

動的タイムワーピング距離を用いたX線天文データの類似検索

林 史尊 *1 天笠 俊之 *2 *3 北川 博之 *2
海老沢 研 *3 中平 聡志 *4

- *1 筑波大学 情報学群 情報科学類
- *2 筑波大学 システム情報系 情報工学域
- *3 宇宙航空研究開発機構 宇宙科学研究所
宇宙科学情報解析研究系
- *4 理化学研究所 基幹研究所 宇宙観測実験
連携研究グループ MAXIチーム

X線アウトバースト現象

- 天体が自身のエネルギーをX線として短期間かつ大量に放出する現象

天体名：GX 339-4
観測期間：1650日間
(2006/9/22～2011/3/29)
観測装置：MAXI

X線強度が短期間で大きく上昇している

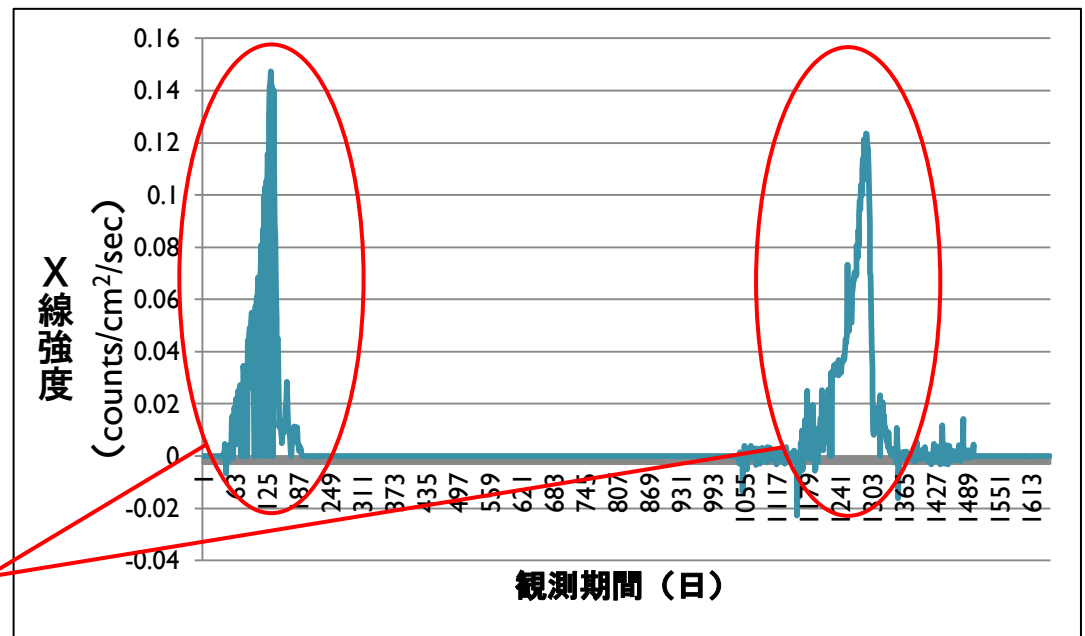


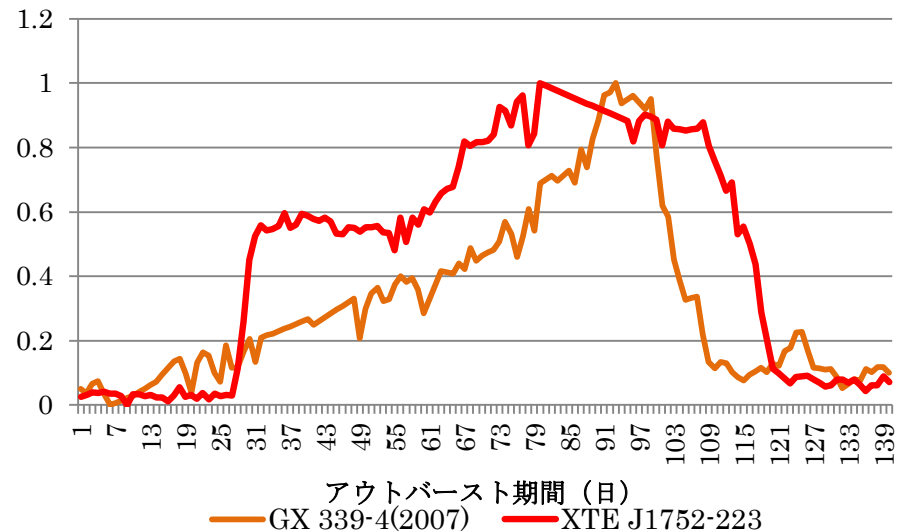
図1.2：X線監視装置から提供される観測データの例

類似するアウトバーストパターンの存在

- X線アウトバースト現象の詳しい仕組みは未解明
- 異なるアウトバースト間でまれに類似した形状のものがみついている
 - 物理過程の類似性が存在する可能性
 - 現象の解明に繋がると期待されている
 - 様々なX線アウトバースト現象から類似パターンを検出したい

図1.3：異なる天体の類似するアウトバーストパターンの例

立ち上がり・立ち下りの傾きに部分的な類似が見られる



研究の目的

- アウトバーストパターンを対象とした類似検索手法を提案し、それらによって類似パターンを発見できることを示す
- アプローチ
 - 観測データからアウトバーストパターンを抽出
 - 動的タイムワーピング法（DTW法）をベースにした類似検索
 - アウトバーストパターンの特徴や類似検索の要求を考慮し、6種類の検索手法を提案
 - 基本手法
 - DTW法
 - Derivative DTW法（DDTW法）
 - 滑り窓法
 - 二分法

表I.1：6種類の検索手法

(全体)	滑り窓法 (部分)	二分法 (部分)
DTW法	DTW法	DTW法
DDTW法	DDTW法	DDTW法

基本手法を利用した部分的類似の検索

アウトバーストパターンの抽出と準備

- 実際の観測データからアウトバーストパターンを抽出
 1. 欠損値の補間（線形補間）
 2. アウトバースト部分の検出
 3. スムージング（線形加重移動平均）
 4. 正規化（値を0～1の間に収める）

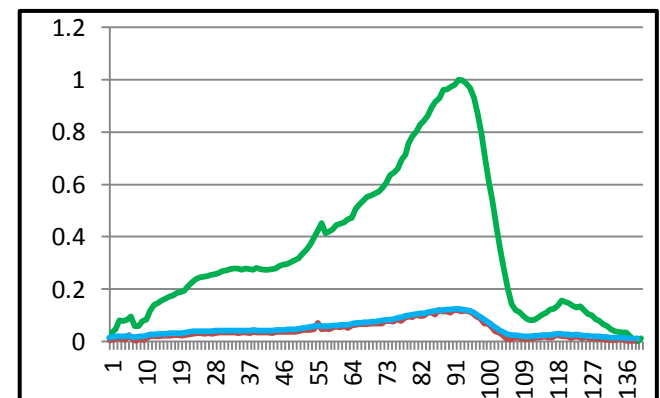
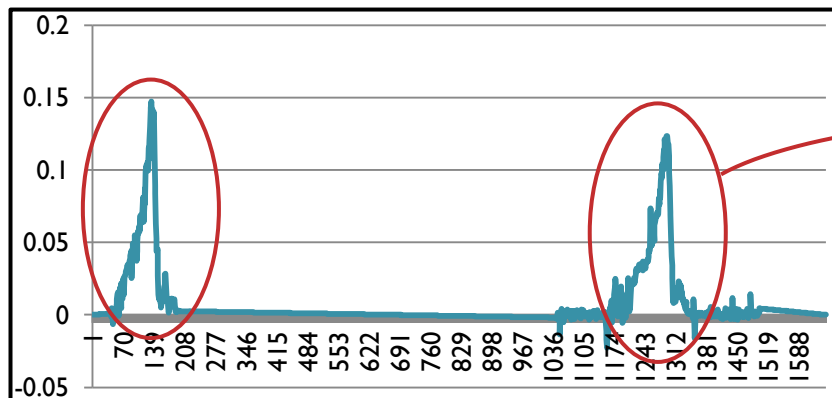


図1.4：アウトバーストパターンの抽出と準備

DTW法 (Dynamic Time Warping)

- 二つの時系列データの最適なマッチングを動的計画法に基づいて求め、類似度を計算する手法 (H. Sakoe, S. Chiba. 1978)
 - 異なる長さの時系列データにも適用可能
 - 点の対応付けごとに距離を求め、全ての距離を累積することで類似度が求められる
 - この累積距離をDTW距離と呼ぶ

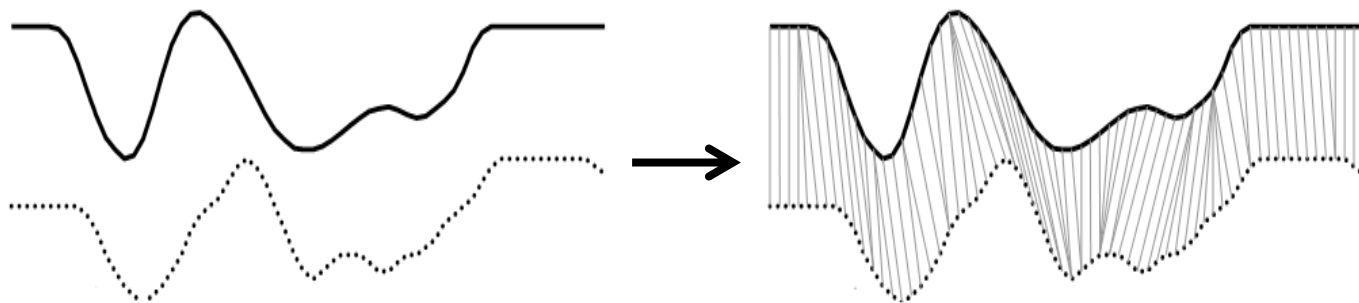


図3.1 : DTW法を用いた二つの時系列データの点の対応付けの例

DDTW法 (Derivative DTW)

- 変化の様子に着目した類似検索
- DTW法の改良手法の一つ (E. J. Keogh, Michael J. 2001)
 - ある時系列データに対し直接DTW法を適用するのではなく、その時系列データから導出されるデータ値の上昇・下降の度合を示した時系列データに対してDTW法を適用

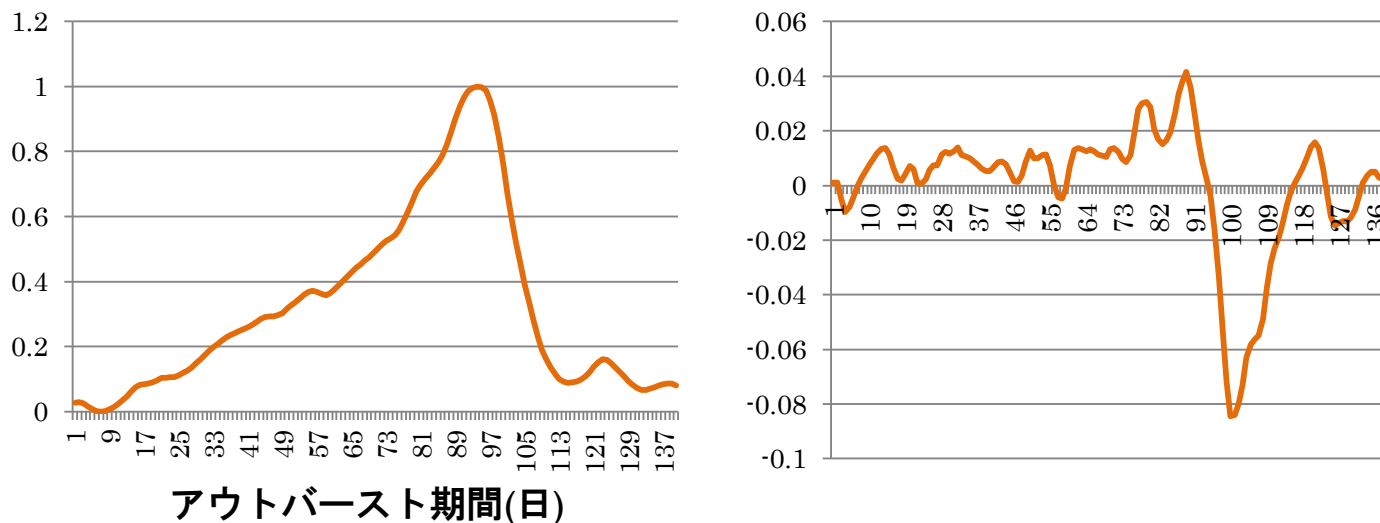


図3.2 : DDTW用変換データの例
左が変換前、右が変換後のデータ

滑り窓法 (滑り窓を用いた部分系列の類似度計算)

- 部分的類似を持つパターンの発見
 - パターンの一部に限定してDTW法, DDTW法を適用
- 滑り窓 (Sliding Window) を用い以下を繰り返す
 - 滑り窓に切り取られた部分区間内でDTW距離を計算
 - 滑り窓をずらす
- 窓幅とずらし幅はデータ長に比例して伸縮
 - 二つのデータ長が異なる場合でも, ずらす回数を固定にする

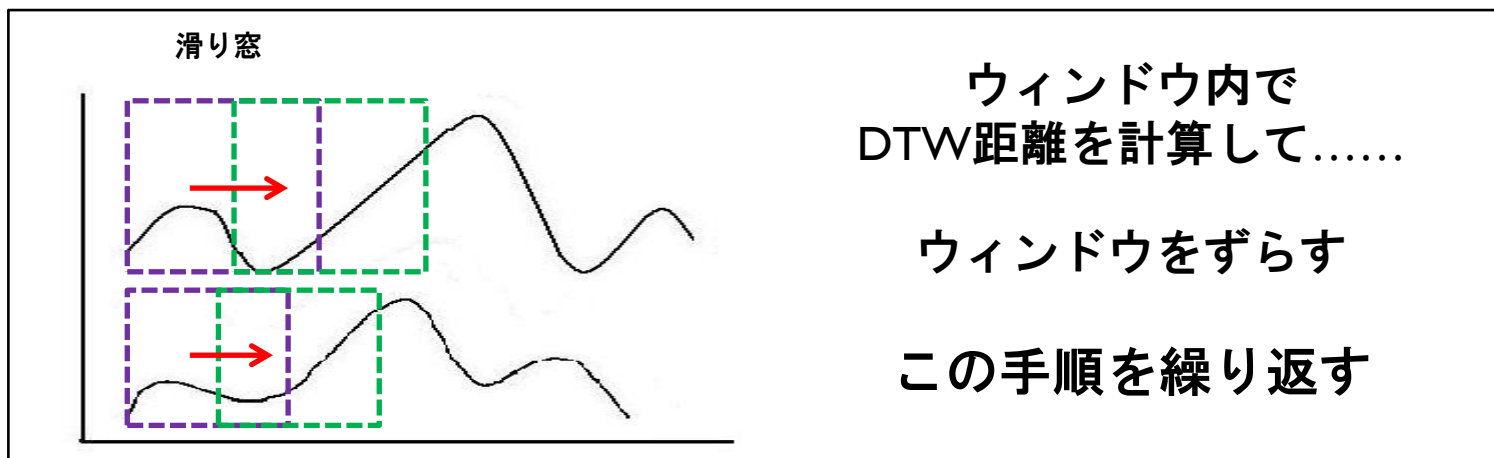


図3.3 : 滑り窓を用いたDTW距離の計算

二分法 (パターンを頂点で二分した類似度計算)

- 立ち上がり, 立ち下がりの様子に着目した類似検索
 - アウトバーストの頂点 (最大値) でデータを二分割し, 前部, 後部のそれぞれでDTW距離を計算

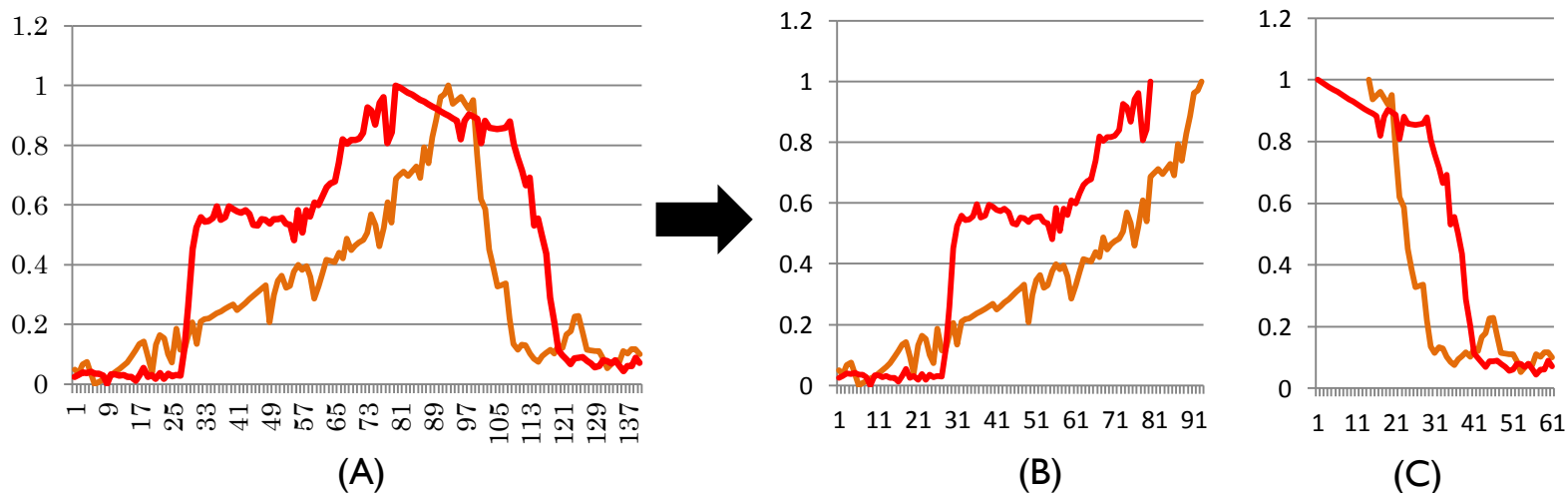


図3.4：二分法の例

(A) 元のパターン (B) 立ち上がり部分のパターン
(C) 立ち下がり部分のパターンを左右反転したもの

データセット

- NASA の管理する衛星 Swift に設置されたX線監視装置 BAT が観測する天体の内，一定以上のX線強度を持つ155個の天体の観測データを使用
 - 1個の天体の観測データから複数のアウトバーストが検出された場合，発生期間の長い二つのみを抽出
 - 計算時間が膨大になってしまうことを防ぐため
 - 抽出されたアウトバーストパターンは合計153個

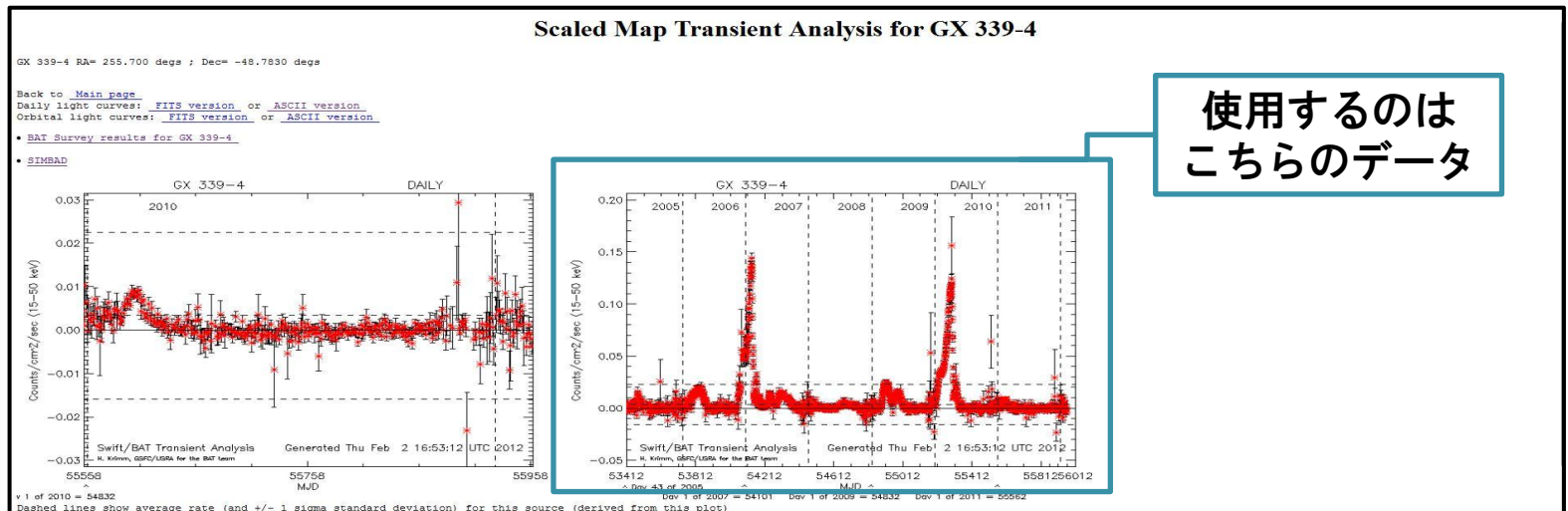


図4.1：BATから提供される観測データの例（Webページのスクリーンショット）

実験内容

- 目的
 - 提案した類似検索手法により，類似するアウトバーストパターンが検出できることを示す
- 三種類の実験
 - 実験1：単純なDTW法，DDTW法の適用
 - 実験2：滑り窓法を用いたDTW法，DDTW法の適用
 - 実験3：二分法を用いたDTW法，DDTW法の適用
- 評価
 - 結果を類似度順に並べ，上位20件について実際の波形を目視で確認

実験1

- 目的
 - 類似度の高いアウトバーストパターンが検出できているか確認する
- 方法
 - 全てのデータに対し、相互の類似度をDTW法，DDTW法によりそれぞれ求める
 - 類似度の高いペア20件について、実際の波形を目視で確認し、類似しているかを判定

実験 1 の結果

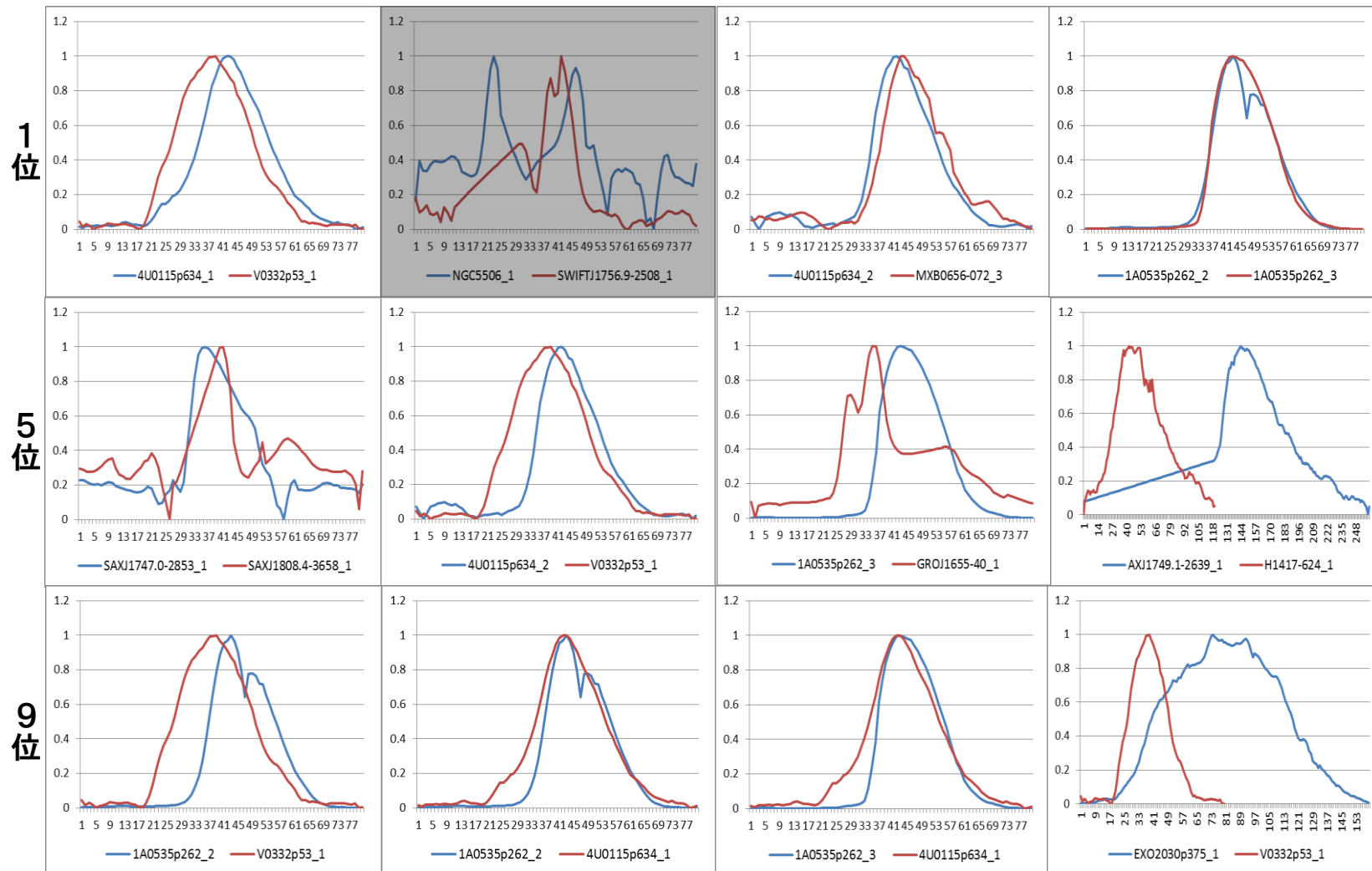


図4.2.1 : 実験 1 の結果 (DTW法)

実験 1 の結果

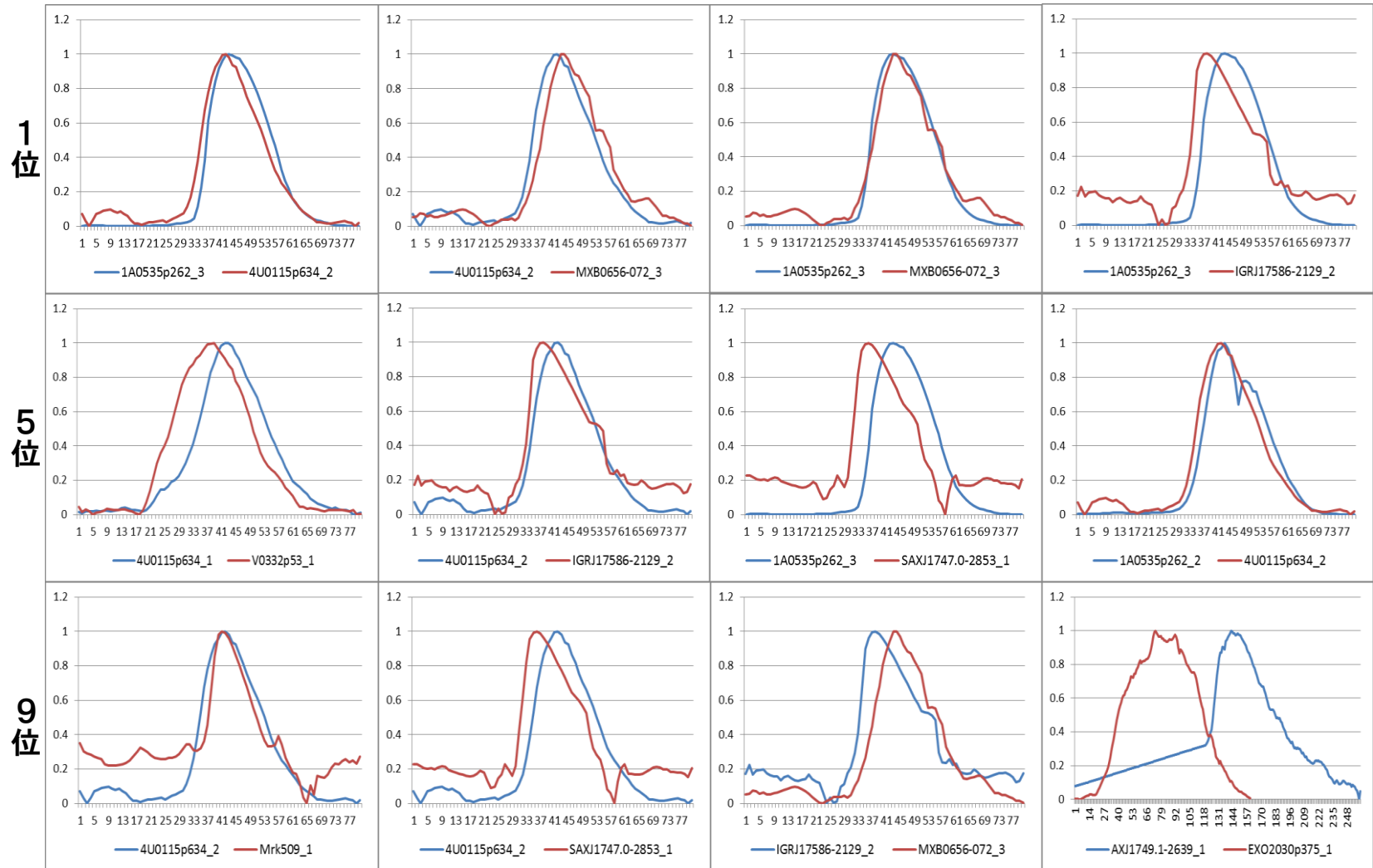


図4.2.2：実験 1 の結果 (DDTW法)

実験2

- 目的

- 滑り窓法を用いて部分的類似を含むパターンが検出できているか確認する

- 方法

- 全てのデータに対し、相互の類似度を滑り窓法（DTW法、DDTW法を用いる）により求める
- 指定したしきい値以下のDTW距離を持つ区間を部分類似区間とし、部分類似区間を三つ以上持つペアを抽出
- 類似部分区間のDTW距離の平均値を求め、平均値が小さいペア20件について、実際の波形を目視で確認し、部分的類似が発見できているかを判定

実験2の結果

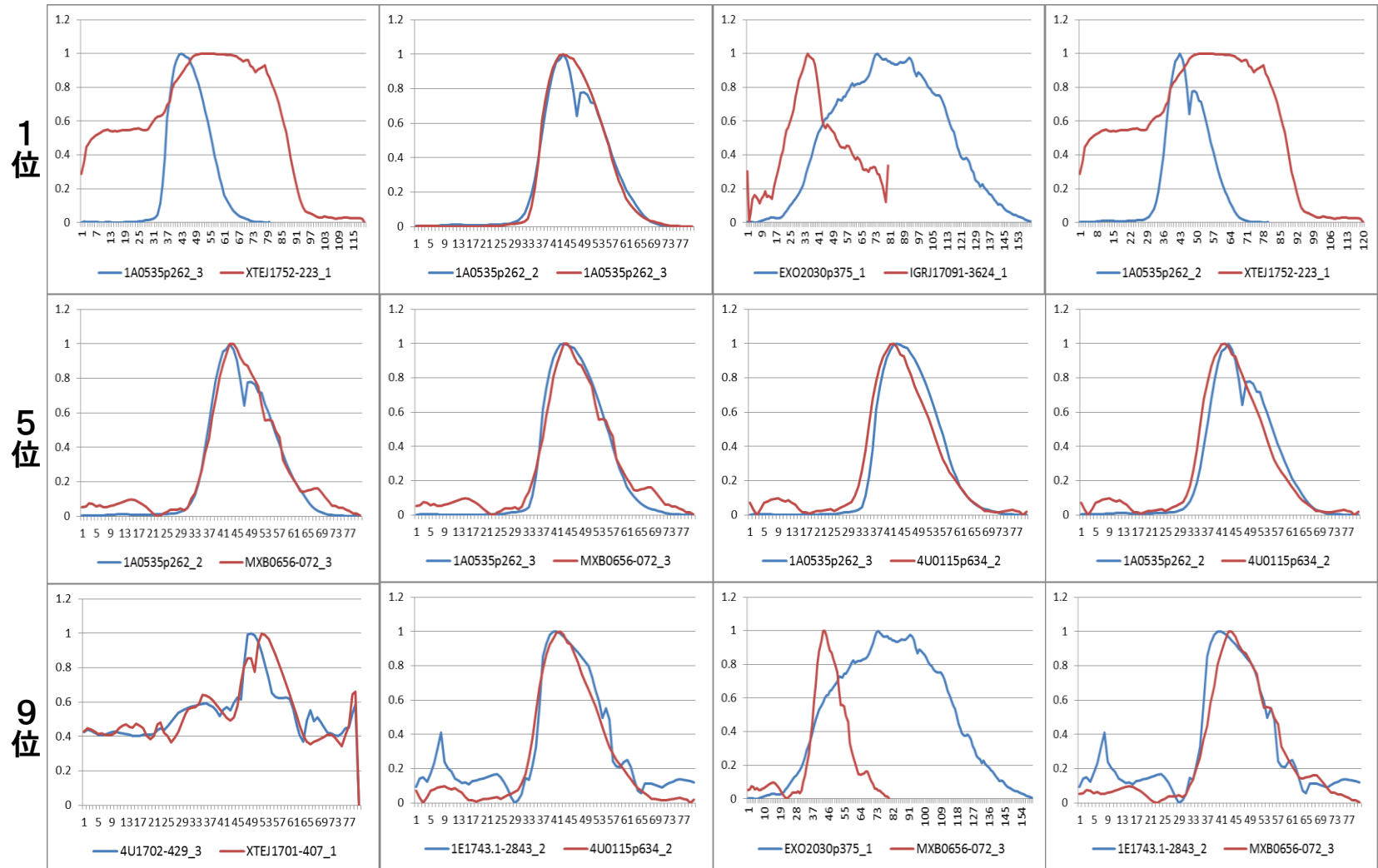


図4.3.1：実験2の結果（DTW法）

4. 評価実験 - 実験2

実験2の結果

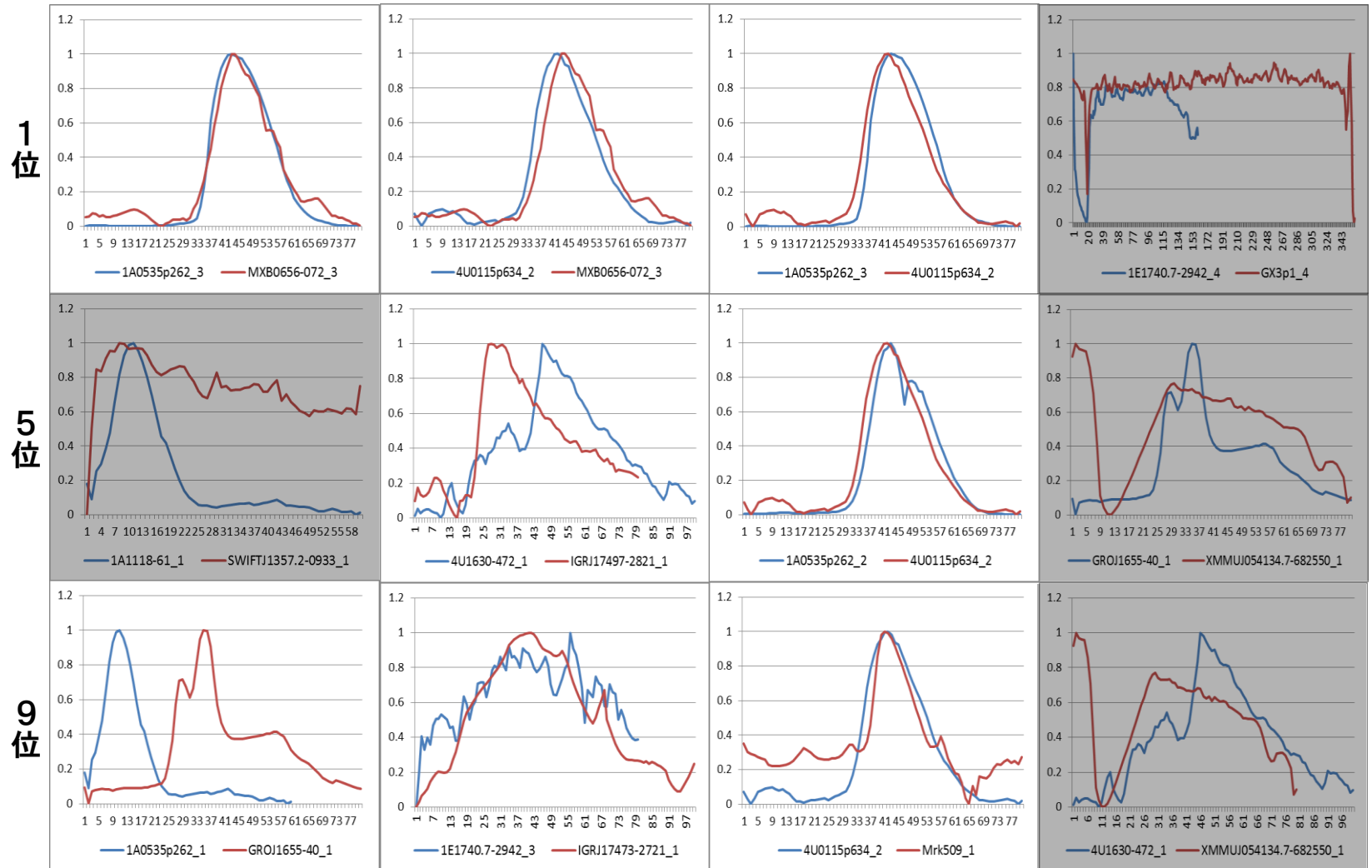


図4.3.2：実験2の結果（DDTW法）

実験3

- 目的
 - 二分法を用いて、立ち上がり(下がり)部分が類似するアウトバーストパターンが検出できているか確認する
- 方法
 - 全てのデータに対し、相互の類似度を二分法（DTW法，DDTW法を用いる）により求める
 - 立ち上がり部分，立ち下がり部分のそれぞれで，類似度の高いペア20件について実際の波形を目視で確認し，類似しているかを判定

実験3の結果

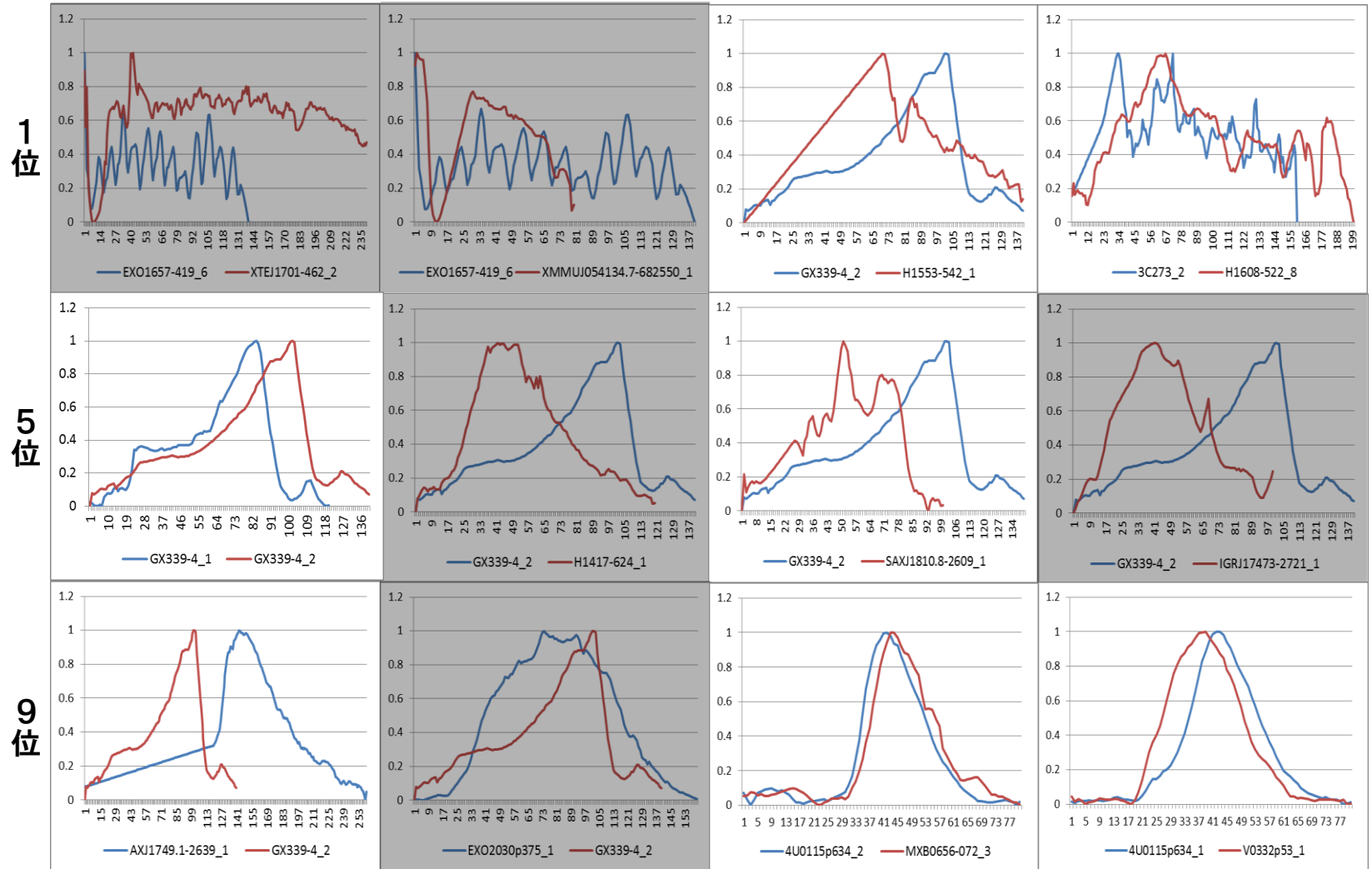


図4.4.1：実験3の結果（DTW法，立ち上がり部分）

実験3の結果

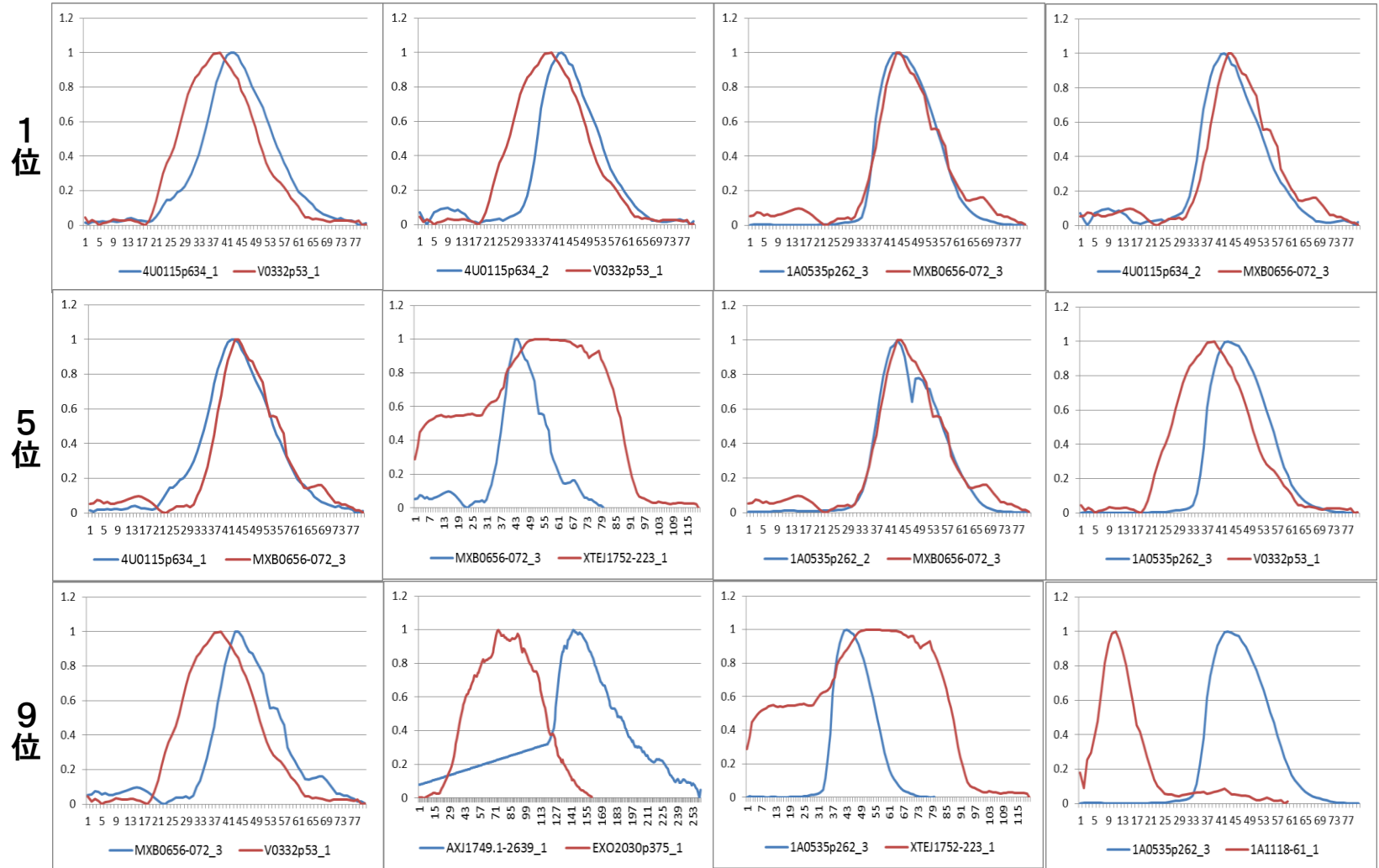


図4.4.1：実験3の結果（DTW法，立ち下がり部分）

実験3の結果

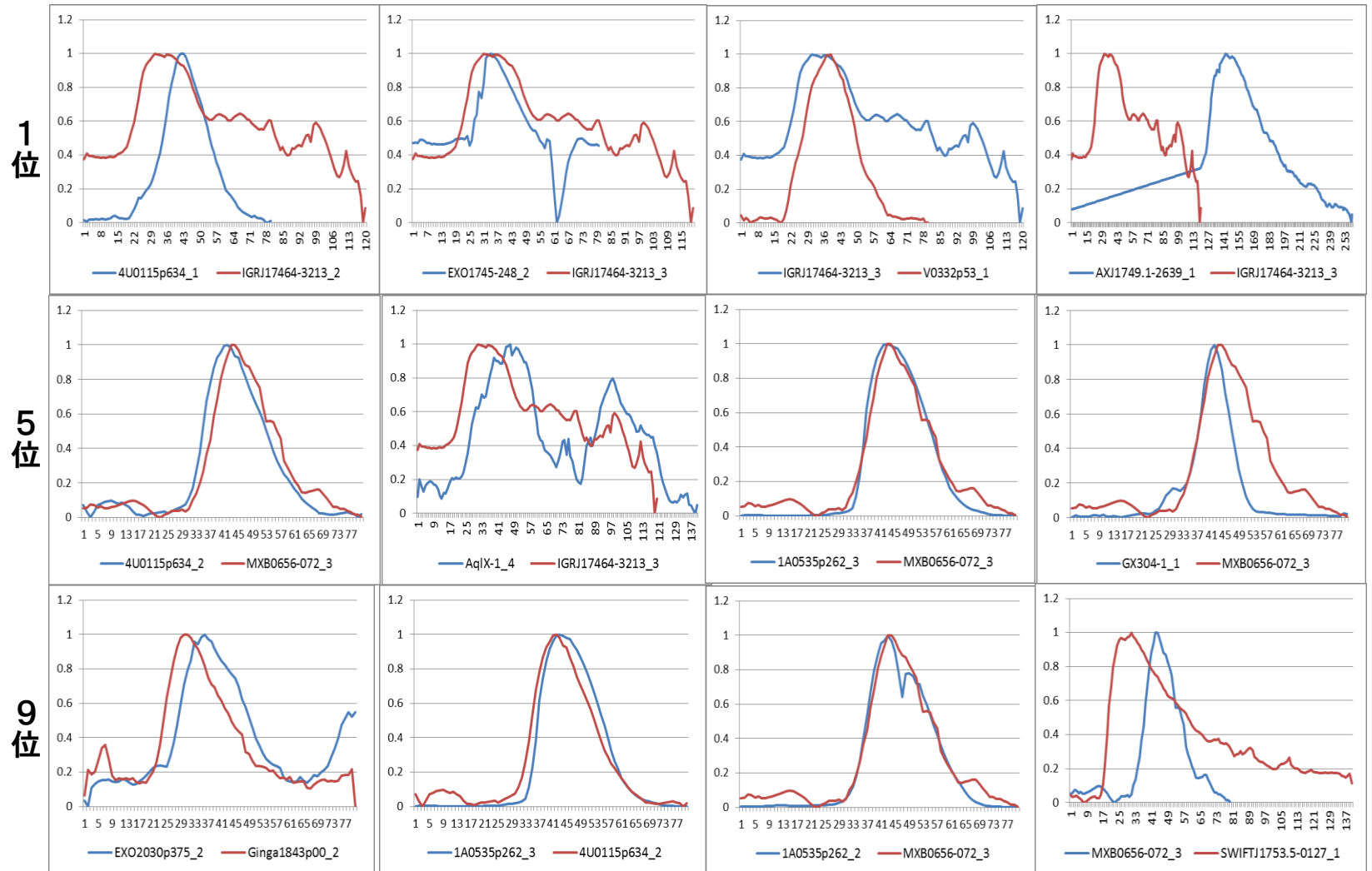


図4.4.2：実験3の結果（DDTW法，立ち上がり部分）

実験3の結果

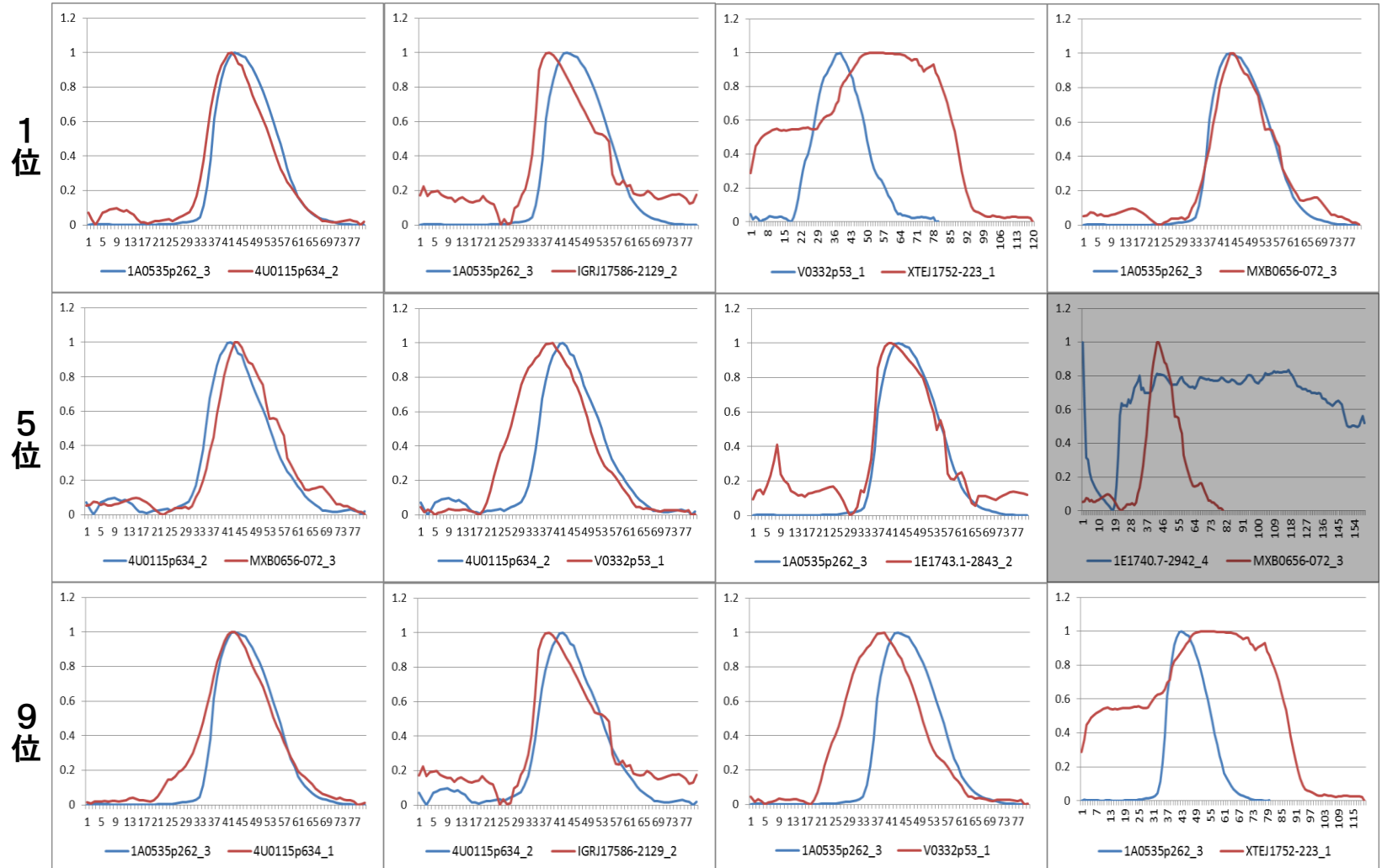


図4.4.2：実験3の結果（DDTW法，立ち下がり部分）

全体の結果

表4.1：各実験における上位20件中の類似判定件数

実験の種類		類似判定件数 (20件中)
単純 (DTW法)		18
単純 (DDTW法)		20
滑り窓法 (DTW法)		18
滑り窓法 (DDTW法)		13
二分法 (DTW法)	立ち上がり部分	13
	立ち下がり部分	20
二分法 (DDTW法)	立ち上がり部分	20
	立ち下がり部分	18

まとめ

- **本研究では，X線アウトバーストの類似パターンを検出する手法を提案し，それらを使って実際に類似パターンを発見できるか検証した**
 - 実験1で，DTW法とDDTW法を用い，全体的に類似したパターンを発見できることを示した
 - 実験2で，滑り窓法を用い，部分的類似を含むパターンを発見できることを示した
 - 実験3で，二分法を用い，立ち上がり，立ち下がり部分の類似を持つパターンを発見できることを示した

今後の課題

- アウトバーストパターンの抽出方法を改良
- 実験結果を専門家に送って意見を求め、要求に対する最も適切な手法を検討
- 二つのX線観測データを入力とし、アウトバーストパターンの類似検索を行うアプリケーションを作成
 - Webサービスとして提供