

# 地震波形の類似度評価におけるワッサーズタイン計量の特性とセンブランス解析への適用

奈良 樹<sup>1</sup>・後藤 浩之<sup>2</sup>

<sup>1</sup> 学生会員 京都大学大学院 工学研究科都市社会工学専攻 (〒611-0011 京都府宇治市五ヶ庄)

E-mail: nara@catfish.dpri.kyoto-u.ac.jp

<sup>2</sup> 正会員 京都大学准教授 防災研究所 (〒611-0011 京都府宇治市五ヶ庄)

E-mail: goto@catfish.dpri.kyoto-u.ac.jp (Corresponding Author)

地震波形の類似度評価は地震工学および地震学における重要な要素の一つである。近年、地震波形の類似度評価の新たな指標としてワッサーズタイン計量を用いることが提案されている。ある問題においてワッサーズタイン計量が有効であることは示されているが、地震工学の分野において波形の類似度評価の指標としてワッサーズタイン計量を適用した研究はなされていない。本研究では、波形の類似度指標として旧来から用いられている平均二乗誤差、および新たな指標であるワッサーズタイン計量の性質をそれぞれ詳細に調べた後、センブランス解析に対してワッサーズタイン計量を適用し、平均二乗誤差を用いる旧来の手法との比較を行った。

**Key Words:** Wasserstein metric, MSE, seismic waveform, similarity, semblance analysis

## 1. はじめに

2つの地震波形の振幅や位相の特徴がどの程度類似しているか（以下、地震波形の類似度と呼ぶ）を定量的に評価することは、地震工学および地震学における重要な要素の一つである。

地盤構造や震源過程を推定するための逆解析では、あるパラメータから作成される人工波形と実際の観測波形との類似度を評価する必要がある。例えばFWI (Full Waveform Inversion) は、仮定した地下構造から人工波形を作成して実際の観測波形との類似度を評価し、その類似度が最大となるように地下構造のパラメータを更新していくことで実際の地下構造を推定する手法である<sup>1)</sup>。また、震源インバージョンは、仮定した震源断層モデルから作成した人工波形と実際の観測波形との類似度を評価し、その類似度が最大となるようにモデルパラメータを更新していくことで、実際にすべりが生じた領域、すべり量、すべり継続時間、破壊伝播速度を推定する手法である<sup>2)</sup>。観測波形同士の類似度を評価することもある。例えばセンブランス解析は、地震計アレイを構成する複数の観測点で得られた観測波形同士で類似度を評価することで、波群の到来方向と伝播速度を推定する手法である<sup>3)</sup>。

地震波形の類似度評価の指標として、一般には波形の平均二乗誤差が使用される。また、震源インバージョンでは波形の包絡線の平均二乗誤差も組み合わせて使用される<sup>4)</sup>。低周波数側では波形そのものの平均二乗誤差が、高周波数側では波形の包絡線の平均二乗誤差が類似度評価の指標として使用される。しかし、どの周波数帯で指標を切り替えるべきかは明確にされていない。

近年、地震波形の類似度評価の新たな指標として、最適輸送理論の距離関数の一つであるワッサーズタイン計量を用いることが提案されている<sup>5)</sup>。地震波形の類似度評価の指標としてワッサーズタイン計量を用いると、FWIにおいて有効であることが示されている<sup>6)</sup>。しかし、地震波形の類似度評価の指標として、波形やその包絡線の平均二乗誤差、ワッサーズタイン計量を用いた場合で、それぞれどのような性質があるかについて詳細に議論した研究はなされていない。また、FWI以外に地震工学の分野において波形の類似度評価の指標としてワッサーズタイン計量を適用した研究もなされていない。

そこで本研究では、地震波形の類似度評価の指標として、波形やその包絡線の平均二乗誤差、ワッサーズタイン計量を用いた場合でそれぞれどのような性質があるのかについて、様々なパターンの地震波形で考察する。ま

た、その考察を踏まえ、大阪市福島区の地震計アレイにおいて観測された 2018 年大阪府北部の地震の波形を用いてセンブランス解析を行い、ワッサーズタイン計量を用いた場合の手法と平均二乗誤差を用いる旧来の手法との比較を行う。

## 2. ワッサーズタイン計量の概要

本研究では、ワッサーズタイン計量を用いて地震波形の類似度評価を行う。本章では、最適輸送理論の概要とワッサーズタイン計量の定義、および地震波形に対してワッサーズタイン計量を適用するための変換方法を説明する。

### (1) 最適輸送理論

ワッサーズタイン計量は最適輸送理論の中で定義される距離関数である。まずこの最適輸送理論の概要を説明する。最適輸送理論とは、物質をある場所からある場所へ輸送するときの最小コストとそのときの輸送経路を求めるための理論体系である。この理論はフランスの数学者、ガスパール・モンジュによって提唱された<sup>9)</sup>。本研究では、1次元空間上での最適輸送理論を考える。

まずは最適輸送理論の簡単な例を考える。一次元空間上に相異なる位置 $p_1, p_2, \dots, p_n$ に同じ重さの物質があり、それを相異なる位置 $q_1, q_2, \dots, q_n$ に移動させるとする。 $p_i$ から $q_j$ への移動には $p_i$ と $q_j$ の距離 $d(p_i, q_j)$ に比例したコスト $cd(p_i, q_j)$ がかかる( $c$ は正の定数)。全部で $n!$ 通りの輸送方法があるが、その中に輸送コストが最小となる最適な輸送があると考えられる。

この問題について、 $n$ を限りなく大きくし、物質が連続的に分布している場合を考える。密度が一定で連続的に分布している物質の集合 $D_1, D_2$ を考え、それぞれの集合の物質の総量は等しいとする。 $c(x, y)$ を位置 $x$ の物質を位置 $y$ に輸送するのに要する物質単位量あたりのコストとし、 $D_1$ から $D_2$ へ物質をすべて輸送する場合について、次の全単射条件式(1)(2)と物質保存条件式(3)を満たすような輸送写像 $T$ を考える。

$$x \neq x' \Rightarrow T(x) \neq T(x') \quad (1)$$

$$T(D_1) = \{T(x) | x \in D_1\} = D_2 \quad (2)$$

$$u \subset D_1 \Rightarrow |T(u)| = |u| \quad (3)$$

ただし、 $u$ は集合 $D_1$ の部分集合である。この場合での最適輸送問題は、この条件のもとで総コスト

$C(T) = \int_{D_1} c(x, T(x)) dx$ を最小にする輸送写像 $T$ を求める問題となる。

さらに、位置によって物質の密度が異なる場合の輸送問題について考える。物質が一次元の空間 $X$ に連続的に分

布しているとし、輸送の前後で総質量は不変とする。簡単のために物質の総質量をどちらも1とする。このとき物質の分布を確率測度とみなすことができ、輸送前の分布を $\mu$ 、輸送後の分布を $\nu$ と表すことにする。 $d\mu(x)$ 、 $d\nu(y)$ はそれぞれ位置 $x, y$ における物質の質量を表す。 $\int f(x) dx = 1$ 、 $\int g(y) dy = 1$ を満たすような非負値関数 $f(x)$ 、 $g(y)$ を考え、 $f(x)$ が輸送前の位置 $x$ における密度を、 $g(y)$ が輸送後の位置 $y$ における密度を表すとすると、

$$d\mu(x) = f(x) dx \quad (4)$$

$$d\nu(y) = g(y) dy \quad (5)$$

が成立する。このとき、 $\mu$ を $\nu$ へ移す輸送法を $X \times X$ 上の確率測度 $\pi$ で表し、 $d\pi(x, y)$ を $x$ から $y$ へ輸送される物質の質量と捉える。すると $x$ から輸送される質量は $d\mu(x)$ 、 $y$ へ輸送される質量は $d\nu(y)$ なので、

$$\int_{y \in X} d\pi(x, y) = d\mu(x) \quad (6)$$

$$\int_{x \in X} d\pi(x, y) = d\nu(y) \quad (7)$$

が成立する。この式(6)、(7)を満たす確率測度 $\pi$ を $\mu, \nu$ のカップリングと呼び、カップリングがなす空間を $\Pi(\mu, \nu)$ と書く。また、物質1個を $x$ から $y$ へ輸送するためのコスト $c(x, y)$ は $X \times X$ 上の関数とみなせる。そして $x$ にある物質をすべて $y$ へ輸送するには $c(x, y) d\pi(x, y)$ のコストがかかるから、 $\mu$ を $\nu$ へカップリング $\pi$ で輸送するための総コストは、

$$C(\pi) = \int_{X \times X} c(x, y) d\pi(x, y) \quad (8)$$

と表される。よって、密度が位置によって変動し、連続的に物質が分布している場合の最適輸送問題は、「 $\Pi(\mu, \nu)$ において $C$ の最小値を求め、その最小値を達成するカップリングを具体的に求めよ」となる。最適輸送問題の解の存在性について議論する必要があるが、総コスト $C(\pi)$ の最小値とそれを達成する元が存在し、さらにそれらの一意性についても示されている<sup>9)</sup>。

### (2) ワッサーズタイン計量の定義

式(8)においてコスト関数を $c(x, y) = |x - y|^p$ とし、最適輸送問題の解として一般の $p$ 次ワッサーズタイン計量 $W_p(\mu, \nu)$ は次式のように定義される。

$$W_p(\mu, \nu) = \left\{ \min_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\pi(x, y) \right\}^{\frac{1}{p}} \quad (9)$$

本研究では $p = 2$ のとき、すなわち $(x - y)^2$ をコスト関数としたときの最適輸送問題の解である2次ワッサーズタイン計量 $W_2$ を用いる。

物質を $\mu$ から $\nu$ へ輸送するときの輸送写像を $T$ とする。 $T$ によって $\mu$ にあった物質はすべて $\nu$ へ輸送される。この

とき輸送前のある位置 $A$ にあった物質は位置 $T(A)$ へ輸送されると考えられるので、輸送前の質量 $\mu[A]$ と輸送後の質量 $\nu[T(A)]$ に等しい。この質量保存条件を満たした輸送写像 $T$ の集合を $\tau(\mu, \nu)$ とし、輸送写像 $T$ を用いて2次ワッサースタイン計量 $W_2$ は次のように書ける。

$$W_2(\mu, \nu) = \left\{ \min_{T \in \tau(\mu, \nu)} \int_{\mathbb{R}^d} |x - T(x)|^2 d\mu(x) \right\}^{\frac{1}{2}} \quad (10)$$

ここで、確率密度関数 $f(x)$ 、 $g(y)$ を用いれば、式(10)は次のようにも書ける。

$$W_2(f, g) = \left\{ \min_T \int_X |x - T(x)|^2 f(x) dx \right\}^{\frac{1}{2}} \quad (11)$$

なお本論文では以降、断りのない場合は式(11)の両辺を二乗した $W_2^2$ をワッサースタイン計量と呼ぶことにする。

目的関数 $\int_X |x - T(x)|^2 f(x) dx$ が最小となるときに輸送写像 $T(x)$ は、 $f(x)$ 、 $g(y)$ の累積分布関数を $F(x)$ 、 $G(y)$ とすれば、

$$T(x) = G^{-1} \circ F \quad (12)$$

が成り立ち、ワッサースタイン計量は

$$W_2^2(f, g) = \int_X |x - G^{-1}(F(x))|^2 f(x) dx \quad (13)$$

と与えられることが知られている<sup>7)</sup>。 $F = s$ の変数変換を施すと、式(13)は次のように変形できる。

$$W_2^2 = \int_0^1 |F^{-1}(s) - G^{-1}(s)|^2 ds \quad (14)$$

本論文では以上を踏まえて、式(14)によりワッサースタイン計量 $W_2^2$ を具体的に計算する。

### (3) 地震波形の変換方法

上述のようにワッサースタイン計量が累積分布関数の逆関数を用いて求められることから、正負どちらの値も取りうる地震波形 $f(t)$ 、 $g(t)$ を非負の値 $f_p(t)$ 、 $g_p(t)$ に変換し、さらにそれを正規化して確率密度関数 $\tilde{f}(t)$ 、 $\tilde{g}(t)$ に変換する必要がある。

非負の値に変換する手法はいくつか提案されているが、ここでは次式の変換方法<sup>8)</sup>を用いる。

$$f_p(t) = \ln(e^{af(t)} + 1) \quad (15)$$

ここで、 $a$ は次式で定義される定数である

$$a = \frac{3.0}{\max|f(t)|} \quad (16)$$

この $f_p(t)$ を次式を用いて正規化し、確率密度関数とみなす。

$$\tilde{f}(t) = \frac{f_p(t)}{\int f_p(t) dt} \quad (17)$$

類似度を評価する対象の地震波形 $g(t)$ についても同様の変換を行うことにより、式(14)に基づいてワッサース

タイン計量 $W_2^2$ を計算することができる。

## 3. 平均二乗誤差とワッサースタイン計量の比較

本章では、地震波形の類似度評価の指標として、波形やその包絡線の平均二乗誤差（以下、MSEと表記する）を用いた場合と、ワッサースタイン計量を用いた場合でそれぞれどのような特性があるのか、様々なパターンの地震波形で考察する。

### (1) 単一リッカー波の場合

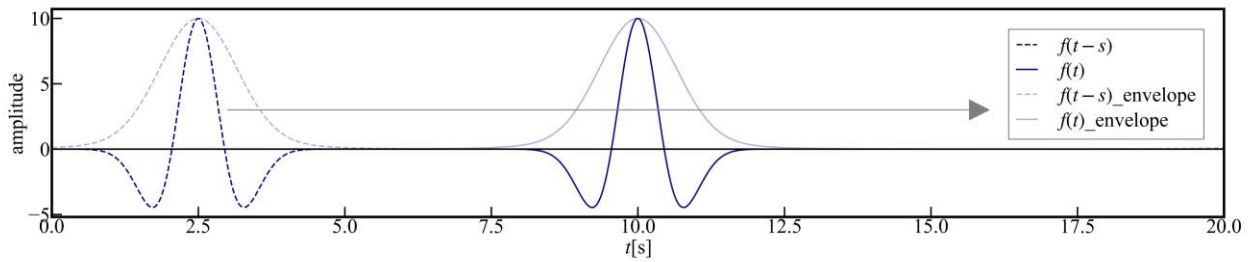
まずは、フェーズの明瞭な波形について比較するために、図-1のような0.5[Hz]の単一リッカー波を考え、時間ずれ $s$ に対して各類似度指標がどのような値を示すか考察する。これは、低周波数成分が卓越する波形を比較するケースに対応した例とみなせる。ノイズなしの波形（図-1(a)）および平均0のピンクノイズが加えられた波形（図-1(b)）について検討した。ただし、ノイズには0.1[Hz]~1.0[Hz]の周波数成分のみを持つようにバンドパスフィルタをかけている。また、それぞれの波形の包絡線もともに図-1に描画している。

2つの波形が、時間ずれ $s$ の大きさによってどれだけ類似度が変動するか考察する。ノイズなし、およびノイズの標準偏差が5、10、15、20、25、30の場合について、類似度評価の指標としてMSE、包絡線のMSE、 $W_2^2$ を用いた結果をそれぞれ図-2、3、4に示す。いずれの類似度指標も0に近いほど二つの波形間の類似度が大きいと評価し、値が0であれば波形は完全に一致する。

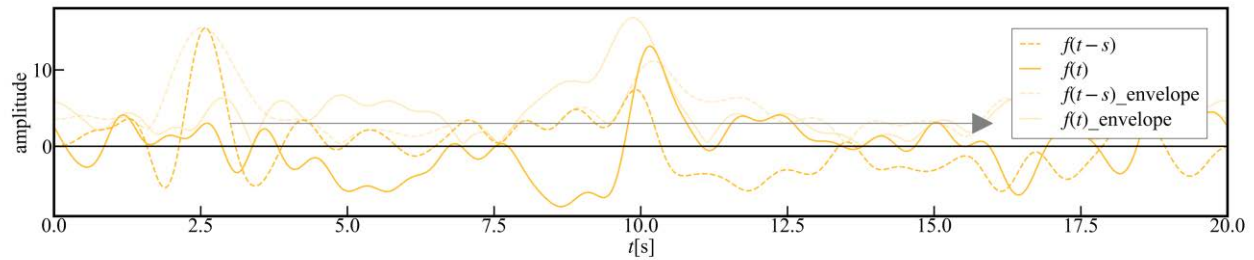
MSE（図-2）は、ノイズが小さいとき時間ずれがほぼ0[s]で類似度が最大となる一方、ノイズが大きくなると類似度が最大となる時の時間ずれ $\tau$ が0[s]からかなり乖離することが分かる。これは図-1(b)にみられるように、ノイズの長周期成分の凹凸が偶然に合致するためと考えられる。包絡線のMSE（図-3）は、ノイズが大きくなるにつれて類似度が最大となる時の時間ずれ $\tau$ が0[s]から少しずつ乖離していく傾向があることが分かる。したがって、ノイズが大きくない条件下ではMSEは包絡線のMSEよりも適切に時間ずれを特定できると考えられ、1章で述べたような、低周波数成分の波形では一般的に波形そのものの平均二乗誤差が類似度の指標として用いられることの理由を示しているといえる。ワッサースタイン計量 $W_2^2$ （図-4）は、ノイズが小さいときに時間ずれが0[s]のときにほぼ類似度が最大となっていることが分かるが、ノイズが少しでも大きくなると類似度が最大となる時の時間ずれ $\tau$ が0[s]から乖離することがわかる。しかし、図-5からわかるように、ノイズが小さいときには時間ずれ $s$ [s]に対する $W_2^2$ はほぼ凸形状を描

き、これが MSE や包絡線の MSE と異なる点である。実際に、同一の波形の時間シフトにおいて  $W_2^2$  が完全な凸形状になることについては厳密な証明がなされている<sup>9)</sup>。

波形に付加するノイズをランダムに変えながら各ケースについて 100 回繰り返し、類似度指標が最大となる時間ずれ  $\tau$  が時間ずれ 0[s] とどの程度乖離があるかを統計



(a) ノイズなし



(b) ノイズあり(標準偏差 30)

図-1 単一のリッカー波の時間シフト

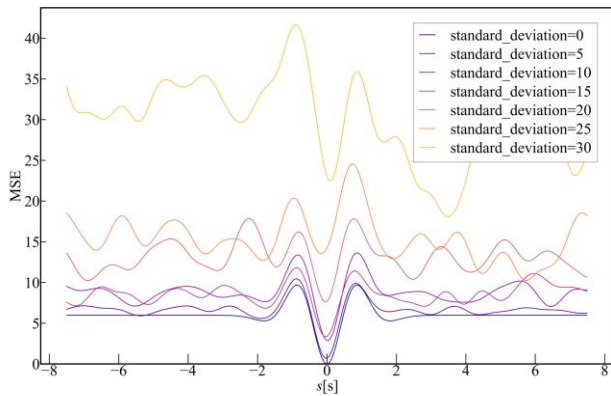


図-2 図-1 の  $f(t)$  と  $f(t-s)$  の類似度(MSE)と時間ずれ  $s$  の関係

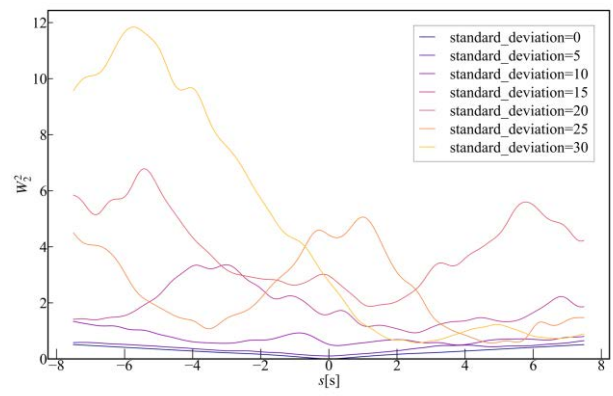


図-4 図-1 の  $f(t)$  と  $f(t-s)$  の類似度( $W_2^2$ )と時間ずれ  $s$  の関係

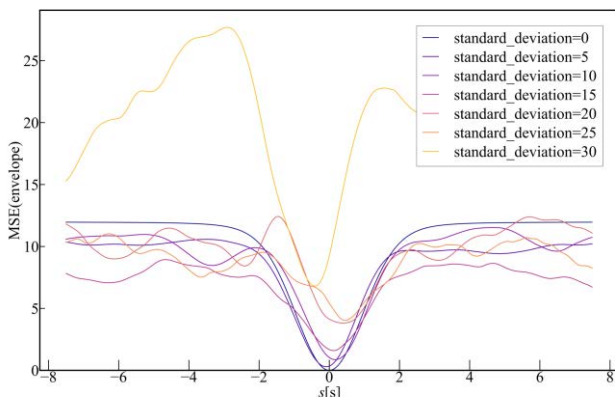


図-3 図-1 の  $f(t)$  と  $f(t-s)$  の類似度(包絡線の MSE)と時間ずれ  $s$  の関係

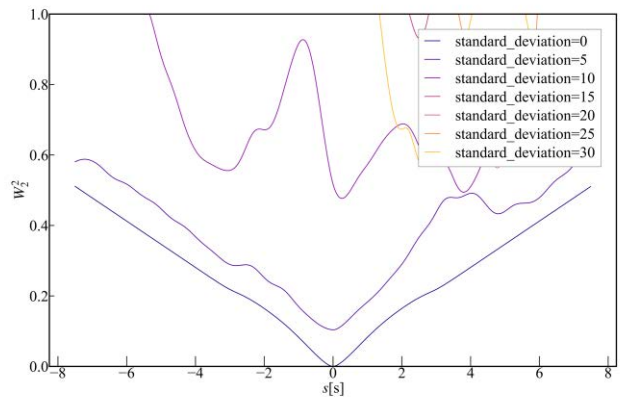


図-5 図-4 の  $0 \leq W_2^2 \leq 1$  の拡大図

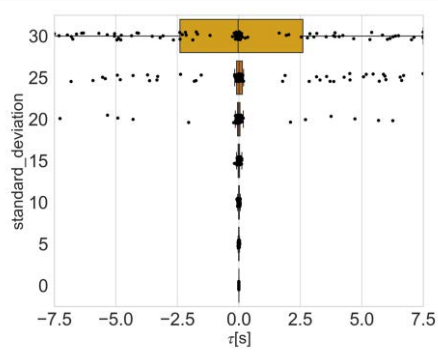


図-6 MSEが最小(類似度が最大)となる時間ずれ $\tau$ とノイズの標準偏差の関係

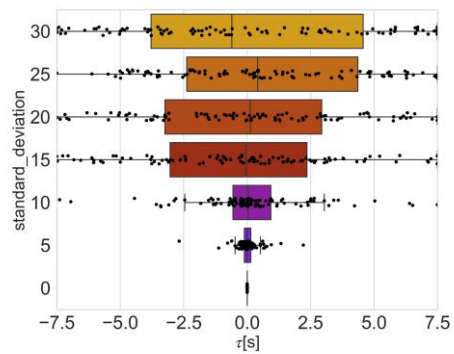


図-8  $W_2^2$ が最小(類似度が最大)となる時間ずれ $\tau$ とノイズの標準偏差の関係

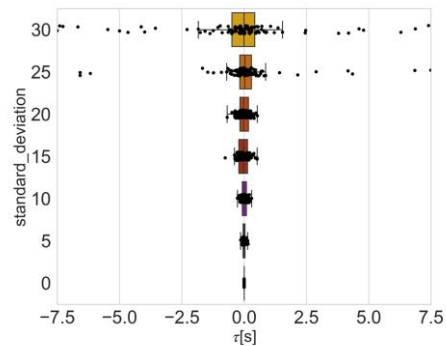


図-7 包絡線のMSEが最小(類似度が最大)となる時間ずれ $\tau$ とノイズの標準偏差の関係

的に考察する。図-6, 7, 8にその結果を示す。横軸は類似度が最大となる時間ずれ $\tau$ 、縦軸はノイズの標準偏差の大きさである。また、プロットされている黒点は各試行結果を、箱ひげ図は試行結果の分布に対応したものを

示している。図-6, 7 から、ノイズが小さいときにはMSEの方が包絡線のMSEよりも類似度が最大となる時間ずれ $\tau$ が0[s]に近い傾向があるが、ノイズの標準偏差が30のときは包絡線のMSEの方が $\tau$ のばらつきは小さくなる傾向にあることが分かる。一方、 $W_2^2$ は標準偏差が10以上でばらつきが大きくなる傾向にある。

## (2) 重ね合わせたリッカー波の場合(パターン1)

次にフェーズの変化が密な波形について考えるために、図-9のような2~4[Hz]のリッカー波を複数重ね合わせた波形を考え、(1)と同様に時間ずれ $s$ に対して各類似度指標がどのような値を示すか考察する。これは高周波数成分が卓越する波形を比較するケースに対応した例とみなせる。ノイズなしの波形(図-9(a))および平均0のピンクノイズを加えた波形(図-9(b))について検討した。ただし、ノイズには0.5[Hz]~10[Hz]の周波数成分のみを持

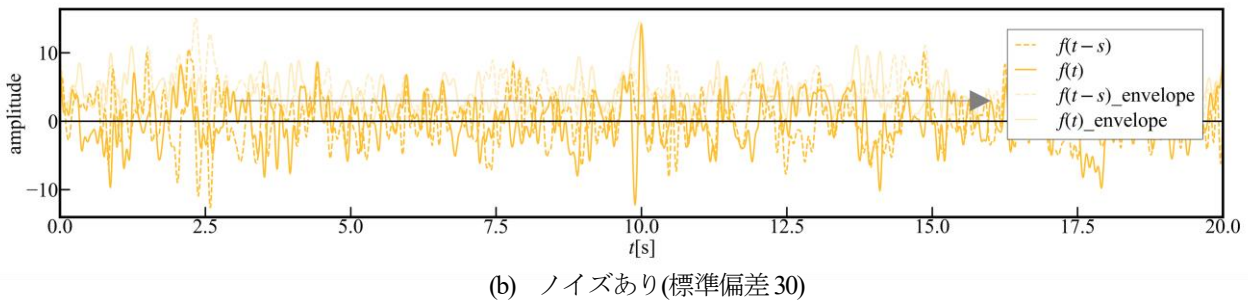
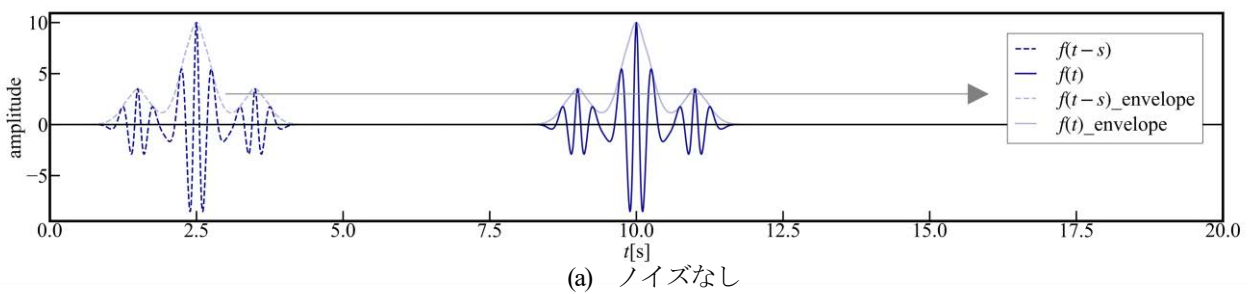


図-9 重ね合わせたリッカー波(パターン1)の時間シフト

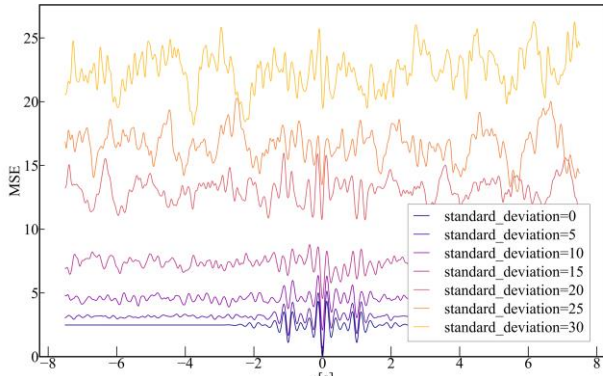


図-10 図-9の  $f(t)$  と  $f(t-s)$  の類似度(MSE)と時間ずれ  $s$  の関係

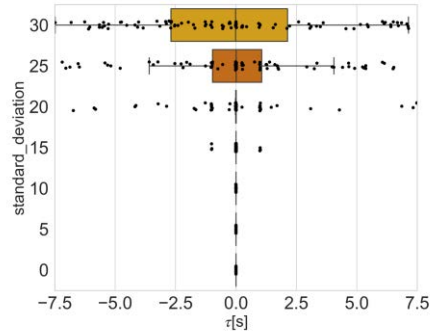


図-13 MSEが最小(類似度が最大)となる時間ずれ  $\tau$  とノイズの標準偏差の関係

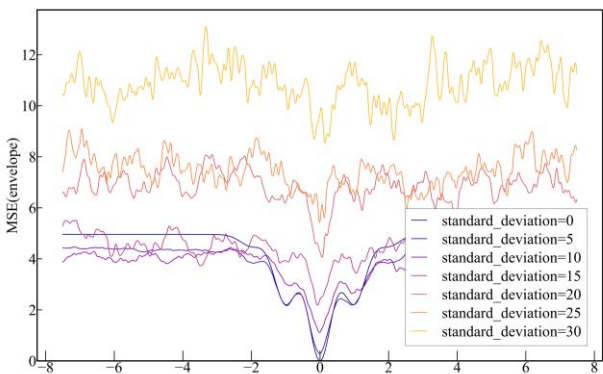


図-11 図-9の  $f(t)$  と  $f(t-s)$  の類似度(包絡線のMSE)と時間ずれ  $s$  の関係

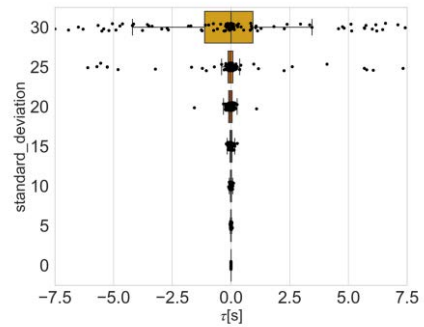


図-14 包絡線のMSEが最小(類似度が最大)となる時間ずれ  $\tau$  とノイズの標準偏差の関係

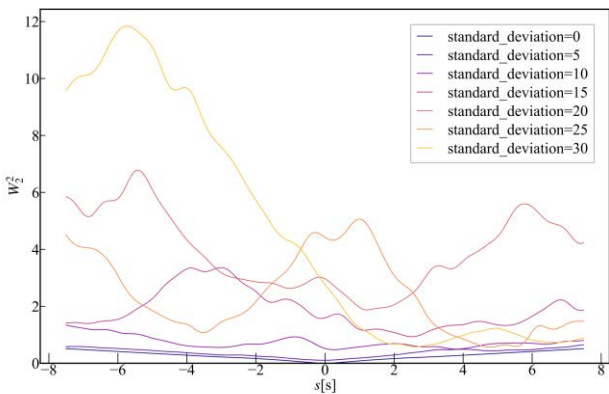


図-12 図-9の  $f(t)$  と  $f(t-s)$  の類似度( $W_2^2$ )と時間ずれ  $s$  の関係

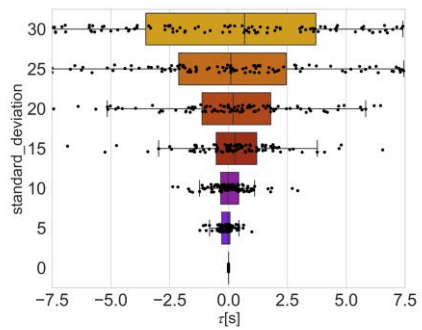


図-15  $W_2^2$ が最小(類似度が最大)となる時間ずれ  $\tau$  とノイズの標準偏差の関係

つようにバンドパスフィルタをかけている。また、それぞれの波形の包絡線もともに図-9に描画している。

2つの波形が、時間ずれ  $s$  の大きさによってどれだけ類似度が変動するか考察する。ノイズなし、およびノイズの標準偏差が 5, 10, 15, 20, 25, 30 の場合について類似度評価の指標として MSE, 包絡線の MSE,  $W_2^2$  を用

いた結果を、それぞれ図-10, 11, 12に示す。

波形に付加するノイズをランダムに変えながら各ケースについて 100 回繰り返し、類似度指標が最大となる時間ずれ  $\tau$  が時間ずれ 0[s] とどの程度乖離があるかを統計的に考察する。図-13, 14, 15 にその結果を示す。図-13, 14 から、ノイズが小さいときには MSE の方が包絡

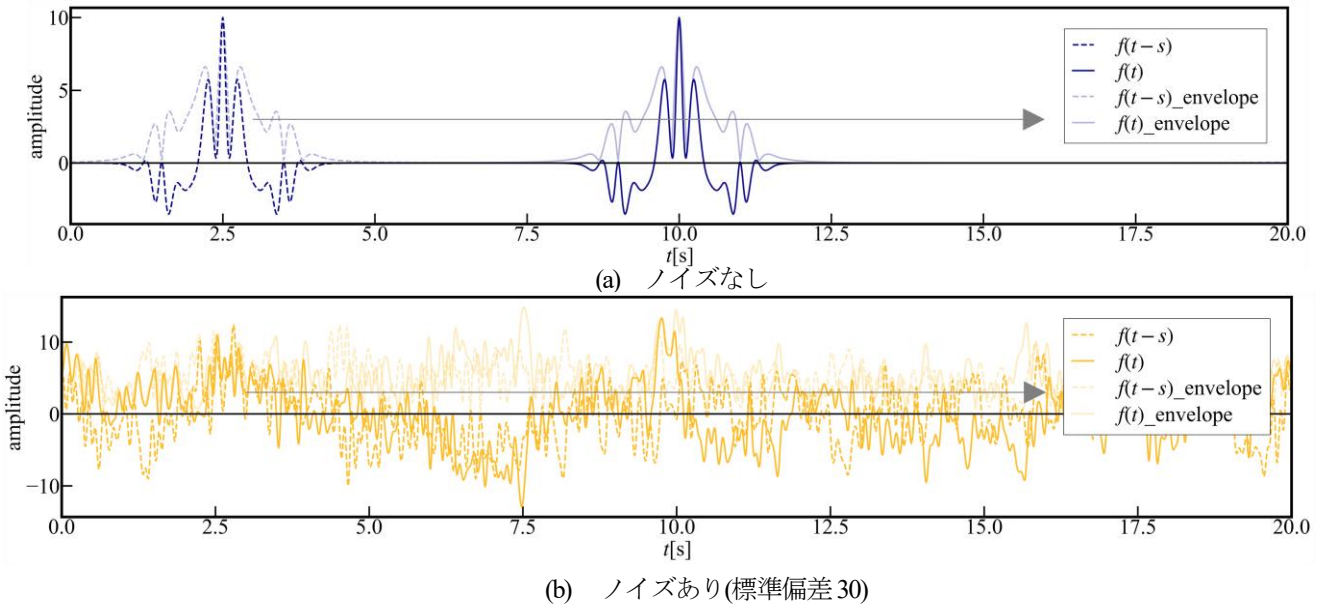


図-16 重ね合わせたリックカー波(パターン2)の時間シフト

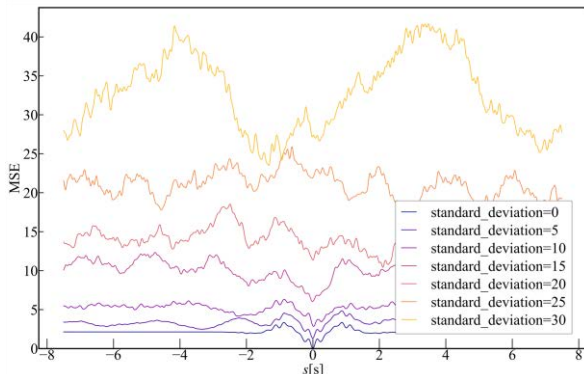


図-17 図-16 の  $f(t)$  と  $f(t-s)$  の類似度(MSE)と時間ずれ  $s$  の関係

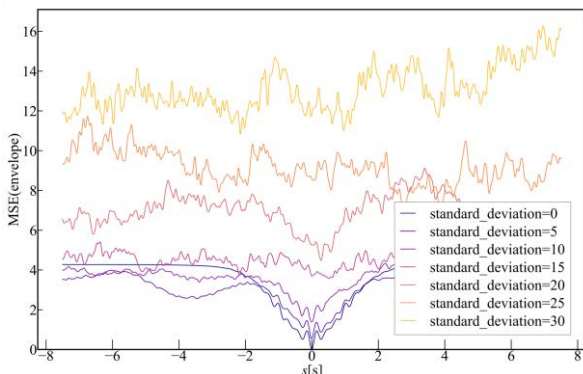


図-18 図-16 の  $f(t)$  と  $f(t-s)$  の類似度(包絡線のMSE)と時間ずれ  $s$  の関係

線の MSE よりも類似度が最大となる時刻ずれ  $\tau$  が 0[s] に近い傾向があるが、ノイズの標準偏差が 25 を越えると包絡線の MSE の方がばらつきは小さくなる傾向にあることが分かる。一方で、 $W_2^2$  はノイズの標準偏差が 5 以上でばらつきが大きい傾向にあることが分かる。

### (3) 重ね合わせたリックカー波の場合(パターン2)

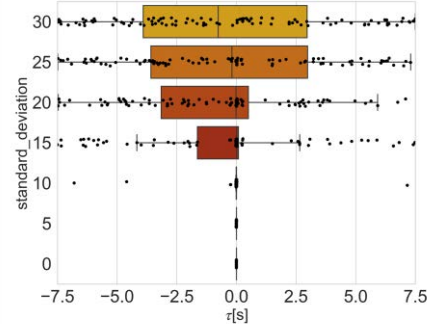


図-20 MSE が最小(類似度が最大)となる時間ずれ  $\tau$  とノイズの標準偏差の関係

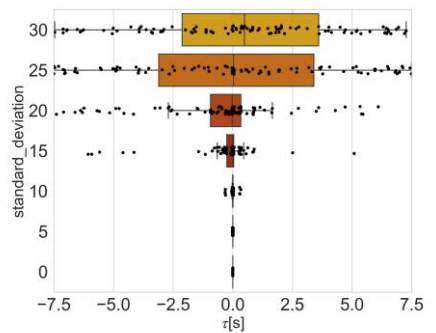


図-21 包絡線の MSE が最小(類似度が最大)となる時間ずれ  $\tau$  とノイズの標準偏差の関係

次に高周波数成分と低周波数成分をどちらも含む波形について考えるために、図-16 のような (2) で使用した 2~4[Hz] の波形に (1) で使用した 0.5[Hz] の波形を重ね合わせた波形を考え、(1)、(2) と同様に時間ずれ  $s$  に対して各類似度指標がどのような値を示すか考察する。ノイズなしの波形(図-16(a)) および平均 0 のピンクノイズを加えた波形(図-16(b)) について検討した。ただし、ノ

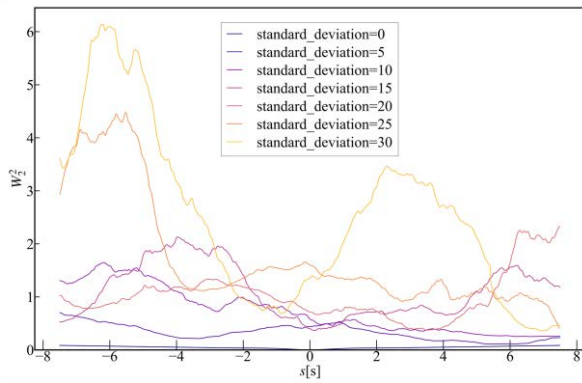


図-19 図-16 の  $f(t)$  と  $f(t-s)$  の類似度 ( $W_2^2$ ) と時間ずれ  $s$  の関係

イズには 0.1[Hz]~10[Hz]の周波数成分のみを持つようにバンドパスフィルタをかけている。また、それぞれの波形の包絡線もともに図-16に描画している。

2つの波形が、時間ずれ  $s$  の大きさによってどれだけ類似度が変動するかを考察する。ノイズなし、およびノイズの標準偏差が 5, 10, 15, 20, 25, 30 の場合について、類似度評価の指標として MSE, 包絡線の MSE,  $W_2^2$  を用いた結果を、それぞれ図-17, 18, 19 に示す。

波形に付加するノイズをランダムに変えながら各ケースについて 100 回繰り返し、類似度指標が最大となる時間ずれ  $\tau$  が時間ずれ 0[s] とどの程度乖離があるかを統計的に考察する。図-20, 21, 22 にその結果を示している。図-20, 21 から、ノイズが小さいときには MSE の方が包絡線の MSE よりも類似度が最大となる時刻ずれ  $\tau$  が 0[s] に近い傾向があるが、ノイズの標準偏差が 15 を越えると包絡線の MSE の方がばらつきは小さくなる傾向にあることが分かる。一方、 $W_2^2$  は標準偏差が 5 以上でばらつきが大きい傾向にあることがわかる。

#### (4) 実際の地震波形の場合

最後に、実際の地震波形について、(1), (2), (3) と同様に時間ずれ  $s$  に対して各類似度指標がどのような値を示すかを考察する。この地震波形は大阪市福島区内にある関西地震観測研究協議会の地震観測点福島(FKS)<sup>10</sup> で観測された 2018 年 6 月 18 日の大阪府北部の地震の EW 成分の速度波形である。ノイズなしの波形 (図 23(a)) および平均 0 のピンクノイズを加えた波形 (図 23(b)) について検討した。ただし、ノイズには 0.1[Hz]~10[Hz]の周波数成分のみを持つようにバンドパスフィルタをかけている。また、それぞれの波形の包絡線もともに図-23に描画している。

2つの波形が、時間ずれ  $s$  の大きさによってどれだけ類似度が変動するかを考察する。ノイズなし、およびノイズの標準偏差が 40, 80, 120, 160, 200 の場合について、

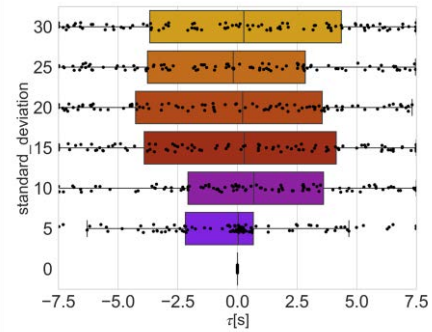


図-22  $W_2^2$  が最小(類似度が最大)となる時間ずれ  $\tau$  とノイズの標準偏差の関係

類似度評価の指標として MSE, 包絡線の MSE,  $W_2^2$  を用いた結果を、それぞれ図-24, 25, 26 に示す。

$W_2^2$  は、ノイズが小さいときには明瞭に凸形状となり、ノイズが大きいかにも全体的な傾向としては凸形状となっていることが分かる。この点が(1), (2), (3)のケースと異なる点である。

波形に付加するノイズをランダムに変えながら各ケースについて 50 回繰り返し、各類似度指標が最大となる時間ずれ  $\tau$  が時間ずれ 0[s] とどの程度乖離があるかを統計的に考察する。図-27, 28, 29 にその結果を示している。図-27, 28 から、ノイズが小さいときは MSE も包絡線の MSE もともに類似度が最大となる時間ずれ  $\tau$  はほぼ 0[s] と一致するが、ノイズが大きくなるとどちらの指標もばらつくことが分かる。ただし、包絡線の MSE の方がばらつきは小さい傾向にある。一方で  $W_2^2$  は、ノイズが小さいときには MSE や包絡線の MSE に比べてややばらつきはあるが、ノイズが大きいかには MSE や包絡線の MSE に比べてばらつきは小さい傾向にあることが分かる。

#### (5) 解析結果に対する考察

(1)-(4)の解析結果を踏まえると、大きく 3 点のことが分かった。1 点目は、波形の周波数成分に関わらず、ノイズが小さいときには、包絡線の MSE よりも MSE の方が類似度が最大となる時間ずれ  $\tau$  が 0[s] に近い傾向となり、ノイズが大きくなると、MSE より包絡線の MSE の方が類似度が最大となる時間ずれ  $\tau$  がばらつきにくい傾向にあることが分かった。つまり、波形の類似度指標として MSE と包絡線の MSE のどちらを用いるべきかの判断基準として、一般的には周波数成分の大小が用いられるが、今回の解析結果によると、比較する地震波形にどれだけ大きさのノイズが加わっているかを基準にするのが適切であるということが分かった。しかし、実際に加わっているノイズの大きさを正確に調べ

ることは困難であり、MSE と包絡線の MSE のどちらを用いるべきかの適切な判断は難しいといえる。一方、 $W_2^2$ については、MSEや包絡線のMSEに比べて、ノイズが大きくなると類似度が最大となる時間ずれが 0[s]から

大きく乖離する傾向にあるということが分かった。

2点目は、(4)の実際の地震波形のケースのように考える時間ウィンドウが長いとき、 $W_2^2$ は MSE や包絡線の MSE に比べて、ノイズが大きくなっても類似度が最大

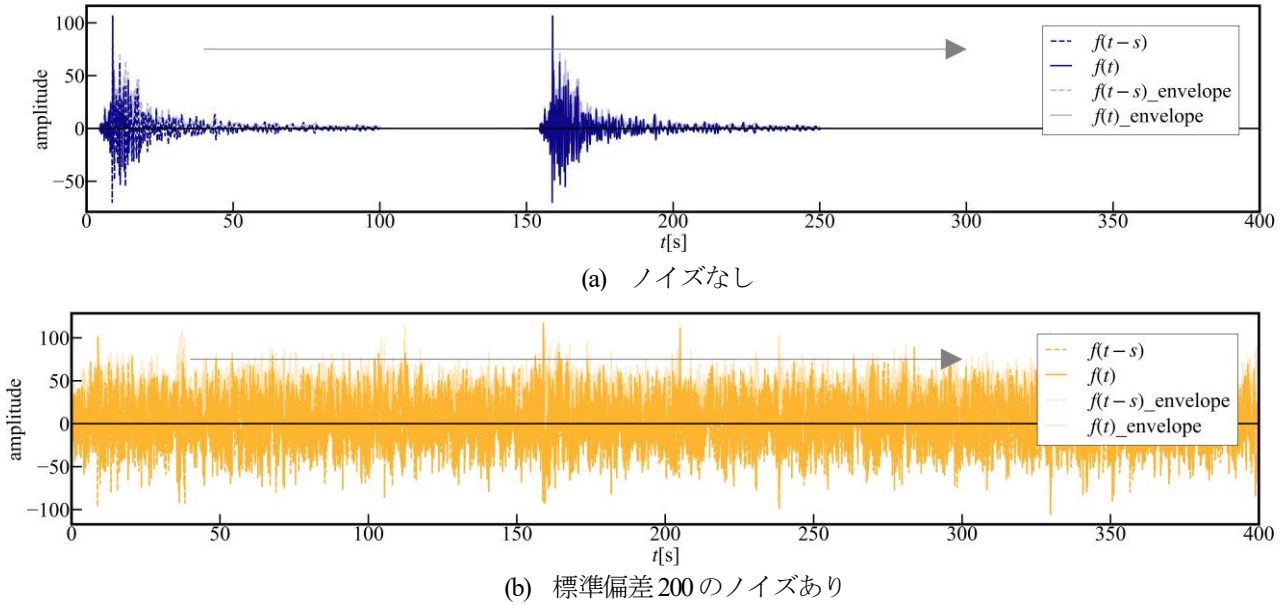


図-23 実際の地震波形の時間シフト

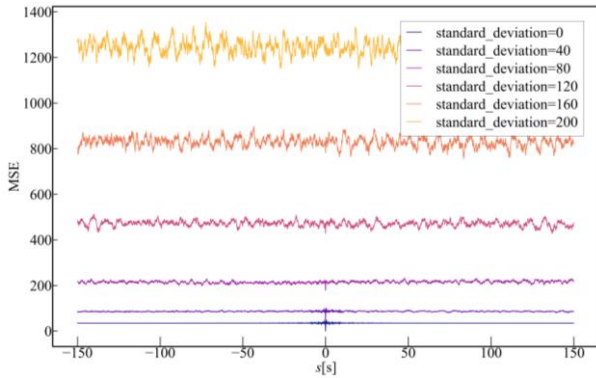


図-24 図-23の $f(t)$ と $f(t-s)$ の類似度(MSE)と時間ずれ $s$ の関係

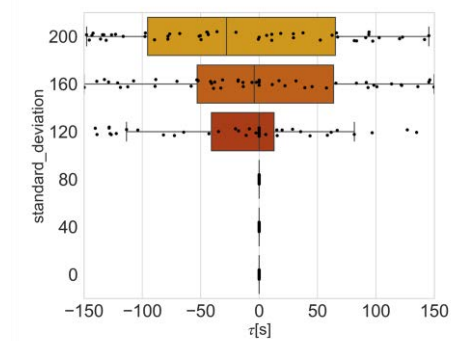


図-27 MSE が最小(類似度が最大)となる時間ずれ $\tau$ とノイズの標準偏差の関係

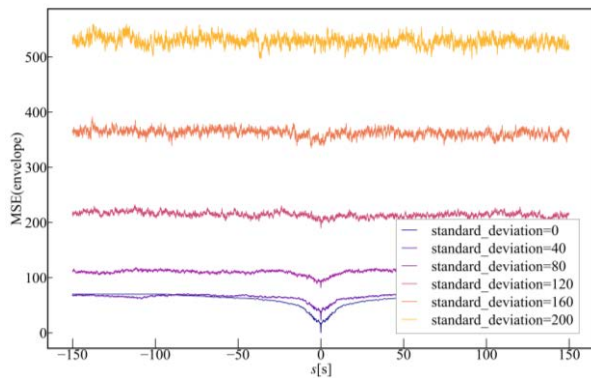


図-25 図-23の $f(t)$ と $f(t-s)$ の類似度(包絡線のMSE)と時間ずれ $s$ の関係

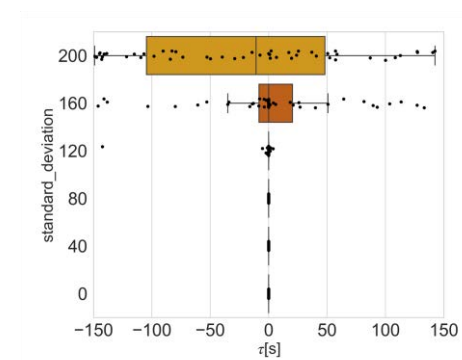


図-28 包絡線のMSEが最小(類似度が最大)となる時間ずれ $\tau$ とノイズの標準偏差の関係

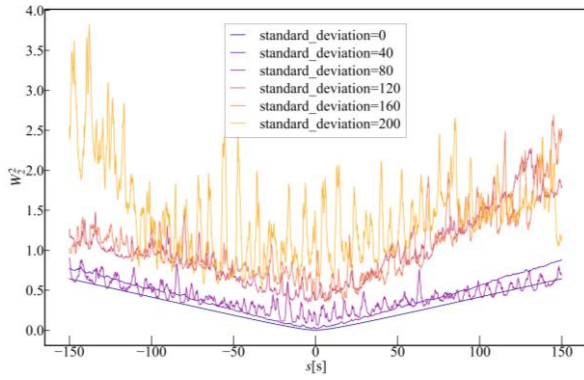


図-26 図-23の $f(t)$ と $f(t-s)$ の類似度( $W_2^2$ )と時間ずれ $s$ の関係

となる時間ずれ $t$ がばらつきにくいという傾向がある。

3点目は、波形の周波数成分の大小に関わらず、ノイズが小さいときには、時間ずれ $s$ に対してMSEは細かく振動するような挙動を示すが、 $W_2^2$ は比較的滑らかに下に凸の形状を描くことが分かった。これは時間ずれ $s$ に対して示す挙動として、MSEに比べ $W_2^2$ の方が極小値を発生させにくいということであり、FWIなどのインバージョンで有用な理由を示している。しかし、ノイズが大きくなると $W_2^2$ も細かく振動するようになることも分かった。

#### 4. センブランス解析へのワッサースタイン計量の適用

本章では、地震波形の類似度評価が実際に必要となるセンブランス解析に対して、ワッサースタイン計量を適用することを考える。センブランス解析の概要、対象とした地震計アレイおよび地震の説明をした後、3章での解析結果およびその考察を踏まえ、センブランス解析におけるワッサースタイン計量を用いた場合の手法と平均二乗誤差を用いる旧来の手法との比較を行う。

##### (1) センブランス解析

センブランス解析とは、地震計アレイにおいて観測された地震波形に対して波群の到来方向と伝播速度の組をさまざまな値で仮定し、地震波形の類似度を評価することで最適な波群の到来方向と伝播速度の組を発見する解析手法をいう。具体的には、地震計アレイにおいて観測されるそれぞれの地震波形についてある基準となる観測点を決め、地震計アレイから見て波群の到来方向と伝播速度の2つのパラメータを仮定し、パラメータから計算される基準点からのそれぞれの観測点への走時差を求める。次にその走時差だけ地震波形をずらしたのちすべての観測波形について類似度を評価する。このとき正しく

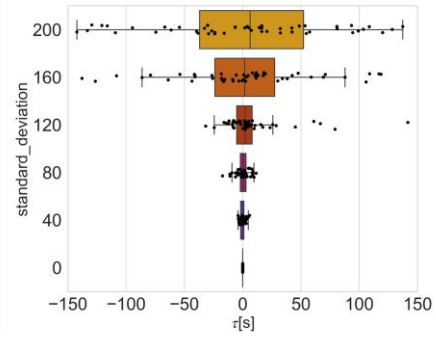


図-29  $W_2^2$ が最小(類似度が最大)となる時間ずれ $t$ とノイズの標準偏差の関係

走時差が求まっているときに類似度が最大となり、波群の真の到来方向と伝播速度を推定できる。

一般的に、センブランス解析では次のように定義されるセンブランス係数 $S_e$ が用いられる<sup>3)</sup>。

$$S_e = \frac{\sum_{k=1}^N \left\{ \sum_{i=1}^M f_i \left( t_k + \frac{x_i}{c} \right) \right\}^2}{M \sum_{k=1}^N \sum_{i=1}^M \left\{ f_i \left( t_k + \frac{x_i}{c} \right) \right\}^2} \quad (18)$$

ここで、地震計の総数を $M$ 、時間ウィンドウ内の総データ数を $N$ 、地震計の番号を $i$ 、データの番号を $k$ 、基準となる点 $i_0$ と地震計 $i$ の距離を伝播方向ベクトルに投影したときの長さを $x_i$ 、地震波の伝播速度を $c$ 、地震計 $i$ の時間 $t_k$ におけるデータを $f_i(t_k)$ としている。このセンブランス係数 $S_e$ の取りうる値の範囲は $0 \leq S_e \leq 1$ であり、類似度が高いと1に近づく。ここで、式(18)は帰納的に次式のように変形できる。

$$S_e = \frac{\sum_{k=1}^N \left\{ \sum_{i=1}^M f_i \left( t_k + \frac{x_i}{c} \right) \right\}^2}{\sum_{k=1}^N \left\{ \sum_{i=1}^M f_i \left( t_k + \frac{x_i}{c} \right) \right\}^2 + \sum_{k=1}^N \sum_{j=2}^M \sum_{i=1}^{j-1} \left\{ f_i \left( t_k + \frac{x_i}{c} \right) - f_j \left( t_k + \frac{x_j}{c} \right) \right\}^2} \quad (19)$$

この式の分母に着目すると、2つの波形の時系列データの差の二乗を計算する項が出現しており、 $S_e$ はMSEと等価な性質を持つことが分かる。

例として、図-30のような簡単な地震計アレイを考える。地震計が全部で $M$ 個あるとし、そのうちの地震計 $i$ と地震計 $j$ について着目する。ある基準点 $i_0$ を定め、 $i_0$ と各地震計との距離が分かっているとす。このとき、波群の到来方向を $\phi$ とすれば、 $x_i$ や $x_j$ は $\phi$ の関数となる。つまり、センブランス解析において波群の到来方向 $\phi$ と伝播速度 $c$ の2つのパラメータを変化させることによって、走時差を任意に表現することができる。地震計アレイでは一般的に非常に高密度に地震計を配置しているため、どの地震計においても $\phi$ と $c$ の値は共通であると仮定できる。地震計 $i$ 、 $j$ だけでなく、 $M$ 個の地震計すべての組み合わせについて考えれば、式(18)からセンブラン

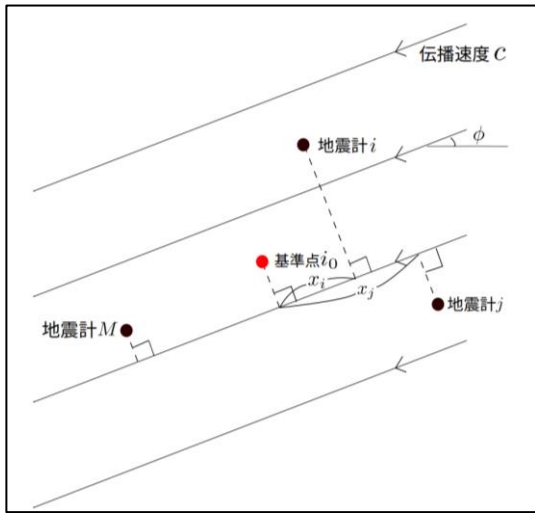


図-30 地震計アレイの概略図

ス係数 $S_e$ を計算することができる。

ここで、センブランス係数 $S_e$ の代わりに、ワッサースタイン計量を波形の類似度指標として用いることを考える。まず、ある時間ウィンドウ内の総データ数を $N$ とし、地震計番号が $i, j$ の2つの地震計で観測された地震波形の類似度がワッサースタイン計量を用いて $W_{2,N}^2(f_i(t), f_j(t))$ と表されるとする。ここで波群の到来方向 $\phi$ と伝播速度 $c$ による時間ずれを考慮すると、

$W_{2,N}^2(f_i(t + \frac{x_i}{c}), f_j(t + \frac{x_j}{c}))$ と表される。地震計の総数が $M$ であるとき、すべての地震計の組み合わせでワッサースタイン計量を計算し、その平均を $W$ とすれば、

$$W = \frac{\sum_{p=1}^M \sum_{i=1}^{M-p} W_{2,N}^2(f_{i+p}(t + \frac{x_{i+p}}{c}), f_i(t + \frac{x_i}{c}))}{M C_2} \quad (20)$$

となる。この $W$ をセンブランス解析における新しい類似度指標として用いる。

## (2) 解析対象の地震計アレイ

本論文で使用した地震計アレイは、大阪市福島区内に設置された地震計アレイ<sup>10)</sup>である。合計4台の地震計で構成され、関西地震観測研究協議会の地震観測点福島(FKS)<sup>10)</sup>の周辺に3台の地震計が設置されている。図-31にFKS周辺の地震計の設置位置を示す。地震計のサンプリング周波数は100[Hz]となっており、GPSによって時刻が同期されている。

## (3) 解析対象の地震

今回センブランス解析の対象としたのは、2018年6月18日7時58分に発生した大阪府北部の地震とする。この地震は当該地震計アレイが観測を開始してから最大の平均最大加速度を記録した地震である。解析には速度波形のEW成分を用いた。地震計アレイの各観測点で観測された速度波形のEW成分を図-32に示している。なお

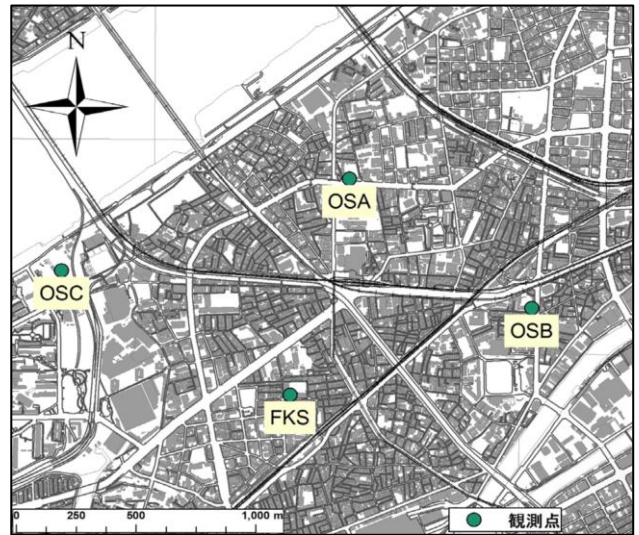


図-31 地震計の設置位置

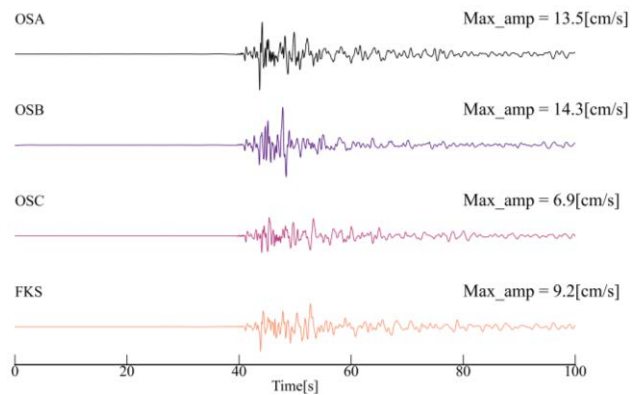


図-32 大阪府北部の地震における速度波形(EW成分)

時間軸の0[s]は7時58分00秒を表している。

## (4) 解析結果の描画方法

前述のように、センブランス解析は伝播速度 $c$ と波群の到来方向 $\phi$ のパラメータを変化させて行うものである。解析結果はスローネスの二次元平面上にコンターマップで描画する。スローネスとは、波数を角速度で除した次元が[s/m]の物理量である。波群の波数ベクトルの $x$ 成分、 $y$ 成分をそれぞれ $k_x, k_y$ とすると、波数 $k$ 、角速度 $\omega$ 、波速 $c$ の間に $c = \frac{\omega}{k}$ の関係があることから、 $k =$

$\sqrt{k_x^2 + k_y^2}$ より、

$$\frac{1}{c} = \sqrt{\left(\frac{k_x}{\omega}\right)^2 + \left(\frac{k_y}{\omega}\right)^2} \quad (21)$$

の関係がある。ここで、スローネスの $x$ 成分は $\frac{k_x}{\omega}$ 、 $y$ 成分は $\frac{k_y}{\omega}$ である。式(21)から、 $\frac{k_x}{\omega}, \frac{k_y}{\omega}$ 平面上に半径が $\frac{1}{c}$ の円を描けることが分かる。また、波群の到来方向 $\phi$ について

では、

$$\cos \phi = \frac{\frac{k_x}{\omega}}{\sqrt{\left(\frac{k_x}{\omega}\right)^2 + \left(\frac{k_y}{\omega}\right)^2}} \quad (22)$$

$$\sin \phi = \frac{\frac{k_y}{\omega}}{\sqrt{\left(\frac{k_x}{\omega}\right)^2 + \left(\frac{k_y}{\omega}\right)^2}} \quad (23)$$

によって定めることができる。つまり、スローネスに関する直交座標系 $\left(\frac{k_x}{\omega}, \frac{k_y}{\omega}\right)$ を与えると、伝播速度 $c$ の逆数と到来方向 $\phi$ に関する極座標系 $\left(\frac{1}{c}, \phi\right)$ に変換できることが分かる。

続いて、解析結果の見方を図-33を例として説明する。図-33の白丸で囲まれた箇所が波形の類似度の最も大きい箇所であるとする。伝播速度 $c$ の値を示した同心円の中心を基準点 $i_0$ の位置とみなし、 $i_0$ から東向きベクトルを基準として到来方向 $\phi$ を与える。白丸で囲まれた箇所が $\left(\frac{1}{c}, \phi\right)$ であるとき、その時間ウィンドウでは波群が到来方向 $\phi$ で、かつ伝播速度 $c$ で地震計アレイに到達したと考えることができる。

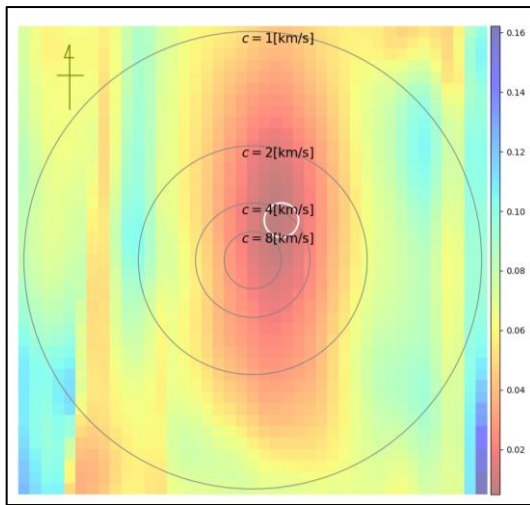


図-33 解析結果の見方

### (5) 解析結果

解析の対象とする時間は図-32の40[s]-50[s]とする。50[s]以降も地震波が到達していることは明らかであるが、周囲の盆地境界で表面波が励起され、盆地内に残留するという大阪平野特有の特徴（実際に、2018年6月18日大阪府北部の地震において、関口らは強震動シミュレーションを用いて、盆地内のあちこちで後続波が生成され、様々な方向に伝播して大阪平野内に複雑な波動場が

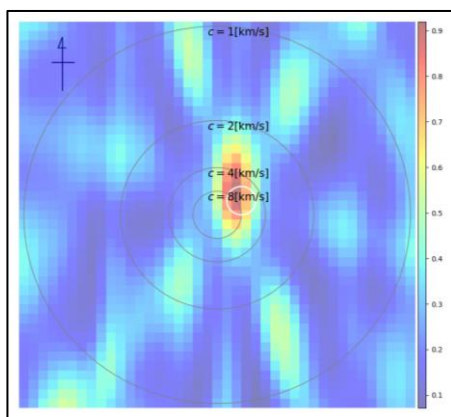
生成される様子を再現しており、その後続波の再現波形は大阪平野内の観測波形とよく合致している<sup>11)</sup>を考慮すると、50[s]以降は様々な方向から波群が到来し、センブランス解析が困難である可能性がある。この地震の震源は当該地震計アレイから見ておおよそ北東方向である大阪府高槻市であり、実体波は北東方向から到来すると予想できるため、実体波が卓越していると考えられる40[s]-50[s]に着目することとした。センブランス解析の時間ウィンドウ幅は2.0[s]としており、1.0[s]ずつ注目する時間ウィンドウをずらしながらセンブランス解析を行う。

図-34から図-42が40[s]-50[s]におけるセンブランス解析の結果である。類似度が大きくなるほど赤色に近づき、類似度が小さくなるほど青色に近づくように図示している。ただし、センブランス係数 $S_e$ はその値が大きいほど類似度が大きく、式(21)により定義される $W$ はその値が小さいほど類似度が大きいため、カラーバーの色と数値の大小とが逆であることに注意されたい。

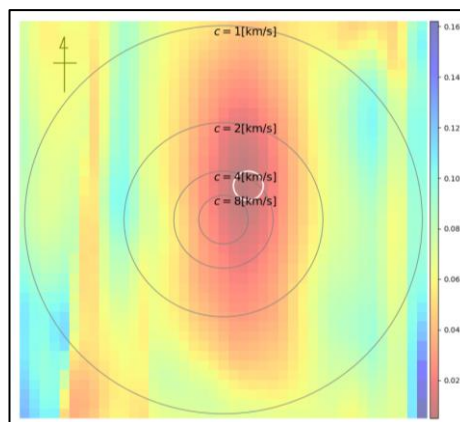
### (6) 解析結果の妥当性の検討と考察

波群の到来方向は、全体の傾向として $S_e$ 、 $W$ ともにおおよそ北東方向に類似度の大きな箇所が求められており、震源方向が当該地震計アレイから見て北東方向であることから妥当な結果といえる。また、伝播速度は40[s]-45[s]の間で $S_e$ 、 $W$ ともに4-8[km/s]を示している。この時間ウィンドウではP波・S波が卓越していると考えられ、入射各を考慮した見かけの速度と考えれば、適切なオーダーの値と考えられる。

続いて、類似度指標として $S_e$ と $W$ をそれぞれ用いたときの違いを考察する。全体的な傾向として $S_e$ は極値が複数出現していることが分かる。解が正しく求まっていると考えられる時間ウィンドウもあるが、図-39~図-42では極値の影響により合理的でない解となっている。一方、 $W$ は極値が少なく大域的な解の判別が可能で、合理的な解を示していることが分かる。これは3章でも述べたような、ノイズが小さいときに、時間ずれ $s$ に対してMSEは細かく振動するような挙動を示すが、 $W_2^2$ は比較的滑らかに下に凸の形状を描くという性質に対応する。この地震計アレイは非常に高密度な位置関係であり、各地震計で観測される地震波は表層地盤や減衰の影響の違いが小さいと考えられるため、各地震計で観測された地震波形同士のノイズは小さいと考えられる。そのため、このような結果が得られたと考察できる。一方、地震計アレイの密度が疎である場合や、各地震計の設置位置で地盤特性がそれぞれ大きく異なる場合はノイズが大きくなるため、 $W$ を用いた場合でも妥当な解を得ることは困難であると予想できる。

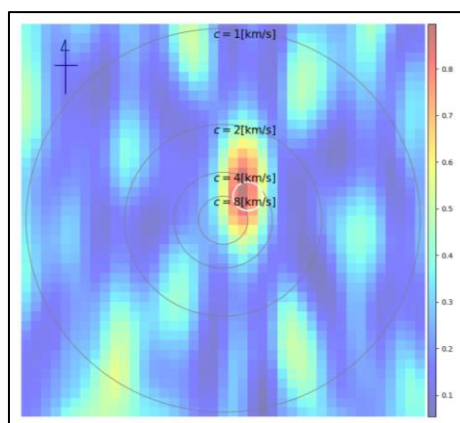


(a)  $S_e$

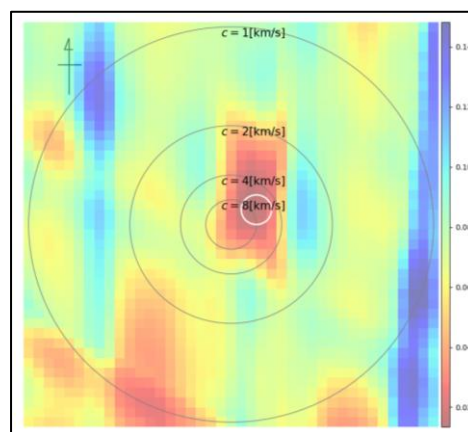


(b)  $W$

図-34 時間ウィンドウ 40[s]-42[s]の解析結果

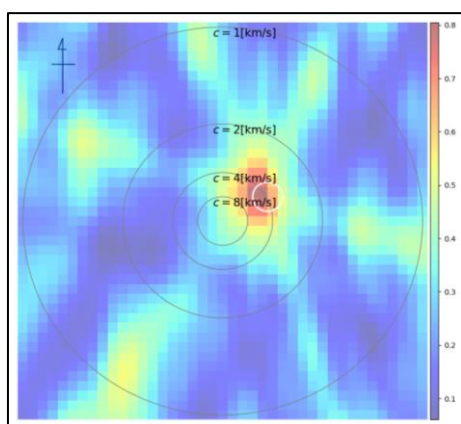


(a)  $S_e$

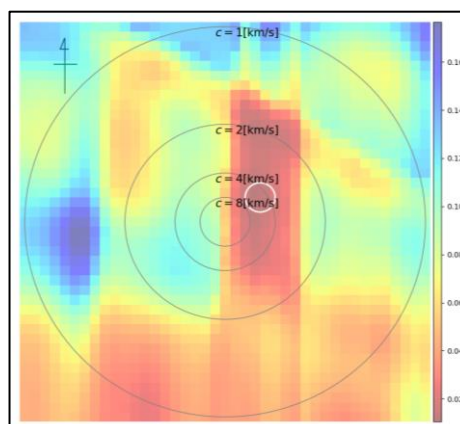


(b)  $W$

図-35 時間ウィンドウ 41[s]-43[s]の解析結果

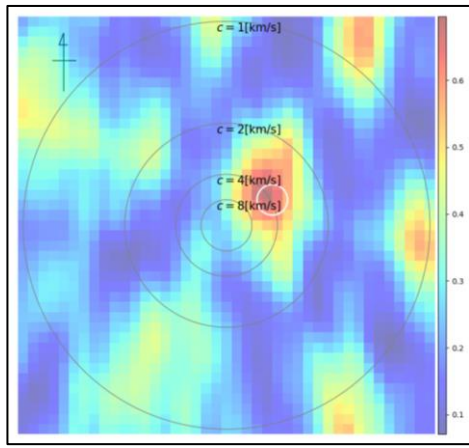


(a)  $S_e$

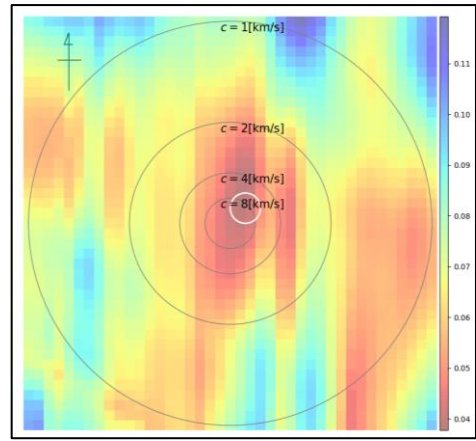


(b)  $W$

図-36 時間ウィンドウ 42[s]-44[s]の解析結果

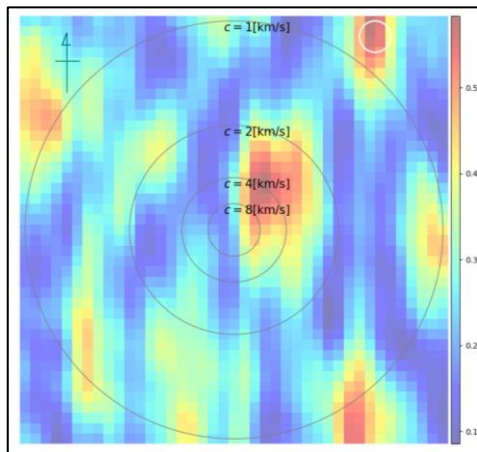


(a)  $S_e$

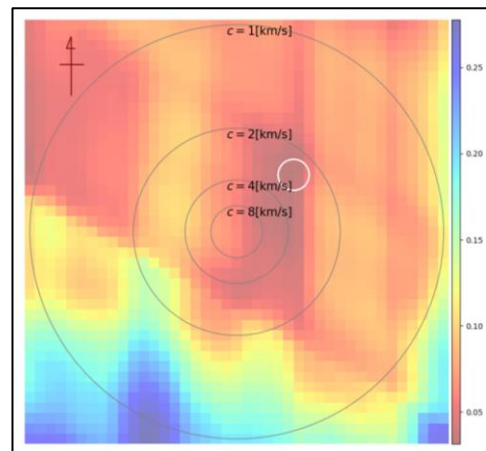


(b)  $W$

図-37 時間ウィンドウ 43[s]-45[s]の解析結果

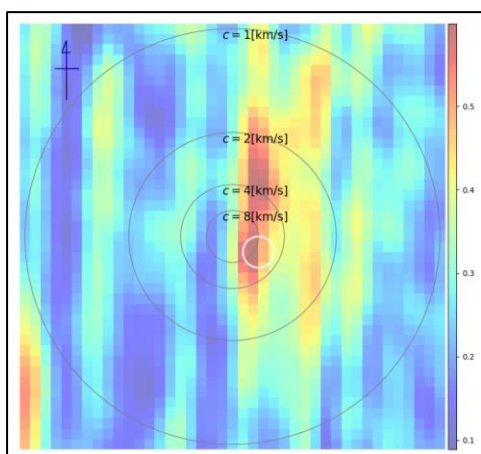


(a)  $S_e$

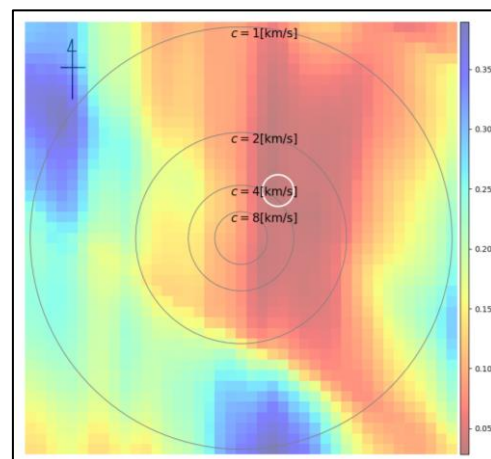


(b)  $W$

図-38 時間ウィンドウ 44[s]-46[s]の解析結果

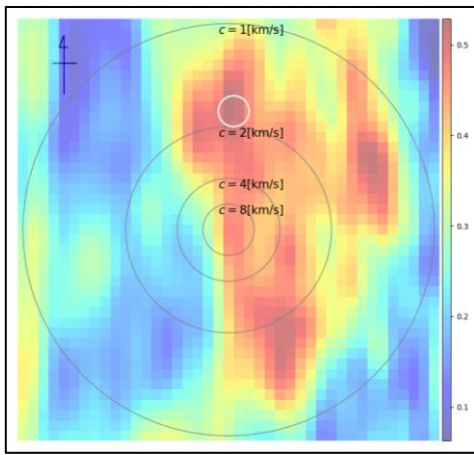


(a)  $S_e$

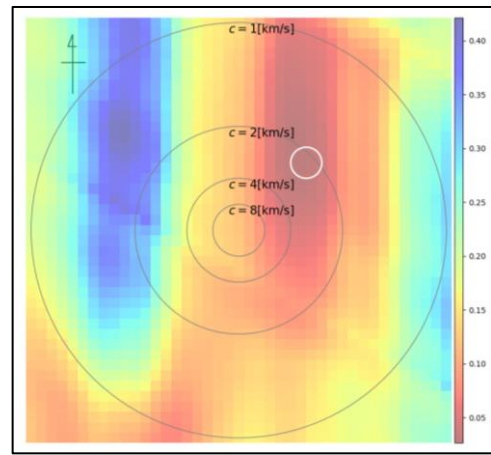


(b)  $W$

図-39 時間ウィンドウ 45[s]-47[s]の解析結果

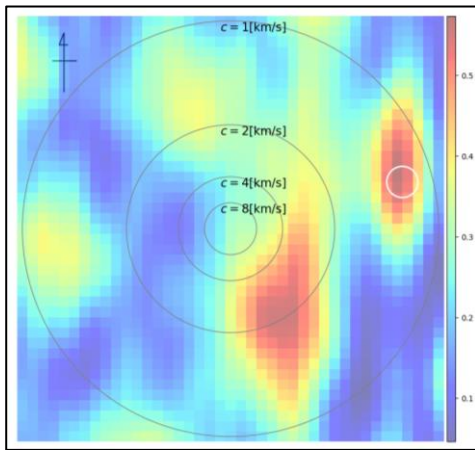


(a)  $S_e$

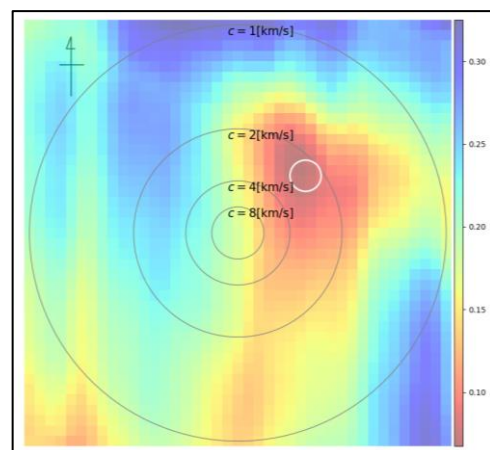


(b)  $W$

図-40 時間ウィンドウ 46[s]-48[s]の解析結果

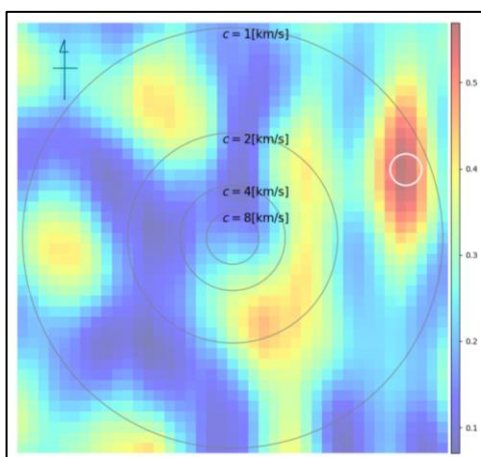


(a)  $S_e$

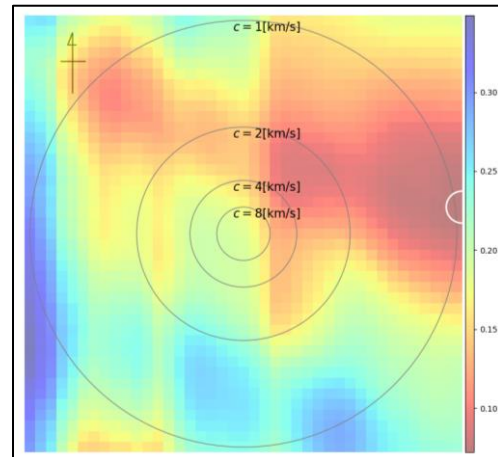


(b)  $W$

図-41 時間ウィンドウ 47[s]-49[s]の解析結果



(a)  $S_e$



(b)  $W$

図-42 時間ウィンドウ 48[s]-50[s]の解析結果

## 5. 結論

本研究ではまず、地震波形の類似度評価の指標として、波形やその包絡線の平均二乗誤差、ワッサースタイン計量を用いた場合でそれぞれどのような性質があるのか調べるため、様々なパターンの地震波形で具体的に解析を行った。ノイズが小さいときには波形の平均二乗誤差を、ノイズが大きいときには包絡線の平均二乗誤差を用いることで、適切に時間ずれを特定できる可能性があることがわかった。一方、ワッサースタイン計量は考える時間ウィンドウが長く、かつノイズが大きいときには波形や包絡線の平均二乗誤差に比べて適切に時間ずれを特定できる可能性があることがわかった。また、平均二乗誤差は時間ずれに対して、ノイズが小さくても細かく振動するような挙動を示すが、ワッサースタイン計量は比較的滑らかに下に凸の形状を描くことがわかった。

続いて、大阪市福島区の地震計アレイにおいて観測された2018年大阪府北部の地震の波形を用いてセンブランス解析を行い、ワッサースタイン計量を用いた場合の手法と平均二乗誤差と等価なセンブランス係数を用いる旧来の手法との比較を行った。センブランス係数を用いた場合では極値の影響により合理的でない解を示す時間ウィンドウがあった一方、ワッサースタイン計量を用いた場合では極値が少なく大域的な解の判別が可能で、合理的な解を示すことが明らかとなった。

本研究では波形の類似度評価の各指標の調べるために周波数成分の大小とノイズの大小に着目して複数のパターンで解析を行ったが、さらにパターンを増やして解析を行うほか、波形を構成する他の要素に着目して解析を行う必要がある。また、ワッサースタイン計量を地震波形に適用する際の正規化手法については、本研究で用いた手法以外にも提案されている手法があり、その場合で

も同様の解析を行って特徴を調べる必要がある。

## REFERENCES

- 1) Virieux, Jean, et al. "An introduction to full waveform inversion." *Encyclopedia of exploration geophysics*. Society of Exploration Geophysicists, 2017. R1-1.
- 2) 八木勇治. "震源過程解析手法の開発." 地震 第2輯 61.Supplement (2009): 297-307.
- 3) N. S. Neidell and M. Turhan Taner. "Semblance and other coherency measures for multichannel data." *Geophysics* 36.3 (1971): 482-497.
- 4) K.Asano and T.Iwata. "Source model for strong ground motion generation in the frequency range 0.1–10 Hz during the 2011 Tohoku earthquake." *Earth, planets and space* 64.12 (2012): 1111-1123.
- 5) Bjorn Engquist and Brittany D Froese. Application of the Wasserstein metric to seismic signals. *arXiv preprint arXiv:1311.4581*, 2013.
- 6) Cedric Villani. *Optimal transport: old and new*, Vol. 338. Springer, 2009.
- 7) Cedric Villani. *Topics in optimal transportation*, Vol. 58. American Mathematical Soc., 2021.
- 8) Bjorn Engquist and Yunan Yang. Optimal transport based seismic inversion: Beyond cycle skipping. *arXiv preprint arXiv:2002.00031*, 2020.
- 9) Bjorn Engquist, Brittany D Froese, and Yunan Yang. Optimal transport for seismic full waveform inversion. *arXiv preprint arXiv:1602.01540*, 2016.
- 10) 関西地震観測研究協議会. 観測点情報, 2022. <http://www.ceorka.org/>. (2022年1月28日閲覧)
- 11) T.Iwata H.Sekiguchi, K.Asano. Strong Ground Motion Simulation in Osaka Basin, Japan, for the 2018 Northern Osaka Prefecture Earthquake. In *Proceeding of 17th World Conference on Earthquake Engineering*, No. C001162, Sendai, Japan, September 13th to 18th 2020.

(Received ??, ????)

(Accepted ??, ????)

## NATURE OF WASSERSTEIN METRIC ON WAVEFORM SIMILARITY EVALUATION AND APPLICATION TO SEMBLANCE ANALYSIS

Tatsuki NARA, Hiroyuki GOTO

Similarity evaluation of seismic waveforms is one of the most important aspects in earthquake engineering and seismology. Recently, Wasserstein metric has been proposed as a new index for similarity evaluation of seismic waveforms, and although Wasserstein metric has been shown to be effective in a certain field, there has been no researches on its application as an index for waveform similarity evaluation in earthquake engineering. In this study, the characteristics of the mean square error (MSE), which has been conventionally used as a measure of waveform similarity, and the Wasserstein metric, which is a new measure, are investigated in detail, and then Wasserstein metric is applied to semblance analysis and compared with the conventional method using MSE.