同じホワイトノイズ・ランダムウォークノイズ成分を持つが SSE 成分を全く含まないデータに対して, ICA を適用した. この結果, SSE 成分が全く含まれていないにも関わらず, 抽出成分は SSE のすべり時間関数と 0.9 以上の高い相関係数を示した. これは, ランダムウォークノイズ成分が SSE 成分に比べて大きなパワーを持つ場合, 抽出成分のみで結果を評価しようとする誤検出を引き起こしやすいことを示す. 一方で, 応答行

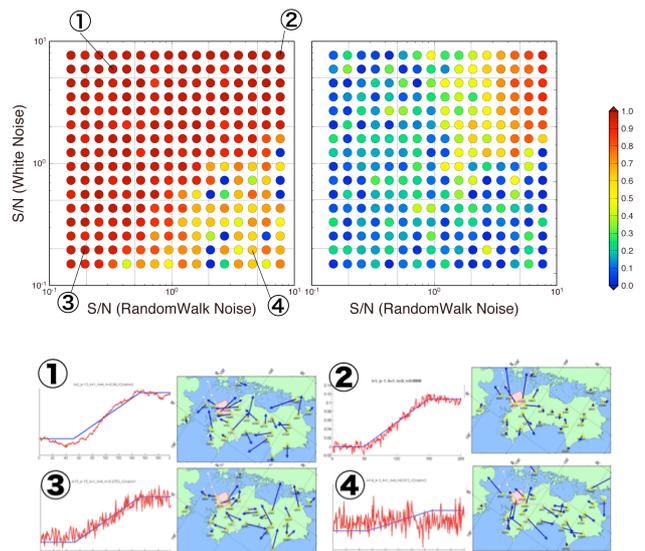


図 6 データ長 200 時間, 独立成分数 4 とした場合の ICA 適用結果例. (左上) ホワイトノイズ (縦軸) ランダムウォークノイズ (横軸) に対する SSE 成分の S/N 比に対して抽出した独立成分と仮定した SSE のすべり時間関数との相関係数. (右上) 同じく推定した応答行列と仮定した SSE による傾斜応答の相関係数. (下) 抽出結果例.

列と仮定した傾斜応答の相関係数にはその様な傾向は見られず, 応答行列のみ或いは応答行列と抽出成分の両方を評価することにより, 精度良く, かつ安定して SSE 成分を検出することが出来ることが分かった.

図 7 に各ノイズ強度で 100 通りの人工データを作成し, ICA を適用した結果を示す. 応答行列の再現性は, 平均値で見るとホワイトノイズに対しては S/N 比 1 以上, ランダムウォークノイズに対しては 2 以上で相関係数 0.5 以上の成分を抽出できている. 一方で相関係数の分散を見ると, ランダムウォークノイズに対して S/N 比が 5 を超えていてもホワイトノイズの S/N 比が 1 程度の場合には 0.1 程度の大きな値を持っている. この様な領域では安定して SSE 成分を抽出することが困難であることが分かる.

また SSE の継続時間は変化させずに, データ長のみを変えて ICA を適用したところ, 短いデータ長の方が高い相関係数で検出出来ることが分かつ

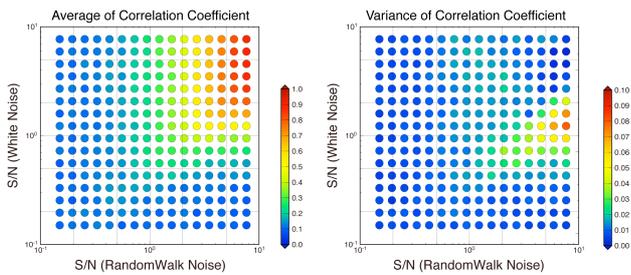


図 7 データ長 200 時間，独立成分数 3 とした場合の各ノイズ強度について 100 通りのデータセットに ICA を適用した結果. (左) 各ノイズ強度における推定した応答行列と仮定した SSE 傾斜応答の相関係数の平均値. (右) 同じく分散.

た. 一方で，仮定する独立成分数によって ICA の適用結果にそれほど大きな変化は見られなかった.

(1-5) ICA による SSE の検知能力

図 8 に各ノイズ強度について，推定した応答行列と仮定した SSE 応答の相関係数が上から 90% 目の値を示す. これは各ノイズ強度に対して，ICA が安定して保証できる相関係数を示している. また，各ノイズ強度を Hi-net 傾斜データの代表的な値 (ホワイトノイズ : $3.0E-8$ rad ; ランダムウォークノイズ : $3.0E-9$ rad/hour^{0.5}) とした場合の，対応する SSE の規模，継続時間も示す. さらに SSE の典型的なモーメント解放率の範囲 ($1.0E+12 \sim 1.0E+13$ Nm/s)，既存の手法 (Kimura et al., 2011) による検知限界，これまでに手動解析により推定されている SSE の震源モデルパラメータも合わせて示す.

相関係数 0.5 以上の範囲に注目すると，Mw 5.8 以上の SSE は安定して検出出来ると考えられる. また Kimura et al. [2011] の検知限界とこの範囲はおおよそ一致しており，ICA をそのまま適用するだけでも同等の性能を持った検出手法を開発できうる.

(1-6) 結果のまとめと考察

本研究で得られた結果から，ICA による SSE 成分の抽出能力は，仮定する独立成分数にはそれほど依存せず，ノイズのパワーやテストデータ長の

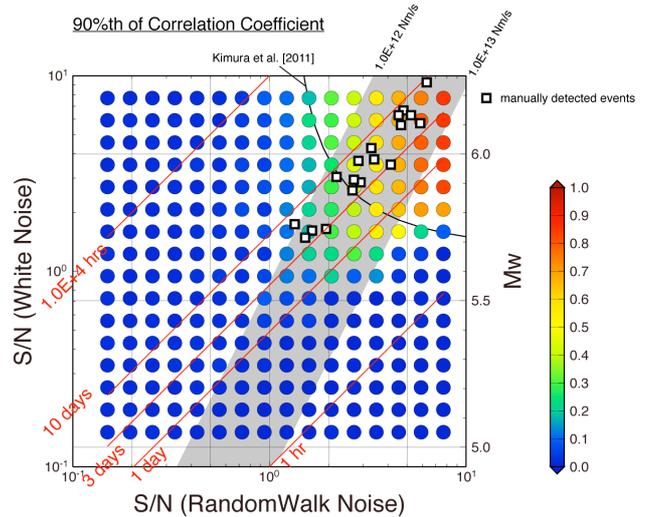


図 8 データ長 200 時間，独立成分数 3 とした場合の各ノイズ強度について 100 通りのデータセットに ICA を適用した結果. 相関係数が大きい方から 90% 目の値を示す. ホワイトノイズ強度，ランダムウォークノイズ強度をそれぞれ $3.0E-8$ rad, $3.0E-9$ rad/hour^{0.5} とした場合の対応する SSE の規模を縦軸右側，継続時間を赤線，モーメント解放率 $1.0E+12 \sim 1.0E+13$ Nm/s の範囲を灰色領域で示す. Kimura et al. [2011] による検知能力を黒線，手動で検出されている SSE の規模，継続時間を白抜き四角で合わせて示す.

影響をより強く受けることがわかった. また，SSE 発生の有無の評価基準として，独立成分の波形は信頼性が低く，応答関数 (分離行列) の方が優れていることがわかった. 特に，SSE を模した step function を含んでいないテストデータからも，ランダムウォーク成分の影響により step 状の独立成分が抽出されてしまうことがわかった. これは，fastICA が非ガウス性を評価するアルゴリズムであることと，step function の確率密度関数は非正規性が高いことと関係していると思われる.

ノイズの影響については，ホワイトノイズよりランダムウォークノイズの方が，平均的な抽出能力に与える影響が大きいことがわかった. ランダムウォークノイズによってはテストデータ平均値が大きくシフトしてしまうことがあり，それによ

って非ガウス性の評価が困難になるためと思われる。また、テストデータ長が短いほど抽出能力が高くなる傾向も、時系列長が短いほど平均値シフトの影響を受けにくいことを反映していると考えられる。一方で、応答行列の相関係数の分散を調べたところ、ランダムウォークノイズの S/N 比が大きくてもホワイトノイズの S/N 比が 1 程度では、抽出能力にばらつきが生じるという結果を得た。これは、抽出能力の再現性を低下させる主要な要因は、SSE よりも長周期のトレンド成分ではなく、SSE より短周期のノイズ成分が寄与していることを示唆している。

実際に観測される SSE の継続時間は、数日～数年とイベントによって大きく異なる。数日程度の短期的 SSE の検出に関しては、データを切り出すウィンドウを数十日程度とし、適切なノイズ除去をすることで、Kimura et al. [2011] と同等以上のパフォーマンスで検知できると思われる。

(2) 当初計画の達成状況について

(2-1) 当初開発スケジュールと遂行状況: 90%

当初計画の研究スケジュールとして、当初の 4 ヶ月間 (4~7 月) でシステムのプロトタイプ実装と実験計画の確定、続く 4 ヶ月間 (8~11 月) で九州大学 PRIMERGY CX-400 への移植と並列化処理の実装、最後の 4 ヶ月間で CX-400 での本格的な解析実施および結果の検証を予定していた。実際の実施状況としては、九州大学情報基盤研究開発センターのサポートもあり、ほぼ当初の研究スケジュール通りに実施することができた。

(2-2) 研究目標の達成状況: 70%

当初の研究計画に基づき、現実的かつ可能な限り多様な条件において ICA による SSE の検知能力の評価を行った。当初の計画では計 640 通りのパラメータにおいて検出能力評価を行う予定であった。しかし、九州大学情報基盤研究開発センターの Fujitsu PRIMERGY CX400 では最大 1,152 コアが使用可能であること、同じノイズレベルでの検出試験においても乱数の種によって結果が異なるこ

とが実施途中で判明したことから、各パラメータに対し 100 通りの乱数の種を用いて検出試験を行うこととした。最終的には 921,600 通りという当初研究計画を遥かに上回る規模で詳細な基礎データを収集することができた。その結果、検出能力を低下させる主要因は長周期の非定常成分であることがわかった。また、ホワイトノイズの影響が大きい状況では、SSE の検知能力が不安定になりうる (精度のバラツキが大きくなる) ことも見えてきた。これらの結果は今後、実観測データを使った SSE のリアルタイム監視や ICA の拡張を行う上で重要な指針を与えるものである。

テストデータ長については当初 80 通りの長さで試験を実施することを想定していたが、抽出能力のデータ長に対する依存性が比較的単純であることが途中で見えてきたことから、最終的には 9 通りのテストデータ長で試験を行えば十分であると判断した。また、独立成分数の選択について、交差検定法などの統計的な方法の実用性を検証する予定であったが、独立成分数に対する結果の依存性があまり見られなかったため、この検証は省略した。さらに、SSE 発生の有無の判別は独立成分の波形に基づいて行うことを想定していたが、この評価基準が適切ではないことが判明したため、応答関数の相関係数によって抽出能力を評価する方針に切り換えた。このように、部分的にパラメータサーベイを実施しながら研究計画を柔軟に見直すことができたのは、システムのプロトタイプを MATLAB で実装したことによって、初期の開発コストを低減できたことが大きい。

一方で、当初研究計画では実観測された地殻変動データに ICA を用いた検出法を適用し、既存手法との比較を行う予定であったが、抽出能力を左右する要因の分析に注力したため、実施期間中に達成することができなかった。また、適当なパラメータ群のクラスを自動的に絞り込んで行くメタ・ヒューリスティックな方法の構築も検討していたが、SSE 成分抽出の安定性の検討まで行ったため、その開発には至らなかった。

4. 今後の展望

(1) 長周期の非定常成分の除去

本研究で得られた結果により、ICA による SSE の検知を阻害し得る最も大きな要因は、ランダムウォークノイズ成分のような長周期の非定常成分であることがわかった。今後、長周期成分を模した時系列モデルの導入などにより、ICA の前処理で長周期の非定常成分を除去するフィルターを導入することで、短期的 SSE の検知能力はさらに向上する可能性がある。

(2) 短周期のノイズ成分の除去

抽出能力の分散から、ホワイトノイズのような短周期のノイズ成分の存在により、抽出能力の再現性が低下することがわかった。長周期のトレンド成分の影響が小さい状況では、移動平均や低域通過フィルターなどを前処理として用いることで、比較的容易に検出能力を高い状態で安定化させることが可能と思われる。

(3) SSE の空間特性を取り入れた ICA の拡張

本研究において収集した詳細な基礎データにより、ICA を用いた信号分離によって地球物理学的知見に基づく断層モデルを用いなくても、傾斜変動データ中の SSE を検知でき得るという結果を得た。ICA は、信号源（潜在変数）および信号の伝搬過程（潜在構造）両方の情報が未知という状況を想定している。今後、SSE の検知能力をさらに向上させるためには、SSE の応答特性に関する緩やかな拘束条件（事前情報）を、統計的独立性の評価関数に導入する形で ICA を拡張することが考えられる。それにより、地球物理学的なモデルを直接用いた方法を上回る検知能力の達成が期待される。

(4) 実観測された傾斜変動データへの適用

実際の傾斜変動データには、モデル化するのが困難な降雨などの気象擾乱の影響が重畳している。このような気象擾乱の影響がある実観測データに対し、ICA を用いた SSE 検出法のパフォーマンスを既存手法と比較する必要がある。

(5) オンライン SSE 監視システムへの拡張

実際の傾斜変動データはリアルタイムで取得さ

れるため、SSE をオンラインで抽出することができれば、平時には見られない地殻変動を早期に検知することが可能となる。本研究の結果から、ICA の適切なタイムウィンドウは SSE の継続時間によって異なると考えられるため、実際の運用では複数のタイムウィンドウで最新のデータを切り出し、それぞれのデータに ICA を適用することが現実的である。その後、各タイムウィンドウで推定された独立成分の応答関数から、各独立成分の「SSE らしさ」を評価する手続きが必要になる。その際、ICA の応答関数から「SSE らしさ」を適切に評価する判別器の開発が課題である。

また、SSE などの地殻活動の高精度自動検出手法の開発は、手動での逐次解析が困難な大規模データの解析に必要不可欠である。SSE に伴う変動は、本研究で対象とした傾斜データのみではなく GNSS やひずみ計など多種かつ大量のデータで捉えられており、これらの同時解析が今後の課題である。本研究では、様々なノイズ特性の人工データについて ICA の性能評価を網羅的に行った。従って、様々な種類の地殻変動データについてそのノイズ特性を評価さえ出来れば、本研究の結果から ICA による抽出性能を評価可能であり、今後の重要な指標となる。

5. 研究成果リスト

(1) 学術論文（投稿中のものは「投稿中」と明記）

(2) 国際会議プロシーディングス

(3) 国際会議発表

(4) 国内会議発表

●徳永旭将, 木村武志, 稲津大祐, 中村和幸, 2013, 独立成分分析の網羅的適用による地殻変動データからのスロースリップイベント成分の抽出, 日本地震学会秋季大会, P2-06.

●徳永旭将, 木村武志, 稲津大祐, 中村和幸, 2013, 独立成分分析を用いた地殻変動データからのスロースリップイベント成分抽出, 日本測地学会第 120 回講演会, 2.

(5) その他（特許, プレス発表, 著書等）