

平成 30 年度 修士論文

ニューラルネットワークを用いた  
走時トモグラフィの改良に関する研究

首都大学東京大学院

都市環境科学研究科

都市基盤環境学域

17885403 金本龍之介

指導教官 小田義也 准教授

# 目次

## 第1章 緒論

- 1.1 背景・目的
- 1.2 本論文の構成

## 第2章 従来法と課題点

- 2.1 屈折法地震探査
- 2.2 走時トモグラフィ
- 2.3 非線形最小二乗法による解析概要と課題点
  - 2.3.1 非線形最小二乗法による解析概要
  - 2.3.2 非線形最小二乗法を用いた解析の課題への取り組み
- 2.4 走時トモグラフィで用いられるニューラルネットワーク
  - 2.4.1 ニューラルネットワーク
  - 2.4.2 階層型ニューラルネットワーク
  - 2.4.3 ニューロン（ユニット）の入力時の処理
- 2.5 階層型ニューラルネットワークの学習
- 2.6 ニューラルネットワークを用いた走時トモグラフィ
- 2.7 ニューラルネットワークを用いた走時トモグラフィの課題

## 第3章 ニューラルネットワークを用いた走時トモグラフィの改良

- 3.1 使用したニューラルネットワーク
- 3.2 本研究にて変更したパラメータ
  - 3.2.1 学習回数
  - 3.2.2 伝達関数
  - 3.2.3 重みの初期値
  - 3.2.4 入力の初期値

## 第4章 数値解析

- 4.1 解析領域
- 4.2 誤差の評価
- 4.3 解析結果
  - 4.3.1 学習回数の比較

#### 4.3.2 重みの初期値を変更したときの学習回数比較

(1)  $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布を用いたとき

(2)  $\sqrt{\frac{2}{n}}$ の標準偏差を持つガウス分布を用いたとき

#### 4.3.3 伝達関数を変更したときの学習回数比較

(1) 一様乱数を用いたとき

(2)  $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布を用いたとき

(3)  $\sqrt{\frac{2}{n}}$ の標準偏差を持つガウス分布を用いたとき

#### 4.3.4 入力の変換を導入したとき

(1) パターンコーディングの導入

#### 4.4 断層構造での結果

#### 4.5 二進数の導入

#### 4.6 考察

## 第5章 結論

### 5.1 まとめ

# 第1章 緒論

## 1.1 背景・目的

屈折波地震探査は土木分野において広く利用されている物理探査手法の一つで、坑道掘削前後の状況調査などに適用されている。屈折法地震探査のデータ解析法としては、これまでにはぎ取り法が使われてきたが、近年ではトモグラフィ解析も一般的となっている。走時トモグラフィの解析は、震源から観測点まで弾性波が伝播する時間（走時）をデータとして、X線 CT スキャンと同様の原理で地下構造を推定するものである。走時トモグラフィにおける最適化手法として、主に、非線形最小二乗法が用いられているが、解析領域を適切なサイズで離散化する必要があることや、正解に十分近い初期モデルを与えないと局所解に陥りやすいという特徴があることから、解析の際にはこれらの適切な設定が重要な課題となる。

この課題を解消するため、解析領域を離散化する必要がなく、初期モデルを与える必要のない方法として、ニューラルネットワークを用いた走時トモグラフィが提案された。しかし、現在のニューラルネットワークを用いた走時トモグラフィでは、断層など不連続な地下構造を評価することが困難であった。

そこで本研究では、ニューラルネットワークの学習回数、伝達関数、重みの初期値、そして、入力パターンに着目し、従来の走時トモグラフィの改良を行い、数値解析によりその効果を評価した。

## 1.2 本論文の構成

本論文は全 6 章で構成されている。以下に本論文の構成とその概略を述べる。

第 1 章では、緒論として研究の背景と目的について論述した。

第 2 章では、本研究で対象としている屈折法地震探査およびニューラルネットワークを用いた走時トモグラフィの解析方法について述べるとともに、ニューラルネットワークを用いた走時トモグラフィの課題について整理した。学習には、誤差情報を出力層から入力層に向かって伝播させて重みを修正していく誤差逆伝播法を用いた。誤差逆伝播法で求められる重みの修正量は理論走時と観測走時をそれぞれ差の二乗を誤差として、その誤差を重みで偏微分したものに学習率を乗算したものである。

第 3 章では、本研究の改良手法について論述した。本研究では、学習回数の検討を行った。そして、ニューロン同士が信号を送る際に用いられる伝達関数をシグモイド関数から ReLU 関数に変更することを試みた。またシグモイド関数、ReLU 関数のそれぞれで重みの初期値の与え方について検討した。従来法では重みの初期値として一様乱数を与えていたが、本研究では、Glorot, et, al.(2010)<sup>1)</sup>、He, et, al.(2015)<sup>2)</sup> らの方法を参照して、前層のノード数を  $n$  として、それぞれ平均を 0、標準偏差を  $\sqrt{1/n}$ 、 $\sqrt{2/n}$  の正規分布に従う乱数を重みの初期値として使用した。入力パターンの変更においては、空間座標値を直接入力する



従来法と、パターンコード化して入力する方法を採用した。

第4章では、3章で述べた改良について、数値実験を行いその効果を評価した。数値実験は、水平方向に24km、深さ方向に4kmの二次元領域とし、速度推定点は深さ方向に1km間隔、水平方向に2km間隔の計65点とした。モデルは成層構造、解析領域の中心付近に断層を想定した速度不連続を設けたもの（以下、断層構造1）、そして、解析領域の端部に断層を想定した速度不連続を設けたもの（以下、断層構造2）の3モデルである。学習回数、伝達関数、重みの初期値、そして、入力パターンについては3章で述べたものを用い、数値実験の結果、速度変化がない部分では、正解値の速度と出力値の速度のぶれはあるが、正解値にはない速度不連続を検出するような誤りはなかった。しかし、成層構造、断層構造ともに、サンプルによっては速度不連続部分において速度が連続的に変化する結果となり、不連続を明確に再現することはできなかった。また、断層構造1よりも断層構造2においてその傾向は顕著であった。これは、解析領域端部の波線密度が解析領域中央部に比べて小さいことが原因であると考察した。

第5章では、本研究で得られた成果をまとめ、今後の課題について論述した。

## 第2章 従来法の手法と課題点

### 2.1 屈折法地震探査

屈折法地震探査は、地表で人工地震を発生させ、これを地表に直線的に配置した受信器で観測する。観測データのうち初動走時、すなわち直接波と地中の速度の異なる地層境界で屈折して地表に伝播した屈折波を解析することにより地下の速度構造を推定する探査手法である。この探査手法で求められる弾性波速度は地盤強度との相関が良いため、岩種と推定された弾性波速度から岩盤分類や地山区分の情報として利用されており、土木構造物設計・施工の際の良質な情報となっている。これらの特徴が評価され、屈折法地震探査は土木分野で広く利用されている。この探査手法で得られたデータの解析にははぎ取り法（萩原の方法）が多く利用されていたが、近年では、トモグラフィ解析も利用されている。

屈折法地震探査の基本原理解説であるスネルの法則（図 2.1）について述べる。この法則は臨界屈折を説明する際に必要である。第 1 層（地表側）から波動が入射する際の入射角を  $\theta$ 、屈折角を  $\theta'$  とするとき、 $\sin \theta / \sin \theta' = V_1 / V_2$  が成り立つ。 $V_1$  は第 1 層での速度で、 $V_2$  は第 2 層での速度を表す。入射角が大きくなると、屈折角も大きくなり、やがて  $90^\circ$  となり地表境界面に沿って波が伝播する。屈折角が  $90^\circ$  となる入射角を臨界角と呼び、 $\sin 90^\circ = 1$  であるので、 $\sin \alpha = V_1 / V_2$  となる。

屈折法地震探査の概要を紹介する。2 層で構成される水平成層構造を考えたとき、地表に近い第 1 層の速度を  $V_1$ 、第 1 層の下部にある第 2 層の速度を  $V_2$  とする。ここで  $V_1 < V_2$  とする（一般的な地層速度は深部の方が速い。仮に  $V_2 < V_1$  である場合は屈折法地震探査を適用できない）。初動走時を震源からの距離ごとにプロットした走時曲線図を図 2.2 に示す。 $T_1$  は第 1 層中を最短距離で伝播する直接波の走時、 $T_2$  は第 1 層と第 2 層の境界面で臨界屈折して境界面を伝播し、再び臨界角屈折をして地表で観測される屈折波の走時である。臨界距離より近傍では直接波が初動走時となり、臨界距離より遠方では屈折波が初動走時となる。図 2.2 中の  $T_1$ 、 $T_2$  の傾きは各層を伝播する弾性波速度の逆数（スローネス）である。また、 $T_2$  の切片から層厚がわかる。今回は最も簡単な構造で説明をしたが、実際に探査を行う際の速度構造は不明であり、水平成層構造である可能性も無いに等しいため  $T_2$  は直線にならない。そのため、はぎ取り法やトモグラフィ解析が行われる。

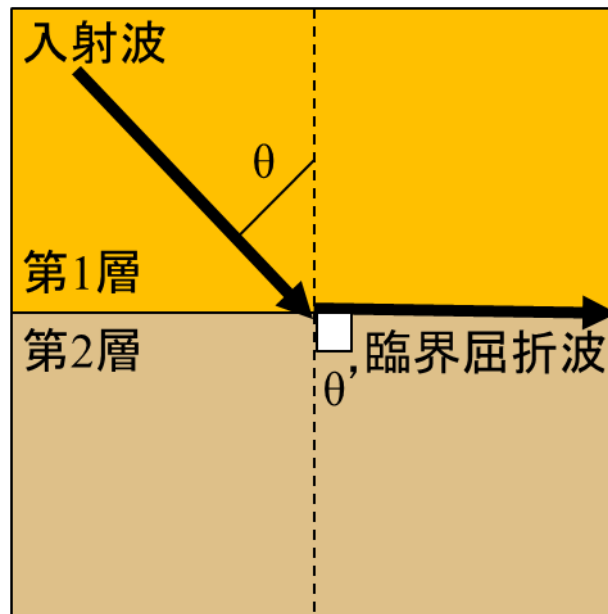


図 2.1 スネルの法則

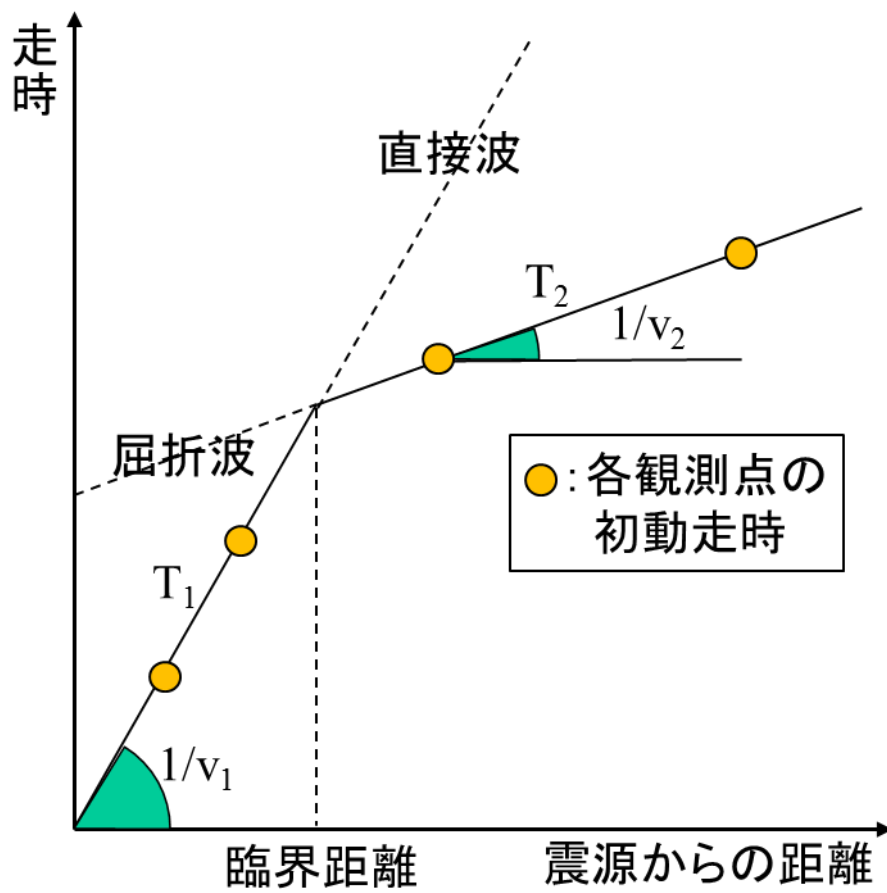


図 2.2 走時曲線図

## 2.2 走時トモグラフィ解析

走時トモグラフィ解析は、X線CTスキャンの対象を人体ではなく、地下に対して行っているものである。図2.3右図で表されるX線CTスキャンとは、360度全方向からX線撮影を行い、被写体の内部の構造を把握するものである。図2.3左図は震源から受信点までの波線経路を図示化したものである。屈折法データを用いた走時トモグラフィ解析は、観測走時と理論走時の誤差を最小にする地下構造モデルを逆解析により求める解析法であり、逆解析には非線形最小二乗法を用いることが多い。しかし、非線形最小二乗法を用いた逆解析には2つの課題がある。

1つ目は、解析領域内を離散化する必要があるという点である。人体に行うX線CTスキャンと異なり、走時トモグラフィは解析領域外の360度全方位に震源と受信点を設置することが困難なため、解析領域内の波線密度に差が発生する。この状況で格子点に分割するサイズは、波線密度の小さい領域に合わせて設定されるため、必然的に格子点間隔が大きくなる。その結果、波線密度が大きな領域でも解析結果の分解能が低くなる。様々な波線密度の領域が存在する中で、最適な分割サイズを決定することは容易ではない。

2つ目は、適切な初期モデルを与えなくてはならない点である。一般的に初期モデルが正解値から大きく外れている場合、解析の結果が局所解に陥ることがある。図2.4（春日部、2016）<sup>3)</sup>は最小二乗法を用いた数値実験の例で、正解モデルとはかけ離れた初期モデルを与えた際に解析結果が正解モデルを表現できなかった様子を表している。初期モデルを正解モデルに近い値で与えることが可能であればこの課題は解消されるが、周辺地盤の地下構造を把握し初期モデルを与える事を考えれば、容易ではない。

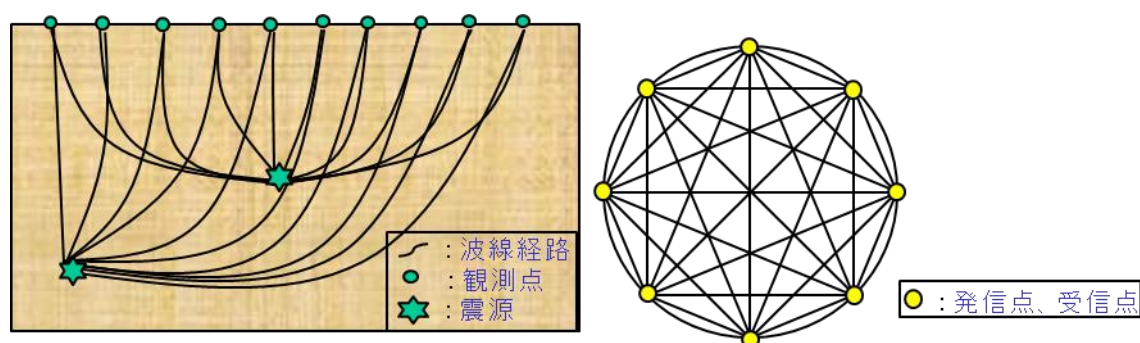


図 2.3 震源から受信点までの波線経路（左）と X 線 CT スキャンの模式図（右）

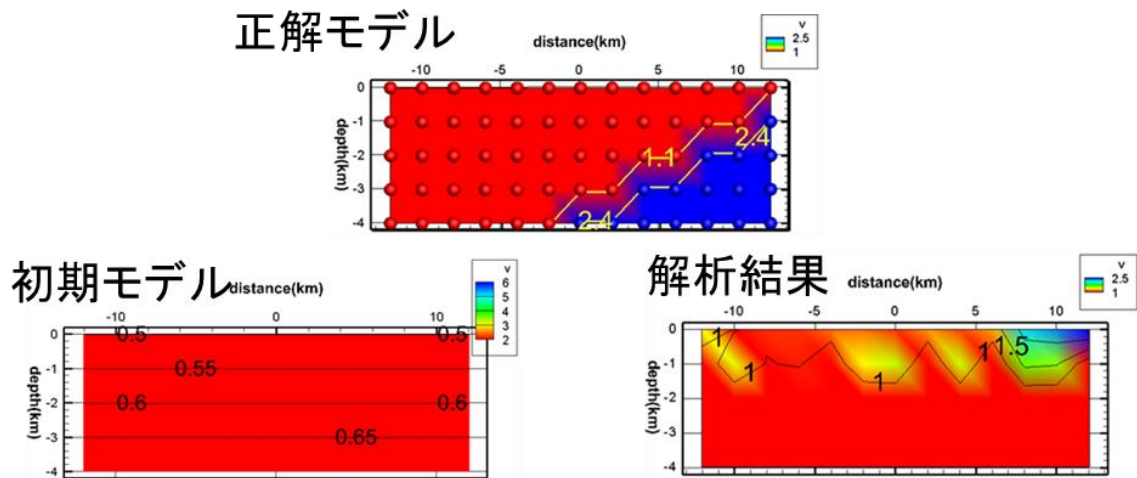


図 2.4 最小二乗法での解析例（春日部、2016）<sup>3)</sup>

## 2.3 非線形最小二乗法による解析概要と課題

本節では、走時トモグラフィ解析の逆解析に非線形最小二乗法を用いた場合の解析概要、既往の研究での課題解消への取り組み、そして、既往の研究結果を紹介する。

### 2.3.1 非線形最小二乗法によるトモグラフィ解析

非線形最小二乗法による解析手順を以下に示す。

1. 解析領域を離散化し、格子点に任意の弾性波速度を入力し、初期モデルを作成する。
2. 震源から観測点までの波線の到達時間（理論走時）を計算により求める（波線追跡）。
3. 波線追跡によって求めた理論走時と観測走時の誤差が減少するように速度構造を修正する。
4. 2-3 の工程を与えられた回数行うか、理論走時と観測走時の誤差が一定値以下になったら解析を終了する。

解析フローチャート（図 2.5）は以下の通りである。

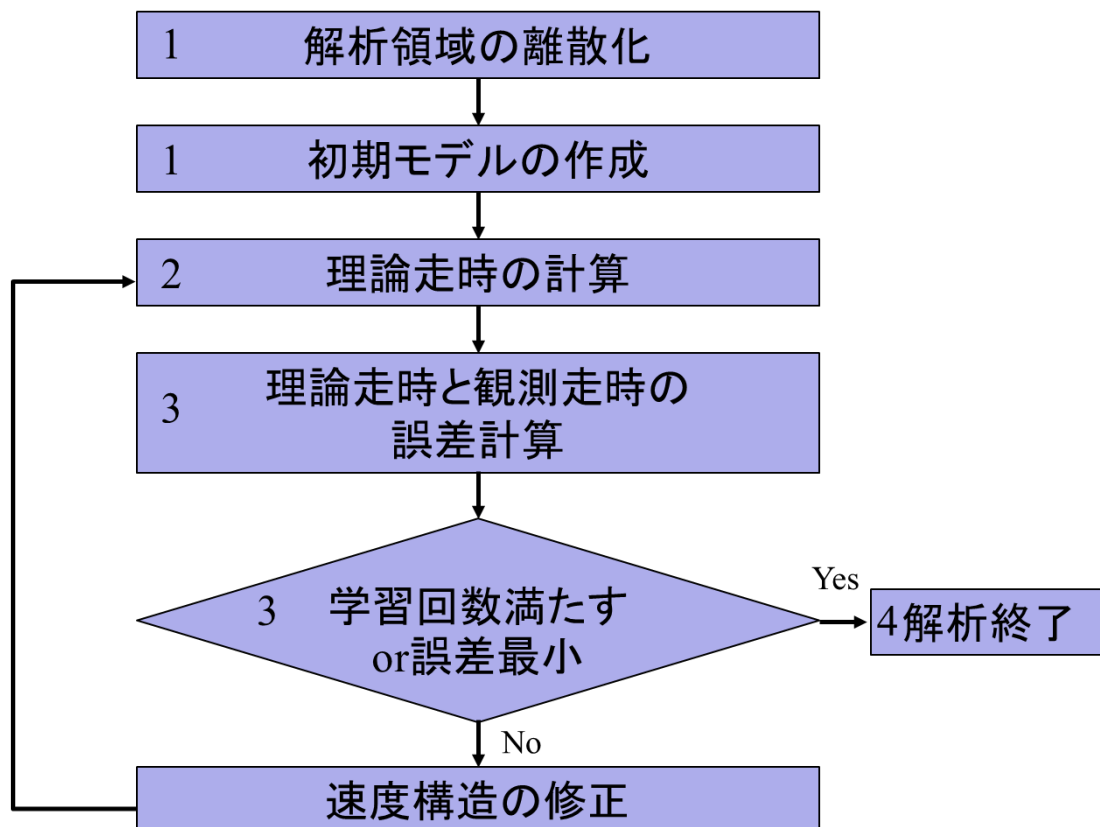


図 2.5 解析フローチャート

### 2.3.2 非線形最小二乗法の課題への取り組み

2.2 で述べたとおり、非線形最小二乗法を解析に用いた際に 2 つの課題がある。

1 つ目の課題である離散化における分割サイズの決定方法として、モデルの隣り合う各グリッドに速度変化を与えたモデルを真の速度構造として逆解析を行うチェッカーボードレボリューションテスト(Humphreys and Clayton,1988)<sup>4)</sup>を活用する方法がある。チェッカーボードレボリューションテストを様々な格子点間隔で実施し、その中から最も解像度の良い格子点間隔を選出することで、最適な格子点間隔を決定することができる。

2 つ目の課題の適切な初期モデルを与える必要がある点について、春日部(2016)<sup>3)</sup>は、3 つの異なる初期モデルを用意し、それぞれの解析結果を比較した。以下に結果(図 2.6)を示す。図 2.6a が正解モデル、図 2.6b I、b II、b III はそれぞれ初期モデル、図 2.6c I、c II、c III はそれぞれ解析結果となっている。初期モデルは I、II、III の順に正解モデルに近づいている。図 2.6 から初期モデルが正解値から離れるに連れ、正解モデルの再現度が低くなることが判った。実際の地下構造探査においては、正解値に近い初期モデルを設定するのは容易なことではない。そこで、白石ほか(2010)<sup>5)</sup>では、初期モデル依存性の問題に対して、ランダムで出力された多数の初期モデルを用いて、それぞれ

出力されたモデルの平均値を取ったとき、およそ何個ランダムな初期モデルを用意すれば安定した解が得られるのか検証した。その結果、100 個程度では不安定、200 から 300 個にかけて解の収束が安定すると結論づけられている。つまり、1 つの初期モデルに対する逆解析に時間を要するような解析ケースでは、安定した解を得るために多くの時間が必要となる。

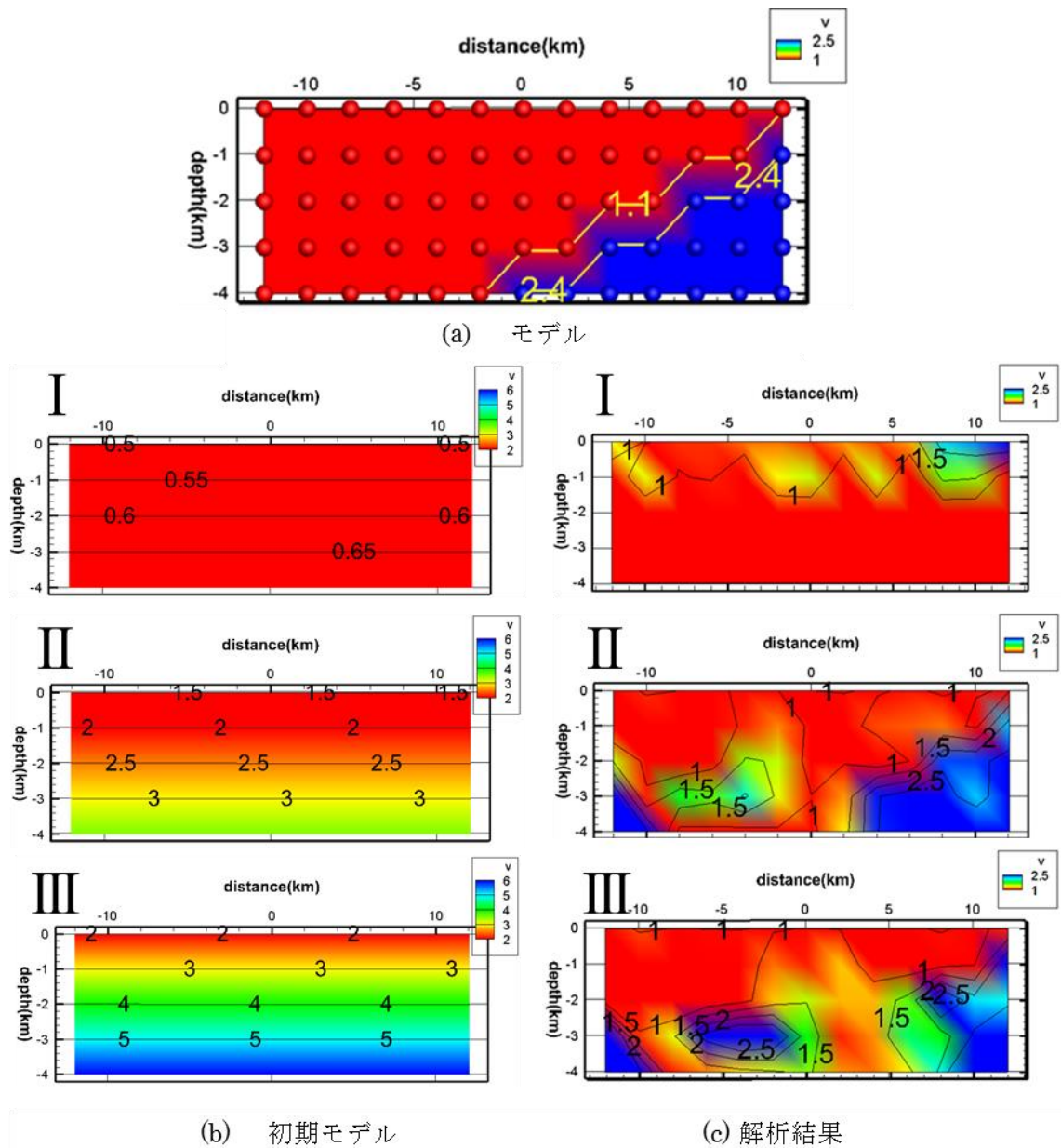


図 2.6 最小二乗法での初期モデルの検討例（春日部、2016）<sup>3)</sup>

## 2.4 走時トモグラフィで用いられるニューラルネットワーク

先に述べた非線形最小二乗法を用いた場合の課題を解消するために、解析領域を離散化す



る必要がなく、初期モデルも与える必要がないニューラルネットワークを用いた走時トモグラフィが提案された(春日部、2016)<sup>3)</sup>。この方法は、任意座標を入力、その点のスローネスを出力とするニューラルネットワークを設計し、波線経路上の標本点（あるいはサンプリングポイント）を利用して学習を行うことで、波線密度に応じた分解能で地下構造を推定できるという特徴を持っている。

#### 2.4.1 ニューラルネットワーク

ニューラルネットワークを構成するニューロン（図 2.7）は、人間の脳の神経回路の仕組みを模したモデルである。ニューロンは、細胞体、軸索、そして樹状突起から構成されている。細胞体がニューロンの本体、樹状突起が入力端子、そして軸索が出力端子の役割を果たしている。シナプスはニューロンからの信号を他のニューロンへ伝達する役割がある。ニューロン単体では至極少ない処理能力しかないが、シナプスによって結び付けられて多数集まることで複雑な処理も可能にしている。このニューロンを数理モデル化したものの組み合わせがニューラルネットワークである。

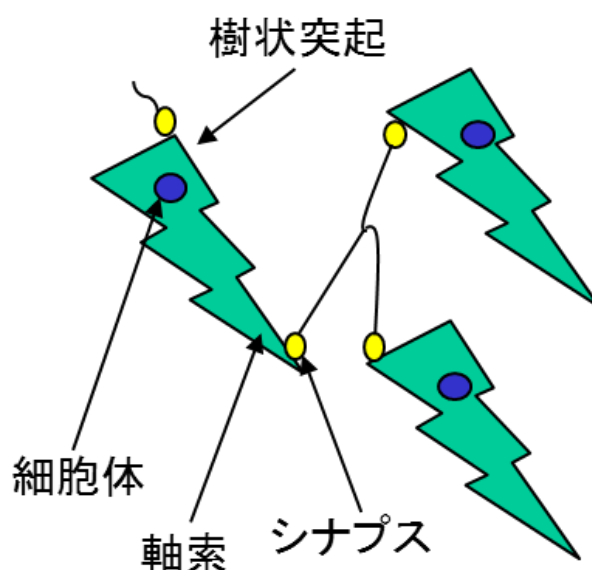


図 2.7 ニューロンの構成

#### 2.4.2 階層型ニューラルネットワーク

階層型ニューラルネットワーク（図 2.8）はニューロンで層を構成し、ニューロンが信号を入力層から出力層に向かって伝達していくネットワークである。ネットワークへの入力をする第 1 層を入力層、ネットワークからの出力をする最終層を出力層、そして、中間の層を隠れ層と呼ぶ。ニューロンによる情報の伝達は、隣の層のいずれかのニューロンに行い、同じ層同士、層を飛ばしての伝達、また、逆向き方向の伝達は禁止されている。一般的に隠



れ層の数が多いほどネットワークの能力は高いといわれるが、ただ多くしても処理時間を増やすだけで無駄になるため、処理の目的に合わせて必要最低限の隠れ層を設けるのが良い。

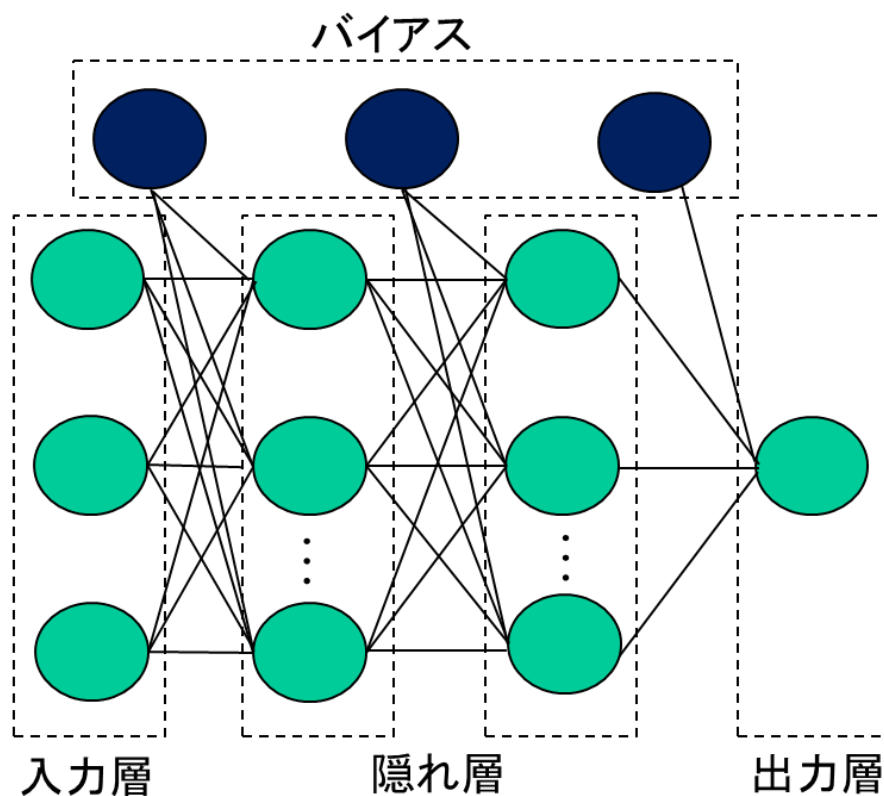


図 2.8 階層型ネットワーク

#### 2.4.3 ニューロン（ユニット）の入出力処理

図 2.9 はある層のユニットが前の層のユニットからの入力を受けて、次の層のユニットへ出力する様子を図示したものである。

ユニットの入力は、前層  $i$  番目のユニットからの入力信号を  $X_i$ 、入力信号に対する重みを  $W_i$  とし、前層の全ユニットからの入力信号と重みの積和にバイアス値  $W_0$  を加えたものとなる。各ユニットは前層からの入力を伝達関数に代入して出力する。また、ニューラルネットワークの学習に用いる伝達関数は非線形関数でなければならない。仮に伝達関数に線形関数

$$f(x) = wx \quad (2.1)$$

を用いたとする。これが 3 層重なったとすると、

$$y = f(f(f(x))) \quad (2.2)$$

となり、

$$y = w * w * w * x \quad (2.3)$$

(2.3)のように出力層と入力層のみで表現できてしまうため、隠れ層を設けている意味がなくなってしまうため伝達関数は非線形関数とする。

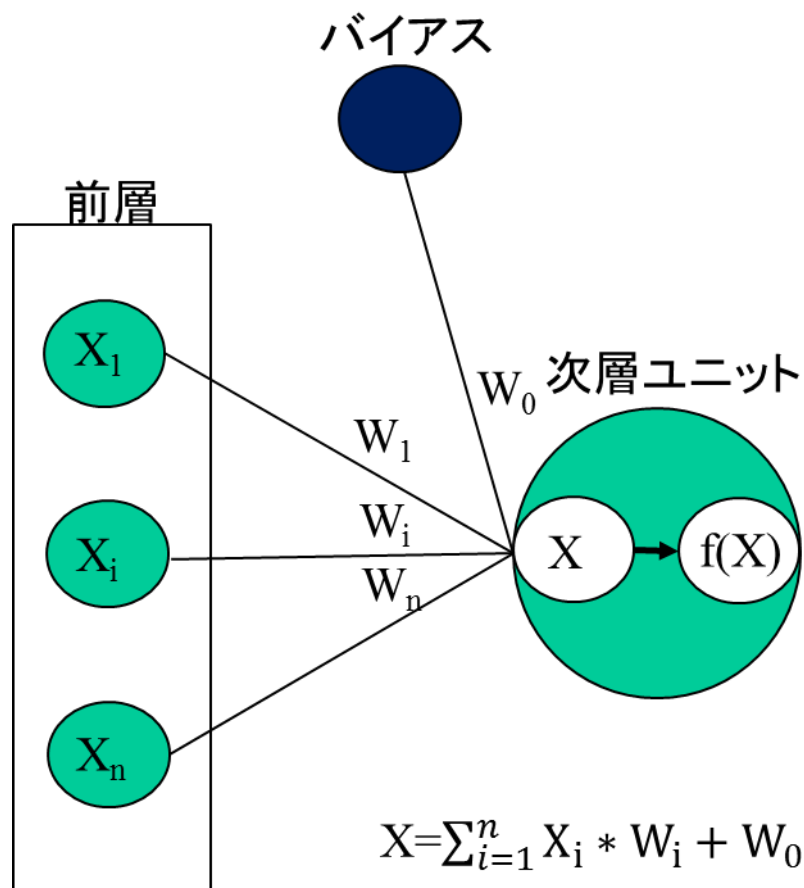


図 2.9 ニューロンのモデル化

## 2.5 階層型ニューラルネットワークの学習

階層型ニューラルネットワークの学習には誤差逆伝播(Rumelhart et al.,1986)<sup>6)</sup>を用いた。誤差逆伝播法とは、ニューラルネットワークの順伝播により求められた出力値と、正解値との二乗誤差を比較し、誤差を 0 に収束させるよう重みやバイアスを修正していく学習方法

である。

入力層以外の各層ユニットは、前層の全てのユニットから（出力×重み）という入力を受け取って、その総和にバイアス値を加算した後に、伝達関数へ代入して、結果を出力する。以下に春日部(2016)<sup>3)</sup>に従い重みの修正量の導出過程を述べる。

伝達関数にシグモイド関数を用いた場合のときを考える。 $i_j^k$ 、 $o_j^k$ をそれぞれ第  $k$  層の第  $j$  ユニットの入力の総和、出力とし、 $W_{ij}^{k-1k}$ を第  $k-1$  層の第  $i$  ユニットの第  $k$  層の第  $j$  ユニットの重み、 $W_{0j}^k$ を第  $k$  層の第  $j$  ユニットのバイアス値とすると

$$i_j^k = \sum_i W_{ij}^{k-1k} o_i^{k-1} + W_{0j}^k \quad (2.3)$$

$$o_i^k = f(i_j^k) \quad (2.4)$$

である。出力層からの出力と正解である教師信号  $ts_j$  との誤差の 2 乗和は、

$$E = \sum_j (o_j^m - ts_j)^2 \quad (2.5)$$

である。重み  $W$  の修正量  $\Delta W$  は  $E$  の  $W$  についての偏微分であるから、 $\varepsilon$  を正の微少量として、

$$\Delta W_{ij}^{k-1k} = -\varepsilon \frac{\partial E}{\partial W_{ij}^{k-1k}} \quad (2.6)$$

で表せる。

$$\begin{aligned} \frac{\partial E}{\partial W_{ij}^{k-1k}} &= \left( \frac{\partial E}{\partial i_j^k} \right) \left( \frac{\partial i_j^k}{\partial W_{ij}^{k-1k}} \right) \\ &= \left( \frac{\partial E}{\partial i_j^k} \right) o_i^{k-1} \end{aligned} \quad (2.7)$$

であり、 $k \neq m$  のときは、

$$\begin{aligned} \frac{\partial E}{\partial W_{ij}^{k-1k}} &= \left( \frac{\partial i_1^{k+1}}{\partial i_1^{k+1}} \right) \left( \frac{\partial i_1^{k+1}}{\partial o_j^k} \right) \left( \frac{\partial o_j^k}{\partial i_j^k} \right) \\ &= \sum_1 \left( \frac{\partial E}{\partial i_1^{k+1}} \right) W_{ji}^{kk+1} f'(i_j^k) \end{aligned} \quad (2.8)$$

となる。

また  $k = m$  のときは、

$$\frac{\partial E}{\partial i_j^k} = 2(o_j^m - ts_j)f'(i_j^m) \quad (2.9)$$

となる。

$\frac{\partial E}{\partial i_j^k} = d_j^k$  とおき、修正量は式(2.8)で計算できる。

$$\begin{aligned} \Delta W_{ij}^{k-1k} &= -\varepsilon d_j^k o_j^{k-1} \\ d_j^k &= (\sum_1 W_{j1}^{kk+1} d_1^{k+1}) f'(i_j^k) \\ d_j^m &= 2(o_j^m - ts_j)f'(i_j^m) \end{aligned} \quad (2.10)$$

ここで

$$f(i_j^k) = f(x) = \frac{1}{1+\exp(-x)} = o_j^k \quad (2.11)$$

だから

$$f'(x) = f(x)\{1 - f(x)\} = o_j^k(1 - o_j^k) \quad (2.12)$$

となる。実際には、 $\Delta W$ は収束を進めるために

$$\Delta W_{ij}^{k-1k}(s+1) = -\varepsilon d_j^k o_j^{k-1} + \alpha \Delta W_{ij}^{k-1k}(s) \quad (2.13)$$

で計算する。ここで、 $s$ は修正の回数、 $\alpha$ は正の微小量である。

誤差逆伝播法で学習を進めるには、上記の工程で重みの更新を行い、更新された重みでニューラルネットワークの順伝播を行い、出力から速度構造モデルを再度作成させる流れである。

## 2.6 ニューラルネットワークを用いた走時トモグラフィ

ニューラルネットワークを用いた走時トモグラフィによる解析手順（図 2.10）を以下に示す。

1. ニューラルネットワークに重みを与え、順伝播により初期モデルを作成する。
2. 波線をランダムに 1 本選出し、波線追跡を行う。
3. 波線経路上の標本点の空間座標値をニューラルネットワークの入力層に入力し、標本点

のスローネスを出力させる。

4. スローネスを波線に沿って積分することにより理論走時を算出する。
5. 理論走時と観測走時の誤差が収束するように重みの修正を行い、修正された重みを用いて順伝播により新たなモデルを作成する。
6. 2～5 の工程を決めた回数（学習回数）を行い学習が終了する。
7. 学習を終えたニューラルネットワークに任意の空間座標値を入力し、スローネスを出力させ、速度に変換することで解析領域の速度構造を推定する。

なお、波線追跡には Pseudo-Bending method(Um and Thurber,1987)<sup>7)</sup>を用いた。

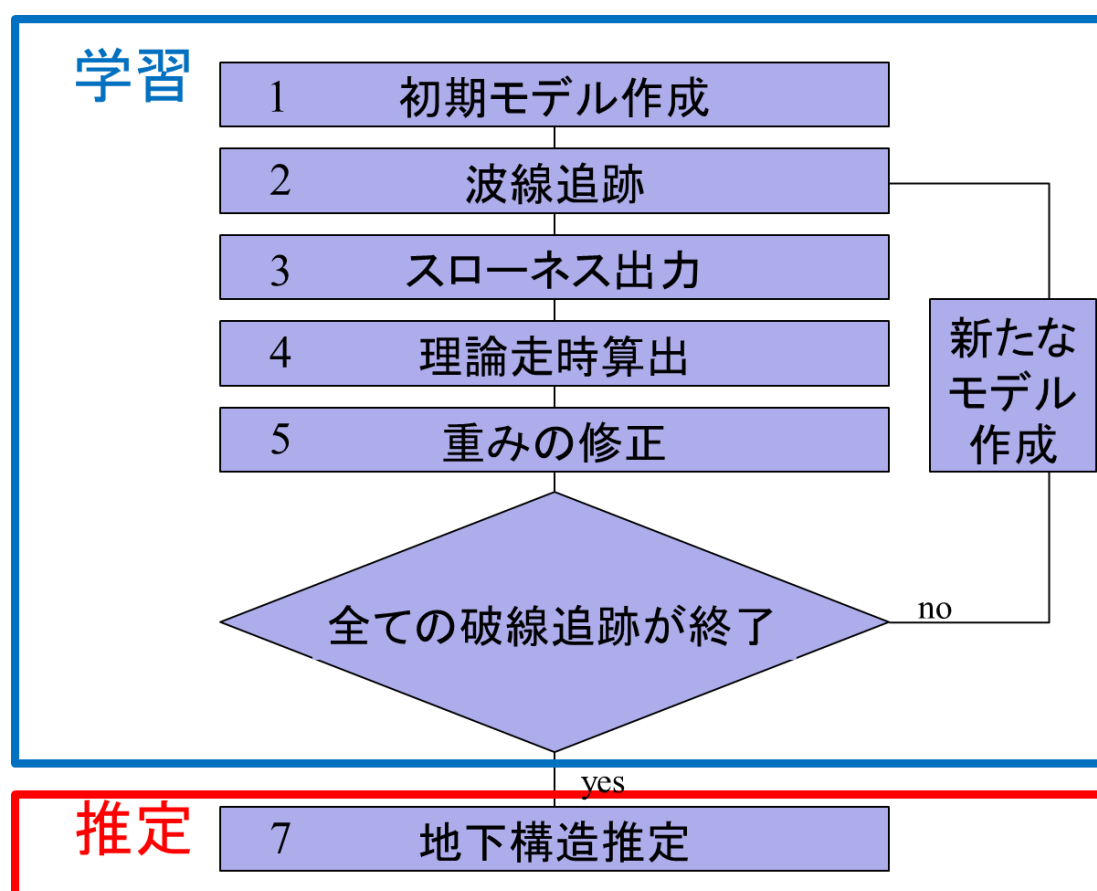


図 2.10 解析手順

## 2.7 ニューラルネットワークを用いた走時トモグラフィの課題

既往の研究からニューラルネットワークを用いた走時トモグラフィには、不連続な構造に対して解析精度が良くない。これは、ニューラルネットワークへの入力に近い値だと出力値も自ずと近い値が出力されてしまうからである。図 2.11 は春日部(2016)<sup>3)</sup>の結果である。地表付近に観測点と震源を水平方向に設置し、全ての波線をランダムに選出して、2.6で述べた解析をした結果であるが、解析結果が正解値とかけ離れてしまうという課題があ

る。

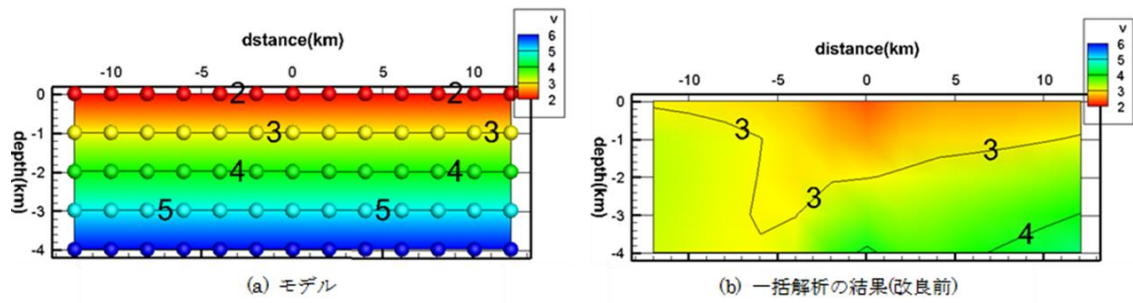


図 2.11 解析結果（春日部、2016）<sup>3)</sup>

ニューラルネットワークを用いた走時トモグラフィ解析では、全ての波線からランダムにひとつ選び、その走時データを学習に用いている。図 2.11 のような結果となったのは、波線が解析領域の下部まで到達しておらず、波線が通っていない箇所の推定ができなかったことが原因のひとつとして考えられる。この傾向は解析領域下部に急な速度変化を有している地下構造だとさらに顕著になる。図 2.11 の結果では乱数を与えて 1 回のみの出力であるが、乱数を複数回出力すると良好な結果を得られる場合がある。そのため、ニューラルネットワークを用いた走時トモグラフィにおいてパラメータの改良を行った上で、乱数を複数回発生させてサンプルを複数取るとは検討の余地があると考えている。

### 第3章 ニューラルネットワークを用いた走時トモグラフィの改良

#### 3.1 使用したニューラルネットワーク

本研究で用いるニューラルネットワーク（図 3.1）は、春日部（2016）<sup>3)</sup> で用いられているものと同様のニューラルネットワークを使用した。パターン認識に適しているといわれている階層型ネットワークを用い、入力層に空間座標地を入力し、出力層では弾性波速度の逆数（スローネス）を出力させ、対象領域内の弾性波速度の 2 次元分布を再構成する。ネットワークの構造は、入力層を 3 ユニット、隠れ層を 2 層で共に 30 ユニット、出力層を 30 ユニットとした。

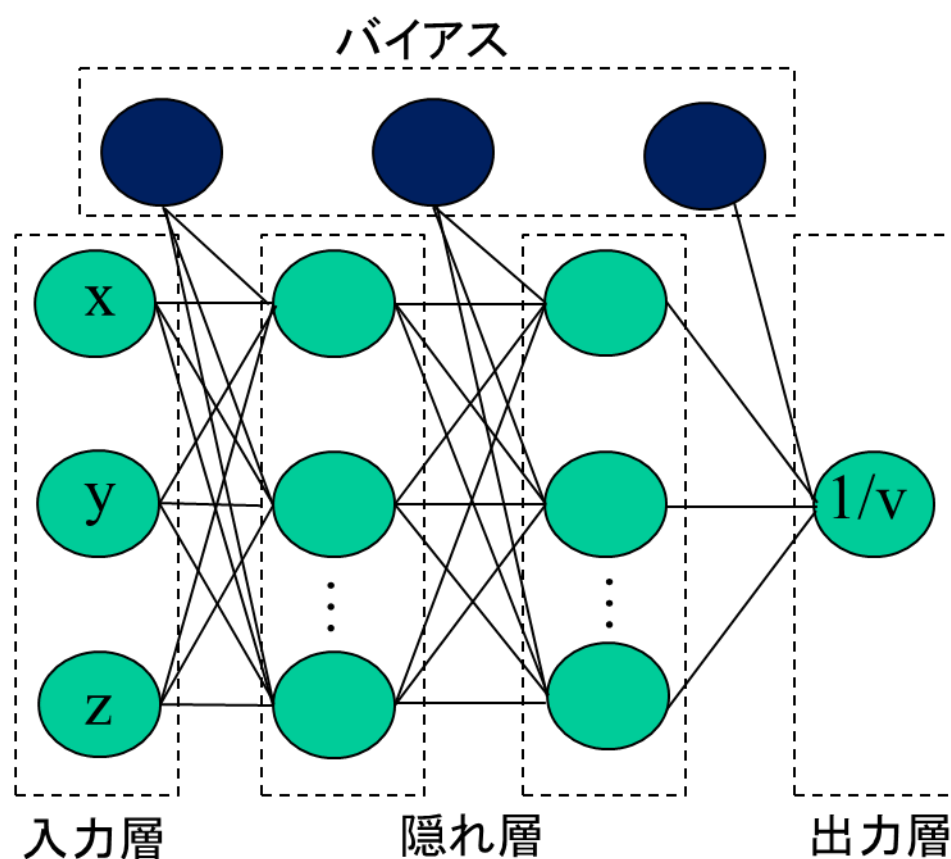


図 3.1 本研究でのネットワーク構造

## 3.2 本研究にて変更したパラメータ

本研究では 2.7 で述べた課題を踏まえて、解析結果を向上させるべく、学習回数、伝達関数、重みの初期値、そして、入力 of 初期値に関して改善の余地があると考えた。3.3.1～3.3.4 でそれぞれについて述べる。

### 3.2.1 学習回数

春日部 (2016) <sup>3)</sup> にて学習回数は一律 1000 回と設定されていた。しかし、このパラメータが本問題に対して、過学習、または、さらに学習を進めることにより結果が向上するのかわという議論はされていない。そこで本研究では、学習回数を 10、100、1000、2000、回の 4 パターンで行い、それぞれの学習回数で汎化能力を比較した。

### 3.2.2 伝達関数

本研究では、春日部 (2016) <sup>3)</sup> で用いられていたシグモイド関数(式 3.1)

$$f(x) = \frac{1}{1+e^{-x}} \quad (3.1)$$

に加え、近年ニューラルネットワークでの学習に多く用いられている ReLU(Rectified Linear Unit)関数(Xavier, Glorot.et,al.,2011)<sup>8)</sup>(式 3.2)

$$f(x) = \max(0, x) \quad (3.2)$$

を使用した。ただし ReLU 関数は、ニューラルネットワークの隠れ層のみで使用し、出力層では恒等関数(式 3.3)

$$f(x) = x \quad (3.3)$$

を用いた。シグモイド関数はニューラルネットワークの伝達関数として広く用いられてきた。誤差逆伝播による学習においてはシグモイド関数の微分

$$f'(x) = (1 - f(x))(f(x)) \quad (3.4)$$

を用いて学習を進める。入力する値が大きい場合や小さい場合は、出力の値すなわち、誤差の修正量が 0 に近似し、また、最大でも  $x = 0$  のときに出力は 0.25 となる。すなわち、重みの修正量が 1 回の学習に対して小さすぎて、学習の進みが非常に遅い。そして、隠れ層の数を増やすと、さらに勾配が消失して収束が遅くなる。これに対して、ReLU 関数の微分は入力が正の数であれば 1 となり、勾配消失の心配がない。しかし、入力が負の数であると、勾



配は 0 となり重みの修正が止まる。

本研究では、シグモイド関数、ReLU 関数の双方を用いてそれぞれの傾向、精度について比較した。ただし、ReLU 関数を用いたとき、順伝播の出力が 0 であると速度が $\frac{1}{0}$ となり、解析自体が停止してしまうため、下限値を 0.01 と設定した。

### 3.2.3 重みの初期値

本研究では、春日部 (2016) <sup>3)</sup> で用いられていた範囲-0.5~0.5 の一様乱数に加え、伝達関数にシグモイド関数を用いるときに推奨されている、前層のユニット数を  $n$  として、 $\frac{1}{\sqrt{n}}$  の標準偏差を持つガウス分布に従う乱数 (Glorot, et, al., 2010)<sup>1)</sup>、そして、伝達関数に ReLU 関数を用いるときに推奨されている、前層のユニット数を  $n$  として、 $\sqrt{\frac{2}{n}}$  の標準偏差を持つガウス分布に従う乱数 (He, et, al., 2015)<sup>2)</sup> の 3 パターンを重みの初期値として与えた。ReLU 関数の際は、 $x$  が負の値をとると出力が 0 になるため、シグモイドに比べてより広い広がりを持たせるという意味で $\sqrt{\frac{2}{n}}$ の標準偏差になっている。また、標準偏差が 0.01 であると全く学習ができないと斎藤 (2016) <sup>9)</sup> が示している。本研究においては、 $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布で第 1 層目が $\frac{1}{\sqrt{n}} = 0.5$ 、第 2 層目以降が $\frac{1}{\sqrt{n}} \approx 0.180$ 、 $\sqrt{\frac{2}{n}}$ の標準偏差を持つガウス分布で第 1 層目が $\sqrt{\frac{2}{n}} \approx 0.707$ 、第 2 層目以降が $\sqrt{\frac{2}{n}} \approx 0.254$ となった。ガウス分布に従う乱数の作成は、一様乱数を出力した後にボックス=ミュラー法(Box and Muller, 1958)<sup>10)</sup>を用いてガウス分布に変換した。一様乱数を任意のガウス分布に変換する式は以下の通りである。

$$Z_1 = a\sqrt{-2\log X}\cos 2\pi Y + b \quad (3.5)$$

$$Z_2 = a\sqrt{-2\log X}\sin 2\pi Y + b \quad (3.6)$$

確率変数  $X$ 、 $Y$  が互いに独立で、範囲 0~1 の一様乱数のとき、 $a$ 、 $b$  にそれぞれ指定した標準偏差、平均値を代入することで、標準偏差  $a$ 、平均  $b$  のガウス分布に従う乱数 $Z_1$ 、 $Z_2$  がそれぞれ作成できる。

### 3.2.4 入力 of 初期値

ニューラルネットワークの入力に空間座標値を用いると、隣り合う推定点のように近い

座標値を入力した際、出力値も似たような値が出力される傾向にある。この性質は、速度構造が比較的滑らかに変化するモデルであれば問題はないが、断層など速度不連続が存在する場合、表現することができない。そこで本研究では、空間座標値を直接入力する手法に加え、空間座標値をパターンコード化してユニットに入力する手法（パターンコーディング）の導入を試みた。

## 第4章 数値解析

### 4.1 解析領域

数値解析では、成層構造と断層構造を対象にし、2次元構造の解析を行った。解析領域は水平方向に 24km、深さ方向に 4km とした。震源と観測点は 0.5km 間隔で 48 点設置した。弾性波速度の推定点間隔は水平方向に 2km、深さ方向に 1km とした (図 4.1)。

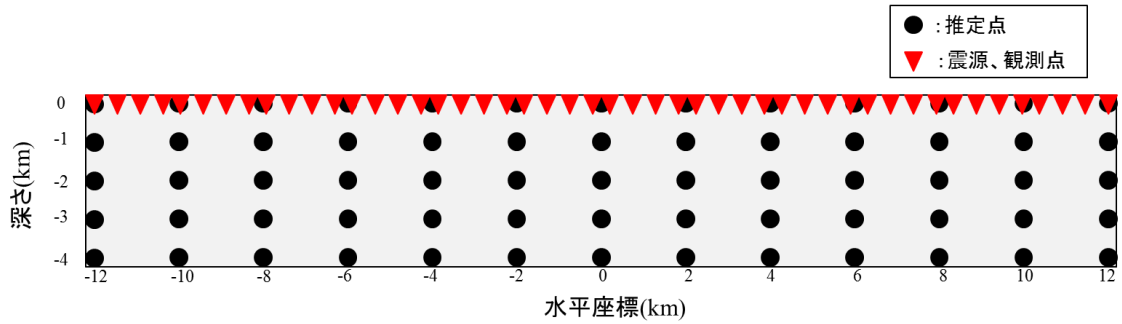


図 4.1 解析領域図

### 4.2 誤差の評価

本研究では、出力された地下構造が正解値と比較してどれほど近似しているか定量的に評価する目的で、推定点における弾性波速度の平均誤差を  $E_{v_p}$  と定義した。波線の走時残差は RMS 残差(平均二乗誤差)で評価する。

$$E_{v_p} = \frac{\sum_{i=1}^N |V_{p_{true}}^i - V_{p_{est}}^i|}{N} [km/sec] \quad (4.1)$$

$$RMS\text{残差} = \sqrt{\frac{\sum_{p=1}^P (t_{obs}^p - t_{cal}^p)^2}{P}} [sec] \quad (4.2)$$

ここで  $N$  は予測点数であり、 $V_{p_{true}}^i$  は  $i$  番目の予測点上の弾性波速度である。 $V_{p_{est}}^i$  は  $i$  番目の予測点におけるニューラルネットワークの出力値である。 $P$  は波線経路数であり、 $t_{obs}^p$  と  $t_{cal}^p$  はそれぞれ  $P$  番目の波線での観測走時と理論走時である。

### 4.3 解析結果

ニューラルネットワークの学習は、学習回数によって大きく解析精度が変化してくる。本研究における学習回数は、全波線の学習を1回として、10、100、1000、2000回の4パターンで行った。正解値モデルは、深さが1kmずつ深くなるほど、弾性波速度が1km/sずつ速くなる水平成層構造とした(図4.2)。

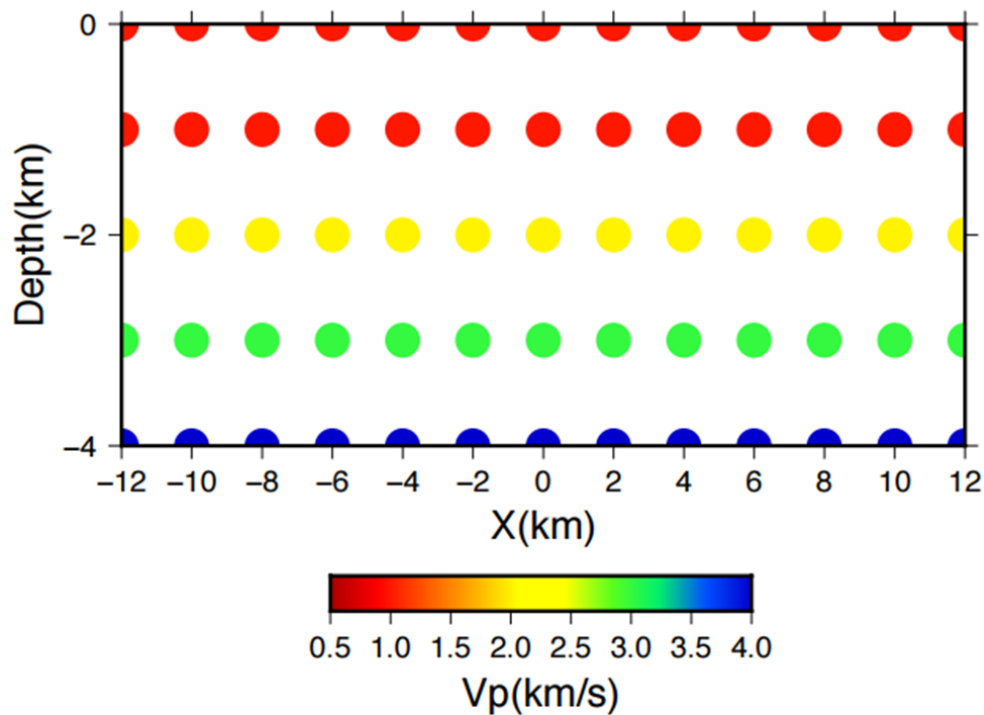
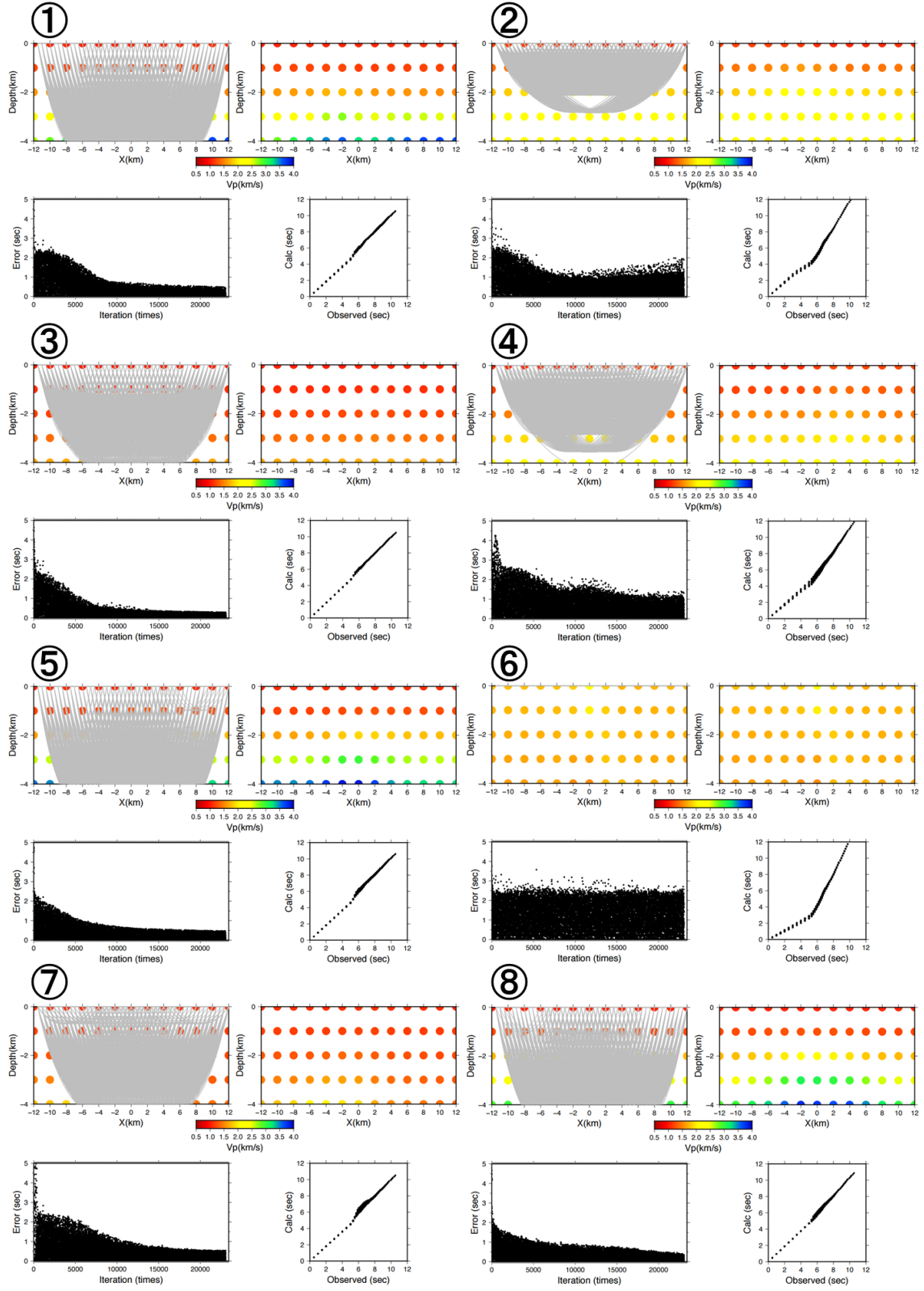


図 4.2 水平成層構造

#### 4.3.1 学習回数の比較

重みの初期値は範囲-0.5~0.5の一様乱数、伝達関数にシグモイド関数、そして、入力には空間座標値を用いたニューラルネットワークを使用して、学習回数の違いによる精度の変化を比較した。図4.3に学習回数10回の解析結果を示す。図4.3には解析結果である速度構造と合わせて、誤差の収束状況、理論走時と観測走時の関係も示している。また、初期重みをランダムで与えているため、それぞれ解析結果が異なる。従って、同じ解析を10回行った。得られたデータから誤差の最小値、最大値、そして、平均値を算出した(表4.1)。



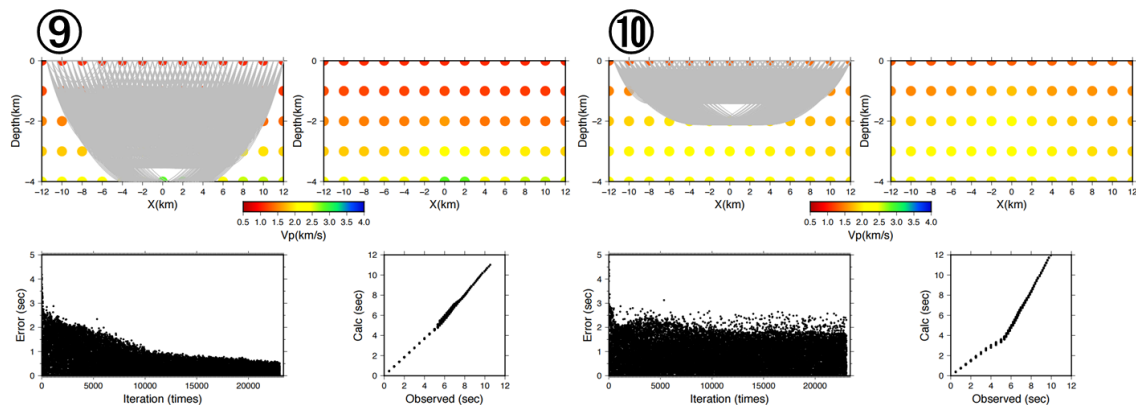


図 4.3 学習回数 10 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

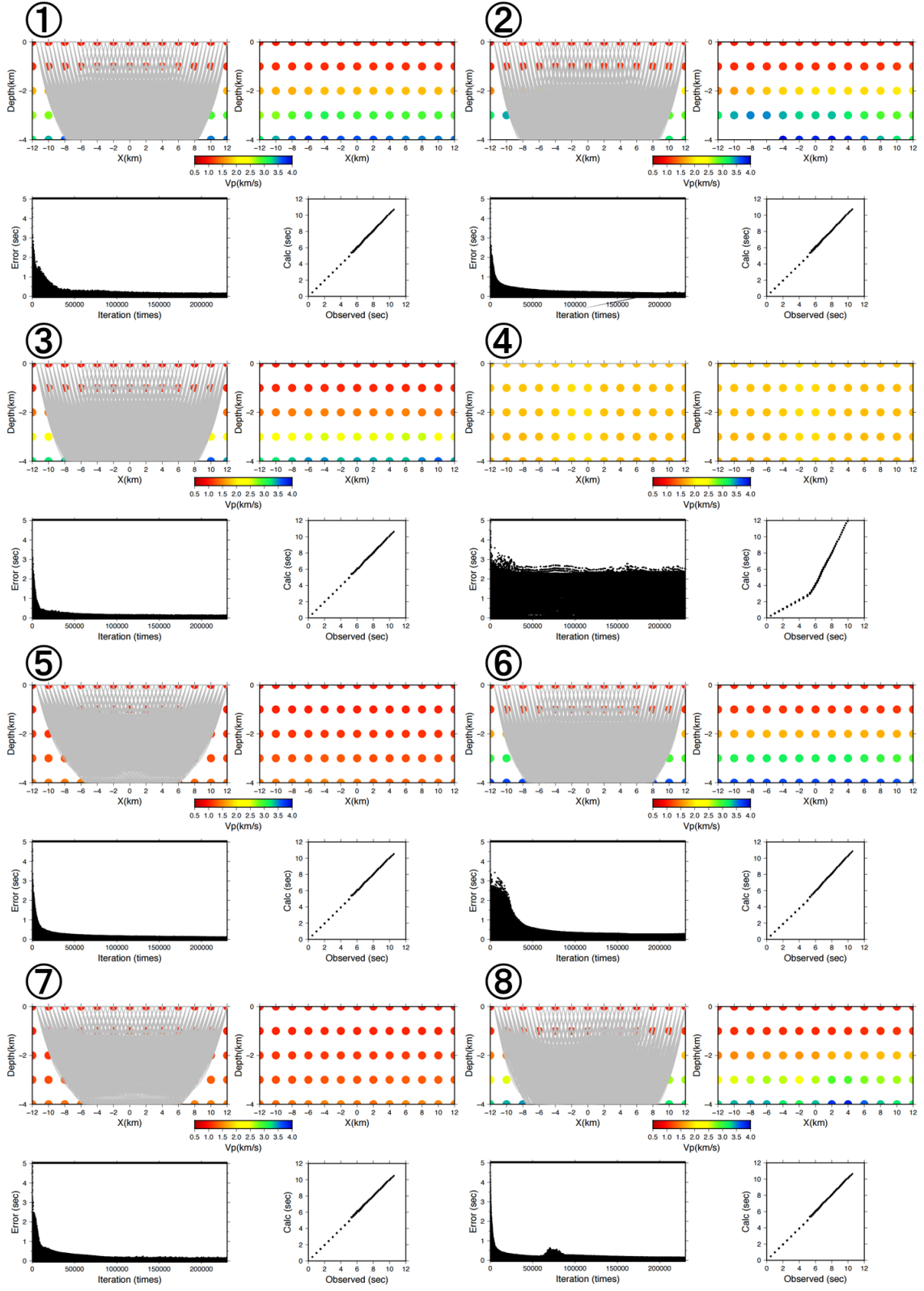
表 4.1 学習回数 10 回の平均誤差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
1	0.38	0.16
2	0.66	0.72
3	0.97	0.11
4	0.77	0.53
5	0.27	0.19
6	1.10	1.40
7	0.97	0.22
8	0.29	0.18
9	0.69	0.25
10	0.80	0.97
平均	0.69	0.47

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
最小	0.27	0.11
最大	1.10	1.40
平均	0.69	0.47

10 サンプルの解析結果を比べると、正解モデルに近いものと、そうでないものの差が大きい。これは学習回数が少ないことにより、収束しきれていないパターンがあるためである。特徴としては走時が 6 秒を超える波線に関して誤差の修正ができていない傾向がある。また、結果が良好だったものについては、ランダムで与えた重みにより与えられた初期モデルが正解値に近かったためだと思われる。

次に学習回数が 100 回の際の結果を図 4.4 と表 4.2 に示す。



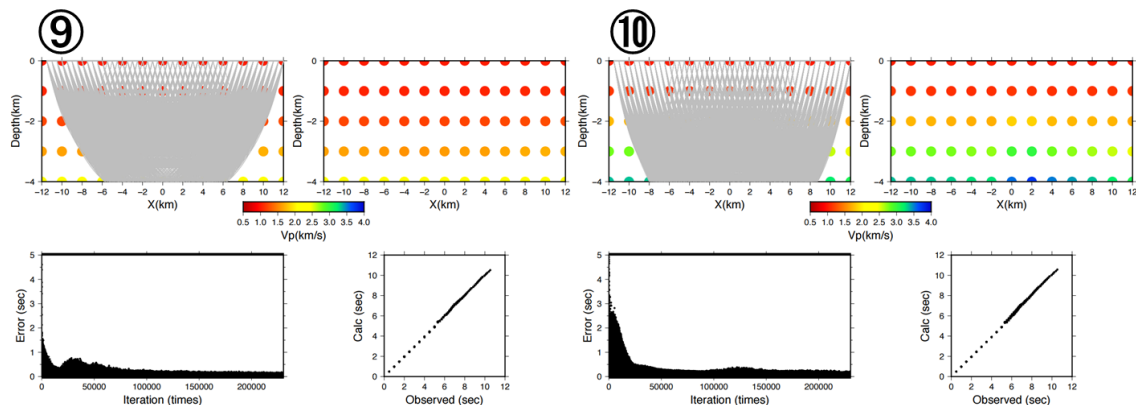


図 4.4 学習回数 100 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.2 学習回数 100 回の平均誤差 $E_{v_p}$ と RMS 残差

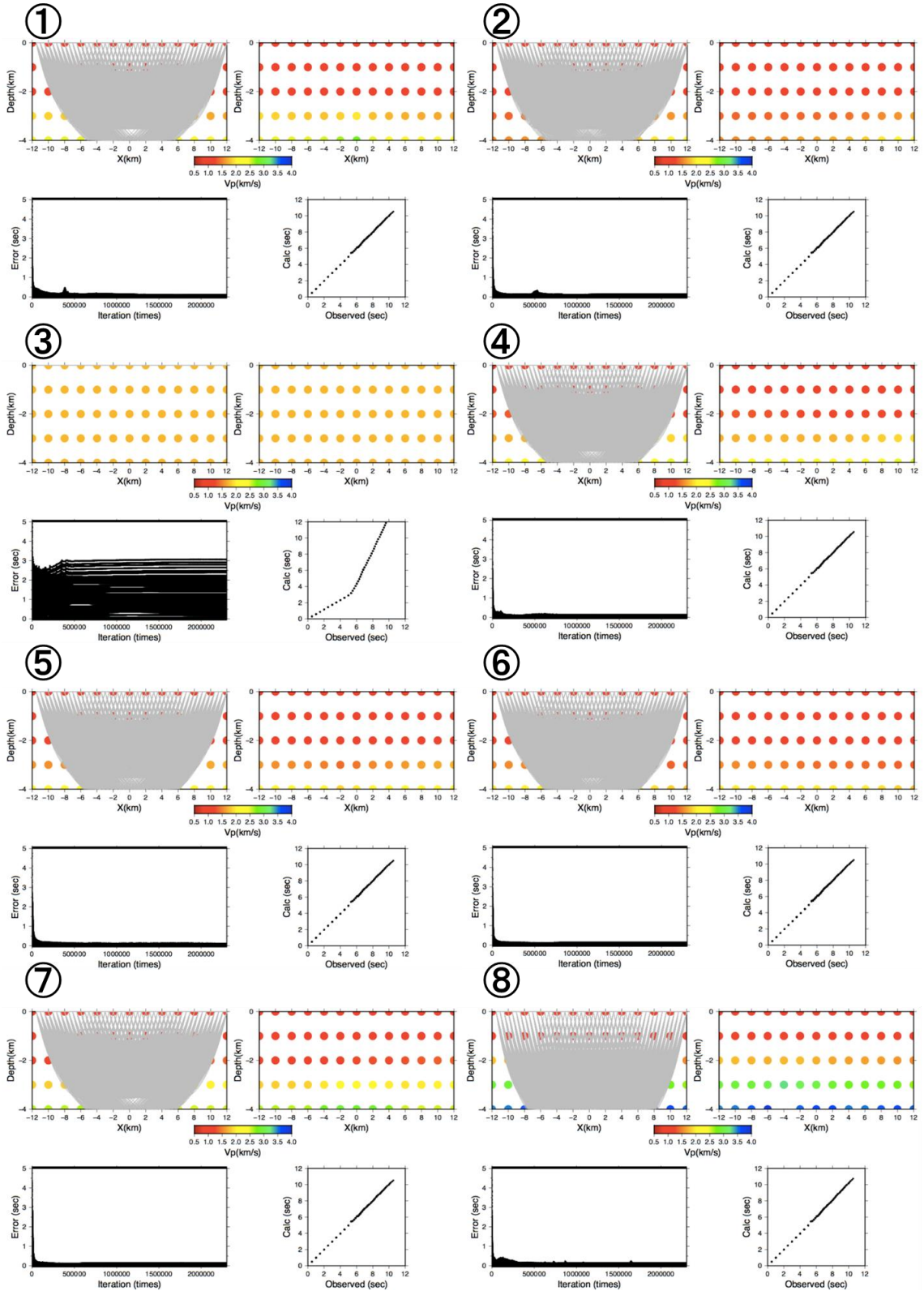
	$E_{v_p}[km/sec]$	$RMS$ 残差[sec]
1	0.19	0.08
2	0.18	0.10
3	0.36	0.06
4	1.06	1.40
5	1.06	0.05
6	0.18	0.13
7	1.04	0.05
8	0.29	0.08
9	0.81	0.05
10	0.26	0.09
平均	0.54	0.21

	$E_{v_p}[km/sec]$	$RMS$ 残差[sec]
最小	0.18	0.05
最大	1.06	1.40
平均	0.54	0.21

この結果も、正解値に近いものとそうでないものの差は激しいが、学習回数が 10 回のときに比べ、最小、最大、平均値のどれをとっても誤差が小さかった。しかし、図 4.4 の⑤や⑦のように、誤差は収束しているにもかかわらず、地下構造を推定できていないサンプルがあった。

続いて学習回数 1000 回の様子を図 4.5、表 4.3 に示す。





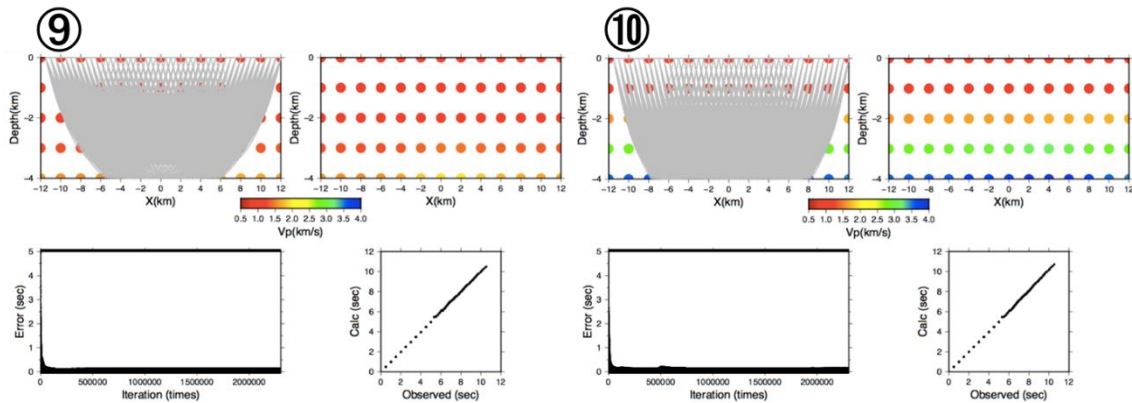


図 4.5 学習回数 1000 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

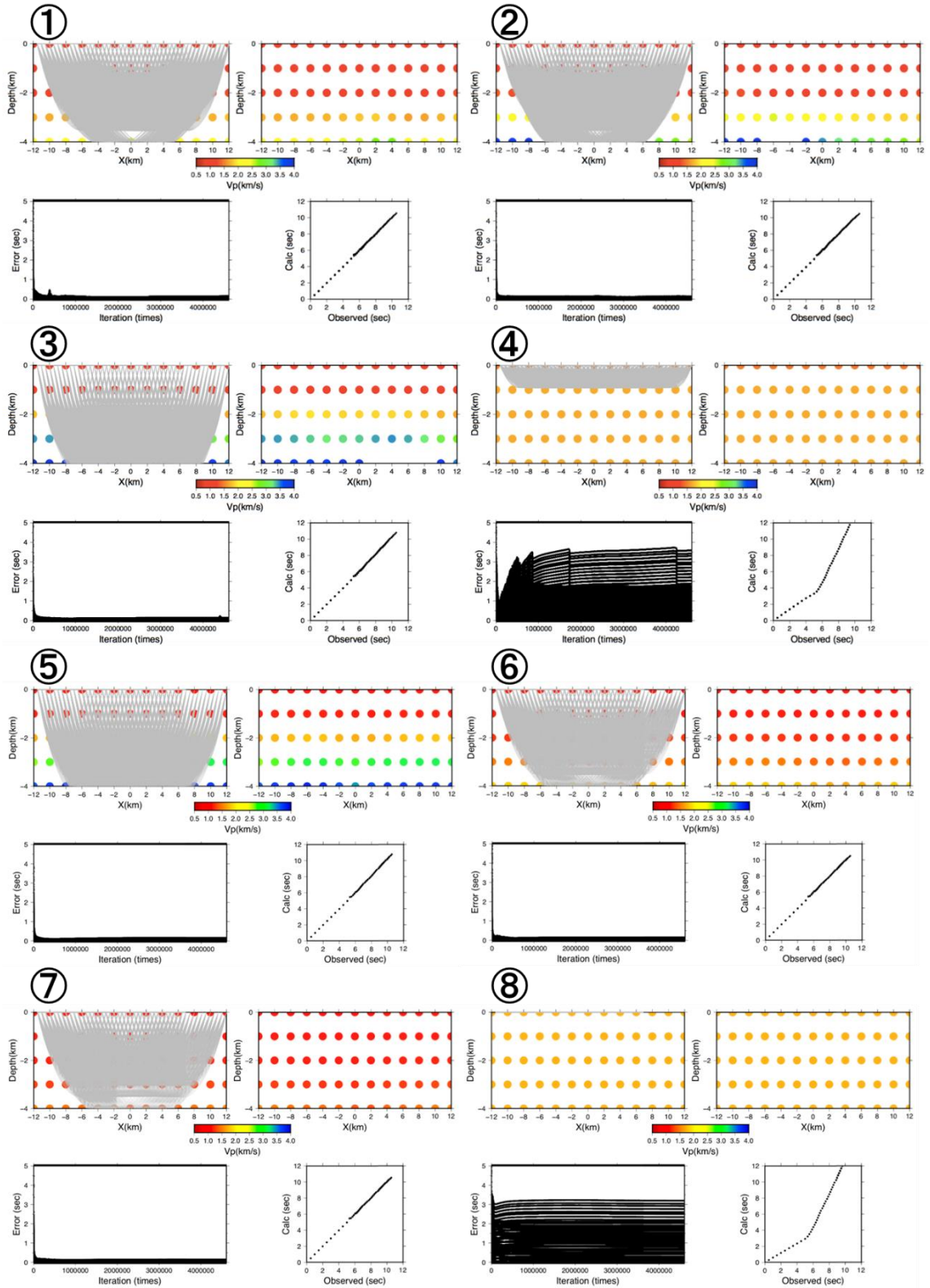
表 4.3 学習回数 1000 回の平均誤差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
1	0.73	0.04
2	0.95	0.04
3	1.06	1.35
4	0.76	0.04
5	0.84	0.04
6	0.90	0.04
7	0.63	0.04
8	0.18	0.08
9	0.98	0.04
10	0.18	0.08
平均	0.72	0.18

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
最小	0.18	0.04
最大	1.06	1.35
平均	0.72	0.18

学習回数 100 回のときに比べて、**RMS** 残差は減少したが、平均誤差は悪化した。推定された地下構造を確認しても、学習回数 100 回のときに比べ、ほとんど推定できていないサンプルも多かった。

最後に学習回数 2000 回の様子を図 4.6、表 4.4 に示す。



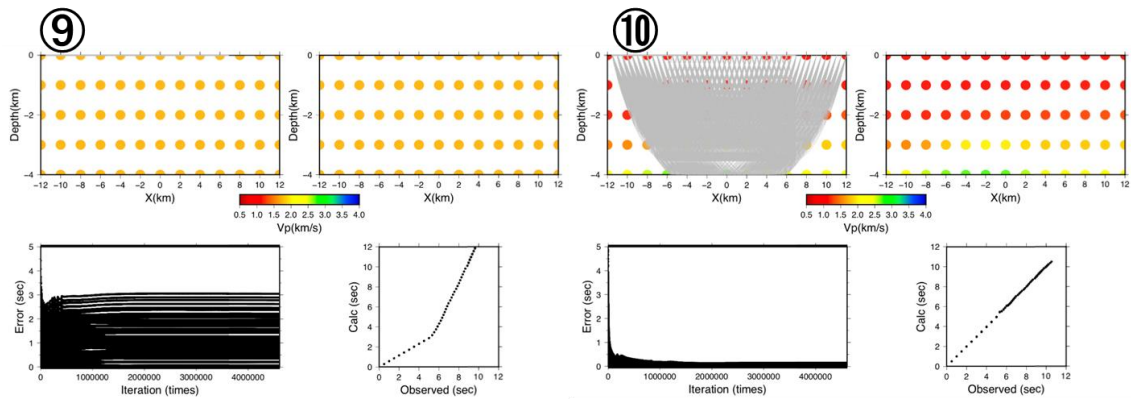


図 4.6 学習回数 2000 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.4 学習回数 2000 回の平均誤差 $E_{v_p}$ と RMS 残差

	$E_{v_p}[km/sec]$	RMS残差[sec]
1	0.76	0.04
2	0.46	0.05
3	0.15	0.10
4	1.02	1.21
5	0.15	0.08
6	0.94	0.04
7	1.08	0.04
8	0.01	1.34
9	1.05	1.35
10	0.71	0.04
平均	0.63	0.43

	$E_{v_p}[km/sec]$	RMS残差[sec]
最小	0.01	0.04
最大	1.08	1.35
平均	0.63	0.43

学習回数 1000 回に比べて、平均誤差は減少したが、RMS 残差は大幅に悪化した。図 4.6 からわかるように誤差が収束に向かわないものも数多かった。明らかに過学習と言えるだろう。よって、今回の観測ジオメトリでは、重みの初期値は範囲・0.5～0.5 の一様乱数、伝達関数にシグモイド関数、そして、入力には空間座標値を用いたニューラルネットワークの場合、学習回数 100～1000 回が妥当であると考えられる。

#### 4.3.2 重みの初期値を変更したときの学習回数比較

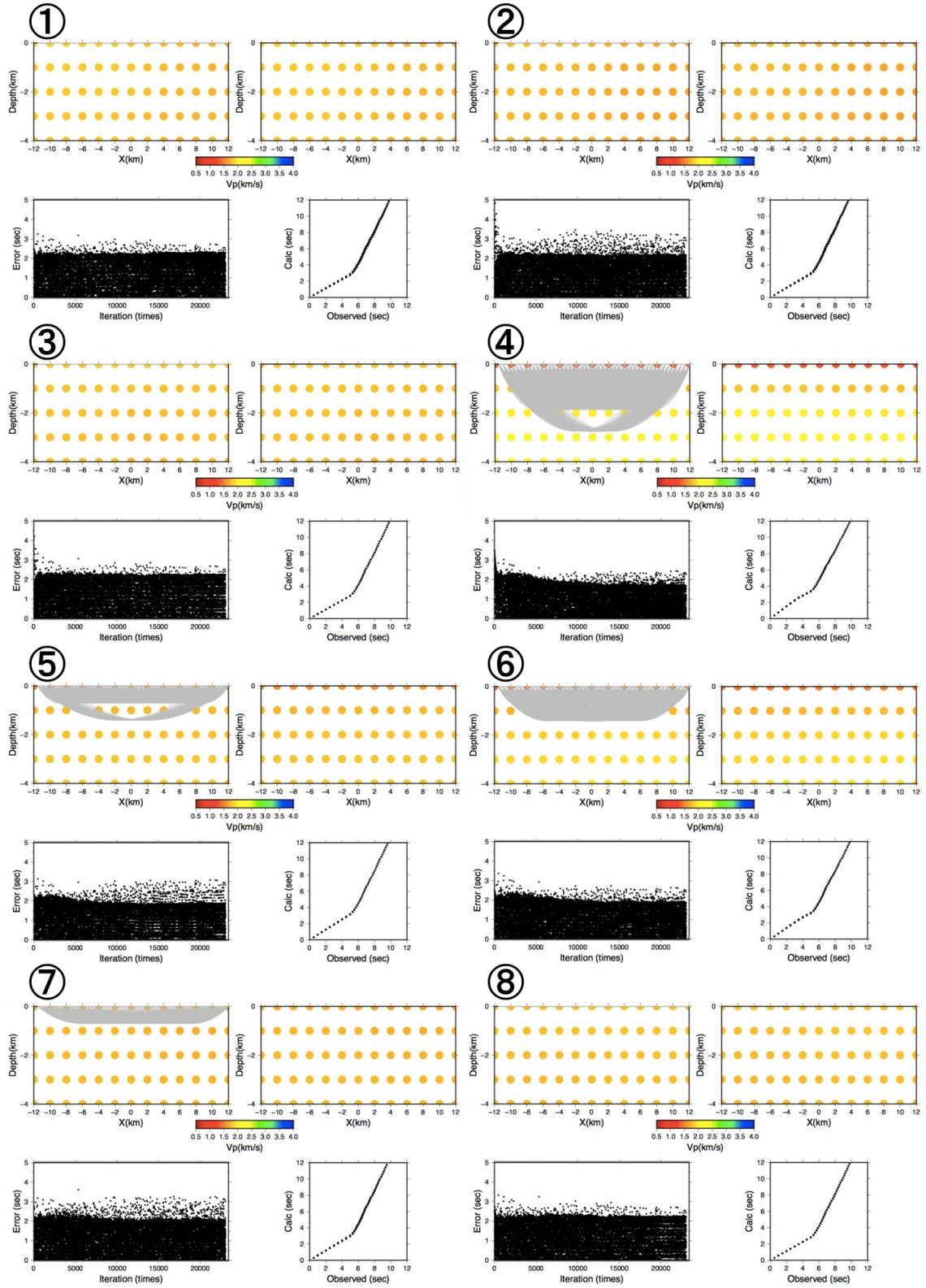
4.3.1 での解析から、重みの初期値のみ変更して、学習回数の比較を行った。伝達関数と

してシグモイド関数を使用する場合は、 $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布の重みが適していると言われているが、一様乱数の場合や、 $\sqrt{\frac{2}{n}}$ の標準偏差を持つガウス分布を用いたときの場合と比べてどのような特徴が見られるのか考察する。

(1)  $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布を用いたとき

以下に学習回数 10 回の様子を図 4.7、表 4.5 に示す。





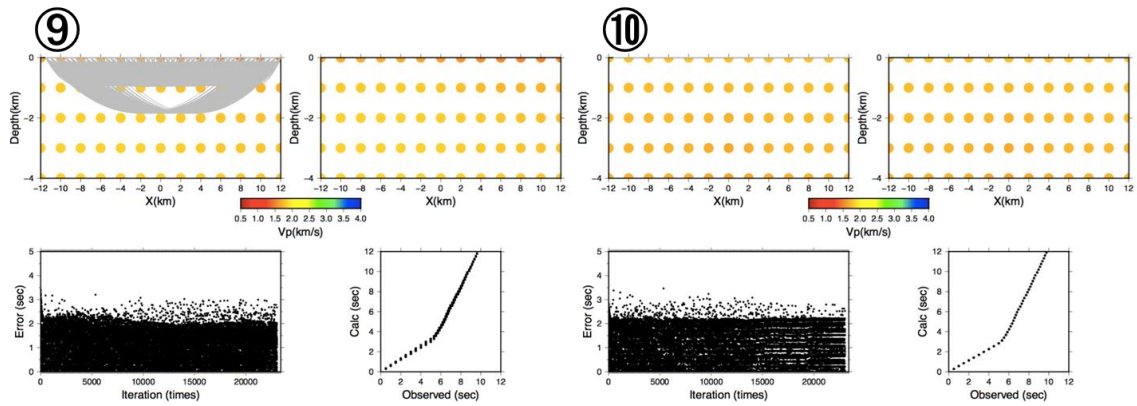


図 4.7 学習回数 10 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

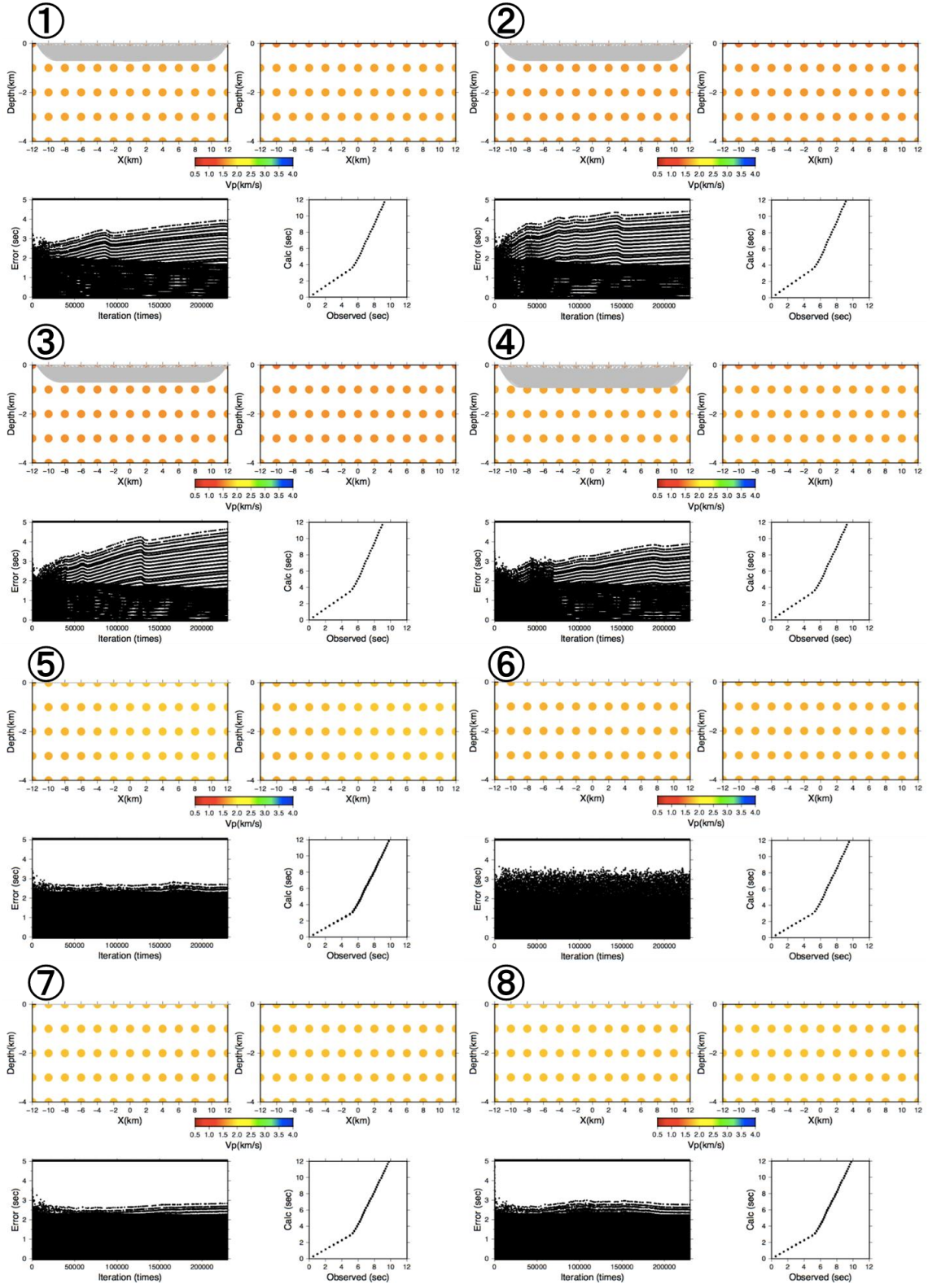
表 4.5 学習回数 10 回の平均誤差 $E_{v_p}$ と RMS 残差

	$E_{v_p}[km/sec]$	RMS残差[sec]
1	1.06	1.38
2	1.07	1.34
3	1.06	1.38
4	0.83	1.01
5	0.99	1.23
6	0.91	1.14
7	1.03	1.30
8	1.06	1.29
9	0.98	1.23
10	1.07	1.36
平均	1.01	1.27

	$E_{v_p}[km/sec]$	RMS残差[sec]
最小	0.83	1.01
最大	1.07	1.38
平均	1.01	1.27

一様乱数を用いた場合に比べて平均誤差が約 1.5 倍、RMS 残差が 3 倍弱という結果となった。ほぼ全てのサンプルに共通していることは、波線が解析領域浅部に集中して、深部をほとんど通過していないことである。誤差の収束もされていないことから、一様乱数と違い、学習回数を多く取らないと良好な結果にならない可能性がある。

次に学習回数 100 回の様子を図 4.8、表 4.6 に示す。





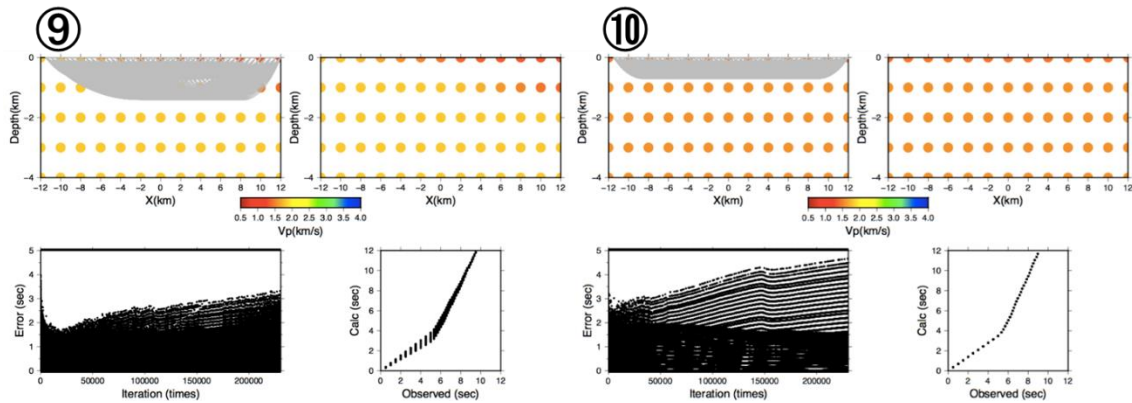


図 4.8 学習回数 100 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.6 学習回数 100 回の平均誤差 $E_{v_p}$ と RMS 残差

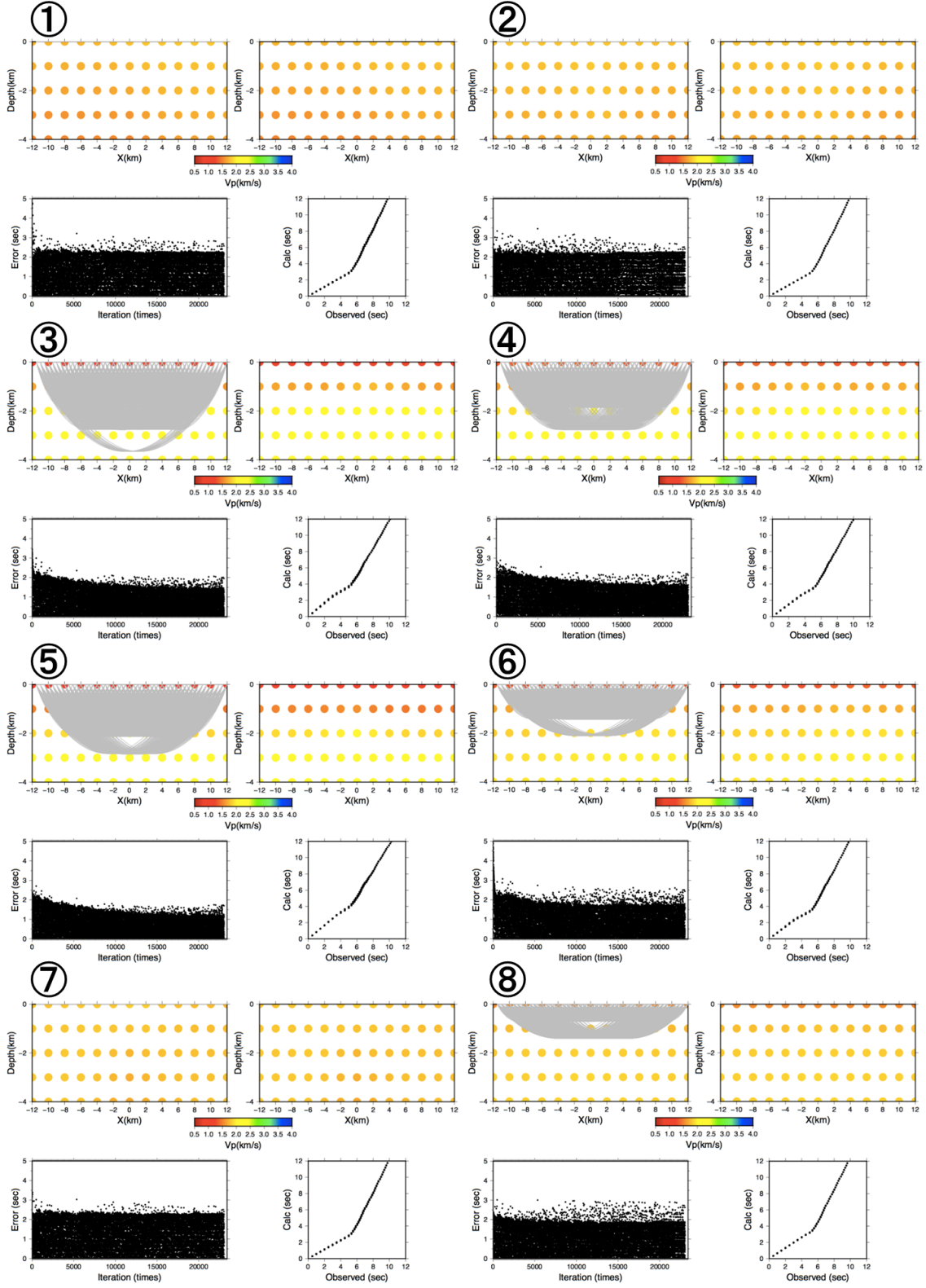
	$E_{v_p}[km/sec]$	RMS残差[sec]
1	1.04	1.25
2	1.05	1.33
3	1.06	1.38
4	1.04	1.26
5	1.05	1.37
6	1.06	1.34
7	1.05	1.36
8	1.05	1.37
9	0.95	1.19
10	1.06	1.38
平均	1.04	1.32

	$E_{v_p}[km/sec]$	RMS残差[sec]
最小	0.95	1.19
最大	1.06	1.38
平均	1.04	1.32

学習回数が 100 回になると、誤差が発散に向かっているものも多く見られた。波線もほとんど解析領域を通っておらず、結果も芳しくない。前述のように重みの標準偏差が 0.01 であると全く学習は行われない。今回の標準偏差は 0.180 と 0.5 である。標準偏差が 0.01 の 10 倍以上の値であるが、本研究のケースでは標準偏差が 0.180 程度であっても学習が進まなかった。学習回数が 1000 回、2000 回の場合も同様に精度は悪かった。なお、学習回数 1000 回、2000 回の結果は付録に添付する。

(2)  $\sqrt{\frac{2}{n}}$ の標準偏差を持つガウス分布を用いたとき

シグモイド関数を伝達関数として使用する際に、 $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布の重みを使用することが推奨されているが、良好な結果は出力されなかった。標準偏差の大きさに着目し、本来は ReLU 関数を伝達関数として使用する際に推奨されている $\sqrt{\frac{2}{n}}$ の標準偏差を持つガウス分布に従う乱数の重みの初期値として与え、伝達関数としてシグモイド関数を用いた場合の解析を行った。図 4.9、表 4.7 に学習回数 10 回の解析結果を示す。



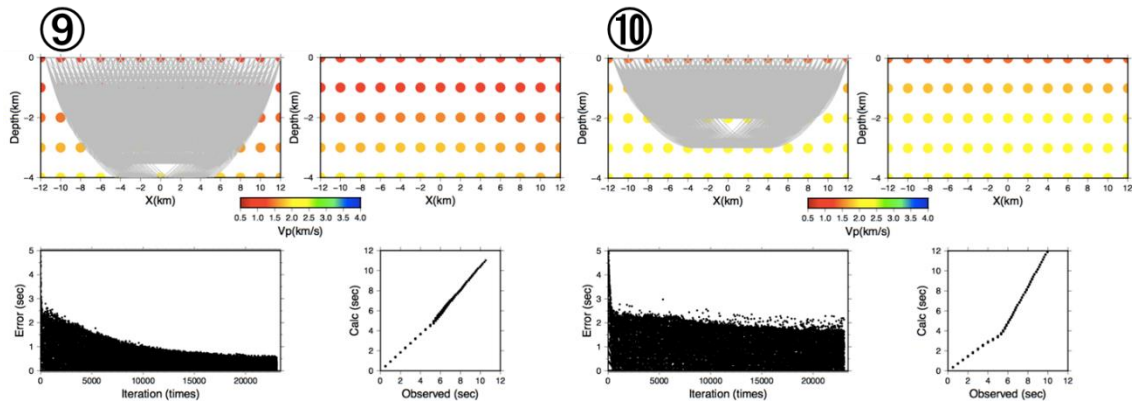


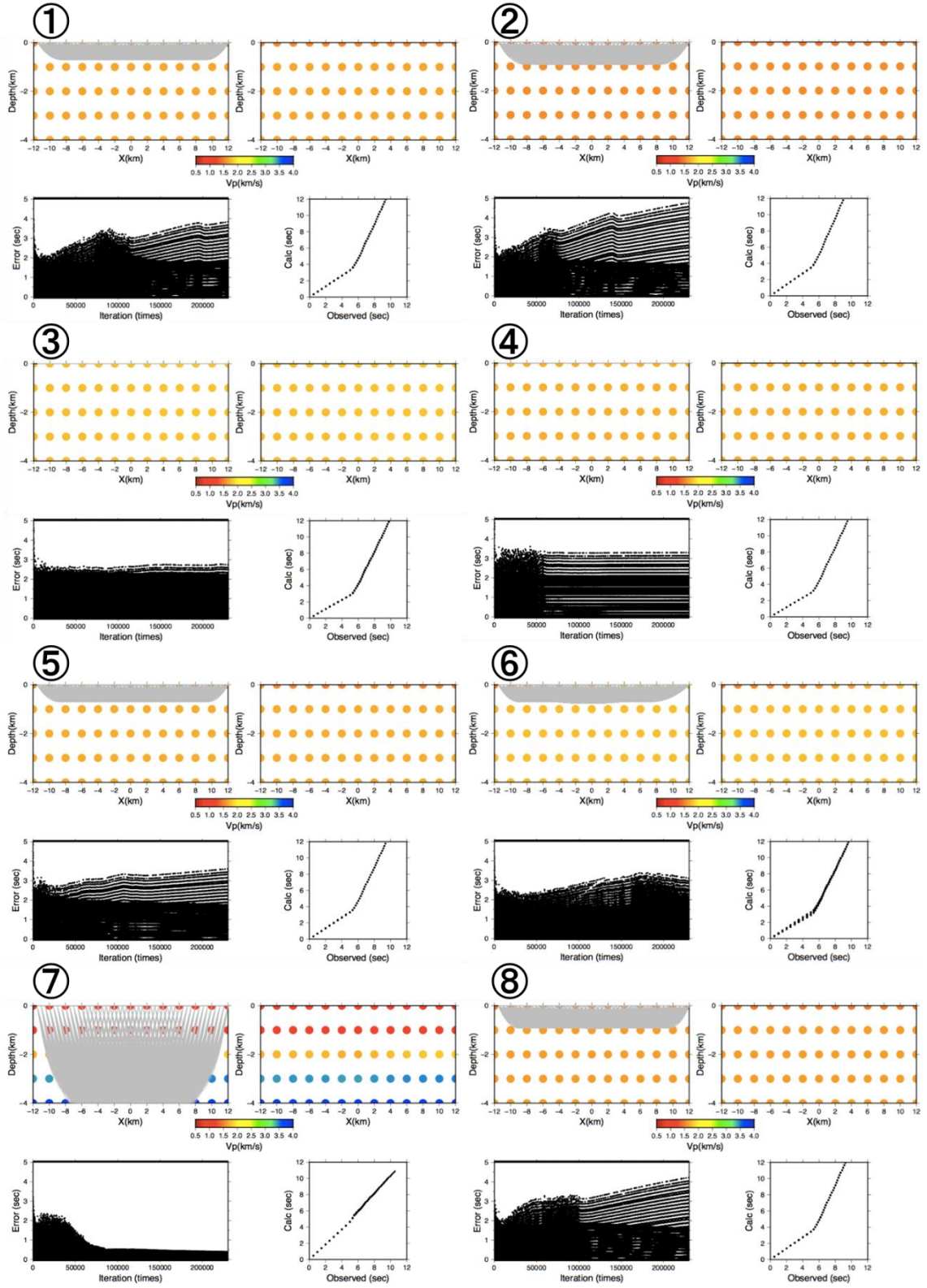
図 4.9 学習回数 10 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.7 学習回数 10 回の平均誤差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	RMS残差[sec]
1	1.13	1.06
2	1.09	1.38
3	0.71	0.83
4	0.78	0.94
5	0.68	0.72
6	0.84	1.02
7	1.05	1.38
8	0.95	1.18
9	0.85	0.27
10	0.81	0.99
平均	0.89	0.98

	$E_{vp}[km/sec]$	RMS残差[sec]
最小	0.68	0.27
最大	1.13	1.38
平均	0.89	0.98

(1)の学習回数 10 回のときは全てのパターンで誤差が収束しない、もしくは、わずかに収束するのみであり、平均誤差の平均が 1.01、RMS 残差の平均が 1.27 であった。対して、⑨のように誤差が収束しているものも見られ、平均誤差の平均が 0.89、RMS 残差の平均が 0.98 となり結果は向上した。しかし、図 4.9 の速度構造を見ると、正解モデルが再現できているとは言い難い。 $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布を用いたときの学習回数 10 回のときに比べて誤差の収束が進んでいるため、同じ重みを用いて学習回数 100 回でも同様に検証した。図 4.10、表 4.8 にその結果を示す。



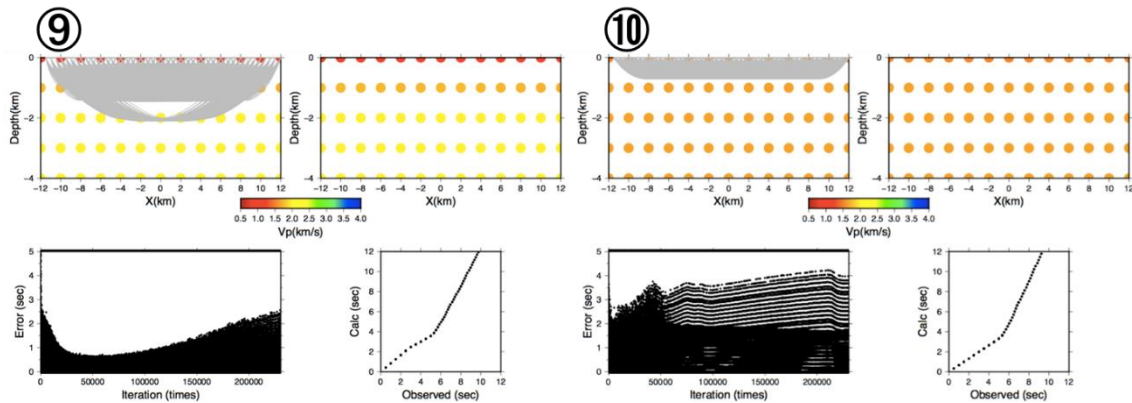


図 4.10 学習回数 100 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.8 学習回数 100 回の平均誤差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
1	1.04	1.26
2	1.07	1.41
3	1.05	1.37
4	1.06	1.34
5	1.03	1.23
6	1.01	1.25
7	0.24	0.18
8	1.05	1.29
9	0.79	0.94
10	1.05	1.28
平均	0.94	1.15

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
最小	0.24	0.18
最大	1.07	1.41
平均	0.94	1.15

⑦のサンプルだけは概ね正解値を推定できているが、他のサンプルは明らかに誤差が発散している。 $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布の重みの標準偏差 0.180、0.5 からそれぞれ標準偏差を約 1.4 倍したが、誤差が発散した。学習回数が 1000 回、2000 回の場合も同様に精度は悪かった。なお、学習回数 1000 回、2000 回の結果は付録に添付する。

#### 4.3.3 伝達関数を変更したときの学習回数比較

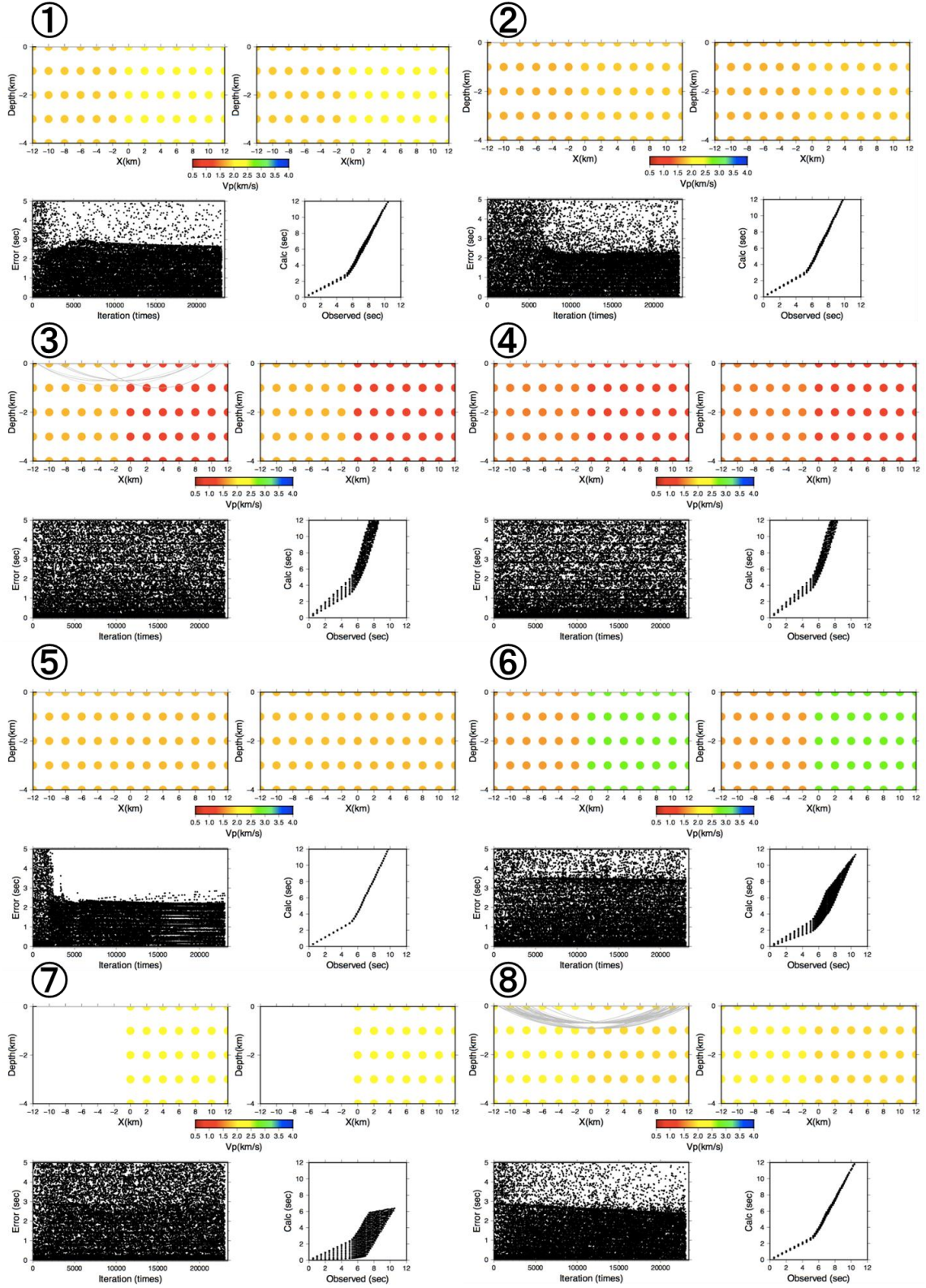
本研究では伝達関数として、シグモイド関数のほかに、ReLU 関数を用いた。本研究で

ReLU 関数を用いる際の出力層への伝達関数は出力値がスローネスを表し、負の数や 0 が出力されることは許されていないため、シグモイド関数を利用した。また、本研究では ReLU 関数からの出力上限値を 10 とし、シグモイド関数への入力は-3~3 に制限した。

(1) 一様乱数を用いたとき

図 4.11、表 4.9 に解析結果を示す。







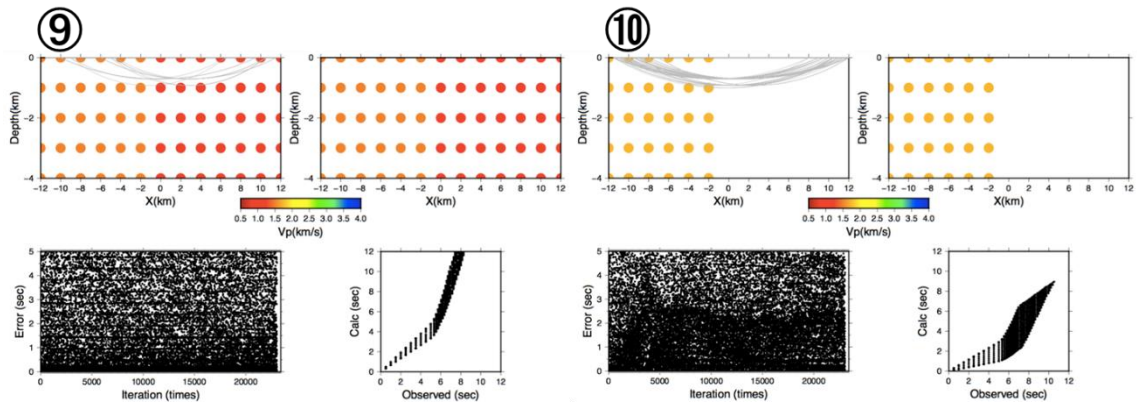


図 4.11 学習回数 10 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.9 学習回数 10 回の平均誤差 $E_{vp}$ と RMS 残差

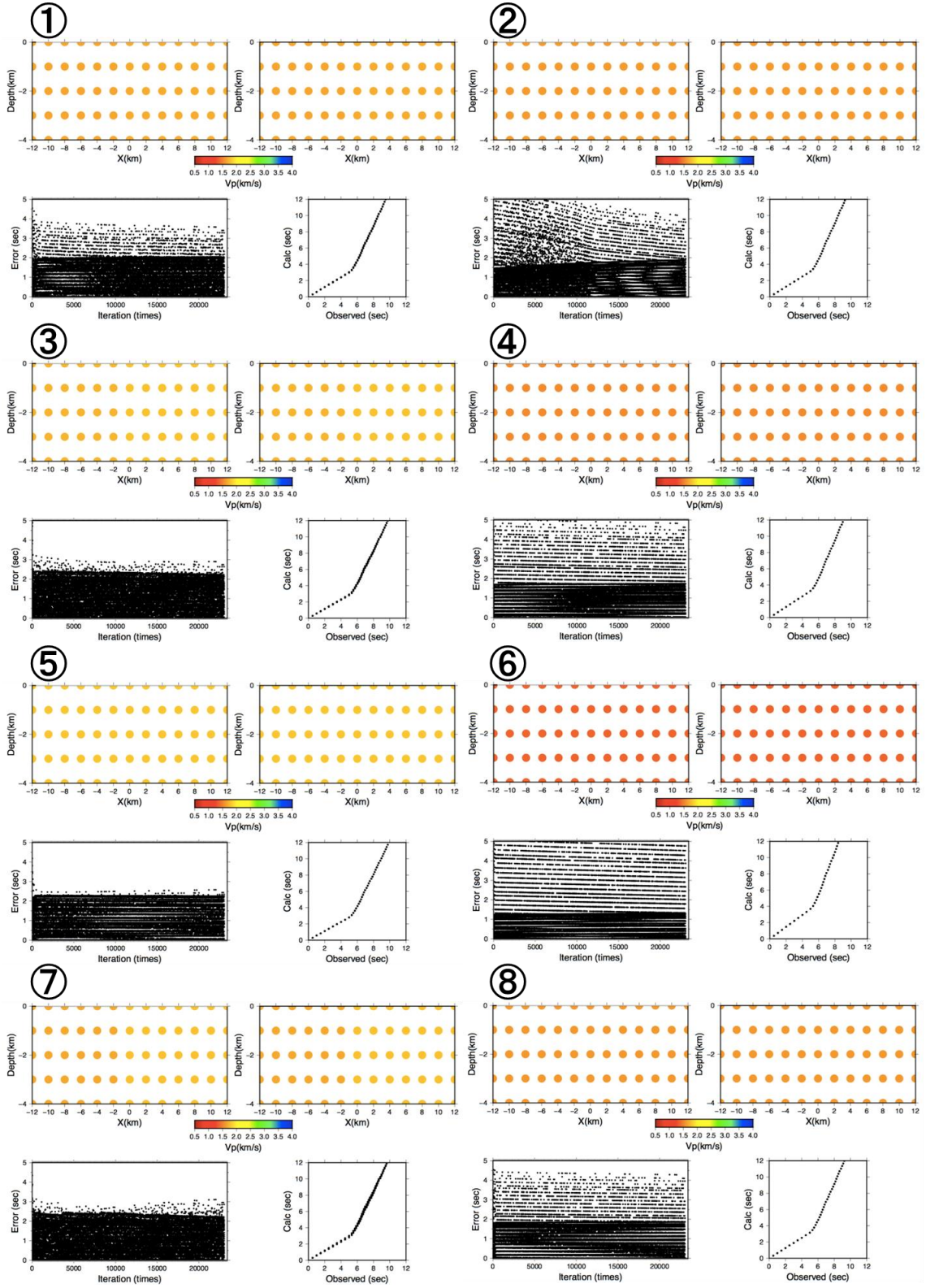
	$E_{vp}[km/sec]$	RMS残差[sec]
1	1.02	1.54
2	1.05	1.36
3	1.13	2.47
4	1.15	2.73
5	1.05	1.36
6	1.12	1.87
7	9.27	3.55
8	1.01	1.55
9	1.15	2.73
10	2.01	2.68
平均	2.00	2.18

	$E_{vp}[km/sec]$	RMS残差[sec]
最小	1.01	1.36
最大	9.27	3.55
平均	2.00	2.18

誤差がわずかに収束しているものも散見されたが、地下構造を推定できているとはいえなかった。

(2)  $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布を用いたとき

図 4.12、表 4.10 に解析結果を示す。



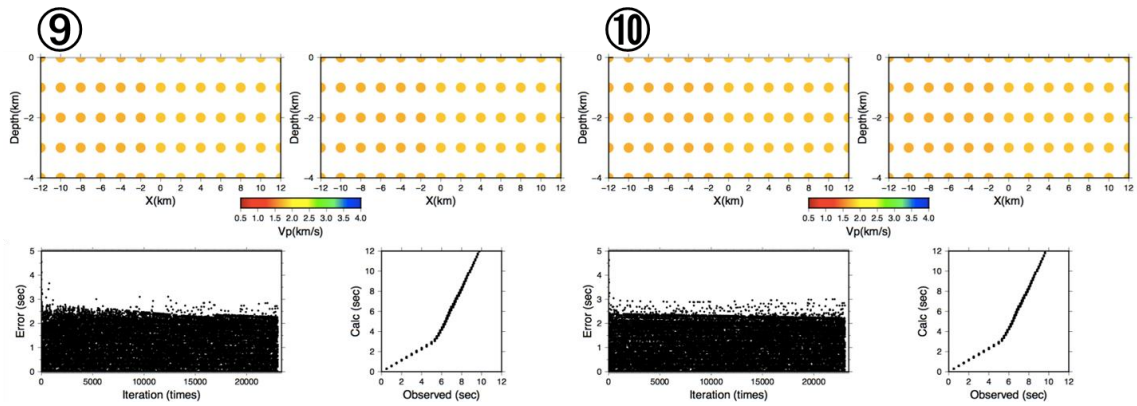


図 4.12 学習回数 10 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.10 学習回数 10 回の平均誤差 $E_{v_p}$ と RMS 残差

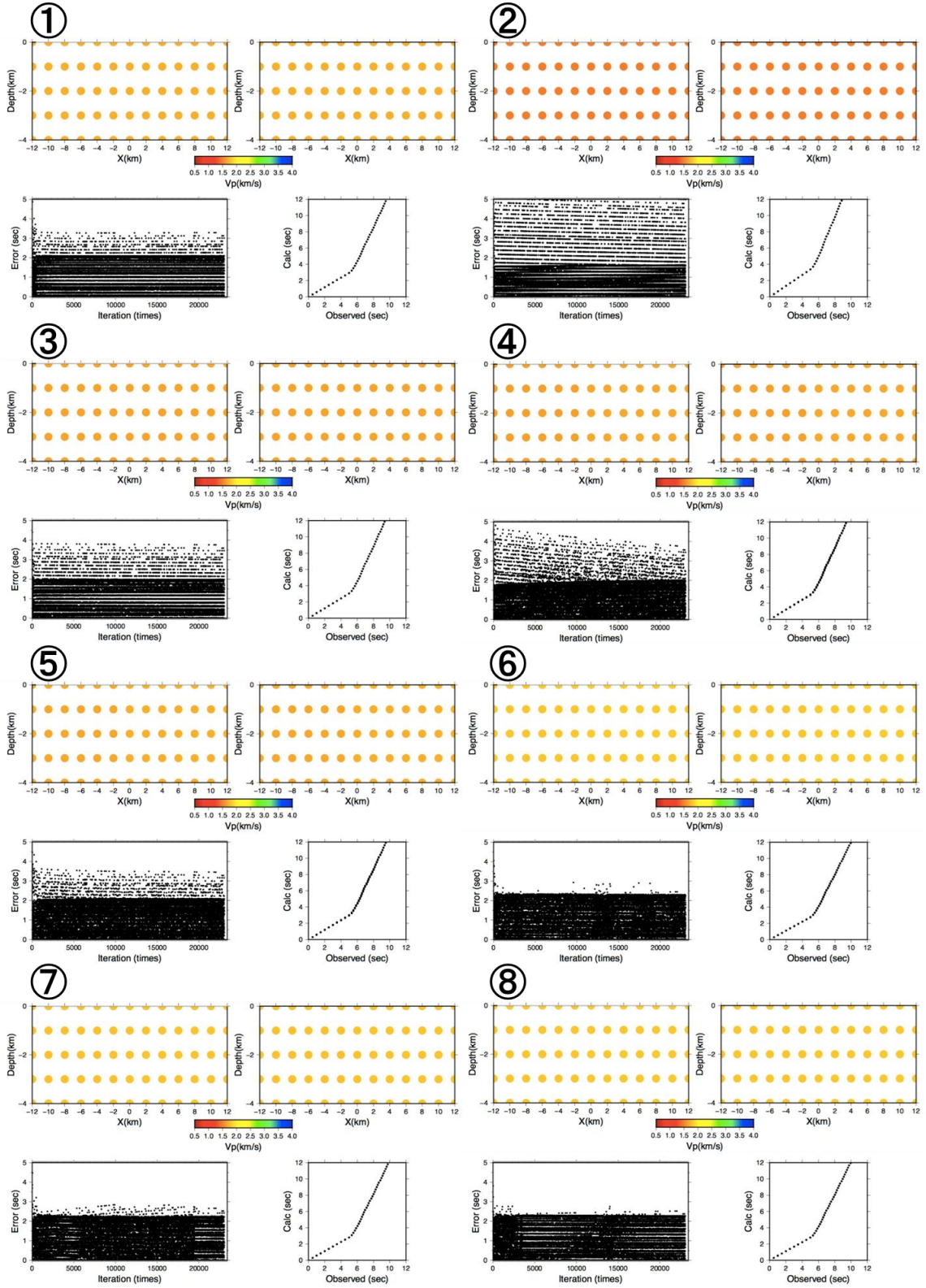
	$E_{v_p}[km/sec]$	RMS残差[sec]
1	1.07	1.34
2	1.08	1.38
3	1.05	1.36
4	1.09	1.47
5	1.04	1.38
6	1.13	1.96
7	1.05	1.35
8	1.08	1.39
9	1.05	1.36
10	1.05	1.35
平均	1.07	1.43

	$E_{v_p}[km/sec]$	RMS残差[sec]
最小	1.04	1.34
最大	1.09	1.96
平均	1.07	1.43

重みに一様乱数を与えた場合と比べても誤差の収束状況は悪化した。そして、求められた地下構造も均質構造に近い構造となっていた。

### (3) $\sqrt{\frac{2}{n}}$ の標準偏差を持つガウス分布を用いたとき

図 4.13、表 4.11 に解析結果を示す。





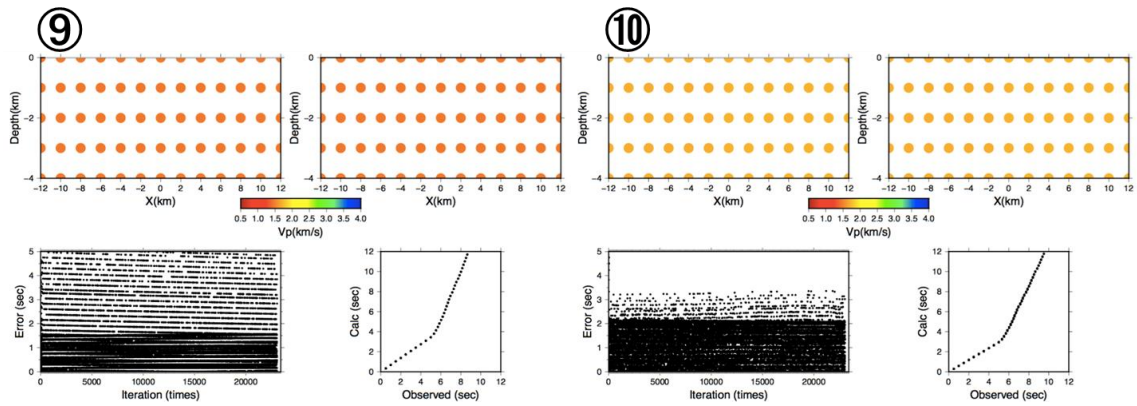


図 4.13 学習回数 10 回の 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.11 学習回数 10 回の平均誤差 $E_{v_p}$ と RMS 残差

	$E_{v_p}[km/sec]$	RMS残差[sec]
1	1.06	1.34
2	1.10	1.58
3	1.07	1.35
4	1.07	1.35
5	1.06	1.34
6	1.04	1.40
7	1.04	1.37
8	1.04	1.39
9	1.11	1.65
10	1.06	1.34
平均	1.07	1.41

	$E_{v_p}[km/sec]$	RMS残差[sec]
最小	1.04	1.34
最大	1.11	1.65
平均	1.07	1.41

$\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス分布を用いたときと同様に誤差が収束せず、地下構造も均質構造に近いものとなった。

#### 4.3.4 入力の変換を導入したとき

2.7 で述べたようにニューラルネットワークには、入力の値が近いと出力の値も近くなる傾向があり、不連続な地下構造を推定する上では深刻な問題となる。そこで本研究では、入力について改良を試みた。以下にその手法を述べる。

### (1) パターンコーディングの導入

パターンコーディングは、予測点の空間座標値をあるルールに則って、2 値パターンに変換する手法である。本研究で用いている多層パーセプトロンに適用することで関数近似器として高い性能が得られていることが知られている。(丹野ほか、2013) <sup>11)</sup> パターンコーディングの具体的な手法は、野中ほか (2011) <sup>12)</sup> によると以下の通りである。

(1) 変数値に関する情報がパターン全体に広く分散し、かつ全体として十分な冗長性をもつこと。これは分散表現の定義のようなものであるが、分かりやすいというならば、パターンの成分のうち、どの一つまたは数個を見ても変数値が特定できないが、全体の半分程度を見れば必ず特定できるということである。

(2) 変数値が連続的に変化するにつれてコードパターンが徐々に変化すること。また、変数値が近いほどコードパターン間の相関が高く、遠いほど低くなること。これによって、非局所的な汎化が生じることが期待できる。

(3) 1 と -1 の成分が常にほぼ同数であること。これにより、第 3 層において常に約半数の素子が不感化されて 0 を出力することになる。

(4) 十分に離れた変数値を表すコードパターン間の相関が 0 であること。あるサンプルを学習したとき、無関係な点への影響（干渉）が小さくなる。

(1)、(2)に関しては確実に満たす必要があり、(3)、(4)が満たしていることが望ましい。図 4.14 で示されているのはパターンコーディングの様子を図示化したものである。上記の方法で  $P_1$  のパターンを決定した後、いくつかの素子の符号を逆転させて  $P_2$  のパターンとし、同様の作業を  $P_3$ 、 $P_Q$  と行う。また図 4.15 にはパターンコーディングを適用した多層パーセプトロンを示す。

本研究では、伝達関数にシグモイド関数、重みの初期値に一様乱数、そして、学習回数を 10 回で解析を行いその結果を図 4.16、表 4.12 に示す。

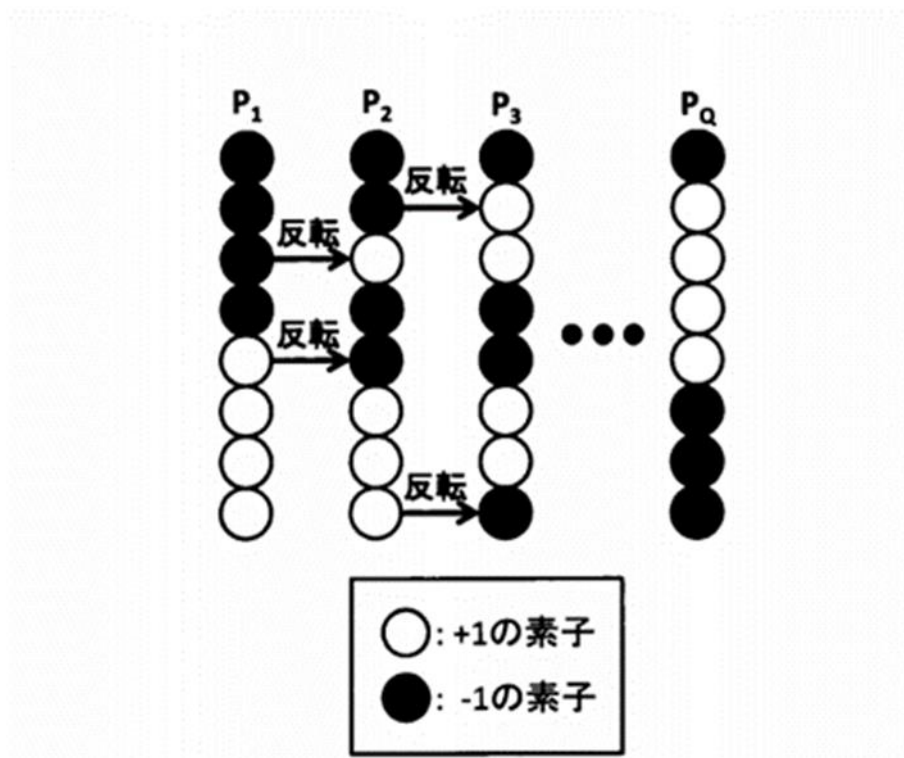


図 4.14 パターンコーディング例(丹野ほか、2013)<sup>13)</sup>

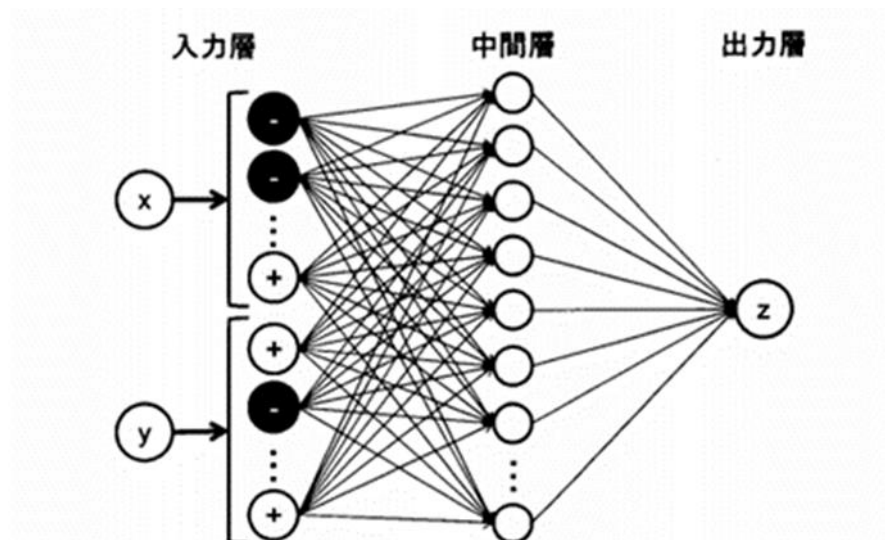


図 4.15 パターンコーディングを適用した多層パーセプトロン(丹野ほか、2013)<sup>11)</sup>

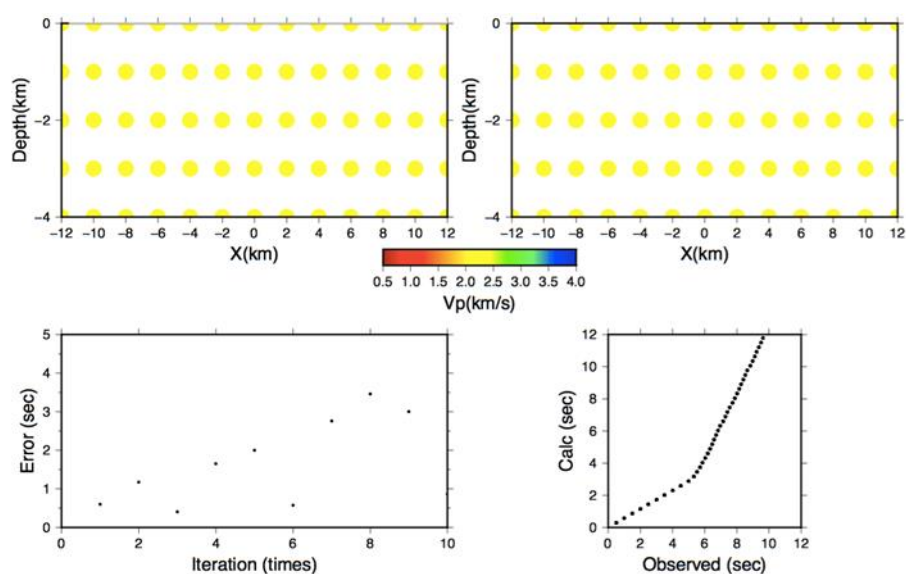


図 4.16 学習回数 10 回の解析結果（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.12 学習回数 10 回の平均誤差 $E_{v_p}$ と RMS 残差

$E_{v_p}[km/sec]$	RMS残差[sec]
1.03	1.35

パターンコーディングを導入して解析を試みたが、その出力値は均質構造となり、解析領域の浅部にしか波線が通らなかった。誤差の学習（図 4.16 下段左）に関しては、本来であれば 23040 回誤差の計算がされるところが、10 回のみの計算となり、その 10 回分を見ても誤差は発散に向かっている。

#### 4.4 断層構造での結果

ここでは、地下構造が不連続な場合として、図 4.17、図 4.18 に示すような断層構造を対象に検討を行った。これまでの検討の結果、最も精度の良かった、伝達関数にシグモイド関数、重みの初期値に範囲-0.5～0.5 の一様乱数、空間座標値入力、そして、学習回数 1000 回で解析を行った。



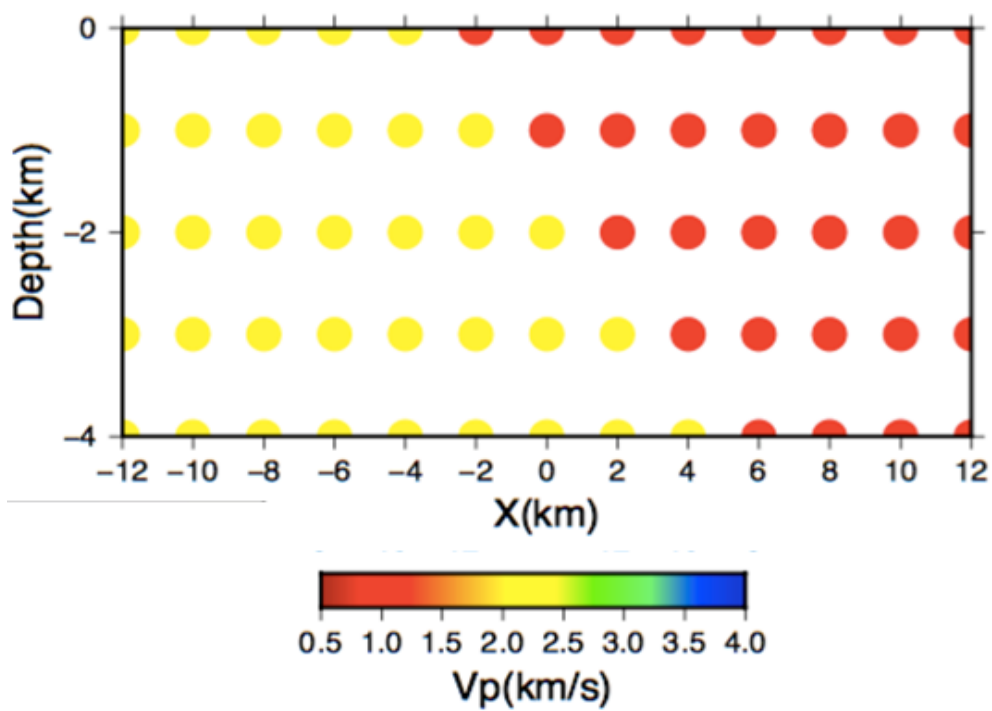


図 4.17 解析領域の中央に断層がある構造

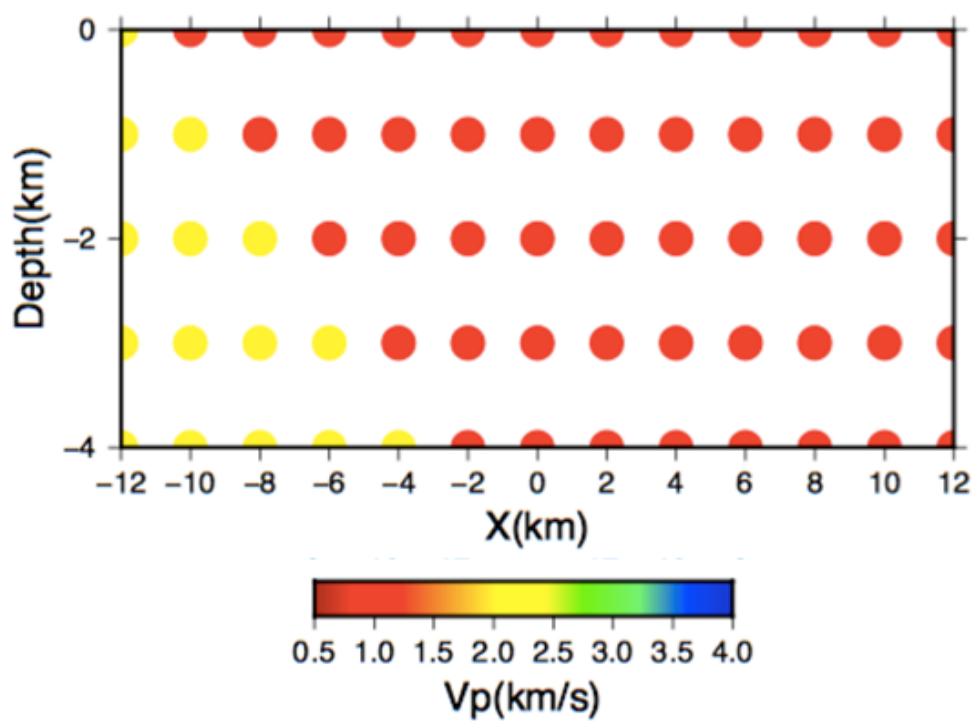
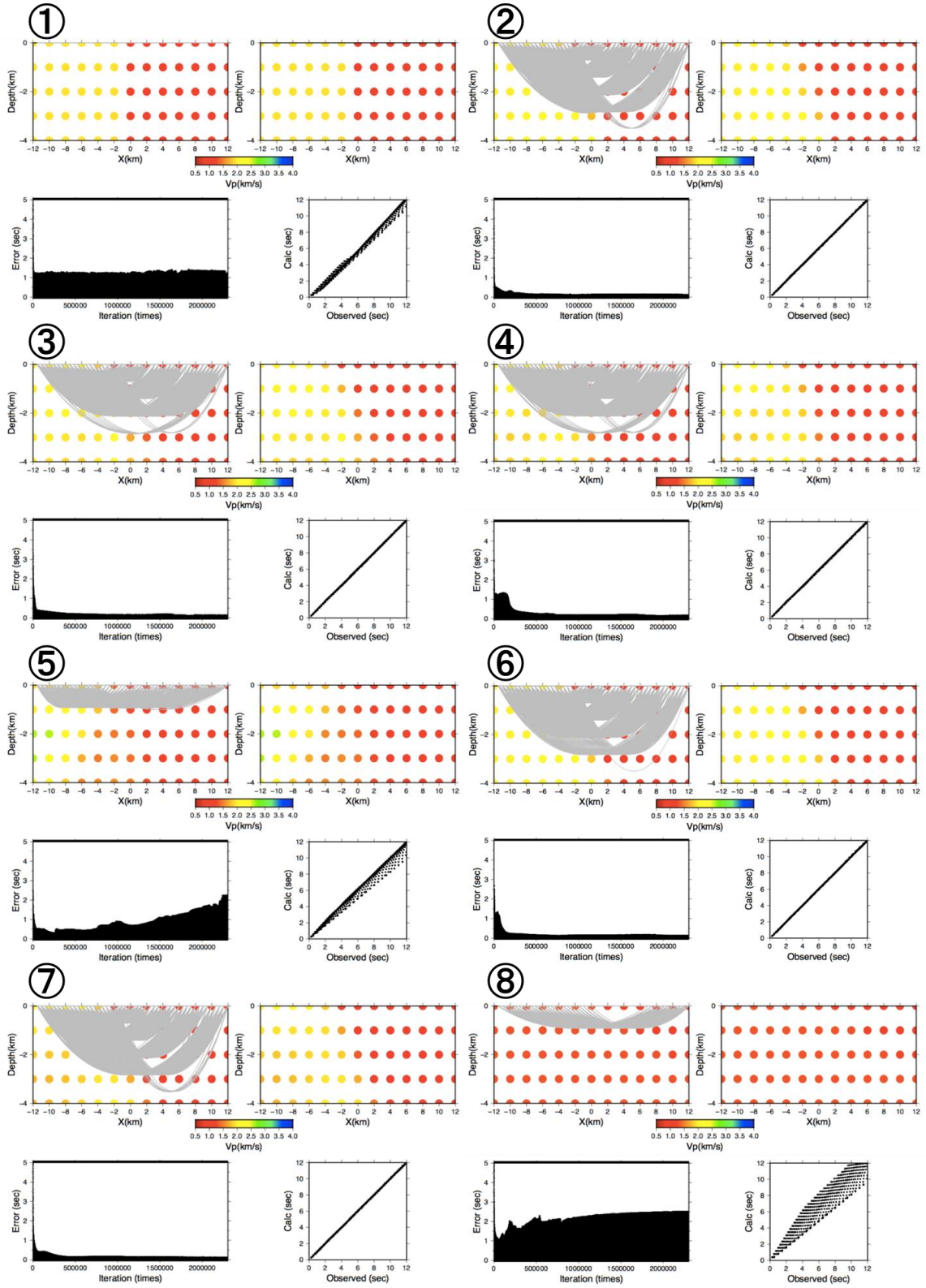


図 4.18 解析領域の端部に断層がある構造

- (1) 断層構造が解析領域の中央にある場合

断層構造が解析領域の中央にある場合の結果を図 4.19、表 4.13 に示す。



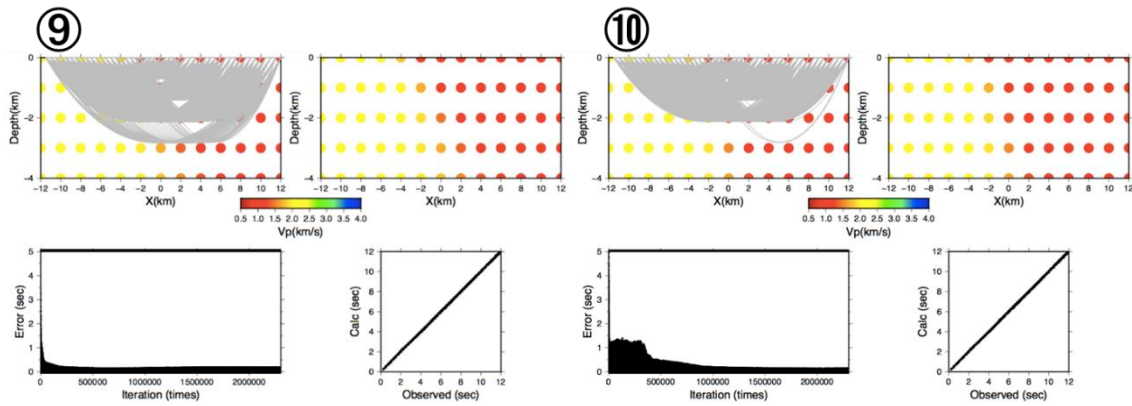


図 4.19 学習回数 1000 回での 10 サンプル分データ（上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関）

表 4.13 学習回数 1000 回での断層構造の平均誤差 $E_{vp}$ と RMS 残差

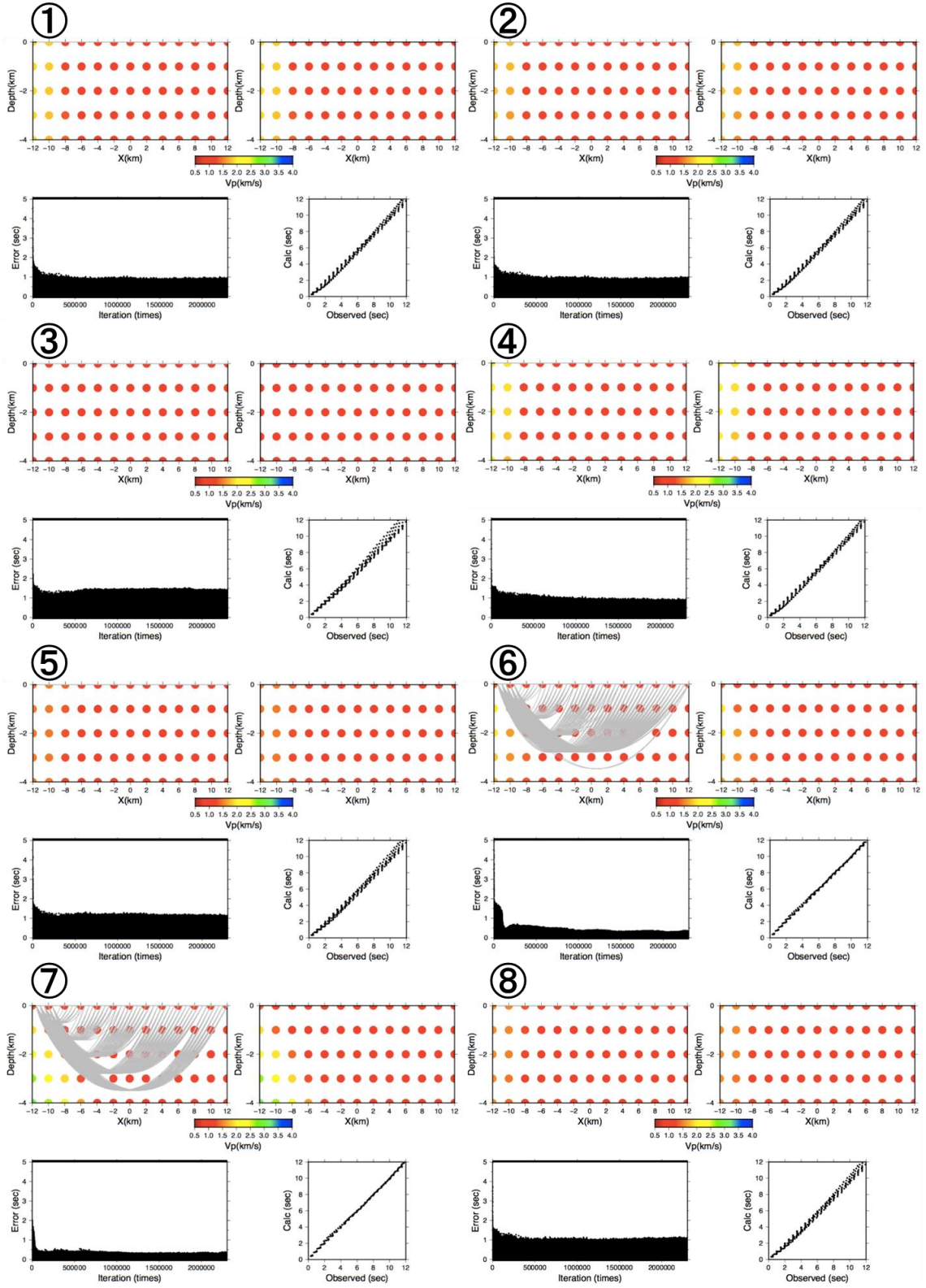
	$E_{vp}[km/sec]$	RMS残差[sec]
1	0.17	0.40
2	0.09	0.03
3	0.08	0.04
4	0.11	0.05
5	0.29	0.62
6	0.08	0.03
7	0.12	0.03
8	0.50	1.12
9	0.07	0.04
10	0.09	0.04
平均	0.16	0.24

	$E_{vp}[km/sec]$	RMS残差[sec]
最小	0.07	0.03
最大	0.50	1.12
平均	0.16	0.24

平均誤差、RMS 残差共に良好な結果を示した。実際の推定された地下構造を見ると、断層が明確に表れていた。また、数は少ないが誤差が発散してしまう場合もあった。

## (2) 断層が解析領域の端部にある場合

断層構造が解析領域の端部にある場合の結果を図 4.20、表 4.14 に示す。



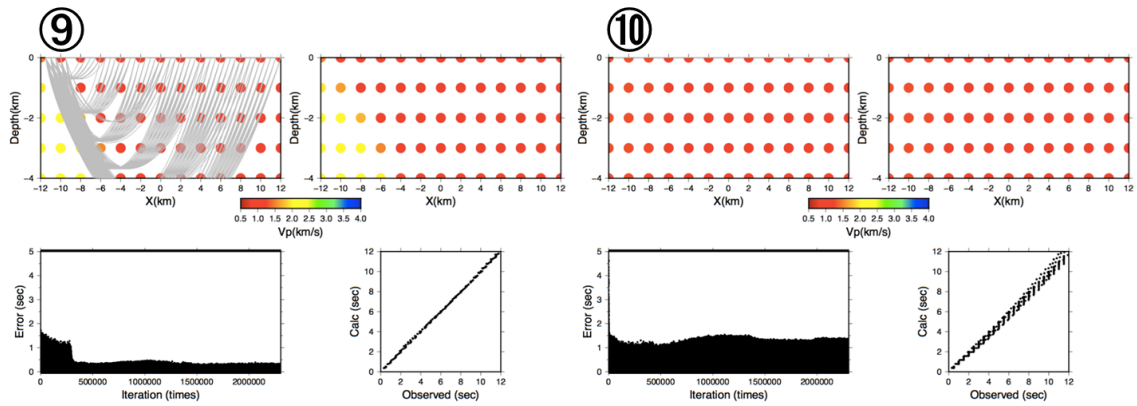


図 4.20 学習回数 1000 回での 10 サンプル分データ 上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差の収束状況、下段右：理論走時と観測走時間の相関)

表 4.14 学習回数 1000 回での断層構造の平均誤差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	RMS残差[sec]
1	0.13	0.28
2	0.15	0.29
3	0.21	0.37
4	0.13	0.28
5	0.19	0.32
6	0.13	0.09
7	0.12	0.08
8	0.18	0.31
9	0.09	0.06
10	0.20	0.35
平均	0.15	0.24

	$E_{vp}[km/sec]$	RMS残差[sec]
最小	0.09	0.06
最大	0.21	0.37
平均	0.15	0.24

平均誤差、RMS 残差共に解析領域の中央に断層がある場合と概ね同じ値であった。しかし、解析領域の中央に断層がある場合と比べると、波線が解析領域を通過しているサンプルが数少なく、断層が明確に表れていなかった。また、走時誤差に関しても 0 秒に近づくものが目に見えて少なかった。

#### 4.5 二進数の導入

本研究ではパターンコーディングを用いて入力についての検討を行ったが、誤差の学習面で解析が上手く進まなかった。そこで、新たな入力についての検討として、入力の空間座



標値を二進数に変換して一桁ずつユニットの入力とする方法を提案する。

本研究で適用する場合以下の手法となる。

空間座標値を 2 進数に変換する際、10 進数の小数から変換するより、10 進数の整数から変換する方が容易である。本研究で与えられている解析領域の推定点座標は、小数第 4 位まで使用している。この場合は 10000 倍して、6 桁の整数として 2 進数への変換を行う。しかし、6 桁の整数を 2 進数変換すると、入力層のユニット数が多くなってしまう。必要最小限の桁数を確認するために、1 桁ずつ削り、それぞれの桁で波線が滑らかに描けているかを検証した (図 4.21)。震源は解析領域の端部の点で、観測点は震源から最も近傍の点、解析領域中心付近の点、そして、解析領域逆側の端部の点で検証した。その結果、解析領域中心、逆側端部においては桁数を減らして波線を描写してもほとんど変化は見られなかったが、最も近傍の点では、1 桁減らすことで波線の形そのものが変化してしまうため、本研究では、6 桁の整数を 2 進数へ変換することとした。解析領域内の推定点で 2 進数表記した際に、最大の桁数となる座標は、-11.5125 である。この座標値を 2 進数表記すると、32 桁となるため、入力のユニット数は  $x$ 、 $y$ 、 $z$  座標それぞれで 32 個ずつとする。2 進数表記した際に 32 桁に満たないものは、あまったユニットに 0 を入力する。

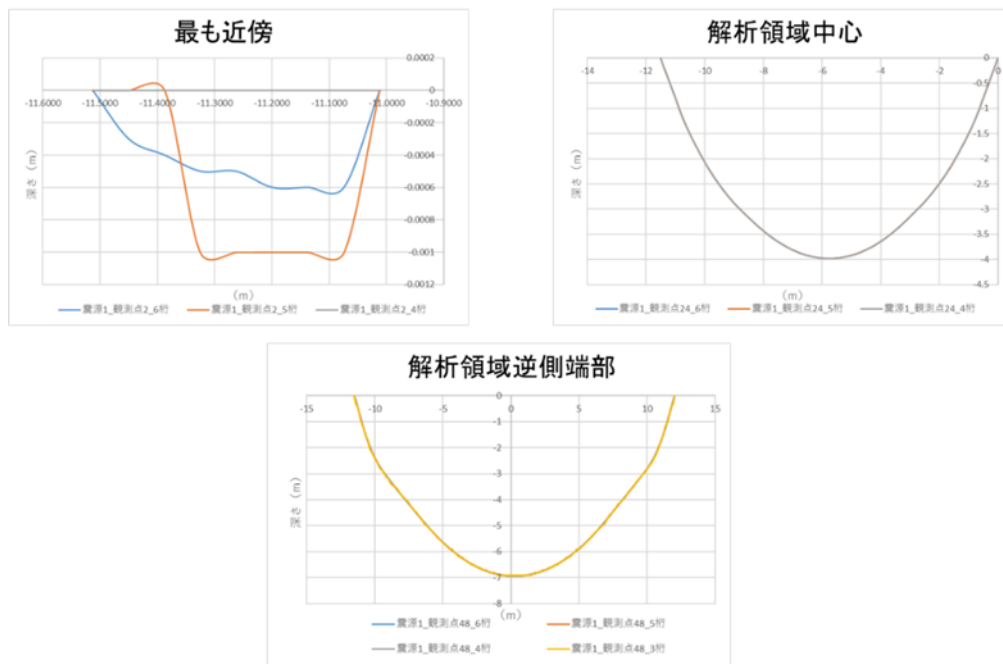


図 4.21 波線比較

#### 4.6 考察

既往の研究から学習回数、重みの初期値、伝達関数、そして入力の与え方を変更して解析精度の向上を試みた。各パラメータを変更したときの結果から考察をまとめる。

概ね適切な学習回数を求めることができたがその範囲は 100~1000 回と幅広く、さらに

細かく区切り判断していくことが求められる。平均誤差や RMS 残差を比較すると平均誤差は学習回数 100 回のほうが優秀で、RMS 残差は学習回数 1000 回のほうが優秀な値を示した。これは、学習回数 100 回程度であると、過学習に陥ることは少なく、間違った方向に学習を進めていたとしても局所解に落ちきることがないまま学習が終わる。一方、学習回数が 1000 回であると、正しい学習が行えている場合は学習回数が 100 回のときに比べさらに良好な結果を示し、間違った学習を進めている場合はさらに結果が悪化する。本研究では同じ学習回数で 10 サンプルとったが、このサンプル数では何パーセントの割合で間違った学習を進めているのかを把握できない。そのため、さらに細かい区切りでの学習回数の設定と、さらに多くのサンプルをとることが必要であると考ええる。

重みの初期値を変更した際には、ほとんどのサンプルで誤差が収束しない現象が起こった。重みの初期値を変更した理由は、偏った重みを持っていると勾配消失に陥る可能性がある、この問題を解消するために重みの分布に適度な広がりを持たせる必要があるためである。標準偏差が 0.01 になると学習が全く行われれないという問題はあるが、本研究での標準偏差は 0.01 から十分に離れておりこの問題とは無関係だと考えられる。従来ニューラルネットワークの入力には 0~1 の値が用いられるが、本研究においては入力に空間座標値を用いており、その範囲は-12~12 となる。この問題を解消するためには、入力に用いられている空間座標値を 0~1 に正規化することが効果的だと考えられる。

伝達関数の変更においては、ReLU 関数を用いた場合、図 4.11 で示している結果が特に解析領域の中央で弾性波速度が分断される出力であった。この結果から言えることは、ニューラルネットワークが順伝播する際に、入力層への入力値が正か負かが決まるだけでそれぞれ決まった値に陥り、学習を繰り返してもその局所解から抜け出すことができなかったということである。本研究では、伝達関数に ReLU 関数を用いた際、重みの初期値として一様乱数を用いたことで重みパラメータに偏りが生まれ局所解に陥ったと見られる。上記の重みの問題が解決できなかったため重みパラメータに一様乱数を用いたが、空間座標値を正規化するなどの対策をとり、課題を解消した上で ReLU 関数に適した重みパラメータを与えることで改善できると考えられる。

本研究の結果を比べる中で、解析精度の悪いサンプルの理論走時と観測走時の相関に着目すると走時が 6 秒を超える辺りから相関が悪くなる現象が見られる。これは、今回用いた水平成層モデルにおける初動走時が直接波から屈折波に変わる走時が 6 秒であると考えられる。解析精度の悪いサンプルの解析精度を向上させるために、観測走時データの走時が 6 秒以内のものと 6 秒以上のものに分けて二段階で解析を行うことが有効であると考えられる。

## 第5章 結論

### 5.1 まとめ

本研究ではニューラルネットワークを用いた走時トモグラフィの改良を試みた。様々なパラメータを変更し数値解析を行った結果、伝達関数にシグモイド関数、重みの初期値には一様乱数、入力には空間座標値を用いたもので、学習回数は100～1000回の際の解析精度が良好であった。しかし、解析結果には地下構造を良く評価できているものと良くないもののばらつきが激しかった。本研究では10サンプルでの結果であったが、今後はさらにサンプル数を取ることで安定性が増すことが予想される。また、解析精度が最も良好であったパラメータを用い、正解モデルを成層構造から断層構造に変更し数値解析を行った結果、平均誤差が大幅に修正された。ただ、本研究では断層と考えるものを斜線で表したが、結果は垂直線で速度構造を分断している形になった。これは本研究の解析は波線が通っている範囲で速度構造を推定できるため、断層構造の地表に近い箇所は推定できているが、深い箇所は推定できていると言えない。

今後は、解析領域内に波線を通す方法を確立するとともに、4.6で述べた本研究で対応できなかった問題について解決することが課題となる。



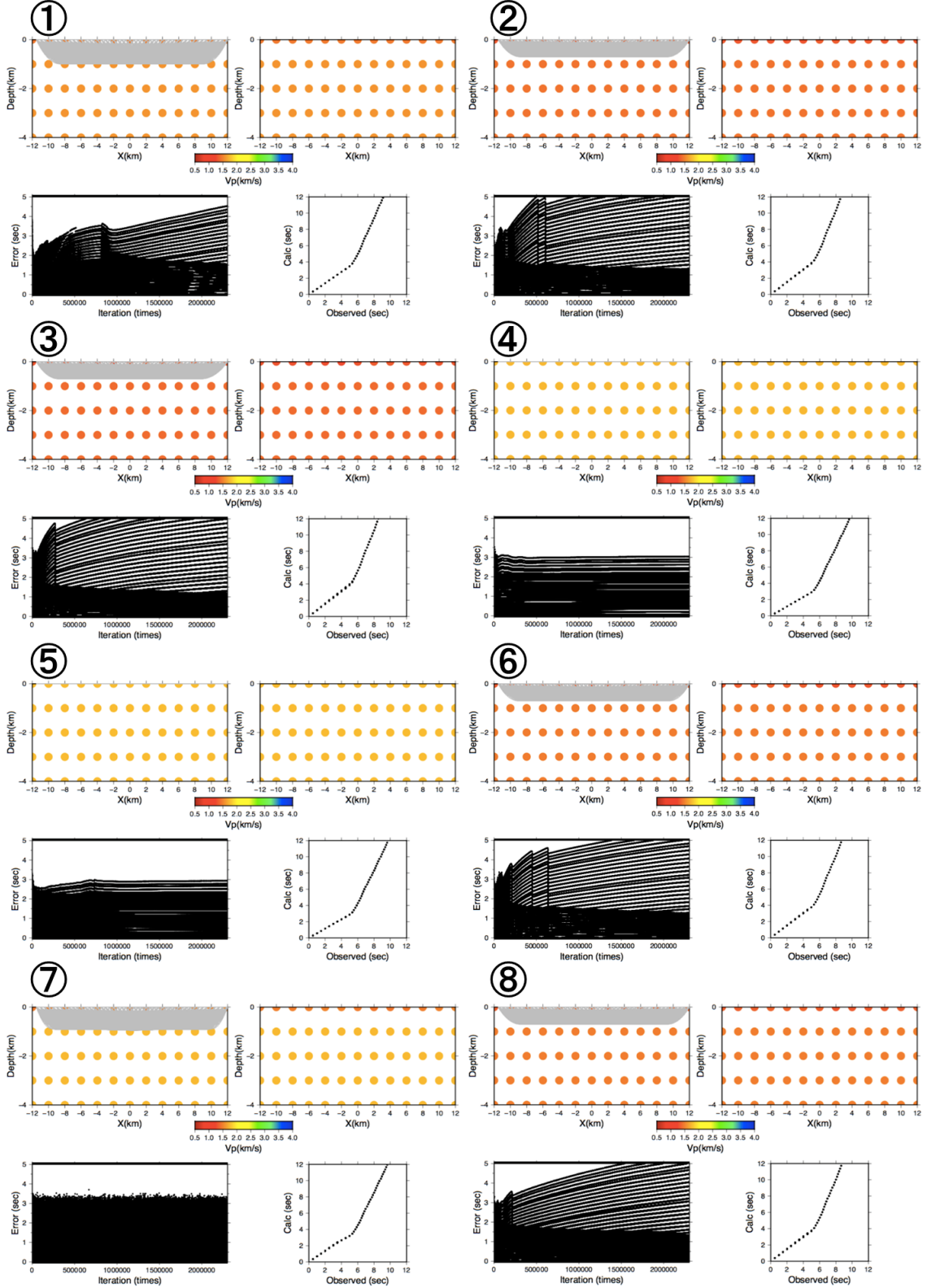
## 謝辞

本論文をまとめるにあたり、ご指導頂きました小田義也准教授、吉嶺充俊准教授、深田地質研究所松岡俊文様、磯真一郎様に深く感謝の意を表します。また、小田義也准教授には特に熱心なご指導を頂き、深く感謝申し上げます。さらに、同研究室の学生諸君の惜しみない協力には非常に助けられました。ありがとうございます。

最後に、ご協力頂きました全ての方々に感謝の意を表し、謝辞とさせていただきます。

2019 年 2 月 金本龍之介

# 付録



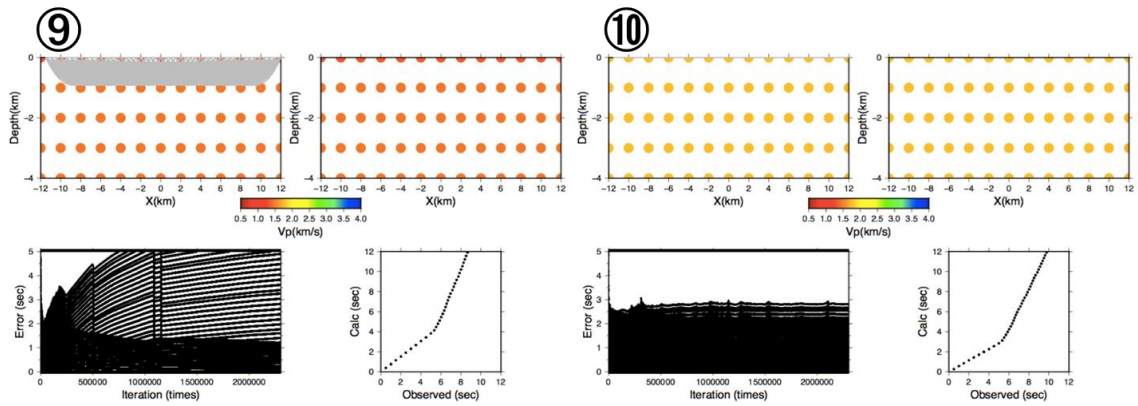
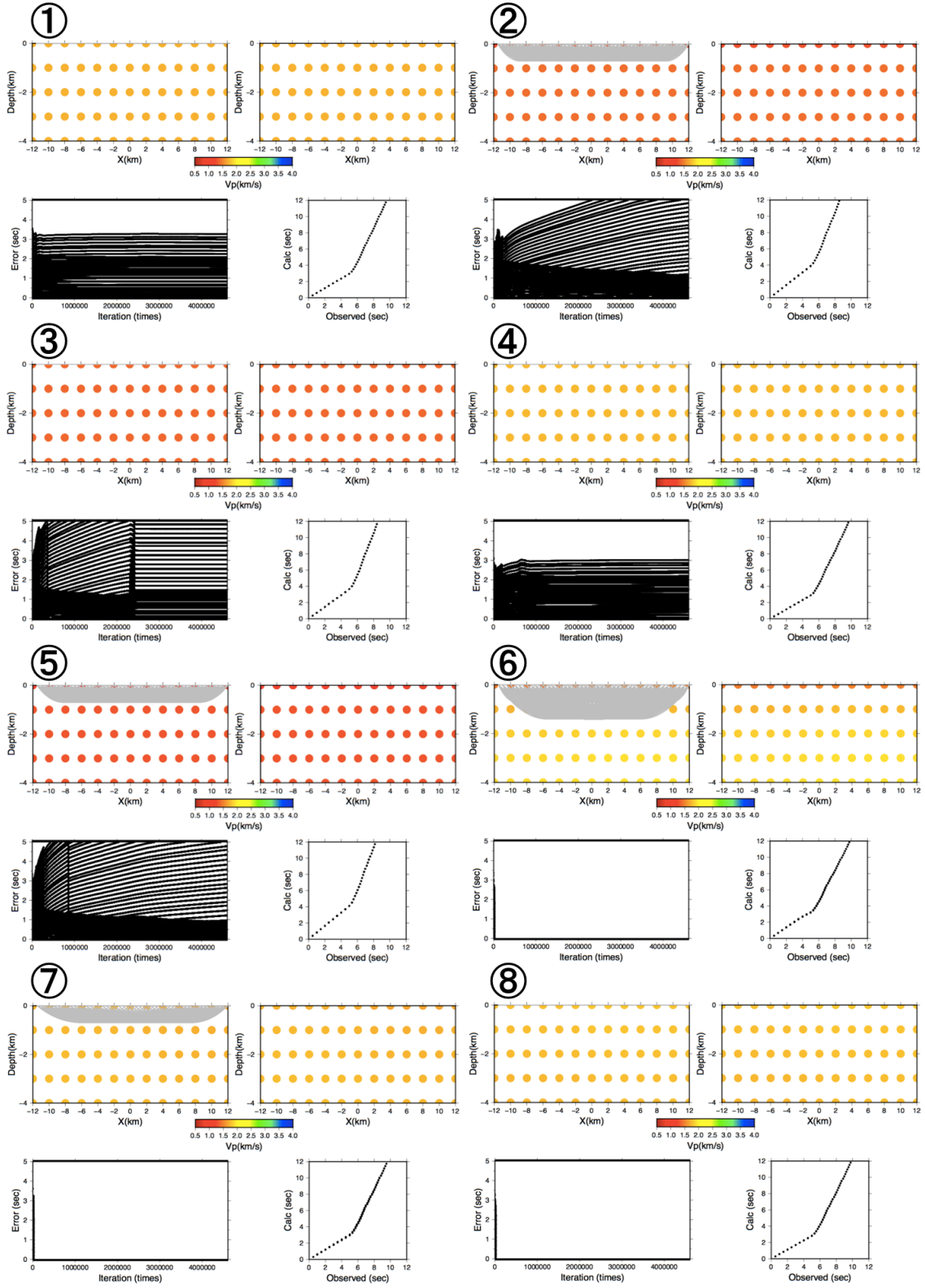


図 重みの標準偏差に $\sqrt{1/N}$ の標準偏差を用いた学習回数 1000 回の 10 サンプルデータ  
(上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差  
の収束状況、下段右：理論走時と観測走時間の相関)

表 重みの標準偏差に $\sqrt{1/N}$ の標準偏差を用いた学習回数 1000 回での断層構造の平均誤  
差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	RMS残差[sec]
1	1.06	1.36
2	1.09	1.73
3	1.10	1.86
4	1.05	1.35
5	1.05	1.35
6	1.08	1.59
7	1.00	1.20
8	1.08	1.59
9	1.08	1.66
10	1.05	1.36
平均	1.06	1.51

	$E_{vp}[km/sec]$	RMS残差[sec]
最小	1.00	1.20
最大	1.10	1.86
平均	1.06	1.51



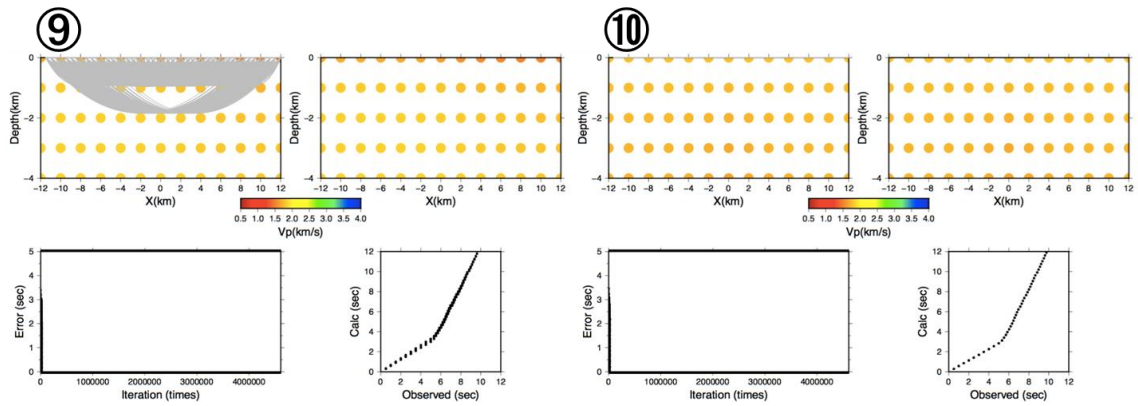
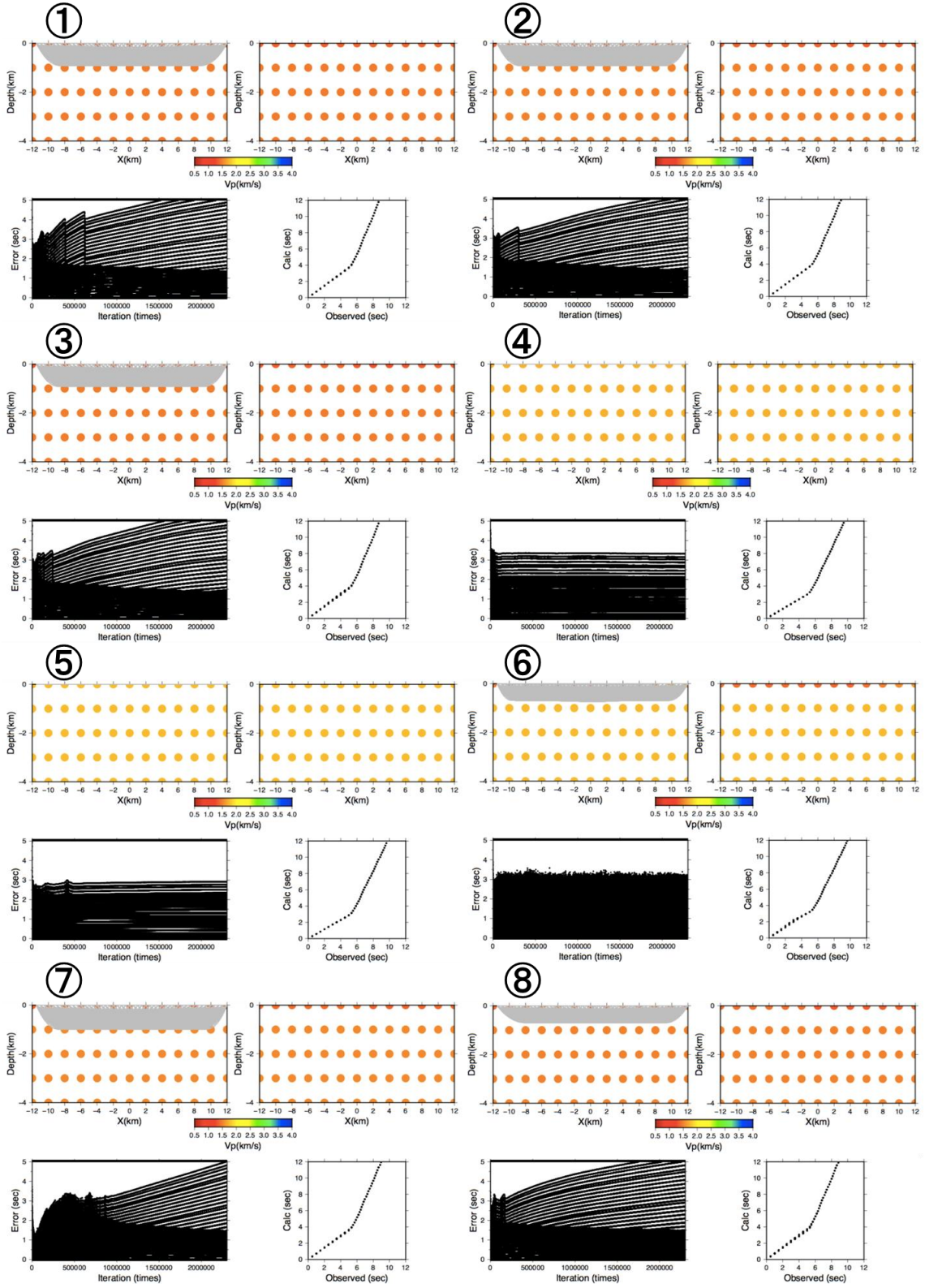


図 重みの標準偏差に $\sqrt{1/N}$ の標準偏差を用いた学習回数 2000 回の 10 サンプルデータ  
(上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差  
の収束状況、下段右：理論走時と観測走時間の相関)

表 重みの標準偏差に $\sqrt{1/N}$ の標準偏差を用いた学習回数 2000 回での断層構造の平均誤  
差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
1	1.06	1.34
2	1.09	1.80
3	1.13	1.91
4	1.05	1.35
5	1.12	2.29
6	0.91	1.14
7	1.03	1.30
8	1.06	1.37
9	0.98	1.23
10	1.07	1.36
平均	1.05	1.51

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
最小	0.91	1.14
最大	1.13	2.29
平均	1.05	1.51





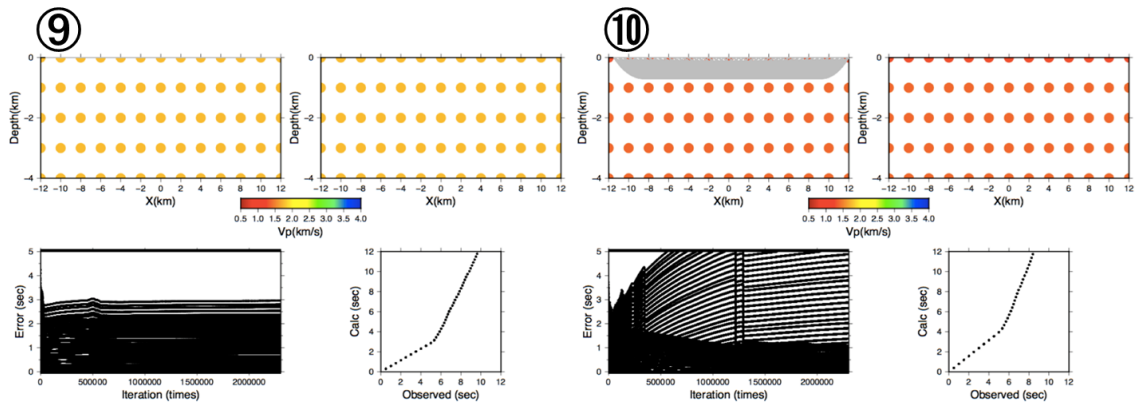
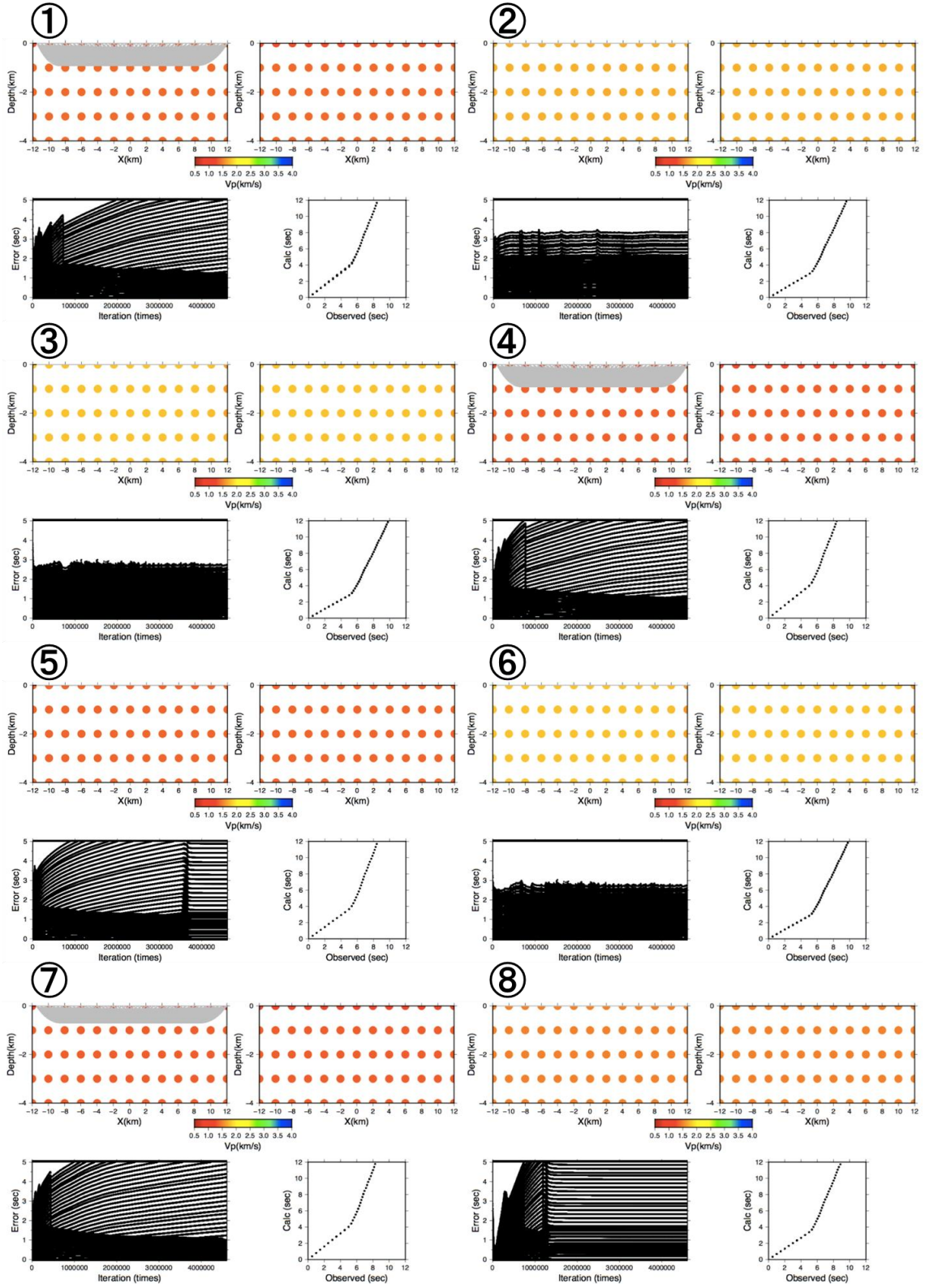


図 重みの標準偏差に $\sqrt{(2/N)}$ の標準偏差を用いた学習回数 1000 回の 10 サンプルデータ  
 (上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差  
 の収束状況、下段右：理論走時と観測走時間の相関)

表 重みの標準偏差に $\sqrt{(2/N)}$ の標準偏差を用いた学習回数 1000 回での断層構造の平均誤  
 差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
1	1.09	1.63
2	1.07	1.53
3	1.08	1.59
4	1.06	1.34
5	1.05	1.35
6	0.99	1.19
7	1.07	1.45
8	1.08	1.54
9	1.05	1.35
10	1.10	1.90
平均	1.06	1.49

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
最小	0.99	1.19
最大	1.10	1.90
平均	1.06	1.49





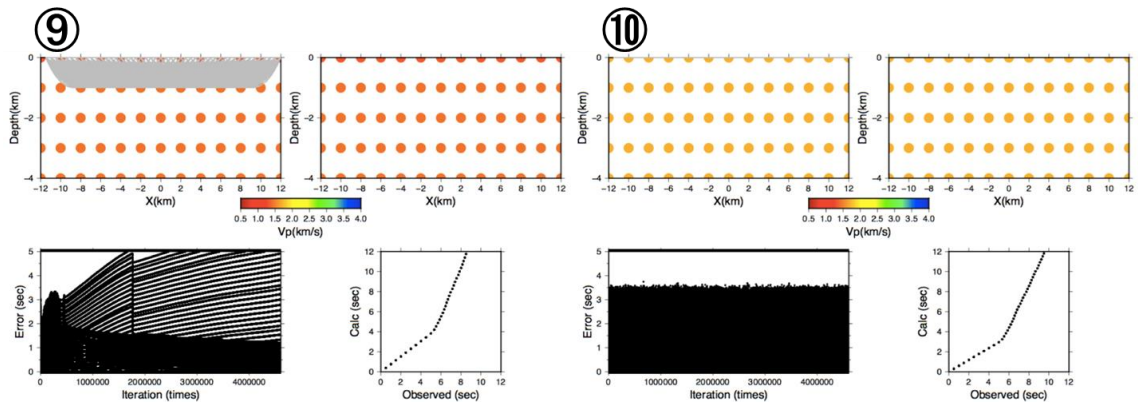


図 重みの標準偏差に $\sqrt{(2/N)}$ の標準偏差を用いた学習回数 2000 回の 10 サンプルデータ  
(上段左：解析結果の速度構造と波線経路、上段右：解析結果の速度構造、下段左：誤差  
の収束状況、下段右：理論走時と観測走時間の相関)

表 重みの標準偏差に $\sqrt{(2/N)}$ の標準偏差を用いた学習回数 2000 回での断層構造の平均誤  
差 $E_{vp}$ と RMS 残差

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
1	1.10	1.88
2	1.06	1.34
3	1.04	1.37
4	1.11	2.02
5	1.13	1.92
6	1.04	1.37
7	1.11	2.10
8	1.10	1.55
9	1.09	1.75
10	1.06	1.34
平均	1.09	1.66

	$E_{vp}[km/sec]$	$RMS$ 残差[sec]
最小	1.04	1.34
最大	1.13	2.10
平均	1.09	1.66

#### 参考文献

- 1) Xavier, Glorot. and Yoshua, Bengio. (2010) : Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS2010). Society for Artificial Intelligence and Statistics.
- 2) Kaming, He., Xiangyu, Zhang., Shaoqing, Ren. and Jian, Sun (2015) : Delving Deep into Rectifiers Surpassing Human-Level Performance on ImageNet Classification. In 1026-1034,
- 3) 春日部拓也(2016) : 屈折法地震探査データの自動解析に関する研究、首都大学東京大学院修士論文
- 4) Humphreys, E. and R. W. Clayton(1988) : Adaptation of back projection Tomography to seismic travel time problems, J.Geophys.Res., 93, 1073-1085,
- 5) 白石和也、阿部進、岩崎貴哉、斉藤秀雄、佐藤比呂志、越谷信、加藤直子、新井隆太、川中卓(2010): 屈折初動走時トモグラフィ解析における初期モデルランダム化による解の信頼性評価、社団法人 物理探査学会 第 63 巻第 4 号、pp.345-356
- 6) Rumelhart, David E., Hinton, Geoffrey E., Williams, Ronald J(1986) : Learning representations by back-propagating errors., Nature323(6088): 533-536,
- 7) Um, J., and C.H. Thurber(1987) : A fast algorithm for two-point seismic ray tracing. Bull. Seismol. Soc. Am., 77, 972-986,
- 8) Xavier, Glorot., Antoine, Bordes., Yoshua, Bengio(2011) : Deep Sparse Rectifier Neural Networks., Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics(AISTATS-11)15: 315-323,
- 9) 斎藤康毅 (2016) 『ゼロから作る Deep Learning-Python で学ぶディープラーニングの理論と実装』 オライリージャパン.
- 10) G.E.P. Box., and M.E. Muller(1958) : A note on the generation of random normal deviates. Annals Math. Stat, Vol. 29, pp. 610-611,

- 11) 丹野智博、堀江和正、小林高彰、森田昌彦 (2013) : ニューラルネットによるパターン分類におけるパターンコーディングの効果、一般財団法人 電子情報通信学会
- 12) 野中和明、田中文英、森田昌彦(2011) : 階層型ニューラルネットの 2 変数関数近似能力の比較、電子情報通信学会論文誌 D Vol.J94-D No.12 pp.2114-2125