



Toward a variational assimilation of polarimetric radar observation in a convective scale NWP model

Guillaume THOMAS, Jean-François MAHFOUF, and Thibaut MONTMERLE

Centre National de Recherches Météorologiques-Groupe de Modélisation et d'Assimilation pour la Prévision,
Toulouse, France

Correspondence: Guillaume THOMAS(guillaume.thomas@umr-cnrm.fr)

Abstract. This paper presents the potential of non-linear and linear versions of an observation operator for simulating polarimetric variables observed by weather radars. These variables, deduced from the horizontally and vertically polarised backscattered radiations, give information about the shape, the phase and the distributions of hydrometeors. Different studies in observation space are presented, as a first step toward their inclusion in a variational data assimilation context, which is not treated here. Input variables are prognostic variables forecasted by the AROME-France Numerical Weather Prediction (NWP) model at convective scale, including liquid and solid hydrometeor contents. A non-linear observation operator, based on the T-matrix method, allows to simulate the horizontal and the vertical reflectivities (Z_{HH} and Z_{VV}), the differential reflectivity Z_{DR} , the specific differential phase K_{DP} and the copolar correlation coefficient ρ_{HV} . To assess the uncertainty of such simulations, perturbations have been applied on input parameters of the operator, such as dielectric constant, shape and orientation of the scatterers. Statistics of innovations, defined by the difference between simulated and observed values, are then performed. After some specific filtering procedures, shapes close to Gaussian have been found for both reflectivities and for Z_{DR} , contrarily to K_{DP} and ρ_{HV} . A linearised version of this observation operator has been obtained by its Jacobian matrix estimated with the finite difference method. This step allows to study the sensitivity of polarimetric variables to hydrometeor content perturbations, in the model geometry as well as in the radar one. The polarimetric variables Z_{HH} and Z_{DR} appear to be good candidates for hydrometeor initialisation, while K_{DP} seems to be useful only for rain contents. Due to the weak sensitivity of ρ_{HV} , its use in data assimilation is expected to be very challenging.

1 Introduction

For a couple of decades, convective scale Numerical Weather Prediction (NWP) models have been developed to forecast mesoscale meteorological phenomena such as storms, wind gusts and fog, which can represent important socio-economic threats. Nowadays, most of operational convective scale NWP models have fine, km scale, horizontal resolutions (see review by Gustafsson et al. (2018)). In the present study, the AROME-France model from Météo France (Seity et al., 2011) is used with a resolution of 1.3 km (Brousseau et al., 2016). This high resolution allows, in addition to a fully non-hydrostatic compressible set of equations, an explicit representation of the deep moist convection and related dynamical parameters. As such models are run over a specific geographical region, initial conditions and lateral boundary conditions are required. Ducrocq et al.



25 (2002) showed that accurate initial conditions can be more important than lateral boundaries to obtain skillful forecasts with such Limited Area Models (LAMs). Several methods exist to provide initial conditions but, as explained by Gustafsson et al. (2018), better performances are obtained when a convective scale data assimilation step is considered, compared to an initial state downscaled from a global model. In addition to observations that are representative of larger scales, observations at fine spatial resolutions and high sample frequencies are required in order to get an accurate representation of the dynamics occurring
30 at these small scales (Benjamin et al., 2016; Brousseau et al., 2016). It is the case for data produced by weather radars: with a kilometric or finer resolution and few minutes temporal sampling, they are able to provide information about the intensity of precipitating systems through the horizontal reflectivity and about their dynamics from Doppler radial winds.

The dual-polarization radar technology allows to go further in the description of precipitating systems. Seliga and Bringi (1976) were ones of the first to investigate the capabilities of polarimetric radars for a better understanding and representation
35 of precipitating systems. Since then, numerous studies have shown the interest of Dual POLarized (DPOL) variables to improve storm description and related processes. In a first paper, Kumjian (2013a) firstly describes the DPOL variables, their characteristics and ranges of values while, in a second one (Kumjian, 2013b), he explains their usefulness for the detection of meteorological phenomena, such as hail, supercells or bright bands. DPOL variables can also be used to control the radar data quality as, for example, the determination of echo type, using a combination of several polarimetric variables. Gourley et al.
40 (2007) use a fuzzy logic algorithm to distinguish meteorological echoes from non-meteorological ones, which is necessary for improving Quantitative Precipitation Estimation (QPE). Detection of the major hydrometeor type in meteorological echoes can also be done with fuzzy logic algorithms, as proposed by Al-Sakka et al. (2013) for example. Another application of DPOL variables, for QPE improvements, is to use new relationships between Z_{HH} , DPOL variables and rain rate, as in the "Joint polarization experiment" (Ryzhkov et al., 2005).

45 Nowadays however, only the horizontal reflectivity and the Doppler wind are operationally exploited in the retrieval of initial conditions of NWP models. From a Data Assimilation (DA hereafter) point of view, the challenge is to extract useful information about the main control variables from Z_{HH} , which is an indirect observation of model variables. At Météo-France, a two-step method is operationally performed in the AROME-France model (Wattrelot et al., 2014): pseudo-profiles
50 of relative humidity are firstly retrieved from Z_{HH} using a 1D Bayesian inversion (following Caumont et al. (2010)), and then are assimilated as pseudo-atmospheric soundings in a three-dimensional Variational (3DVar) system. In this approach, the complex linearization of the reflectivity observation operator is avoided. However, the prognostic hydrometeor variables are not initialized and they evolve with respect to the analyzed thermodynamical conditions at the beginning of the model integration. When coupled with the assimilation of Doppler winds, storm dynamics and precipitation forecasts are clearly improved, especially when low level wind convergence is sampled (Montmerle and Faccani, 2009). Such procedure is also
55 used operationally at JMA (Ikuta and Honda, 2011) and by some countries of the HIRLAM community (Ridal and Dahlbom, 2017). Other NWP models (e.g UKV at the Met-Office or HRRR at NOAA) make use of Z_{HH} (or radar-based precipitation rate analysis in the case of the UKV) through Latent Heat Nudging procedures (see again Gustafsson et al. (2018) for details). A range of studies have been also undertaken to assimilate Doppler wind and Z_{HH} using methods based on Ensemble Kalman Filter (EnKF) (e.g Tong and Xue (2005), Dowell et al. (2011) or Bick et al. (2016)), mostly for case studies. These methods



60 avoid the linearization of observation operators and can quite straightforwardly consider hydrometeors in the control variable, but they are particularly prone to sampling noise.

In this paper, preliminary work is presented in order to prepare the assimilation of DPOL variables in a convective scale variational DA system. Such system is based on the minimization of a cost function, which is composed of two terms defining distances (i) between the model state and a background and (ii) between the model state in the observation space and the 65 observations. The second term requires non-linear (NL hereafter) observation operators in order to retrieve the model equivalent of every observation at their locations. Statistics between observed and simulated values (called innovations) are used at this stage to quantify the performance of the model in this particular space and to perform quality controls in order to produce innovation distributions that are close to a Gaussian shape, such conditions leading to optimal variational DA results. In this study, the NL observation operator described by Augros et al. (2015) is used to simulate DPOL variables from the AROME-70 France model. This operator is based on the T-matrix approach, which describes scattering by particles (Waterman, 1965). This approach has been used in several studies. For instance, in Bringi et al. (1986), it allows to study the melting of graupels by simulating the differential reflectivity Z_{DR} . It is also used in the observation operator proposed by Jung et al. (2008), with a one moment bulk microphysical scheme, to simulate all DPOL variables. A more complex observation operator has then been proposed by Ryzhkov et al. (2011) with a spectral microphysical scheme. Even though it leads to more physically coherent 75 results, its computational cost is not yet compatible with operational NWP requirements.

When error Gaussianity and operator linearity are respected, the cost function of a variational DA system is close to a quadratic function for which the minimum can be easily obtained by e.g the method of least squares. The estimation of its gradient, which needs linearized versions of the observation operators, is then required. For operators related to precipitation, this is not straightforward as cloud microphysical processes are often highly non-linear due to the presence of on/off switches 80 (Sun, 2005). Furthermore, Errico et al. (2007) pointed out that these non-linearities can severely affect the analysis. Such difficulties explain why the 1D+3D Var approach has been initially preferred in AROME-France as discussed above. In their first attempt to assimilate Z_{HH} in a 4D Var, Sun and Crook (1997) found indeed better results when simply retrieving the rain mixing ratio from an empirical relationship with Z_{HH} instead of directly assimilating Z_{HH} by using a NL observation operator. This approach based on empirical relationships has been more recently extended to solid precipitating species by Gao 85 and Stensrud (2011). Other attempts of direct assimilation of Z_{HH} with encouraging results have been made in 3D-Var (Wang et al., 2013b) and 4D-Var (Wang et al. (2013a); Sun and Wang (2013)). Nevertheless, no operational applications have been performed yet, particularly because only warm microphysical processes are considered. In this paper, before trying to linearize the highly NL DPOL observation operator, its Jacobians have been computed in order to study the sensitivity of simulated polarimetric variables to hydrometeor content perturbations.

90 The main goal of this paper is to study an observation operator of DPOL variables in order to determine its properties and suitability for DA, especially for hydrometeor contents initialisation in the variational context of the AROME-France convective scale NWP model. No assimilations are thus performed yet and only results in the observation space are discussed at this point. The behaviour of the operators presented in this paper in a variational DA system will be the focus of a future paper. Section 2 firstly describes the NL observation operator, a quantification of its errors, and examples of DPOL variables



95 simulation for different meteorological situations simulated by AROME-France. In section 3, innovation statistics are discussed and used to perform quality controls on polarimetric observations. Finally, section 4 focuses on the DPOL observation operator Jacobians to determine the validity of the linearity hypothesis, and to quantify the sensitivity of DPOL variables to the various simulated hydrometeor contents. Conclusions and perspectives from this study are given in section 5.

2 A non-linear polarimetric radar observation operator (HDPOL)

100 2.1 HDPOL description

2.1.1 Generalities

The \mathcal{H}_{DPOL} observation operator has been developed by Augros et al. (2015), and only the main characteristics are summarized here. It uses the T-Matrix method (Mishchenko and Travis, 1994) to compute the backscattering coefficients according to frequency, temperature and type of hydrometeors. The microphysical scheme used to predict hydrometeor contents 105 is the one from the AROME model. This scheme, called ICE3, is a one moment microphysical scheme with water vapour and five hydrometeors species: cloud droplets, rain, snow, pristine ice and graupel (Canaux et al., 1994; Pinty and Jabouille, 1998). In the present study, only the last four have been used for the DPOL simulations, in addition to a melting species which has the characteristics of melting graupel. This melting species represents the sum of the three solid hydrometeor contents when 110 temperature is above 0°C. In this microphysical scheme, the Particle Size Distribution (PSD) of each hydrometeor is expressed as the product between the total number concentration N_0 and generalized Gamma distribution. The slope parameter used to characterize the PSD shapes depends upon the hydrometeor content M (expressed in $kg.m^{-3}$), this last being the ratio between the hydrometeor content (express in $kg.kg^{-1}$) and the density of an air parcel. The parameters describing the hydrometeor PSDs 115 are given in Table 1 of Caumont et al. (2006).

Two other parameters are required for the backscattering coefficient computation: the hydrometeor shape and the dielectric 115 constant. The latter, which describes how a material reacts to the application of electrical field, is simulated by the Debye model for raindrops (Caumont et al., 2006), and by the Maxwell-Garnet mixing formula for ice particles (Ryzhkov et al., 2011). This last formula allows to consider solid hydrometeors as ice particules with air inclusions. For melting hydrometeor 120 species, dielectric constant is computed with a weighted Maxwell-Garnet mixing formula (Matrosov, 2008), which permits to consider melting species as liquid water inclusions in ice and as ice inclusions in liquid water, depending upon liquid water and graupel fractions. For hydrometeor diameters lower than 0.5 mm, all particles are considered as sphericals (axis ratio of 1). For larger diameters, the axis ratios depend upon the hydrometeor types. Rain drops are described as spheroids, with an aspect ratio depending on diameter, in order to account for the flattening which is proportional to their size (Brandes et al., 2002). Concerning snow particles, a spheroid shape is also assumed with axis ratios linearly decreasing from 1 to 0.75 when the particle diameter increases from 0.5 mm to 8 mm. For higher diameters, the minimum value of 0.75 is kept. Same 125 characteristics are used for graupel and for melting hydrometeor species, but with axis ratios linearly decreasing from 1 to 0.85.



Finally, pristine ice is simulated as spherical. All these parameters have been proposed by Augros et al. (2015) as a result of sensitivity studies.

2.1.2 Simulated DPOL variables

Once the hydrometeors characteristics are defined, the T-matrix method is used to compute the back-scattering coefficients for 130 different values of particle diameters, radar elevations, temperatures and water contents (as listed in Table 1 of Augros et al. (2015)). These coefficients are then integrated over diameters and stored in look-up tables to speed up computations. They are then used for the computation of the four DPOL variables of interest from the following equations:

$$Z_{HH,VV} = 10\log(Z_{hh}, Z_{vv}) = 10\log \left(10^{18} \frac{4\pi\lambda^4}{\pi^5 |K_w|^2} \sum_{i=1}^n \int_{D_{min}}^{D_{max}} |S_{HH_i, VV_i}^b(D)|^2 N_i(D) dD \right) \quad (1)$$

where Z_{HH} ¹ and Z_{VV} represent respectively the horizontal and vertical reflectivities (in dBZ), Z_{hh} (Z_{vv}) the horizontal 135 reflectivity (vertical reflectivity) expressed in $mm^6 \cdot m^{-3}$, λ the wavelength (in m), $|K_w|^2$ the dielectric factor, function of the dielectric constant, $N_i(D)$ the number of particles with a diameter D for the hydrometeor type i and $S_{HH_i}^b, S_{VV_i}^b$ the horizontal and vertical backscattering coefficients respectively, b exponent standing for "backward".

$$Z_{DR} = 10\log\left(\frac{Z_{hh}}{Z_{vv}}\right) \quad (2)$$

Z_{DR} being the differential reflectivity (in dB). This variable brings information on target aspect ratio and phase. It can be 140 explained by its dependence upon the ratio between the horizontal and vertical reflectivity when expressed in $mm^6 \cdot m^{-3}$. For spherical hydrometeors (*i.e.* with equivalent horizontal and vertical cross sections), Z_{DR} is equal to 0 dB. This variable will be positive (negative) when the hydrometeor horizontal dimension is larger (smaller) than the vertical one. Z_{DR} is very sensitive to the hydrometeor dielectric factor: liquid hydrometeors will have a higher Z_{DR} value than the solid ones with similar shape and size distribution.

145 ρ_{HV} expresses the copolar correlation coefficient:

$$\rho_{HV} = \frac{\left| \sum_{i=1}^n \int_{D_{min}}^{D_{max}} S_{HH_i}^b(D) \times S_{VV_i}^b(D) N_i(D) dD \right|}{\sqrt{\sum_{i=1}^n \int_{D_{min}}^{D_{max}} |S_{HH_i}^b(D)|^2 N_i(D) dD \times \sum_{i=1}^n \int_{D_{min}}^{D_{max}} |S_{VV_i}^b(D)|^2 N_i(D) dD}} \quad (3)$$

This quantity gives information on homogeneity. When a large variety of hydrometeor sizes, shapes, phases and orientations are represented in the observed volume, ρ_{HV} values will be close to 0.

¹The reader should notice that "horizontal reflectivity" (Z_{HH}) in this paper relates to the horizontal equivalent reflectivity.



The specific differential phase K_{DP} (in $^{\circ}.km^{-1}$), is defined as:

$$150 \quad K_{DP} = 10^3 \lambda \frac{180}{\pi} \sum_{i=1}^n \int_{D_{min}}^{D_{max}} \Re(S_{HH_i}^f - S_{VV_i}^f) N_i(D) dD \quad (4)$$

where $\Re(S_{HH_i}^f - S_{VV_i}^f)$ expresses the difference between the real part of the forward horizontal and vertical scattering coefficients (the f exponent meaning "forward"). This polarimetric variable expresses the phase difference between the horizontal and vertical polarized electromagnetic (EM) wave between a specific distance. In the case of spherical hydrometeors, the same amount of matter will be crossed by these two waves. Therefore, no phase difference will be observed. For non spherical particles, the horizontally and vertically polarized EM wave will have to cross different amounts of matter, which will cause a phase difference. Because this variable only depends upon the phase difference and not upon the cross-section, it is not affected by attenuation, nor by geographical masks.

2.2 Illustration on a case study

To assess qualitatively the ability of \mathcal{H}_{DPOL} to simulate DPOL variables, PPIs (Plan Position Indicators) of the different 160 DPOL variables are compared for one particular meteorological case. On the 10th of October 2018, an important convective event strokes the South of France, mainly because of a strong south-westerly flow that took place off the coast over the Mediterranean sea. It represented an important source of humidity and, because of low-level convergence due to orography, strong precipitation occurred over the Var, Bouches-du-Rhône and Gard departments. Such kind of meteorological events are quite common over the Mediterranean region. For instance, Llasat et al. (2010) report that, between 1990 and 2006, 185 165 flash-flood events occurred around the Mediterranean basin, and about half of them happened during the autumn season. The meteorological event which took place on the 10th of October 2018, has produced more than 100 mm of rain in 24 h over a large area of the Var department, and locally more than 150 mm, with flash floods causing two casualties. In addition to the heavy precipitation, strong wind gusts up to 100 km/h have been observed. Meteorological fields from a 1 h AROME-France forecast, valid at 14 UTC, have been used as input to \mathcal{H}_{DPOL} .

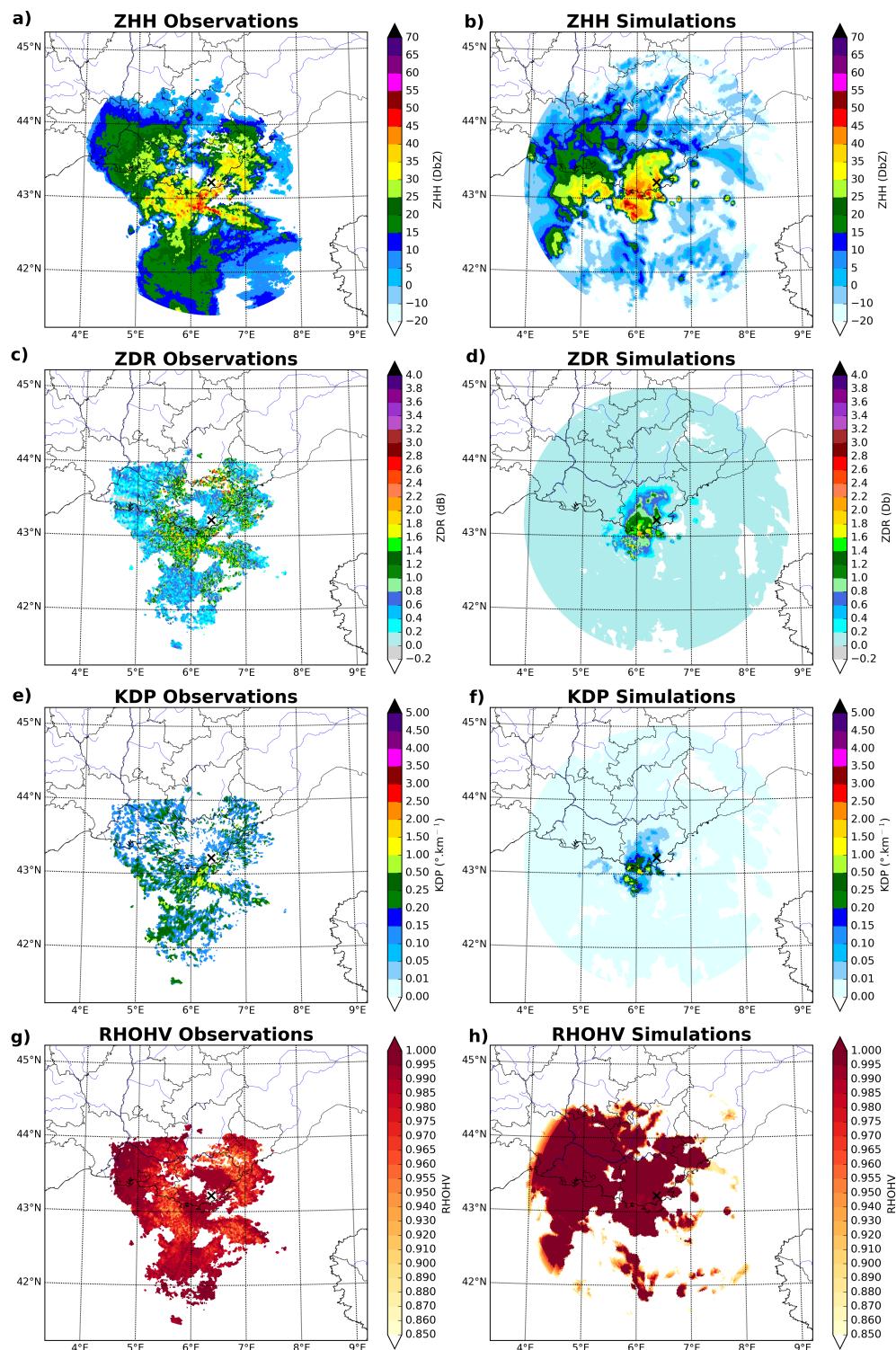


Figure 1. Observed (left) and simulated (right) Z_{HH} (a and b), Z_{DR} (c and d), K_{DP} (e and f) and ρ_{HV} (g and h) for the Collobrières radar the 2018-10-10 at 14 UTC (elevation: 2.2°).



170 Fig. 1 represents the observed and the simulated DPOL variables during this event, using the S-Band Collobrières radar
 located along the French Mediterranean coast (indicated by a dark cross), for an elevation angle of 2.2° . The simulated hy-
 droeteor mixing ratios from the 1h AROME-France forecast have been interpolated on the radar beam for the same elevation
 angle (Fig. 2). All observations have been filtered with the methodology described in section 3.1. This particular radar is part
 of the French radar network ARAMIS (Tabary, 2007), which, in 2019, is composed of 31 weather radars, 28 having a DPOL
 175 capacity.

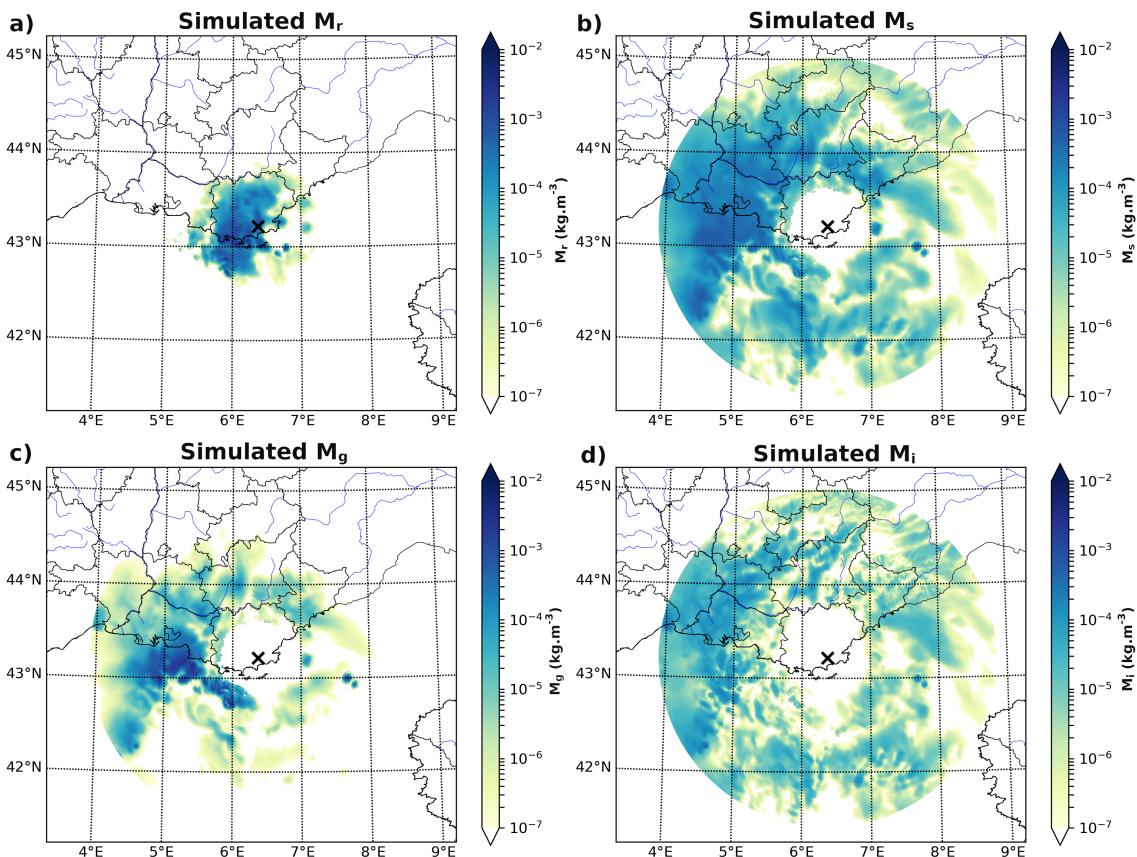


Figure 2. Hydrometeor mixing ratios from a 1h AROME-France forecast valid the 2018-10-10 at 14 UTC for (a) rain, (b) snow, (c) graupel and (d) pristine ice.

180 The Z_{HH} observations show a narrow band of very high values over the sea near the radar, reaching locally more than 50 dBZ. An area of medium to high values (20 to 40 dBZ) is located inland in the north-west quadrant, while an extended stratiform area of Z_{HH} values above 15 dBZ is located more offshore. When examining the simulations, high values of Z_{HH} with comparable values are also present close to the radar, however covering a wider area than in the observations. The inland area of Z_{HH} is well represented, but with lower values than observed. Finally, for the southern part of the precipitating system far from the radar, Z_{HH} is clearly underestimated. Since solid hydrometeors are present in this area, as displayed in Fig. 2,



either AROME-France did not simulated sufficient amount of it, or such underestimation comes from the observation operator. A combination of these two hypotheses could also explain these differences.

185 In agreement with areas of large Z_{HH} values, observed Z_{DR} can reach locally between 2.0 to 2.8 dB (Fig. 1c). Elsewhere, values are predominantly lower than 1 dB, with many small spots of higher values. Considering the simulations (Fig. 1d), the range of values are close to the observations in the area where simulated rain prevails. Nevertheless, as for simulated Z_{HH} and contrarily to what is observed, the range of values decreases with distance, which is typical of an evolution from convective liquid precipitation to more spherical solid hydrometeors. As solid hydrometeors are simulated in those areas (see again Fig. 2), comparison to observations clearly reflects the inability of the ICE3 microphysical parameterisation to represent the observed 190 variability of hydrometeors, particularly in the ice phase.

195 The observed and simulated specific differential phase K_{DP} are compared in Figs. 1e,f. In both cases, values up to $1.5^\circ \cdot \text{km}^{-1}$ are locally displayed in the main convective area close to the radar, which is characterized by intense rainfall. As for Z_{DR} , the area of largest simulated Z_{HH} values is associated with significant K_{DP} values. Elsewhere however, K_{DP} values are close to zero. In general, these high amounts of null or close to zero values for K_{DP} or Z_{DR} are associated with locations where the simulated Z_{HH} is lower than 20 dBZ, corresponding to small amounts of hydrometeor contents simulated by AROME-France.

200 Here, only co-polar correlation coefficient values higher than 0.85 are displayed, lower ones being associated with non-meteorological echoes. Concerning the observations (Fig. 1g), the ρ_{HV} values are very close to 1 in the area of large Z_{HH} values, mostly composed of liquid hydrometeors. Then, a ring of values between 0.9 and 0.98 is displayed, denoting the solid hydrometeors melting within the so-called bright band. Far from the radar, ρ_{HV} values increase up to 1, indicating more 205 homogeneous solid hydrometeor distributions. In the simulation (Fig. 1h), most of the areas where Z_{HH} is above 20 dBZ are associated with a ρ_{HV} close to 1, corresponding to very homogeneous scenes. Furthermore and contrarily to what has been observed, the melting layer is not visible in the simulation. Far from the radar and in regions where the hydrometeor contents are low, the simulated ρ_{HV} decreases significantly, reaching values less than 0.1 (not shown). It corresponds to areas where snow and ice contents are similar (see respectively Fig. 2b and Fig. 2d). Due to this specific condition, each hydrometeor 210 type influences equally the computation of ρ_{HV} . Nevertheless, due to the large differences between the characteristics of each hydrometeor type (listed in Section 2.1.1), a large non-homogeneity is induced and leads to low ρ_{HV} values. Nevertheless, as such values are usually associated with non-meteorological echoes, these non-realistic simulated values are discarded. These 215 results show the strong limitation of \mathcal{H}_{DPOL} for simulating realistic values of ρ_{HV} .

220 Considering all case studies (Table 1), it was generally found that realistic simulations of DPOL variables can be obtained especially in regions where liquid precipitation occur. In presence of solid hydrometeors, the simulation of Z_{DR} , K_{DP} and ρ_{HV} increases in complexity and comparisons with observation show large differences. It also appears that simulated Z_{HH} can be underestimated when only solid hydrometeors are present. The same patterns have been obtained for both S-band and C-band radars simulations, confirming the results of Augros et al. (2015). These misrepresentations are probably partly due to hypotheses done in \mathcal{H}_{DPOL} on the hydrometeor shapes and aspect ratios, on their PSDs and on their dielectric constants. 225 Indeed, these specified parameters might not be adequate for all meteorological situations. This is especially true for solid hy-



drometeors, for which the simulation of DPOL variables is particularly complex and dependent on hydrometeor characteristics simulated by the ICE3 microphysical scheme that does not describe them with enough details.

2.3 Assessment of model errors

In order to describe the uncertainties associated with the simulation of DPOL variables, the impact of changes to three 220 \mathcal{H}_{DPOL} main input variables has been examined. These variables are the dielectric constant, the hydrometeor aspect ratio and the hydrometeor oscillation with respect to the horizontal plane. For each type of hydrometeors, these parameters have been tested independently by applying a perturbation to the default value. For each parameter change, the look-up tables have been recomputed and a new simulation performed. Six different meteorological cases sampled by S-band radars of the French 225 ARAMIS network (Tabary, 2007) (Table 1) have been chosen and, coupled with a set of 23 different configurations, it leads to a total of 138 different simulations. Standard deviations have been computed for each combination of input parameters listed in Table 2 and for each radar elevation. It can be noticed that the oscillation parameter, for which physically realistic values described in Ryzhkov et al. (2011) have been used, has not been perturbed for primary ice which is represented by spheres in ICE3. For the dielectric constant and the hydrometeor aspect ratios, positive and negative relative perturbations have been considered.

Table 1. Simulated meteorological cases used to assess simulation uncertainties. The two radars of interest are operating in S-band. Hours are expressed in UTC.

DATE	RADAR
19 October 2017 06:00	NIMES
05 February 2018 19:00	NIMES
29 May 2018 16:00	NIMES
09 August 2018 06:00	NIMES
07 October 2018 03:00	COLLOBRIERES
10 October 2018 14:00	COLLOBRIERES

230 The results are displayed in Fig. 3 for each DPOL variable. The first noticeable information about Z_{HH} uncertainties are the two quasi-linear tendencies observed near 0 and 1 dBZ. The first one, around 0.1 dBZ, is induced by the perturbation of the rain aspect ratios (not shown). It was found that this behaviour comes from thresholds present in the computation of the backscattering coefficients. The second quasi-linear signal, around 0.9 dBZ, appears to be mostly dependent upon the perturbation of graupel dielectric constant, without being constrained by a threshold. Overall, the uncertainty on Z_{HH} 235 appears to be mostly dependent upon the representation of the three types of solid hydrometeors. Indeed, each uncertainty value higher than 0.2 dBZ in Fig. 3a is associated to a perturbation of a parameter used to represent solid hydrometeors, especially the dielectric constant. For Z_{DR} , a maximum spread around 0.6 dB is displayed. It also comes from thresholds in the backscattering coefficient computations for rain aspect ratio. Overall, there is more variability for cases where the simulated Z_{DR} are below 0.5 dB, which expresses a higher sensitivity of \mathcal{H}_{DPOL} to parameters describing frozen hydrometeors or small



Table 2. Parameters modified to study uncertainties in the \mathcal{H}_{DPOL} operator. * LD stands for Linear Decrease and expresses the linear aspect ratio decrease with the hydrometeor diameter increase, between the two threshold diameter values; MG stands for Maxwell-Garnett.

	Default Configuration			Perturbed Configurations		
	Dielectric constant	Aspect Ratio	Oscillation	Dielectric constant	Aspect Ratio	Oscillation
Rain	Debye	Brandes	0°	Debye ± 5%	Brandes ± 20%	5°; 10°
Snow	MG*	If $d < 0.5$ mm : 1.0 ; if $0.5 \geq d < 8.0$: LD* ; else : 0.75	0°	MG* ± 5%	Snow default ± 20%	10°; 20°
Ice	MG*	1.0	0°	MG* ± 5%	Ice default ± 20%	—
Graupel	MG*	If $d < 0.5$ mm : 1.0 ; if $0.5 \geq d < 10.0$: LD* ; else : 0.85	0°	MG* ± 5%	Graupel default ± 20%	10°; 20°

240 raindrops. The raindrop aspect ratio parameter explains the major part of the uncertainties appearing on Z_{DR} simulation. Nevertheless, a small part of these uncertainties can also be explained by the dielectric constant of solid hydrometeors. Very small uncertainties have been found for ρ_{HV} simulations. Indeed, no matter the perturbed parameter, the highest uncertainty is lower than 1.10^{-3} . Concerning K_{DP} , a quasi-linear threshold of sensitivity is displayed. As for the spread distribution of the other DPOL variables (excepted ρ_{HV}), it comes from the rain aspect ratio. This parameter also explains the major part of the 245 variability of K_{DP} uncertainties.

The results of these sensitivity tests show that Z_{HH} is sensitive to assumptions made on the simulation of the back-scattering coefficients for the different hydrometeor types, especially for the solid ones. Indeed, it appears that Z_{HH} uncertainties are, for their major part, explained by the hypotheses done on the dielectric constant of solid hydrometeors. Then, this could explain the underestimations found for simulated Z_{HH} in presence of solid hydrometeors (Fig. 1). On the contrary, the other DPOL 250 variables appear to be less sensitive to choices made for solid hydrometeors than for liquid ones. Even if DPOL variables are highly influenced by hydrometeor dielectric constants, they are also known to be dependent upon hydrometeor shapes. For Z_{DR} and K_{DP} , uncertainties have been found in the simulations for raindrop aspect ratio perturbations while, for solid hydrometeors, no uncertainties or very small ones have been found by perturbing their aspect ratios. These results can be explained by the one moment microphysical scheme ICE3 that characterizes hydrometeors and related processes. Indeed, the 255 PSDs are deduced from the hydrometeor contents and from constant parameters in order to characterise the generalized Gamma distributions (see again Table 1 in Caumont et al. (2006)). Concerning the hydrometeor shapes, oblate spheroids appear to be a good approximation of raindrop shapes while it might be very limiting for solid hydrometeors that exhibit a large diversity of shapes. For example, Liu (2008) proposed the use of 11 solid particle shapes along with the Discrete Dipole Approximation (DDA) method in order to compute more realistic scattering. Clearly, the T-matrix method is comparatively limited, as it is 260 only applicable on spheres and rotationally symmetrical particles.

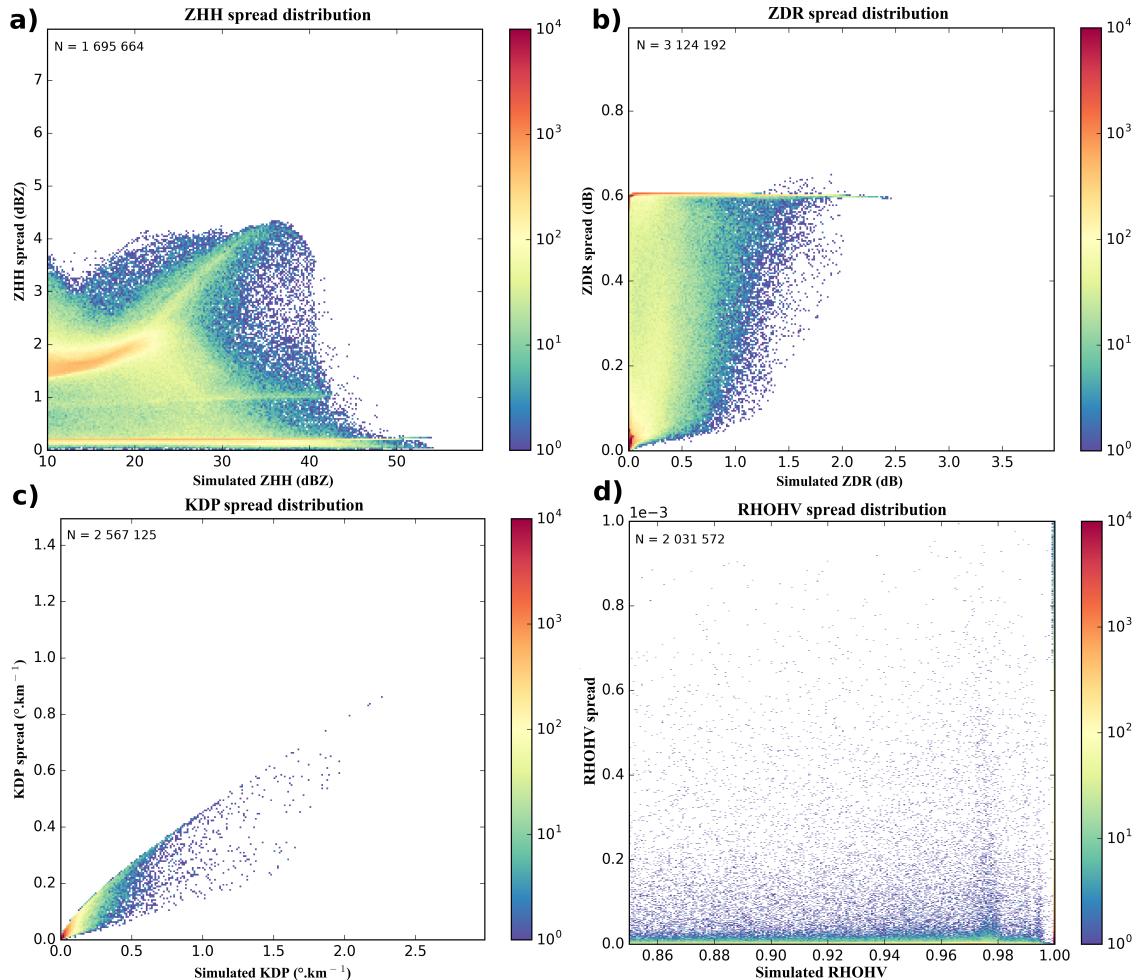


Figure 3. Simulation uncertainty for a) Z_{HH} , b) Z_{DR} , c) K_{DP} and d) ρ_{HV} . N represents the sample size.

A similar study to the one presented here with S-band radars has been conducted with 11 different meteorological cases, but with C-band radars (not shown). Comparable results have been obtained, the spread being nevertheless slightly larger for Z_{HH} values higher than 30 dBZ, for Z_{DR} values higher than 1 dB and for the total range of K_{DP} values. These results highlight the dependence of simulated DPOL variables upon the wavelength. As suggested by the range of values affected by a larger spread, Mie diffusion occurs for large hydrometeors with C-band radars.

Nevertheless, despite choices done in \mathcal{H}_{DPOL} and, as discussed in section 2.2, the polarimetric observation operator is able to simulate DPOL variables in the presence of liquid and solid hydrometeors, as shown for instance in Fig. 1d for Z_{DR} . As discussed at the end of the next section, model errors that have been quantified in this sensitivity study will be considered for specifying a proxy of the observation error standard deviations, in addition to measurement and representativeness errors.



Table 3. Meteorological cases selected to study the innovation statistics of DPOL variables. Hours are expressed in UTC.

DATE	RADAR
17 January 2018 03:00	TRAPPES
14 February 2018 13:00	TOULOUSE
14 February 2018 13:00	GREZES
26 May 2018 12:00	MOMUY
06 June 2018 00:00	TRAPPES
29 July 2018 07:00	PLABENNEC
10 October 2018 14:00	COLLOBRIERES
24 April 2019 14:00	NANCY
25 April 2019 13:00	TOULOUSE
08 May 2019 04:00	GREZES
10 May 2019 16:00	TOULOUSE
10 May 2019 19:00	TRAPPES

270 Such values could then be used as diagonal elements of an observation error covariance matrix **R** necessary for variational DA studies.

3 Statistics of innovations

As explained previously, the optimality of variational DA requires Gaussianity of errors. For that purpose, innovation statistics (differences between observation and model counterparts) are examined. An ad-hoc quality control could then be 275 defined in order to improve Gausiannity. In this study, such statistics have been computed for 12 contrasted meteorological cases, encompassing convective and stratiform precipitation. Among those cases, only the Collobrières case has been observed by a S-band radar, while the others have been sampled by C-band radars (see Table 3). The geographical radar location is given in Fig. 1 from Tabary (2007).

3.1 Data pre-processing

280 Several filters are applied on the observations, principally to remove non meteorological echoes and regions of too low signal-to-noise ratio (SNR). Non-meteorological echoes are filtered using an echo type determination algorithm developed by Gourley et al. (2007). A second filter removes possible remaining ones by excluding pixels for which ρ_{HV} values are lower than 0.85. The third filter uses a threshold on SNR values. Tabary et al. (2013) explain that polarimetric variables are very sensitive to noise and, for safety reasons, all pixels with a SNR value lower than 15 dB are discarded. Finally, a median filter is 285 applied to remove all residual isolated noisy data.

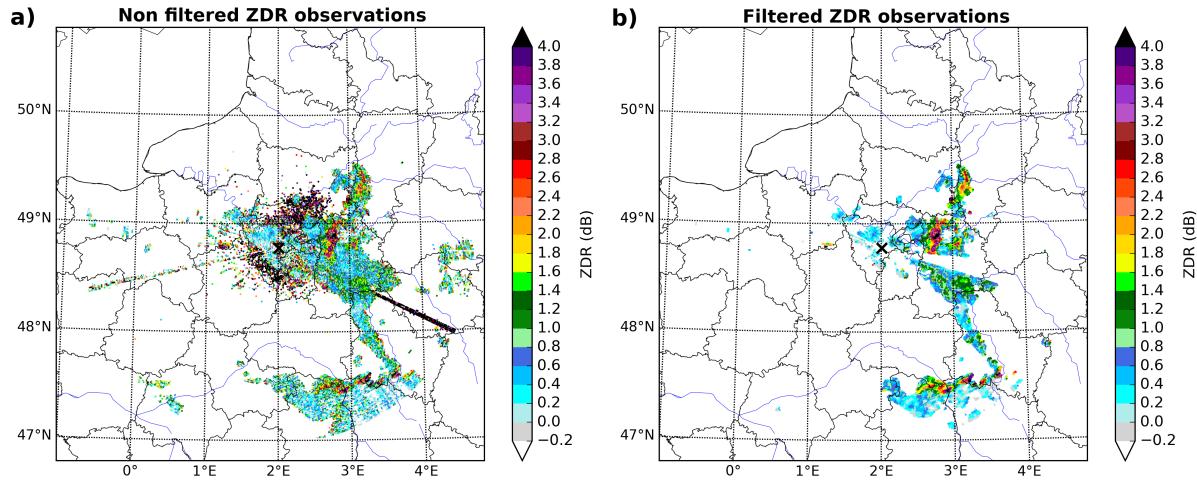


Figure 4. Filtering effects on Z_{DR} observed by the Trappes radar during a convective event on 2018-06-0600 UTC, for an elevation of 0.4° .

Fig. 4 shows the effects of these filters on Z_{DR} values observed during a convective event that occurred in the Paris area. In the left panel, ground clutters, characterized by high Z_{DR} values, are present close to the radar. Medium to very strong values can also be observed along two different azimuths. These patterns are often due to WLAN (Wireless Local Area Network) interferences. In the right panel, those patterns are not present anymore, thanks to the application of the various filters. One can notice that other features, located far from the radar, have also been removed. They correspond to data with SNR values lower than 15 dB.

3.2 Results

In order to quantify the effect of the filters, innovations have been computed on non-filtered and filtered observations. Fig. 5a shows that filtering the Z_{HH} observations lead to a small decrease of the bias and standard deviation values, from 3.22 to 3.18 dBZ and from 11.44 to 11.15 dBZ respectively. Concerning Z_{DR} (Fig. 5b), the filtering leads to strong changes in the innovation statistics. Indeed, the bias decreases from 0.37 to 0.33 dB while the standard deviation drops from 1.26 to 0.55 dB. For K_{DP} (Fig. 5c), filters do not influence innovation statistics. The use of filters mostly affects the negative part of ρ_{HV} innovations (Fig. 5d), by removing simulated values close to one. Overall, these results show small modifications of innovations statistics, except for Z_{DR} . For the latter, quality controls appear to be critical and allow to discard about 43% of spurious Z_{DR} observations. The innovation distributions appear to have a Gaussian shape for Z_{HH} and Z_{DR} while it is not the case for K_{DP} and ρ_{HV} .

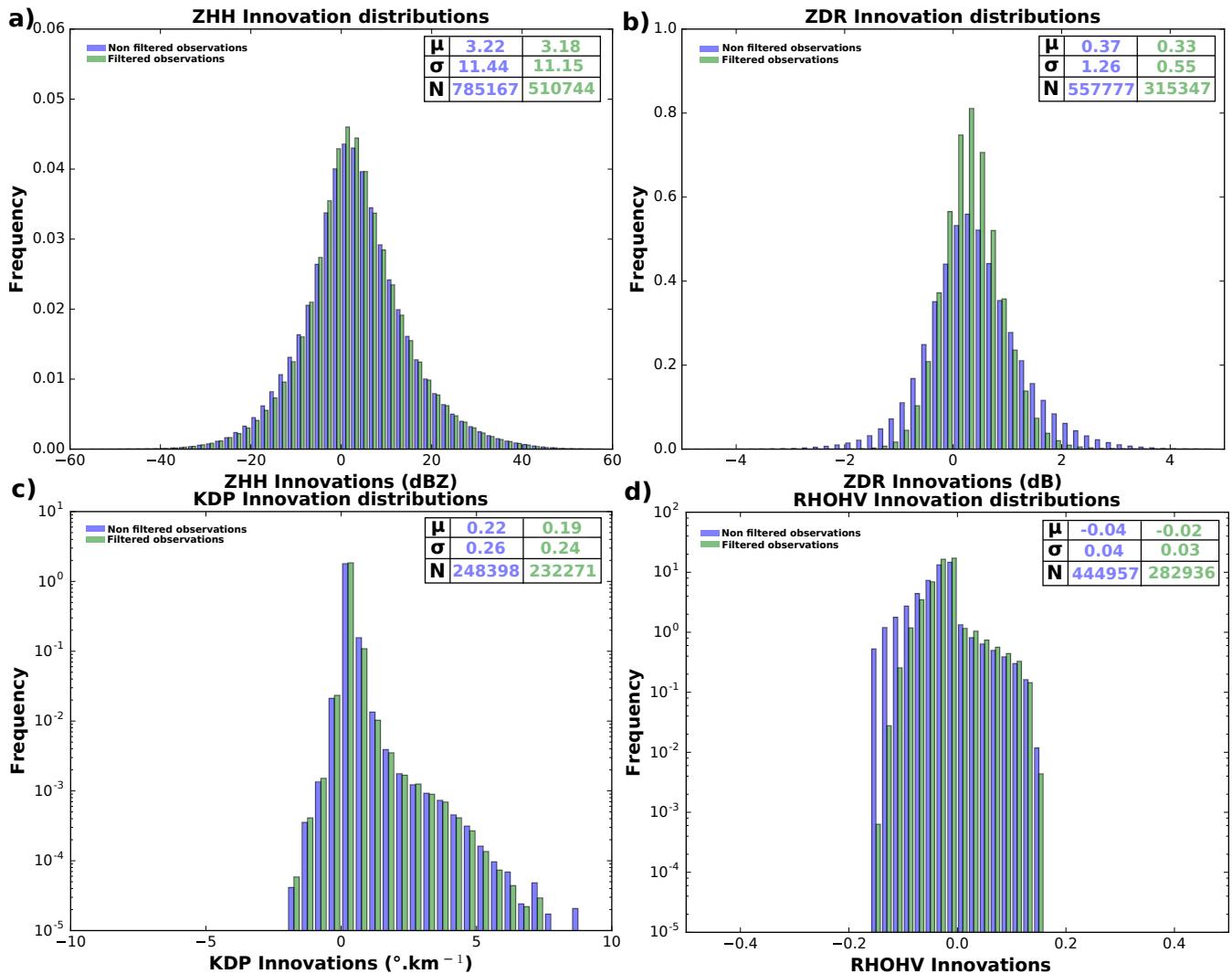


Figure 5. Innovations without (blue histograms) and with (green histograms) filters on observations for: a) Z_{HH} , b) Z_{DR} , c) K_{DP} and d) ρ_{HV} . μ represents the innovation bias, σ the innovation standard deviation and N the sample size

In order to study if these conclusions depend on the hydrometeor phase, innovation statistics have been computed for different vertical levels. Fig. 6 represents such distributions over altitude for the studied DPOL variables, when filters are applied. The Z_{HH} innovation distribution (Fig. 6a) exhibits a positive bias which tends to slightly increase with altitude. It expresses underestimations done in presence of solid phase hydrometeors that have been already noticed, especially for pristine ice which is present at high altitudes. One can notice a small asymmetry present in the innovation distribution below 4 km. Indeed, in this part of the atmosphere, a larger number of innovation values are represented in the distribution between 40 to 60 dBZ than in the symmetrical negative part. Depending upon the meteorological situation, this range of altitudes correspond to the melting



layer. Large positive innovations in this particular area can highlight melting layer misrepresentations in \mathcal{H}_{DPOL} which lead
310 to Z_{HH} underestimations. Same phenomenon is present for other DPOL variables for this range of altitudes, especially for the
differential reflectivity. About Z_{DR} (Fig. 6b), a positive bias indicates an underestimation in the simulations. Nevertheless, for
altitudes higher than 10 km, the bias drops to nearly 0 dBZ values. Concerning K_{DP} (Fig. 6c), innovations show a small bias
315 which slightly increases with altitude. Below 7 km, the innovation spread shows underestimations and overestimations done
in simulations while above 7 km, innovations are always positive, indicating systematic K_{DP} underestimations with \mathcal{H}_{DPOL}
at those levels, in presence of solid hydrometeors. Contrarily, ρ_{HV} innovation distributions (Fig. 6d) show a negative bias,
indicating overestimations in the simulations.

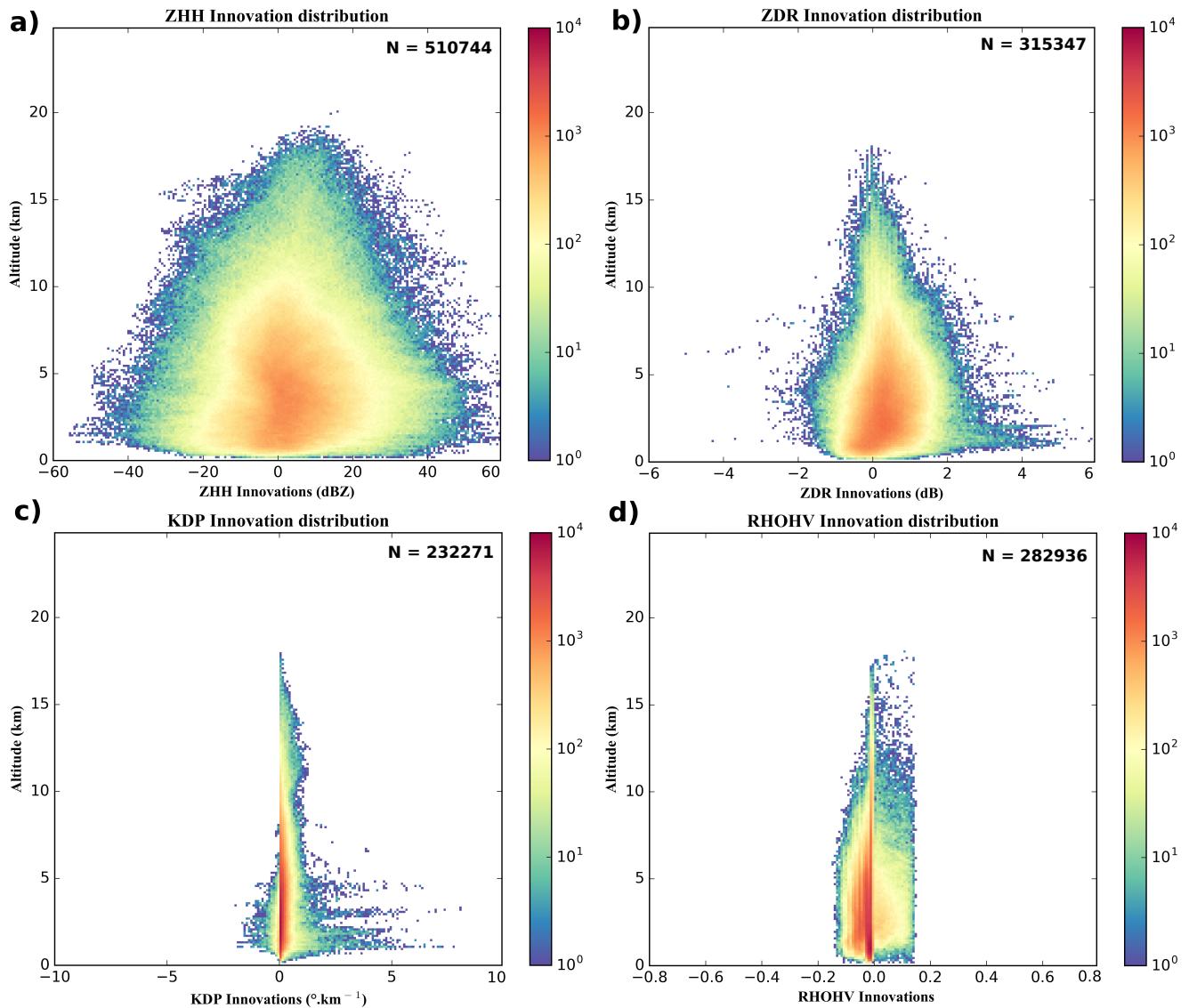


Figure 6. Innovation distributions over altitude for a) horizontal reflectivity (Z_{HH}), b) differential reflectivity (Z_{DR}), c) specific differential phase (K_{DP}) and d) co-polar coefficient (ρ_{HV}). N indicates the sample size for each DPOL variables.

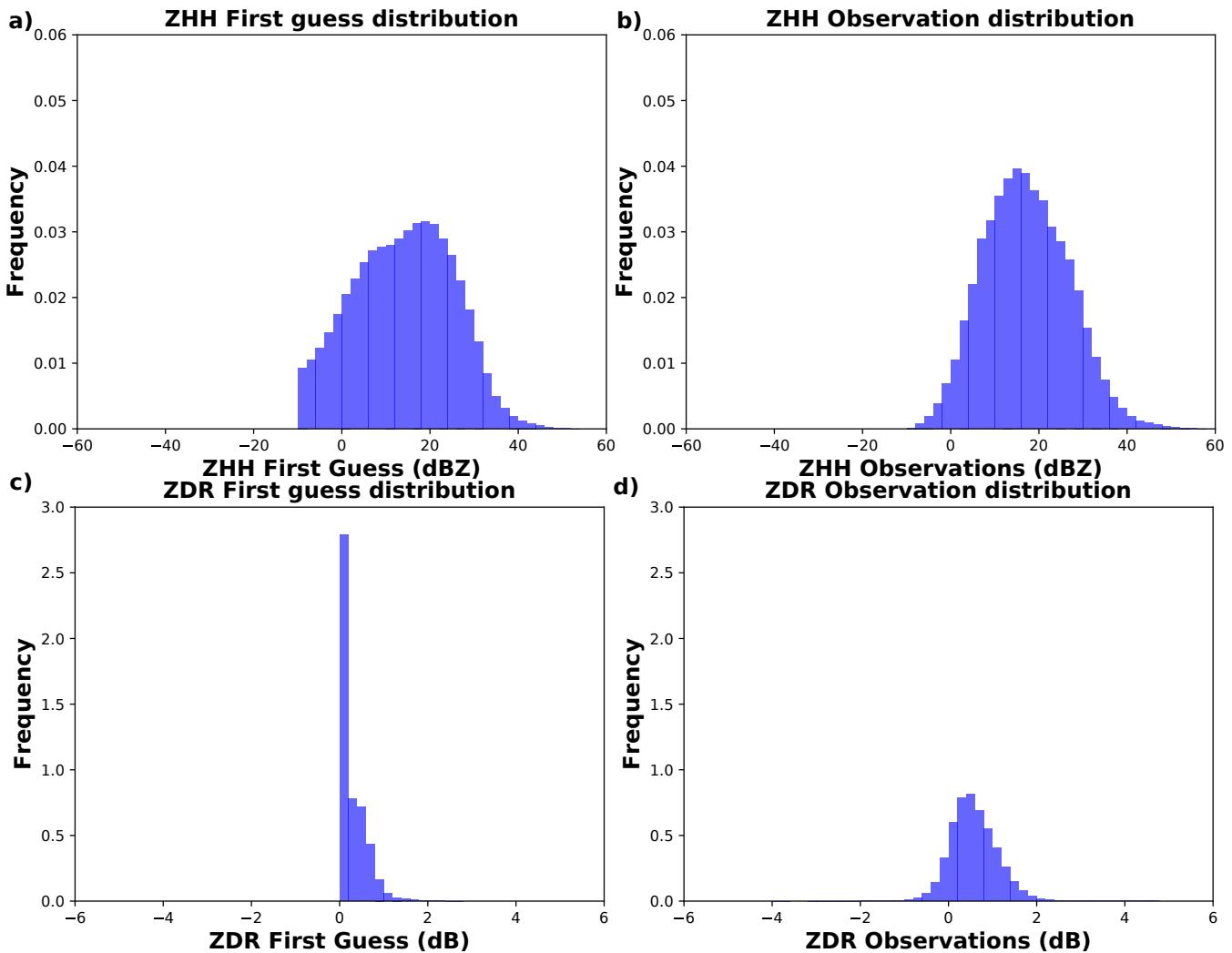


Figure 7. First guess (left panels) and observation distributions (right panels) for Z_{HH} (upper panels) and Z_{DR} (bottom panels).

To better understand the behaviour of innovations, separated distributions of simulated and observed values are examined. Fig. 7 represents the distributions of first-guess and observation distributions for Z_{HH} and Z_{DR} . Concerning Z_{HH} , first guess and observation distributions (Figs. 7a and b respectively) look very similar for values above 20 dBZ, which shows the \mathcal{H}_{DPOL} capacity to simulate such variable in presence of medium to heavy precipitation. Between 10 dBZ and 20 dBZ, the number of observations is larger than the simulated ones, which reflects the \mathcal{H}_{DPOL} underestimation of Z_{HH} in the solid phase already pointed out. Between 0 and -10 dBZ, the number of simulated values is higher than the observed ones, denoting \mathcal{H}_{DPOL} capacity to simulate small values in the presence of very low hydrometeor contents. Simulated values below -10 dBZ, which are generally close to the radar SNR, cannot be considered for the assimilation and thus have been discarded.



325 Concerning Z_{DR} , first guess and observation distributions (Fig. 7d and e respectively) appear to be quite different. For positive Z_{DR} values, both distributions are similar, especially for small values. Indeed, the higher the Z_{DR} value is, the larger is the difference between first guess and observation distributions. It comes from the complexity of Z_{DR} simulations, especially in the presence of solid hydrometeors, where large underestimations occur. In addition, a large fraction near 0 dB in the simulations is not represented in the observations. Finally, the negative Z_{DR} values in the observations have not been simulated by
 330 \mathcal{H}_{DPOL} . They correspond to hydrometeors with larger vertical axes than horizontal ones. As such hydrometeor shape is not represented in \mathcal{H}_{DPOL} , simulated Z_{DR} cannot reach negative values. Physically, such values are usually associated to particular situations in convective events which can cause preferential vertical orientation of solid hydrometeors, as electrification processes (Kumjian, 2013a).

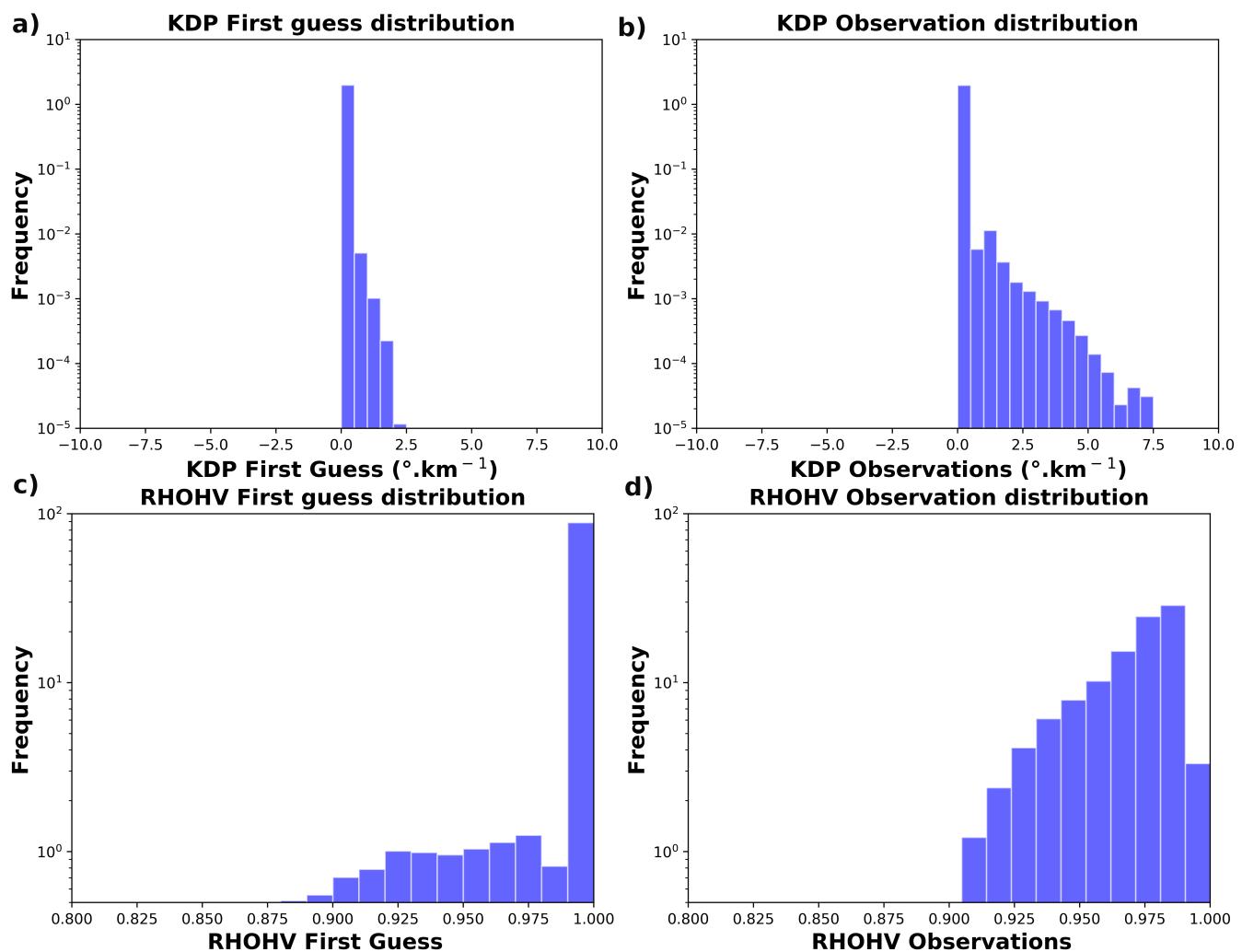


Figure 8. Same as Fig. 7 for K_{DP} (upper panels) and for ρ_{HV} (bottom panels)



Fig. 8 is similar to Fig. 7, but for K_{DP} and ρ_{HV} and only in the presence of liquid hydrometeors. K_{DP} first guess distribution 335 (8a), in comparison to the observations distribution (8b) emphasizes the large underestimations of the simulations. Indeed, the largest simulated K_{DP} values is about $2.0^\circ.km^{-1}$ while the maximal observed one reaches $7.5^\circ.km^{-1}$. This leads to an innovation distribution which is far from Gaussian, with a strong positive bias (see Fig. 5c). To be able to assimilate this variable, a strict data selection should be done. Regarding ρ_{HV} , most of simulated values are very close to 1.0 (Fig. 8d) while the observed values range between 0.90 and 1.0. These results reinforce the lack of variability of simulated ρ_{HV} values found 340 in Section 2. It leads to an innovation distribution which is far from Gaussian (see Fig. 5d).

These innovation statistics can also be used to define an approximation of observation standard deviations. Indeed, Errico et al. (2000) explain that, if innovation PDFs are Normally distributed:

$$\sigma_d^2 = \sigma_o^2 + \sigma_b^2 \quad (5)$$

σ_d^2 being the variance of the innovation PDFs, σ_o^2 and σ_b^2 being respectively the observation and the background error variances.

345 In order to obtain a very first approximation of the observation standard deviation σ_o , it can be assumed that σ_o and σ_b are equivalent. In such conditions:

$$\sigma_o = \frac{\sigma_d}{\sqrt{2}} \quad (6)$$

Values of $\sigma_o(Z_{HH}) = 7.88$ dBZ, $\sigma_o(Z_{DR}) = 0.39$ dB, $\sigma_o(K_{DP}) = 0.17^\circ.km^{-1}$ and $\sigma_o(\rho_{HV}) = 0.02$ have been found. Nevertheless, such values must be refined, especially for K_{DP} and ρ_{HV} for which Gaussianity have not been found in the 350 innovation PDFs.

4 Polarimetric variable Jacobians

4.1 Perturbation size determination

As explained in the introduction, the adjoint of the linearized observation operator is required in the formulation of the gradient of the cost function. \mathcal{H}_{DPOL} being an observation operator which deals with cloud microphysical processes, numerous 355 highly non-linear processes are present. In this study, the linearized version of \mathcal{H}_{DPOL} , noted \mathbf{H} , has been estimated by his Jacobians, computed through the use of the finite difference method:

$$\mathbf{H}(x) = \frac{\partial \mathcal{H}}{\partial M} \approx \frac{\mathcal{H}(M + \delta M) - \mathcal{H}(M)}{\delta M} \quad (7)$$

\mathcal{H} representing the non linear version of \mathcal{H}_{DPOL} and δM a perturbation of the hydrometeor content M .



Then, the Jacobian matrix can be estimated as follow:

$$360 \quad \begin{pmatrix} \frac{\partial \mathcal{H}_1}{\partial M_{i,1}} & \cdots & \frac{\partial \mathcal{H}_l}{\partial M_{i,1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathcal{H}_1}{\partial M_{i,k}} & \cdots & \frac{\partial \mathcal{H}_l}{\partial M_{i,k}} \end{pmatrix} \quad (8)$$

k representing the model level number and l the radar elevation and $M_{i,k}$ is hydrometeor content (in $kg.m^{-3}$) associated for type i .

First of all, it is important to evaluate the validity of the linear regime, according to the size of the perturbation δM . Duerinckx et al. (2015) proposed a method for examining the difference of computations between equivalent positive and negative perturbations. They explain that, as long the problem stays in a linear regime, this difference should remain close to zero. In this study, hydrometeor contents can span a wide range of values (several orders of magnitude). As a consequence, the perturbations are chosen as a fraction of the hydrometeor content (instead of a fixed value). The optimal value of perturbation for the Jacobians has been estimated in the model space. As \mathcal{H}_{DPOL} computes the DPOL variables independently for each pixel of the domain before being interpolated on the radar beam, only the diagonal elements of the Jacobian matrix are examined, since pixels, in model space, are uncorrelated both horizontally and vertically. These computations have been done for several profiles but, for illustration, only a single profile associated with convective precipitation is presented. Fig. 9 shows how the optimal rain content fraction δM_r has been chosen to compute the Jacobians $\delta Z_{HH}/\delta M_r$. One can notice that the optimal rain content fraction lies between $10^{-2} g.m^{-3}$ and $10^{-6} g.m^{-3}$.

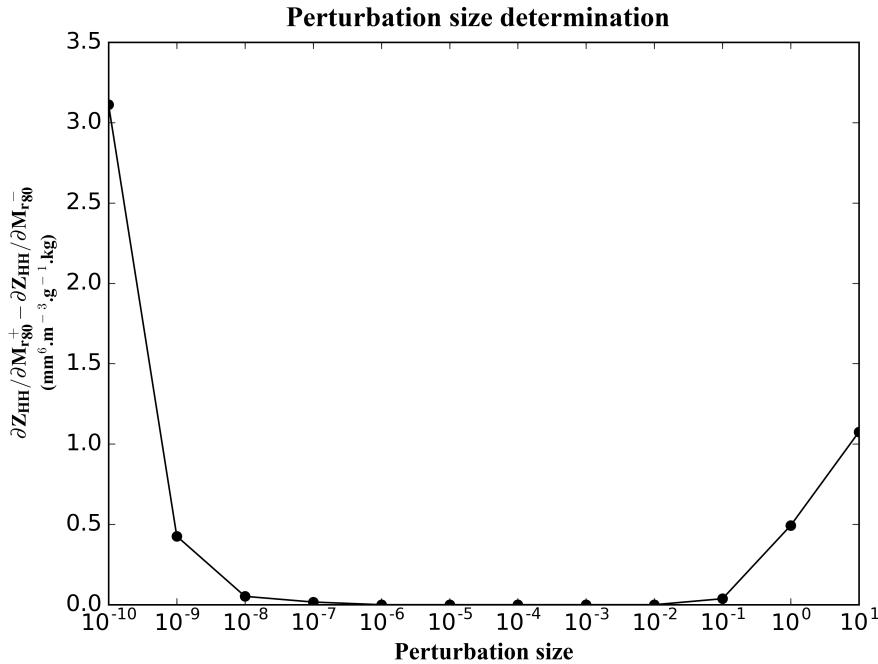


Figure 9. Difference between $\partial Z_{HH} / \partial M_r$ at model level 80 (about 1 km of altitude), estimated with positive and negative perturbations as a function of the perturbation size δM_r

For each hydrometeor type, a single fraction of hydrometeor content has been determined for all DPOL variables, by selecting the highest optimal fraction size in common between the four DPOL variables. It was found that the optimal fraction size is 10^{-5} g.m^{-3} for rain and primary ice contents, while 10^{-4} g.m^{-3} is more suitable for snow and graupel contents.

4.2 Jacobian profiles in the model space

The information provided by the DPOL variables depends upon the interaction of the different hydrometeors scanned by the radar beam. A primary step towards understanding DPOL variable Jacobians is to exclude the radar beam effect. In that case, it is proposed to first consider the diagonal elements of the Jacobian matrix computed in model space. The perturbation used at each level in the Jacobian computation is applied as follows:

$$M_k^* = M_k + M_k \delta M \quad (9)$$

M_k^* representing the perturbed hydrometeor content at level k , M_k the hydrometeor contents at level k and δM the optimal fraction of hydrometeor contents previously chosen. The product $M_k \delta M$ represents the perturbation applied at level k . Once the Jacobians are computed, a normalisation by 10% of the hydrometeor profile is applied. This procedure allows a comparison between Jacobians of a given DPOL variable for different hydrometeor types.

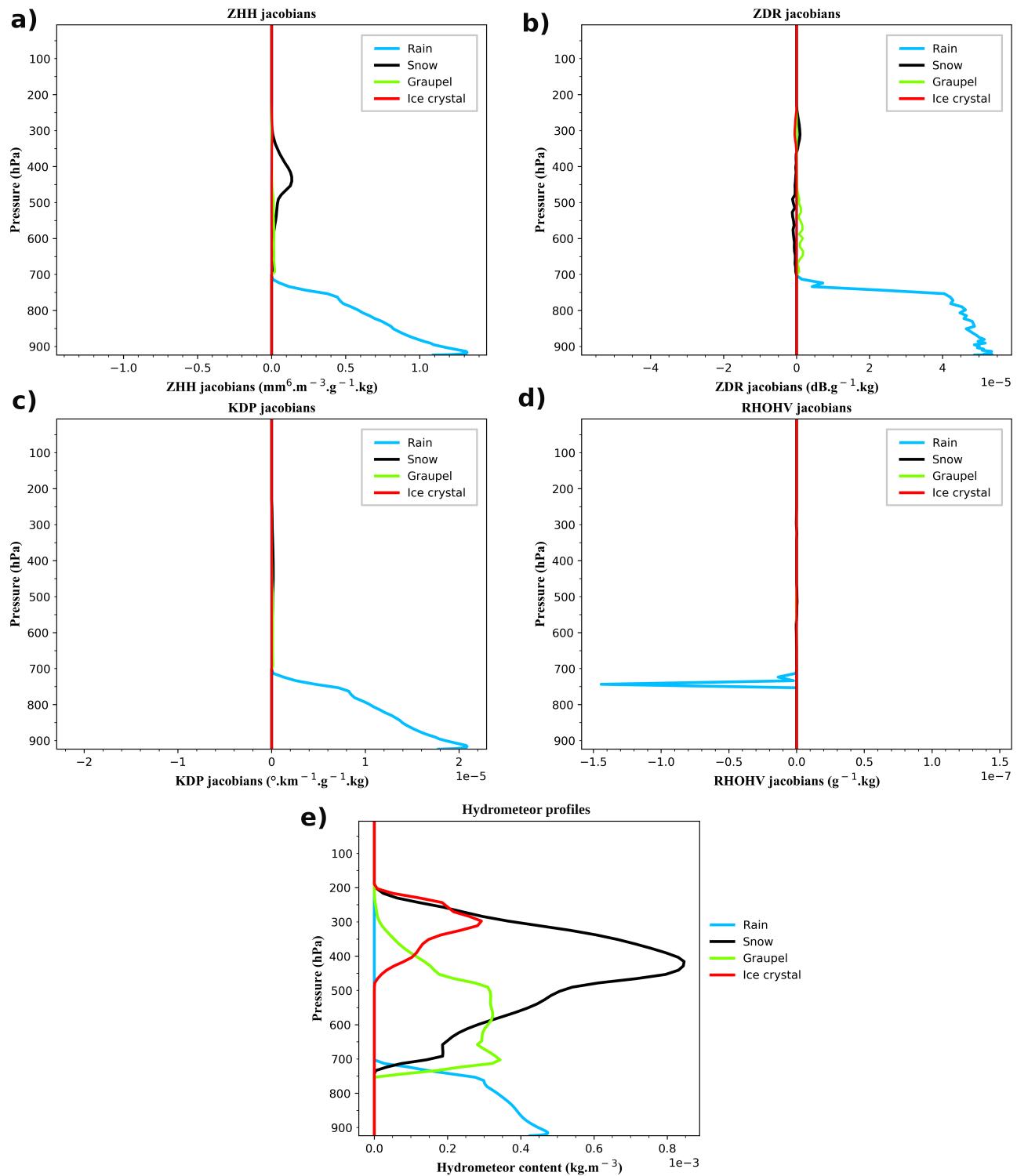


Figure 10. a) Z_{HH} Jacobians, b) Z_{DR} Jacobians, c) K_{DP} Jacobians and d) ρ_{HV} Jacobians, e) represent hydrometeors content profiles, associated with a convective event which stroke the Hérault and Gard departments on the 2017-10-19-0600 TU. The Jacobians presented here are normalised by 10 percent of the hydrometeor contents.



Fig. 10 presents the DPOL variables Jacobians computed from a particular profile extracted within a convective cell forecasted by AROME-France (Fig. 10e). It shows that an increase in hydrometeor content leads to higher Z_{HH} Jacobian values² (Fig. 10a). It is explained by the fact that reflectivity is proportional to the total hydrometeor cross-section (see Eq. 1). Nevertheless, due to different dielectric constants, Jacobian values are different according to the hydrometeor type. Indeed, Z_{HH} appears to be more sensitive to rain content perturbations while its sensitivity to snow content perturbations is about one order of magnitude lower. Z_{HH} sensitivities to graupel and pristine ice perturbation is even lower.

Contrary to Z_{HH} which mostly depends on the total cross-section and the dielectric factor, other DPOL variables are also strongly dependent upon hydrometeor characteristics, such as their shape or their orientation. Z_{DR} for instance, as previously explained, depends upon hydrometeor dielectric constant and shape, as well as the proportion of each type of hydrometeors with respect to the total hydrometeor content. Fig. 10b shows Z_{DR} Jacobians for different hydrometeor content perturbations. Raindrops being simulated as oblate spheroids, their larger horizontal cross-section compared to the vertical one leads to positive Z_{DR} Jacobians for a rain content perturbation (Fig. 10b). For other hydrometeor content perturbations, Jacobian values can be negative. Such values can be observed for ice content perturbation around 300 hPa and for snow content perturbation between 400 hPa and 700 hPa. For pristine ice, the negative Jacobian values are due to the increase of spherical particles in the presence of snow (see Fig. 10e). It causes a small decrease of the proportion of non-spherical particles in the total hydrometeor content and then, a decrease of Z_{DR} . On the contrary, an increase of snow content in the same part of the atmosphere causes positive Jacobian values due to the increase of non spherical particles in the total hydrometeor content. From 300 hPa to approximately 600 hPa, the graupel content is increasing while snow amounts are increasing between 300 hPa and 400h Pa and then, decreasing (Fig. 10e). Between these pressure levels, the Z_{DR} Jacobian associated with a snow content perturbation is decreasing and reaches negative values. The graupel dielectric constant being higher than for snow, it becomes predominant in the Jacobian values. So, even if there is an increase in snow content, which is characterized by flatter particles than graupel, their presence leads to a reduction of Z_{DR} when it is not the prevailing hydrometeor type. One can notice that the negative values of $\partial Z_{DR} / \partial M_s$ between 300 hPa and 600 hPa show vertical oscillations which are likely due to changes in proportions of the different hydrometeor types. Below 600 hPa, the snow content becomes lower than the graupel one. Since an increase in snow adds flatter particles, it also leads to an increase in Z_{DR} Jacobian values in this part of the atmosphere. Approximatively below 700 hPa, rain content is increasing and, because of the large predominance of liquid water dielectric constant over the one from other hydrometeor types, the $\partial Z_{DR} / \partial M_s$ values drop to zero, even though snow is still present near the melting layer. Overall, as for Z_{HH} , the highest Jacobian values are obtained for rain content perturbations, due to the large difference in dielectric constant between liquid water and solid hydrometeors.

For K_{DP} , sensitivity is found for rain content perturbations, but the one for solid hydrometeors is negligible (Fig. 10c). Indeed, a rain content perturbation will lead to an increase in amount of matter crossed by radar pulses and then, to a K_{DP} increase. The same results are not obtained for the other hydrometeor contents because of the smaller associated dielectric constant. Concerning ρ_{HV} , no sensitivity have been found, except for rain content perturbations in the melting layer.

²Jacobian study has been done with linear reflectivity units in order to stay closer to a linear regime than with the use of logarithmic reflectivity units.



420 To conclude this section, it has been found that DPOL variables are more sensitive to rain content perturbations than to other hydrometeors, mainly because of large values of liquid water dielectric constant. Another important information is that the most sensitive DPOL variable appears to be the horizontal reflectivity Z_{HH} , followed by the differential reflectivity Z_{DR} and then by the specific differential phase K_{DP} . The co-polar correlation coefficient ρ_{HV} has very small sensitivities to rain content perturbations only. Moreover, since strongly non-Gaussian innovation statistics have been noticed in Section 3.2, this
425 quantity can be hardly used in data assimilation with the current observation operator and cloud microphysical scheme.

4.3 Jacobian matrix in the observation space

430 As observations are not available on the model grid, the NL observation operator has to compute the model equivalent in the observation space. To do so, after an horizontal interpolation of the model profiles to the observation location, \mathcal{H}_{DPOL} computes the DPOL variables on the model profiles and then, interpolate them over the main lobe of the radar beam (see Wattrelo et al. (2014)). Contrary to Jacobians computed in the model space, the one obtained in the observation space are
435 represented by a full Jacobian matrix. It has been computed for the four DPOL variables, but only results for Z_{HH} and Z_{DR} are shown and discussed here. Comparable conclusions can be drawn for the other variables.

Fig. 11 displays such complete $\partial Z_{DR} / \partial M_r$, for the profile displayed in Fig. 10, located at 80 km from the radar. The presence of rain sensitivity between the ground and approximatively 700 hPa is consistent with the hydrometeor content
440 profiles (Fig. 10e). Nevertheless, it can be seen that the values are now split over the two radar elevations which sample rain. Indeed, for a rain content perturbation applied at 800 hPa for example, the impact on the Jacobian values is noticed over the two first radar elevations, with a larger impact on the 0.6° elevation. This behaviour is explained by the Gaussian shape used to represent the main lobe of the radar beam. In that way, a perturbation applied near of the center of the radar main lobe will have a more important impact on the Jacobian than if applied on its sides.

445 An interesting feature is also present on the Z_{DR} Jacobians for a rain content perturbation. Indeed, Z_{DR} is known to increase when the scanned atmosphere is composed of non spherical particles. However, the Jacobian values around 700 hPa indicate that a positive rain content perturbation leads to a small decrease of the Z_{DR} value. As rain water content is small, this is actually caused by the addition of very small rain drops. Indeed, as described in Section 2.1.1, the hydrometeor content is the variable which influences the particule size distribution through the shape distribution parameter. For small hydrometeor contents, a majority of small particles are considered as spherical. In consequence, with small rain contents, a positive perturbation will cause the addition of spherical or nearly spherical particles in the scanned volume and then, a decrease of the Z_{DR} value.

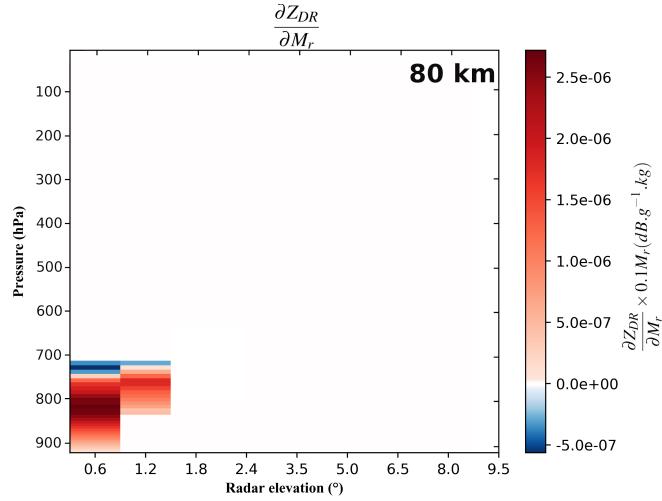


Figure 11. Z_{DR} Jacobian for a rain content perturbation associated to the AROME hydrometeor profile shown in Fig. 10 located at 80 km from the radar. The Jacobians are normalised by 10 percent of the hydrometeor content.

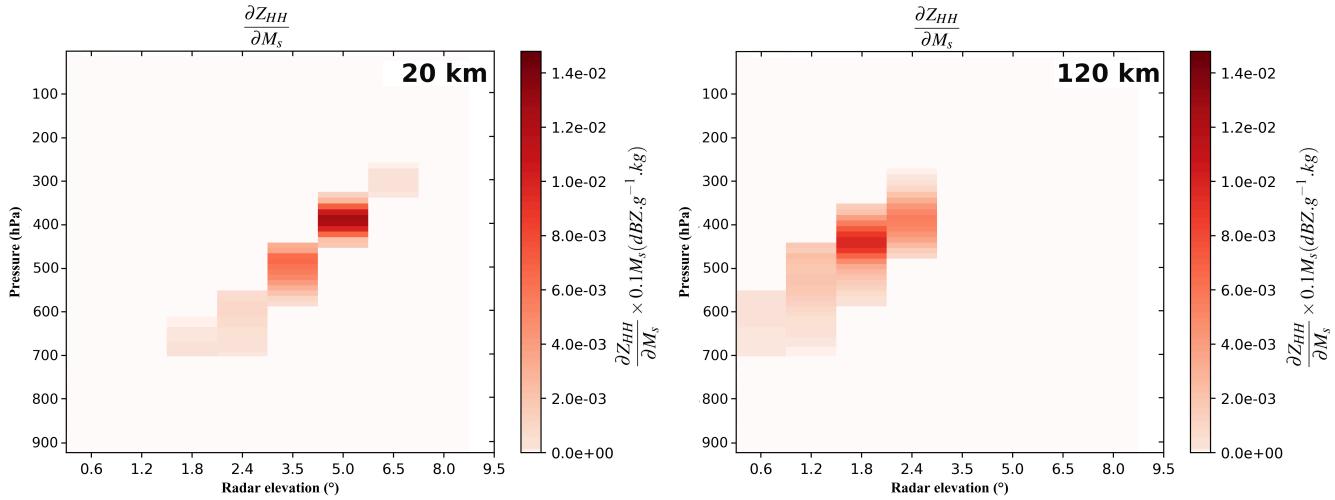


Figure 12. Z_{HH} Jacobian for a snow content perturbation associated to the AROME hydrometeor profile shown in Fig. 10 artificially located at (a) 20 km, and at b) 120 km from the radar. The Jacobians are normalised by 10 percent of the hydrometeor content.

Another important parameter to consider when dealing with radar geometry is the distance to the radar. The radar beam being represented as a cone (see Fig. 3 in Wattrelot et al. (2014)), the width of the beam and the altitude of the sampled volume are proportional to the distance to the radar. To quantify this effect on the Jacobian values, the convective hydrometeor content profile has been artificially placed at 20 km and 120 km from the radar. Fig. 12 presents such Jacobians for Z_{HH} with respect to a snow content perturbation. Firstly, no matter the distance to the radar, the positive snow content perturbation



leads to an increase of Z_{HH} , related to the increase of the total cross-section. Concerning the radar geometry, two effects due to the distance to the radar (beam width and altitude) are observed. The first one is the beam width enlargement. For an elevation of 2.4° , Z_{HH} sensitivity information lies between about 700 hPa and 550 hPa (Fig. 12a) at 20 km from the radar
455 while, at 120 km, it lies between 480 hPa and 280 hPa. The side effect of radar beam broadening is a sensitivity reduction due to a repartition of the same amount of information in a larger volume. Indeed, at 20 km, the highest Jacobian value is about $2.10^{-2} dBZ.g^{-1}.kg \times 0.1M_s$ (Fig. 12a) while at 120 km, it drops to $8.10^{-3} dBZ.g^{-1}.kg \times 0.1M_s$ (Fig. 12b). The second effect of the radar geometry is related to the altitude. Indeed, the farer the observation from the radar is, the higher in the atmosphere it is. This effect is visible in Fig 12. At 20 km from the radar (Fig 12a), the elevation angle 6.5° is low enough to
460 get information in the snow region. Nevertheless, at 120 km from the radar and with an elevation angle of 6.5° , the radar beam is located aloft the snow region.

5 Conclusions

This paper focused on studying operators required for the variational assimilation of polarimetric variables from ground based weather radars in convective scale NWP models. For that purpose, a radar observation operator \mathcal{H}_{DPOL} , based on
465 the T-matrix theory, has been used for the simulation of the following polarimetric variables: horizontal reflectivity Z_{HH} , differential reflectivity Z_{DR} , specific differential phase K_{DP} and co-polar correlation coefficient ρ_{HV} . To simulate these variables, \mathcal{H}_{DPOL} uses hydrometeor contents (rain, snow, graupel and pristine ice) from the AROME-France model. It has been found that more realistic simulations are obtained in the presence of liquid hydrometeors, especially for K_{DP} . To investigate
470 the complexity of DPOL variable simulations, parameters used to characterise hydrometeors in the T-matrix method have been perturbed, such as hydrometeor aspect ratios, dielectric constant or oscillation. A weak sensitivity of the simulations to those parameters has been found, excepted for the dielectric constants of solid hydrometeors in the case of simulated Z_{HH} , and for the rain aspect ratio for Z_{DR} and K_{DP} .

Even if polarimetric radars are able to detect fine spatial structures, filters need to be applied in order to remove non-meteorological data, as well as the possible noise. A positive effect of these filters has been found on innovation statistics for
475 the four DPOL variables computed for twelve different meteorological cases, with reductions of biases and standard deviations. Nevertheless, only Z_{HH} and Z_{DR} innovations distributions appear to be close to a Gaussian shape. Innovation distributions as a function of altitude show the complexity of simulations in presence of solid hydrometeors, but also for levels where melting layer can be encountered.

A linearised version of the polarimetric observation operator has been evaluated by computing its Jacobians with the finite
480 difference method. The results show that polarimetric variables are more sensitive to rain content perturbations than to solid hydrometeor ones, especially because of their different dielectric constants. The Jacobian computation also supports the fact that Z_{HH} appears to be the most sensitive variable to hydrometeor content perturbations, followed by Z_{DR} . Small K_{DP} sensitivity to rain content contents has been found, while no sensitivity has been detected for ρ_{HV} . Then, radar measurement geometry has been considered to study DPOL variables sensitivities. Long distances between radar and the profile of interest



485 decrease the sensitivity due to the beam broadening, but also induce sensitivities at higher altitudes due to the radar elevation angle.

The present results show that only some DPOL variables appear to be promising for the initialisation of hydrometeor contents through variational data assimilation. Among them, the horizontal reflectivity Z_{HH} and the differential reflectivity Z_{DR} are good candidates. The specific differential phase K_{DP} might also be useful for rain. Nevertheless, the simulation of the polarimetric variables for certain type of precipitation or meteorological cases remains difficult. The main reason comes from the 490 ICE3 one-moment microphysical scheme that has been used both in the calibration of the T-Matrix and in the AROME-France NWP model from which the simulations have been performed. In this microphysical scheme, the generalized gamma distributions, used to describe hydrometeor distributions, have shapes which are only driven by the hydrometeor content. DPOL variables being very sensitive to hydrometeor size distributions, such microphysical scheme appears to be limiting. Another 495 limitation is the use of a single particle shape affected by an axis ratio, while DPOL variables are known to be sensitive to hydrometeor shapes. A two moment microphysical scheme coupled with more complex hydrometeor shapes and scattering computation method as DDA (Discrete Dipole Approximation) proposed by DeVoe (1964) could lead to large improvements.

Despite the encountered difficulties for K_{DP} and ρ_{HV} simulations, assimilation tests should be run for Z_{HH} and Z_{DR} for all types of hydrometeors, while K_{DP} could be used for rain content initialisation only. This will be done in a future 500 study performed in a 1D-Var DA system, in which both non-linear and linear operators presented here will be exploited. Quantification of errors in \mathcal{H}_{DPOL} and the study on innovation statistics that have been presented in this paper will also be very useful for characterizing observation errors. Nevertheless, these values constitute first approximations which need to be diagnosed with a more objective method, as the one proposed by Desroziers et al. (2005). In that context, the impact of the DPOL assimilation for analyzing hydrometeor contents, as well as temperature and humidity will be studied in this framework.

505 *Code availability.* The polarimetric observation operator is available in the MESO-NH NWP model. It is a non-hydrostatic research model under the CeCILL-C free licence. For more information, please consult MESO-NH website (<http://mesonh.aero.obs-mip.fr/mesonh54>).

Data availability. The Météo France polarimetric radar data are accessible against licence-fee only. For more information, please consult the following webpage: https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=105&id_rubrique=35

510 *Author contributions.* All the authors conduct the research and analysis. Guillaume Thomas wrote the manuscript. All the authors contribute to the manuscript improvement.

Competing interests. The authors declare that they have no conflict of interest.



Acknowledgements. This research has been conducted during the G. Thomas' PhD, funded by Météo France. The authors would like to thank Sylvain CHAUMONT for the providing of polarimetric radar data, and Béatrice FRADON for her advices and her carefulness about those data. The authors would also thank Eric WATTRELOT for numerous advices given the first year of this PhD. A special aknowledgement also
515 has to be done to Maud MARTEL who helped with the polarimetric data decoding, provided advices on weather radar technology and who accepted to provide an internal revision of this manuscript.



References

Al-Sakka, H., Boumahmoud, A.-A., Fradon, B., Frasier, S. J., and Tabary, P.: A New Fuzzy Logic Hydrometeor Classification Scheme Applied to the French X-, C-, and S-Band Polarimetric Radars, *Journal of Applied Meteorology and Climatology*, 52, 2328–2344, 520 <https://doi.org/10.1175/jamc-d-12-0236.1>, 2013.

Augros, C., Caumont, O., Ducrocq, V., Gaussiat, N., and Tabary, P.: Comparisons between S-, C- and X-band polarimetric radar observations and convective-scale simulations of the HyMeX first special observing period, *Quarterly Journal of the Royal Meteorological Society*, 142, 347–362, <https://doi.org/10.1002/qj.2572>, 2015.

Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., Olson, J. B., James, E. P., Dowell, D. C., Grell, 525 G. A., Lin, H., Peckham, S. E., Smith, T. L., Moninger, W. R., Kenyon, J. S., and Manikin, G. S.: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh, *Monthly Weather Review*, 144, 1669–1694, <https://doi.org/10.1175/mwr-d-15-0242.1>, 2016.

Bick, T., Simmer, C., Trömel, S., Wapler, K., Franssen, H.-J. H., Stephan, K., Blahak, U., Schraff, C., Reich, H., Zeng, Y., and Potthast, R.: Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale, *Quarterly Journal of the Royal Meteorological Society*, 142, 1490–1504, <https://doi.org/10.1002/qj.2751>, 2016.

530 Brandes, E. A., Zhang, G., and Vivekanandan, J.: Experiments in Rainfall Estimation with a Polarimetric Radar in a Subtropical Environment, *Journal of Applied Meteorology*, 41, 674–685, [https://doi.org/10.1175/1520-0450\(2002\)041<0674:eirewa>2.0.co;2](https://doi.org/10.1175/1520-0450(2002)041<0674:eirewa>2.0.co;2), 2002.

Bringi, V. N., Rasmussen, R. M., and Vivekanandan, J.: Multiparameter Radar Measurements in Colorado Convective 535 Storms. Part I: Graupel Melting Studies, *Journal of the Atmospheric Sciences*, 43, 2545–2563, [https://doi.org/10.1175/1520-0469\(1986\)043<2545:mrmicc>2.0.co;2](https://doi.org/10.1175/1520-0469(1986)043<2545:mrmicc>2.0.co;2), 1986.

Brousseau, P., Seity, Y., Ricard, D., and Léger, J.: Improvement of the forecast of convective activity from the AROME-France system, *Quarterly Journal of the Royal Meteorological Society*, 142, 2231–2243, <https://doi.org/10.1002/qj.2822>, 2016.

540 Caniaux, G., Redelsperger, J.-L., and Lafore, J.-P.: A Numerical Study of the Stratiform Region of a Fast-Moving Squall Line. Part I: General Description and Water and Heat Budgets, *Journal of the Atmospheric Sciences*, 51, 2046–2074, [https://doi.org/10.1175/1520-0469\(1994\)051<2046:ansots>2.0.co;2](https://doi.org/10.1175/1520-0469(1994)051<2046:ansots>2.0.co;2), 1994.

Caumont, O., Ducrocq, V., Delrieu, G., Gosset, M., Pinty, J.-P., Parent du Chatelet, J., Andrieu, H., Lemaitre, Y., and Scialom, G.: A Radar Simulator for High-Resolution Nonhydrostatic Models, *Journal of Atmospheric and Oceanic Technology*, 23, 1049–1067, <https://doi.org/10.1175/JTECH1905.1>, <http://journals.ametsoc.org/doi/abs/10.1175/JTECH1905.1>, 2006.

545 Caumont, O., Ducrocq, V., Wattrelot, E., Jaubert, G., and Pradier-Vabre, S.: 1D+3DVar assimilation of radar reflectivity data: a proof of concept, *Tellus A*, 62, 173–187, <http://dx.doi.org/10.1111/j.1600-0870.2009.00430.x>, 2010.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 131, 3385–3396, <https://doi.org/10.1256/qj.05.108>, 2005.

Dowell, D. C., Wicker, L. J., and Snyder, C.: Ensemble Kalman Filter Assimilation of Radar Observations of the 8 May 2003 Oklahoma 550 City Supercell: Influences of Reflectivity Observations on Storm-Scale Analyses, *Monthly Weather Review*, 139, 272–294, <https://doi.org/10.1175/2010mwr3438.1>, 2011.



Ducrocq, V., Ricard, D., Lafore, J.-P., and Orain, F.: Storm-Scale Numerical Rainfall Prediction for Five Precipitating Events over France: On the Importance of the Initial Humidity Field, *Weather and Forecasting*, 17, 1236–1256, [https://doi.org/10.1175/1520-0434\(2002\)017<1236:ssnrf>2.0.co;2](https://doi.org/10.1175/1520-0434(2002)017<1236:ssnrf>2.0.co;2), 2002.

555 Duerinckx, A., Hamdi, R., Mahfouf, J.-F., and Termonia, P.: Study of the Jacobian of an extended Kalman filter for soil analysis in SUR-FEXv5, *Geoscientific Model Development*, 8, 845–863, <https://doi.org/10.5194/gmd-8-845-2015>, 2015.

Errico, R. M., Fillion, L., Nychka, D., and Lu, Z.-Q.: Some statistical considerations associated with the data assimilation of precipitation observations, *Quarterly Journal of the Royal Meteorological Society*, 126, 339–359, <https://doi.org/10.1002/qj.49712656217>, 2000.

Errico, R. M., Bauer, P., and Mahfouf, J.-F.: Issues Regarding the Assimilation of Cloud and Precipitation Data, *Journal of the Atmospheric Sciences*, 64, 3785–3798, <https://doi.org/10.1175/2006jas2044.1>, 2007.

Gao, J. and Stensrud, D. J.: Assimilation of Reflectivity Data in a Convective-Scale, Cycled 3DVAR Framework with Hydrometeor Classification, *J. Atmos. Sci.*, 69, 1054–1065, <http://dx.doi.org/10.1175/JAS-D-11-0162.1>, 2011.

560 Gourley, J. J., Tabary, P., and du Chatelet, J. P.: A Fuzzy Logic Algorithm for the Separation of Precipitating from Non-precipitating Echoes Using Polarimetric Radar Observations, *Journal of Atmospheric and Oceanic Technology*, 24, 1439–1451, <https://doi.org/10.1175/jtech2035.1>, 2007.

Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., Linskog, M., Barkmeijer, J., Dahlbom, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., Alexander, C., Dowell, D., Liu, S., Ikuta, Y., and Fujita, T.: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres, *Quarterly Journal of the Royal Meteorological Society*, 144, 1218–1256, <https://doi.org/10.1002/qj.3179>, 2018.

570 Ikuta, Y. and Honda, Y.: Development of 1D+4DVAR data assimilation of radar reflectivity in JNoVA, *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Model.*, 41, 2011.

Jung, Y., Zhang, G., and Xue, M.: Assimilation of Simulated Polarimetric Radar Data for a Convective Storm Using the Ensemble Kalman Filter. Part I: Observation Operators for Reflectivity and Polarimetric Variables, *Monthly Weather Review*, 136, 2228–2245, <https://doi.org/10.1175/2007mwr2083.1>, 2008.

575 Kumjian, M.: Principles and applications of dual-polarization weather radar. Part I: Description of the polarimetric radar variables, *Journal of Operational Meteorology*, 1, 226–242, <https://doi.org/10.15191/nwajom.2013.0119>, 2013a.

Kumjian, M.: Principles and applications of dual-polarization weather radar. Part II: Warm- and cold-season applications, *Journal of Operational Meteorology*, 1, 243–264, <https://doi.org/10.15191/nwajom.2013.0120>, 2013b.

580 Liu, G.: A Database of Microwave Single-Scattering Properties for Nonspherical Ice Particles, *Bulletin of the American Meteorological Society*, 89, 1563–1570, <https://doi.org/10.1175/2008bams2486.1>, 2008.

Llasat, M. C., Llasat-Botija, M., Prat, M. A., Porcú, F., Price, C., Mugnai, A., Lagouvardos, K., Kotroni, V., Katsanos, D., Michaelides, S., Yair, Y., Savvidou, K., and Nicolaides, K.: High-impact floods and flash floods in Mediterranean countries: the FLASH preliminary database, *Advances in Geosciences*, 23, 47–55, <https://doi.org/10.5194/adgeo-23-47-2010>, 2010.

585 Matrosov, S.: Assessment of Radar Signal Attenuation Caused by the Melting Hydrometeor Layer, *IEEE Transactions on Geoscience and Remote Sensing*, 46, 1039–1047, <https://doi.org/10.1109/tgrs.2008.915757>, 2008.

Mishchenko, M. I. and Travis, L. D.: T-matrix computations of light scattering by large spheroidal particles, *Optics Communications*, 109, 16–21, [https://doi.org/10.1016/0030-4018\(94\)90731-5](https://doi.org/10.1016/0030-4018(94)90731-5), 1994.

Montmerle, T. and Faccani, C.: Mesoscale Assimilation of Radial Velocities from Doppler Radars in a Preoperational Framework, *Monthly Weather Review*, 137, 1939–1953, <https://doi.org/10.1175/2008mwr2725.1>, 2009.



590 Pinty, J. and Jabouille, P.: A mixed-phase cloud parameterization for use in mesoscale non-hydrostatic model: simulations of a squall line and of orographic precipitations, in: Conf. on Cloud Physics, pp. 217–220, American Meteorological Society Everett, WA, 1998.

Ridal, M. and Dahlbom, M.: Assimilation of Multinational Radar Reflectivity Data in a Mesoscale Model: A Proof of Concept, *Journal of Applied Meteorology and Climatology*, 56, 1739–1751, <https://doi.org/10.1175/jamc-d-16-0247.1>, 2017.

Ryzhkov, A., Pinsky, M., Pokrovsky, A., and Khain, A.: Polarimetric Radar Observation Operator for a Cloud Model with Spectral Micro-physics, *Journal of Applied Meteorology and Climatology*, 50, 873–894, <https://doi.org/10.1175/2010jamc2363.1>, 2011.

595 Ryzhkov, A. V., Schuur, T. J., Burgess, D. W., Heinselman, P. L., Giangrande, S. E., and Zrnic, D. S.: The Joint Polarization Experiment: Polarimetric Rainfall Measurements and Hydrometeor Classification, *Bulletin of the American Meteorological Society*, 86, 809–824, <https://doi.org/10.1175/bams-86-6-809>, 2005.

600 Seity, Y., Brousseau, P., Malardel, S., Hello, G., BÃ©nard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective scale operational model, *Monthly Weather Review*, 139, 976–991, <https://doi.org/10.1175/2010MWR3425.1>, <http://journals.ametsoc.org/doi/abs/10.1175/2010MWR3425.1>, 2011.

Seliga, T. A. and Bringi, V. N.: Potential Use of Radar Differential Reflectivity Measurements at Orthogonal Polarizations for Measuring Precipitation, *Journal of Applied Meteorology*, 15, 69–76, [https://doi.org/10.1175/1520-0450\(1976\)015<0069:puordr>2.0.co;2](https://doi.org/10.1175/1520-0450(1976)015<0069:puordr>2.0.co;2), 1976.

Sun, J.: Convective-scale assimilation of radar data: Progress and challenges, *Quarterly Journal of the Royal Meteorological Society*, 131, 605 3439–3463, <https://doi.org/10.1256/qj.05.149>, 2005.

Sun, J. and Crook, N.: Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint. Part I: Model Development and Simulated Data Experiments, *Journal of Atm. Sciences*, 54, 1642–1661, 1997.

Sun, J. and Wang, H.: Radar Data Assimilation with WRF 4D-Var. Part II: Comparison with 3D-Var for a Squall Line over the U.S. Great Plains, *Monthly Weather Review*, 141, 2245–2264, <https://doi.org/10.1175/mwr-d-12-00169.1>, 2013.

610 Tabary, P.: The New French Operational Radar Rainfall Product. Part I: Methodology, *Weather and Forecasting*, 22, 393–408, <https://doi.org/10.1175/waf1004.1>, 2007.

Tabary, P., Fradon, B., and Boumahmoud, A.-A.: La polarimétrie radar à Météo-France, *La Météorologie*, 8, 59, <https://doi.org/10.4267/2042/52055>, 2013.

Tong, M. and Xue, M.: Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments, *Monthly Weather Review*, 133, 1789–1807, <https://doi.org/10.1175/mwr2898.1>, 2005.

615 Wang, H., Sun, J., Fan, S., and Huang, X.-Y.: Indirect Assimilation of Radar Reflectivity with WRF 3D-Var and Its Impact on Prediction of Four Summertime Convective Events, *Journal of Applied Meteorology and Climatology*, 52, 889–902, <https://doi.org/10.1175/jamc-d-12-0120.1>, 2013a.

Wang, H., Sun, J., Zhang, X., Huang, X.-Y., and Auligné, T.: Radar Data Assimilation with WRF 4D-Var. Part I: System Development and Preliminary Testing, *Monthly Weather Review*, 141, 2224–2244, <https://doi.org/10.1175/mwr-d-12-00168.1>, 2013b.

620 Waterman, P.: Matrix formulation of electromagnetic scattering, *Proceedings of the IEEE*, 53, 805–812, <https://doi.org/10.1109/proc.1965.4058>, 1965.

Wattrelot, E., Caumont, O., and Mahfouf, J.-F.: Operational Implementation of the 1D+3D-Var Assimilation Method of Radar Reflectivity Data in the AROME Model, *Monthly Weather Review*, 142, 1852–1873, <https://doi.org/10.1175/mwr-d-13-00230.1>, 2014.