Impact of cloud horizontal inhomogeneity and directional sampling on the retrieval of cloud droplet size by the POLDER instrument

H. Shang\textsuperscript{1,2}, L. Chen\textsuperscript{1}, F. M. Bréon\textsuperscript{3}, H. Letu\textsuperscript{1,4}, S. Li\textsuperscript{1}, Z. Wang\textsuperscript{5}, and L. Su\textsuperscript{1}

\textsuperscript{1}State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China
\textsuperscript{2}College of Resources and Environment, University of the Chinese Academy of Sciences, Beijing, China
\textsuperscript{3}Laboratoire des Sciences du Climat et de l’Environnement, UMR CEA-CNRS-UVSQ, Gif-Sur-Yvette, France
\textsuperscript{4}Research and Information Center, Tokai University, Tokyo, Japan
\textsuperscript{5}Satellite Environment Center, Ministry of Environmental Protection, Beijing, China

Correspondence to: L. Chen (chenlf@radi.ac.cn)

Received: 17 May 2015 – Published in Atmos. Meas. Tech. Discuss.: 1 July 2015
Revised: 4 November 2015 – Accepted: 10 November 2015 – Published: 24 November 2015

Abstract. The principles of cloud droplet size retrieval via Polarization and Directionality of the Earth’s Reflectance (POLDER) requires that clouds be horizontally homogeneous. The retrieval is performed by combining all measurements from an area of 150 km $\times$ 150 km to compensate for POLDER’s insufficient directional sampling. Using POLDER-like data simulated with the RT3 model, we investigate the impact of cloud horizontal inhomogeneity and directional sampling on the retrieval and analyze which spatial resolution is potentially accessible from the measurements. Case studies show that the sub-grid-scale variability in droplet effective radius (CDR) can significantly reduce valid retrievals and introduce small biases to the CDR (\textasciitilde 1.5 \mu m) and effective variance (EV) estimates. Nevertheless, the sub-grid-scale variations in EV and cloud optical thickness (COT) only influence the EV retrievals and not the CDR estimate. In the directional sampling cases studied, the retrieval using limited observations is accurate and is largely free of random noise.

Several improvements have been made to the original POLDER droplet size retrieval. For example, measurements in the primary rainbow region (137–145$^\circ$) are used to ensure retrievals of large droplet (>15 \mu m) and to reduce the uncertainties caused by cloud heterogeneity. We apply the improved method using the POLDER global L1B data from June 2008, and the new CDR results are compared with the operational CDRs. The comparison shows that the operational CDRs tend to be underestimated for large droplets because the cloudbow oscillations in the scattering angle region of 145–165$^\circ$ are weak for cloud fields with CDR $> 15$ \mu m. Finally, a sub-grid-scale retrieval case demonstrates that a higher resolution, e.g., 42 km $\times$ 42 km, can be used when inverting cloud droplet size distribution parameters from POLDER measurements.

1 Introduction

Liquid water clouds cover approximately 20–30\% of the globe and play an important role in the Earth’s radiation balance (Zeng et al., 2011). One of the key radiative properties of liquid water clouds is the droplet size distribution, which is represented by the droplet effective radius (CDR) and the effective variance (EV) (Bréon and Doutriaux-Boucher, 2005; Bréon and Colzy, 2000; Hansen and Travis, 1974). The observation of CDR and EV not only has a significant influence on the modeling of clouds’ climate feedbacks (Stubenrauch et al., 2013; Dandin et al., 1997) but is also meaningful to aerosol–cloud–precipitation interaction research (Penner et al., 2004; Shang et al., 2014). Satellite retrievals of the CDR and EV have been used to extend the detailed knowledge gained from specific cloud campaign studies to the larger spatial and temporal scales that are relevant to climate (Liu et al., 2002, 2011; Bréon, 2006).

The remote sensing of the droplet size distribution using passive sensors is achieved by either the bi-spectral re-
flectance method or the multi-angular polarized reflectance method. The bi-spectral reflectance method (Nakajima and King, 1990) is based on the measured reflectance in the visible (0.8 µm) and near-infrared bands (either 1.6, 2.1, or 3.7 µm), which are jointly sensitive to cloud optical thickness and particle size. In operation, the measured reflectances in the visible and near-infrared bands are compared to the reflectances in a pre-calculated look-up table (LUT) to derive the cloud optical thickness and the CDR simultaneously. This method has been used to derive daytime CDR from the Global Retrieval of ATSR Cloud Parameters and Evaluation (ATSR-GRAPE) (Sayer et al., 2011), the Pathfinder Atmospheres – Extended (PATMOS-x) (Walther and Heidinger, 2012), the Visible Infrared Imaging Radiometer Suite (VIIRS) (Walther et al., 2013), the National Oceanic and Atmospheric Administration/Advanced Very-High-Resolution Radiometer (NOAA/AVHRR) (Nakajima and Nakajima, 1995; Kawamoto et al., 2001) and the Moderate-Resolution Imaging Spectroradiometer (MODIS) (Baum et al., 2000; Platnick et al., 2003; Sayer et al., 2011; Walther and Heidinger, 2012).

Nighttime cloud properties have been derived from reflected moonlight (0.7 and 3.75 µm) as measured by the VIIRS instrument (Walther et al., 2013).

Although the bi-spectral method is well established and widely used, it still suffers from significant limitations. In particular, it cannot provide useful information on the EV which is assumed in the computation of the LUT (Painemal and Zuidema, 2011). For instance, the MODIS algorithm assumes EV = 0.1. In reality, the cloud EV changes with cloud type (Zhang, 2013), and the EV of many stratuscumulus cloud fields has been found to be equal to or less than 0.05 (Bréon and Doutriaux-Boucher, 2005). In addition, the LUT is calculated for a plane-parallel cloud field using a 1-D radiative transfer model that does not consider the influence of the 3-D structure of clouds. However, Horváth (2004) found that no more than 17% of the pixels containing marine liquid water clouds are plane-parallel objects, suggesting that the retrieval error arising from the solar-viewing geometry cannot be neglected (Zinner et al., 2010; Wolters et al., 2010; Vant-Hull et al., 2007; Di Girolamo et al., 2010). The effects of cloud horizontal homogeneity also make the retrieval more complex. Marshak et al. (2006) found that ignoring the cloud variability at the sub-pixel scale results in underestimates of the CDR, while ignoring cloud inhomogeneity at scales exceeding the pixel scale can lead to overestimates. It is found that the vertical structure induced by drizzle and 3-D radiative effects operate together to cause dramatic differences between the 1.6, 2.1, and 3.7 µm retrievals (Zhang et al., 2012; Zhang, 2013; Nakajima et al., 2010a, b; Nagao et al., 2013). In addition, the water vapor absorption within a cloud and the presence of an absorbing aerosol layer above a cloud leads to a positive bias in the retrieval (Alexandrov et al., 2012a; Coddington et al., 2010; Haywood et al., 2004).

The multi-angular polarized method (Bréon and Goloub, 1998) was developed for the Polarization and Directional-
clouds may not satisfy the homogeneity assumption (Schutgens and Roebeling, 2009). Moreover, the coarse resolution limits the usage in certain aerosol–cloud interaction studies (Sekiguchi, 2003; Gryspeerdt et al., 2014). Based on the above analysis, the following two questions are the focal points of this study.

1. What are the effects of cloud horizontal inhomogeneity on the CDR and EV estimates from multi-angle polarized observations?

2. Can droplet size distributions be retrieved with reduced directional sampling from POLDER measurements, which would allow a better spatial resolution?

To answer these questions, both POLDER-like data simulated with the RT3 model and POLDER L1B measurements were used. The remainder of this paper is organized as follows: in Sect. 2, POLDER’s polarized CDR and EV retrieval procedure is described. The retrieval cases, which aim to examine the effects of cloud horizontal inhomogeneity on the CDR, EV, and cloud optical thickness (COT), are presented in Sect. 3. Then, the impact of directional sampling on the retrieval is presented, considering both the performances at different wavelengths and the noise in the inversion. Sub-grid-scale retrieval cases based on POLDER measurements are provided in Sect. 4, and the study is summarized in Sect. 5.

2 Data and methods

2.1 POLDER-like observations simulated with the RT3 model

The polarized radiative transfer (RT3) model (Evans and Stephens, 1991, 2010) is used to solve the plane-parallel case of polarized monochromatic radiative transfer for isotropic media, in which the particle properties and the scattering processes are restricted by Hovenier et al. (2014). Both solar and thermal sources of radiation are considered.

The extinction coefficients, scattering coefficients, and the phase matrix information of liquid cloud particles are precalculated based on the Lorenz–Mie code of Bohren and Huffman (1983), where the phase matrix information is described in the form of coefficients for Legendre polynomial expansions (Hansen and Travis, 1974; Hovenier, 2012; Hovenier and Van der Mee, 1983). This information and the cloud layer height are loaded into RT3 from a precalculated file. The polarization in the case of aerosol and molecular layers is treated the same as that of cloud layers (Kotchenova et al., 2006). To represent the multiple scattering process, the RT3 model adopts the doubling and adding technique (Hansen and Travis, 1974; Hansen and Hovenier, 1971). The underlying ground surface in RT3 can be modeled as Lambertian or Fresnel surfaces (Kotchenova and Vermote, 2007). The angular field of the radiation is expressed with a Fourier series in azimuth angle and discretized zenith angle based on the method of (Ishimaru et al., 1984). This enables fast and accurate simulations of multi-directional polarized radiance (expressed as a four-vector of Stokes parameters \( I, Q, U, V \)) at a particular optical depth or height. Notably, the scattering angle simulated for a certain pixel in RT3 is no more than 32. We calculated the POLDER-like reflectances along the rainbow region with a 1° interval using cubic spline interpolation.

In this study, the RT3 model was used to simulate the POLDER-like reflectances (with Eq. 1) of a cloud field with a known droplet size distribution. The simulations are applied using three polarized channel wavelengths (490, 670, and 865 nm). The underlying ground surface is treated as a Lambertian surface with a constant albedo of 0.02. The atmosphere (Table 1) was assumed to consist of three plane-parallel layers (Cheng et al., 2008). We simulated reflectances using cloud fields of 0.5 km geometric thickness with various values of cloud optical thickness \( \tau \) and cloud heights of 3 km. The CDR values of the clouds ranged from 5 to 20 \( \mu \)m in 1 \( \mu \)m increments, and the EV values were 0.01, 0.02, and 0.05. Three COT values (1, 5, and 10) were considered in this study, and the contributions from the underlying surface and the aerosol and molecule layers are negligible (Coddington et al., 2010; Goloub et al., 2000). The Rayleigh optical thickness for different wavelengths was set according to the results of Bodhaine et al. (1999). The solar zenith angle was assumed to be 20°, and the relative azimuth angle was 180° (the sensor was in front of the sun), which avoids

Table 1. The atmospheric layers and components assumed in the RT3 model for simulating the POLDER-like polarized signal. Lambertian surfaces are treated in the simulation with constant albedo value. The \( \tau_{\text{molecular}} \) and \( \tau_{\text{clouds}} \) represent the optical thickness of aerosol and cloud layers, respectively.

<table>
<thead>
<tr>
<th>Layers</th>
<th>Components</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface</td>
<td>Ocean, Molecules,</td>
<td>Albedo = 0.02</td>
</tr>
<tr>
<td>0–3 km</td>
<td>aerosols</td>
<td>( \tau_{\text{molecular}} = 0.051, 0.014 ) and 0.051 for 490, 670, and 865 nm wavelengths, respectively. ( \tau_{\text{aerosol}} = 0.3 )</td>
</tr>
<tr>
<td>3–3.5 km</td>
<td>Liquid water clouds</td>
<td>( \tau_{\text{clouds}} = 1, 5, 10 )</td>
</tr>
<tr>
<td>3.5–15 km</td>
<td>Molecules</td>
<td>( \tau_{\text{molecular}} = 0.105, 0.029 ) and 0.104 for 490, 670, and 865 nm wavelengths, respectively.</td>
</tr>
</tbody>
</table>

www.atmos-meas-tech.net/8/4931/2015/ Atmos. Meas. Tech., 8, 4931–4945, 2015
Potential error induced by the rotation to the scattering plane (Hansen and Travis, 1974; Alexandrov et al., 2012a).

2.2 POLDER multi-angle polarized observations

POLDER is a French sensor run by CNES (Centre National d’Etudes Spatiales) and was launched on PARASOL in December 2004. It is a multi-spectral imaging polarimeter composed of a 2-D charged coupled device (CCD) detector array, a rotating wheel with spectral filters and polarizers, and wide field-of-view (1800 km) telecentric optics (Parol et al., 2004; Tanré et al., 2011). The CCD detector array provides a spatial resolution of approximately 6 km at nadir. When it passes over a target, the POLDER instrument acquires up to 16 successive multi-angle measurements of both the total and polarized reflected solar radiances in nine narrow channels with center wavelengths of 443, 490, 565, 670, 763, 765, 865, 910, and 1020 nm. For three of the nine spectral bands (490, 670, and 865 nm), a polarizer is added to the filters to assess the degree of linear polarization and the polarization direction, and are related to the Stokes vectors of the polarization (Leroy et al., 1997).

In December 2009, PARASOL was maneuvered out of the Afternoon Constellation, known as the “A-Train” (Nakajima et al., 2010b; Stephens et al., 2002). The POLDER measurements before 2009 agree well with those of other sensors (Stubenrauch et al., 2013; Zeng et al., 2011). Therefore, the POLDER L1B radiance and L2 CDR products in 2008 were used in this study. The polarized reflectance \( R_p \) adopted in this study was normalized as follows:

\[
R_p = -\frac{4}{\mu_s} \frac{Q}{E_0} \left( 1 + \frac{\mu_s}{\mu_v} \right),
\]

where the vector \( Q \) is described with respect to the scattering plane defined by the solar and viewing directions. \( E_0 \) (\( W \cdot m^{-2} \)) is the extraterrestrial solar irradiance, and \( \mu_s \) and \( \mu_v \) are the cosines of the solar and viewing angles, respectively. The polarized reflectance is positive when the polarization direction is orthogonal to the scattering plane and negative when it is parallel to that plane. Note that most scattering and reflection processes generate a polarization perpendicular to the scattering plane; thus, the definition of \( R_p \) used here results in positive values. This choice of sign was adopted by Bréon and Doutriaux-Boucher (2005) and is opposite to that proposed by Alexandrov et al. (2012a). The normalization term in the parentheses is used so that, in the single scattering approximation, \( R_p \) is proportional to the polarized scattering phase function.

2.3 POLDER operational cloud droplet retrievals

The CDR and EV are retrieved at a spatial resolution of approximately 100 km x 100 km and are not part of the standard POLDER/PARASOL cloud parameters. However, the CDR retrieval is very precise and accurate when the requirements for narrow size distributions and homogeneous distributions are met (Bréon and Goloub, 1998). Both daily and monthly CDR data are available from http://www.icare.univ-lille1.fr/drupal/parasol/overview_product.

2.4 Retrieval method

The first step of the retrieval algorithm is to calculate the LUT of the polarized phase functions \( P \) for various CDR and EV values. The CDR \( (r_{\text{eff}}) \) and EV \( (v_{\text{eff}}) \) we adopted in this study are defined as follows (Hansen and Travis, 1974):

\[
r_{\text{eff}} = \int_{r_1}^{r_2} \pi r^3 n(r) dr / \int_{r_1}^{r_2} \pi r^2 n(r) dr,
\]

\[
v_{\text{eff}} = \int_{r_1}^{r_2} (r - r_{\text{eff}})^2 \pi r^2 n(r) dr / \int_{r_1}^{r_2} \pi r^2 n(r) dr,
\]

where \( n(r) dr \) is the number of particles per unit volume with a radius between \( r \) and \( r + dr \), and \( r_1 \) and \( r_2 \) are the smallest and largest particles in the size distribution, respectively. The cloud droplets are parameterized in the form of a gamma distribution (Lafrance et al., 2002) with the following form:

\[
\frac{nr}{\Gamma[(1-2v_{\text{eff}})/v_{\text{eff}}]} \cdot e^{-r_{\text{eff}}/v_{\text{eff}}},
\]

\[
C = \int \left[ \frac{n(r) dr}{\Gamma[(1-2v_{\text{eff}})/v_{\text{eff}}]} \right] / \Gamma[(1-2v_{\text{eff}})/v_{\text{eff}}],
\]

where \( \Gamma \) is the Gamma function, and \( N \) is the total number of particles per unit volume. We adopted the Lorenz–Mie code of Bréon and Doutriaux-Boucher (2005) in calculating the LUT of polarized phase functions \( P \). The angular resolution of the LUT was 0.5\(^\circ\) for the rainbow scattering angles, the CDR values ranged from 5 to 20 \( \mu \)m with 0.5 \( \mu \)m increments, and the EV values were 0.01, 0.02, 0.05, and 0.1. Although the operational algorithm used the measurements in the scattering angle range of 145–180\(^\circ\), we also used those in the primary rainbow region (137–145\(^\circ\)). These angles were used to increase the retrieval accuracy of large CDRs (15–20 \( \mu \)m). In addition, we added the polarized phase functions for EV = 0.1 in the LUT, allowing the EV to be retrieved over a broader range. In Fig. 1a, the angular positions of the maxima and minima in the polarized scattering phase function are sensitive to the CDR. In Fig. 1c, the polarized reflectance local maxima and minima are sensitive to the EV. In Fig. 1b, the angular distance between maxima decreases with increasing wavelength. For example, twice as many oscillations occur at 490 nm as at 865 nm.

The second step is to fit the observed polarized reflectances (OPRs) with the polarized phase functions \( P \) in the LUT, which is based on the following expression of Bréon and Doutriaux-Boucher (2005):

\[
R_p (\lambda) = a \cdot P_p (\Theta, r_{\text{eff}}, v_{\text{eff}}, \lambda) + b \cdot \cos^2 (\Theta) + c,
\]
The EV to be retrieved over a broader range; and (3) using

\( \cos^2(\Theta) \) instead of \( \Theta \) in the second term in Eq. (6) to match the POLDER measurements and the pre-calculated phase functions.

3 Retrieval tests using RT3 simulations

3.1 Sub-grid-scale variability in the CDR and EV

The impact of sub-grid-scale variability in the CDR and EV is assessed through a modeled cloud field consisting of several equal-area subregions with a constant COT but with variable CDR and EV values. Importantly, the mixture of two or more gamma size distributions (subregions) is not another gamma size distribution, and the mean droplet effective radius and variance of the combined distributions is not simply the average of the effective radii and variances of the subregions. The mean droplet effective radius and variance were calculated using the method of Alexandrov and Lacis (2000) and Alexandrov et al. (2012b). Retrievals from a heterogeneous cloud field are presented in Fig. 2. The polarized reflectances were simulated at 865 nm. In Fig. 2a, we assumed that one-third of a POLDER pixel was covered by a cloud with CDR = 10 \( \mu \)m and EV = 0.01, another third was covered by a cloud with CDR = 15 \( \mu \)m and EV = 0.01, and the remaining third was covered by a cloud with CDR = 20 \( \mu \)m and EV = 0.01. Similarly, in Fig. 2b, half of a POLDER pixel was assumed to be covered by a cloud with CDR = 10 \( \mu \)m and EV = 0.01, and the other half was covered by a cloud with CDR = 20 \( \mu \)m and EV = 0.01. The mean effective radii and variances for the mixtures in Fig. 2a, b, and c are 17.07 \( \mu \)m and 0.06, 18.00 \( \mu \)m and 0.06, and 15 \( \mu \)m and 0.03, respectively. The examples in Fig. 2a and b show that the retrieved CDR values based on the mean reflectance of inhomogeneous pixels tend to be smaller (~1.5 \( \mu \)m) than the mean of the sub-pixel CDRs. Furthermore, the retrieved EV was larger than that at the sub-pixel level, because the averaging of the signal from different CDRs reduces the amplitude of the polarized reflectance oscillations. Figure 2c assumes the same CDRs but different EV values (0.01, 0.02, and 0.05) within three sub-pixels; the retrieved CDR is accurate and the retrieved EV is close to the mean of the sub-pixel EV values.

The scattering angle range used in the operational POLDER procedure is 145–165° and does not include the primary rainbow region of 137–145°. To further assess the information content of the primary rainbow structure for the retrieval, more cases were examined with respect to CDR variability. Each case was retrieved twice, using either the 137–165° or the 145–165° scattering angle ranges. The POLDER-like polarized reflectances used in each retrieval are with a directional interval of 0.2°. As shown in Table 2, more valid retrievals are received from the former group (137–165°) than the latter group (145–165°), and the CDR is underestimated by 8.2 \( \mu \)m for the case of “15 + 20” in
Figure 2. The retrievals from a heterogeneous cloud field with constant $COT = 5$ and variable CDR and EV values. The dashed lines indicate the separate rainbow structures for sub-grid-scale cloud fields. Three equal-area subparts with CDR = 10, 15, and 20 $\mu$m were considered in (a); two equal-area subparts with CDR = 10 and 20 $\mu$m were considered in (b); three equal-area subparts with EV = 0.01, 0.02, and 0.05 were considered in (c); the blue line represents the rainbow structure for the heterogeneous cloud field; the red line depicts the best fit.

Table 2. Retrievals from a heterogeneous cloud field with variable CDRs using POLDER-like polarized reflectances (865 nm) in 137–165° and 145–165° ranges, respectively. In all cases, the EV in the sub-scale cloud and the COT were assumed to be 0.01 and 5, respectively. The “+” indicates the equal share of the CDRs in the cloud fields. The mean CDR and EV indicate the effective radii and variances for the combined droplet size distributions. The CDR and EV estimates are restricted with $T_1 > 0.978$ and $T_2 < 0.01$.

<table>
<thead>
<tr>
<th>Combined CDRs ($\mu$m)</th>
<th>Sub-scale EV</th>
<th>Mean CDR ($\mu$m)</th>
<th>Mean EV</th>
<th>Retrievals of 137–165°</th>
<th>Retrievals of 145–165°</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CDR ($\mu$m)</td>
<td>EV</td>
</tr>
<tr>
<td>$5 + 10$</td>
<td>0.01</td>
<td>9.00</td>
<td>0.06</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$5 + 15$</td>
<td>0.01</td>
<td>14.00</td>
<td>0.06</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$5 + 20$</td>
<td>0.01</td>
<td>19.12</td>
<td>0.04</td>
<td>13.0</td>
<td>0.1</td>
</tr>
<tr>
<td>$10 + 15$</td>
<td>0.01</td>
<td>13.46</td>
<td>0.04</td>
<td>13.0</td>
<td>0.1</td>
</tr>
<tr>
<td>$10 + 20$</td>
<td>0.01</td>
<td>18.00</td>
<td>0.06</td>
<td>16.5</td>
<td>0.1</td>
</tr>
<tr>
<td>$15 + 20$</td>
<td>0.01</td>
<td>18.20</td>
<td>0.03</td>
<td>17.5</td>
<td>0.05</td>
</tr>
<tr>
<td>$5 + 10 + 15$</td>
<td>0.01</td>
<td>12.70</td>
<td>0.11</td>
<td>12.0</td>
<td>0.1</td>
</tr>
<tr>
<td>$5 + 10 + 20$</td>
<td>0.01</td>
<td>16.92</td>
<td>0.13</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$5 + 15 + 20$</td>
<td>0.01</td>
<td>17.35</td>
<td>0.08</td>
<td>17.5</td>
<td>0.05</td>
</tr>
<tr>
<td>$10 + 15 + 20$</td>
<td>0.01</td>
<td>17.07</td>
<td>0.06</td>
<td>16.0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

the latter group. These results demonstrate that the retrievals with the primary rainbow measurements are more reliable. In addition, the CDR estimates of the former group are close to the mean radii, with biases of less than 1.5 $\mu$m. Regarding the EV estimates, both the retrievals of the two groups have considerable biases with no identifiable trends.

In conclusion, the heterogeneity in the cloud field CDR significantly reduces valid droplet size distribution retrievals, and introduces uncertainties to its mean estimate when using the operational procedure. However, the impact of this variability is very much reduced when using the information content of the primary rainbow (i.e., the angular range 137–145°).

3.2 Sub-grid-scale variability in COT

The current POLDER size distribution retrieval procedure adopts a relatively coarse resolution that may introduce errors by simply assuming that a cloud is homogeneous within an area of 150 km × 150 km. We considered a heterogeneous cloud with a constant CDR and variable optical depth to investigate whether the variability in the COT affects the POLDER CDR retrievals. The polarized reflectance for heterogeneous cloud fields with COT = 1, 5, and 10 and CDR = 5, 11, and 16 $\mu$m were simulated using the RT3 model for the wavelength 865 nm; and the POLDER-like polarized reflectances used in each retrieval are with a directional interval of 0.2°. The size distribution retrieval results are presented in Table 3. The results indicate that, in most cases, the COT variability has a negligible impact on the CDR estimate, although it does affect the retrieved EV values. Variability in COT in a cloud field changes the amplitude of the rainbow structure, which is similar to the impact of the EV, whereas the angular positions of the peak values are insensitive to changes in the COT. Figure 3 illustrates the results of COT = 1, 5, and 10 for CDR = 11 $\mu$m and EV = 0.02. All areas (different COT values) generate a polarized reflectance with a similar angular shape. The mixed curve agrees well with the best fit curve, indicating that the retrieval is very accurate.

In conclusion, our simulations indicate that the spatial variability of the cloud optical thickness has no discernable
Table 3. Retrievals from a heterogeneous cloud field with variable COT using POLDER-like polarized reflectances in the 137–165° range. In all cases, the EV in the sub-grid-scale cloud was assumed to be 0.02, and the CDR values included 5, 11, and 16 µm. The “+” indicates the equal share of the COTs in the cloud fields.

<table>
<thead>
<tr>
<th>Input cloud optical thickness</th>
<th>Actual CDR (µm)</th>
<th>Actual EV</th>
<th>Retrieved CDR (µm)</th>
<th>Retrieved EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 + 5</td>
<td>5</td>
<td>0.02</td>
<td>5</td>
<td>0.05</td>
</tr>
<tr>
<td>5 + 10</td>
<td>5</td>
<td>0.02</td>
<td>5</td>
<td>0.05</td>
</tr>
<tr>
<td>1 + 10</td>
<td>5</td>
<td>0.02</td>
<td>5</td>
<td>0.05</td>
</tr>
<tr>
<td>1 + 5 + 10</td>
<td>5</td>
<td>0.02</td>
<td>5</td>
<td>0.05</td>
</tr>
<tr>
<td>1 + 5</td>
<td>11</td>
<td>0.02</td>
<td>11</td>
<td>0.02</td>
</tr>
<tr>
<td>5 + 10</td>
<td>11</td>
<td>0.02</td>
<td>11</td>
<td>0.02</td>
</tr>
<tr>
<td>1 + 10</td>
<td>11</td>
<td>0.02</td>
<td>11</td>
<td>0.02</td>
</tr>
<tr>
<td>1 + 5 + 10</td>
<td>11</td>
<td>0.02</td>
<td>11</td>
<td>0.02</td>
</tr>
<tr>
<td>1 + 5</td>
<td>16</td>
<td>0.02</td>
<td>16</td>
<td>0.02</td>
</tr>
<tr>
<td>5 + 10</td>
<td>16</td>
<td>0.02</td>
<td>16</td>
<td>0.02</td>
</tr>
<tr>
<td>1 + 10</td>
<td>16</td>
<td>0.02</td>
<td>16</td>
<td>0.02</td>
</tr>
<tr>
<td>1 + 5 + 10</td>
<td>16</td>
<td>0.02</td>
<td>16</td>
<td>0.02</td>
</tr>
</tbody>
</table>

impact on the CDR estimate when using the retrieval method based on the polarized reflectance. However, this variability does have some impact on the retrieval of the effective variance.

3.3 The number of observations used in the retrieval

In this study, we assume that there is an angular criterion (the angular number in the scattering angles between 137 and 165°) for the accurate POLDER cloud droplet size retrieval. When the angular sampling of a given resolution satisfies the criterion, the resolution is considered to be efficient for the retrieval. Additionally, we assume that the POLDER measurements are distributed across the same interval in the rainbow region. Based on these assumptions, the angular criterion directly determines the optimal resolution of POLDER’s CDR and EV retrievals.

To understand whether a resolution is sufficient for POLDER droplet size retrievals, the effects of directional sampling were investigated. By controlling the number of POLDER-like polarized reflectance observations (N) used in the retrieval, the relationship between the retrieval accuracy and the number of observations can be analyzed. Before the simulation, we compared the directional distribution of the scattering angle (SA) with real POLDER data and found that the measurements had nearly the same SA interval, ranging from 4 to 10°, which was determined by the solar-viewing geometry, the time interval between successive acquisitions, and satellite speed. The satellite speed and period of acquisition are constant, but the solar zenith angle varies along the orbit; therefore, the viewing geometry in the cloudbow view direction varies along the orbit. In the simulated cases described below, we assume that the POLDER multi-directional observations cover the entire cloudbow angular range and are evenly distributed. We also assumed that the measurements are not affected by noise induced by the instrument and the spatial structure of clouds. In each retrieval test, the observations were evenly distributed over the rainbow scattering range (137–165°).

Comparisons of the retrievals at 865 nm with the actual CDR and EV values are shown in Fig. 4; N was increased from 5 to 100 in these simulations. The results suggest that the correlation ($R^2$) and the RMSE do not change after N exceeds 12. Each plot contains 16 retrievals, with the actual CDR ranging from 5 to 20 µm. As the number of observations was increased from 5 to 8, an increase in the $R^2$ between the CDR retrievals and the actual values was observed; the retrieval accuracy of the EV also improved for larger N. The $R^2$ and the RMSE for N > 12 are stable at 0.99 and 0.13, respectively.

As expected, large uncertainties occurred in the CDR and EV retrievals with fewer than eight observations; robust re-
Comparisons of the retrieved CDR and EV with the actual values assumed in the RT3 modeling. In all cases, the actual EV and the COT were assumed to be 0.05 and 5, respectively. $N$ in every plot represents the number of POLDER-like reflectances employed in the inversion. From the top left to the bottom right, $N$ increases from 5 to 12. As $N$ increases from 12 to 100, the $R^2$ and the RMSE are stable at 0.99 and 0.13, respectively. The case for $N=80$ is given last. The color of the points represents the EV results.

Figure 4. Comparisons of the retrieved CDR and EV with the actual values assumed in the RT3 modeling. In all cases, the actual EV and the COT were assumed to be 0.05 and 5, respectively. $N$ in every plot represents the number of POLDER-like reflectances employed in the inversion. From the top left to the bottom right, $N$ increases from 5 to 12. As $N$ increases from 12 to 100, the $R^2$ and the RMSE are stable at 0.99 and 0.13, respectively. The case for $N=80$ is given last. The color of the points represents the EV results.

Retrievals are found when 10 or more observations are provided. Figure 5 further exhibits the fitting results for retrieval cases with insufficient ($N=5$) and sufficient ($N=12$) observations. In Fig. 5a and b, the observed rainbow structures coincide well with the best fit curves; however, the retrievals are not reliable. Figure 5c and d present reliable fitting curves when $N=12$. The observed rainbow structure and the best fit curves do not perfectly coincide with each other because of uncertainties induced by the spline interpolation. These results confirm that the observed rainbow structures can be precisely represented when sufficient observations are provided.

3.4 $N$ requirements at different wavelengths

The oscillations in the rainbow structure are known to vary with wavelength, which may cause the number of observations required at different wavelengths to vary. To compare the retrievals based on the 490, 670, and 865 nm wavelengths, we retrieved the CDR and EV values at these three wavelengths. At the wavelengths of 670 and 865 nm, the RMSE values between the 16 group retrievals (the actual CDRs ranged from 5 to 15 µm) and the three EV values were analyzed. The analysis for CDRs between 16 to 20 µm is missing because the 490 nm Legendre series for cloud fields with large droplets exceeded the limitations of the RT3 model.

Figure 6a, b, and c present RMSE comparison results at three wavelengths for a cloud field with EVs of 0.01, 0.02, and 0.05, respectively. The better performance for 490 nm than for 670 nm is attributable to the absence of large droplets (> 15 µm). When the large droplets were removed from the comparison, the performances in three wavelengths were similarly good. At least 11 observations were needed for the 670 nm wavelength to maintain an RMSE less than 1; however, the number of observations needed at 865 nm was 8. This difference is caused by the higher number of oscillations associated with the increase in wavelength; therefore, more observations are required to fit the curve. As EV increased, the slopes of the oscillations became gentler; thus, fewer observations were required. For a limited number of observations, the wavelengths of 865 nm or 670 nm were found to be more reliable than 490 nm, especially for large droplets. The EV retrieval accuracies for the three wavelengths share a similar trend with the CDR retrievals: fewer observations were required at larger wavelengths. A comparison of the retrievals using the three wavelengths with 12 observations with the actual value EV = 0.02 is presented in Table 4.

In conclusion, the accuracy of the retrieval results increases with an increase in the wavelength. Thus, for POLDER, the 865 nm channel is better suited than the other two channels when limited angular observations are employed. However, the shorter wavelength channel does pro-
Table 4. Comparison of the CDR and EV retrievals for three wavelengths using 12 observations with the actual EV = 0.02; the COT was assumed to be 5 in all cases. The absence of retrievals for CDR = 16–20 µm is due to the Legendre series for cloud fields having effective radii greater than 15 µm, which exceeds the limitations of the RT3 model.

<table>
<thead>
<tr>
<th>Actual CDR (µm)</th>
<th>490 nm</th>
<th>670 nm</th>
<th>865 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CDR (µm)</td>
<td>EV</td>
<td>CDR (µm)</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0.02</td>
<td>5.5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0.02</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>0.02</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0.02</td>
<td>8.5</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0.02</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0.05</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>0.05</td>
<td>11.5</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>0.05</td>
<td>12.5</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>0.05</td>
<td>14</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>0.05</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>0.05</td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td>16.5</td>
<td>0.02</td>
<td>15.5</td>
</tr>
<tr>
<td>17</td>
<td>19</td>
<td>0.05</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>19.5</td>
<td>0.05</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>0.05</td>
<td>19.5</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>0.05</td>
<td>20</td>
</tr>
</tbody>
</table>

Figure 6. Comparisons of RMSE for the CDR retrievals using different numbers of observations at wavelengths of 490, 670, and 865 nm and for EVs of 0.01, 0.02, and 0.05, respectively. The COT was assumed to be 5 in all cases.

3.5 Noise in the retrievals

The theoretical analysis in Sects. 3.1 and 3.2 indicates that the inversion is stable if a sufficient number of observations is provided. However, the POLDER observations are not always optimal, and noise may be introduced by instruments, the transmission of the signal, and/or the inhomogeneous structure of the cloud fields. This noise can be modeled as pseudo-random errors added on the multi-angle polarized reflectances (Fougnie et al., 2007; Cairns et al., 2003).

To analyze the noise sensitivity of the droplet size retrieval, we added Gaussian noise with increasing standard deviation to the simulated polarized reflectances (van Diedenhoven et al., 2012). Retrievals at 865 nm with actual CDR values of 5–20 µm and actual EV value of 0.05 were evaluated. Because the influence of Gaussian noise is related to the number of observations, we performed each retrieval three times, with 9, 12, and 20 observations in the rainbow region. As shown in Fig. 7, the retrievals are essentially unaffected by random noise when the standard deviations are less than 10% of the signal. The RMSE values and the maximum errors (not shown) of the CDR estimates are no more than 0.05 and 1 µm, respectively. For noise contributions larger than the 10% of the signal, the RMSE values of the CDR estimate increase rapidly. The instrument- and calibration-related noise in the POLDER-measured polarized reflectance values is considered to be within 3% (van Diedenhoven et al., 2012; Fougnie et al., 2007). These results confirm that the inversion is robust as long as the noise contributions do not exceed 10% of the valid signal, even if the observations are limited.
4 POLDER retrieval results

4.1 Comparisons with the operational CDRs

We applied the size distribution retrievals to the improved method using the global POLDER L1B data from June 2008. The retrievals were performed with scattering angles of 137–165°, polarized reflectances at the 865 nm wavelength, and a resolution of 150 km × 150 km. To make the comparison more convenient and reliable, an inversion was performed according to the geolocations of operational size distribution products. We adopted the inversion quality index (Qual) to evaluate our retrievals because the Qual is used in the operational procedure to include the bias generated from the variability of the signal. The Qual is the ratio between the variability in the signal and the RMSE of the fitting. A larger Qual value means a smaller fitting bias. Detailed information on Qual can be found in Bréon and Doutriaux-Boucher (2005). Strict conditions ($T_1 > 0.98$ and $\text{Qual} > 6$) were used to select the new retrievals, and the comparisons with the operational CDRs are shown in Fig. 8. The CDRs derived from the improved method are well correlated with the operational CDRs of 5 to 14 µm. However, the correlation decreases for CDR values greater than 15 µm, with the operational CDRs underestimated by 2–4 µm.

The difference associated with CDRs (> 15 µm) between the two approaches can be explained by the absence of measurements from the primary rainbow region (137–145°) in the operational retrieval procedure. According to Fig. 1a, the cloudbow oscillations (137–145°) for CDR > 15 µm are more pronounced than those for smaller CDRs (< 15 µm), and the cloudbow oscillations (145–165°) attenuated faster for larger CDRs. For example, the cloudbow oscillations for CDR = 19 µm are very weak for scattering angles exceeding 155°, although intense oscillations are present in the cloudbow for CDR = 5 µm, even at a scattering angle of 170°. Table 5 presents comparisons between CDRs, estimated using POLDER-like measurements for the 137–165° and 145–165° scattering angle ranges. The retrievals for CDR < 15 µm are accurate in both groups. The CDRs derived from the 145–165° scattering angle range tend to be underestimated. However, in the 137–165° range, the retrievals are more reliable across the entire CDR range. Furthermore, other factors may explain the difference between our retrievals and the operational retrievals, including the criteria used for selecting the best fit results and the number of wavelengths employed in the retrievals. However, we do not believe these differences introduce large biases because we added the Qual factor in the retrieval to preserve the accuracy of our results. The measurements at the wavelengths 490 and 679 nm are not used in our method because Sect. 3.4 demonstrates that the retrieval is accurate when applied at 865 nm.

4.2 Sub-grid-scale retrievals

The above analysis revealed that the proposed algorithm is robust when sufficient observations are obtained in the rainbow region and that heterogeneity in clouds can lead to biased estimates of the CDR and EV distributions. It is necessary to examine the operational retrievals at a higher resolution and compare the results with the operational products. A case study was conducted with a resolution of 7 × 7 pixels for the data from 1 June 2008. We divided the grid (150 km × 150 km) into 36 sub-grids to derive detailed CDR information at the sub-grid scale. In accordance with previous analyses, retrievals can be conducted with approximately 10 observations at the scattering angles of 137–165° at the wavelength of 865 nm. Considering instrumental noise, we applied the retrieval to observations that encompassed the entire rainbow region; the number of polarized reflectances...
Table 5. Comparisons between CDRs estimated using POLDER-like measurements for scattering angle ranges of 137–165° and 145–165°, respectively. The retrievals were applied using 40 measurements at a wavelength of 865 nm. The actual EVs used include 0.01, 0.02, and 0.05.

<table>
<thead>
<tr>
<th>Actual CDR (µm)</th>
<th>CDRs retrieved for angles of 137–165° (µm)</th>
<th>CDRs retrieved for angles of 145–165° (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EV = 0.01</td>
<td>EV = 0.02</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>13.5</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>15.5</td>
<td>15.5</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>18</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

exceeded 15. The grid and sub-grid retrievals at 70.5° W, 172.1° N are shown in Fig. 9a and b. In the operational CDR products, the grid-scale CDR and EV were 10.5 µm and 0.05, respectively. The same CDR was retrieved using our algorithm, although the retrieved EV was 0.1. The different EV value was caused by the omission of the polarized phase functions for EV = 0.1 from the operational procedure. According to the sub-grid-scale retrievals shown in Fig. 9b, 24 effective sub-grid-scale retrievals, with CDRs ranging from 6 to 14.5 µm were derived. The sub-grid-scale average of the CDR and EV values were 10.54 µm and 0.09, which was very close to the grid-scale retrieval result. However, we emphasize that the results presented herein are based on a particular case that is not representative of all POLDER observations. The impact of drizzle in this region is unclear and must be elucidated in future research.

5 Conclusions

In this study, our retrieval algorithm is based on the ideas of Bréon and Goloub (1998) and Bréon and Doutriaux-Boucher (2005), but it includes improvements in the use of measurements of the primary rainbow region (137–145°) to provide more reliable large droplet (> 15 µm) retrievals. The POLDER global L1B data from June 2008 were used to derive the CDRs via the improved method for comparison with the operational CDR products. The CDRs derived using the improved method correlate well with the operational products for the CDRs of 5 to 14 µm; however, for the CDRs of 15 to 19 µm, the operational CDRs are underestimated by 2–4 µm. The retrieval cases using the POLDER-like measurements show similar results. These biases can be explained by the absence of measurements from the primary rainbow region (137–145°), which becomes significant because the cloudbow oscillations are more pronounced for large droplets (CDR > 15 µm) than for smaller droplets.

Based on the modeled POLDER-like polarized reflectances at wavelengths of 490, 670, and 865 nm, the impacts of cloud horizontal inhomogeneity and directional sampling on the retrieval of cloud droplet size by the POLDER instrument were investigated. The sub-grid-scale variability in the CDR reshapes the observed rainbow structures and results in a lot of retrievals being inaccessible. However, the variability in the CDR biases both the CDR and EV estimates, and the associated uncertainties are greater when not including the primary rainbow measurements in the retrieval. However, the sub-grid-scale variability in the EV and COT affects the EV retrievals and does not exert discernable impact on the CDR estimates. Therefore, higher-resolution retrievals provide much more successful droplet size distribution estimates and reduce the biases introduced by the effects of horizontal inhomogeneity in clouds.

To understand whether a resolution is sufficient for POLDER droplet size retrievals, the effects of directional sampling were investigated. The case studies showed that the algorithm is robust when sufficient measurements are provided, and the required number of measurements was found to decrease as the wavelength increased from 490 to 865 nm, which is determined by the cloudbow oscillations at those wavelengths. Furthermore, the retrieval is robust as long as
the random noise is no greater than 10\% of the signal. Finally, a case study demonstrated that the droplet size distribution retrieval can be performed at 7 × 7 pixels. The results suggest that the POLDER size distribution retrieval algorithm can be applied at a high resolution and that substantial uncertainties due to cloud horizontal inhomogeneity exist.

Several questions require further investigation. For example, if the cloud size distribution is inferred at a higher resolution, will the bias between the MODIS and POLDER CDRs still exist? A long series of POLDER CDRs should be derived at high resolutions, and the POLDER and MODIS CDR products should be compared. In addition, the theoretical minimal number of measurements needed for droplet size distribution retrievals should be estimated to help guide instrument design and improvement.

Acknowledgements. This work was supported by the CAS/SAFEA International Partnership Program for Creative Research Teams (grant no. KZZD-EW-TZ-09), the National Science and Technology Ministry (grant no. 2014BAC21B03) and the National Natural Science Foundation of China (grant no. 41471367). The data used in this paper were derived from the CNES/POLDER instrument onboard NASDA/ADEOS. The authors thank the Interactions Clouds Aerosols Radiations Etc (ICARE) thematic center for processing the POLDER data. We are grateful to K. F. Evans at the University of Colorado for the radiative transfer 3 and Lorenz–Mie codes.

Edited by: A. Kokhanovsky

References

Liu, Y., Wu, W., Jensen, M. P., and Toto, T.: Relationship between cloud radiative forcing, cloud fraction and cloud albedo,


