

# トラッキングデータを用いた サッカーの試合における 戦況変化の抽出

第5回スポーツデータ解析コンペティション  
サッカートラッキング部門

東京大学大学院 工学系研究科  
神谷啓太、中西航、泉裕一郎

# 分析の背景／目的および方針

# 背景と目的

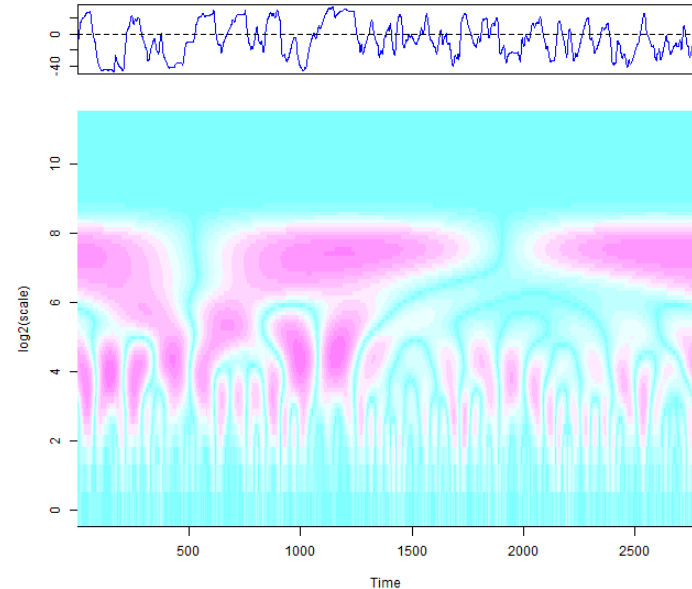
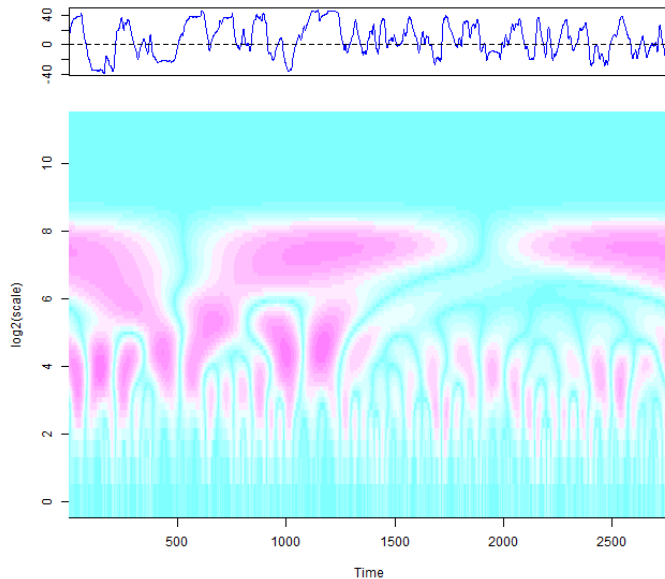
- 背景
  - サッカーの試合：22人の選手がピッチ上で複雑に動いている
  - トラッキングデータには22人の複雑な動きがすべて含まれている
- 既往研究(Kijima et al., 2014)
  - 「複雑に見えるサッカーの試合も、実は同じような上下動が様々な時間スケールで繰り返されているものである」
- たしかにそのような気もするけれど、一方で、観戦者が感じる「戦況」には試合中で変化がある
  - 攻勢であったのに、いつの間にか守勢に転じていた
  - 停滞していた状況が、一気に動き出した
- 目的
  - トラッキングデータを用いて、このような戦況変化の抽出を目指す
  - 具体的には、
    - (1)分析に使う変数を選定する
    - (2)Change Finderを適用する
    - (3)結果を解釈する



適用結果を変数選定にフィードバック

# 戦況を知るには

- 抽出したい「戦況」というものの自体を把握するために、選手位置の時系列データに対して基礎分析を行った
  - グラフを眺める
  - 統計量(平均、分散など)を求める
  - 連続ウェーブレット変換で周波数帯に分解する

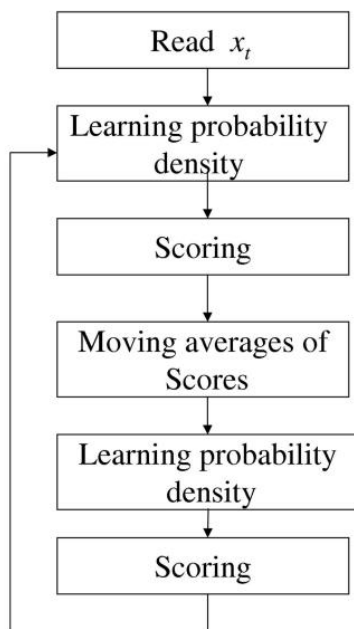
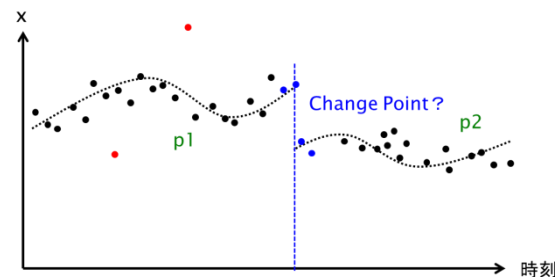


- ただ単に見ていてもよく分からない(ビッグデータにありがちなこと)
- そこで、普通の時系列データ同様に自己回帰することを考える

# Change Finderを用いた変化点検出

5

- 行いたいこと
  - 位置データなどの時系列入力変数を自己回帰する
  - この回帰状況をみながら戦況の変化を抽出したい
- 分かりにくい時系列データからも変化検出が行える手法：  
Change Finder (Takeuchi and Yamanishi, 2006)が適用できる
- Change Finderの概略
  - 回帰する変数は多次元でも可能(VAR過程に適用可能)



時系列データ $x$ を $K$ 次の

ベクトルAR過程でモデル化：
$$x_t = \sum_i^K w_i(x_{t-i} - \mu) + \mu$$

$x$ を観測する度に確率密度関数 $p_t$ を学習  
対数損失を逐次算出

対数損失の $T$ 次移動平均を算出し、  
新たな時系列データ $y_t$ を作成

$y_t$ をVAR過程でモデル化

確率密度関数 $q_t$ を学習しつつ、対数損失を逐次算出

対数損失の $T'$ 次移動平均を変化点スコアとして出力

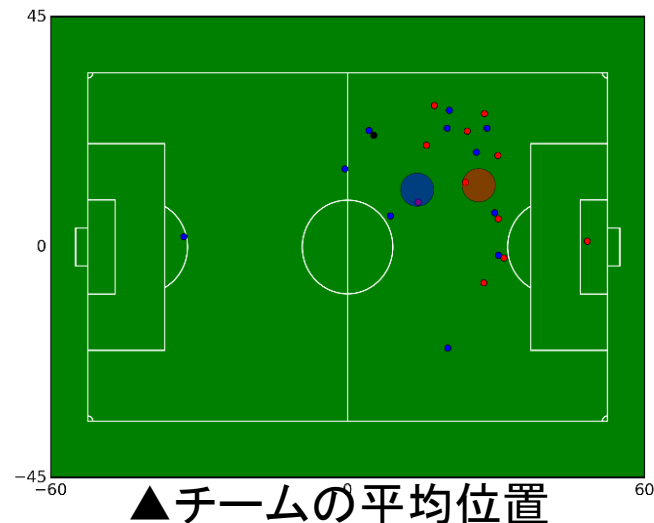
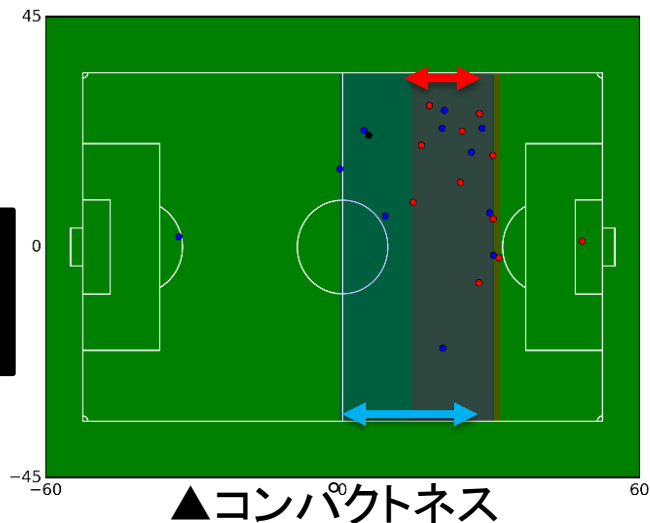
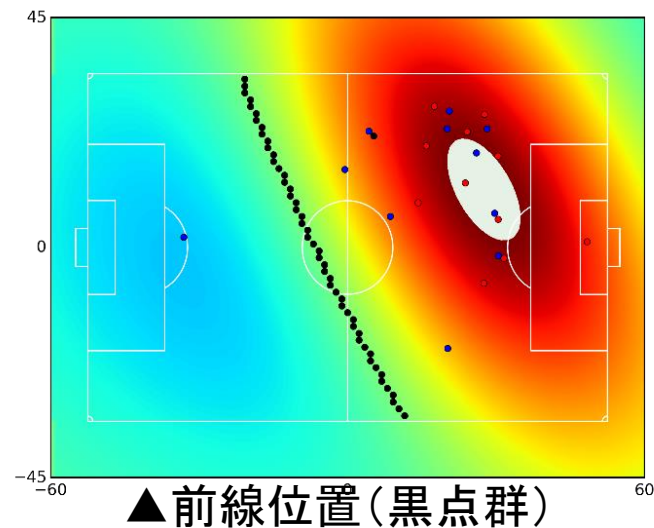
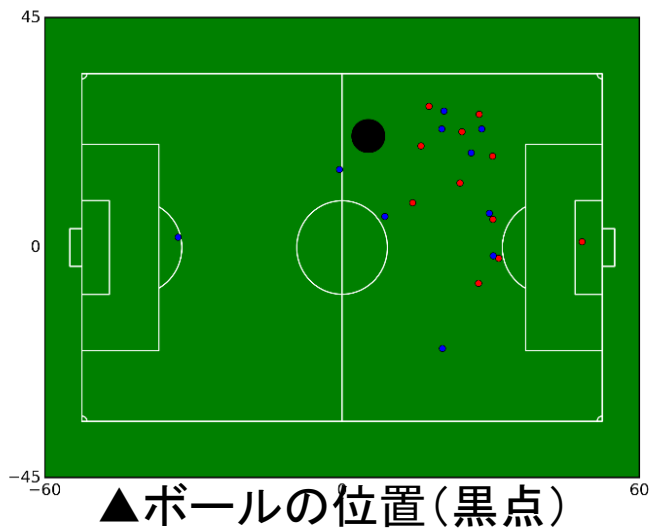
- (1) 分析に使う変数を選定する
- (2) Change Finderを適用する

# 入力変数の検討

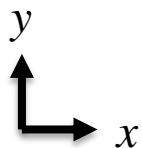
- 単純なChange Finderの適用
  - 入力変数にチーム11人全員の座標
  - ピッチ上22人全員の座標
  - などなど…
- 闇雲に行ってもあまりうまくいかない
- 基礎分析で行った連続ウェーブレットの結果を再検討
  - 「選手位置」は同チームの選手であればポジションに関係なく同傾向
  - 「平均位置」のようなもので代表させて良いのでは？
- それよりも、サッカーの複雑さは、選手相互間やボールとの関係から発生しているのではないか？
- いろいろな変数を作成し、それぞれ基礎分析
  - 「ボールの座標」(ボールタッチデータより作成)
  - 「オフサイドラインの座標」「一番前線にいる選手の座標」
  - 「一番前線の選手の座標」と「オフサイドラインの座標」との距離(コンパクトネス)
  - 「ボールから1番目、…11番目に近い選手までの距離」

# 作成した変数の説明（一例）

8



赤:AWAY  
青:HOME





- 変数の選定
  - すべて1秒ごとのデータに加工
  - ボール位置のX座標
  - HOMEチーム平均位置のX座標
    - HOMEチームとAWAYチームとのデータの相関が高いためHOMEのみ
  - 前線位置のX座標(作成方法は既往研究(Kijima et al., 2014)に従う)
  - HOMEチーム・AWAYチームの「一番前線の選手の座標」と「オフサイドラインの座標」との距離(コンパクトネス)
- 最終的なChange Finderのパラメータ設定
  - 次数 $k$ は通常のVAR過程を行った場合のAICにより決定(=5)
  - 次元数 $d$ は上記で選んだ変数の数(=5)
  - 忘却率 $r$ や平滑化パラメータ $T$ は試行錯誤して決定

## Change Finder のパラメータ

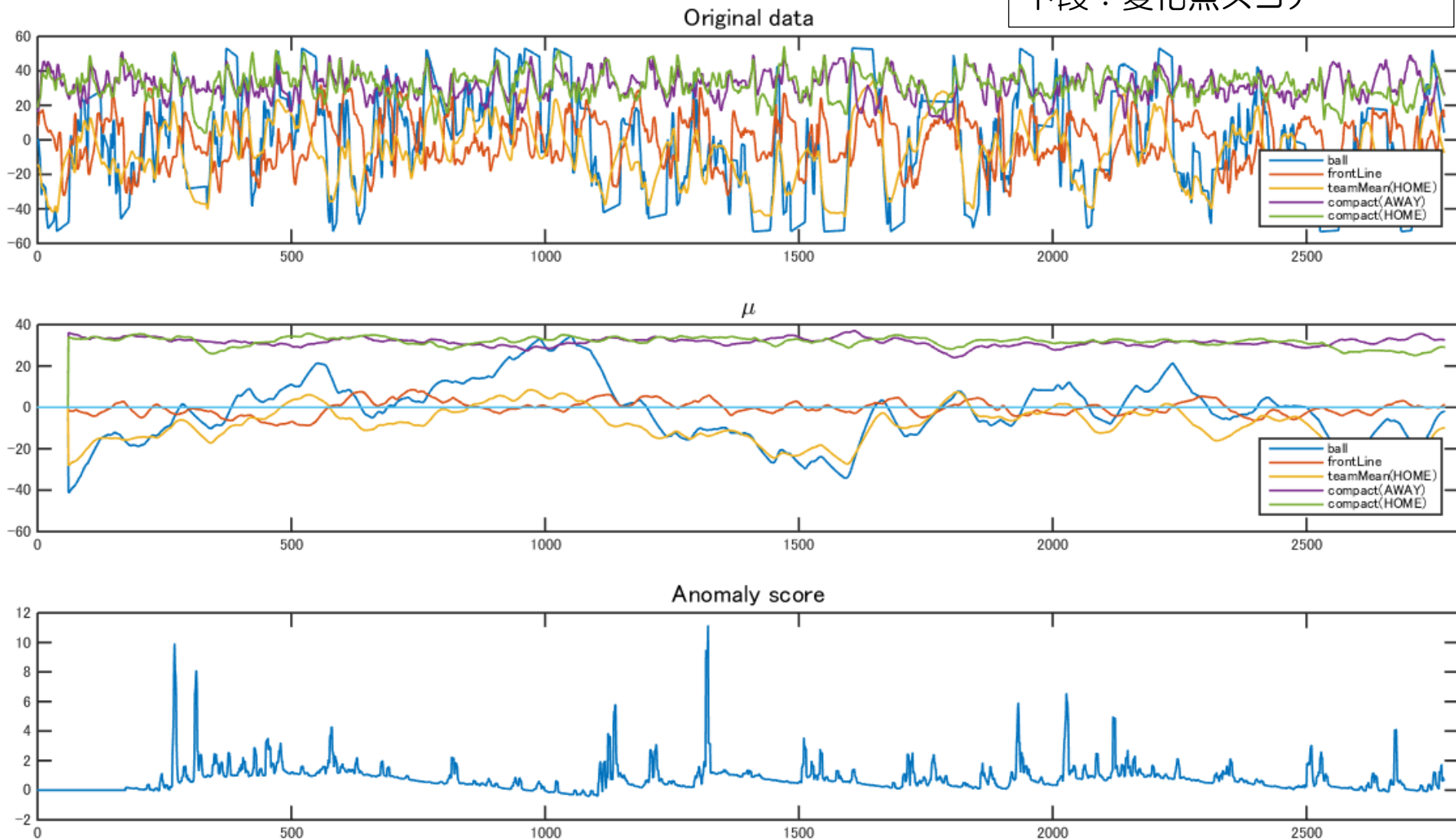
VAR次数	入力次元数	忘却率 $r$	平滑化窓幅 $T$	平滑化窓幅 $T'$
5	5	0.01	50	5

# 変化点の検出結果 1

10

- 鹿島ー湘南（2015/3/14）前半

上段：入力時系列  
中段：パラメータ $\mu$ の推定値  
下段：変化点スコア

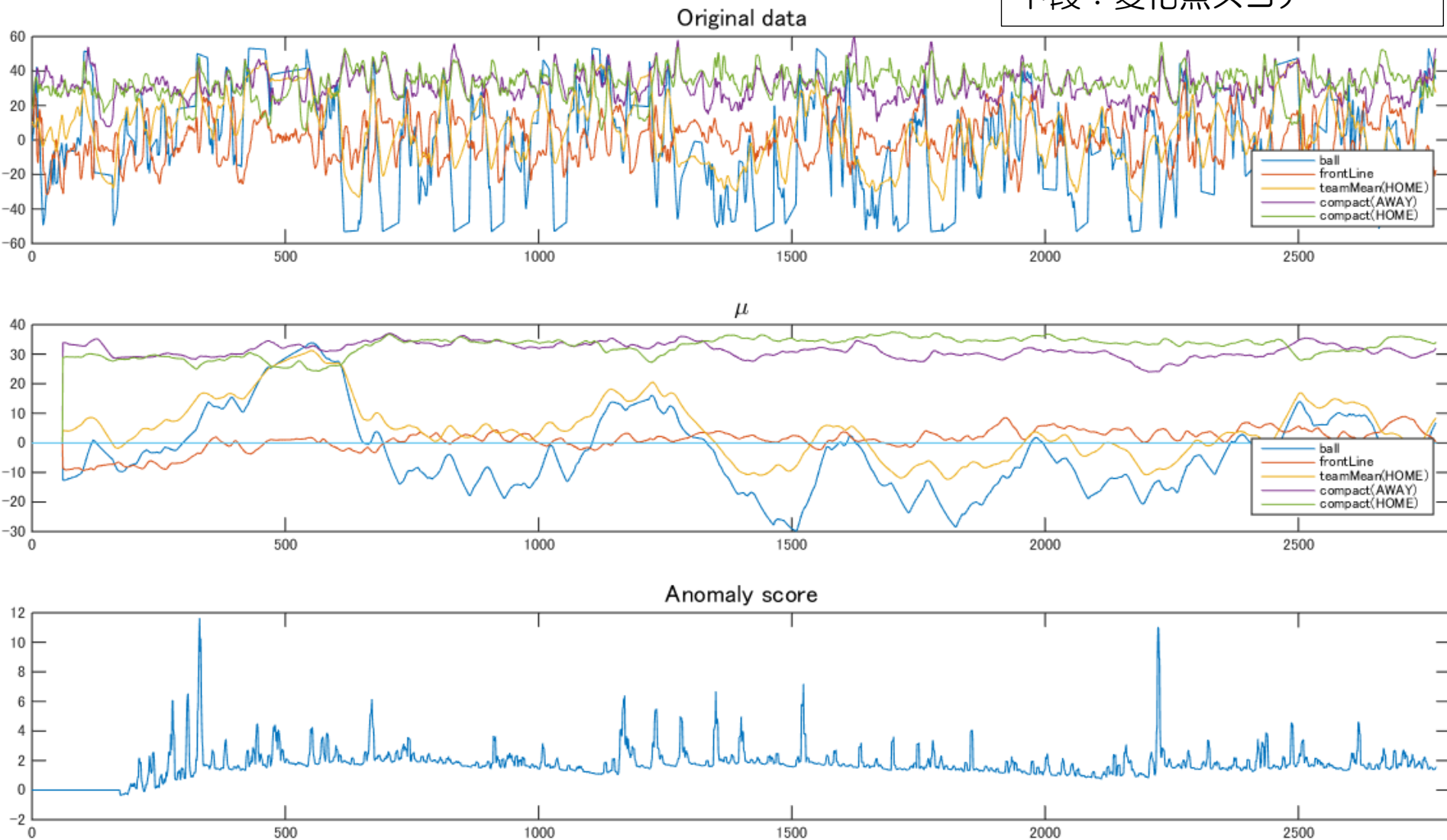


# 変化点の検出結果2

11

- 鹿島ー湘南（2015/3/14）後半

上段：入力時系列  
中段：パラメータ $\mu$ の推定値  
下段：変化点スコア

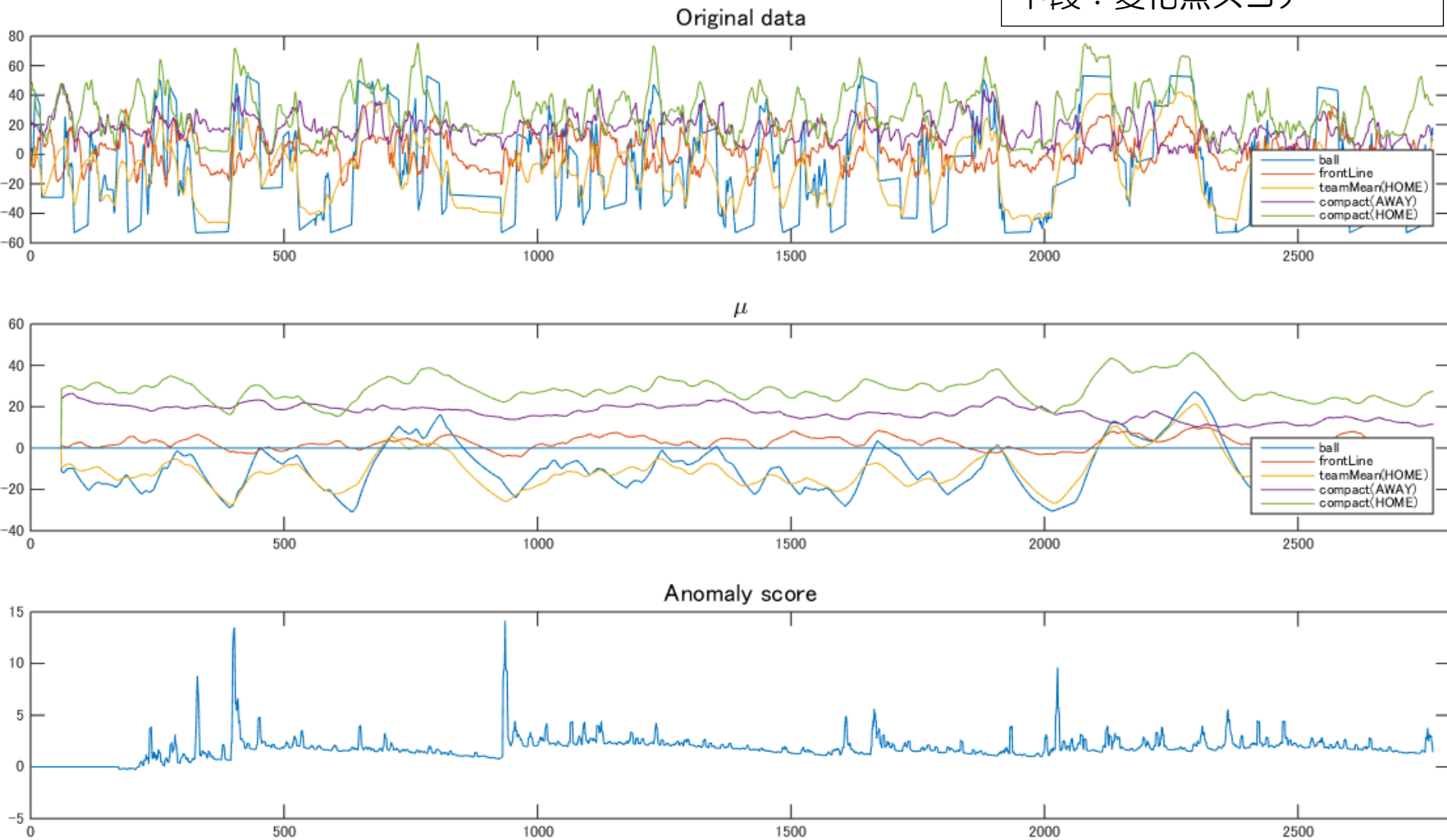


# 変化点の検出結果 3

12

- 松山ー湘南 (2015/6/27) 前半

上段：入力時系列  
中段：パラメータ $\mu$ の推定値  
下段：変化点スコア

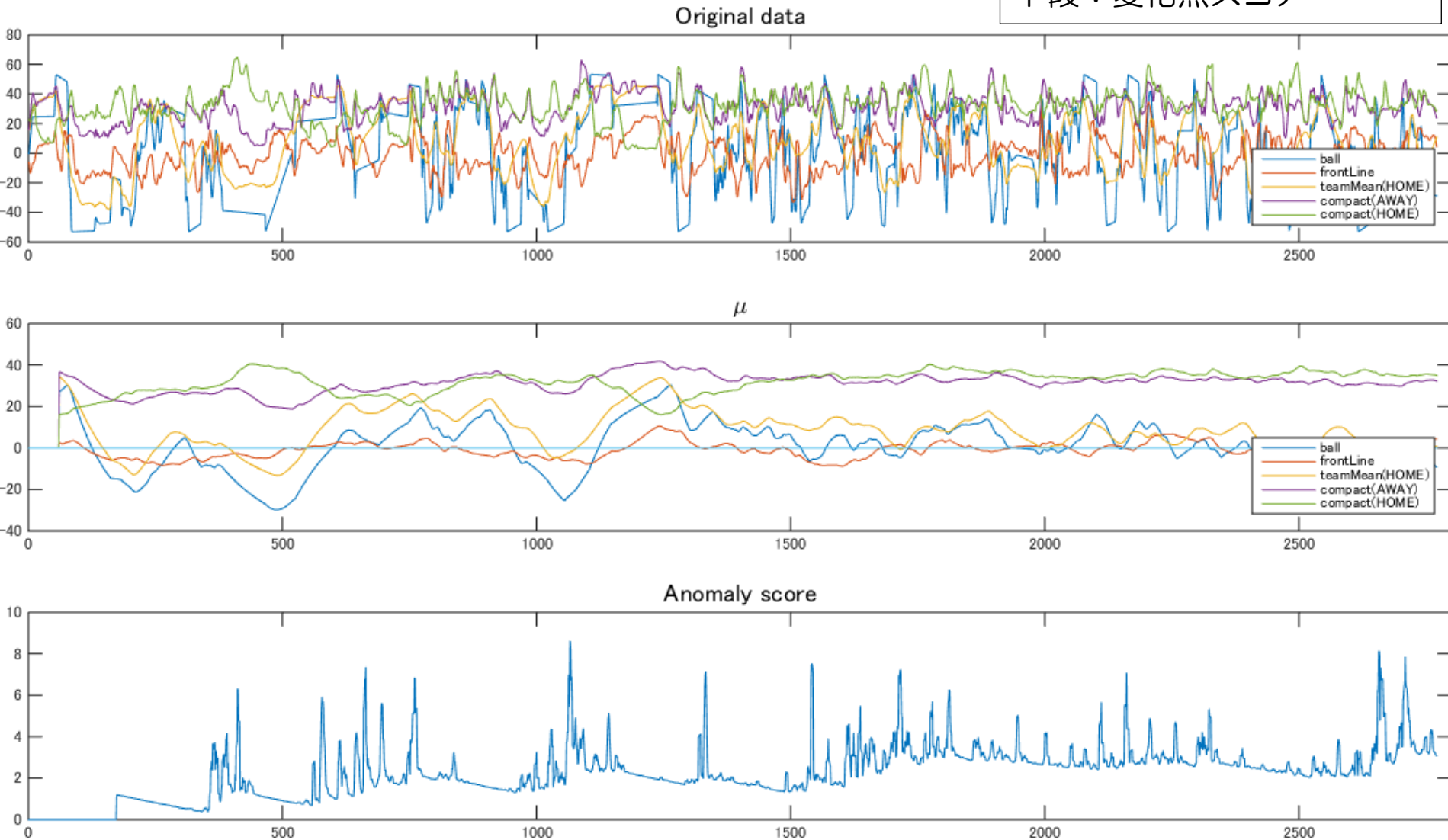


# 変化点の検出結果4

13

- 松山ー湘南（2015/6/27）後半

上段：入力時系列  
中段：パラメータ $\mu$ の推定値  
下段：変化点スコア

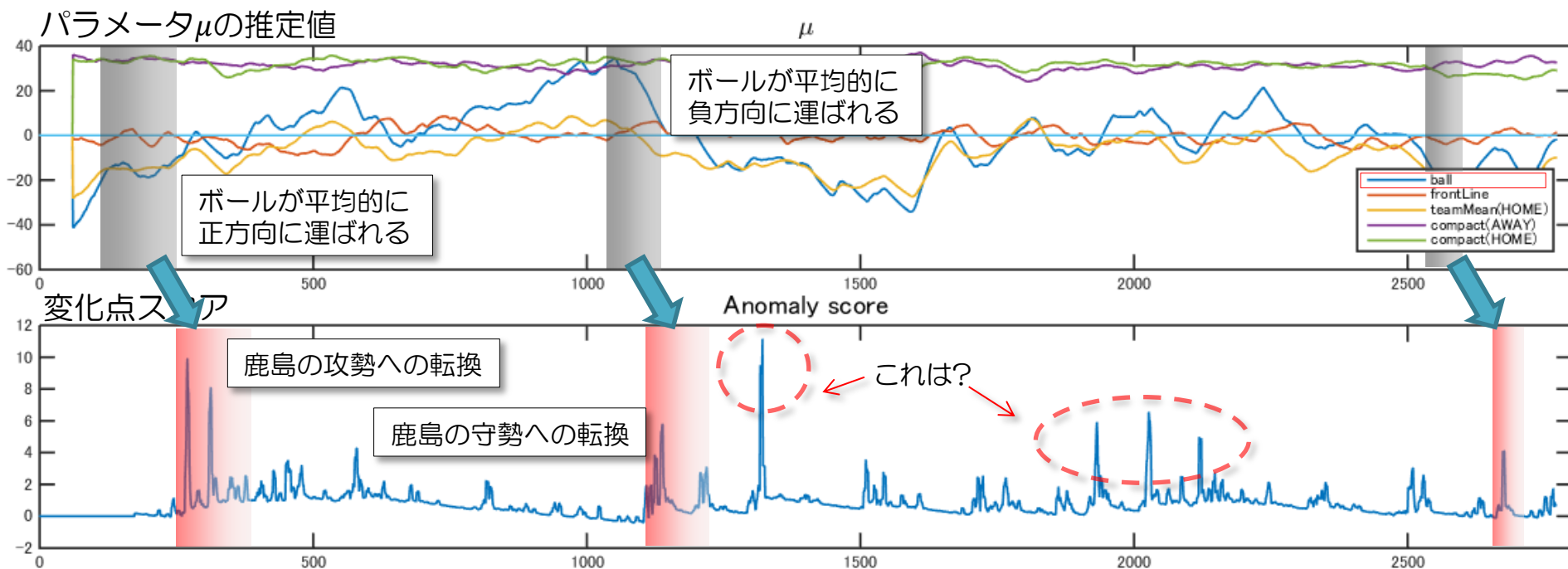


## (3) 結果を解釈する

# 結果の解釈

～鹿島・湘南戦前半を例に～

15

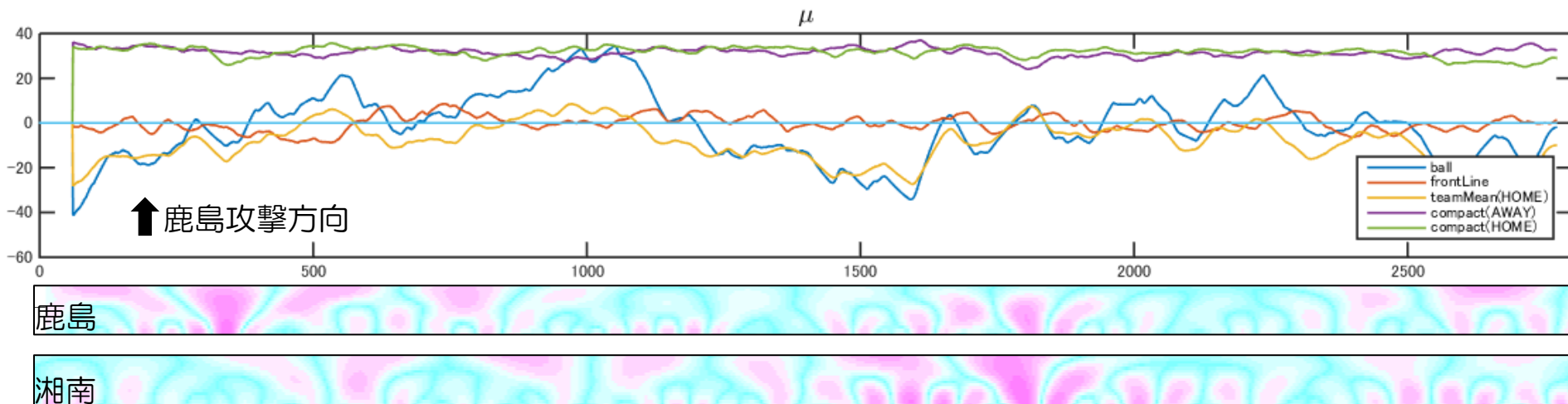


- パラメータ $\mu$ の推定値うち、ボールのX座標に注目
  - VARモデルの式より、この値は「自己回帰分を除いた、ボールが存在する平均的なX座標」と解釈できる
  - ボールの平均的な位置は、チーム間の攻守状況を反映しているのでは
- ここで検出される変化点は、ボールの位置で説明できるような攻守の切り替わりに対応している
- しかし、これではすべての変化点を説明できない

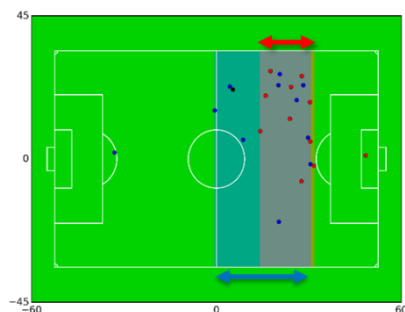


# コンパクトネスとの関係

16

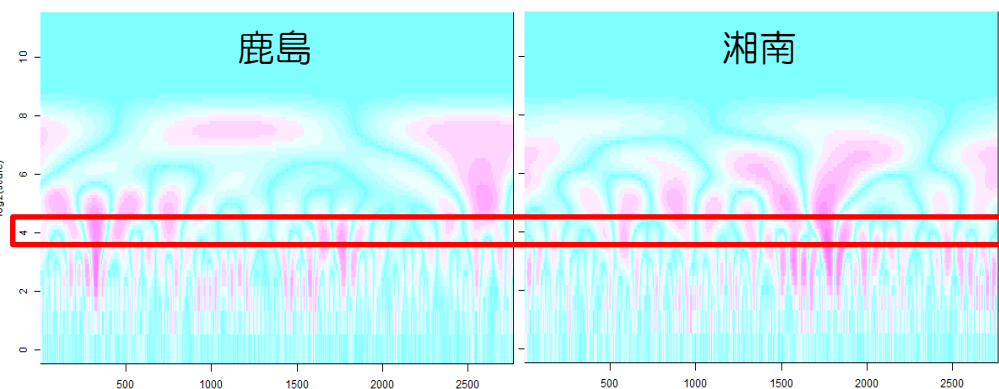


▲ 1/256Hz付近の反応パターンを抜粋



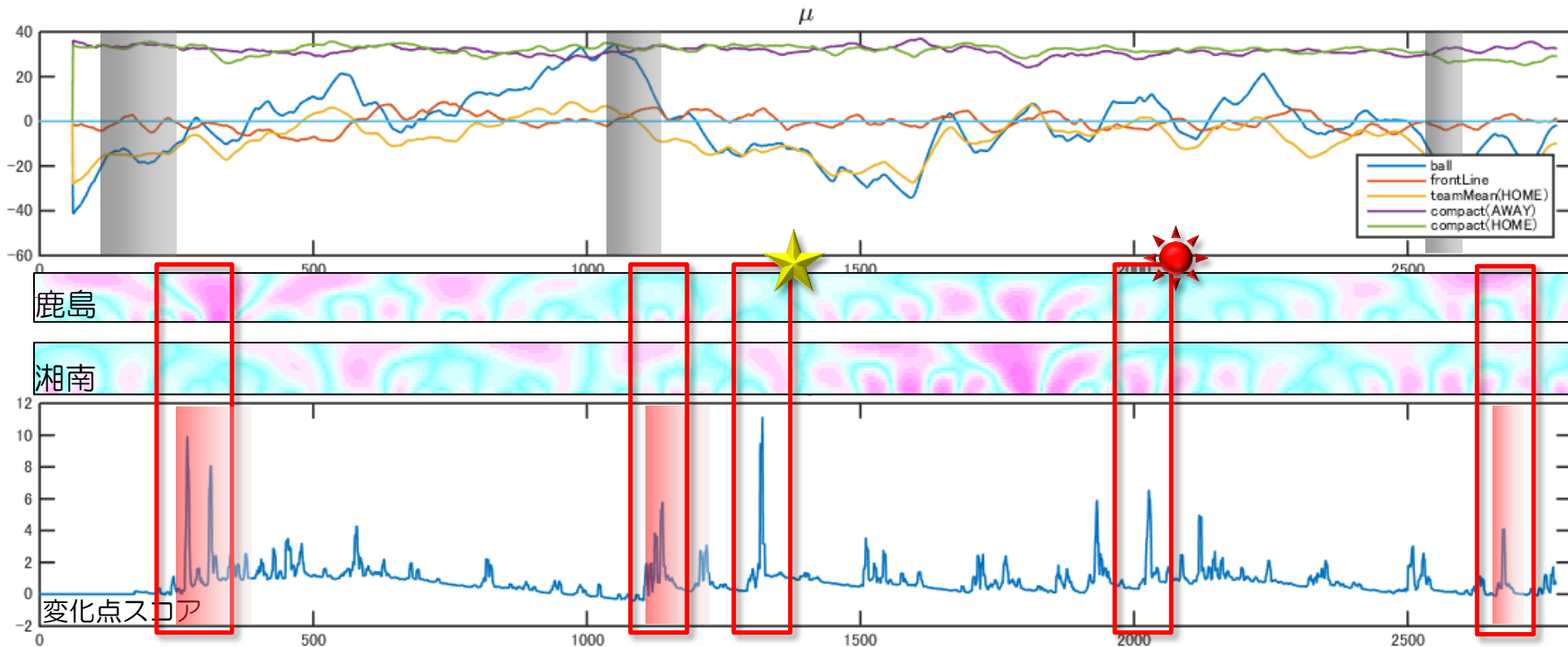
▶ 「コンパクトネス」の連続ウェーブレットへの反応パターン

コンパクトネス＝  
(最前線) - (オフサイドライン)



- 入力変数のウェーブレットを全て見たところ、コンパクトネスが1/256Hz付近ではチーム間で逆の反応パターンをとってる傾向にあった
- この周波数帯でのコンパクトネスが、攻守の切り替わりを表しているのでは？
  - (攻勢) ボールを奪取した後、最前線が急激に、オフサイドラインが徐々に上がる⇒高反応
  - (守勢) 守りに転じた時は、全体として徐々に下がり始める⇒低反応





- コンパクトネスで変化点の解釈を考えた場合：
  - ボールの位置変化から明らかだった前述の変化点も含み、攻守も整合的
  - 変化点★は、ボールの位置変化をよく見ると、膠着状態から攻めが実現した時の戦況変化であったと解釈できる
- 検知された変化点はどれも攻守に関する戦況変化を説明できそう
  - 変化点★について、簡単には分からない戦況変化である可能性
    - 実際の試合映像などで確認したい

# まとめ

＜トラッキングデータを用いて、サッカーの試合における戦況変化の抽出を行った＞

- 分析に用いる変数について、基礎分析や既往研究に基づき検証し、選定した
- Change Finderを適用し、変化点の抽出を行った
- 検出された変化点のうち複数は、実際の攻守に関する戦況変化と対応していた
- 簡単には分からない戦況変化も検出できる可能性が示唆された

＜今後の方策＞

- 変化点が検出された原因について、ARモデルのパラメータを精査する
- 実際の試合映像などで試合状況について確認したい

# 参考文献

- A. Kijima, K. Yokoyama, H. Shima and Y. Yamamoto, “Emergence of self-similarity in football dynamics,” *The European Physical Journal B*, vol. 87, no. 2, pp. 1-6, 2014 .
- Takeuchi, J., and Yamanishi, K., “A Unifying Framework for Detecting Outliers and Change Points from Time Series,” *IEEE Trans. on Knowledge and Data Engineering*, 18(4), pp.482-492, 2006.
- 山西健司「データマイニングによる異常検知」共立出版、2009.