

# 非負値行列因子分解を用いた 南海トラフ巨大地震による津波浸水分布の空間分布特性の評価

Non-negative Matrix Factorization of Spatial Distribution of Tsunami Inundation Depth  
Caused by the Nankai Megathrust Earthquakes

高橋 幸宏<sup>1</sup>, 能島 暢呂<sup>2</sup>

Yukihiro TAKAHASHI<sup>1</sup> and Nobuoto NOJIMA<sup>2</sup>

<sup>1</sup> 岐阜大学大学院 工学研究科 博士後期課程

Doctoral Program in Graduate School of Engineering, Gifu University

<sup>2</sup> 岐阜大学 工学部, 教授

Professor, Department of Civil Engineering, Gifu University

We evaluate spatial distribution characteristics of tsunami inundation depth by using “non-negative matrix factorization (NMF)”. Numerical examples are shown for tsunami inundation depth caused by the Nankai Megathrust Earthquakes. The spatial bases obtained by NMF well represent the regional distribution characteristics of inundation depth associated with tsunami fault settings. While the bases obtained by singular value decomposition used in the previous study contain negative values, the NMF bases are easy to interpret because of their non-negative nature, resulting in a concise representation of the original inundation depth. Such an advantage can be enhanced in terms of sparsity by applying non-negative double singular value decomposition for matrix initialization in NMF.

**Keywords:** Nankai megathrust earthquakes, tsunami, inundation depth, spatial distribution characteristics, non-negative matrix factorization, non-negative double singular value decomposition

## 1. はじめに

内閣府に設置された「南海トラフの巨大地震モデル検討会<sup>1)</sup>」では、南海トラフ巨大地震のうち、発生頻度は極低頻度であるが最大級の津波被害が予測される津波高・浸水深分布を想定および公開している<sup>2)</sup>。この想定では11ケースの津波断層モデルが設定され、想定範囲は九州から東北までの広範囲に及ぶ。一方、文部科学省に設置された地震調査研究推進本部では、南海トラフ地震で348,345種類の地震パターン（波源断層モデルの組み合わせ）を対象とした確率論的津波ハザード評価を行っている<sup>3)</sup>。また防災科学技術研究所では、極低頻度の巨大地震を含む確率論的津波ハザード評価を行っている<sup>4)</sup>。

津波解析および津波リスク評価における不確実性の扱いは重要な課題であり、これまでも多くの研究が行われている。例えば、断層長さや走向の違いが津波水位や波高の空間分布などに及ぼす影響の評価<sup>5)</sup>や、東北地方の津波痕跡高を用いた津波解析の計算水位のばらつきの定量的評価<sup>6)</sup>、断層面のすべり量を確率量とした津波高の不確実性における定量的評価<sup>7)</sup>、応答曲面を用いた津波高のシミュレーションによる最大津波高の不確実性の定量化<sup>8)</sup>および主成分分析による地点間相関の分析<sup>9)</sup>などである。

一方、複数ケースの津波浸水深分布における空間的なばらつきと空間相関（空間相関特性）を定量的に明らかにするため、特異値分解によるモード分解手法が提案されている<sup>10)</sup>。これは地震動分布の空間相関特性の評価手法<sup>11)–13)</sup>を、津波浸水深分布に応用したものである。南海トラフの巨大地震モデル検討会<sup>1)</sup>における11ケースの浸水深分布に対してモード分解が適用された。その結果、モード1の空間相関特性は平均的な浸水深分布の分布形状を反映し、モード2以上では超大すべり域を含む大すべり域の配置条件を反映していることが明らかとなった。

ここで、特異値分解では、互いに無相関なモードの固有ベクトルとして直交基底が得られるため、各モードの空間相関を独立した成分として表現できる。しかしながら、分解された基底が負値を含み、合成の際の重み係数もまた負値を含むため、対象とする物理量によっては、分解・合成の結果の解釈および意味付けが困難になることがある。一方、非負値行列因子分解<sup>14)</sup>は非負行列を非負基底の加算のみで近似表現する方法であり、非負値の物理量の分解・合成において、解釈や意味付けが容易となり、ゼロ要素が多くスパース性が高い結果が得られれば、特に有効性を発揮することが知られている。

以上の背景のもとで本研究では、複数ケースの津波浸水深分布における空間分布特性の評価手法として、非負

値行列因子分解を適用する方法を提案する。津波浸水深分布はすべて非負値であり、非負値行列因子分解の適用条件に合致することから、その空間分布特性の評価に有効であることが期待される。得られる各基底は、対象とする複数ケースを加算的に合成するための要素分布であり、津波断層モデルとの比較を通じて、浸水深分布の影響要因を明確化できるという利点がある。

ここで非負値行列因子分解の初期化手法についてはこれまでに多くの手法が提案されている。本研究では、標準的な一様乱数による初期化手法と、先行研究<sup>10)</sup>で用いた特異値分解を基とした非負2重特異値分解<sup>15)</sup>による初期化手法の2つを適用して浸水深の空間分布特性を評価する。具体的には、先行研究<sup>10)</sup>と同様に南海トラフ巨大地震による11ケースの津波浸水深分布に非負値行列因子分解を適用して、基底ベクトルによって浸水深の空間的な分布特徴をマップ表現するものであり、初期化手法の違いが評価結果に及ぼす影響について明らかにする。

## 2. 非負値行列因子分解

### (1) 津波による浸水深分布の非負値行列因子分解

浸水深を変量  $x$  ( $\geq 0$ ) で表し、浸水深分布 (評価地点数  $M$ ) を  $M$ 次元の特徴ベクトル  $\mathbf{x} = (x_1, \dots, x_M)^T$  で表す。異なる津波断層パラメータによる  $N$  ケースの浸水深分布が得られているものとし、それぞれの浸水深分布の  $\mathbf{x}$  を列要素とした  $M$  行  $N$  列 ( $M \gg N$ ) の非負行列  $\mathbf{X}$  ( $i$  行  $j$  列要素は評価地点  $i$  のケース  $j$  の浸水深) を定義する。

$$\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N) = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{M1} & \cdots & x_{MN} \end{pmatrix} \quad [1]$$

式[1]の行列  $\mathbf{X}$  に非負値行列因子分解 (NMF: Non-Negative Matrix Factorization)<sup>14)</sup> を適用すると、 $\mathbf{X}$  は近似的に非負行列  $\mathbf{W}$  と  $\mathbf{H}$  に分解される。

$$\mathbf{X} = \mathbf{WH} = \begin{pmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{M1} & \cdots & w_{Mk} \end{pmatrix} \begin{pmatrix} h_{11} & \cdots & h_{1N} \\ \vdots & \ddots & \vdots \\ h_{k1} & \cdots & h_{kN} \end{pmatrix} \quad [2]$$

ここに  $k$  は、 $k \leq \min\{M, N\} = N$  を満たす任意の基底数であり、基底番号  $l$  は  $1 \leq l \leq k$  となる。式[2]の行列  $\mathbf{W}$ ,  $\mathbf{H}$  の意味は、下記のとおりである。

- 行列  $\mathbf{W}$ : 全要素が非負値である  $M$  行  $k$  列の基底行列。  
第  $l$  列の基底ベクトル  $\mathbf{w}_l = (w_{1l}, \dots, w_{il}, \dots, w_{Ml})^T$  は、基底  $l$  における評価地点の特徴量の空間分布を規定しており、そのマップ表示により特徴量の空間分布を可視化できる。
- 行列  $\mathbf{H}$ : 全要素が非負値である  $k$  行  $N$  列の係数行列。  
第  $j$  列の結合係数ベクトル  $\mathbf{h}_j = (h_{1j}, \dots, h_{lj}, \dots, h_{kj})^T$  は、ケース  $j$  における各基底の重みを表す。

非負値の線形結合  $\mathbf{WH}$  は行列  $\mathbf{X}$  の  $k$  次元での低次元近似であるといえる。また、 $\mathbf{H}$  は非負制約によってスパース性が現れやすい。以上の「非負基底の加算」と「 $\mathbf{H}$  のスパース性」により、NMF では  $\mathbf{W}$  および  $\mathbf{H}$  の解釈が容易になるという特長がある。

NMF では、一般的に基底数  $k$  の増加とともに行列  $\mathbf{X}$  の近似精度は向上する。しかし、 $\mathbf{X}$  の特徴抽出という点において、 $\mathbf{W}$  が張る  $k$  次元のベクトル空間で  $\mathbf{X}$  を表現する必要があるため、適切な基底数を探索的に決定する必要がある。

NMF の数値計算には統計ソフト R<sup>16)</sup> のパッケージ NMF<sup>17)</sup> を用い、行列  $\mathbf{W}$ ,  $\mathbf{H}$  に初期値を与え、更新式による収束計算を繰り返すことで  $\mathbf{X}$  の近似を行う。更新式は Brunet et al. の手法<sup>18)</sup> より、カルバック・ライブラー情報量を規準とした最適化問題として次式で定義されている。

$$h_{ij} \leftarrow h_{ij} \frac{\sum_{i=1}^M w_{il} x_{ij} / (\mathbf{WH})_{ij}}{\sum_{i=1}^N w_{il}} \quad [3]$$

$$w_{ij} \leftarrow w_{ij} \frac{\sum_{j=1}^N h_{ij} x_{ij} / (\mathbf{WH})_{ij}}{\sum_{j=1}^N h_{ij}} \quad [4]$$

NMF の近似精度は  $\mathbf{W}$ ,  $\mathbf{H}$  の初期値にも依存するため、本研究では、標準的な一様乱数による手法と、非負2重特異値分解 (NNSVD: Non-negative Double Singular Value Decomposition)<sup>15)</sup> による手法の2種類を用いる。以下にそれらについて記す。

### (2) 一様乱数による行列 $\mathbf{W}$ と $\mathbf{H}$ の初期化

行列  $\mathbf{W}$ ,  $\mathbf{H}$  の初期行列を  $\mathbf{W}_0$ ,  $\mathbf{H}_0$  とする。一様乱数による初期化では、これらの各要素は、 $\mathbf{X}$  の全要素の最大値  $\max\{\mathbf{X}\}$  を用いて、区間  $[0, \max\{\mathbf{X}\}]$  に連続分布した一様乱数によって決定される。これより  $\mathbf{W}_0$  および  $\mathbf{H}_0$  は試行ごとに異なった値をとるため、その収束結果である  $\mathbf{W}$ ,  $\mathbf{H}$  も一意には定まらない。このため一様乱数による初期化では、複数回の試行から最適な  $\mathbf{W}$ ,  $\mathbf{H}$  を選択する方法がとられる。しかし、選択結果も分解結果の一つであり、一様乱数ごとに分解結果の特徴が異なることに留意する必要がある。

### (3) 非負2重特異値分解による行列 $\mathbf{W}$ と $\mathbf{H}$ の初期化

NNSVD<sup>15)</sup> による初期化手法を以下に記す。まず行列  $\mathbf{X}$  に特異値分解<sup>19)</sup> を適用する。

$$\mathbf{X} = \mathbf{UDV}^T$$

$$= \begin{pmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{M1} & \cdots & u_{MN} \end{pmatrix} \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_N \end{pmatrix} \begin{pmatrix} v_{11} & \cdots & v_{1N} \\ \vdots & \ddots & \vdots \\ v_{N1} & \cdots & v_{NN} \end{pmatrix}^T \quad [5]$$

$$\mathbf{U}^T \mathbf{U} = \mathbf{I} \quad \mathbf{V}^T \mathbf{V} = \mathbf{V} \mathbf{V}^T = \mathbf{I} \quad [6]$$

ここでモード番号を  $p$  ( $1 \leq p \leq \min\{M, N\} = N$ ) とする。式[5], [6]の行列  $\mathbf{U}$ ,  $\mathbf{D}$ ,  $\mathbf{V}$ ,  $\mathbf{I}$  の意味は、下記のとおりである。

- 行列  $\mathbf{U}$ :  $M$  次元空間のうち  $N$  次元部分空間を張る正規直交基底によって構成される  $M$  行  $N$  列の行列。第  $p$  列の左特異ベクトル  $\mathbf{u}_p = (u_{1p}, \dots, u_{ip}, \dots, u_{Mp})^T$  は、モード  $p$  における評価地点の特徴量の空間分布と空間相関を規定しており、そのマップ表示により評価地点間の空間的な相関関係を可視化できる。
- 行列  $\mathbf{D}$ :  $N$  行  $N$  列の対角行列 ( $\mathbf{D}$  の階数は  $\text{rank}(\mathbf{D}) = \min\{M, N\} = N$ )。  $p$  行  $p$  列要素の特異値  $d_p$  (固有値の正の平方根) は、ケースに依存しないモード  $p$  の重みを表す。
- 行列  $\mathbf{V}$ :  $N$  次元空間を張る正規直交基底によって構成される  $N$  行  $N$  列の行列。第  $p$  列の右特異ベクトル  $\mathbf{v}_p = (v_{1p}, \dots, v_{jp}, \dots, v_{Np})^T$  は、モード  $p$  における各ケースの重みを表す。
- 行列  $\mathbf{I}$ :  $N$  行  $N$  列の単位行列。

以上の特異値分解において、分解結果である  $\mathbf{U}$ ,  $\mathbf{D}$ ,  $\mathbf{V}$  は一意に定まる。また式[5]の  $\mathbf{UDV}^T$  の計算により、特異ベクトルにおける符号の任意性は相殺される。

次に行列  $\mathbf{U}$ ,  $\mathbf{V}$  における  $l$  列の特異ベクトル  $\mathbf{u}_l$ ,  $\mathbf{v}_l$  を、正要素  $\mathbf{u}_{+l}$ ,  $\mathbf{v}_{+l}$  と負要素  $\mathbf{u}_{-l}$ ,  $\mathbf{v}_{-l}$  に分離する。

$$\mathbf{u}_{+l} = \max\{\mathbf{u}_l, \mathbf{0}\}, \quad \mathbf{u}_{-l} = \min\{\mathbf{u}_l, \mathbf{0}\} \quad [7]$$

$$\mathbf{v}_{+l} = \max\{\mathbf{v}_l, \mathbf{0}\}, \quad \mathbf{v}_{-l} = \min\{\mathbf{v}_l, \mathbf{0}\} \quad [8]$$

ここで  $\mathbf{u}_{+l}, \mathbf{u}_{-l}$  は  $M$  次元ベクトルであり、 $\mathbf{v}_{+l}, \mathbf{v}_{-l}$  は  $N$  次元ベクトルである。  $\mathbf{0}$  はゼロベクトル (式[7]では  $M$  次元、式[8]では  $N$  次元) であり、 $\max$  および  $\min$  はベクトル要素ごとに適用されるものとする。符号別の特異ベクトルにおけるノルムの積を  $\sigma_{+l}, \sigma_{-l}$  と表す。

$$\sigma_{+l} = \|\mathbf{u}_{+l}\| \cdot \|\mathbf{v}_{+l}\| \quad [9]$$

$$\sigma_{-l} = \|\mathbf{u}_{-l}\| \cdot \|\mathbf{v}_{-l}\| \quad [10]$$

NMF における行列  $\mathbf{W}, \mathbf{H}$  は非負基底で構成されるため、NNDSVD では正規直交基底である特異ベクトル  $\mathbf{u}_l, \mathbf{v}_l$  から非負基底 (斜交基底)  $\mathbf{u}'_l, \mathbf{v}'_l$  を抽出する。  $\mathbf{u}'_l, \mathbf{v}'_l$  を次式で定義する。

$$(\mathbf{u}'_l, \mathbf{v}'_l, \sigma'_l) = \begin{cases} (\mathbf{u}_{+l}, \mathbf{v}_{+l}, \sigma_{+l}) & \text{if } \sigma_{+l} \geq \sigma_{-l} \\ (-\mathbf{u}_{-l}, -\mathbf{v}_{-l}, \sigma_{-l}) & \text{otherwise} \end{cases} \quad [11]$$

初期行列  $\mathbf{W}_0$  の第  $l$  列ベクトルを  $\mathbf{w}_{0l}$  と表し、初期行列  $\mathbf{H}_0$  の第  $l$  行ベクトルを  $\mathbf{h}_{0(l)}$  と表した場合、これらを上記の  $\mathbf{u}'_l$  および  $\mathbf{v}'_l$  より次式で定義する。

$$\mathbf{w}_{0l} = \sqrt{d_l \sigma'_l} \cdot \frac{\mathbf{u}'_l}{\|\mathbf{u}'_l\|} \quad [12]$$

$$\mathbf{h}_{0(l)} = \sqrt{d_l \sigma'_l} \cdot \frac{\mathbf{v}'_l}{\|\mathbf{v}'_l\|} \quad [13]$$

ここで NNDSVD の基底番号は特異値分解のモード番号に対応させることとする。

以上の NNDSVD による計算手順を経て、初期行列  $\mathbf{W}_0, \mathbf{H}_0$  を設定する。上述の通り、特異値分解の分解結果は、

本質的には一意に定まる (符号の任意性を除く) ため、2. (2) の一様乱数と異なり、NNDSVD による初期化およびその収束結果  $\mathbf{W}, \mathbf{H}$  は一意に定まる。また式[7], [8]で定義された  $\mathbf{W}_0, \mathbf{H}_0$  のゼロ要素は、更新式[3], [4]による収束結果  $\mathbf{W}, \mathbf{H}$  でもゼロ要素として保持されるため、NNDSVD では、一様乱数よりも分解結果のスパース性が期待できる。

### 3. 南海トラフ巨大地震による津波の浸水深分布

本研究では、先行研究<sup>10)</sup>と同様に南海トラフの巨大地震モデル検討会の第二次報告<sup>1)</sup>で公表されたデータを用いて検討を行う。

#### (1) 津波断層モデル

南海トラフの巨大地震モデル検討会<sup>1)</sup>では、 $M_w=9$  クラスの巨大地震の中でも最大クラスの津波を引き起こす津波断層モデルとして、基本的な検討ケース (以下、基本ケース) である 5 ケース、その他派生的な検討ケース (以下、派生ケース) である 6 ケースの計 11 ケース (図 1, 表 1) を検討している (ケース数  $N=11$ )<sup>2)</sup>。基本ケースでは、津波断層モデル中に大すべり域と超大すべり域が 1 箇所ずつ設定されており、全 5 ケースで南海トラフ沿いを網羅的に検討している。また、破壊開始点は大すべり域の中心付近の深さ 20km に設定されている。なお、ケース 2 では過去の地震 (1944 年昭和東南海地震, 1946 年昭和南海地震) を基に紀伊半島の塩沖沖に破壊開始点が設定されている。派生ケースでは、すべり域中に分岐断層

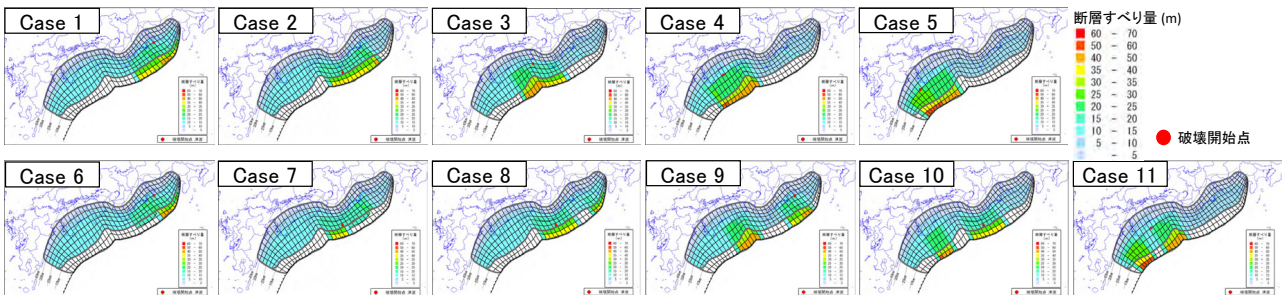


図 1 南海トラフ巨大地震の津波断層モデル (文献<sup>2)</sup>より引用, 編集)

表 1 南海トラフ巨大地震の津波断層パラメータ (文献<sup>2)</sup>を基に作成)

タイプ	Case	大すべり域, 超大すべり域		$M_0$ [Nm]	$M_w$	平均すべり量 [m]	派生の対象ケース (破壊開始点の位置)
		位置	配置パターン				
基本的な検討ケース	Case 1	駿河湾~紀伊半島沖	大すべり域+超大すべり域	$6.1 \times 10^{22}$	9.1	10.3	
	Case 2	紀伊半島沖	大すべり域+超大すべり域	$6.3 \times 10^{22}$	9.1	10.7	
	Case 3	紀伊半島沖~四国沖	大すべり域+超大すべり域	$6.4 \times 10^{22}$	9.1	10.9	
	Case 4	四国沖	大すべり域+超大すべり域	$6.4 \times 10^{22}$	9.1	10.8	
	Case 5	四国沖~九州沖	大すべり域+超大すべり域	$6.3 \times 10^{22}$	9.1	10.7	
その他派生的な検討ケース	Case 6	駿河湾~紀伊半島沖	大すべり域+ (超大すべり域, 分岐断層)	$5.4 \times 10^{22}$	9.1	9.0	Case 1 (Case 1)
	Case 7	紀伊半島沖	大すべり域+ (超大すべり域, 分岐断層)	$5.3 \times 10^{22}$	9.1	8.8	Case 2 (Case 2)
	Case 8	駿河湾~愛知県東部沖 三重県南部沖~徳島県沖	大すべり域+超大すべり域 大すべり域+超大すべり域	$6.2 \times 10^{22}$	9.1	10.4	Case 1, 2 (Case 2)
	Case 9	愛知県沖~三重県沖 室戸岬沖	大すべり域+超大すべり域 大すべり域+超大すべり域	$6.7 \times 10^{22}$	9.1	11.3	Case 1, 3 (Case 1)
	Case 10	三重県南部沖~徳島県沖 足摺岬沖	大すべり域+超大すべり域 大すべり域+超大すべり域	$6.4 \times 10^{22}$	9.1	10.8	Case 2, 4 (Case 2)
	Case 11	室戸岬 日向灘	大すべり域+超大すべり域 大すべり域+超大すべり域	$6.6 \times 10^{22}$	9.1	11.2	Case 4, 5 (Case 4)

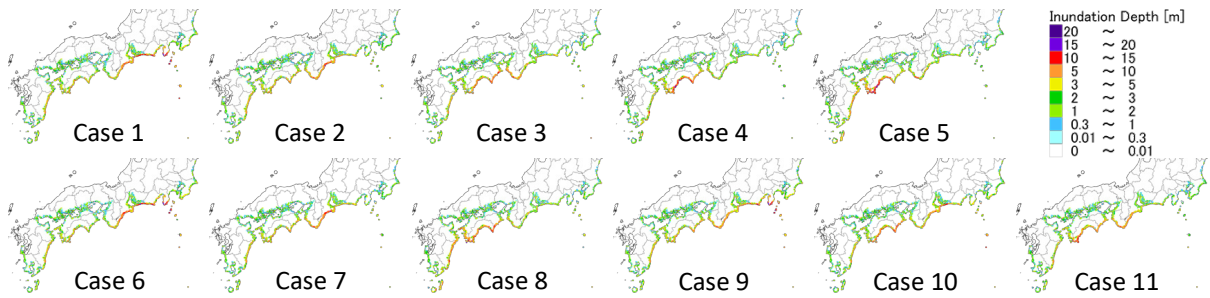
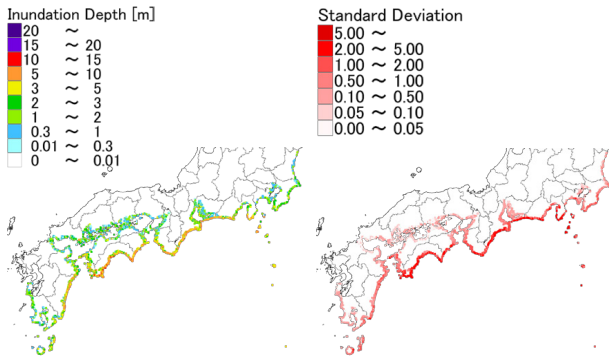


図2 南海トラフ巨大地震による津波の浸水深分布（文献<sup>1)</sup>を基に作成）



(a) 平均浸水深 (b) 標準偏差  
図3 11 ケースの浸水深の平均値と標準偏差の分布

を考慮した2ケース（ケース6,7）と大すべり域と超大すべり域が2箇所設定された4ケース（ケース8~11）が検討され、破壊開始点は派生の対象である基本ケースと同位置に設定されている。

## (2) 津波の浸水深分布

浸水深は、沿岸部および陸域の遡上部にて10mメッシュ単位で評価され、その数値は0.01m単位である。浸水深の推計に際して、津波断層は破壊開始点から破壊が伝播するモデルであり、潮位は満潮位（「平成24年気象庁潮位表」の各地の年間最高潮位を参考に設定<sup>2)</sup>）として、家屋などの障害物は粗度係数で表現されている。堤防（水門を含む）がある場合は「地震動により堤防が機能しなくなる条件」と「津波が堤防を越えることにより当該堤防が機能しなくなる条件」の2つが検討されている。本研究では、先行研究<sup>10)</sup>と条件を合わせるため、後者による浸水深分布を用いる。

評価対象領域は日本平面直角座標第02系（九州の大部分）から第09系（関東周辺から宮城県）であり、全11ケースで共通して浸水深が評価されているメッシュ数は約664万である。先行研究<sup>10)</sup>では、その中から1000データごとに1データを抽出した6644メッシュを対象に空間相関特性が評価されている。本研究ではNMFの計算の都合上、全ケースで0.00mと評価された3地点を除いた $M=6641$ メッシュを分析対象とした。

本研究で用いた11ケースの浸水深分布を図2に示す。これより、超大すべり域を含む大すべり域（以下、大すべり域）の近傍を中心とした太平洋沿岸部で、浸水深の値が大きいたことが読み取れる。全11ケースの平均浸水深と標準偏差の分布を図3に示す。平均浸水深と標準偏差はともに、太平洋沿岸部で大きな値を示し、内海になるにつれ値は小さくなっている。

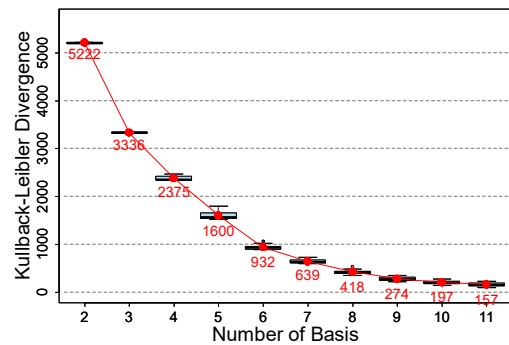


図4 NMF-Rにおける基底数別の初期値の違いによる近似精度の箱ひげ図（各基底における初期化の試行回数：20回、赤点および赤色の数字：20回の試行における平均値）

## 4. 非負値行列因子分解（初期化手法：一様乱数）による浸水深の空間分布特性の評価

本章では、図2の浸水深分布に対して、初期化手法として一様乱数を用いた非負値行列因子分解（以下、NMF-R）を適用する。

### (1) 非負値行列因子分解（初期化手法：一様乱数）における基底数と初期値の違いによる近似精度の変動

先述の通り、NMFでは、基底数によって行列 $\mathbf{X}$ に対する行列 $\mathbf{WH}$ の近似精度は変動する。また、NMF-Rでは、初期値 $\mathbf{W}_0, \mathbf{H}_0$ が試行によって異なるため、収束結果の $\mathbf{WH}$ も一意に定まらず、近似精度が変動する。そこで基底数 $k$ を $2 \leq k \leq 11$ と設定し、それぞれの基底数で20種類の一様乱数を与えて収束計算を行った。これより得た、 $\mathbf{X}$ に対する基底結合 $\mathbf{WH}$ の近似精度（カルバック・ライブラー情報量による残差）の変動を図4に示す。ここで、図中の各基底における赤点および赤色の数字は20回の試行における近似精度の平均値を表す。これより、近似精度の変動は、初期値（一様乱数）の違いよりも基底数の違いによる影響が大きいことが読み取れる。また、基底数の増加に伴う近似精度の向上の効果は、基底数6と7の間で小さくなっていることが読み取れる（カルバック・ライブラー情報量の減少分：基底数4~5~6~7~8の間でそれぞれ775, 668, 293, 221）。

### (2) 非負値行列因子分解（初期化手法：一様乱数）による浸水深の空間分布特性

基底数 $k=5\sim7$ （近似精度の向上の効果が小さくなる直前の $k=6$ とその前後）におけるNMF-Rの分解結果を比較した（紙面の都合上、図示を省略）。基本ケース1~5に着目した場合、 $k=6$ では2つの基底で近似できるのに対し、



$k=5$  および  $7$  では  $3$  つを要したことから、よりシンプルな  $k=6$  を採用した。NMF-R によって得た基底ベクトル  $\mathbf{w}$  の空間分布を図 5 に示し、結合係数ベクトル  $\mathbf{h}$  を図 6 に示す。基底ベクトルの空間分布における支配要因について、結合係数ベクトルを津波断層モデルと関連付けながら、以下に考察する。

### a) 基底 1

基底ベクトルは愛知県から静岡県太平洋沿岸部に大きな値が分布している。結合係数ベクトル (図 6) より基底 1 は、当該地域の近傍に大すべり域が設定された基本ケース 1 とその派生ケースの一部 (ケース 6, 8) で比較的大きな値をとる。これより、基底 1 は愛

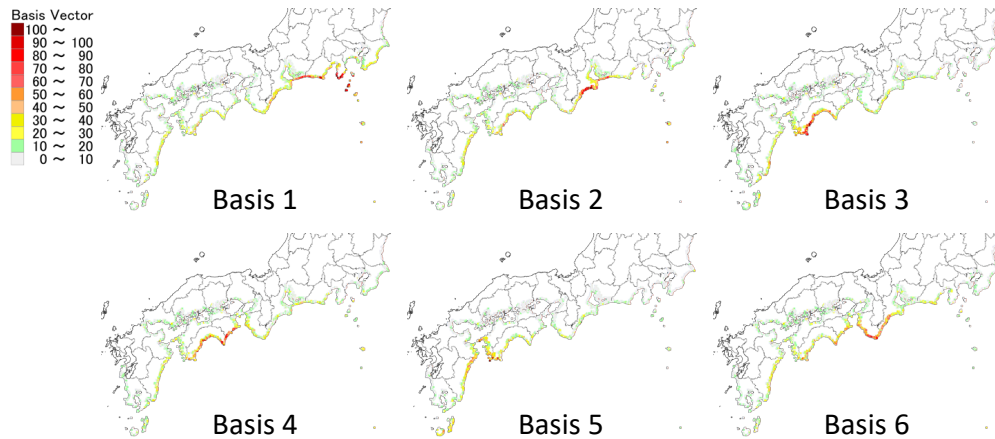


図 5 基底ベクトルの空間分布 (NMF-R,  $k=6$ )

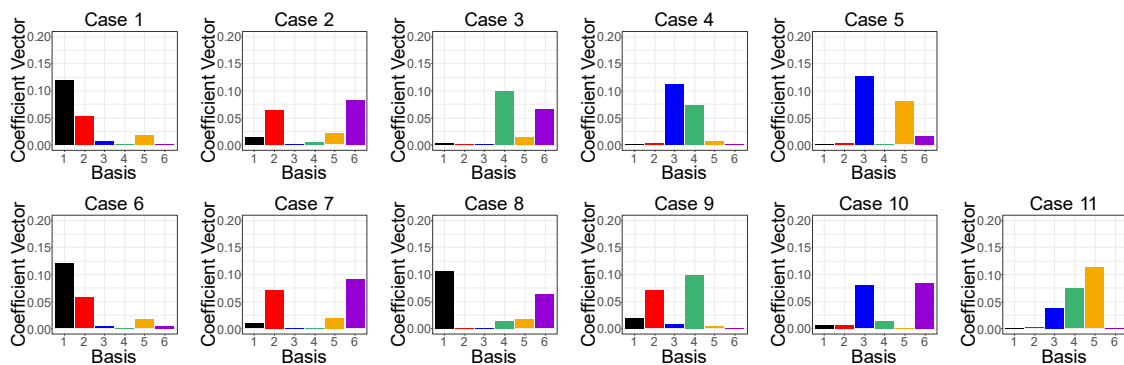


図 6 結合係数ベクトル (NMF-R,  $k=6$ )

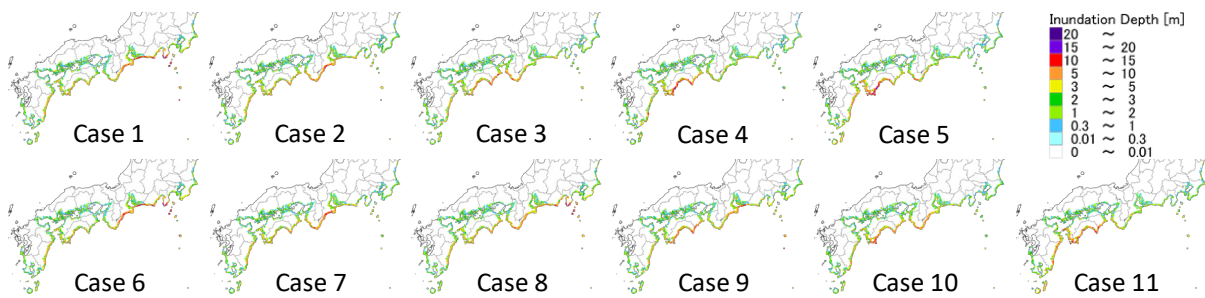


図 7 基底結合した近似浸水深分布 (NMF-R,  $k=6$ )

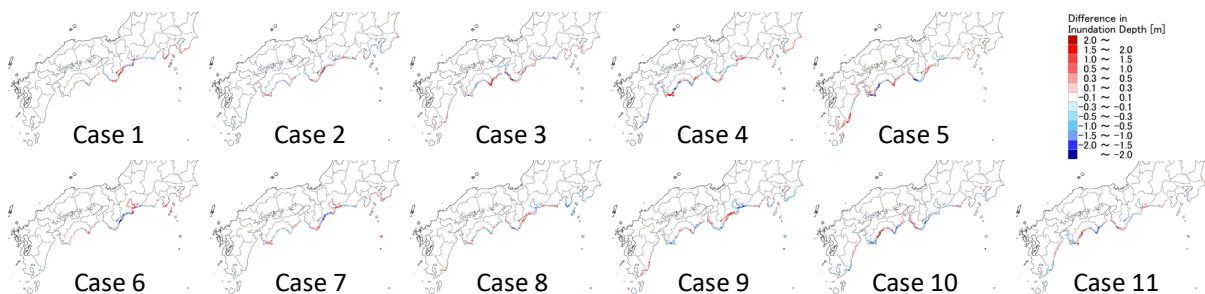


図 8 基底結合 (NMF-R,  $k=6$ ) と元データとの浸水深差の分布

知県～静岡県の近傍における大すべり域の配置条件を反映していると考えられる。

### b) 基底 2

基底ベクトルは三重県から愛知県の太平洋沿岸部で大きな値が分布する。結合係数ベクトル (図 6) より基底 2 は、当該地域の近傍に大すべり域が設定された基本ケース 1, 2 とその派生ケースの一部 (ケース 6, 7, 9) で寄与している。これより、基底 2 では、三重県～愛知県の近傍における大すべり域の配置条件を反映していると考えられる。

基底 3～6 も同様に大すべり域の配置条件を反映していると考えられる。なお、基底 3 では「四国西部」、基底 4 では「四国東部」、基底 5 では「九州～四国西部」、基底 6 では「紀伊半島」の近傍における大すべり域の配置条件であると考えられる。

結合係数ベクトルより、基本ケース 1～5 は主に 2 つの基底によって近似されており、派生ケース 6～11 は派生元の基本ケースを構成する基底によって近似されていることが読み取れる。しかし、行列  $\mathbf{H}$  に含まれるゼロ要素は多いとは言えず、NMF の特長としての  $\mathbf{H}$  のスパース性が十分に発揮されていない。つまり NMF-R による近似モデルは、多くの基底が複雑に寄与しており、分解・合成の意味付けが不明瞭になっている。

### (3) 非負値行列因子分解 (初期化手法: 一様乱数) の基底結合による近似浸水深分布

図 7 に NMF-R の基底結合による近似浸水深分布を示し、図 8 に近似浸水深分布 (図 7) から元の浸水深分布 (図 2) を引いた浸水深差の分布を示す。これより近似浸水深分布と元の浸水深分布に明瞭な空間分布の違いは見られない。近似浸水深と元の浸水深の相関係数は、0.982～0.996 であり、高い相関関係にある。また 2 乗平均平方誤差 (RMSE; Root Mean Squared Error) は、0.217～0.456 であ

った。これらより、基底数 6 の NMF-R による低次元近似は十分に高い精度で元の浸水深分布を表せているといえる。

## 5. 非負値行列因子分解 (初期化手法: NNDSVD) による浸水深の空間分布特性の評価

本章では、図 2 の浸水深分布に対して、初期化手法として NNDSVD を用いた非負値行列因子分解 (以下、NMF-NNDSVD) を適用する。

### (1) 特異値分解による浸水深分布の正規直交基底

浸水深分布 (図 2) に対して特異値分解を適用して得た左特異ベクトルの空間分布を図 9 に示し。右特異ベクトルを図 10 に示す。ここで図 9, 図 10 では、モード 6 までの結果を示す。NNDSVD では、図 9 の左特異ベクトルを用いて式[12]より行列  $\mathbf{W}_0$  を構成し、図 10 の右特異ベクトルを用いて式[13]より行列  $\mathbf{H}_0$  を構成する。

先行研究<sup>10)</sup>では、対数変換および標準化を行った浸水深分布に対して特異値分解を適用した。なお、これらの変換により非負値条件が満たされなくなるため、本研究ではこれらの変換を行っていない。このため、先行研究<sup>10)</sup>の特異ベクトルと本研究の特異ベクトルは異なるが、大局的な傾向は同様である。これより、モード 1 は浸水深分布の平均的な分布傾向を表し、モード 2 以上は大すべり域の配置条件を表していると考えられる<sup>10)</sup>。特異値分解による空間相関特性の評価の詳細については、先行研究<sup>10)</sup>を参照されたい。ここで、モード 1～6 のモード合成によって元の浸水深分布を低次元近似したところ、近似浸水深分布と元の浸水深分布の相関係数は全ケースで 1.000 であり、ほぼ完全相関であった。

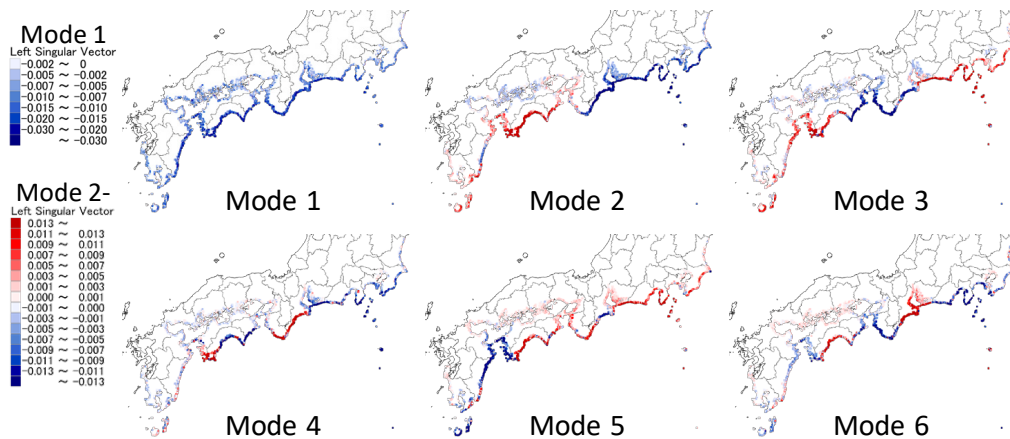


図 9 特異値分解におけるモード 6 までの左特異ベクトルの空間分布



図 10 特異値分解におけるモード 6 までの右特異ベクトル

(2) 非負値行列因子分解（初期化手法：NNDSVD）における基底数の違いによる近似精度の変動

全 11 ケースの浸水深分布に対して、基底数を  $2 \leq k \leq 11$  と設定した NMF-NNDSVD をそれぞれ適用した。その基底結合の近似精度を図 11 に示す。NMF-R による近似精度（図 4、図 11 の黒色の破線）と比較すると、全般的に近似精度は NMF-R のほうが高く、その傾向は基底数の増加に従って顕著となる。これについては、基底の特徴を踏まえて、5. (5) で考察を行う。

(3) 非負値行列因子分解（初期化手法：NNDSVD）による浸水深の空間分布特性

NMF-NNDSVD の適用にあたっては、既往研究<sup>10)</sup>および NMF-R との比較を目的として、基底数  $k=6$  を採用し、

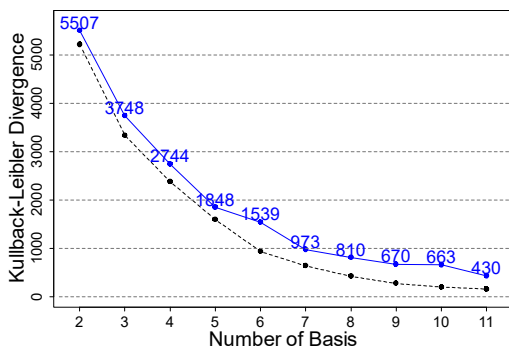


図 11 NMF-NNDSVD における基底数の違いによる近似精度の変動（黒色の破線：図 4 の NMF-R における近似精度）

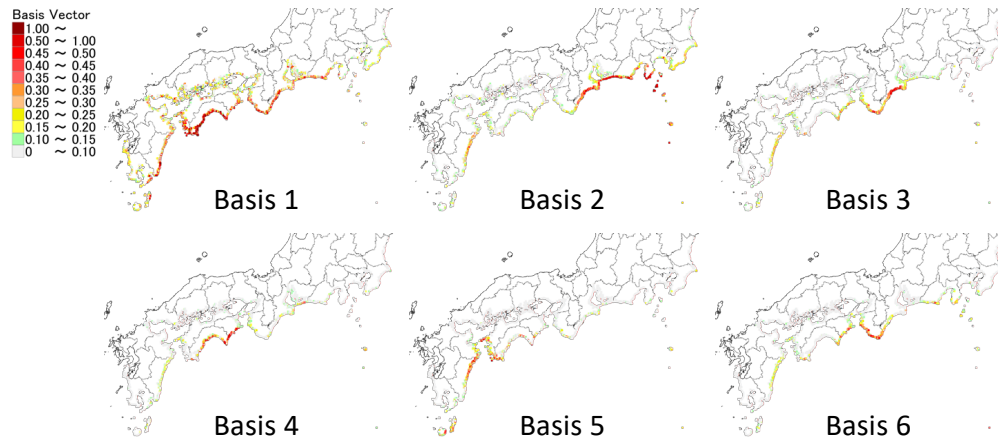


図 12 基底ベクトルの空間分布 (NMF-NNDSVD,  $k=6$ )

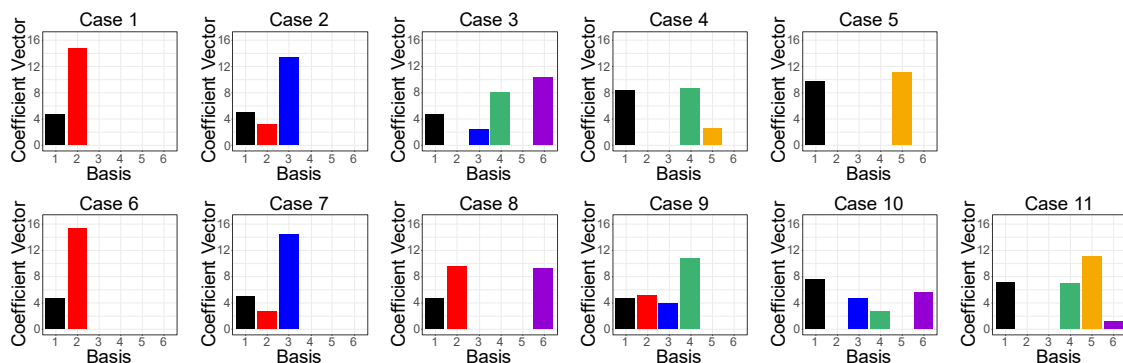


図 13 結合係数ベクトル (NMF-NNDSVD,  $k=6$ )

NMF-NNDSVD を適用した。NMF-NNDSVD より得た基底ベクトル  $\mathbf{w}$  の空間分布を図 12 に示し、結合係数ベクトル  $\mathbf{h}$  を図 13 に示す。基底ベクトルおよび結合係数ベクトルより、浸水深の空間分布特性の支配要因を考察する。

a) 基底 1

基底ベクトルは太平洋沿岸部のほぼ全域、特に四国（主に高知県）とその周辺で大きな値が分布し、内海になるにつれ値は小さくなっている。この分布傾向は、図 2 の浸水深分布で全般的に見られた。結合係数ベクトル（図 13）より基底 1 は、全 11 ケースで寄与している。また、大すべり域が高知県の近傍に設定された基本ケース 4, 5 とその派生ケース（ケース 10, 11）の結合係数は、他のケースよりも値が大きい。これより、基底 1 は対象領域全体（特に四国西部）での浸水深分布の規模を反映していると考えられる。

基底 1 は全ケースに寄与しているのに対し、基底 2～6 では、基底によって寄与するケースが異なっている。これより基底 1 が領域全体をカバーし、基底 2～6 は各ケースを特徴づけるものであると考えられる。

b) 基底 2

基底ベクトルは紀伊半島東部とそれ以東で大きな値が分布している。結合係数ベクトル（図 13）より基底 2 は、当該地域の近傍に大すべり域が設定された基本ケース 1, 2（特にケース 1）とその派生ケースの一部（ケース 6～9）で寄与している。これより、基底 2 は紀伊半島東部とそれ以東における大すべり域の配置条件を反映していると考えられる。

基底 3～6 も基底 2 と同様に大すべり域の配置条件を反映していると考えられる。なお、基底 3 では「紀伊半

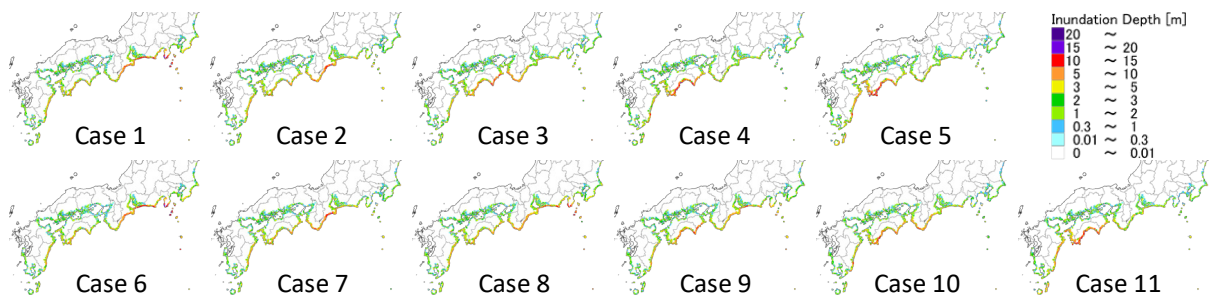


図 14 基底結合した近似浸水深分布 (NMF-NNDSVD,  $k=6$ )

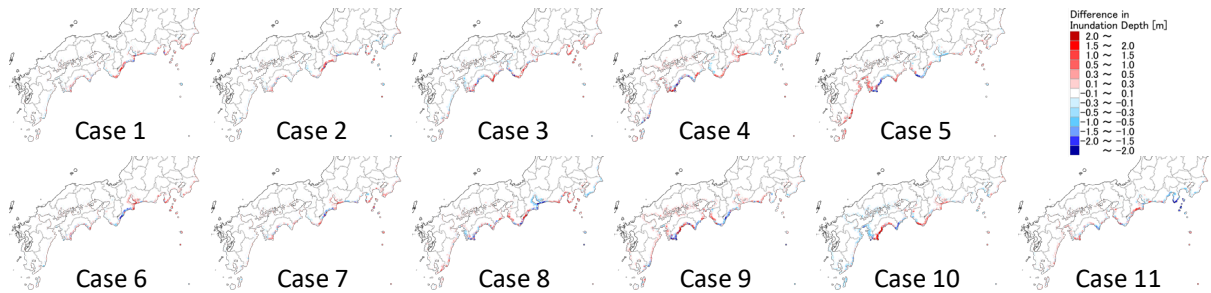


図 15 基底結合 (NMF-NNDSVD,  $k=6$ ) と元データとの浸水深差の分布

表 2 非負値行列因子分解と特異値分解による浸水深分布の影響要因

	非負値行列因子分解 (基底数6)		特異値分解	
	NMF-R (初期化手法: 一様乱数)	NMF-NNDSVD (初期化手法: 非負2重特異値分解)		
基底1	愛知～静岡における大すべり域の配置条件	対象領域全体 (特に四国西部) における浸水深分布の規模 ※全ケースに寄与する基底	平均的な浸水深分布形状 ※全ケースに寄与するモード	モード1
基底2	三重～愛知における	紀伊半島東部とそれ東における大すべり域の配置条件	紀伊半島を境にした大すべり域の配置条件	モード2
基底3	四国西部における	紀伊半島における	紀伊半島周辺とその両側での	モード3
基底4	四国東部における	四国東部における	対象領域を4分割した (九州沖, 四国沖, 紀伊半島沖, 愛知・静岡沖)	モード4
基底5	九州～四国西部における	九州～四国西部における	分割された領域における	モード5～11
基底6	紀伊半島における	紀伊半島西部における	※左特異ベクトル: 次数だけ断層の走向方向に分割 ※高次モード: 各ケースの細部の調整項	

島」, 基底 4 では「四国東部」, 基底 5 では「九州～四国西部」, 基底 6 では「紀伊半島西部」の近傍における大すべり域の配置条件であると考えられる。

各ケースは基底 1 とその他の基底によって近似されている。基底 1 の結合係数の大きさによって対象領域全体における浸水深レベルが表現され, 基底 2～6 によって各ケースの浸水深の空間的な特徴 (対象領域全体の中で特に浸水深の値が大きい領域) が表現されている。また, 行列  $H$  の結合係数ベクトルはそれぞれ 2～4 つのゼロ要素を含み, スパース性が向上している。これより, 各ケースは少数の基底によって近似されていることが読み取れる。

#### (4) 非負値行列因子分解 (初期化手法: NNDSVD) の基底結合による近似浸水深分布

図 14 に NMF-NNDSVD の基底結合による近似浸水深分布を示し, 図 15 に近似浸水深分布 (図 14) から元の浸水深分布 (図 2) を引いた浸水深差の分布を示す。これより, NMF-R と同様に, 基底結合による浸水深分布と元の浸水深分布とは空間分布に明瞭な違いは見られない。近似浸水深と元の浸水深の相関係数は, 0.969～0.995 であった。RMSE は 0.236～0.605 であった。これらより, 基底数 6 の NMF-NNDSVD における行列  $W, H$  は行列  $X$  を高精度で低次元近似しているといえる。

#### (5) 異なる初期化手法を用いた非負値行列因子分解の比較

以上の基底数 6 における NMF-R と NMF-NNDSVD による浸水深の空間分布特性の評価結果を表 2 にまとめる。参考までに, 図 9 の特異値分解による浸水深の空間相関特性の評価結果についても記す。なお, NMF-R による基底は試行により異なった結果が得られるため, 順序には任意性がある。一方, NMF-NNDSVD による基底は特異値分解のモード順に対応付けられている。

まず NMF により分解された基底について比較する。NMF-R を用いた場合は, いずれの基底も, 津波断層モデルの大すべり域の分布に深く関連したものとなっている。一方, NMF-NNDSVD を用いた場合は, 基底 1 は対象領域全体をカバーする空間分布を表し, 全ケースに寄与している。基底 2～6 については, 津波断層モデルの大すべり域の分布に関連した基底となっている。NMF-R では NMF-NNDSVD の基底 1 (領域全体をカバーし, 全ケースに寄与) のような基底は現れていない。また NMF-NNDSVD では, 領域全体をカバーする基底 1 と各ケースを特徴づける基底 2～6 によって, 結合係数行列  $H$  のスパース性が高く, 分解結果の解釈を容易としている。

一方, 特異値分解におけるモードは, 式[5]に示されるように固有値の降順であるため, 高次モードになるほどモードの重みは減少して重要度が低下し, 左特異ベクトルの空間分布 (図 9) は複雑な形状となり, 解釈が困難



となる<sup>10)</sup>。これに対して NMF では、式[2]の結合係数ベクトルで表される基底の重みは、各ケースを特徴づけるものであり、相対的な重要度に差異はなく、基底の空間分布が複雑化することもないため、分解・合成の解釈が容易である。特に NMF-NNDSVD では、基底の空間分布が局在化され、特徴が単純化されている。

次に分解結果の安定性について比較する。NMF-R では初期値が一様乱数によって決定されるため、試行ごとに分解結果が異なる。一方、NMF-NNDSVD では初期値が特異値分解によって一意に定まるため、分解結果もまた一意に定まり、安定的な結果が得られる。

最後に NMF の近似結合の精度について比較する。NMF の近似浸水深分布と元の浸水深分布の相関係数は、NMF の初期化手法に依らずほぼ同程度であり、強い正相関(0.969 以上)であった。一方、カルバック・ライブラー情報量による残差や RMSE を基準とした場合、NMF-R が優位であった。ただしこれはスパース性とトレードオフであり、図 6 に示したように NMF-R による基底が各ケースに複雑に寄与していることに起因する。

## 6. おわりに

本研究では、複数ケースの浸水深分布の空間分布特性の評価手法として非負値行列因子分解(NMF)を提案した。NMF による評価手法では、先行研究<sup>10)</sup>の特異値分解を用いた評価手法の問題点(負値を含む基底の合成の意味付けの困難さ)を「非負基底の加算のみによる合成」によって解決している。数値計算例として、南海トラフ巨大地震において想定される 11 ケースの津波浸水深分布<sup>2)</sup>に対して、基底数を 6 とした NMF を適用した。NMF における初期化手法として、一様乱数と非負 2 重特異値分解(NNDSVD)の 2 種類の手法を用いた。得られた知見を以下にまとめる。

- (1) 初期化手法に一様乱数を用いた NMF (NMF-R) によって得られた基底は、超大すべり域を含む大すべり域の配置条件を反映している。なお、各基底は大すべり域の配置条件が深く関連したケースに寄与している。
- (2) 初期化手法に NNDSVD を用いた NMF (NMF-NNDSVD) によって得られた基底 1 は全ケースに寄与しており、その空間分布は対象領域全体的特徴量を示し、領域全体での浸水深分布の規模を反映している。基底 2~6 は NMF-R と同様に大すべり域の配置条件を反映している。
- (3) 結合係数行列  $\mathbf{H}$  のスパース性に関しては、NMF-NNDSVD は NMF-R よりも優位であり、基底に基づく各ケースの解釈が容易である。空間分布特性の抽出に関しても、特異値分解により初期値が一意に規定される NMF-NNDSVD の方が、一様乱数に基づく NMF-R よりも安定的な結果が得られる。
- (4) NMF-NNDSVD と NMF-R による近似結合を行い、それぞれ元の浸水深分布の近似精度を比較した。相関係数は両初期化手法で同程度であるが、カルバック・ライブラー情報量や RMSE に関しては多くの基底が複雑に寄与する NMF-R が優位であり、スパース性とトレードオフの関係にあることがわかった。

本研究では 11 ケースの浸水深分布を対象としたが、浸水深の空間分布特性と津波断層パラメータとの関連性<sup>20)</sup>や空間分布を網羅するために必要な基底数などについて

は、さらに検討が必要である。防災科学技術研究所の津波ハザードステーション(J-THIS)<sup>21)</sup>では、「波源断層を特性化した津波の予測手法(津波レシピ)<sup>22)</sup>」に基づき、すべり角やすべり量、震源域の組み合わせなどが異なる多種多様な津波断層モデルによる浸水深分布が想定されている。この浸水深分布が利用可能になれば、さらに多様な浸水深分布を対象とした検討が可能になるため、今後の課題としたい。

## 謝辞

本研究は JSPS 科研費 JP20H02413 の助成を受けたものである。また本研究で使用した浸水深分布は、南海トラフの巨大地震モデル検討会<sup>1)</sup>の第二次報告<sup>2)</sup>における成果として、内閣府よりデータ提供を受けたものである。記して、謝意を表する次第である。

## 参考文献

- 1) 内閣府：南海トラフの巨大地震モデル検討会、<http://www.bousai.go.jp/jishin/nankai/model/> (最終アクセス 2021 年 2 月 9 日)
- 2) 南海トラフの巨大地震モデル検討会：(第二次報告)津波断層モデル編—津波断層モデルと津波高・浸水域等について—, 2012, 100p.
- 3) 地震調査研究推進本部地震調査委員会：南海トラフ沿いで発生する大地震の確率論的津波評価, 2020, 78p.
- 4) (国研)防災科学技術研究所：南海トラフ沿いの地震に対する確率論的津波ハザード評価—第一部 本編—, 防災科学技術研究所研究資料, 第 439 号, 2020.
- 5) 今村文彦, 首藤伸夫：津波高さに及ぼす断層パラメータ推定誤差の影響. 海岸工学論文集, Vol. 36, pp.172-182, 1989.
- 6) 栗田哲史, 松山昌史, 内野大介：東北地方太平洋沖地震津波の痕跡値を用いた津波解析のばらつき評価, 土木学会論文集 B2(海岸工学), Vol. 39, No. 2, pp.I-216-I\_220, 2013.
- 7) 安田誠宏, 丸山拓真, Katsuichiro Goda, 森信人, 間瀬肇：確率津波モデルを用いた南海トラフ巨大地震津波の不確実性評価, 土木学会論文集 B2(海岸工学), Vol. 71, No. 2, pp.I-295-I\_300, 2015.
- 8) 小谷拓磨, 高瀬慎介, 森口周二, 寺田賢二郎, 福谷陽, 大竹雄, 野島和也, 桜庭雅明：応答曲面を用いた数値解析援用確率論的津波ハザード評価, 土木学会論文集 A2(応用力学), Vol. 72, No. 1, pp.I\_1381-I\_1385, 2016.
- 9) Kotani, T., Moriguchi, S., Terada, K., Takase, S., Otake, Y., Fukutani, Y. and Sakuraba, M.: A Tsunami Risk Analysis Considering Correlations Among Coastal Cities, 7<sup>th</sup> Asia Conference on Earthquake Engineering, Paper No.ACEE0095, 2018.
- 10) 高橋幸宏, 能島暢呂：南海トラフ巨大地震による津波の浸水深分布の空間相関特性の評価とシミュレーション, 地域安全学会論文集, No.36, 論文番号 8, pp.1-8, 2020.3.
- 11) 能島暢呂, 久世益充, LE QUANG DUC：シナリオ地震動予測地図の特異値分解によるモード分解と地震動分布のシミュレーション, 日本地震工学会論文集, Vol. 18, No. 2, pp.2\_95-2\_114, 2018.
- 12) 能島暢呂, 久世益充, 高橋幸宏：シナリオ地震動予測地図のモード分解による予測震度分布の空間相関特性の分析, 日本地震工学会論文集, Vol. 19, No. 1, pp.1\_121-1\_135, 2019.

- 13) 能島暢呂, 高橋幸宏: 複数ケースの震源パラメータ設定がシナリオ地震動予測地図の空間相関特性に及ぼす影響について, 東濃地震科学研究所報告, No. 45, pp.53-66, 2020.
- 14) 亀岡弘和: 非負値行列因子分解, 計測と制御, Vol. 51, No. 9, pp. 835-844, 2012.
- 15) Boutsidis, C. and Gallopoulos, E.: SVD based initialization: A head start for nonnegative matrix factorization, Pattern Recognition, Vol. 41, pp. 1350-1362, 2008.
- 16) R Core Team: R: A language and environment for statistical computing, 2020, <https://www.R-project.org/>. (最終アクセス: 2021年2月9日)
- 17) Gaujoux, R. and Seoighe, C.: A flexible R package for nonnegative matrix factorization, BMC Bioinformatics, Vol. 11, Article number 367, pp. 1-9, 2010.
- 18) Brunet, J. P., Tamayo, P., Golub, T. R. and Mesirov, J. P.: Metagenes and molecular pattern discovery using matrix factorization, Proceedings of the National Academy of Sciences of the United States of America, Vol. 101, No. 12, pp. 4164-4169, 2004.
- 19) 金谷健一: これなら分かる応用数学教室 最小二乗法からウェブレット, 共立出版, 2003, 270p.
- 20) 外里健太, 小谷拓磨, 波多野僚, 高瀬慎介, 森口周二, 寺田賢二郎, 大竹雄: 数値解析結果の空間モード分解による津波のリスク評価, 日本計算工学会論文集, Paper No. 20200003, 2020.
- 21) (国研)防災科学技術研究所: 津波ハザードステーション (J-THIS) : <http://www.j-this.bosai.go.jp/> (最終アクセス: 2021年2月9日)
- 22) 地震調査研究推進本部地震調査委員会: 波源断層を特性化した津波の予測手法 (津波レシピ), 2017, 33p.

(原稿受付 2020.8.23)  
(登載決定 2021.1.9)