

## 解 説

# ベイズ分類を用いた二重偏波気象レーダーによる降水種別判別 — クラスター分析による教師データの作成 — Hydrometeor classification algorithm for dual polarization weather radar using Bayes classifier — An approach for fast training data creation by cluster analysis —

小池 哲司<sup>1</sup>

### 要 旨

本報告は平成 28・29 年度の観測部観測技術開発課題として実施し、二重偏波気象レーダーの利用技術としての降水種別判別のアルゴリズムの開発について記したものである。降水種別判別アルゴリズムは本手法を元に平成 30 年度以降も改良が続けられている。

平成 27 年度末から二重偏波化の運用が開始された空港気象ドップラーレーダー (DRAW) のデータを元に教師データを作成し、ベイズ分類により雨・雪・湿雪・雹・霰などの降水種別への判別を行った。判別の教師データは、クラスター分析による手法により効率良く作成している。

降水種別判別の活用例として雹の判別域とレーダー雷解析及び LIDEN データとの比較を行い、雹判別域は発雷域とよく対応しており、発雷予測への活用に有効であることを示すことができた。

### 1. はじめに

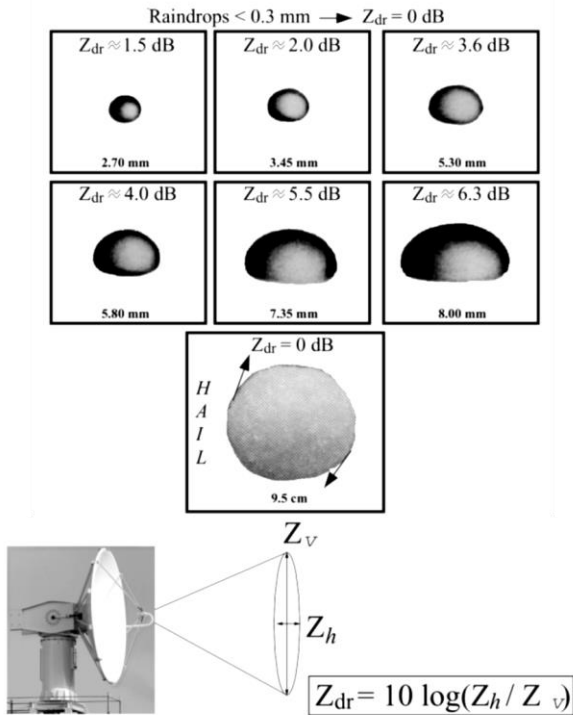
平成 28 年 3 月から空港気象ドップラーレーダー (DRAW) に二重偏波気象レーダーが導入され、令和 2 年 3 月から一般気象レーダーにも二重偏波気象レーダーの導入が始まった。二重偏波気象レーダーは水平と垂直の異なる偏波の電波を放射し、これまでの気象レーダーと比べて、降水粒子の扁平度と均一性に関する情報を新たに観測することが可能であり、そこから降水種別を推定することができる。第 1.1 図に雨滴及び雹の図を示

す。雨滴 (raindrops) は径が大きくなるにつれて空気抵抗により形が扁平する。一方、雹 (hail) は径が大きくなっても丸く、雪も固形のためほとんど扁平は無い。また融解層 (ブライトバンド) は、均一性に関する二重偏波情報を用いることにより判別が可能である。

### 2. DRAW における二重偏波情報の観測

DRAW は C バンドの気象レーダーであり、5 分ごとに複数仰角の PPI 走査 (定仰角) と 1 方位角の

<sup>1</sup> 元観測部観測課観測システム運用室 (提出時 大阪管区気象台気象防災部地震火山課)  
(令和 4 年 2 月 14 日発行)



第 1.1 図 降水粒子（雨滴と雹）の大きさとその扁平の様子（Bringi and Chandrasekar, 2001）

RHI 走査(定方位角)の観測を行う。PPI の仰角のうち最低仰角(0.7°)の観測は 1 分ごとに実施する。本報告では PPI のデータのみを取り扱っており、レーダーの画像もレーダーサイト近傍では観測高度は地表に近く、サイトから離れるにつれ観測高度が高くなる。PPI データは観測範囲 120 km(距離分解能 150 m), 方位分解能 0.7°の 800×512 のデータセットである。

DRAW の二重偏波情報の観測例を第 2.1 図に示す。水平偏波反射強度( $Z_{hh}$ )はこれまでのレーダー反射強度と同じである。反射因子差( $Z_{dr}$ )と偏波間位相差変化率( $K_{dp}$ )は降水粒子の扁平率に関する量であるが、特性が異なる(付録 1 参照)。偏波間相関係数( $\rho_{hv}$ )は観測する降水粒子の均一性を表している。注意点として、強雨域のレーダーサイト側からの背面(白点線内)は降雨減衰のため $Z_{dr}$ が過小になっており、補正が必要である。また、第 2.1 図の $\rho_{hv}$ は多くの領域で通常の雨のためマゼンダ( $\rho_{hv} \approx 1$ )であるが、レーダーサイト近傍では $\rho_{hv}$ が低い。MTI 処理(Moving Target Indication, ドップラー速度が 0 の成分を除去することによりク

ラッターを抑圧する)により $Z_{hh}$ はクラッターの影響を除去しているが、 $\rho_{hv}$ には影響が残っており、観測高度が低いところで $\rho_{hv}$ が低いのはこのためで、教師データの作成には留意が必要である。

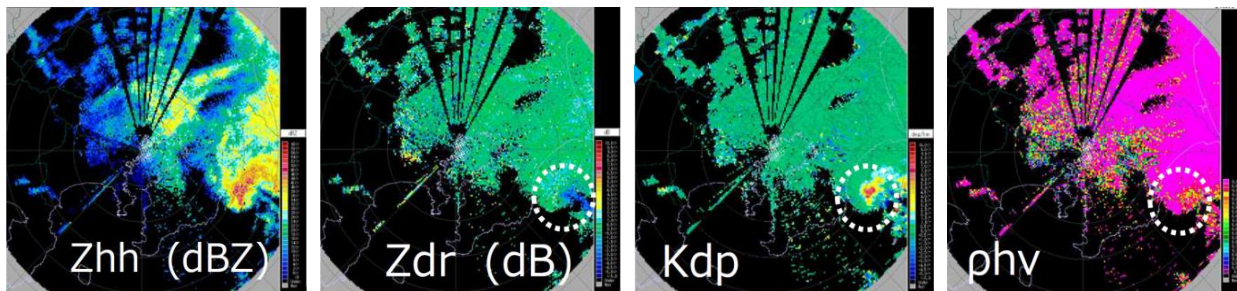
### 3. 二重偏波情報による降水種別の分類

二重偏波情報による降水種別分析の分類方法として、初期の研究ではファジィ論理による手法が良く使われていた。ファジィ論理による方法は二重偏波情報や気温の一変数ごとに対応する降水粒子の存在範囲を決める手法であり、その存在範囲はメンバーシップ関数と呼び、この作成は比較的容易である。第 3.1 表に Keenan(2003)のファジィ論理による手法の際に用いる二重偏波情報の範囲を示す。

ファジィ論理による方法のほかにベイズ分類による方法もよく用いられている。ファジィ論理の方法が一変数の手法であるのに対し、ベイズ分類では多変数での確率密度分布を用いる手法である。第 3.1 図に Marzano *et al.*(2008)によるベイズ分類で用いる各降水粒子の確率密度分布の例を示す。Marzano *et al.*(2008)の手法は $[Z_{hh}, Z_{dr}, K_{dp}, T]$ の 4 次元ガウス分布による確率密度分布モデルを用いている。ベイズ分類はファジィ論理よりも判別精度が良いが、多次元確率密度分布モデルの作成はファジィ論理のメンバーシップ関数の作成よりも作業量はかなり多い。

ベイズ分類は機械学習の「教師あり学習」にあたり(第 3.2 図)、分類のための教師データを必要とする。

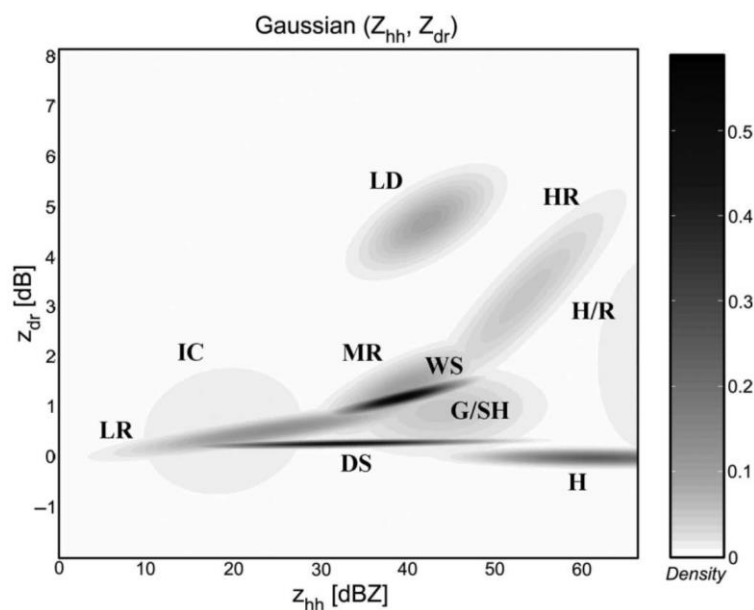
これまでの二重偏波情報を用いた降水種別判別における教師データの作成は、光学式ディストロメーターなどの地上観測やビデオゾンデなどの高層観測との比較で作成していたために、教師データを作成するためのサンプル数が少なく緻密な教師データの作成は難しかった。またレーダーの周波数だけでなく製造型が異なっても二重偏波情報は特性が異なり、教師データをそのまま流用することは出来ない。これは二重偏波情報による降水種別判別の開発を難しくしている一因である。



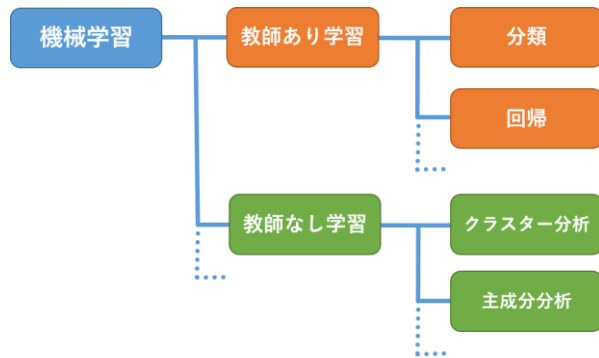
第 2.1 図 羽田 DRAW での二重偏波情報の観測例 (2016 年 9 月 22 日 13:35 (0435UTC)). 円点線は発達した積乱雲による強雨域でメソサイクロンを伴っている. 降雨減衰により, 強雨域のレーダーサイトからの背面は  $Z_{dr}$  はマイナスになり,  $\rho_{hv}$  は  $S/N$  が悪くなり 1 より低くなっている. (各二重偏波情報については付録 1 に解説)

第 3.1 表 ファジィ論理による各降水種別に対する二重偏波情報のとり得る範囲 Keenan (2003) より一部要素を抜粋.

	$Z_{HH}$ (dBZ)	$Z_{DR}$ (dB)	$\rho_{HV}(0)$	$K_{DP}$ (deg km <sup>-1</sup> )	Temperature (°C)
Drizzle	10-25	0.2 to 0.7	>0.97	0 to 0.06	>-10
Rain	25 to 60	0.5 to 4	>0.95	0 to 20	>-10
Snow (dry, low density)	-10 to 35	-0.5 to 0.5	>0.95	-1 to 1	< 0
Snow* (dry, high density)	-10 to 35	0.0 to 1	>0.95	0 to 0.4	<0
Snow (wet, melting)	20 to 45	0.5 to 3	0.5 to 0.9	0 to 1	0 to 5
Graupel, dry	20 to 35	-0.5 to 1	>0.95	0 to 1	<0
Graupel, wet	30 to 50	-0.5 to 2	>0.95	0 to 3	-15 to 5
Hail, small < 2 cm wet	50 to 60	-0.5 to 0.5	0.92 to 0.95	-1 to 1	-15 to 20
Hail, large > 2 cm wet	55 to 65	-1 to 0.5	0.90 to 0.92	-1 to 2	-25 to 20
Rain and hail	45 to 80	-1 to 6	>0.9	0 to 20	-10 to 25



第 3.1 図 Marzano *et al.* (2008) による各降水粒子の  $(Z_{hh}, Z_{dr})$  平面上における確率密度分布モデル (LR : light rain, MR : middle rain, HR : heavy rain, G/SH : graupel/small hail, H/R : hail/rain, H : hail, DS : dry snow, WS : wet snow, IC : ice crystal, LD : large drops)



第 3.2 図 機械学習の系統分類

この教師データの作成のために、レーダーデータ以外の比較観測からではなく、二重偏波気象レーダーデータの「教師なし学習」から教師データの作成を行った研究がある。

Grazioli *et al.* (2015) は X バンド二重偏波気象レーダーの観測データを高度なクラスター分析的手法により、観測データ側から教師データの作成を行い、その報告では、雨・雪・融解層を良く分離できているように見える。しかし、DRAW データを用いて一般的なクラスター分析により調査したが、クラスター分析の手法が違うためか融解層を分離できるような解析は再現できなかった。

Wen *et al.* (2015) は観測データからクラスター分析や混合ガウスモデル (GMM : Gaussian Mixture Model) により代表値を抽出し、さらに非階層的クラスター分析により教師データを作成している。この手法は DRAW への適用も可能であるが、これらの作業が事例ごとに手作業で行われており、教師データのサンプル数がかなり少ないのが難点である。そこで、手作業をなるべく少なくする手法を考案することにした。

#### 4. 本報告による教師データの作成手法

本報告では以下の方法により、ベイズ分類の教師データとなるような降水粒子の確率密度分布モデルを作成した。詳細は付録 2 に示すが、概略は以下のとおりである。

- (a) DRAW で観測した二重偏波情報のデータセットから GMM により代表値を抽出する。
- (b) (a) で作成した代表値を非階層的クラスター分析である K-means 法により複数 (10) の領

域に分割する。

- (c) (b) で作成した領域について、他資料を参考にしながらその領域に対応する降水種別を同定する。

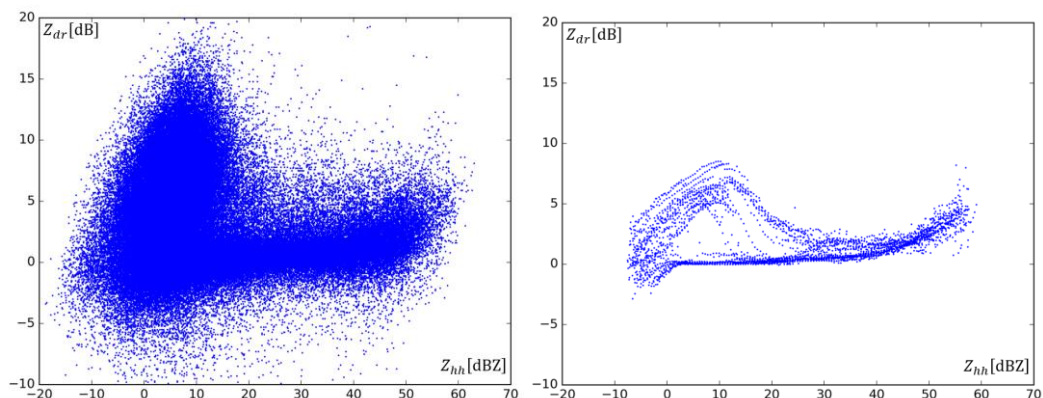
- (d) (c) で同定した領域についてカーネル密度推定 (Kernel Density Estimation : KDE) により確率密度分布を求める。

(a) の GMM は複数のガウス分布の当てはめを行うもので、クラスター分析の一種にも数えられる。本手法ではそのガウス分布のピーク位置をその分布に従うデータの代表値としている。GMM は EM アルゴリズム (反復法による最尤推定法の一つ) を用いているためピーク精度の安定性が良いものの、収束は遅く計算には時間がかかる。当てはめるガウス分布の個数はベイズ情報量基準 (Bayesian Information Criterion : BIC) で決める。第 4.1 図にその代表値の抽出例を示す。左はエコーデータから算出した  $Z_{hh} - Z_{dr}$  プロットで、右はそれに対し GMM により抽出した代表値である。

$Z_{dr} \leq 0$  [dB],  $Z_{hh} = 0 \sim 40$  [dBZ] の雨や  $Z_{dr} > 0$  [dB],  $Z_{hh} > 40$  [dBZ] の強雨,  $10 > Z_{dr} > 0$  [dB],  $Z_{hh} < 20$  [dBZ] の非降水エコーの特徴を良くとらえたプロットを描いている。

(b) では教師なし学習として非階層的クラスター分析を使用した。一般的にクラスター分析と言えば樹形図 (デンドログラム) に代表される階層的クラスター分析を思い浮かべることが多いが、階層的クラスター分析は大規模な計算に向いていないため、非階層的クラスター分析、その中でも大規模データの処理に適した K-means 法を採用した。非階層的クラスター分析は事前に分割するクラスター数を決めておく必要があり、本手法では 10 に固定している。クラスター数が少ない場合は異なる種類の降水粒子のデータが一つのクラスター内に入る可能性が高くなり、クラスター数が多い場合は逆に一つの降水種別の分布を作成するのに多くのクラスターを合成して作成することになり作業の手間になる。試行錯誤していくうちにクラスター数が 10 に落ち着いた。

(c) では分類したクラスターがどの降水種別に対応するか地上観測や他論文等を参考に同定を



第 4.1 図 2016 年 7 月 14 日 0455UTC の羽田 DRAW の  $Z_{hh} - Z_{dr}$  プロット (ただし  $0^{\circ}\text{C}$  面から 1 km より下の雨域のデータ)。(左) 元データ (右) GMM による代表値抽出後。

行った。ただし、周波数帯 (S, C, X) が異なれば  $Z_{dr}$  や  $K_{dp}$  の特性が大きく異なり、機器が異なれば  $\rho_{hv}$  の特性も異なる。また参考にした論文として C バンドの Marzano *et al.* (2008) や Al-Sakka *et al.* (2013) を挙げる。Al-Sakka *et al.* (2013) の題名に Fuzzy がついているが、教師データは  $Z_{hh} - Z_{dr}$ ,  $Z_{hh} - K_{dp}$ ,  $Z_{hh} - \rho_{hv}$  の各二重偏波情報の変数間関係を 2 次元でモデル化しており、本報告の手法に近い。Marzano *et al.* (2008) と Al-Sakka *et al.* (2013) は教師データを T-matrix 法 (Mishchenko and Travis, 1998) により作成している。T-matrix 法は各降水粒子の粒径分布モデルを与え、二重偏波情報の数値計算を行う手法である。計算コードは公開されており、DRAW の周波数帯でも計算して確認している。T-matrix 法の結果と DRAW のデータから (a) の手法により得た代表値を比較すると、DRAW データの方が強雨 ( $Z_{hh} > 40[\text{dBZ}]$ ) で  $K_{dp}$  が小さい傾向にある。そのため降水粒子の同定は、他論文をそのまま参考にするにはせず、降水粒子間の二重偏波情報の関係の傾向を見て行っている。

### 5. 降水種別判別で用いるベイズ分類

本報告ではベイズ分類を用いて降水種別判別を行うが、ベイズ分類に使う変数は、二重偏波情報 ( $Z_{hh}, Z_{dr}, K_{dp}, \rho_{hv}$ ) とそのテクスチャ解析 (第 6.3 節で解説,  $(\sigma(Z_{dr}), \sigma(\rho_{hv}), \sigma(\psi_{dp}))$ ) 及び各二重偏波データの観測高度と  $0^{\circ}\text{C}$  面高度の予報値の偏差

$\Delta z_0$  である。現状では  $0^{\circ}\text{C}$  面高度は数値予報から算出する。変数

$$\mathbf{x} = [Z_{hh}, Z_{dr}, K_{dp}, \rho_{hv}, \sigma(Z_{dr}), \sigma(\rho_{hv}), \sigma(\psi_{dp}), \Delta z_0]$$

を観測したとき、これが降水種別  $C_i$  である確率  $P(C_i|\mathbf{x})$  はベイズの定理により以下のように書くことができる。

$$P(C_i|\mathbf{x}) \propto P(\mathbf{x}|C_i) \times P(C_i)$$

ここで  $P(C_i|\mathbf{x})$  は事後確率、 $P(\mathbf{x}|C_i)$  は尤度関数、 $P(C_i)$  は事前確率と呼ばれる。尤度関数は降水種別  $C_i$  で  $\mathbf{x}$  となりうる確率で、教師データの確率密度分布そのものである。ベイズ分類での降水種別判別は、判別する降水粒子ごとに右辺の値を計算し最大のものに対応する降水粒子とするものである。

尤度関数、すなわち教師データの確率密度分布モデルの作成は Grazioli *et al.* (2015) のように多次元のガウス分布モデルにすると簡単に計算できるが、この確率密度分布がガウス分布で近似可能かは分からない。少なくとも  $Z_{hh} - K_{dp}$  の関係は非線形に変化するため難しいと考える。そのため、確率密度分布のモデルはグリッドモデルによる作成を行った。 $P(\mathbf{x}|C_i)$  は  $\mathbf{x}$  が 8 変数のため、グリッドモデルを作成する際には 8 次元のグリッドが必要である。しかし、高次元グリッドは次元数が増えるほど大量のメモリが必要なため、厳密な手法ではないが

$$P(\mathbf{x}|C_i) \sim \prod_j P(Z_{hh}, X_j|C_i)$$

とし、各変数と  $Z_{hh}$  の 2 次元へ射影させ、最終的な



確率分布はそれを掛け合わせることで求めることとした。また、 $\Delta z_0$ には数値予報の予測誤差が含まれるため、確率密度分布モデルを作成した時点で既に誤差を含んでいる。そのため尤度関数に $\Delta z_0$ を含めず、 $\Delta z_0$ を考慮した事前確率で対応している。

6. ベイズ分類のための確率密度分布モデル  
や事前確率の作成方法

6.1 K-means 法によるクラスター分析とクラ  
スターの降水種別同定

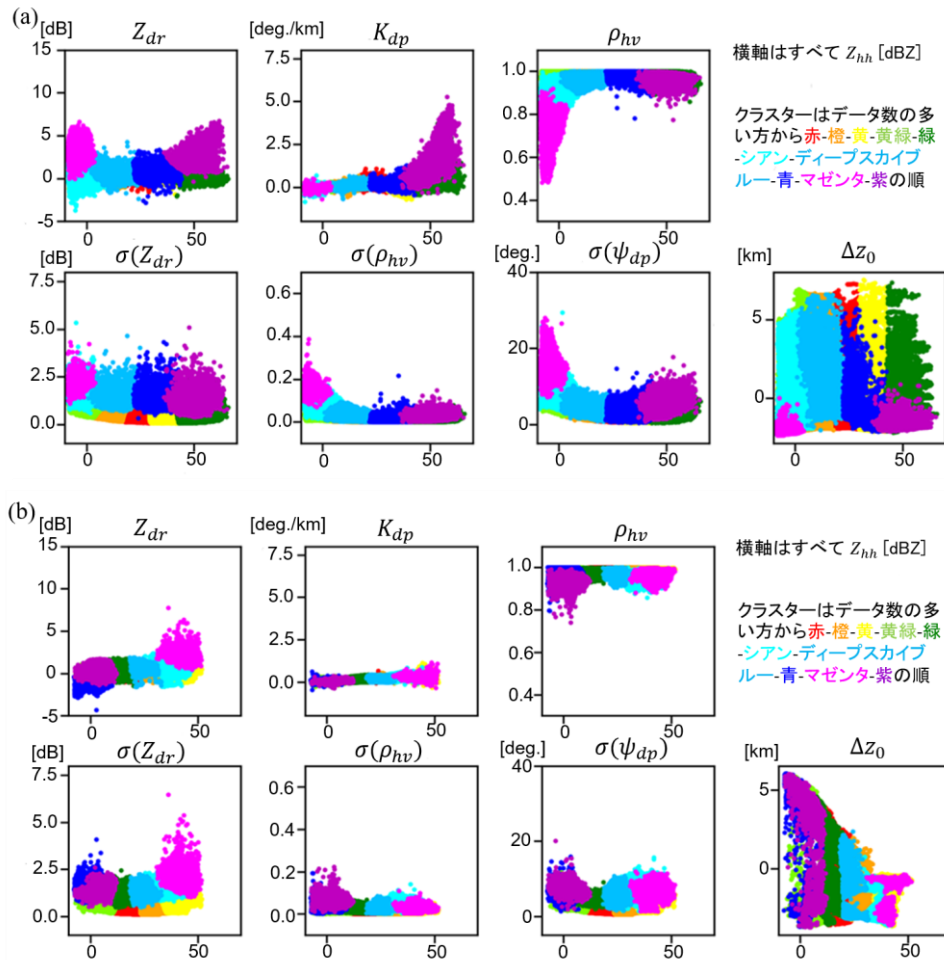
K-means 法によるクラスター分析は、第 4 章(a)の GMM により抽出した代表値を 1-3 時間程度足し合わせて解析するデータとした。層状性降水などではデータ数が多すぎて K-means 法でも取り扱

うことが難しくなるため 1 時間程度にし、不安定性降水等ではデータ数が少ないため 3 時間程度としている。また、雹などの特異な降水粒子のデータが推定される領域や方位角・仰角で区切ることは行わなかった。これはクラスター分析には複数種類のデータが含まれている方が、かえって分類しやすいという考え方のためである。

クラスター分析に使用する変数は $\Delta z_0$ を除いた

$$\mathbf{x} = [Z_{hh}, Z_{dr}, K_{dp}, \rho_{hv}, \sigma(Z_{dr}), \sigma(\rho_{hv}), \sigma(\psi_{dp})]$$

の 7 変数とする。各変数に対してスケーリングを行っているが、クラスター分析はスケーリングが異なると、分類結果に影響を及ぼすことがあるため、解析対象ごとに標準正規化はせず個別のスケーリングを行っている（付録 2）。



第 6.1 図  $Z_{hh}$  (横軸) と  $Z_{dr}$ ,  $K_{dp}$ ,  $\rho_{hv}$ ,  $\sigma(Z_{dr})$ ,  $\sigma(\rho_{hv})$ ,  $\sigma(\psi_{dp})$ ,  $\Delta z_0$  (縦軸) の GMM 代表値抽出後データに対するクラスター分析。(a, 上) 2017 年 4 月 15 日 15:00-17:00 (b, 下) 2017 年 5 月 26 日 10:00-11:00. 成田 DRAW の GMM による代表値を抽出した後のデータセットを K-means 法によるクラスター分析により分類。クラスターはデータの個数の多い方から赤-橙-黄-黄緑-緑-シアン-ディープスカイブルー-青-マゼンタ-紫の順。



に A・B だけでは分類ミスになる事例が見られた。反射強度  $Z_{hh}$  はかなり大きい地上の降水量は多くはないことから融解層やみぞれと推定されるが、偏波間相関係数  $\rho_{hv}$  が 1 に近い（融解層データでは異例）例が見られた。この事例を集め湿雪 C とした。この湿雪 C は⑥霰や⑦雹に近い分布をしているが、あくまでも二重偏波情報のデータセットから抽出したものであるため、粒子の物理的状態は不明である。(b)の  $\Delta z_0$  のグラフのマゼンダとシアンのカラスタには  $\Delta z_0 = 0[m]$  付近とその下の塊が含まれる。 $\Delta z_0 = 0[m]$  付近は融解過程の降水と推定されるが、その下は不明である。この不明な塊は、クラスタ分析ではまれな現象ではなく層状性降水でよく見られるものである。

雪に関する降水粒子の同定では、地上で降雪が観測される事例がほとんどなく、融解層より上層のデータのクラスタのパターンから通常の雪（乾雪）よりも  $Z_{dr}$  が大きくばらつきがみられる Ice Crystals と未識別のカラスタも分けている。

### 6.2 数値予報の気温データの利用と事前確率

各降水種別の確率密度分布は近い分布との重なりがあるため、二重偏波のみでは降水種別判別は困難であり、数値予報の気温データを利用する必要がある。過去の研究では

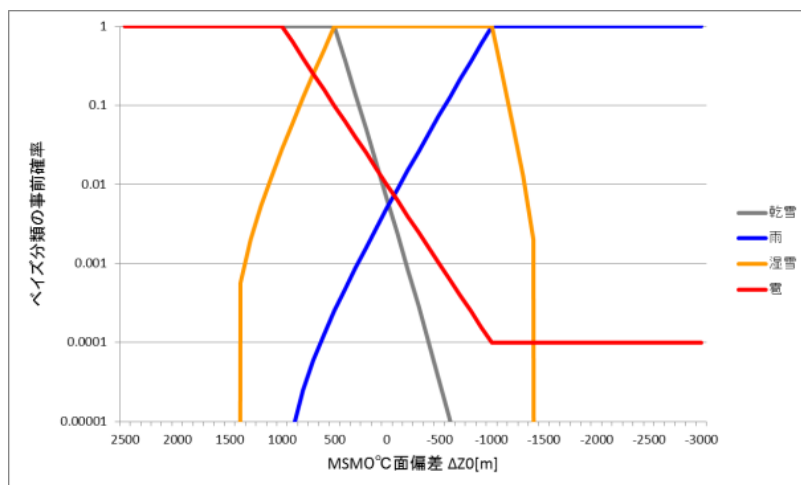
- ・ 気温  $T[K]$  または  $[^{\circ}C]$
- ・  $0^{\circ}C$  面高度 ( $z_0$ ) との差  $\Delta z_0 = z - z_0 [m]$

の 2 通りの利用がある。本報告では  $\Delta z_0$  を使用し

ている。これはレーダー雷解析（笠原，2011）の中間プロダクトとして MSM 等温度面高度データが DRAW のセンターシステム内で作成されており、ルーチン化を考えると都合が良いからである。また  $T[K]$  と  $\Delta z_0 [m]$  とでは  $\Delta z_0$  の方が計算は容易である。しかし  $\Delta z_0$  は逆転層で  $0^{\circ}C$  面高度が複数ある場合や地上でも氷点下になる場合では適用できない問題がある。後者の場合には外挿補間により  $\Delta z_0$  を求めている。冬季、特に北海道などの高緯度地域では地上でも氷点下を下回る場合が多い。このような場合では外挿による計算は精度が無いため  $\Delta z_0$  の採用は難しい。また逆転層の場合のことも考慮して、 $T$  を用いた検証を行っておく必要があると考えている。

第 5 章で述べたとおり  $\Delta z_0$  は数値予報から算出したものであるため予測誤差を含んでおり、ベイズ分類の直接の利用は事前確率を決める際の利用になる。

降水種別を四つのグループに分け、各グループの事前確率を第 6.2 図に示す。横軸は  $\Delta z_0 [m]$  で縦軸は事前確率を対数座標でプロットしている。雨は  $0^{\circ}C$  面の下面、乾雪は  $0^{\circ}C$  面の上面、湿雪は  $0^{\circ}C$  面付近に分布しており、この確率分布は実データを元にして経験的に調整したものであるが、湿雪の  $0^{\circ}C$  面偏差の範囲を  $-1,500[m] < \Delta z_0 < 1,500[m]$  としたのは、降水粒子の物理特性のみで決めたものではなく、予測誤差を含んだ分布を考慮したものであることに注意されたい。



第 6.2 図 降水種別グループごとのベイズ分類の事前確率



### 6.3 テクスチャ解析の利用

レーダーにおいてテクスチャ解析とえば、合成開口レーダーなどの解析で良く使われる2次元的な分布や模様の特徴量を示すことが多い。しかし、気象レーダーの降水種別解析ではある格子点の周囲との差の偏差を指すことが多い。定義の一例として *Gourley et al. (2007)* の場合を示す。二重偏波情報の変数を  $X(r, \theta)$  とすると、ある格子点  $(r_0, \theta_0)$  のテクスチャ  $\sigma_x(r_0, \theta_0)$  は

$$\sigma_x(r_0, \theta_0) = \sqrt{\frac{\sum_{r \neq r_0} \sum_{\theta \neq \theta_0} (X(r, \theta) - X(r_0, \theta_0))^2}{N}} \quad (1)$$

と定義される。 $\rho_{hv}$  や  $\psi_{dp}$  は低仰角においてクラッターによる外れ値が多くその影響を受けるため、標準偏差と中央絶対偏差との類推から本報告では以下のような定義を用いている。

$$\sigma_x(r_0, \theta_0) = \text{Median}(|X(r, \theta) - X(r_0, \theta_0)|) \quad (2)$$

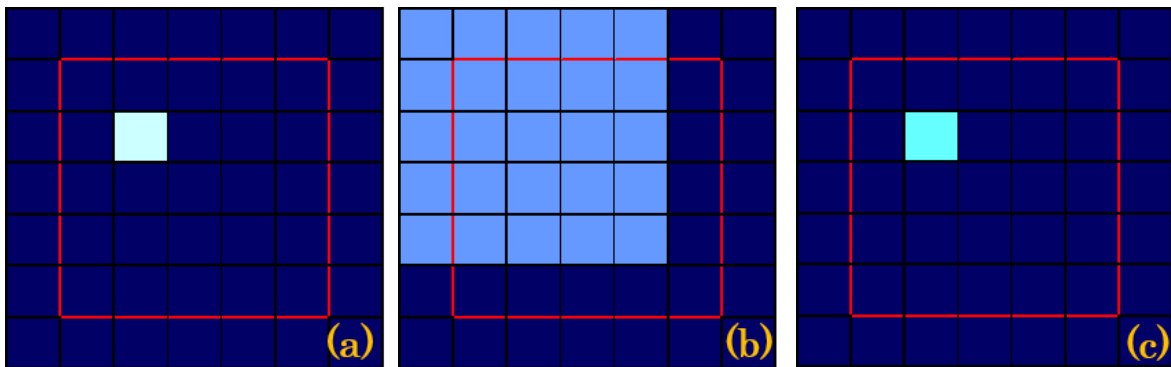
この2式による違いについて第6.3図を用いて説明する。(a)はこれから計算する二重偏波情報の格子を示し、外れ値以外の値の変動はほとんどないとする。テクスチャ計算を行う範囲を5×5マスの範囲とする。また上から3行目、左から3列目に外れ値を入れておく。(1)式の定義では外れ値はその関係する範囲に影響を与えるためテクスチャは(b)のようになる。(2)式の定義では外れ値の影響はほとんど受けず(c)のようなテクスチャ

になり精度的に改善される。

本報告の降水種別判別で利用するテクスチャは  $Z_{dr}$ ,  $\rho_{hv}$ ,  $\psi_{dp}$  の三つから求めたものである。この内  $\psi_{dp}$  は受信信号偏波間位相差と呼ばれるもので、折返し補正処理などにより  $\phi_{dp}$  を計算し、その動径 ( $r$ ) 方向の変化率が  $K_{dp}$  である。また  $Z_{dr}$ ,  $\psi_{dp}$  は角度 ( $\theta$ ) 方向で値が大きく変化するため、 $\theta$  は固定して動径 ( $r$ ) 方向のみのテクスチャを計算している。動径 ( $r$ ) 方向の振幅は  $\pm 1$  km 動かし、 $\rho_{hv}$  の角度 ( $\theta$ ) 方向は 1 km に相当する角度で、レンジごとに振幅を調整している。

### 6.4 クラッター判別

エコー判別として特に二重偏波気象レーダーで重要視されているものに、グラントクラッターやシークラッターの判別がある。他研究では降水種別判別でこれらのクラッター判別もすることが多い。二重偏波 DRAW では信号処理部中にクラッター除去技術である MTI や二重偏波情報を用いた簡易なクラッター判別処理が実装されており、一次データの時点でほぼクラッターは除去されている。そのため本報告の降水種別判別ではクラッター判別処理に関する機能は含めなかった。ただし、MTI 処理後のレーダーエコーは  $\rho_{hv}$  が極端に低くなっており、本報告の降水種別判別ではほとんど  $\text{⑰NOISE}$  として判別される。



第 6.3 図 テクスチャ計算の模式図。(a) 二重偏波情報  $X(r, \theta)$  の格子、上から 3 行目、左から 3 列目に外れ値を入れる。(b) *Gourley et al. (2007)* の定義によるテクスチャ。(c) 第 6.3 節 (2) 式の中央値を使った定義によるテクスチャ

## 7. 降水種別判別の事例とその検証

### 7.1 層状性降水での融解層検出

降水種別判別の例として、まず層状性降水で雨・融解層・雪の3層が分離できているかの確認を行う。第7.1図に関西 DRAW の仰角 5.1°PPI データでの例を示す。PPI データであるため画像の中心にあるレーダーサイト付近が地表に近く、レーダーから離れる図の外側に行くにつれて観測高度は高くなる。本事例では融解層（ブライトバンド）が明瞭に見える事例で、降水粒子は融解層での反射率が高いため反射強度の高い所がリング状に見えている。降水種別判別でも湿雪カテゴリーを示すオレンジ系の色で判別されていることがわかる。水平反射強度 $Z_{hh}$ を見るとリングの南側は強度が弱くはっきりしないが、降水種別分類ではこの部分でも $\rho_{hv}$ は低くはっきりとした融解層の特徴を示しているため、融解層と判別が可能である。融解層の内側が雨の領域で、外側が雪の領域になるが、おおむね降水種別判別結果もそのように判別されている。判別カテゴリーの⑨湿雪 B と④雨/霰の領域は近接しており、この二つの分類ミスも開発当初の課題であった。この分類ミスを少なくするために、

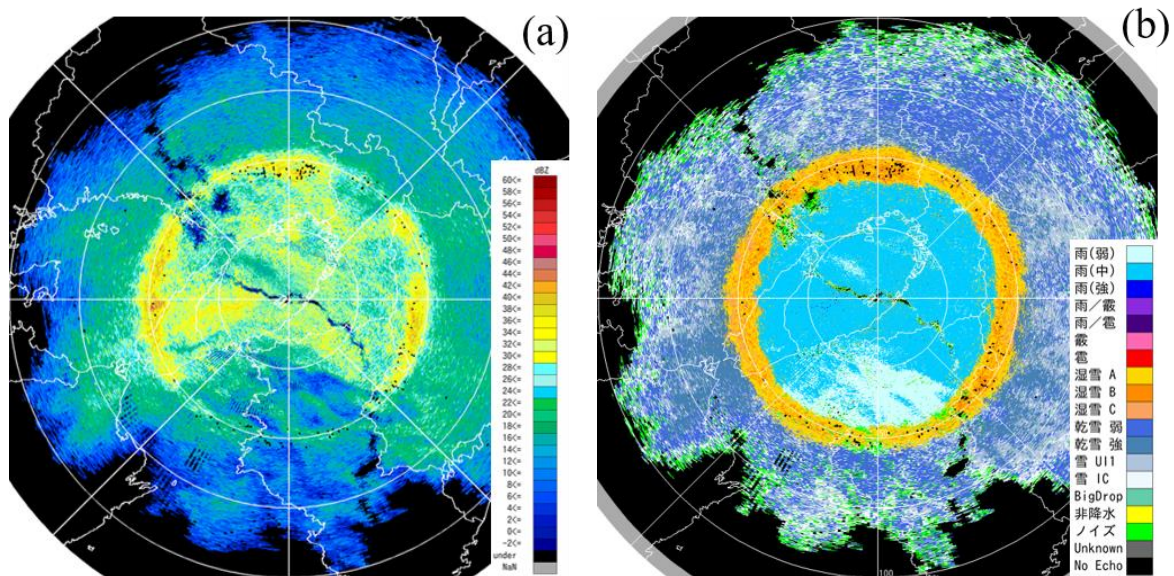
- ・  $K_{dp}$  やテクスチャ解析の導入
- ・ 湿雪のカテゴリーを細かく分ける
- ・ 事前分布の  $\Delta z_0$  を調整する

の3点を行っている。

なお、(a) ではエコーにひびが入ったようにエコーの無い領域が見られるがこれは MTI による影響で、MTI はドップラー速度 0 m/s 近傍のドップラースペクトルを除去し、固定物からのクラッターを除去するが、エコーのひびの箇所はドップラー速度 0 m/s 付近の領域に対応し、MTI 法が過度に適用されたものためである。

### 7.2 霰と雹の判別と発雷域との比較検証

もう一例として、降水種別判別の霰と雹の判別結果と、LIDEN 標定結果及びレーダー雷解析（笠原，2011）とを比較することにより、降水種別判別の霰と雹の判別精度を評価した。通常の雷雲では雲中に霰や雹の存在が示唆されるが、逆に霰や雹が雲中に存在しても発雷するとは限らない。しかし霰と雹の判別域と LIDEN 標定結果との比較はおおむねの判別傾向の評価には役に立つと考えている。

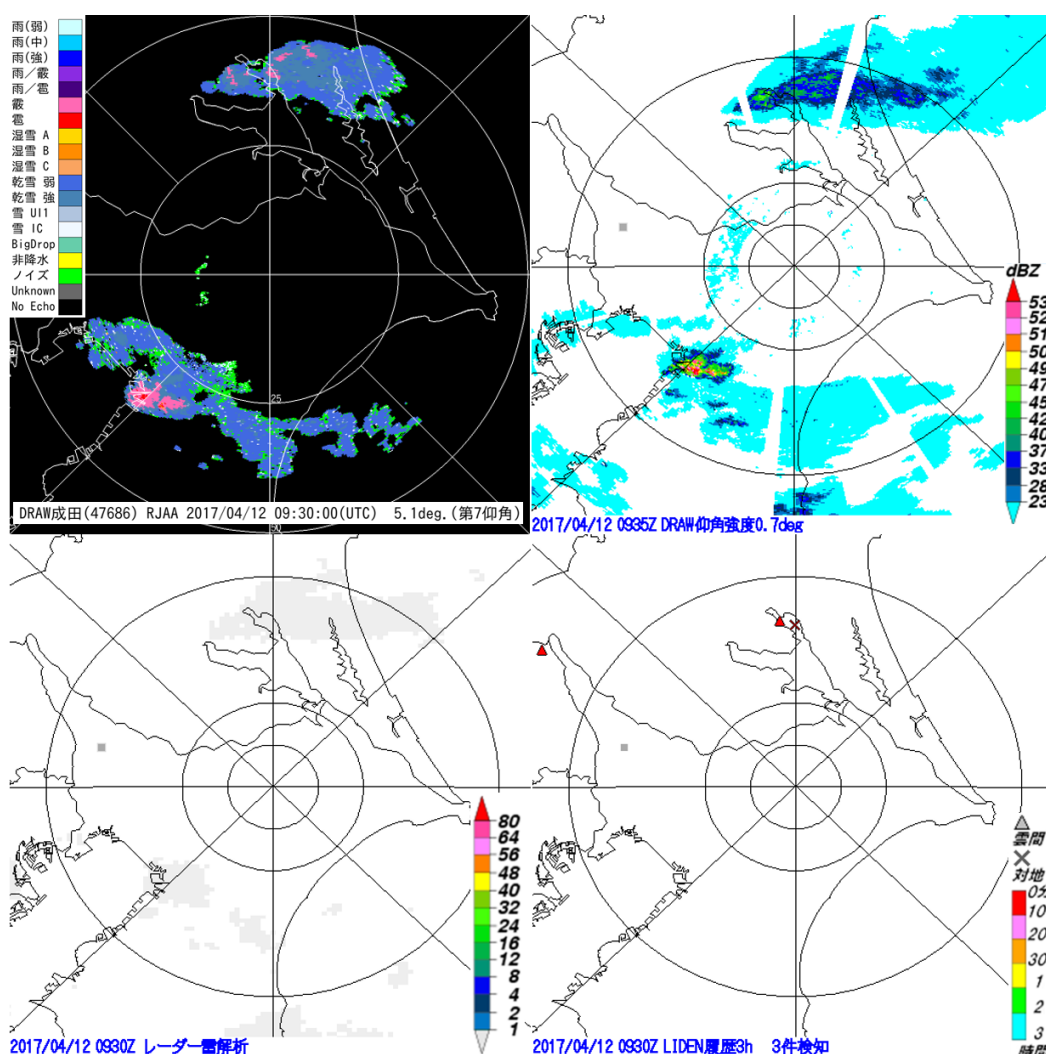


第 7.1 図 関西 DRAW の 2016 年 7 月 8 日 23:00 (1400UTC)。(a, 左) 水平反射強度 $Z_{hh}$  (b, 右) 降水種別判別結果

レーダー雷解析は $-20^{\circ}\text{C}$ で 40 dBZ 以上といった単偏波気象レーダーの雷予測に基づくもので、二重偏波気象レーダー降水種別判別結果の霰・雹判別結果との比較は単偏波と二重偏波での雷予測技術の傾向の違いの調査と言ってよい。

夏場の大規模な熱雷ではレーダー雷解析の雷判定域と本報告の降水種別判別の霰と雹の判別域との違いはあまりなかった。ここでは違いが見られた小規模な発雷の例を第 7.2 図に示す。左上の降水種別判別結果では霞ヶ浦の北側で霰と、千葉市付近で霰及び雹を検出している。LIDEN では霞ヶ浦北部で落雷・雲間雷を一つずつ検出してい

るが、レーダー雷解析では未検出であった。千葉市付近では LIDEN は雷を検出していなかったが、その 50 分後では LIDEN が発雷を捉えており、レーダー雷解析でも事前に雷雲を検出している。二重偏波レーダーではこれまで単偏波レーダーでは識別できなかったより小さい氷粒の霰域を判別することが可能になり、今後の雷予測の精度向上やリードタイムの確保の可能性を示唆していると言える。ただし教師データは成田 DRAW のデータによるものであり、日本海側冬季雷の霰域にも適用可能かはさらなる調査が必要である。



第 7.2 図 2017 年 4 月 12 日 18:30(0930UTC)成田 DRAW での降水種別判別と各種データとの比較。(左上)降水種別判別結果(仰角 $5.1^{\circ}$ ) (右上)レーダーエコー ( $Z_{hh}$ , 仰角 $0.7^{\circ}$ ) (左下)レーダー雷解析によるレーダー雷解析指数 (RLA) (右下) LIDEN による雷検知結果。レーダーエコーの時刻が異なるのは時刻定義や PPI 走査タイミングを考慮し 35 分のデータを使用しているため。

### 7.3 降水種別判別における課題や問題点

本手法では数値予報の0°C面データを使うため、降水種別判別の精度は予測誤差に大きく影響される。また0°C面が地表近くの場合は、地表近くの気温の分布が複雑で数値予報モデルで表現しにくいいため判別精度は特に悪くなる。さらに、使用した数値予報のMSMの水平解像度は5kmでDRAWの格子と比べると粗く、水平面で気温傾度が大きい場合は影響を受け、降水種別判別にMSMの格子の場所で不連続が生じることがある。そのため水平解像度が2kmで地上の局地性の強いデータを同化しているLFMデータの使用が望ましいと言え、今後の調査が必要である。

二重偏波気象レーダーでは、水平偏波と垂直偏波の反射強度の比 $Z_{dr}$ は校正され、通常の雨では $Z_{dr}$ は0dBになる。この $Z_{dr}$ のバイアス変動する場合があります、降水種別判別に影響を与えることから、運用上継続して監視する必要がある。

特に雹を伴うような強エコー域背面の領域では $Z_{hh}$ が減衰し $Z_{dr}$ の精度が下がるため判別精度が悪くなる。そのため強エコー域背面ではどの種別にも当てはまらない⑩Unknownで判別されることが多い。また、レーダーサイト直上に強雨域がある場合はレドームに水膜が生じその影響により $Z_{hh}$ や $Z_{dr}$ の品質に影響を与える。

本報告の教師データ作成では成田DRAWのデータを元に作成している。この成田DRAWの教師データを用いて羽田・関西DRAWの降水種別判別の確認は行っているが、降水粒子の粒径分布は地域や季節によって差があるため、成田DRAWで作成した教師データは他地域で合わない可能性がある。

今後一般レーダーが二重偏波化されていくが、DRAWのレーダーとはハードウェアや信号処理仕様等の差異により、特に $\rho_{hv}$ の特性が異なる可能性があり、新たに教師データを作成する必要がある。

### 8. まとめ

二重偏波情報による降水種別判別で技術上困難な問題として、分類に用いる教師データ作成の問題があり、従来の地上やゾンデデータから作成する方法は大変手間がかかるのが難点であった。

その教師データを機械学習の教師なし学習の一つであるクラスター分析(K-means法とGMM)を駆使することにより、効率良く作成できた。この教師データを元にベイズ分類の手法で降水種別判別を行い評価した結果、霰・雹検出などでは良好な結果が得られたが、数値予報の0°C面が地表付近にある場合では降水種別判別の精度が悪い場合があることや成田・羽田・関西以外の地域では未検証のこともあり、全国的な運用を目指して今後さらなる改良や検証が必要である。

### 謝辞

二重偏波気象レーダーのデータは従来の単偏波の気象レーダーデータと比べデータ量が1桁程度増え、その1次データの準備のために当時の観測部航空気象観測整備運用室及び成田・東京・関西の各航空地方気象台の皆様には大変なご労力を頂いた。また観測整備計画課の梶原佑介氏、気象技術開発室の南雲信宏氏には本稿執筆に際し助言を頂いた。この場をお借りし感謝申し上げます。

### 参考文献

- Al-Sakka, H., A.-A. Boumahmoud, B. Fradon, S. J. Frasier, and P. Tabary (2013) : A new fuzzy logic hydrometeor classification scheme applied to the French X-, C-, and S-band polarimetric radars. *J. Appl. Meteorol. Climatol.*, **52**, 2328-2344.
- Bringi, V. N. and V. Chandrasekar (2001) : Polarimetric Doppler Weather Radar: Principles and Applications. Cambridge University Press.
- Gourley, J. J., P. Tabary, and J. Parent du Chatelet (2007) : A fuzzy logic algorithm for the separation of precipitating from

- nonprecipitating echoes using polarimetric radar observations. *J. Atmos. Ocean. Technol.*, **24**, 1439-1451.
- Grazioli, J., D. Tuia, and A. Berne (2015) : Hydrometeor classification from polarimetric radar measurements: a clustering approach. *Atmos. Meas. Tech.*, **8**, 149-170.
- 笠原真吾 (2011) : 雷ナウキャストにおける雷の解析・予測技術と利用方法. 測候時報, **78**(3), 95-140.
- Keenan, T. (2003) : Hydrometeor classification with a C-band polarimetric radar. *Aust. Met. Mag.*, **52**, 23-31.
- Marzano, F. S., D. Scaranari, M. Montopoli, and G. Vulpiani (2008) : Supervised classification and estimation of hydrometeors from C-band dual-polarized radars: A Bayesian approach. *IEEE Trans. Geosci. Remote Sens.*, **46**, 85-98.
- Mishchenko, M. I. and L. D. Travis (1998) : Capabilities and limitations of a current FORTRAN implementation of the T-matrix method for randomly oriented, rotationally symmetric scatterers. *J. Quant. Spectrosc. Radiat. Transfer*, **60**(3), 309-324.
- Thompson, E. J., S. A. Rutledge, B. Dolan, V. Chandrasekar, and B. L. Cheong (2014) : A dual-polarization radar hydrometeor classification algorithm for winter precipitation. *J. Atmos. Ocean. Technol.*, **31**, 1457-1481.
- Wen, G., A. Protat, P. T. May, X. Wang, and W. Moran (2015) : A cluster-based method for hydrometeor classification using polarimetric variables. Part I: Interpretation and analysis. *J. Atmos. Ocean. Technol.*, **32**, 1320-1340.



付録 1 二重偏波レーダーで観測する二重偏波情報一覧

下記に二重偏波レーダーで観測する二重偏波情報の一覧を記す。ただし気象庁の運用する二重偏波レーダーで観測しないものは除く ( $Z_{hv}$ ,  $LDR$  など)。

$Z_{hh}$ [dBZ] 水平偏波反射強度	水平偏波の反射強度。これまでの単偏波の気象レーダーの反射強度と同じもの。Z-R関係式により降水強度推定を行うときに用いる。反射強度は雨滴による散乱・減衰のため距離方向に徐々に減衰する(降雨減衰)が、二重偏波レーダーでは偏波間位相差 $\phi_{dp}$ を用いて減衰量を補正できる。
$Z_{vv}$ [dBZ] 垂直偏波反射強度	垂直偏波の反射強度。扁平な降水粒子は水平偏波より散乱断面積が小さいため、 $Z_{hh}$ より小さい値を示す。通常は $Z_{dr}$ を計算するために用いる。
$Z_{dr}$ [dB] 反射因子差	水平偏波と垂直偏波の反射強度の比。反射強度の単位は dBZのため $Z_{dr} = Z_{hh} - Z_{vv}$ 。扁平度を表す二重偏波情報で、 $Z_{dr}$ がプラスは横に、マイナスは縦に降水粒子が扁平していることを表す。
$\psi_{dp}$ [deg]受信信号偏波間位相差	水平偏波と垂直偏波の位相差、水平偏波が遅れる方をプラスとする。値は $0^{\circ}$ - $360^{\circ}$ で折り返しに注意を要するため使用することは少ない。 $\phi_{dp}$ の計算に用いる。
$\phi_{dp}$ [deg] 偏波間位相差	$\psi_{dp}$ に折り返し補正を施すとともに、ミー散乱(大きな粒子による電磁波散乱)による影響等を除去したもの。水平に扁平した降水粒子をレーダーからの電磁波が通過すると、水平偏波が遅れるため $\phi_{dp}$ は増加し、垂直に扁平した降水粒子を通過すると減少する。 $K_{dp}$ を計算するのに用いるほか、降雨減衰の影響を受けない位相の性質を利用し、降雨減衰の補正に利用可能。
$\rho_{hv}$ 偏波間相関係数	水平・垂直偏波間の相関係数でレーダービーム内の降水粒子の均一性(雨と雪などの異なる粒子が混合していないか)を示す。雨や雪の単一な降水であれば係数は高く1に近い値をとり、融解層のように混在する領域では係数は0.95-0.85程度の低い値を示す。
$K_{dp}$ [deg/km] 偏波間位相差変化率	$\phi_{dp}$ の動径方向の傾きで、レーダービーム内の降水粒子の扁平度に対応するが、降水強度に対して非線形性が強く、扁平度の小さい弱い降水や氷粒子に対しては感度が無い。降雨減衰の影響が無いため高精度の降水強度の推定に利用されている。そのため強雨域に $K_{dp}-R$ 関係を用い、弱雨域には $Z_{hh}-R$ 関係を降水強度推定に用いることが多い。

付録 2 非階層クラスター分析による降水種別判別の教師データ作成

①データの準備

降水種別判別に用いるデータは $Z_{hhc} \cdot Z_{arc} \cdot \rho_{hv} \cdot K_{dp} \cdot \sigma(Z_{dr}) \cdot \sigma(\rho_{hv}) \cdot \sigma(\psi_{dp}) \cdot \Delta z_0$ の 8 種である。このうち $Z_{hhc}$ と $Z_{arc}$ は $Z_{hh}$ と $Z_{dr}$ から降雨減衰補正をしたもので、 $\sigma(\cdot)$ のデータは中央絶対偏差によるテクスチャである。 $\Delta z_0$ は数値予報の 0 °C面高度とレーダー観測格子高度との差である。

②GMM による代表値抽出

PPI レーダーデータ 1 枚ごとに、 $Z_{hhc}$ に対し 4 dBZ の範囲を抽出し、8 種のデータについて 8 次元混合ガウスモデル (Gaussian Mixture Model, GMM) を適用する。ガウシアンモデルの数は 1-3 とし、数の決定はベイズ情報量基準 (Bayesian Information Criterion, BIC) で行う。ガウシアンセンターをその代表値とする。また $Z_{hhc}$ で 4 dBZ の範囲を抽出する作業は 0.5 dBZ ずつずらしていき、-20-70 dBZ まで行う。具体例は本文第 4.1 図に示してある。

③データのスタッキング

②で抽出した代表値データを数時間足し合わせる。目安として DRAW のレーダーレンジ全面にエコーがかかるような層状性の降水では 1 時間以内、発雷があるような対流性降水の場合は 3 時間程度である。

④K-means 法による非階層的クラスター分析とその結果に対する推定降水粒子の同定

③のデータに対し K-means 法を用い複数のクラスターに分割する。ただしデータとして用いる 8 変数のデータの中で $\Delta z_0$ は K-means 法の計算に使用しない。クラスター分析の変数として使用しないが、 $\Delta z_0$ によりクラスターの降水粒子の状態が水、氷/雪もしくはその融解したものかの推定を行う。また以下の表と数式によるスケールリングを施してからクラスター分析を行い、その後逆演算をする。

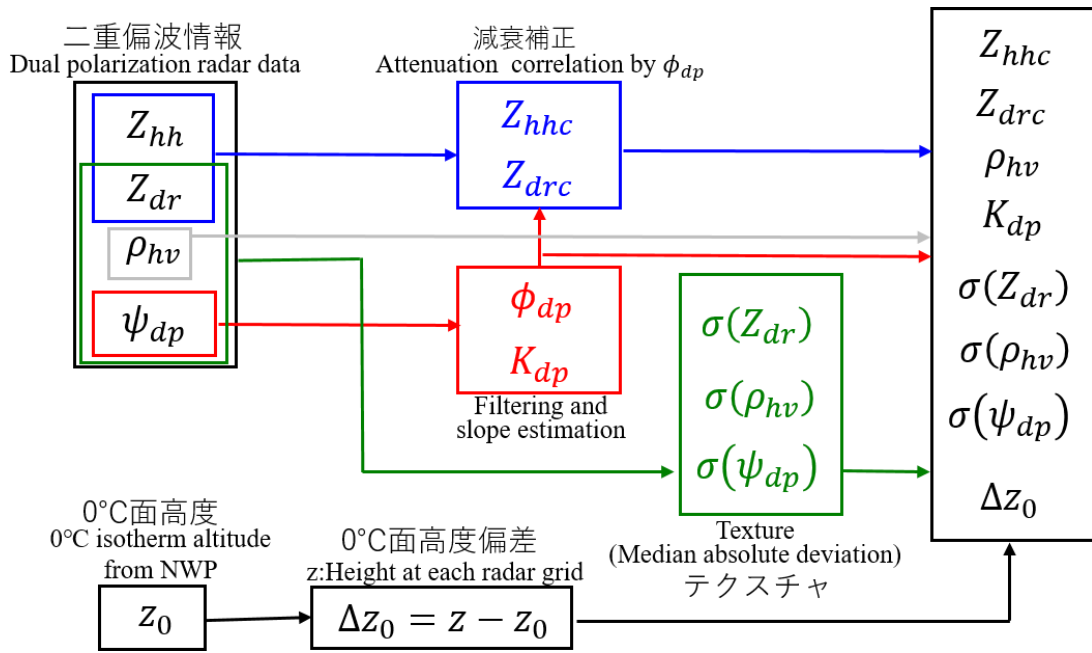
	Upper	Lower
Zhh	50[dBZ]	10[dBZ]
Zdr	5[dB]	-1[dB]
Kdp	7.5[deg/km]	-2.5[deg/km]
$\rho_{hv}$	1.0	0.7
$\sigma(Zdr)$	4[dB]	0[dB]
$\sigma(\psi_{dp})$	40[deg]	0[deg]
$\sigma(\rho_{hv})$	0.5	0

$$Z_{scaled} = \frac{Z - Z_{lower}}{Z_{upper} - Z_{lower}}$$

クラスター分析の結果のクラスターがどの降水種別なのか同定しておく必要がある。この同定の判断材料としては、上記の $\Delta z_0$ のほか、クラスターのデータ数及び先行研究の結果や T-matrix 法によるシミュレーション結果などを利用する。

⑤確率密度分布の推定

降水種別との紐づけが済んだクラスターデータをカーネル密度推定法(Kernel Density Estimation)により確率密度分布に変換する。④と同様に変数は $\Delta z_0$ を除いた七つの変数 $Z_{hhc} \cdot Z_{arc} \cdot \rho_{hv} \cdot K_{dp} \cdot \sigma(Z_{dr}) \cdot \sigma(\rho_{hv}) \cdot \sigma(\psi_{dp})$ である。



Filtering using the centers of Gaussian Mixture Model (GMM) to 8 variables for each PPI data. The number of components is from 1 to 3 selected by Bayesian information criterion (BIC).

Stacking the filtered data (gaussian center) for 1 to 3 hours.

K-means clustering to 10 groups for 7 variables ( $Z_{hhc}, Z_{drc}, \rho_{hv}, K_{dp}, \sigma(Z_{dr}), \sigma(\rho_{hv}), \sigma(\psi_{dp})$ ). Identifying a cluster with a specific precipitation type.

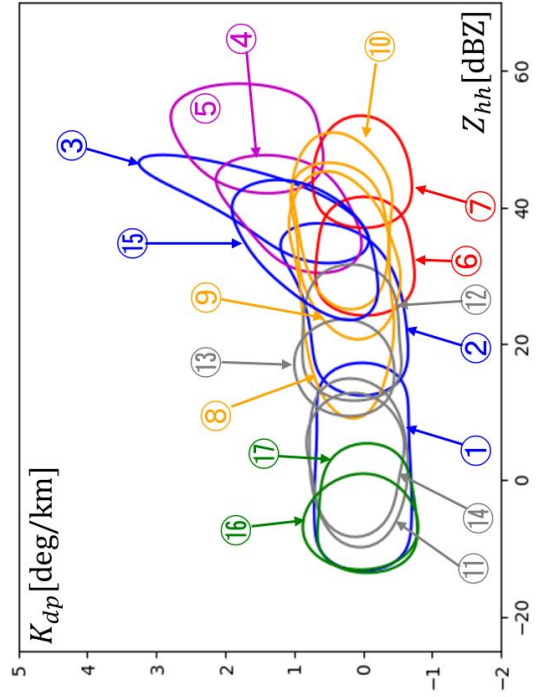
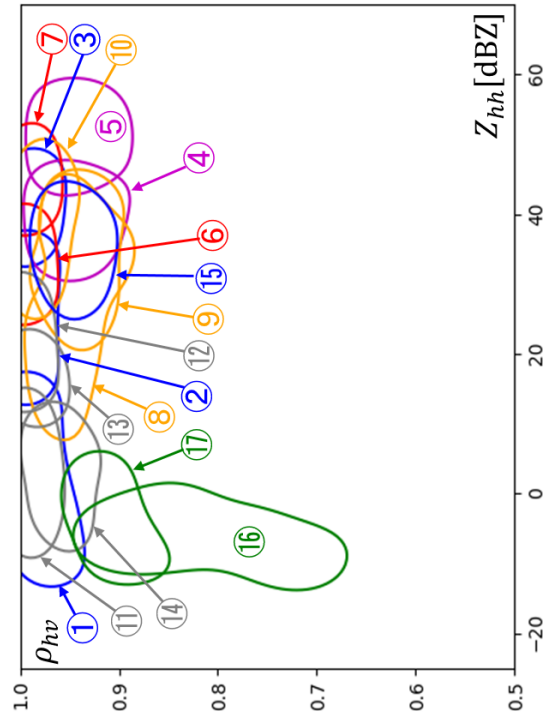
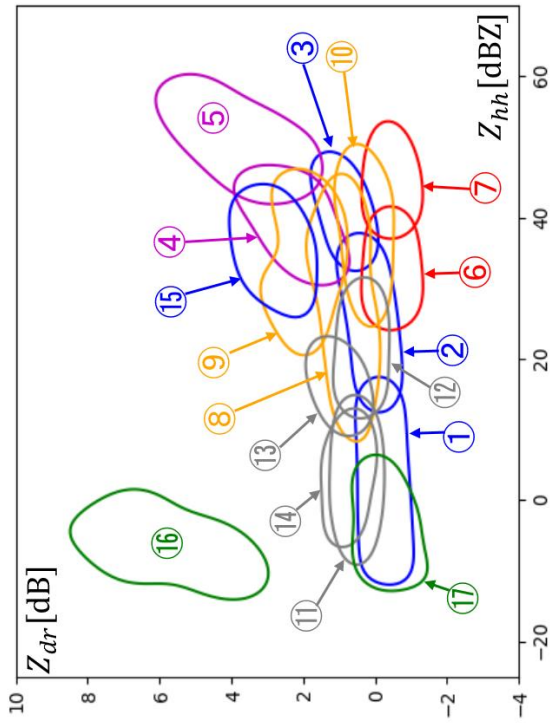
Estimating the probability density from the identified cluster using Kernel Density Estimation(KDE).

降水種別判別の教師データ作成模式図

Flow diagram of training data creation for Hydrometeor classification algorithm

付録 3 降水種別の確率分布 (2次元に射影したもので 50%等値線, 番号は第 6.1 表もしくは付録 4 と同じ).

Appendix 3 2D projections of the 7D probability density distributions of precipitation types (50 % contour). The contour index corresponds to the index of Table 6.1 or Appendix 4.



付録 4 本手法で分類する降水種別の一覧

Appendix 4 List of the classified precipitation

Precipitation types 降水種別	種別の特徴ならびに判別するための判断材料	二重偏波パラメータ					ρ hv 均一性	温度特性
		Zhh[dBZ] サイズ	Zdr[dB] 偏平度	Kdp[deg/km] 偏平度	Zdr≠0 偏平度	Kdp≠0 偏平度		
① 雨(弱) Weak Rain	種別の特徴ならびに判別するための判断材料							
② 雨(中) Rain	水平反射強度 Zhh<20[dBZ]、降水強度にして mm/hr<0.5 の弱い雨	Zhh<20	Zdr≠0	Kdp≠0	Zdr≠0	Kdp≠0	ρ hv≠1	雨
③ 雨(強) Heavy Rain	水平反射強度 Zhh=20~40[dBZ]、降水強度にして 0.5~12[mm/hr] の中程度の雨	20<Zhh<40	0<Zdr<2	0<Kdp<2	0<Zdr<2	0<Kdp<2	ρ hv≠1	雨
④ 雨/霰 Rain/Graupel	水平反射強度 Zhh>40[dBZ]、降水強度にして >12[mm/hr] の強い雨で Kdp が大	Zhh>40	0<Zdr<5	0<Kdp<10	0<Zdr<5	0<Kdp<10	ρ hv≠1	雨
⑤ 雨/雹 Rain/Hail	霰と雨の混合状態もしくは濡れた雹、水平反射強度 Zhh<50[dBZ] のもの	35<Zhh<50	0<Zdr<3	0<Kdp<1	0<Zdr<3	0<Kdp<1	0.9<ρ hv<1	雨
⑥ 霰 Graupel	霰と雨の混合状態もしくは濡れた雹、水平反射強度 Zhh>50[dBZ] のもの、Kdp は雨(強)より小さい	Zhh>50	0<Zdr<8	-2<Kdp<3	0<Zdr<8	-2<Kdp<3	0.8<ρ hv<1	雨
⑦ 雹 Hail	丸い氷の粒子で Zdr≠0[dB]、水平反射強度 Zhh<45[dBZ]	30<Zhh<45	Zdr≠0	Kdp≠0	Zdr≠0	Kdp≠0	ρ hv≠1	乾雪
⑧ 湿雪 A Wet Snow A	丸い氷の粒子で Zdr≠0[dB]、水平反射強度 Zhh>45[dBZ]	Zhh>45	Zdr≠0	Kdp≠0	Zdr≠0	Kdp≠0	ρ hv≠1	雹
⑨ 湿雪 B Wet Snow B	層状性降水等で見られる融解層のうち、反射強度は弱く高度は 0°C 面よりやや高い所に出現するもの	15<Zhh<45	0<Zdr<2	Kdp≠0	0<Zdr<2	Kdp≠0	0.9<ρ hv<1	湿雪
⑩ 湿雪 C Wet Snow C	融解層やみぞれ中の放射強度の強い領域、ただし融解層とみぞれでは若干特性が異なるが合わせた領域にしている	25<Zhh<60	1<Zdr<5	0<Kdp<2	1<Zdr<5	0<Kdp<2	0.7<ρ hv<0.95	湿雪
⑪ 乾雪弱 Weak Dry Snow	AB よりも放射強度が強いが実際の降水量は多くはない、Zdr は小さく ρ hv は 1 に近い、霰・雹のエコーに似る	30<Zhh<60	0<Zdr<1	Kdp≠0	0<Zdr<1	Kdp≠0	ρ hv≠1	湿雪
⑫ 乾雪強 Heavy Dry Snow	層状性降水系の雪で水平反射強度が Zhh<20[dBZ] のもの	Zhh<20	0<Zdr<2	Kdp≠0	0<Zdr<2	Kdp≠0	0.95<ρ hv<1	乾雪
⑬ 雪(未識別1) Snow(identified)	層状性降水系の雪で水平反射強度が Zhh>20[dBZ] のもの	Zhh>20	0<Zdr<2	Kdp≠0	0<Zdr<2	Kdp≠0	0.95<ρ hv<1	乾雪
⑭ Ice Crystals	雪の領域で、湿雪と付随して出現することが多いが ρ hv がほぼ 1 のもの、識別は仮のもの、詳しくは不明	20<Zhh<40	0<Zdr<2	Kdp≠0	0<Zdr<2	Kdp≠0	ρ hv≠1	乾雪
⑮ Big Drops	小さな氷粒で 0<Zdr<5[dB] とばらつきが大きいもの	0<Zhh<25	0<Zdr<5	Kdp≠0	0<Zdr<5	Kdp≠0	0.9<ρ hv<1	乾雪
⑯ 非降水 Biological Scatter	大きい雨滴で体積当たり粒子数が少ない、Zhh は中程度で Zdr と Kdp が雨より大きい、不安定性降水前などにまれに出現	30<Zhh<50	0<Zdr<5	0<Kdp<3	0<Zdr<5	0<Kdp<3	0.9<ρ hv<1	雨+
⑰ NOISE	生物起源と思われる降水を伴わないエコー。Zdr が極端に大きく ρ hv が低い、値のばらつきが非常に大きい。	Zhh<10	Zdr>3	Kdp≠0	Zdr>3	Kdp≠0	0.5<ρ hv<1	雨+
⑱ Unknown	S/N が悪い領域。クラッターによりノイズが大きい、エコーの端や最上層の信号が弱い領域。ばらつきが大きい。	Zhh<10	-2<Zdr<3	Kdp≠0	-2<Zdr<3	Kdp≠0	0.7<ρ hv<1	-
⑲ No Echo	ベイズ分類でどの種別の確率も低かったもの。強雨域の背面やビームカット域など。	-	-	-	-	-	-	-
	レーダーのエコーが無い領域	-	-	-	-	-	-	-