Intelligent Understanding and Processing of Airborne Sense and Avoid Radar Data with Antenna Diversities

Zhengzheng Li[1], Yan Zhang[2]

ABSTRACT. A technology for the knowledge-based detection and classification of aerospace hazards through on-board radar sensor is discussed, the expected hazards are mainly from hydrometeor. The technology takes advantage of fully-diversified radar systems (multiple channels, dual-polarizations and potentially multi-frequencies (S-C-X bands), and is based on a Bayesian-type, Gaussian mixture self-learning scheme. Physical models of target scattering are incorporated into the machine-learning framework, and Monte-Carlo sampling is used to address the large amount of random variables. The algorithms are highly adaptive and able to accommodate different types of hazards. Simulations and ground test results are presented as initial validations.

1. INTRODUCTIONs

Sense-and-Avoid is becoming a key capability for the future operations of unmanned aero-vehicles (UAV). Although the GPS and satellite links can provide relayed ground radar data, the space and time resolutions are not sufficient for critical safety applications. Also, the pilots (human or automatic) need a comprehensive picture of hazard, including both weather and collision objects, in order to make timely and efficient decisions. Using multiple frequency bands, different antenna polarizations and other diversities is the solution to achieve fast-scanning, multiple functions, and deeper insight into hazard physics [1]. On the other hand, the diversities of antenna system raise a true challenge on how to calibrate, interpret, and understand these data intelligently. This paper assumes airborne radar with multiple receive channels and dual linear-polarization is used as the key sensor for hazard detection, and introduces a Gaussian-mixture based detection and classification scheme. The innovative advantages of the data processing scheme originate from a knowledge-driven hazard signature model, which incorporate both the physical models (EM scattering prediction) and the Bayesian probability models. Monte-Carlo sampling is used to handle the large amount of uncertain variables. The techniques developed are validated through realistic, mixed phased weather simulator as well as ground radar data. Overall, an optimized balance among knowledge model complexity and the classification accuracies can be achieved.

2. HYDROMETEOR HAZARD KNOWLEDGE BASE

2.1. The weather and physical scattering models: Monte-Carlo sampling

One well-known difficulty for aerospace hydrometeor hazards detection is the lack of knowledge of radar signatures for different kind of hydrometeor hazards, though there are many existing

[1] Atmospheric Radar Research Center, University of Oklahoma, Zhengzheng.li@ou.edu
[2] Atmospheric Radar Research Center, University of Oklahoma, rockee@ou.edu

results about hydrometeor signature and hydrometeor classification, most of which are for S-band and C-band ground-based radars [2, 3]. There has been no satisfactory model for dual-polarized radar signatures of hydrometeors at X-band, which is widely used by airborne weather radar. The proposed airborne radar system also has to perform in various situations where types, shapes and sizes of hydrometeors may vary significantly, like in summer rain environment or winter storm environment, or in low altitude as well as high altitude, bringing more challenges to the hazard detection system. Since real measurements from radar usually lack of ground truths and clean data (there are more than one type of hydrometer inside a radar resolution volume), radar signatures from simulations has been a major source of knowledge for hydrometeor classification. A single cell Monte-Carlo sampling simulation approach had been developed for this interest. Concept of the simulation is shown in Figure 1.

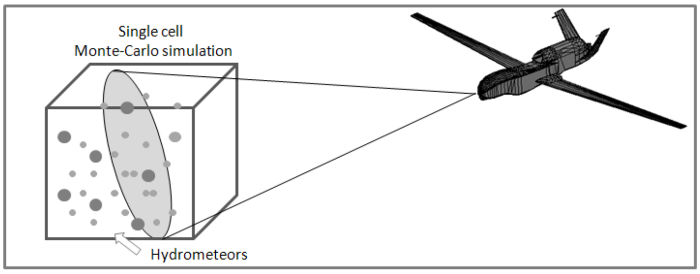


Figure 1. Concept of Single cell Monte-Carlo simulation

Consider a single radar resolution volume (known as *single cell*) that is filled with hydrometeors. The type, total number of hydrometeor and size of a particular particle are controlled by the drop size distribution (DSD). Scattering amplitudes of each particle (hydrometeor) present in the volume can be calculated from T-Matrix ([4]) technique by assuming it has spheroid shape. Therefore, radar return of the whole volume is just the composite of echoes backscattered by all hydrometeors. To make the simulation more realistic, range weighting (caused by pulse shape) and angle weighting (caused by antenna pattern) are also incorporated. Due to limitations of microphysical knowledge and uncertainties of hydrometeors, electromagnetic scattering models and variables including hydrometeor shape, canting angle and drop size distribution (DSD) are allowed to have uncertainties (randomness) to avoid assumptions and losing generality in this study. Detailed description of this simulation and how it corporate with a numerical weather prediction (NWP) model can be found in [5].

Frequency of the simulation is set at 9.41 GHz to match the hardware system under development, and it can be easily changed or extended to dual-frequency applications. Outputs of the simulation are equivalent reflectivity factor at both horizontal and vertical polarization ($Z_{h,v}$), differential reflectivity (Z_{dr}), specific differential phase (K_{dp}), correlation coefficients (ρ_{hv}) and specific attenuation ($A_{h,v}$).

2.2 Scatter plots and statistical interpretations

By artificially setting some physical parameters, we can control the number, size, types and even the melting behavior of the hydrometeors in the volume. For example, we can create a scenario where there is only a certain amount of snowflakes with 20% melting and nothing else in the volume. By generating many scenarios and computing the returns, radar signatures for all the species at different kinds of amounts can be obtained. Since the intercept parameter of DSDs and temperatures are both random variables within a reasonable range, different types of precipitation can also be emulated.

In order to have a statistical significance, a set of 10,000 scenarios for each species including rain, snow, hail, melting snow, melting hail and mixed rain and hail have been carried out. Only one species out of the six species is generated in each scenario. In other words, simulation outputs of each case are 'clear' data for that particular hydrometeor species. Scatter plots of Z_h-Z_{dr} are shown in Figure 2 as an example (More plots can be found in [5], mixed rain and hail is not shown in this figure).

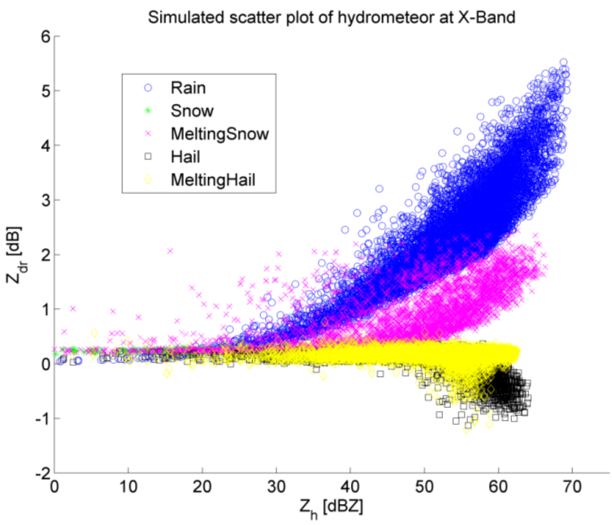


Figure 2. Simulated Scatter plots of Z_h-Z_{dr} at X-Band

3. GAUSSIAN MIXTURE MODEL

Most existing hydrometeor classification techniques are based on fuzzy logic approach, which is flexible and easily adapted to new evidences. However, it fails to model cross correlations among features. For example, in hydrometeor classification, it cannot utilize the underneath relations between reflectivity (Z_h) and differential reflectivity (Z_{dr}) from one observation. Thus, knowledge gained from Monte-Carlo simulations are learned using Gaussian mixture model (GMM) which has been well applied in many areas [6].

As shown in Figure 2, there are five classes of hydrometeors considered. They are *rain, snow, melting snow, hail and melting hail*. Given a radar observation vector X which here includes Z_h,

Z_{dr}, K_{dp}, ρ_{hv}, a Bayesian classifier assigns a hydrometeor class according to its conditional posterior probabilities $p(c_i|X)$ where $i = 1, \ldots, 5$ and c_i is the corresponding hydrometeor class. The classification rule, which is also known as maximum-likelihood rule, is quite intuitive, as the hydrometeor class is obtained by the index c_i that maximizes the conditional posterior probability

$$X \in c_i <=> p(c_i|X) > p(c_j|X) \qquad \forall j \neq i$$

However, conditional posterior probabilities are usually unknown. From Bayes theorem, posterior probabilities can be calculated from prior probabilities and conditional likelihood PDF, which is

$$p(c_i|X) = \frac{p(c_i)p(X|c_i)}{p(X)}$$

Prior probability $p(c_i)$ is set equal here for all classes, even though they are likely different from each other in different temperatures. (Temperature information will be included in future study, since the airborne platform has the advantages of knowing temperature around the radar). $p(X|c_i)$, which is the priori PDF of hydrometeor class c_i, is approximated here using multidimensional Gaussian Mixture Model (GMM) and trained from simulation data.

$$p(c_i|X) = \sum_{m=1}^{M} p_m^{(i)} g(X; \mu_m^{(i)}, \Sigma_m^{(i)})$$

where

$$g(X; \mu_m^{(i)}, \Sigma_m^{(i)}) = \frac{1}{(2\pi)^{(D_0/2)}|\Sigma_m^{(i)}|^{1/2}} e^{-\frac{1}{2}(X-\mu_m^{(i)})^T(\Sigma_m^{(i)})^{-1}(X-\mu_m^{(i)})}$$

M is total number of Gaussian mixtures, here we set $M = 3$. $p_m^{(i)}$ is the prior probability for the m^{th} mixture of class c_i. D_0 is number of dimension of input vector X, here $D_0 = 3$. $\mu_m^{(i)}$ is a column vector ($D_0 \times 1$) containing mean values and $\Sigma_m^{(i)}$ is the covariance matrix ($D_0 \times D_0$) for the m^{th} mixture of class c_i. "T" and "-1" stand for matrix transpose and inverse, respectively.

$p_m^{(i)}$, $\mu_m^{(i)}$ and $\Sigma_m^{(i)}$ for each hydrometeor class are trained by using Expectation-Maximization (EM) algorithm ([6]). Therefore, the classifier assigns each observation a class by maximizing the conditional prior probability, as given in next equation.

$$X \in c_i <=> p(X|c_i) > p(X|c_j) \qquad \forall j \neq i$$

Figure 3 shows the GMM model for rain in Z_h and Z_{dr} dimensions. By comparing it with the scatter plot (figure 2), we can see GMM model well presents the original distribution.

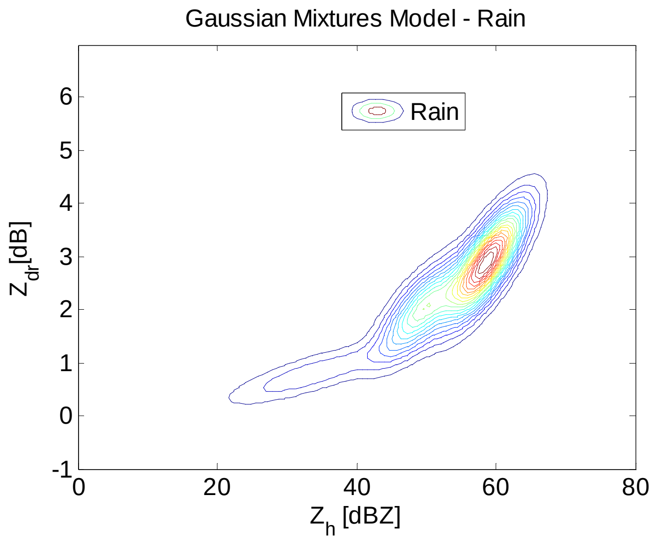


Figure 3. Gaussian mixtures model for rain (Z_h-Z_{dr})

Data from simulation has been randomly divided into two parts. 80% data for training and 20% data for testing. Table.1 lists the accuracy of this classifier. As is shown, the classifier does a good job in distinguishing rain, snow and hail but makes some mistakes when telling if snow/hail is melting. That is probably because no temperature information is used. It is hard to detect melting when dual-polarized radar measurements mainly contain shape information of hydrometeor.

	Rain	Snow	MeltingSnow	Hail	MeltingHail
Rain	**0.9855**	0.0075	0.005	0	0.002
Snow	0	**0.798**	0.202	0	0
MeltingSnow	0.044	0.3515	**0.6045**	0	0
Hail	0	0	0.002	**0.6695**	0.3285
MeltingHail	0.0005	0	0.001	0.161	**0.8375**

Table 1. Classification accuracy of GMM model for airborne radar system (X-band)

4. Attenuation Correction

In X-band, as wavelength is close to diameter of hydrometeors, attenuation effects cannot be neglected. Power measurements such as Z_h, Z_{dr} and ρ_{hv} are significantly biased and needed to be corrected. Phase measurement, in the other hand, is not affected by attenuation. With dual polarization capability, specific differential phase (K_{dp}) can be measured and used to estimate attenuation in both horizontal and vertical polarization. Scatter plots (only Rain is shown in figure 4) show that linear relations between $A_{h,v}$ and K_{dp} can be assumed for all classes as given in the following equation.

$$A_{h,v} = \beta \times K_{dp}$$

β values for all classes can also be learned from simulation outputs. Table 2 gives the RMSE (root mean square error) of attenuation correction for rain, snow(dry and wet) and hail(dry and wet).

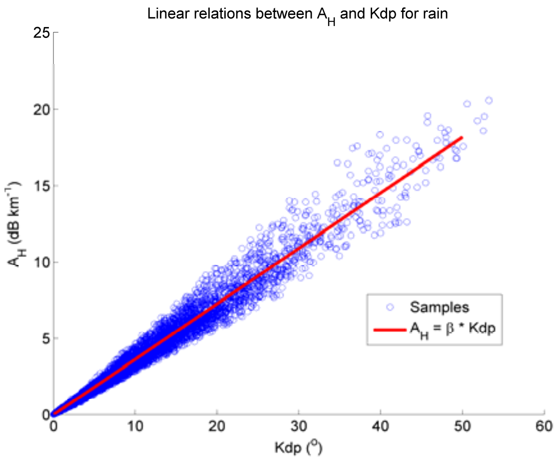


Figure 4. Linear relations between A_h and K_{dp} for rain

	Rain	Snow	Hail
AH	0.6217	1.1082	2.7551
AV	0.6347	1.0240	2.4988

Table 2. RMSE of attenuation correction in dB km^{-1}

5. ATMOSPHERIC HAZARD DETECTION AND CLASSIFICATION

5.1 Example of onboard hail-storm detections

Detection accuracy shown in Table 1 is results for 'clear' data, which means only one species existing in a radar resolution volume. However, in real weather field, there are usually more than one species considering the huge size of one radar resolution volume. Therefore, performance of the system needs to be inspected in simulated weather field where different species mixes with each other. As attenuation in X-band cannot be ignored, how attenuation affects the system also need to be examined. Since we expect the hazards detection system to detect dominant hail species from rain or snow background, only three species including rain, snow and hail are considered (melting hail or melting snow are considered as hail or snow, respectively).

As airplane would fly in different layers in the atmosphere, two cases generated from numerical weather prediction model ([7-9]) are studied. Simulated radar returns includes Z_h, Z_{dr}, K_{dp} and ρ_{hv} (Only plots of Z_h with and without attenuation are shown in Figure 5 and Figure 8) are used as input to the hazards detection system. In case 1, A PPI scan is generated at about 11 km within the stratosphere where only snow and hail exist. Figure 5 shows the weather field for this case. In case two, simulated PPI scan is generated within a melting layer where weather conditions are more complicated. Not only rain but also hail and wet melting hail exist in this layer. Figure 8 show the weather field for case 2.

In case 1, in front of the radar there is a large snow mixing hail area. Although the snow-mixing ratio is not high, at about 15 km ahead of the airplane, there is a region with very high hail density as shown in figure 4. In simulated PPI scans with and without attenuation (figure 6), attenuation effects that lead to 3 to 5 dB difference in reflectivity are shown. As shown in Hazard detection results (Figure 7), the classifier performs very well in this case. When no attenuation in radar returns, the classifier picks up almost all the region where hailstone exists and labels out other snow area. Results completely match the underneath weather field. Even when radar returns are affected by attenuation, though it makes some mistakes (labels some area as rain), the classifier is still able to identify most of the hail regions.

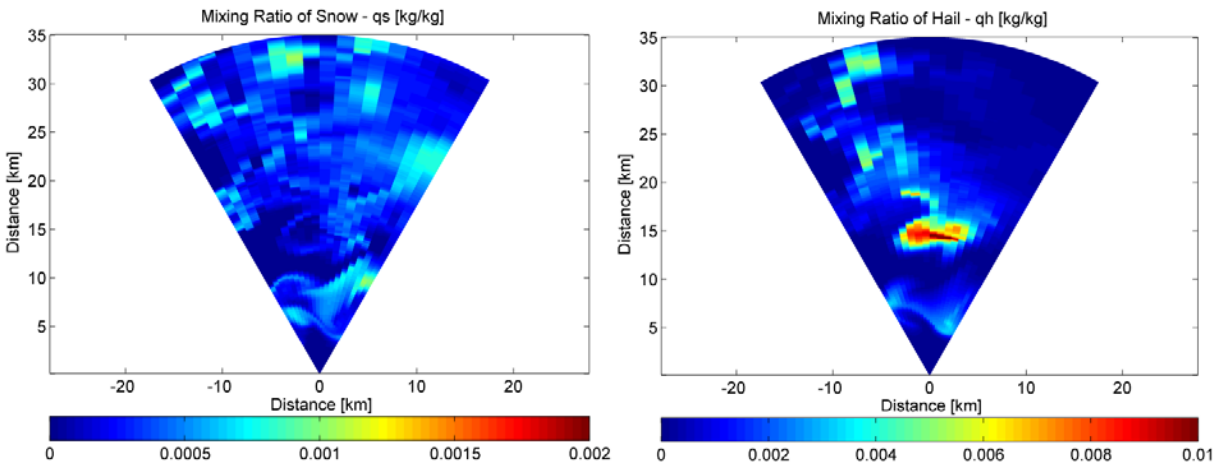


Figure 5. Weather field for case 1

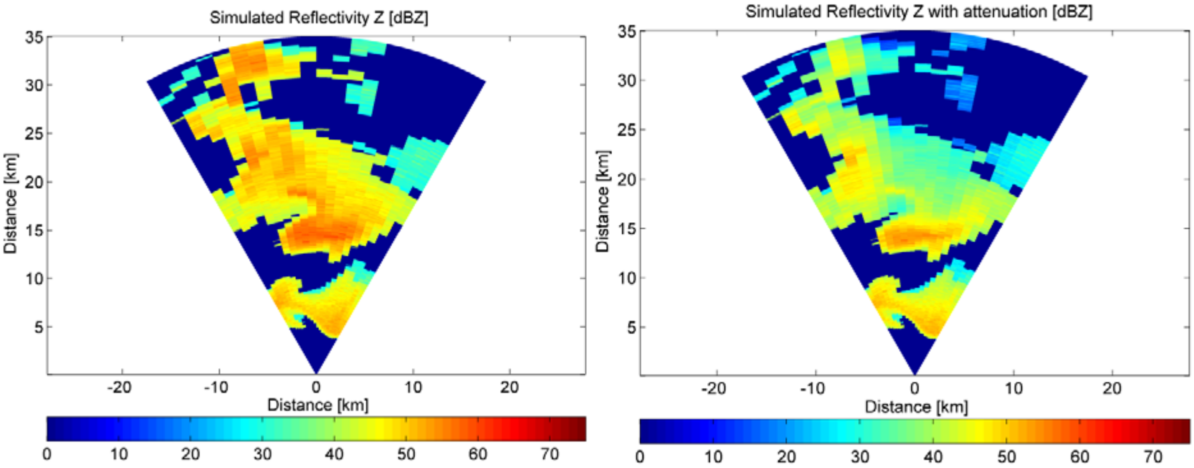


Figure 6. Simulated PPI scan for case 1 (with and without attenuation)

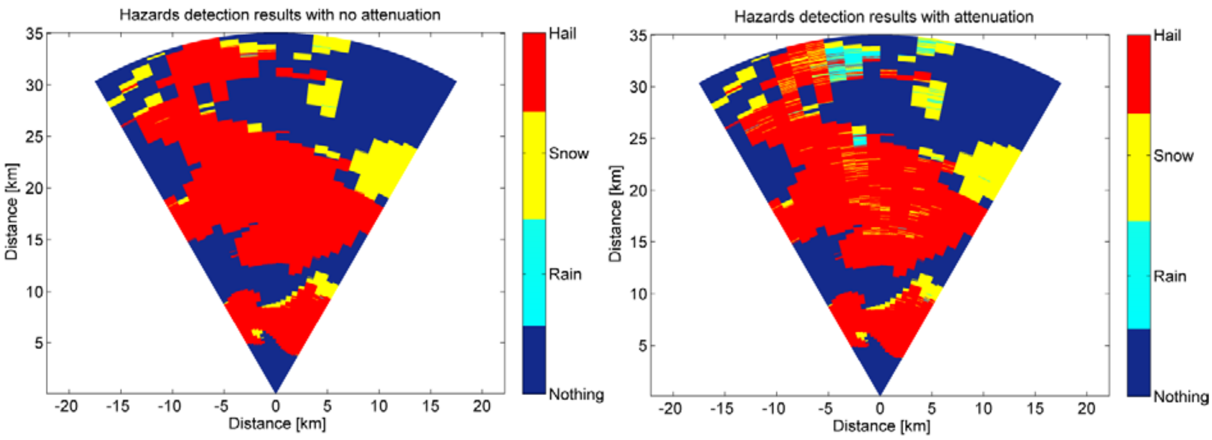


Figure 7. Hazards detection results for case 1 (with and without attenuation)

In case 2, in front of the radar there is a large rain mixing hail area. Rain mainly locates at the left side while hail is everywhere. Both mixing ratios of rain and hail are very high at about 15 km ahead of the airplane as shown in figure 8. Shown in simulated PPI scans with and without attenuation (Figure 8), reflectivity attenuates greatly after the high rain-hail density area. There is about 20 dB difference. As shown in Hazards detection results (figure 10), the classifier also performs well for this case. When no attenuation in radar returns, the classifier identifies most regions where hail is dominant species except the high rain-hail density area. This is acceptable since the model is trained from 'clear' data. It works well in areas where one species is dominant but it is hard to predict when no dominant species exists. More studies on this issue have been performed where both hail and rain are put in the single cell at different mixing levels. We hope to model it using GMM in the further. Performance of classifier is bad when high attenuation exists. After attenuation correction, most highly attenuated area has been recovered. Though it is not perfect, classification result is much better.

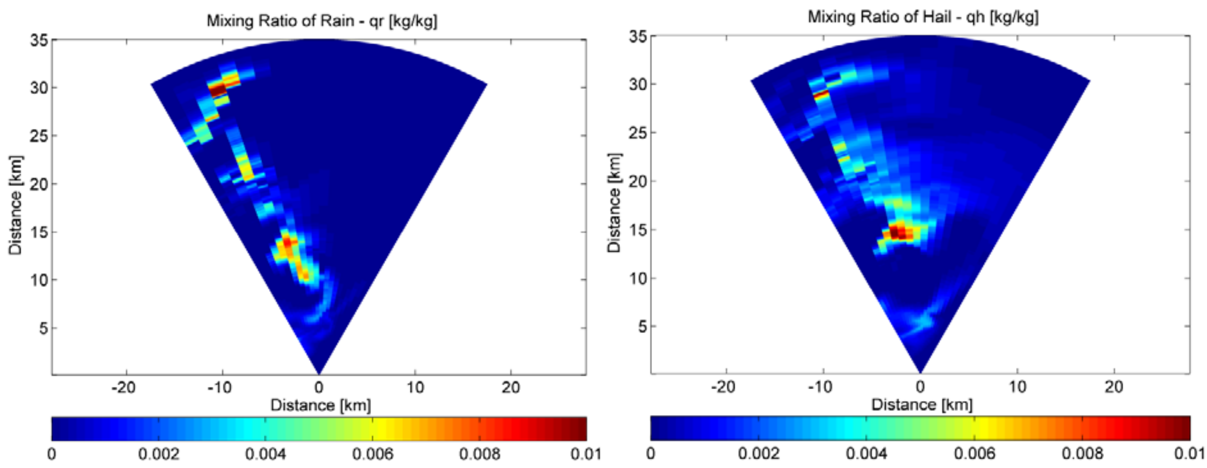


Figure 8. Weather field for case 2

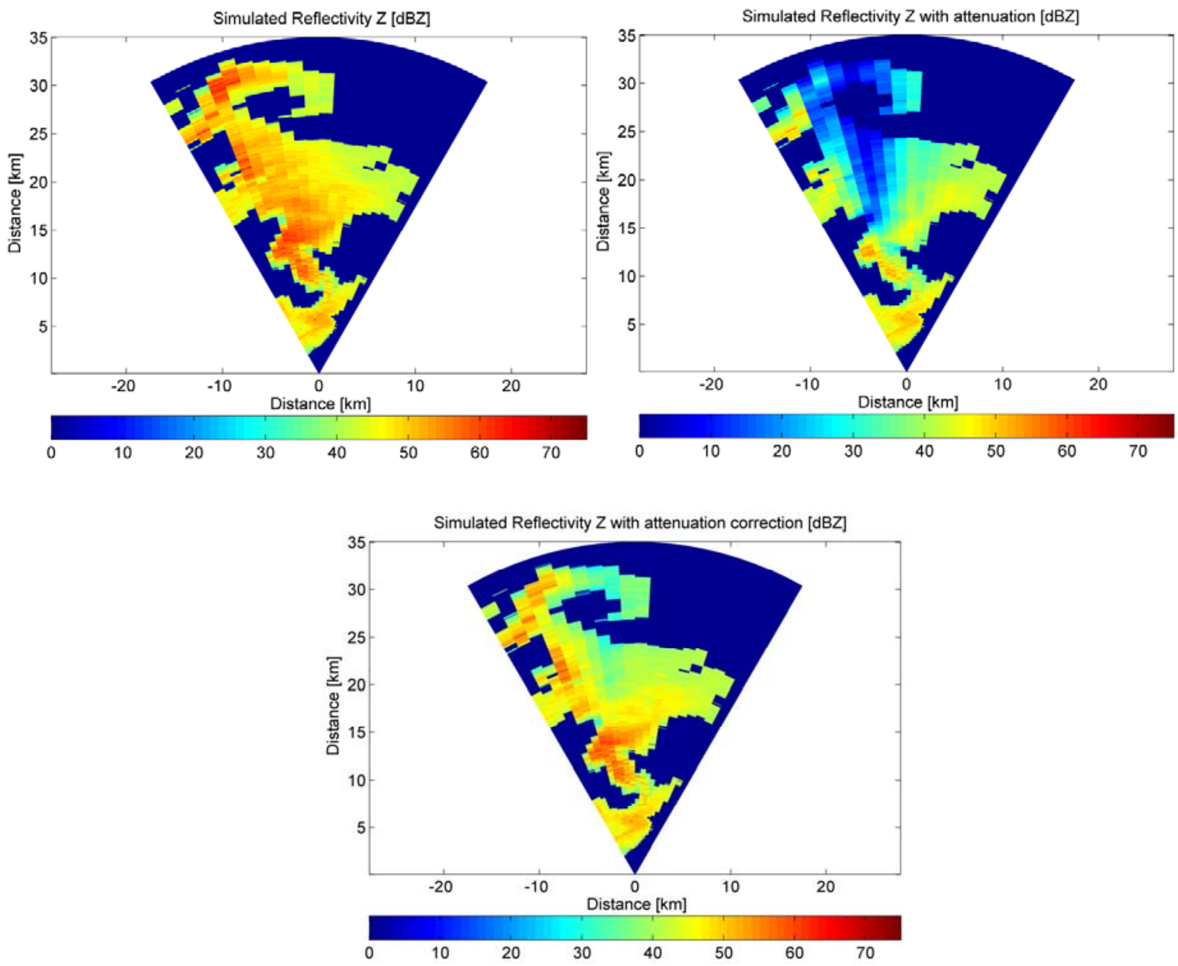


Figure 9. Simulated PPI scan for case 2 (with and without attenuation and after attenuation correction)

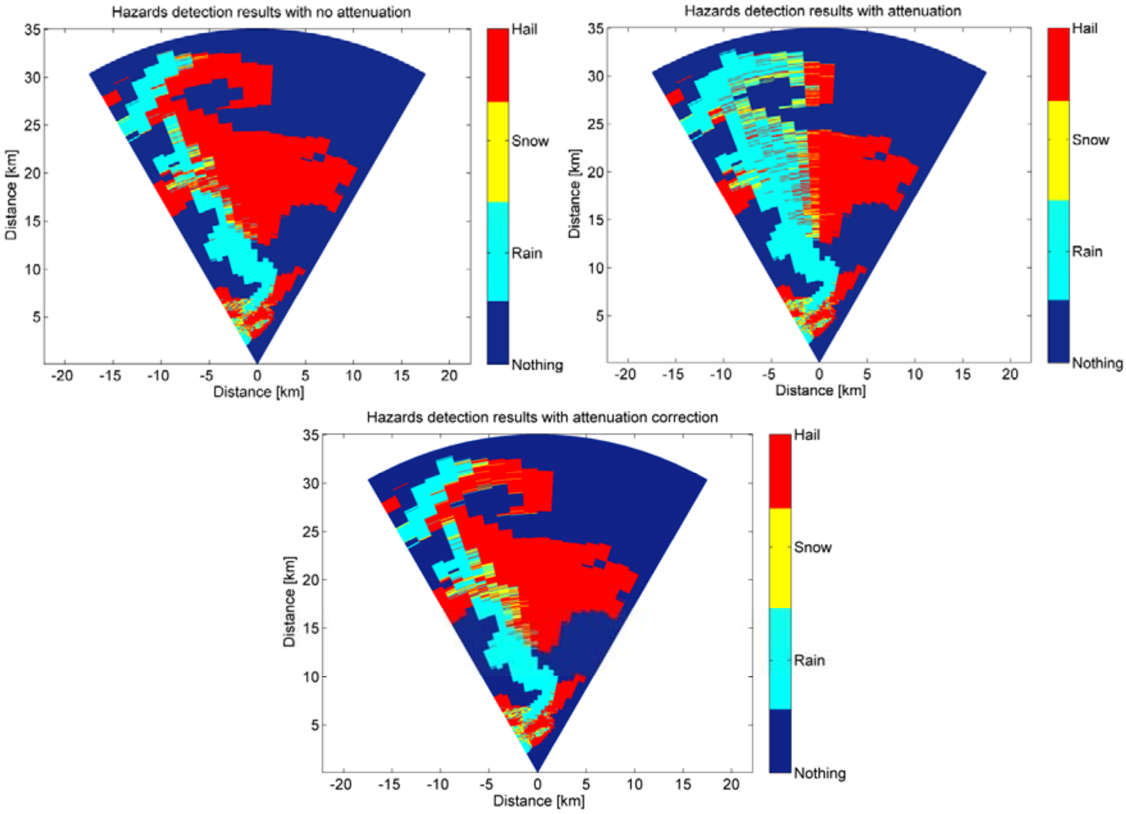


Figure 10. Hazards detection results for case 2(with and without attenuation and after attenuation correction)

5.2 Ground-based OU' data

To further validate knowledge obtained from single cell Monte-Carlo simulation and the GMM hazards detection technique in realistic, noisy environment, ground-based radar observations of a winter storm that happened in Oklahoma are adopted. Those dual-polarization radar observations are from OU-PRIME, which is a high resolution, C-band radar operated by OU-ARRC. According to operating frequency of OU-PRIME, simulations of 10,000 cases have been performed on 5.51GHz without any noise effects. Since phase measurements (K_{dp}) are much easier contaminated by system noise than power measurements (Z_h,Z_{dr},ρ_{hv}), only Z_h,Z_{dr},ρ_{hv} are used as input to the classifier. Following the same steps as stated in section 2, a GMM classifier with 3 mixtures are trained using 80% of the simulation data. The remaining 20% are used for testing. Table.3 lists the accuracy for the classifier whose performance is similar to the one in X-band.

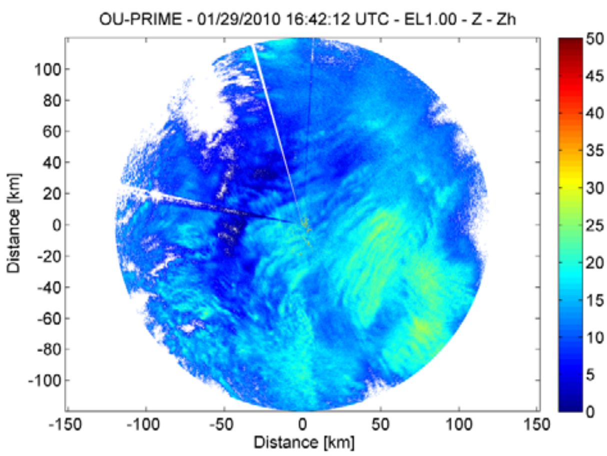
	Rain	Snow	MeltingSnow	Hail	MeltingHail
Rain	**0.9655**	0	0.0340	0	0.0005
Snow	0.0010	**0.7375**	0.2615	0	0
MeltingSnow	0.0645	0.2725	**0.6630**	0	0
Hail	0.0035	0.0035	0.0025	**0.9260**	0.0645
MeltingHail	0.0035	0.0035	0.0020	0.3315	**0.66**

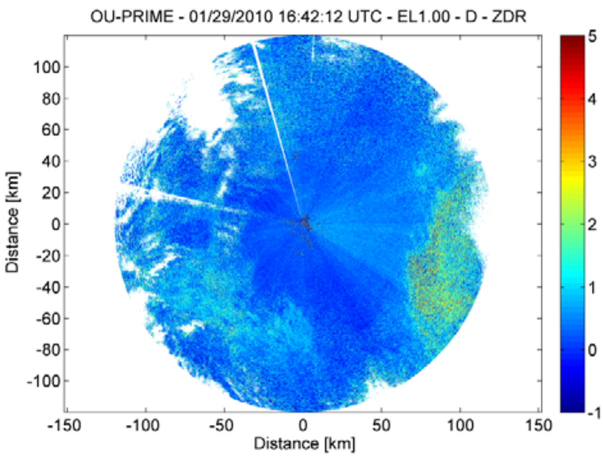
Table 3. Classification accuracy of GMM model for OU-PRIME (C-band)

Figure.11 shows the radar observations and classification results. It was raining then turned into snow that day. As we can see, classification results are reasonable, showing that melting snow and melting hail (ice crystals) everywhere except for low ρ_{hv} areas where SNR are low.

6. CONCLUSIONs

Aerospace hazards mainly hydrometeor classification/detection is discussed. The proposed GMM model performs well in general, even though it may fail to distinguish hail from high density of rain. Single cell Monte-Carlo simulation provides a framework to accommodate various types of aviation hazards beyond hydrometeors. Hard targets such as airplanes and birds can be put into the single cell with hydrometeors, and generate realistic radar signatures for comprehensive monitoring in the future. Attenuation correction techniques, which can also be learned from simulations data, will be incorporated in GMM model to improve system performance in high attenuation situations.





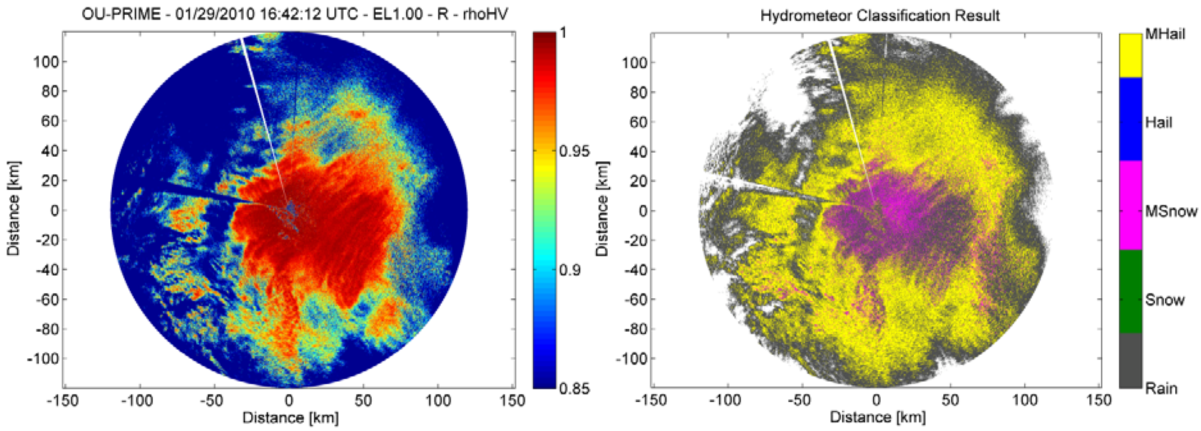


Figure 11. Radar measurements from OU-PRIME and hydrometeor classification result

REFERENCES

[1] Y. Zhang, R. Palmer, G. Z. T.-Y. Yu *et al.*, "Multi-functional Airborne External Hazard Monitoring Radar with Antenna Diversity," *SPIE Remote Sensing Applications for Aviation Weather Hazard Detection and Decision Support conference, San Diego, California, Aug 10-14, 2008*, 2008.

[2] F. S. Marzano, D. Scaranari, and G. Vulpiani, "Supervised Fuzzy-Logic Classification of Hydrometeors Using C-Band Weather Radars," *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING,* vol. 45, no. 11, pp. 3784-3799, 2007.

[3] J. M. Straka, D. S. Zrnić, and A. V. Ryzhkov, "Bulk Hydrometeor Classification and Quantification Using Polarimetric Radar Data: Synthesis of Relations," *Journal of Applied Meteorology,* vol. 39, no. 18, pp. 1341-1372, 2000.

[4] M. I. Mishchenko, "Calculation of the Amplitude Matrix for a Nonspherical Particle in a Fixed Orientation," *Applied Optics,* vol. 39, no. 6, pp. 1026-1031, 2000.

[5] Z. Li, Y. Zhang, G. Zhang *et al.*, "A Microphysics-Based Simulator for Advanced Airborne Weather Radar Development," *Revision I Submitted to IEEE Transactions on Geoscience and Remote Sensing*, 2010.

[6] S. Russell, and P. Norvig, "Artificial Intelligence: A Modern Approach," *Prentice Hall*, 2010.

[7] M. Xue, D. Wang, J. Gao *et al.*, "The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation," *Meteorology and atmospheric physics,* vol. 82, pp. 139-170, 2003.

[8] M. Xue, K. K. Droegemeier, V. Wong *et al.*, "The Advanced Regional Prediction System (ARPS) : A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications," *Meteorology and atmospheric physics,* vol. 76, pp. 143-165, 2001.

[9] M. Xue, K. K. Droegemeier, and V. Wong, "The Advanced Regional Prediction System (ARPS): A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification," *Meteorology and atmospheric physics* vol. 75, pp. 161-193, 2000.