1. Introduction

The variational scheme developed and evaluated by Hogan (2007) for S-band radar for estimating rain rate and detecting hail, is used here specifically for X-band application to correct the measured Z<sub>r</sub> and Z<sub>dr</sub> for attenuation due to mixed phase precipitation (rain mixed with wet ice/hail) which is an especially severe problem for lower power, short range X-band radars. While a number of algorithms for attenuation-correction (based on differential propagation phase) are available at X-band when only rain occurs along the propagation path, there is no stable algorithm as yet when rain is mixed with wet ice/hail. Hence this paper is aimed at applying the variational scheme, which in its formulation, estimates the attenuation due to both rain and hail along the propagation path. We report on several numerical experiments with the variational code to optimize the attenuation-correction in mixed phase precipitation at X-band radar data in a convective storm with rain and wet ice along the path. The preliminary improvements of the variational scheme using CASA X-band radars in Oklahoma) were adapted based on (Z<sub>dr</sub> and Z<sub>r</sub>) to pre-identify hail along the path and thereby initialize the detection of hail in the variational scheme. This was followed by pre-estimating the “reflectivity weighted fraction of ice” using the “deviation from the rain line” methodology as an initialization for the “fraction of ice” in the variational scheme. This scheme then finds an optimal solution for attenuation-corrected Z<sub>r</sub> and Z<sub>dr</sub> due to rain and wet ice along the path. The preliminary improvements of the variational scheme using CASA X-band radar data in a convective storm with rain and wet hail are evaluated.

2. Variational method and forward model (FM)

In Hogan (2007), a method was described, which applies the variational approach to rainfall rate retrieval at S-band from the polarization radar variables reflectivity Z, differential reflectivity Z<sub>dr</sub> and differential propagation phase Φ<sub>dp</sub>. This methodology, also known as “optimal estimation theory”, was used mostly in satellite retrievals, but has only recently been applied to radar applications (e.g., Austin and Stephens (2001), Löhnert et al. (2004)). This method was shown to successfully overcome problems with other techniques, which appear due to inherent measurement fluctuations or “noise” in radar variables (Z<sub>r</sub> and Φ<sub>dp</sub>). The Φ<sub>dp</sub> as the range derivative of an already noisy Φ<sub>dp</sub> can become negative, which is physically impossible in rain. Furthermore, it is difficult to design conventional algorithms to make use of Z<sub>r</sub> and Φ<sub>dp</sub> simultaneously in all rain/hail regimes, so the most appropriate one usually has to be chosen (Hogan (2007)).

The forward model, which is the essence of the variational method, uses the first guess of state vector consisting of the ln(a) for each gate of the beam, where coefficient a is the coefficient between reflectivity Z<sub>r</sub> and the rain rate R:

\[ Z_r = a R^b \]  

(1)

Then these values are used as an input to the forward model to predict the observations at each gate (Z<sub>r</sub> and Z<sub>dr</sub>). The difference between predicted and observed variables is used to change the state vector for better fit with the observations in a least squares sense. This is done by minimization of the cost function, which was defined as

\[ J = \frac{1}{2} \sum_i \left( \frac{Z_{r,i} - Z_{r,\text{pred},i}}{\sigma_{Z_{r,i}}} \right)^2 + \frac{\left( \phi_{i,\text{obs}} - \phi_{i,\text{pred}} \right)^2}{\sigma_{\phi_{i,\text{obs}}}} + \sum_i \left( x_i - x_{i,\text{a priori}} \right)^2 \]  

(2)

where first two summations represent the deviation of the observations Z<sub>r</sub> and Φ<sub>dp</sub> from the values predicted by the forward model Z<sub>r</sub> and Φ<sub>dp</sub>, respectively, and the third summation represents the deviation of the elements of the state vector from some a priori estimate \( x^a \) (a priori \( a=200 \text{ mm}^3 \text{ m}^{-1} \text{ (mm h}^{-1} \text{)}^{1.5} \)). The terms \( \sigma_{Z_{r,i}} \) and \( \sigma_{\phi_{i,\text{obs}}} \) are the root-mean-square observational errors, and \( \sigma_{x,i} \) is the error in the a priori estimate, \( n \) is a set of basis
functions, typically \(\sim m/10\). This minimization process would be repeated until convergence is reached.

This method also can be used to find gates with hail and estimate the fraction of reflectivity due to the hail. If there is a hail segment in the ray, this scheme cannot find a solution for \(\ln(a)\) that, when used in the forward model, can closely predict both \(Z_{dr}\) and \(\Phi_{dp}\), so it is done in 2 passes. The first pass is used for detection of the gates with the hail, and second is used for estimation of the fraction of the measured reflectivity due to hail \(f\). The attenuation-corrected reflectivity is calculated from the observed reflectivity as:

\[
Z_{h\_corrected} = 10^{\frac{0.14}{\text{dBZ}} Z_{h\_observed}}
\]

where \(A_h\) is the total 2-way attenuation at horizontal polarization (path integrated attenuation, PIA) in dB. A similar equation for the corrected \(Z_{dr}\) can be written.

Using S-band data, Hogan (2007) has shown that the optimal estimation scheme produces good results for S-band, but application of it to X-band radar data (CASA radars) does not produce equally good results. The observed radar values of \(Z_{dr}\) and \(\Phi_{dp}\) cannot be predicted by the FM (forward model) well enough in some cases, especially for beams which go through the storm core. The following sections of the paper describe our approach to the improvement of existing variational scheme with the goal to achieve better performance.

3. Variable observational errors in the cost function

One way to improve the performance of the optimal estimation scheme is to adjust the default errors assigned to \(\Phi_{dp}\) and \(Z_{dr}\) values, which are used as an input data into the described algorithm.

For CASA radars, the root-mean-square observational error for \(Z_{dr}\) data, i.e., \(\sigma_{Z\_dr}\) has a default value of 0.5 dB. For low rain rate areas (drizzle) the values of \(Z_{dr}\) are expected to be less than 20 dBZ, and there the observational error should be higher than this default value.

After examination of the CASA radar quality and some numerical experiments it was found that value of \(\sigma_{Z\_dr}\) in first approximation could be changed according to the empirical formula

\[
\sigma_{Z\_dr}(Z_{h}) = \begin{cases} 
0.05Z_{h} & \text{if } Z_{h} < 20 \text{ dBZ} \\
\frac{1}{Z_{h}} & \text{if } Z_{h} > 20 \text{ dBZ} 
\end{cases}
\]

where 0.5 dB is the default value of the observational error. Figure 1 shows the dependency \(\sigma_{Z\_dr}(Z_{h})\).

It should be noted that from theory \(\sigma_{Z\_dr}\) should be dependent on the intrinsic copolar correlation coefficient and the SNR according to eq (6.115) of Bringi and Chandrasekar (2001). The empirical equation in (4) only approximates the dependence on SNR.

The default value of root-mean-square observational error \(\sigma_{\Phi\_dp}\) for \(\Phi_{dp}\), for CASA data is 3 deg. Since \(\sigma_{\Phi\_dp}\) is theoretically related to \(1 - \rho_{hv}^2\) (eq. 6.143 of Bringi and Chandrasekar 2001), after examination of radar data and some numerical experiments it was found that \(\sigma_{\Phi\_dp}\) in the first approximation could be changed according to the formula

\[
\sigma_{\Phi\_dp}(\rho_{hv}) = \begin{cases} 
4.44\rho_{hv} + 5 & \text{if } \rho_{hv} < 0.9 \\
1 & \text{if } \rho_{hv} \geq 0.9 
\end{cases}
\]

where 3 deg is the default value of the observational error for \(\Phi_{dp}\) CASA data.

Figure 2 shows the dependency \(\sigma_{\Phi\_dp}(\rho_{hv})\):

Note that the empirical equation in (5) only approximates the theoretical dependence on \(\rho_{hv}\).

Sample data from the CASA radar from the event of June 10th, 2007, 22:12:57 are used to show the effect of using eqs(4,5), see Fig. 3.
Figures 3 and 4 show that even for gates considered as “good”, i.e., with useful weather information, these observational errors can change significantly from using eqs(4,5). As the result of the above procedure of changing the default values of observational errors $\sigma_{Zdr}$, $\sigma_{\Phi dp}$ leads to the re-balancing the influence of that corresponding variable on the cost function. The overall effect is that corresponding forward-modeled range profiles ($Z'_{dr}$ or $\Phi'_{dp}$) tends to be close to the input variable, as in the case when only one variable ($Z_{dr}$ or $\Phi_{dp}$) was used in the input to the program (these variables can be used as an input to the scheme together or switched off if not available in the radar data).

This is illustrated in the following Fig. 5, which show the gate-by-gate comparison of the CASA KCYR 20070610 dataset used above, data from the 2 deg elevation angle sweep, the beam #318 which were generated using constant observational errors (left panel) and variable observational errors (right panel).

It can be seen that $\Phi_{dp}$ modeled by FM and observed $\Phi_{dp}$ are in much better agreement in the case where variable observational errors were used. In addition, the coefficient $\alpha$ in $A_h-K_{dp}$ relationship

$$A_h = \alpha K_{dp}$$

in this case is $\alpha \approx 0.15$, which is reasonable in rain at X-band.

4. Estimation of reflectivity-weighted fraction of ice in a rain-hail mixture

Another way to improve the optimal estimation algorithm is related to the problem of automatic detection of wet ice and hail in the observed precipitation. In its original form, the algorithm when used with X-band data, in some cases, converges to physically unrealistic set of output variables, which are far from the input and show...
saturation in $\Phi_{dp}$ and $A_h$, where they increase up to the maximum allowed values. It was found that this situation happens mostly for beams going through the core of the storm where occurrence of the hail or wet ice is highly probable.

One can "help" the algorithm by detecting the gates with wet ice and hail and supply them to the program instead of just letting it find these gates by itself. To find the gates where the probability to find hail is high, one can use the $H_{dr}$ concept described by Aydin et al. (1986):

$$H_{dr} = Z_{h} - f(Z_{dr})$$

(7)

where $Z_{h}$ is the measured reflectivity, and $f(Z_{dr})$ defined below:

$$f(Z_{dr}) = \begin{cases} 27; & Z_{h} \leq 0 \text{ dB} \\ aZ_{dr} + 27; & 0 < Z_{h} \leq b \\ 60; & Z_{h} > b \end{cases}$$

(8)

For 3 GHz frequency (S-band) and equilibrium rain drop shape model give values $a=16.5$, $b=2$ dB.

To find the parameters of the $f(Z_{dr})$ at X-band, the variables $Z_{h}$ and $Z_{dr}$ were simulated based on the one minute drop size distribution data from pure rain from a 2D-video disdrometer installed near the CP2 radar (Brisbane, Australia). These variables were simulated for pure rain event assuming the latest information on drop axis ratios and canting angles. One can make a plot similar to the one described in Aydin et al. (1986), but this time for X-band, with the purpose to find the curve which is the rain-hail boundary. This simulated data together with the boundary line is shown in Fig. 6:

---

Fig 5. CASA KCYR 20070610 dataset, the 2 deg elevation angle. Gate-by-gate variables comparison for beams #318 generated using constant observational errors (left panel) and variable observational errors (right panel). Note how $\Phi_{dp}$ goes closer to the observed data, and attenuation $A_h$ has more reasonable values in the right panel.
Even though the data used to create this curve was collected in subtropical coastal environment different from the continental environment of Oklahoma where CASA radars are installed, the boundary line separating rain from the rain-hail mixture should be fairly robust.

For X-band the $f(Z_{dr})$ is found to have the following form:

$$
\begin{align*}
\begin{cases}
27; & Z_{dr} \leq 0 \text{ dB} \\
aZ_{dr} + 27; & 0 < Z_{dr} \leq e \text{ dB} \\
bZ_{dr} + c; & e < Z_{dr} \leq g \text{ dB} \\
60; & Z_{dr} > g \text{ dB}
\end{cases}
\end{align*}
$$

where $a=35.56$, $b=7.23$, $c=39.74$, $e=0.45$ dB, $g=2.8$ dB.

The CASA radar data files include $Z_h$ and $Z_{dr}$ variables corrected for attenuation (by other algorithms), so one can use it to find $H_{dr}$ parameter. It appears that values of $H_{dr}$ found in this manner demonstrate spiky behavior, so one might need to apply FIR (finite impulse response) range filter (in this case FIR filter of order 20 was used) before supplying these values into the optimal estimation algorithm. Fig. 8 demonstrates $H_{dr}$ data calculated for CASA KCYR_20070610-221547 “pure rain” case before and after smoothing by FIR filter.

The FM algorithm was modified to accept the $H_{dr}$ data as an input, i.e. gates with high $H_{dr}$ values (where $H_{dr}$>3 dB) were marked as having hail, and so on the second pass the FM tries to calculate the hail fraction $f$ for these gates, as described in Hogan (2007).

After some experiments it was found that this method of selecting gates with high probability of hail and supplying this information to the FM does not produces sufficiently good results. The hail fraction $f$ calculated by the program still demonstrates spiky, not smooth behavior.

With the purpose of further improving the algorithm in part by recognizing gates with hail, one can use the difference reflectivity factor $Z_{dp}$ as it was proposed by Golestani et al. (1989):

$$
Z_{dp}=10 \log_{10}(Z_h-Z_r), \quad Z_{dr} > Z_r, \text{ mm}^6 \text{ m}^{-3}
$$

Fig. 6 $Z_h$ vs $Z_{dr}$ scatterplot representing simulated data for rain-only case at X-band, and rain-hail boundary line designed for X-band.

Fig. 8. $H_{dr}$ data calculated using $Z_h$, $Z_{dr}$ variables (corrected for attenuation using differential phase as a constraint) for CASA KCYR_20070610-221547. The gates with high probability of hail correspond to the $H_{dr}$ values more than 3.5 dB.
\[ \frac{Z_h}{Z_v} = 1 - 10^{-0.1(\Delta Z)} \]  

(11)

where \( Z_h, Z_v \) – reflectivity in horizontal and vertical polarizations, \( Z \) - total reflectivity, \( Z_h \) is reflectivity due to hail, \( \Delta Z \) is horizontal deviation from the rain line, dB. The \( Z_{dp} \) can be used here to estimate the fraction of reflectivity due to hail in the mixed rain and hail precipitation. The simulated (based on the data from 2D video disdrometer) variables \( Z_h \) and \( Z_v \) for pure rain case at X-band were used to find a “rain line” in \( Z_h \) vs \( Z_{dp} \) space (see Fig. 9).

![Fig. 9 Z_h vs Z_{dp} scatterplot from simulations in rain, and the best fit so-called “rain line”.](image)

The equation for the “rain line” which relates \( Z_h \) and \( Z_{dp} \) is:

\[ Z_{dp} = 1.327Z_h - 19.82 \]  

(12)

Since hail approximately gives \( Z_h = Z_v \) due to nearly spherical shapes, in a rain hail mixture, for a given \( Z_{dp} \) value, one estimates the deviation from the rain line as explained in Chapter 7 of Bringi and Chandrasekar (2001). Even though the error in this initial “guess” of \( f_{ice} \) can be quite high, we expect that the optimal scheme will take as input this first “guess” and converge to an optimal value.

As in the case of \( H_{dr} \), we use the \( Z_h \) and \( Z_{dp} \) corrected for rain attenuation (using differential phase constraint) to calculate \( Z_{dp} \) and fraction of reflectivity due to hail \( f_{ice} \) which is used as the “first guess” for the FM algorithm. The algorithm then re-adjusts the final values of \( f_{ice} \) by minimizing the cost function. One can build the \( f_{ice} \) map for the “mixed phase precipitation” case, and compare it to the calculated and smoothed \( H_{dr} \) values for the same case, as shown in the Fig. 10:

![Fig. 10. Fraction of reflectivity due to ice, \( f_{ice} \), for “mixed phase precipitation” case of April 24, 2007 (KSAO_20070424-172558.netcdf) (top panel), \( H_{dr} (> 3dB) \) is shown in bottom panel. Note that high values of \( H_{dr} \) do not always correspond to the high values of \( f_{ice} \).](image)

One can note that high values of \( H_{dr} \) do not always correspond to the high values of \( f_{ice} \). The map of \( f_{ice} \) looks scattered, for the purposes of achieving better “first guess” it is desirable to use spatial averaging of the \( f_{ice} \) data, by using a 5x5 smoothing window. The same matrix can be applied to the \( H_{dr} \) data (figure 11). It can be seen that spatially smoothed \( H_{dr} \) values (top panel) are better correlated with the spatially smoothed \( f_{ice} \) values (bottom panel). But still, high \( f_{ice} \) values can be seen at gates where \( H_{dr} \) values are low. Based on this, the decision was made to supply the FM algorithm with smoothed values of the reflectivity fraction due to ice \( f_{ice} \) as a “first guess”.

As in the case of \( H_{dr} \), we use the \( Z_h \) and \( Z_{dp} \) corrected for rain attenuation (using differential phase constraint) to calculate \( Z_{dp} \) and fraction of reflectivity due to hail \( f_{ice} \) which is used as the “first guess” for the FM algorithm. The algorithm then re-adjusts the final values of \( f_{ice} \) by minimizing the cost function. One can build the \( f_{ice} \) map for the “mixed phase precipitation” case, and compare it to the calculated and smoothed \( H_{dr} \) values for the same case, as shown in the Fig. 10:
As the result of the aforementioned modifications the FM algorithm becomes more "stable", meaning that there are less beams with saturated values of $\Phi_{dp}$, $A_h$, etc, as shown on the figure 12 (see azimuth sector 280-320° at end of beams in top panel versus bottom panel).

Figure 13 shows the comparison of the different variables for beam #200 of the “mixed precipitation” case of April 24, 2007 obtained using both original (blue lines) and modified (red lines) versions of the FM algorithm. One can see the general effect of the improvements discussed above: in the output of the modified version the differential phase $\Phi_{dp}$ can be followed more precisely, as well as $Z_{dr}$, and $f_{ice}$ looks much less spiky and more smooth, and hail is found in much more gates than before.

5. Sensitivity of the variational scheme to the absolute calibration of the CASA reflectivity input variable

It was stated in Hogan (2007) that, “…We are effectively assuming that, in relative terms, the error in $Z_h$ is much less than the errors in $Z_{dr}$ and $\Phi_{dp}$ so that the retrieval should be forced to be exactly consistent with $Z_h$…” but the program might be sensitive to the absolute accuracy (absolute calibration) of the input variable $Z_h$. For the purpose of testing the sensitivity of the algorithm to the input $Z_h$ variable, the data from CASA radar collected at April 24, 2007 at 17:25:58.
(KSAO_20070424-172558.netcdf file) was selected. It is described on the CASA website as “significant severe wx outbreak... Large hail was prevalent along the line”. It was used as an input to the OES three times: original reflectivity and modified values (increased by 3 dB and decreased by 3 dB). One can compare the maximum values of the \( \Phi_{dp} \) and \( A_h \) variables for these 3 sets of data for each beam of the scan (at the end of the beam, where they become maximum). It can be seen from Fig. 14 that for data sets where \( Z_h \) was increased by 3 dB or even left at the original level, the values of \( \Phi_{dp} \) and \( A_h \) tend to saturate for some beams, especially for ones that go through the core of the storm (beams around #230-250). For the dataset where reflectivity values were decreased by 3 dB there are no saturated \( \Phi_{dp} \) and \( A_h \) data. It has to be noted that experiments were done for reflectivity decreased by 1 and 2 dB, and there were areas of saturated data in the OES output, so -3 dB seems to be the minimum for decrease in input reflectivity to avoid spuriously large \( \Phi_{dp} \) values (at least for this dataset).

One can compare the OES output variables for one beam (#246) which goes through the “problem zone” of the scan, which is the core of the storm, as shown on the Fig. 15 (no \( Z_h \) offset and ± 3 dB). From this figure, note that only one where reflectivity values were decreased by 3 dB (red line) can follow the input data (black line) with sufficient accuracy.

In FM, Z-R relation is used to compute rain rate, and \( K_{dp} \) is calculated and used to compute \( \Phi_{dp} \) values for each gate. So if the input variable \( Z_h \) which is assumed to be measured by CASA radars (to within an uncertainty of 1 dB) is in error (i.e., too “hot” by 3 dB) all calculated variables (like \( \Phi_{dp} \), attenuation \( A_h \), \( A_v \)) achieve unrealistically high values, and FM cannot correct for this even after modifications introduced above.
6. Conclusions

The FM algorithm was adapted from S-band to X-band and applied to the CASA IP-1 dual-polarization radar data. The main goal was the correction of measured reflectivity in mixed phase precipitation (rain+ice) which cannot be achieved using differential phase constraints.

Several modifications to the original scheme were applied, with the general goal to increase OES stability. The principal ones being, (a) adjusting the observational errors based on $Z_h$ and $\rho_{hv}$, (b) provide initial ‘guess’ for fraction of ice, and (c) sensitivity tests for the $Z_h$ offset. Overall, it appears that the optimal estimation scheme can be adapted to X-band data for correction of attenuation due to mixed phase precipitation (rain mixed with wet ice/hail) provided the input data is well-calibrated (system offsets for $Z_h$, $Z_{dr}$ and $\Phi_{dp}$).

References


Fig. 15. Three data sets. Shown are forward-modeled (“fwd”) $\Phi_{dp}$, $A_h$, $Z_{dr}$ variables, compared to the input variables (“clean”).