DEIM Forum 2017 P4-4

時系列 3 次元グリッドデータからのホットスポットの自動抽出・追跡法の開発
-フェーズドアレイ気象レーダデータによる局地的大雨解析への適用-

林 諒[†] 森 啓太[‡] 本田 理恵[§] 佐藤 晋介[¶] 村田 健史[¶] 村永 和哉[♯] 鵜川 健太郎[♯] 佐々 浩司[§] 村田 文絵[§]

† 高知大学理学部応用理学科〒780-8520 高知県高知市曙町 2-5-1

‡ 高知大学院総合人間自然科学研究科 〒780-8520 高知県高知市曙町 2-5-1

§高知大学自然科学系理学部門 〒780-8520 高知県高知市曙町 2-5-1

¶情報通信研究機構 〒184-8795 東京都小金井市貫井北町 4-2-1

♯株式会社セック 〒158-0097　東京都世田谷区用賀 4-10-1

E-mail: †‡§{b133k195, keitmori, honda}@is.kochi-u.ac.jp,

¶♯ {satoh, ken.murata, muranaga, kentaro.ukawa}@nict.go.jp , §{sassa, fumie}@kochi-u.ac.jp

あらまし 時系列 3 次元グリッドデータに含まれる不特定数のホットスポット領域の検出と追跡を自動的に行うための手法の開発を行った．グリッド点のフィールド値がそのホットスポットの観測確率密度に比例するものと仮定して多変量正規分布の混合モデルで表現し，モデルパラメータを EM アルゴリズムによって推定した．さらに，この分布が時間的に連続性であるとの過程に基づいて，前の時間の初期値を次の時間の初期値にゆらぎをもたせて与えることにより，分裂や消滅も視野に入れながら特定のホットスポットの移動の履歴をとらえることを検討した．この手法を人工的に作成したデータで評価し，さらにフェーズドアレイ気象レーダデータからの局地的大雨時の降水コアの検出と追跡に適用した例について紹介する．

キーワード　ホットスポット，オブジェクト抽出，追跡，3 次元，多変量正規分布，フェーズドアレイ気象レーダ

1. はじめに

近年，諸分野で大量のデータがアーカイブされ，これらに機械学習の手法を適用することによって新たな知識やパターンを発見し，価値を生み出すことが期待されるようになっている．こうしたデータは従来，テーブルで表されるものが主流であったが，監視カメラ画像，地球観測衛星による観測画像，2 次元，3 次元センサによるモニタデータなど，時間と空間で変動する複雑なデータも扱われるようになっている．衛星画像や監視カメラデータは，時間とともに変動する定点観測の”時系列”画像としてアーカイブされることが多く，そのようなデータに対する時空間パターンの発見手法が検討されてきた[1][2][3]が，最近開発されたフェーズドアレイ気象データ[4](図 1)のように，3 次元的にサンプリングしたフィールド値が細かい時間間隔でリアルタイムに蓄積されるケースも現れるようになってきた．またさまざまな数値シミュレーションの分野でも，このようなデータは数値計算の結果として広く取得されてきた．一方でデータの解釈には，多くの場合スナップショットとしての空間分布を可視化し，その動的変化を動画で視覚的に確認することが一般的であった．

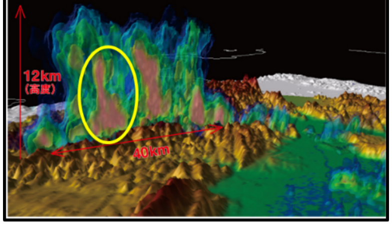

図 1 フェーズドアレイデータに現れた雨雲とオブジェクトの例[4]

こうしたデータには周囲に比べて特異な値をもつ（値が高い，あるいは低い）ホット(コールド)スポットと呼ばれる領域が存在する．ホットスポットは誕生後，一定の生存期間中に発展し，その後消滅していく．このホットスポットを“オブジェクト“として解析することにより，オブジェクトの特徴の変遷として，スカラ・データを要約することができる．気象レーダの例であれば雨雲の追跡や豪雨の前兆パターンなどの新たな知識発見や予測へ応用できると考えられる．

このような問題に対して時系列画像から 2 次元のオブジェクトを抽出・追跡する手法が検討されてきた[1][2][3]．ここでは，オブジェクトを多変量正規分布の混合分布としてモデル化し，最適な成分数をベイズの情報量基準（Bayes Information criteria: BIC）で求めている．また，時間的な連続性に着目して，１つ前の時間の解にラベルをつけて，成分の追加，削除などを行って揺らぎをもたせた複数の初期値を次の時間のモデリングの初期値に用いることによって，成分毎の追跡・分裂・消滅を捉えることを可能にした[2][3]．

本研究では，先行研究[1][2][3]の 2 次元時系列画像に対する混合密度分布推定によるオブジェクト抽出・追跡の手法を 3 次元へと拡張し，さらに原データのフィールド値(レーダ強度，画像の輝度など)を用いることによってオブジェクト抽出の精度向上を検討する．また，この手法を人工的に作成したデータに対して用いて手法の評価を行い，さらに実際に 3 次元気象レーダデータに適用する．

2 章でまずオブジェクト抽出・追跡手法について説明し，3 章ではこの手法を人工データに対し用いて評価，4 章ではこの手法を実際に 3 次元フェーズドアレイ気象レーダのデータに対して実験を行い，その結果について述べる．

2. オブジェクトの抽出・追跡手法

ここで抽出対象とするホットスポットを”オブジェクト”とする．先行研究同様”オブジェクト”は，「生成，移動，分離．消滅などを繰り返しながら時間の経過とともに刻々と姿形を変化させるもの」と定義する．以降，先行研究のオブジェクトのモデルやオブジェクトの抽出・追跡の手法についてレビューしたのち，改良手法について述べる．

2.1. オブジェクトのモデルとパラメータ推定法

原データは 3 次元グリッド上に存在するデータ$I(x_i, y_i, z_i)$として，先行研究[1][2][3]のモデルパラメータ手法の 3 次元的な展開について述べる．原データ$I(x_i, y_i, z_i)$に対してある閾値I_{th}を超えた”座標” d_iをオブジェクトの存在する点の座標集合Dとしてサンプリングするものとする．

$$D = \{d_i | i = 1,2,3,\cdots,n, d_i \in R^M\} \tag{1}$$

$$d_i = \{\left(x_i, y_i, z_i\right)^T | I(x_i, y_i, z_i) > I_{th})\} \tag{2}$$

ここで 3 次元データを取り扱うため$M = 3$である．図 2 のように，これらの座標の塊で表現されるオブジェクトは多変量正規分布の混合モデルを用いて表されるものとする．M次元データに対する多変量正規分布の一般形は以下のようになる．

$$p(d|\theta) = \frac{1}{\left(\sqrt{2\pi}\right)^M \sqrt{|\Sigma|}} exp\left\{-\frac{1}{2}(d-\mu)^T \Sigma^{-1}(d-\mu)\right\} \tag{3}$$

ここで$\theta = (\mu, \Sigma)$は多変量正規分布のパラメータ，$\mu \in R^M$は正規分布の中心ベクトル（平均値），Σは$M\times M$の分散共分散行列とする．オブジェクトの集合は(3)式の重み付き重ね合わせで表すことができる．

$$P(d) = \sum_{j=1}^{K} \omega_j\, p\left(d\middle|\mu_j, \Sigma_j\right) \tag{4}$$

ω_jは重み係数（合計は 1），K は成分数を表す．よって，多変量正規分布のパラメータは以下のようになる．

$$\mu_j = \begin{bmatrix} \mu_{xj} \\ \mu_{yj} \\ \mu_{zj} \end{bmatrix}, \Sigma_j = \begin{bmatrix} \sigma_{xyj} & \sigma_{xyj} & \sigma_{zxj} \\ \sigma_{yxj} & \sigma_{yj} & \sigma_{yzj} \\ \sigma_{zxj} & \sigma_{zyj} & \sigma_{zj} \end{bmatrix} \tag{5}$$

このような仮定のもと，オブジェクトの抽出問題は，Dからのモデルパラメータ$\{\mu_j, \Sigma_j, \omega_j | j = 1,\ldots.K\}$の推定問題に置き換えられる．

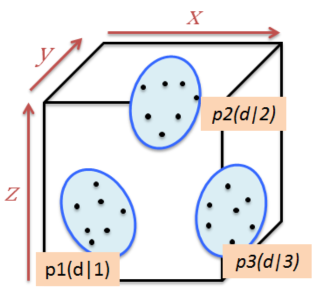

図 2 オブジェクトの混合モデルによる近似

ここで観測値$D = \{d_1 \cdots, d_n\}$はそれぞれどのオブジェクト（多変量正規分布）から発生したものかが未知，すなわち不完全データであるため最尤推定などの手法で直接パラメータを求めることができない．このような場合に有効な手法がEM アルゴリズムである[5]．EM アルゴリズムは，上記のような問題に対して，繰り返しによって逐次モデルの対数尤度を増加させるようにパラメータを推定する手法である．即ち，既に求められたパラメータθから，モデルの対数尤度を増加させるような新しいパラメータθ'を推定するという操作を繰り返すことによって対数尤度$L(\theta) = log \prod_{i=1}^{n} P(d_i|\theta) = \sum_{i=1}^{n} \log P(d_i|\theta)$を最大化するパラメータ$\theta'$を求める．

図 3 にこのようにして得られる多変量正規分布のパラメータの EM アルゴリズムによる計算手順を示す．ここで$z(i,j)$は寄与率であり，観測値d_iがどの成分から発生したかを示す推定量となっている．この値を元に各パラメータの対数尤度を最大にする値を求めるようになっている．

1. $\theta = \{\theta_j = (\mu_j, \Sigma_j, \omega_j) | j = 1,2,\cdots,K\}$をランダムに初期化する．
2. E-Step：
$$z'(i,j) = \frac{\omega_j P(d_i|\theta_j)}{\sum_{j=1}^{K} \omega_j P(d_i|\theta_j)} \tag{6}$$
3. M-Step：
$$\omega_j' = \frac{\sum_{i=1}^{n} z'(i,j)}{n} \tag{7}$$
$$\mu_j' = \frac{\sum_{i=1}^{n} z'(i,j)\, d_i}{\sum_{i=1}^{n} z'(i,j)} \tag{8}$$
$$\Sigma_j' = \frac{\sum_{i=1}^{n} z_j'(i,j)\,(d_i - \mu_j)(d_i - \mu_j)^T}{\sum_{i=1}^{n} z'(i,j)} \tag{9}$$
4. $z(i,j)$が収束しなければ，$\omega_j = \omega_j', \mu_j = \mu_j', \Sigma_j = \Sigma_j'$として 2 に戻る．収束すれば終了する．

図 3 座標に対する EM アルゴリズムによるパラメータ推定

2.2. 観測値の重みを用いたパラメータ推定法

2.1 で述べた通り，先行研究で対象としたデータは，指定した閾値以上の点の座標のみで，オブジェクトの有無のみを表すものとなっていた．しかし，元の観測値自体オブジェクトの中心に向かってピーク状に増大し，値そのものが多変量正規分布で近似可能なものであることも多い．このようなグリッドデータに閾値によるサンプリング処理を行うと，本来観測値が所持していた情報が失われ，観測頻度も一律化されて（1 グリッド点は 1 回しかサンプリングされない）しまい，モデルとして仮定した多変量正規分布からかけ離れてしまう．

図 4 にこの影響を概念的に示す．図 4 の左上のような分布をもったグリッドデータに対して，閾値をこえた座標値のみを取り出すと図 4 の右上のような分布となり，データが本来持っていた情報が失われてオブジェクトの抽出の精度が悪化してしまうと考えられる．また図 4 の下のケースでは，本来 2 つのピークによって示唆されていた 2 つのオブジェクトが，閾値処理によって 1 つに融合してしまう．また，正規分布としては不自然な分布なので逆により多い成分の重ね合わせとして表現しようとする可能性もある．このように閾値処理を行ってデータの座標だけを取り出すことによってオブジェクトの抽出の精度が悪化することが考えられる．よってこの手法を本来の観測値の強度を用いて重み付けできる手法に改良する．

ここで観測値から閾値のオフセット処理をした値を下記のように定義する．

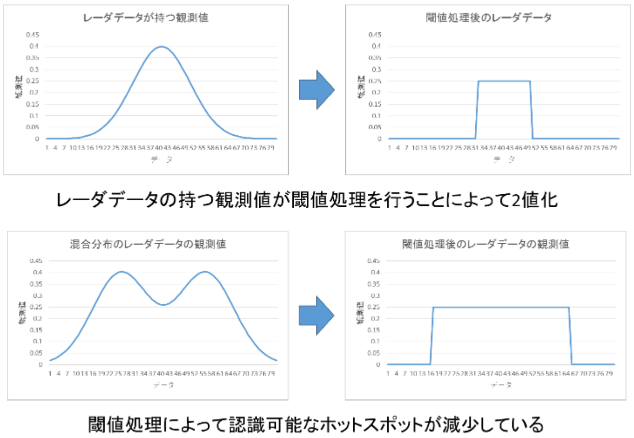

図 4 閾値処理によるモデリングへの影響

$$\delta I_i = I(x_i, y_i, z_i) - I_{th} \tag{1}$$

このオフセット処理はリアルデータに存在するバックグラウンドノイズの除去に必要な処理である．このδI_iが観測確率密度，すなわち観測頻度に比例すると仮定する．図 3 の EM アルゴリズムの M ステップの(7)-(9)式における$\sum_{i=1}^{n} z'(i,j)$を$\sum_{i=1}^{n} z'(i,j)\delta I_i$で置き換えることによって，元データが持っていたフィールド値の情報を観測頻度として利用することができるようになる．この方針に基づいた図 3 の M ステップにおける(7)-(9)式の変更を図 5 に示す．この変更による計算量の増加はなく，同等の計算時間で実行出来るものと考えられる．次章から，両手法を人工データや実データに対して適用し検証することにする．

$$\omega_j' = \frac{\sum_{i=1}^{n} z'(i,j)\,\delta I_i}{\sum_{i=1}^{n} \delta I_i} \tag{2}$$

$$\mu_j' = \frac{\sum_{i=1}^{n} z'(i,j)\, d_i \delta I_i}{\sum_{i=1}^{n} z'(i,j) \delta I_i} \tag{3}$$

$$\Sigma_j' = \frac{\sum_{i=1}^{n} z'(i,j) \left(d_i - \mu_j\right) \left(d_i - \mu_j\right)^T \delta I_i}{\sum_{i=1}^{n} z'(i,j)\, \delta I_i} \tag{4}$$

図 5 原データのフィールド値による重み付けを導入した場合の図 3 の M ステップの(7),(8),(9)式の変更部分

なお，EM アルゴリズムには，大局的最適解を保障せず初期値の選択によって結果が異なる問題がある．また含まれる成分 K をあらかじめ知ることもできない．これらの問題の解決のため，先行研究同様，初期条件では複数の成分数 *Kmin~Kmax* をそれぞれ複数回試行した中から，BIC（ベイズ情報量規準[6]）に基づいて最適なモデルを選択する手法をとる．BIC は次式によって求められる．

$$BIC = -2nL(\theta) + Flogn \tag{5}$$

ここで，$L(\theta)$は対数尤度，Fは自由度，nはデータ数である．3 次元多変量正規分布の場合においては，平均値 3，分散共分散 5（対称性考慮によって 4 成分削減），重み係数 1 の計 10 個のパラメータが存在するが，重み係数の拘束（和が 1）を考慮して，成分数Kに対して，自由度は$10K - 1$となる．

2.3. オブジェクトの追跡

今回取り扱うのは時間とともに連続的に変化するデータであるため，含まれるオブジェクトの分布にも時間的に連続であると仮定できる．オブジェクトはある時点で誕生したのち，移動，変形しながら一定期間存続し，この間，分裂や融合も経験しながら消滅にいたると考えられるため,この効果も取り入れられることが必要である．

これらを考慮したオブジェクトの追跡手法として，前の時刻の解を次の時刻の解に引き継ぐ手法[2][3]が提案されている．図 7 にその概念図を示す．最初の時刻$t = 0$では 2.2.で述べた Kmin〜Kmax 個の K-Means 法で求めた初期値を与え総当り試行で最適解を求めるが，次の時間tからは前の時間$t-1$の解に揺らぎをあたえた複数の解候補を次の時間の初期値として利用する．この際，オブジェクト数も時間の変化とともに変化すると考えられるため，成分数に± 2個の揺らぎを持たせる(ω_jの小さい順に成分を削除，またはω_jの大きい順に解を複製)．また，隣りあう時間のオブジェクトの形状変化が激しくなると解が求められないケースが発生するため，この場合には一旦解のリセットを行い，再度 K-Means 法を用いることで初期値を求め直してここから追跡を再開する．本研究の実験でもこの手法をそのまま用いることにする．

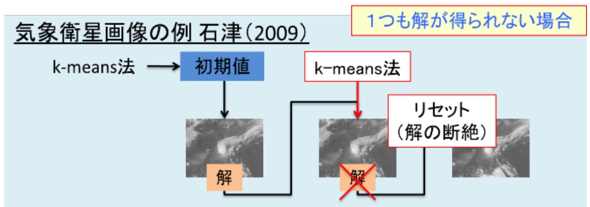

図 6 解の連続性への対策

1	時刻 t=0：	
	1.1	初期値は Kmin~Kmax 個のそれぞれについて K-Means 法によって求める．
	1.2	全ての初期値について EM アルゴリズムでモデルを求める．（複数回思考して BIC 最大の解を求める.）
2	時刻 t=1 以降：	
	2.1	前の時刻の解を元にして，成分を増減（±2 個）した結果を初期値の候補とする．
	2.2	全てについて EM アルゴリズムで海を求め，BIC を計算し，最適解を選択する．時刻を勧め 2.1 へ戻る．
	2.3	どのケースでもパラメータが求まらなかった場合，追跡を中止してリセットし，1.1 へ戻る．

図 7 時間方向への解の追跡手順

3. 人工データを用いた手法の評価

3.1.新・旧手法の比較

前章でのモデルパラメータ決定手法を，人工的に作成したデータに対して用いて実験を行った．人工データは 480×480×45グリッドの範囲のデータを使用し，成分数やオブジェクトの重なり具合などが異なる 8 つのデータを作成した．全ケースの成分数は表 1 に，その一部の分布を図 8，9 に示す．この図は 3 次元データを俯瞰的に示したものとなる．Case1 は縮退のチェックを意図したもので，case2 は単

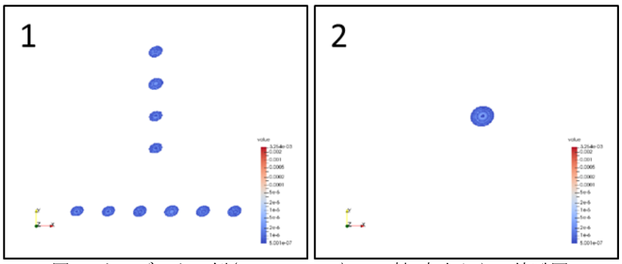

図 8 人工データの例(case1, case2)，Z 軸+方向からの俯瞰図.

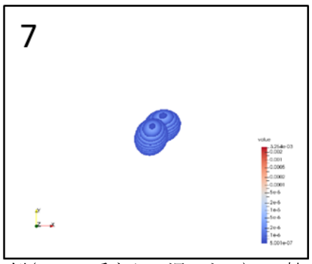

図 9 人工データの例(case7:重なりの深いもの). Z 軸+方向からの俯瞰図.

成分の場合である．また case7 と case 8 ではオブジェクトが重なるようにしている．

手法の評価についてはオブジェクトの成分数とピアソン距離を用いて行なった．ピアソン距離とは確率密度分布間の距離尺度の一つであり，以下の式で計算される．

$$\mathrm{PE}(p||p') = \int p'(x)\left(\frac{p(x)}{p'(x)} - 1\right)^2 dx \qquad (6)$$

式中のpとp'は確率密度関数を表しており，この数値が小さいほどデータ同士の分布が近似しているということになる．今回は元の人工データとオブジェクト抽出結果から再度作成した人工データのピアソン距離を計算し，手法による結果の比較を行った．

両手法の比較の結果も表 1 に示す．観測値を用いない手法を旧手法，観測値を用いた手法を新手法とする．成分数においては，元の人工データと比べて異なった値となったものを赤字で，等しい値となったものを青字で表記している．ピアソン距離では 2 つの手法を比較して結果の良いものを青字，結果の悪いものを赤字で表記している．結果より，原データのフィールド値を用いた手法を用いることで，成分数やピアソン距離において大幅な性能向上を確認することが出来た．しかし，case 1 や case7 においては観測値を用いた手法でも正しくオブジェクトを求められていない．ここで，case 7 では大きくオブジェクトが重なっているため正確なモデルパラメータの推定が難しかったものと考えられる．一方，case 1 での精度の悪化は初期値の決め方によって引き起こされているものであると考え，次の 4.2 で述べる初期値改良を行った．

表 1 人工データからのオブジェクト抽出結果

Case	成分数	オーバラップ	検出成分数		ピアソン距離	
			旧	新	旧	新
1	10		7	7	7.66e-0	3.22e-0
2	1		14	1	4.64e-0	3.19e-4
3	2		2	2	1.89e-0	2.63e-4
4	3		3	3	3.26e-0	1.31e-2
5	4		6	4	3.96e-0	7.13e-4
6	2		6	2	2.09e-0	6.02e-3
7	2	大	30	1	3.21e-0	8.67e-1
8	4	小	26	4	2.25e-0	1.88e-2

3.2.EM アルゴリズムの初期値改良

これまで EM アルゴリズムの初期値には K-Means 法の結果を用いてきた．K-Means 法では初めに各データに対してランダムにクラスタ ID を振り分けた後に，振り分けられたデータをもとに各クラスタ重心を計算する事が多い．この際，データ数が極端に大きいクラスタがあると，その付近に複数のクラスタ重心が求められ，その結果，K-Means 法の計算が局所的最適解に陥り，EM アルゴリズムの初期値として適切でない初期値が受け渡される可能性が出る．

K-Means 法に起因する局所的最適解を避けるため，初期値は基本的にはランダムに与えるものの，互いの近接領域には新しい重心の初期値は設定しない（近接していた場合棄却）制約のもとで決定するように改良した．ただし，初期値設定を禁止する近傍領域の設定には任意性があるため，先見的に求められるデータ中のオブジェクトのサイズなどから決定することが必要である．

実験の結果を図 10 に示す．ここでは，簡易的な可視化のために，各時刻の各オブジェクト点を最も存在する確率が高いオブジェクト ID に以下の式で割り当てる．

$$\mathrm{objectID}(i) = \operatorname*{argmax}_{j} z(i,j) \qquad (1)$$

このオブジェクト ID ごとに閾値を超えたグリッド点に対して異なる色を割り当てる事によって抽出結果を可視化する．初期値改良を行っていない新手法では近接成分の縮退が発生しているのに対し，初期値の改良後は正常に 10 個のオブジェクトが検出されている事が分かる．このように，データ値の重みとしての利用にさらに初期値改良を加えたことでこのようなケースでも妥当な成分数を求めることができるようになったといえる．

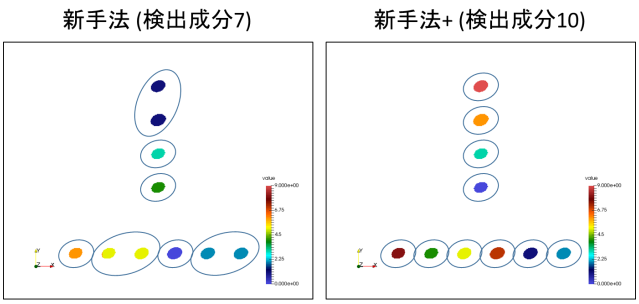

図 10　case1 に対する初期値改良手法を用いたオブジェクト抽出結果．いずれもデータの値を利用する方式（新手法）を用い，初期値は左は従来の k-means 法，右はランダム性を高めた求め方を採用している．各点は最大寄与率を持つオブジェクト ID に応じて色づけしている．

4. 気象レーダデータでの実験

2 章での観測値の有無に関する 2 種類の手法を実際のデータに対して用い，実験を行った．対象としたのは 3 次元フェーズドアレイ気象レーダデータである．

4.1. 対象データ

実験では大阪大学吹田キャンパスに設置されたフェーズドアレイ気象レーダで観測されたレーダ反射強度の 3 次元データを使用した．このデータは 30 秒毎に収集されており，Web 上でも 30 秒毎に更新されるリアルタイムデータを確認することが出来る．これらのデータから今回は 2012 年 7 月 16 時 00 分から 19 時 00 分までに 5 分間隔でサンプリングしたデータを対象とした．表 2 に使用したデータの諸元を，図 11 に観測範囲を示す．

表 2 使用したデータの諸元

項目	値
観測値	レーダ反射強度（dBZ）*
期間, 間隔	2012 年 7 月 16 時 00 分から 19 時 00 分, 5 分間隔
領域	大阪大学吹田キャンパスを中心とした半径 60km 始点（134.867° E, 34.282° N), 高度 0~11.76km
空間解像度	緯度 0.00272765°, 経度 0.00224948°, 高度 0.21km
グリッド数	481×481×45

*降水量に換算すると，30dBZ で 3mm/hr，50dBZ で 50mm/hr 相当

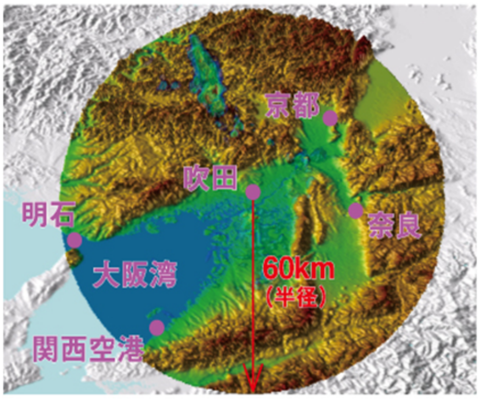

図 11 観測範囲[7]

実験対象のデータが観測された 2012 年 7 月 26 日には，大阪地方で局所的大雨が複数回発生した．これらの豪雨は 3 回にわかれ，まず初めに 16 時 40 分に京都府八幡市及び城陽市に，17 時半頃に大阪府枚方市，京都府京田辺市から精華町に，そして 18 時頃に京都府亀岡市の南部に強い雨をもたらす雲が現れた．これらの豪雨の直前には豪雨の卵であるファーストエコーが発生し，数十分間の間に集中的に雨を降らせた．図 12 に各豪雨のファーストエコーを示す．これらのファーストエコーの強度は 30dBZ 程度である．

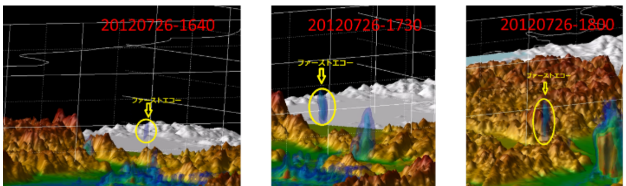

図 12 豪雨前に発生したファーストエコー[7]

図 13 に豪雨の際の典型的なスナップショットを示す．青色の箇所の反射強度は 30dBZ，緑色の箇所は 40dBZ 程度となる．反射強度 30dBZ から 50dBZ 程度で雨雲に相当するオブジェクトが認識できる．実験では閾値として 30-35dBZ を採用した．また予備的な実験として，データのフィールド値（反射強度）による重み付けを行った場合と行わない場合の両者について，全期間について解の継続を行いながらのモデリングを実施した．実験のケース表記については表 3 のとおりである．

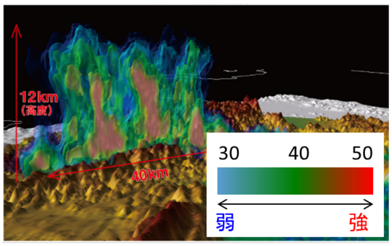

図 13 レーダの反射強度[7]

表 3 検証したケース

Case	閾値(dBZ)	モデリング手法
1	30	座標のみ(旧手法)
2	35	座標のみ(旧手法)
3	30	フィールド値による重み付けあり(新手法)
4	35	フィールド値による重み付けあり(新手法)

4.2.実験結果

まず一部のスナップショットについてモデリングの結果を示す．可視化手法は 2.4 で使用したものと同じである．

図 14 には雨雲が少なく 1 つめの豪雨のファーストエコーの発達時のスナップショットに対応するモデリングの結果を，図 15 は豪雨の発展中のスナップショットを示す．各点は紡錘形の領域毎に色分けされ，多変量正規分布に当たる雲の塊が表現されている．手法による違いはフィールド値による重み付けを行なった手法において，座標のみ用いた手法よりもオブジェクトの検出数が少なくなり，1 つのオブジェクトを複数のオブジェクトに分裂していたものを抑制する効果が見られる．

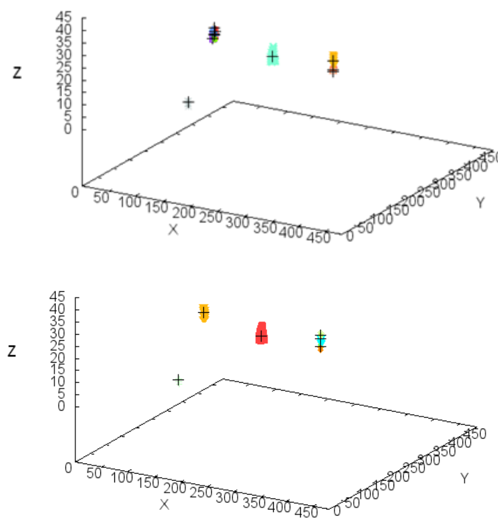

図 14 16 時 30 分のケース 2(上，座標のみ使用(旧手法)，閾値 35dBZ)) ケース 4(下，フィールド値による重み付け(新手法), 閾値 35dBZ)の抽出結果．カラープロットはオブジェクト id 毎に色分けしたオブジェクト点をプロット，黒の十字は抽出されたオブジェクトの中心座標.

また，データのフィールド値が他変量正規分布の値に比例するという仮定についても検証を行った．検証にはオブジェクトの抽出結果である平均，分散，重み係数から混合分布を生成し，これを実際のレーダデータと比較した．図 16 に比較的オブジェクト数の少ない 16 時 40 分の結果を示す．この結果より，レーダデータの反射強度と人工データの確率密度はほぼ比例関係にあり，強度と確率密度が比例するものとし

てこの手法を展開したことは妥当であることが確認できた．

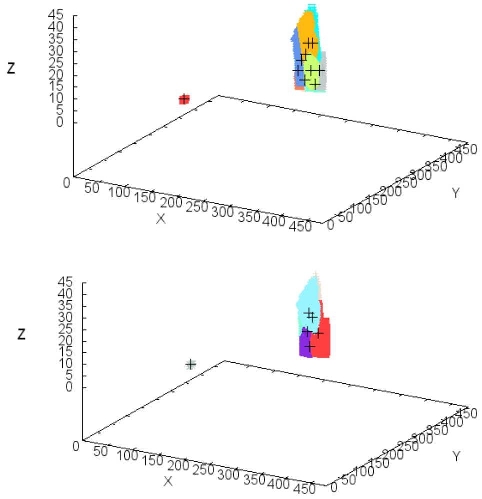

図 15 17 時 45 分のケース 2(上，座標のみ使用(旧手法)，閾値 35dBZ)，ケース 4(下，フィールド値による重み付け(新手法)，閾値 35dBZ)の抽出結果．シンボルは図 14 と同様.

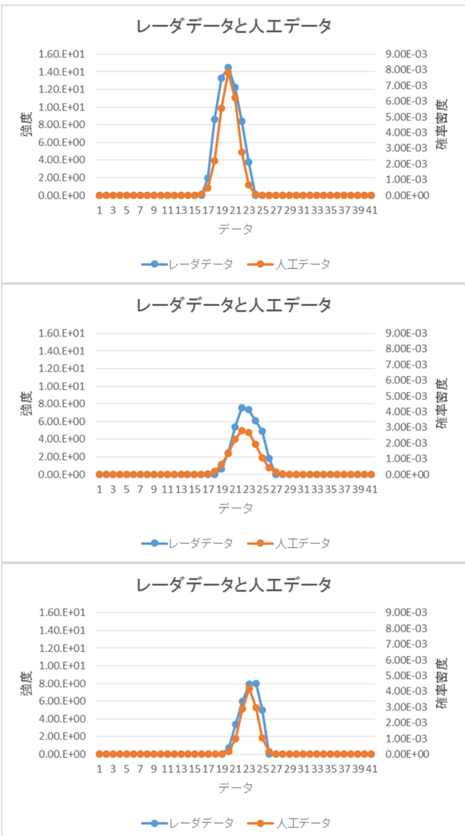

図 16 レーダデータと得られた確率密度分布の比較

最後に，図 17 と図 18 に全てのケースで抽出されたオブジェクト数の時間的推移を豪雨の発生期間とともに示す．縦の実線は専門家によって示されたそれぞれの豪雨のファーストエコーの発生タイミング（赤，青，黄色はそれぞれ第 1 回目，第 2 回目，第 3 回目）であり，同色の長方形のハッチ部分はそれぞれの豪雨の主な継続期間である．30dBZ の結果では豪雨の発生に伴って検出されたオブジェクトの成分が急増し豪雨の収束とともに減少する様子が観測されにくく，35dBZ の結果ではこれが観測しやすいことから，一般的には閾値 35dBZ が望ましいようである．図 18 の結果から，新しい手法では，1 度目の豪雨のファーストエコーが現れた直後の 16 時 40 分以降は強度を用いない手法と比べて成分数が小さくなっていることが確認できる．これは前述したとおり強度を用いたことによる分裂の抑制の効果と考えられる．一方でそれまではむしろ新しい手法の方が大きな成分数が求められていることについては今後さらに追跡過程の精密化の検討も含めて考察を行う必要が有る．

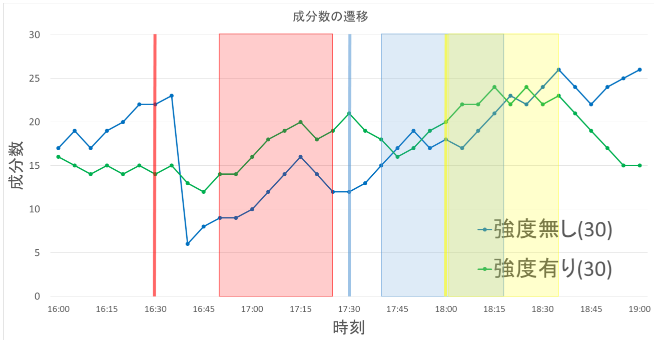

図 17 ケース 1,3 に対して抽出されたオブジェクト数の時間的推移．縦の実線はそれぞれの豪雨のファーストエコー，ハッチ領域は豪雨の主な期間を示し，赤，青，黄色はそれぞれ第 1 回目，第 2 回目，第 3 回目を示す．

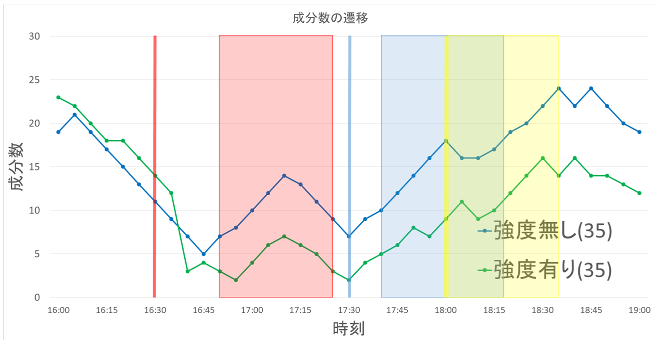

図 18 ケース 1, 3 に対して抽出されたオブジェクト数の時間的推移．他説明は図 17 同様.

5.おわりに

時系列 3 次元グリッドデータからのホットスポットに相当するオブジェクトの自動抽出・追跡法を開発した．この際，2 次元時系列気象画像に適用した手法を 3 次元に拡張し，オブジェクト集合を 3 次元の多変量正規分布の混合密度分布でモデル化して EM アルゴリズムによってパラメータを求める手法をとった．解の追跡には前の時間の解を次回の初期値として揺らぎを持たせて与える手法を採用した．さらに閾値処理後のデータのフィールド値による重み付けを導入することにより，オブジェクトの中心や成分数の決定精度をより向上させることを試み，EM アルゴリズムの初期値に対しても改良を試みた．この手法を人工データに対して評価実験を行い，分裂や縮退をおさえてオブジェクト抽出精度の向上させる効果があることを確認することが出来た．またフェーズドアレイ気象データにも適用して予備的実験を行った．実験の結果，強度を用いることにより，人工データで観察されたのと同様に分裂を抑制することを確認することが出来た．今後，さらに解の追跡精度の向上についても引き続き検討を行い，フェーズドアレイ気象レーダに対して本格的な実験と検証を行い，豪雨の前兆現象の学習や自動的な検出，予測にむけて検討を行う．

参　考　文　献

1. Honda, R., Wang, S., Kikuchi, T. and Konishi, O., Mining of objects from time-series images and its application to satellite weather imagery, Journal of Intelligent Information Science, 19:1, 79-93, 2002
2. 松永知也, 本田理恵, 時系列画像からのオブジェクトベースデータマイニング -オブジェクトの抽出とデータベース化-, DEIM フォーラム 2015, P2-4, pp. 6, 2015
3. 松永知也, 森啓太，本田理恵，時系列画像に含まれるオブジェクト特徴の変遷要約とその可視化, DEIM フォーラム 2016, P1-2, 2016
4. 佐藤晋助，牛尾知雄，水谷文彦，フェーズドアレイ気象レーダの研究開発，NICT News, 2013.1，p.3-5，2013．
5. Dempster, A. P., Laird, N. M., and Rubin, D. B., Maximum likelihood from incompletedata via the EM algorithm. Jornal of the royal statistical society. Series B (methodological), 1-38, 1977.
6. MacQueen, J., Some methods for classification and analysis of multivariate observations. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability (Vol. 1, No. 14), 281-297, 1967, June.
7. National Institute of Information and Communications Technology，日本初「フェーズドアレイ気象レーダ」を開発，http:www.nict.go.jp/press/2012/08/31-1.html，2012．