

**Jean François Ribaud – Pos- Doc - Atividades desenvolvidas para o Projeto
SOS-Chuva no período de Jan/17 até Jan/18**

Desenvolvimento duma nova classificação dos hidrometeoros a partir do radar polarimétrico da banda X para uma melhora previsão das tempestades a curto prazo

Hoje em dia, novos radares equipados com dupla polarização permitem de conseguir mais informações sobre as partículas que constituem as nuvens. Com quatro variáveis (contra somente uma para os radares “clássicos”), esses radares polarímetros podem nos informar sobre o tamanho, a forma, a orientação, e a fase dos hidrômetros (conjunto de partículas de água líquida ou sólida em queda ou suspensão na atmosfera). Desde o surgimento desses radares, várias técnicas foram desenvolvidas para identificar diretamente o tipo do hidrômetro dominante na nuvem. Esta abordagem abre também novas perspectivas em especial para a avaliação das parametrizações em modelos de previsão numérica com alta resolução espacial, a investigação das relações microfísicas, ou ainda melhorar as previsões de curtos prazos (“nowcasting”).

Embora a classificação dos hidrometeoros a partir dos radares com dupla polarização são muito conhecidas, desde os anos 2000, nenhuma classificação foi ainda desenvolvida para as regiões tropicais. Além disso várias pesquisas recentes mostraram que existem diferenças entre as características das nuvens nas estações chuvosa e seca (Cecchini et al. 2017; Jäkel et al. 2017, Machado et al. 2017). Entretanto ninguém até hoje avaliou estas diferenças do ponto de vista da distribuição vertical e espacial dos hidrômetros dominantes. Por todos estes motivos, foi escolhido desenvolver uma nova classificação dos hidrometeoros para os trópicos brasileiros a partir do radar da banda X envolvido no projeto SOS-CHUVA.

Atualmente existem três metodologias para identificar os hidrômetros dos radares com dupla polarização:

1. Booleano. É a primeira metodologia que foi desenvolvida durante os anos 1990-2000 (Bringi et al 1986; Straka et al 1993, Höller et al 1994). Ela consiste em definir basicamente as gamas de valores para cada tipo de hidrômetro e cada variável polarimétrica. Mas esta técnica não leva em conta a possibilidade de que dois hidrômetros possam ser definidos no mesmo espaço polarimétrico, resultando algumas vezes em erros na classificações).
2. Lógica “fuzzy”. Ela representa a maioria das classificações que foram desenvolvidas durante os anos ~ 2000-2010 para os serviços meteorológicos (Vivekanandan et al 1999; Ryzhkov et al 2005; Dolan et al 2009; Al-Sakka et al 2013). Esta metodologia permite precisamente de resolver o problema da superposição booleano, usando suaves transições dos limites nas variáveis do radar a partir de uma combinação das funções Triangulares ou Gaussianas. Essas funções são inferidas através das simulações no T-Matrix e algumas vezes pelas observações in situ. Embora esta técnica é robusta e fácil para implementar em modelos de previsão, ela apresenta certas limitações como: a escolha do número de tipos de hidrometeoros, das simulações de T-Matrix geralmente erradas para a fase sólida e que não considerem a precisão das medições do radar nem a coerência da distribuição espacial dos hidrômetros.
3. Clustering. Nos últimos anos esta técnica parece ser a nova referência (Bechini et al 2014, Grazioli et al 2015, Wen et al 2015, Besic et al 2016). Apesar do custo computacional, esta metodologia usa nenhuma hipótese de base (unsupervised clustering) ou muito pouco (semi-supervised clustering) para procurar em conjunto de dados polarimétricos padrões similares (cluster). No caso do “unsupervised” técnica, o número dos hidrometeoros não são definidos a priori e não se baseiam nas simulações de uma distribuição teórica. O número dos hidrometeoros são definidos a partir dos dados iniciais, e o conteúdo final é manualmente interpretado.

A maioria destas classificações listadas acima usam limites que possam ser específicos para cada hidrometeoro, cada região, ou ainda cada comprimento de onda. A metodologia de “unsupervised clustering” precisamente permite de deixar toda a liberdade ao conjunto de dados polarimétricos sem nenhuma hipótese a

priori. Por exemplo, permitir identificar as colunas convectivas que ocorrem em sistemas severos, não havendo um limite definido entre partículas líquidas e sólidas. Este é caso complexo e frequente onde a água líquida pode atingir vários quilômetros acima da isoterma 0°C.

Das principais metodologias de clustering, foi escolhido a do artigo de Grazioli et al 2015, que se baseia num específico tipo que se chama “Agglomerative Hierarchical Clustering” (Ward et al 1963). Apresenta-se como uma sequência de iterações que cada vez vai agrupar N objetos em n_c clusters a fim de garantir que os objetos de um mesmo cluster apresentam, mais similaridades (físicas) que aqueles pertençam aos outros grupos. No início da metodologia cada objeto corresponde a um cluster ($N = n_c$). Depois de uma iteração, ficam sempre N objetos mas separados desta vez dentro n_c-1 clusters. Essas iterações devem ser repetidas até que fique no final N objetos para somente um cluster. Posteriormente o utilizador poderá escolher quando tiver a “melhor” distribuição entre clusters (por exemplo: 5, 6, ou mais clusters) com ferramentas estatísticas e sua interpretação pessoal.

Os objetos aqui fazem referência a um vector de dimensão cinco, que se define como a combinação das quatro variáveis do radar e uma informação da temperatura (via radiossondagem ou modelo atmosférico):

$$x = \{Z_H, Z_{DR}, K_{DP}, \rho_{HV}, T_C\}$$

com Z_H refletividade horizontal [dBZ], Z_{DR} refletividade diferencial [dB], K_{DP} diferença de fase específica [$^{\circ}$ /km], ρ_{HV} coeficiente de correlação [-], T_C temperatura [$^{\circ}$ C]. Para se conhecer quais clusters são mais similares, foi definido uma métrica (aqui simplesmente euclidiana) e uma regra de aglomeração. Há muitas regras para agrupar os clusters entre eles (simples, completa, mediana, centroide, ponderada, entre outras). Nesse documento apresentamos somente resultados baseados na técnica de Ward, que consiste em reduzir a variância entre objetos dentro do mesmo clusters ou entre clusters.

Os dados polarimétricos usados nesse documento foram coletados com o radar polarimétrico banda X do SOS CHUVA. Como a técnica de clustering precisa de muitos dados para aprender / construir (data-driven) e que no início desse trabalho ainda não

havia os dados de Campinas, escolhemos em primeiro lugar desenvolver a técnica com os dados do projeto ACRIDICON-CHUVA que ocorreu em Manaus em 2014 (Machado et al. 2017). Os dados brutos foram processados da seguinte forma: clutter identificação, Ψ_{DP} correção, K_{DP} cálculo, e atenuação correção (com Z_H e Z_{DR}). Além disso, uma separação entre as regiões estratiformes e convectivas foi adicionada ao conjunto dos dados (Steiner et al, 1995). De fato, há uma diferença entre as duas regiões da nuvens em termos das dinâmicas e microfísicas.

Aqui apresentamos os resultados pela estação húmida de 2014 do experimento Go-Amazon ACRIDICON-CHUVA com o radar polarimétrico banda X que operou próximo de Manaus. Ao longo do período, realizou-se um clustering unsupervised com 50 iterações. Para avaliar a qualidade espacial de cada partição, um teste estatístico, o “elbow” (literalmente “cotovelo” em português) foi realizado e é ilustrado na Figura 1. Ao analisar essa figura, percebe-se que a partir de 4-5 clusters pelas ambas regiões, o ganho torna-se muito fraco. Além disso a análise manual da distribuição espacial dos clusters confirma esse resultado. Assim, o número de classes de hidrometeoros ficou definido como de 4 e 5 clusters pelas regiões estratiformes e convectivas respectivamente.

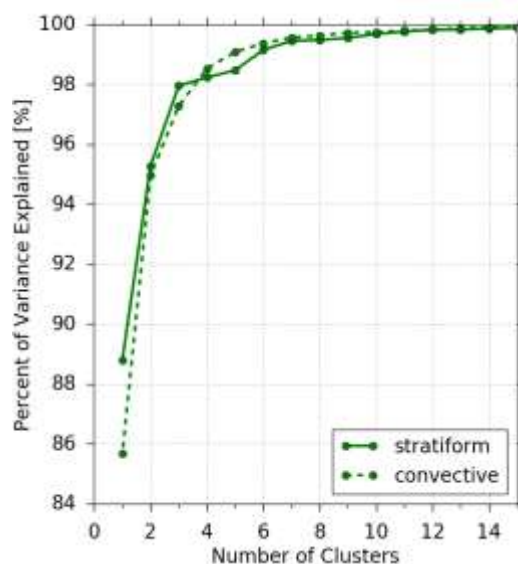


Figura 1: Evolução da variância explicada (%) pelos clusters em função do número dos clusters com os dados da estação chuvosa 2014 e as regiões estratiformes e convectivas. Inicialmente nota-se um ganho muito grande com o aumento das classes, mas em torno de 4-5 classes o ganho é marginal. O número dos clusters foi escolhido nesse ponto.

A figura 2 mostra a distribuição dos clusters para as regiões convectivas ($n_{c_convectiva} = 5$ a 9) na estação chuvosa do ano 2014 e a Figura 3 mostra o RHI apresentando a distribuição vertical dos hidrometeoros dentro da célula convectiva em 7 de fevereiro 2017 no final da tarde. A Figura 4 mostra a matriz de espalhamento comparando cada cluster com os hidrometeoros definidos no artigo de Besic et al 2016 (lógica fuzzy). Ao comparar todas estas informações é possível identificar e interpretar os clusters como tipos de hidrometeoros:

- Cluster $n_c = 5$: de acordo com a Figura 2 esse cluster é definido com seguinte: Z_H [0 ~ 30 dBZ], Z_{DR} [-1 ~ 3 dB], K_{DP} [0 ~ 1 °/km], ρ_{HV} [>0.96] e apresenta uma temperatura positiva (diferencia negativa de altitude entre a medição e a banda brilhante). A partir da Figura 3, o cluster $n_c 5$ parece ser caracterizado na categoria Drizzle (DR) da lógica fuzzy que servirá aqui de referência. Esta alta correlação entre os dois padrões é confirmada com a tabela da Figura 4. De fato, sobre todas as medições selecionadas com a técnica de clustering durante a estação húmida 2014, o cluster $n_c 5$ e o drizzle têm uma concordância de ~ 92%.

- Cluster $n_c = 6$: as maiores diferenças com o cluster precedente são uma refletividade horizontal Z_H mas intensa [20 ~ 50 dBZ] e a refletividade diferencial Z_{DR} [0 ~ 4 dB]. Desta vez, esse cluster 6 apresenta uma correlação óbvia com o tipo Rain (RN) da lógica fuzzy no RHI da Figura 3, mas também ao longo da estação húmida com uma conformidade entre essas duas técnicas maioria de 66%.

- Cluster $n_c = 7$: as distribuições desse cluster são Z_H [20 ~ 50 dBZ], Z_{DR} [-1 ~ 3 dB], K_{DP} [0 ~ 1 °/km], ρ_{HV} [>0.96] e tem uma diferença de altitude positiva referindo-se a um hidrometeoro da fase solida ($T^{\circ}C < 0^{\circ}C$). Na figura 3 podemos ver que esse padrão corresponde, ao tipo Low Density Graupel (LDG) da lógica fuzzy. Sobre todas medições consideradas nesse clustering, a concordância entre o cluster $n_c 4$ e o graupel está de 54% e de 35% com a categoria Wet Snow (WS).

- Cluster $n_c = 8$: esse padrão apresenta principalmente refletividades menores Z_H [15 ~ 30 dBZ] e uma distribuição vertical com temperaturas um pouco mais negativa. Por isso, esse cluster tem uma correlação, mais alta com o tipo dos hidrometeoros que se chama Aggregates (AG ~ 65%).

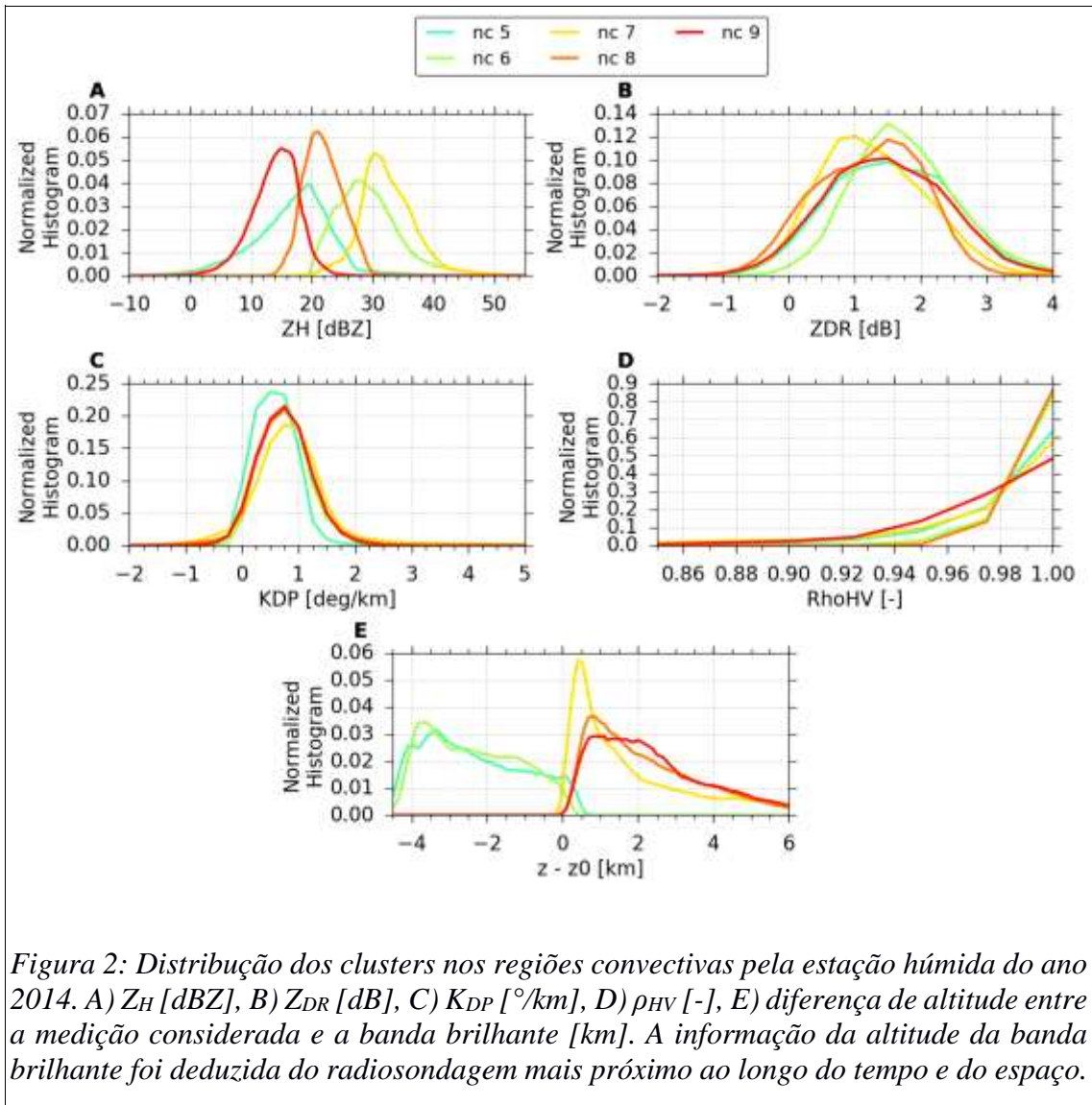
- Cluster $n_c = 9$: as maiores diferenças com o cluster precedente é principalmente na refletividade horizontal mais fraca (Z_H [0 ~ 25 dBZ]). Na figura 2 esse cluster parece ser

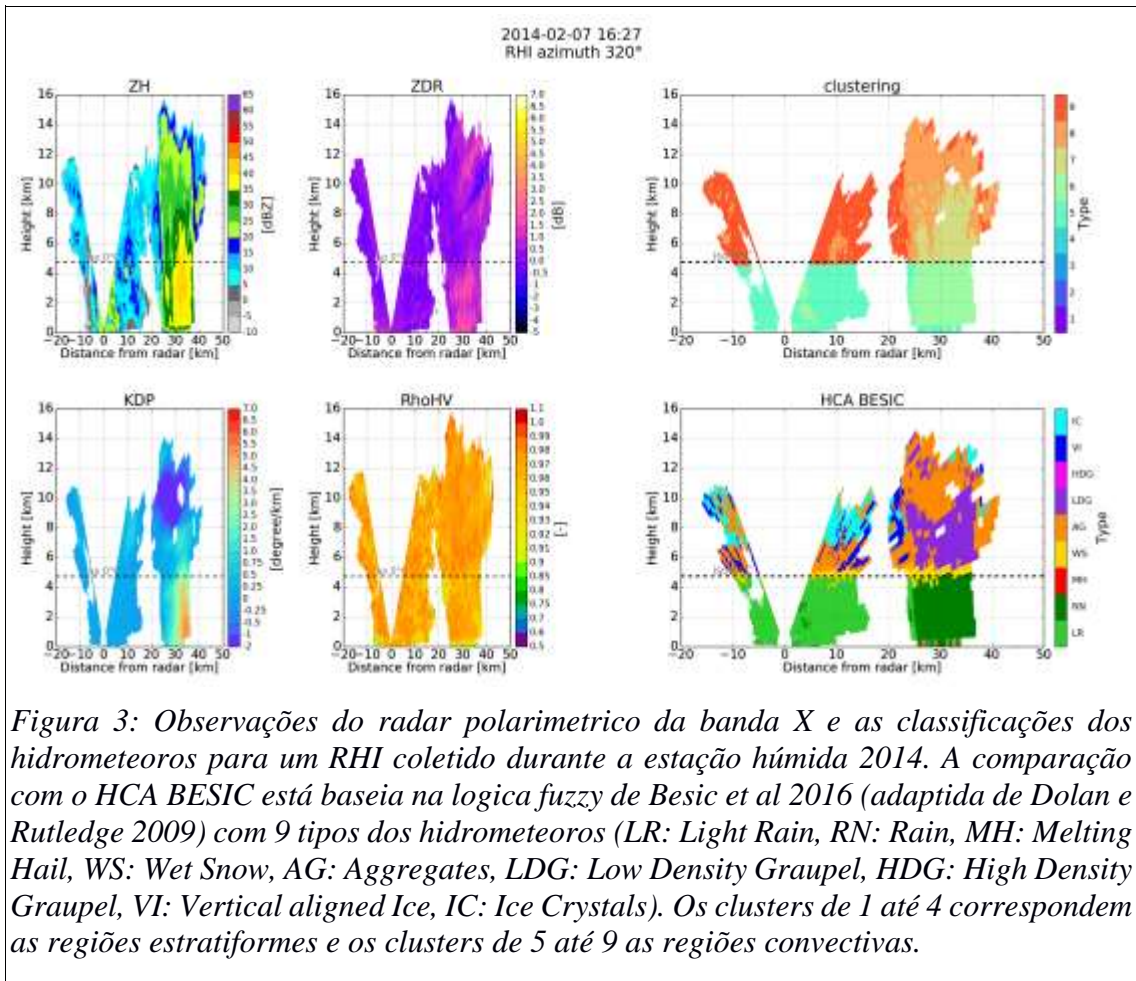
repartido entre os tipos dos hidrometeoros Aggregates (AG) e Ice Crystals (IC) da lógica fuzzy. No final o cluster nc 9 divide-se entre os dois com as correlações de 52% e 22% respectivamente (tabela da Figura 4).

A síntese final do clustering com todos clusters é apresentada na Tabela 5 que mostra as características das variáveis do radar polarimétrico para cada cluster com: a mediana, desvio padrão, e vários percentis.

A análise global desses resultados mostra a boa coerência da metodologia de clustering que permite diferenciar independentemente de hipóteses pré-estabelecidas os variados padrões de tipos de hidrômetros. A distribuição espacial na Figura 3 com o RHI é realista e física para uma célula convectiva. Além disso, a concordância geral entre esse clustering e o artigo de referência que usa uma classificação de lógica fuzzy é forte e distinta.

Finalmente, somente os resultados da estação chuvosa foram mostrados nesse documento. Mas já temos hoje as distribuições dos clusters para as estações seca e chuvosa, assim que para as regiões estratiformes e convectivas. Esses resultados da primeira classificação dos hidrometeoros com um radar polarimétrico da banda X nos trópicos serão objeto de um artigo científico nos próximos meses. Além disso, essa técnica é atualmente em fase de adaptação com os dados do projeto SOS-CHUVA.





TYPE	DZ	RN	MH	WS	DS	LDG	HDG	VI	CR
nc 5	92.45	1.74	0.31	5.35	0.07	0.04	0.0	0.01	0.04
nc 6	28.47	66.57	0.45	4.5	0.0	0.0	0.0	0.0	0.0
nc 7	0.05	0.46	0.0	35.7	8.54	54.32	0.93	0.0	0.0
nc 8	0.04	0.0	0.0	17.11	64.19	16.44	0.0	0.88	1.34
nc 9	0.07	0.0	0.0	13.31	52.54	1.97	0.0	9.49	22.61

Figura 4: Matriz de espalhamento comparando cada cluster com os hidrometeoros definidos de Besic et al 2016.

#	Cl.	Var.	Mean	STD	Q1%	Q5%	Q10%	Q25%	Q50%	Q75%	Q90%	Q95%	Q99%
nc	5	ZH	22.95	6.57	-1.0	5.0	8.5	13.5	18.0	21.0	24.0	26.0	34.0
nc	5	ZDR	1.16	0.91	-0.83	-0.28	-0.05	0.35	1.06	1.69	2.16	2.55	3.34
nc	5	KDP	0.07	0.28	-0.47	-0.26	-0.18	-0.06	0.05	0.17	0.32	0.47	0.96
nc	5	RhoHV	0.97	0.03	0.81	0.91	0.94	0.97	0.98	0.99	1.0	1.0	1.0
nc	5	CS	convf										
nc	6	ZH	36.36	5.6	21.0	22.0	23.5	26.0	29.0	32.5	36.5	40.0	48.0
nc	6	ZDR	1.34	0.78	-0.28	0.11	0.35	0.74	1.21	1.69	2.31	2.71	3.5
nc	6	KDP	0.44	0.86	-0.31	-0.14	-0.07	0.05	0.21	0.46	1.01	1.83	4.73
nc	6	RhoHV	0.98	0.01	0.94	0.96	0.97	0.98	0.99	0.99	1.0	1.0	1.0
nc	6	CS	convf										
nc	7	ZH	35.43	4.48	22.5	26.0	27.5	29.5	32.0	35.5	38.5	40.5	45.5
nc	7	ZDR	0.83	0.8	-1.07	-0.36	-0.13	0.19	0.66	1.29	1.76	2.08	2.71
nc	7	KDP	0.35	0.69	-1.3	-0.64	-0.29	0.08	0.32	0.56	0.91	1.36	3.14
nc	7	RhoHV	0.96	0.04	0.78	0.86	0.91	0.96	0.98	0.99	0.99	0.99	1.0
nc	7	CS	convf										
nc	8	ZH	23.33	3.13	15.5	17.5	18.5	20.0	22.0	24.5	26.5	27.5	29.0
nc	8	ZDR	0.77	0.76	-0.99	-0.52	-0.28	0.11	0.82	1.29	1.61	1.76	2.24
nc	8	KDP	0.24	0.43	-1.12	-0.34	-0.16	0.02	0.21	0.43	0.73	0.97	1.49
nc	8	RhoHV	0.99	0.01	0.95	0.96	0.97	0.98	0.99	1.0	1.0	1.0	1.0
nc	8	CS	convf										
nc	9	ZH	15.77	3.88	3.5	7.5	9.0	12.0	14.5	17.0	18.5	20.0	22.5
nc	9	ZDR	1.08	0.91	-1.07	-0.44	-0.13	0.35	0.98	1.61	2.08	2.47	3.26
nc	9	KDP	0.24	0.37	-0.58	-0.28	-0.17	0.01	0.2	0.44	0.69	0.88	1.31
nc	9	RhoHV	0.97	0.03	0.85	0.91	0.93	0.96	0.98	1.0	1.0	1.0	1.0
nc	9	CS	convf										

Figura 5: características das variáveis do radar polarimétrico para cada cluster com: a mediana, desvio padrão, e vários percentis.

Referências bibliográficas:

- Al-Sakka H, Boumahmoud AA, Fradon B, Frasier SJ, Tabary P. 2013. A new fuzzy logic hydrometeor classification scheme applied to the French X-, C-, and S-band polarimetric radars. *J. Appl. Meteorol. Clim.* 52: 2328–2344.
- Bechini, R. and Chandrasekar, V.: A Semisupervised Robust Hydrometeor Classification Method for Dual-Polarization Radar Applications, *J. Atmos. Ocean. Technol.*, 32, 22–47, 2015.
- Besic, N., Figueras i Ventura, J., Grazioli, J., Gabella, M., Germann, U., and Berne, A.: Hydrometeor classification through statistical clustering of polarimetric radar measurements: a semi-supervised approach, *Atmos. Meas. Tech.*, 9, 4425-4445, <https://doi.org/10.5194/amt-9-4425-2016>, 2016.
- Bringi VN, Rasmussen RM, Vivekanandan J. 1986. Multiparameter radar measurements in Colorado convective storms. Part I: Graupel melting studies. *J. Atmos. Sci.* 43: 2545–2563.
- Micael A. Cecchini, Luiz A. T. Machado, Meinrat O. Andreae, Scot T. Martin, Rachel I. Albrecht, Paulo Artaxo, Henrique M. J. Barbosa, Stephan Borrmann, Daniel Fütterer, Tina Jurkat, Christoph Mahnke, Andreas Minikin, Sergej Molleker, Mira L. Pöhlker, Ulrich Pöschl, Daniel Rosenfeld, Christiane Voigt, Bernadett Weinzierl, and Manfred Wendisch *Atmos. Chem. Phys.*, 17, 10037-10050, <https://doi.org/10.5194/acp-17-10037-2017>, 2017
- Dolan B, Rutledge SA. 2009. A theory-based hydrometeor identification algorithm for X-band polarimetric radars. *J. Atmos. Oceanic Technol.* 26: 2071–2088.
- Grazioli J, Tuia D, Berne A. 2015. Hydrometeor classification from polarimetric radar measurements: A clustering approach. *Atmos. Meas. Tech.* 8: 149–170.
- Höller H, Hagen M, Meischner PF, Bringi VN, Hubbert J. 1994. Life cycle and precipitation formation in a hybrid-type hailstorm revealed by polarimetric and Doppler radar measurements. *J. Atmos. Sci.* 51: 2500–2522.

Jäkel, E., Wendisch, M., Krisna, T. C., Ewald, F., Kölling, T., Jurkat, T., Voigt, C., Cecchini, M. A., Machado, L. A. T., Afchine, A., Costa, A., Krämer, M., Andreae, M. O., Pöschl, U., Rosenfeld, D., and Yuan, T.: Vertical distribution of the particle phase in tropical deep convective clouds as derived from cloud-side reflected solar radiation measurements, *Atmos. Chem. Phys.*, 17, 9049-9066, <https://doi.org/10.5194/acp-17-9049-2017>, 2017.

Machado, L. A. T., Calheiros, A. J. P., Biscaro, T., Giangrande, S., Silva Dias, M. A. F., Cecchini, M. A., Albrecht, R., Andreae, M. O., Araujo, W. F., Arttaxo, P., Borrmann, S., Braga, R., Burleyson, C., Eichholz, C. W., Fan, J., Feng, Z., Fisch, G. F., Jensen, M. P., Martin, S. T., Pöschl, U., Pöhlker, C., Pöhlker, M. L., Ribaud, J.-F., Rosenfeld, D., Saraiva, J. M. B., Schumacher, C., Thalman, R., Walter, D., and Wendisch, M.: Overview: Precipitation Characteristics and Sensitivities to the Environmental Conditions during GoAmazon2014/5 and ACRIDICON-CHUVA, *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/acp-2017-990>, in review, 2017

Ryzhkov AV, Schuur TJ, Burgess DW, Heinselman PL, Giangrande SE, Zrnić DS. 2005. The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification. *Bull. Am. Meteorol. Soc.* 86: 809–824.

Steiner, M., R.A. Houze, and S.E. Yuter, 1995: Climatological Characterization of Three-Dimensional Storm Structure from Operational Radar and Rain Gauge Data. *J. Appl. Meteor.*, 34, 1978–2007.

Straka JM, Zrnić DS. 1993. An algorithm to deduce hydrometeor types and contents from multiparameter radar data. Preprints, 26th Conference on Radar Meteorology, Norman, OK: . 513–515 American Meteorological Society: Boston, MA.

Vivekanandan J, Ellis SM, Oye R, Zrnić DS, Ryzhkov AV, Straka J. 1999. Cloud microphysics retrieval using S-band dual-polarization radar measurements. *Bull. Am. Meteorol. Soc.* 80: 381–388.

Ward, J.: Hierarchical grouping to optimize an objective function, *J. Am. Stat. Assoc.*, 58, 236–244, 1963.

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. Oceanic Technol.*, 32, 1320–1340.