Constraining the Twomey effect from satellite observations: issues and perspectives

Johannes Quaas1, Antti Arola2, Brian Cairns3, Matthew Christensen4, Hartwig Deneke5, Annica M. L. Ekman6, Graham Feingold7, Ann Fridlind3, Edward Gryspeerdt8, Otto Hasekamp9, Zhanqing Li10, Antti Lipponen2, Po-Lun Ma11, Johannes Mülmenstädt1, Athanasios Nenes12,13, Joyce E. Penner14, Daniel Rosenfeld15, Roland Schröder1, Kenneth Sinclair3,16, Odran Sourdeval17, Philip Stier4, Matthias Tesche1, Bastiaan van Diedenhoven3, and Manfred Wendisch1

1Leipzig Institute for Meteorology, Universität Leipzig, Leipzig, Germany
2Finnish Meteorological Institute, Kuopio, Finland
3NASA Goddard Institute for Space Studies, New York, USA
4Department of Physics, University of Oxford, Oxford, UK
5Leibniz Institute for Tropospheric Research, Leipzig, Germany
6Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
7NOAA Earth System Laboratories, Chemical Science Laboratory, Boulder, USA
8Space and Atmospheric Physics Group, Imperial College London, UK
9SRON Netherlands Institute for Space Research, Utrecht, the Netherlands
10Earth System Science Interdisciplinary Center and Department of Atmospheric and Oceanic Science, University of Maryland, College Park, USA
11Pacific Northwest National Laboratory, Richland, USA
12School of Architecture, Civil & Environmental Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
13Institute of Chemical Engineering Sciences, Foundation for Research and Technology Hellas, Patras, Greece
14Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, USA
15Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel
16Department of Earth and Environmental Engineering, Universities Space Research Association (USRA), Columbia, MD 21046, USA
17Université de Lille, CNRS, UMR 8518 – LOA – Laboratoire d’Optique Atmosphérique, Lille, France

Correspondence: Johannes Quaas (johannes.quaas@uni-leipzig.de)

Received: 24 March 2020 – Discussion started: 15 May 2020
Revised: 24 September 2020 – Accepted: 8 October 2020 – Published: 4 December 2020

Abstract. The Twomey effect describes the radiative forcing associated with a change in cloud albedo due to an increase in anthropogenic aerosol emissions. It is driven by the perturbation in cloud droplet number concentration (\(\Delta N_{d,\text{ant}}\)) in liquid-water clouds and is currently understood to exert a cooling effect on climate. The Twomey effect is the key driver in the effective radiative forcing due to aerosol–cloud interactions, but rapid adjustments also contribute. These adjustments are essentially the responses of cloud fraction and liquid water path to \(\Delta N_{d,\text{ant}}\) and thus scale approximately with it. While the fundamental physics of the influence of added aerosol particles on the droplet concentration (\(N_d\)) is well described by established theory at the particle scale (micrometres), how this relationship is expressed at the large-scale (hundreds of kilometres) perturbation, \(\Delta N_{d,\text{ant}}\), remains uncertain. The discrepancy between process understanding at particle scale and insufficient quantification at the climate-relevant large scale is caused by co-variability of aerosol particles and updraught velocity and by droplet sink processes. These operate at scales on the order of tens of me-
tress at which only localised observations are available and at which no approach yet exists to quantify the anthropogenic perturbation. Different atmospheric models suggest diverse magnitudes of the Twomey effect even when applying the same anthropogenic aerosol emission perturbation. Thus, observational data are needed to quantify and constrain the Twomey effect. At the global scale, this means satellite data. There are four key uncertainties in determining $\Delta N_d, \text{ant}$, namely the quantification of (i) the cloud-active aerosol – the cloud condensation nuclei (CCN) concentrations at or above cloud base, (ii) $N_d$, (iii) the statistical approach for inferring the sensitivity of $N_d$ to aerosol particles from the satellite data and (iv) uncertainty in the anthropogenic perturbation to CCN concentrations, which is not easily accessible from observational data. This review discusses deficiencies of current approaches for the different aspects of the problem and proposes several ways forward: in terms of CCN, retrievals of optical quantities such as aerosol optical depth suffer from a lack of vertical resolution, size and hygroscopicity information, non-direct relation to the concentration of aerosols, difficulty to quantify it within or below clouds, and the problem of insufficient sensitivity at low concentrations, in addition to retrieval errors. A future path forward can include utilising co-located polarimeter and lidar instruments, ideally including high-spectral-resolution lidar capability at two wavelengths to maximise vertically resolved size distribution information content. In terms of $N_d$, a key problem is the lack of operational retrievals of this quantity and the inaccuracy of the retrieval especially in broken-cloud regimes. As for the $N_d$-to-CCN sensitivity, key issues are the updraught distributions and the role of $N_d$ sink processes, for which empirical assessments for specific cloud regimes are currently the best solutions. These considerations point to the conclusion that past studies using existing approaches have likely underestimated the true sensitivity and, thus, the radiative forcing due to the Twomey effect.

1 Introduction

Cloud droplets in liquid-water clouds form on cloud condensation nuclei (Aitken, 1880), a subset of the atmospheric aerosol particle population. The formation of cloud droplets in thermodynamic equilibrium is established textbook knowledge (Köhler, 1936). Whether an aerosol particle acts as cloud condensation nuclei (CCN) at a given supersaturation depends on its size and chemical composition, which determine the particle hygroscopicity (Dusek et al., 2006; Ma et al., 2013). If CCN concentrations at one supersaturation level are known, CCN concentrations at other supersaturation levels approximately scale with it according to the Twomey (1959) parameterisation. Here, we implicitly consider a supersaturation level of 0.2% unless otherwise stated. Supersaturation is generated in the large majority of clouds by updraughts. The rare exceptions are formation due to radiative cooling (mainly fog events) or the mixing of cold and dry with warm and moist air masses. Cloud-scale updraughts originate in most cases from turbulence, convection or gravity waves. Updraught velocity, $w$, exhibits a large heterogeneity across temporal and spatial scales (Tontila et al., 2011; Moeng and Arakawa, 2012). For a given probability density function (PDF) of updraughts, in an adiabatic air parcel with no active collision and coalescence, the addition of extra CCN will generally lead to a monotonic increase in cloud droplet number concentration, $N_d$ (Twomey and Warner, 1967). The approximate functional form of the dependence of $N_d$ on CCN concentration is then logarithmic, since the increase in $N_d$ associated with activation of additional aerosol leads to a depletion of the maximum supersaturation (Twomey, 1959).

The CCN concentration in the atmosphere is increased by anthropogenic emission of aerosols and aerosol precursor gases (Boucher et al., 2013). This leads to enhanced $N_d$, unless aerosol particle concentrations are high and updraughts weak (Ghan et al., 1998; Feingold et al., 2001; Reutter et al., 2009). In turn, cloud albedo ($\alpha_c$, the fraction of solar radiative energy reflected back to space by clouds in relation to that incident at the cloud top) increases, as it is a monotonically increasing function of $N_d$. Following Platnick and Twomey (1994) and Ackerman et al. (2000),

$$\left. \frac{\partial \alpha_c}{\partial \ln N_d} \right|_{\text{ant}} = \frac{1}{3} \alpha_c (1 - \alpha_c),$$

a formulation which relies on (i) a two-stream radiative transfer approximation and (ii) the assumption that clouds obey vertical stratification that scales with an adiabatic one and that is horizontally homogeneous. Equation (1) is expressed as a partial derivative: other quantities – notably cloud water path – are considered constant.

These two facts – $N_d$ is a monotonic function of CCN and $\alpha_c$ in the partial-derivative sense is a monotonic function of $N_d$ – imply that the anthropogenic increase in CCN concentrations causes a negative (cooling) radiative forcing due to aerosol–cloud interactions, $RF_{\text{aci}}$ (Boucher et al., 2013), denoted as $F_{\text{aci}}$ (Bellouin et al., 2020b). It can be approximately (neglecting absorption in the column above the cloud after scattering at cloud top) written as (Quaas et al., 2008; Bellouin et al., 2020b)

$$F_{\text{aci}} = F_s^\uparrow \left|_{\text{ant}} \cdot \frac{\partial \alpha_c}{\partial \ln N_d} \cdot \frac{\partial \ln N_d}{\partial \ln a} \cdot \Delta \ln a_{\text{ant}}. \right.$$

with the downward solar radiative flux density (irradiance) above clouds, $F_s^\uparrow$, and a quantitative description of CCN denoted here as $a$. The relative anthropogenic perturbation to $a$ is denoted $\Delta \ln a_{\text{ant}}$. This formulation assumes (i) that only the solar spectrum is relevant, which is well justified for the optically thick, liquid water clouds considered here, since an $N_d$ perturbation only marginally changes the cloud radiative
effect in the terrestrial spectrum of an optically thick cloud and (ii) that there is one liquid water cloud layer that determines the effect so that the problem can be considered purely horizontal in space. In contrast to the formulation by Bellouin et al. (2020b), we consider the problem as horizontally variable in space \((x, y)\) and in time \(t\), i.e. \(R_{\text{f,act}} = R_{\text{f,act}}(x, y, t)\). If Eq. (2) is assessed from temporally sparse satellite data, a proper integration over temporally varying solar zenith angles and cloud diurnal cycles is necessary.

\(R_{\text{f,act}}\) is often referred to as the “Twomey effect” (Twomey, 1974) and also called the “(first) aerosol indirect effect” or “cloud albedo effect” (Lohmann and Feichter, 2001). Atmospheric models simulate a large range for \(R_{\text{f,act}}\) (Gryspeerdt et al., 2020; Smith et al., 2020). It is, thus, necessary to constrain the Twomey effect quantitatively based on observations. Only satellites can provide global observational data that could be used to quantify the global \(R_{\text{f,act}}\) (Stephens et al., 2019).

The Twomey effect has been assessed in many studies (starting with Bréon et al., 2002) in terms of cloud droplet effective radius, \(r_e\), rather than using \(N_d\). This is plausible as, for idealised vertical profiles of droplet size distributions (e.g. vertically constant or adiabatically increasing profiles), cloud optical depth and cloud albedo are easily expressed as, for idealised vertical profiles of droplet size distributions (Stephens et al., 2019).

Among the four factors on the right-hand side of Eq. (2), the first one, \(F_1\), is well quantified for each given latitude, longitude and time. The second one, \(\partial \alpha_c \partial \ln N_d\), can be evaluated using Eq. (1) (Bellouin et al., 2020b; Hasekamp et al., 2019a), or alternatively by radiative-transfer simulations (Mülmenstädt et al., 2019). This implies that the two key problems in determining \(R_{\text{f,act}}\) are the quantification of the anthropogenic perturbation of CCN, \(\Delta \ln \alpha_{\text{ant}}\), and the sensitivity of \(N_d\) to CCN perturbations, \(\beta = \partial N_d/\partial \ln a\) (Feingold et al., 2001). Taken together, this is the distribution of the anthropogenic perturbation of \(N_d\) (here expressed in absolute, not relative, terms):

\[
\Delta N_{d,\text{ant}} = \left. \frac{\partial N_d}{\partial \ln \alpha} \right|_{w} \cdot \Delta \ln \alpha_{\text{ant}} = \beta(w) \cdot \Delta \ln \alpha_{\text{ant}}. \tag{3}
\]

The plausible range of the sensitivity is \(0 \leq \beta \leq 1\), except for heavily polluted situations (where it may become negative; Feingold et al., 2001), or when giant CCN play an important role (Ghan et al., 1998; Morales Betancourt and Nenes, 2014; Gryspeerdt et al., 2016; McCoy et al., 2017) where competition for water vapour during droplet formation is at its strongest. Such conditions represent a significant challenge to models and parameterisations of the process (Morales Betancourt and Nenes, 2014).

The aerosol forcing has to be evaluated at a scale much larger than an individual cloud. One of the key reasons for this is that there is currently no way to use satellite data to determine the anthropogenic fraction of the CCN population for a single air parcel. Methods applying model information, or data-tied approaches such as Bellouin et al. (2013) instead use the scale of model resolution or aggregate data resolution which is typically of the order of \(1^\circ \times 1^\circ\) (or about 100 \(\times\) 100 km\(^2\)). The problem formulated in Eq. (3) then has to be reformulated, using an overbar to denote the averaging over a \(1^\circ \times 1^\circ\) grid box as

\[
\bar{\Delta N_{d,\text{ant}}} = \overline{\int_{\ln a_{\text{ant}}}^{\infty} \frac{\partial N_d}{\partial \ln \alpha} \cdot \mathcal{P}(w) \mathcal{P}(a) \, dw} \cdot \Delta \ln \alpha_{\text{ant}} = \bar{\beta} \cdot \Delta \ln \alpha_{\text{ant}}, \tag{4}
\]

which considers the mean sensitivity of \(N_d\) to CCN, \(\bar{\beta}\), given the probability density function (PDF) of cloud base updraught velocity, \(w\) in the grid box, \(\mathcal{P}(w)\); the PDF of CCN at cloud base within the scene, \(\mathcal{P}(a)\); and the anthropogenic perturbation of the CCN concentration at the grid-box scale, \(\Delta \ln \alpha_{\text{ant}}\). Note in the above equation, \(\bar{\beta}\) is assumed independent of \(\ln \alpha_{\text{ant}}\), which assumes that \(\mathcal{P}(w)\) is independent of cloud properties (primarily, liquid water content), which applies to stratus clouds (Morales and Nenes, 2010) but not in general. Similarly, the covariance of \(\mathcal{P}(w)\) and \(\mathcal{P}(a)\) may not be zero (e.g. Kacarab et al., 2020 – in addition to Bougiatioti et al., 2020). All of the above suggest that observation of \(\bar{\beta}\) at a cloud parcel scale is not directly transferrable to the large scale for an assessment of the Twomey effect. Rather, \(\bar{\beta}\) has to be estimated.

Beyond \(R_{\text{f,act}}\), aerosol–cloud interactions also lead to rapid adjustments: once cloud droplet size distributions are altered due to anthropogenic CCN, cloud microphysical and dynamical processes are modified as well (Albrecht, 1989; Ackerman et al., 2000; Wang et al., 2003; Heyn et al., 2017; Mülmenstädt and Feingold, 2018). Aerosols can induce transitions between cloud regimes, for instance by changing drizzle behaviour (Rosenfeld et al., 2006; Feingold et al., 2010; Wood et al., 2011). The direction and magnitude of these changes depends on the cloud state and regime, because responses to aerosol changes occur due to processes spanning a range from microphysics to the mesoscale (Christensen and Stephens, 2012; Kazil et al., 2011; Wang et al., 2011). These processes include precipitation suppression (Albrecht, 1989),

https://doi.org/10.5194/acp-20-15079-2020

Atmos. Chem. Phys., 20, 15079–15099, 2020
rapid feedbacks involving cloud-top entrainment (Ackerman et al., 2004; Bretherton et al., 2007; Hill et al., 2009; Balkovic et al., 2019) and rapid feedbacks involving cloud lateral entrainment (Xue and Feingold, 2006; Small et al., 2009) as well as responses in dynamics (Xue et al., 2008; Stevens and Feingold, 2009; Wang and Feingold, 2009). If one also considers deep clouds, further intricate cloud adjustments may occur that are not considered here (e.g. Ekman et al., 2011; Fan et al., 2013; Yan et al., 2014). As a result of these adjustment processes, cloud horizontal extent (Gryspeerdt et al., 2016) and liquid water path (Gryspeerdt et al., 2019) respond to perturbations in \( N_d \). The sum of RF_{aci} and the radiative effects of these adjustments is the effective radiative forcing due to aerosol–cloud interactions, ERF_{aci} (Boucher et al., 2013). Based on modelling and data analysis, it is evident that the adjustments and, thus, also ERF_{aci} scale with \( \Delta N_{d, \text{ant}} \) (Bellouin et al., 2020b; Gryspeerdt et al., 2020; Mülmenstädt et al., 2019). Analysis of model data shows that the rapid adjustments due to other contributions (small-scale to mesoscale circulation changes, thermodynamic changes) are small (Heyn et al., 2017; Mülmenstädt et al., 2019). Even so, thermodynamic and dynamic adjustments to aerosol changes can still have an important impact on droplet formation – especially under conditions where droplet formation is largely velocity-limited (Kacarab et al., 2020; Bougiatioti et al., 2020).

Despite the fact that the activation of an individual CCN to form a droplet is well understood in thermodynamic equilibrium (Köhler, 1936), it is not clear how \( N_d \) responds to perturbations of CCN at the scale of a cloudy air parcel, an entire cloud or of a cloud field up to the large scale of the order of \( 1^\circ \times 1^\circ \) as used in Eq. (4). A one-to-one relationship between CCN in the updraught below cumulus and \( N_d \) above the cloud base within the cumulus has been observed (Werner et al., 2014); although even at the cloud updraught scale, this relationship could be a convolution of the effect of CCN on droplet number, vertical velocity variability and lateral entrainment (Morales et al., 2011). At a larger scale, this relation is less pronounced (Boucher and Lohmann, 1995), consistent with the expectation from Eq. (4). In turn, there may be co-variability of updraughts and aerosol concentrations that lead to larger \( \tilde{\beta} \) compared to situations with constant \( w \) (Kacarab et al., 2020; Bougiatioti et al., 2017, 2020).

Ground-based remote-sensing methods provide data to infer the sensitivity term \( \beta \) from long-term observations (Feingold et al., 2003; McComiskey et al., 2009; Schmidt et al., 2015; Liu and Li, 2018). However, this approach is limited to individual sites and cloud regimes. In consequence, when investigating the global radiative forcing relevant for climate studies, the sensitivity term necessarily is derived from satellite remote sensing (Nakajima and Schulz, 2009).

This leads to a number of problems and challenges discussed in more detail in the following sections.

- **Retrieval of CCN.** The first issue is the missing coincidence of cloud and aerosol retrievals. Usually, no aerosol is retrieved below or within clouds. It is thus questionable how representative aerosol in cloudless scenes is for (neighbouring) cloud base CCN. The second issue is the imperfect nature of proxies for CCN. Often the aerosol optical depth (AOD; see below) or a variant thereof is used, which can only imperfectly be related to CCN due to differences in sensitivity and the lack of vertical resolution.

- **Retrieval of \( N_d \).** There are (i) retrieval errors and biases in \( N_d \), which depend on cloud regimes, and (ii) one needs to consider the link between \( N_d \) as formed by CCN activation at cloud base and the retrieved cloud-top \( N_d \). Cloud-top \( N_d \) (\( N_{d, \text{top}} \)) is the one that determines the scattering of sunlight and, thus, is relevant for the top-of-atmosphere cloud radiative effect. It differs from cloud base \( N_d \) (\( N_{d, \text{base}} \)) in conditions where \( N_d \) sinks such as precipitation or mixing play a role. When using \( r_d \) rather than \( N_d \) the additional problem of stratification by retrieved \( L \) arises.

- **Cloud-regime dependence.** Cloud base droplet concentration, \( N_{d, \text{base}} \), is a function of both CCN and updraught, and \( N_{d, \text{top}} \) is further a function of \( N_d \) sinks such as precipitation formation and entrainment mixing. Thus, one needs to understand how the characteristics of \( w \) and its PDF, as well as precipitation and mixing processes, depend on cloud regime and how this may be used for an empirical estimation of \( \tilde{\beta} \).

- **Aggregation scale.** The relation of aggregate quantities is not the same as the aggregate relation, and, thus, one needs to determine how to derive \( \tilde{\beta} \) optimally from remote-sensing data (Grandey and Stier, 2010; McComiskey and Feingold, 2012).

In practical terms, one further needs to assess to which extent a simple scalar sensitivity metric is sufficient, or whether a joint-PDF approach is preferable (McComiskey and Feingold, 2012; Gryspeerdt et al., 2017).

Beyond these questions which are discussed in the following sections, it is necessary to quantify the anthropogenic perturbation to CCN, \( \Delta \ln \tau_{\text{aunt}} \), which is not easily quantified from observations. The key problem is that there is little potential to observe an atmosphere unperturbed by anthropogenic emissions (Carslaw et al., 2013, 2017). Some studies attempt to quantify the anthropogenic perturbation to the column aerosol light extinction, or aerosol optical depth (AOD; \( \tau_d \)), in a data-tied approach (Kaufman et al., 2005; Bellouin et al., 2005, 2013; Kinne, 2019). Such approaches rely on simplifying parameterisations, such as the assumption that small-mode aerosol particles are predominantly anthropogenic. The other option is to estimate it from simulations (Quaas et al., 2009b; Gryspeerdt et al., 2017). There are
some indirect ways to infer the anthropogenic impacts on $N_d$ (Quaas, 2015), such as from trends (Krüger and Graßl, 2002; Bennartz et al., 2011) or periodicity in anthropogenic emissions such as the weekly cycle (Quaas et al., 2009a). Hence, models are involved in determining an anthropogenic perturbation of CCN concentrations, which can even be attempted for individual weather events (Schwartz et al., 2002). In any case, it seems impossible to know the anthropogenic perturbation to the aerosol at the scale of an air parcel; rather it is possible only at larger, aggregate scales. The remainder of this review will focus on the sensitivity term $\bar{\beta}$.

2 Remote sensing of CCN concentrations

The aerosol quantity most accessible to passive satellite remote sensing is AOD (Kaufman et al., 2002). It is derived from the multi-spectral reflectance of the Earth–atmosphere system using the incident solar radiation and retrieving or assuming surface albedo characteristics as well as aerosol absorption coefficient and scattering phase functions. There are four key issues with using the retrieved AOD for estimating the $N_d$ to CCN sensitivity, which will be discussed in the following subsections.

- AOD is the vertical integral of the extinction coefficient. For the sensitivity of $N_d$ to the aerosol, one needs to know the vertical distribution of the CCN concentration, most importantly the CCN at cloud base.

- AOD is an optical integral and does not provide information on the aerosol size distribution and its hygroscopicity. The use of AOD does not isolate aerosol particles that have the size and chemical composition to serve as CCN. It is also affected by aerosol swelling due to hygroscopic growth.

- AOD can be derived only for pixels determined to be cloud-free. The degree to which this correlates with the CCN at the base of (neighbouring) clouds is questionable. In addition, retrieved AOD can show a positive bias due to enhanced reflectance from neighbouring cloudy pixels or due to the lack of detecting spurious clouds in a retrieval scene.

- The optical signal is very weak at low concentrations. Therefore, retrievals become more and more uncertain below a certain aerosol load, especially over land and in situations with variable or uncertain surface albedo.

At aggregate scales, i.e. for monthly averages over regions, AOD from ground-based remote-sensing retrievals (AERONET; Holben et al., 2001) correlates well with CCN surface measurements (Andreae, 2009; Shen et al., 2019). Similar results were also reported for aircraft measurements (Clarke and Kapustin, 2010; Shinozuka et al., 2015). However, at shorter timescales or less spatial aggregation, there are significant deviations from a perfect correlation (Liu and Li, 2014). AOD due to aerosol light extinction is determined by the vertical integral of the extinction cross section, proportional to the vertical integral of the second moment of the aerosol size distribution. In turn, for a given chemical composition of aerosol particles, the CCN concentration is the zeroth moment of the size distribution for particles exceeding a size threshold that depends on supersaturation. In the following, the different problems are discussed in more detail, together with options for a better proxy for CCN from satellite remote sensing.

2.1 Vertical co-location

Stier (2016) investigated the correlation between AOD and CCN as represented in a climate model. He confirmed a mostly positive correlation of the temporal variability of the two quantities, although in some regions the correlation is low or even negative. A key reason for the partly low correlation is the fact that AOD is a vertically integrated quantity and may include aerosol layers that are not interacting with clouds. A similar result was reported from a statistical analysis of satellite data: cloud microphysical parameters correlate well with aerosol properties only if the vertical alignment of the aerosol and cloud layers is accounted for (Costantino and Bréon, 2010, 2013). More recently, Painemal et al. (2020) demonstrate a much higher correlation between $N_d$ and aerosol extinction coefficients below cloud top sampled from satellite lidar compared to $N_d$ vs. AOD. Ship measurements of CCN and microwave-retrieved $N_d$ at cloud base between Los Angeles and Hawaii show a weaker $\beta$ metric as the boundary layer deepens, thus indicating that surface aerosol measurements become less representative for aerosol variability at cloud base as the boundary layer deepens (Painemal et al., 2017), or that the updraughts become high enough to activate smaller aerosols than the accumulation mode. In situ observations suggest that AOD may even be anticorrelated with CCN at cloud base (Kacarab et al., 2020).

A way forward is the use of spaceborne vertically resolved observations such as lidar measurements (Shinozuka et al., 2015; Stier, 2016). The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO; Winker et al., 2009) lidar retrieves aerosol backscatter profiles and thus is capable of identifying aerosol layers (Costantino and Bréon, 2010). Profiles of aerosol particle extinction are inferred from these backscatter profiles by using typical extinction-to-backscatter ratios based on aerosol type. However, the signal is not sensitive to smaller aerosol concentrations, which hampers a quantitative analysis at the large scale (Watson-Parris et al., 2018; Ma et al., 2018). For situations with sufficient aerosol loading for reliable CALIPSO aerosol profile observations, methods for retrieving CCN concentrations from ground-based lidar measurements can be adapted (Feingold and Grund, 1994; Lv et al., 2018; Haarig et al., 2019). These
methods apply empirical extinction-to-particle-concentration relationships to obtain input for CCN concentrations for different aerosol types (Mamouri and Ansmann, 2016). In the future, the EarthCARE satellite mission currently scheduled for launch in 2022 (Illingworth et al., 2015; Hélière et al., 2017) shows promise to extend and improve upon the success of the CALIPSO mission. Its Atmospheric Lidar (ATLID) is a linearly polarised high-spectral-resolution lidar (HSRL) operating at a wavelength of 355 nm. The instrument allows the direct inference of profiles of aerosol backscatter and extinction coefficients, thereby substantially increasing the retrieval accuracy. The direct retrieval of the extinction-to-backscatter (lidar) ratio (Müller et al., 2007) with ATLID (compared to the use of pre-set values in the CALIPSO retrieval: Kim et al., 2018) and the large difference between lidar ratios of aerosols (20–80 sr) and clouds (20–30 sr) are also expected to provide better distinction between optically thin cirrus clouds and aerosols than CALIPSO (Reverdy et al., 2015). While a similar sensitivity to aerosol load is expected for ATLID and CALIOP observations during nighttime, ATLID promises a better daytime sensitivity. EarthCARE is also expected to provide better distinction between optically thin clouds and aerosols than CALIPSO (Reverdy et al., 2015). Airborne measurements have shown that further utilising HSRL at more than one wavelength (extending beyond ATLID) would provide substantial additional information content for retrieving vertically resolved aerosol parameters, especially when combined with polarimeter measurements (Burton et al., 2016). From the passive-remote-sensing perspective, promising results have been obtained for retrievals of aerosol vertical information from near-ultraviolet polarimetry (Wu et al., 2016), although the quality degrades for small aerosol concentrations. Passive observations with high spectral resolution within the oxygen A absorption band around 760 nm can also be used to infer aerosol layer height (Hollstein and Fischer, 2014; Geddes and Bösch, 2015).

In particular, an operational aerosol layer height product is now available from the Tropospheric Monitoring Instrument (TROPOMI) flown on the Sentinel-5p mission (Sanders et al., 2015). Also, a recent study presents promising results based on Orbiting Carbon Observatory 2 (OCO-2) observations (Zeng et al., 2020). In particular, a combination of such approaches, e.g. passive polarimetry and active lidar observations (Stamnes et al., 2018) or multi-angle polarimetry and oxygen A band observations as planned for NASA’s Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission (Remer et al., 2019), shows potential. Retrievals could also combine observations and model adjoints to constrain below-cloud aerosol number, which is directly relevant for aerosol–cloud interactions (Saïde et al., 2012).

In summary, the lack of vertical co-location between retrieved CCN proxy and clouds leads to an underestimate in \( N_d \)–CCN sensitivity (Costantino and Bréon, 2010). Model studies suggest that this bias may be approximately cancelled by a corresponding bias in the anthropogenic component of the cloud base CCN (Gryspeerdt et al., 2017). However, the extent of this cancellation in current observational studies is unknown and requires further investigation. For an accurate estimation of \( \beta \), the use of lidar retrievals seems to be the best way forward, while additional information on the vertical distribution of aerosol can also be gained from present and upcoming passive satellite instruments.

### 2.2 Horizontal co-location

In studies examining \( \beta \) from satellite data, spatial aggregates are considered (i.e. \( \beta \) as in Eq. 4), in which the aerosol retrievals in the cloud-free pixels are averaged at a coarse resolution (such as 1°) and taken to define the relation with \( N_d \) retrievals in the same grid box (Quaas et al., 2008). This assumes that the aerosol population is horizontally homogeneous at such large scales. According to Anderson et al. (2003), this is often the case. It has been confirmed from aircraft data for the stratocumulus cases investigated by Shinozuka et al. (2020). However, CCN is consumed when droplets activate, and aerosol is scavenged when clouds precipitate. Hence, the assumption of aerosol concentration horizontal homogeneity is questionable, at least in precipitating clouds.

It is the aerosol in air masses before cloud particles form that is relevant to compute the aerosol impact on \( N_d \) (Gryspeerdt et al., 2015). In one of the early aerosol–cloud interaction studies from satellite data Bréon et al. (2002) used trajectories to identify cloudless situations in which aerosol retrievals were possible for air masses that later formed clouds. This is a promising solution but it requires much more effort than the simpler co-location assumptions. It also requires reliable, high-resolution information about atmospheric trajectories. Another complication is that the formation rate of secondary aerosol is enhanced by aqueous phase reactions, potentially enhancing aerosol concentrations in the vicinity of clouds (Jeong and Li, 2010). Such trajectory approaches are particularly useful when they exploit the high temporal resolution that is available from geostationary satellites. Aerosol retrievals from geostationary satellites may be combined using trajectory modelling to link these to clouds that form in these air masses (Kikuchi et al., 2018), or also the aerosol retrieval from a polar orbiter could be related to clouds retrieved from geostationary satellites that form in the same air masses (Christensen et al., 2020).

Altogether, the lack of horizontal co-location may imply somewhat too low \( \beta \) due to the potential de-correlation of CCN concentrations and \( N_d \) in situations with spatially heterogeneous aerosol. The consideration of backward trajectory analysis seems the best option to address the issue since there is no solution yet to retrieve aerosols below or within clouds from satellite.
2.3 Hygroscopic growth of aerosol particles

The extinction of solar radiation by aerosol particles is a strong function of the hygroscopic growth of the particles. Haze particles attenuate much more sunlight compared to the same aerosol particle ensemble in dry conditions. AOD is thus heavily influenced by the variability of relative humidity. The light extinction caused by dry particles (at relative humidities below 30%) is much better correlated to CCN concentrations than the extinction of particles at ambient relative humidity (Shinozuka et al., 2015). Liu and Li (2018) showed that using total AOD compared to dry AOD as a CCN proxy when estimating $\beta$ from measurements at different Atmospheric Radiation Measurements (ARM) sites resulted in a 23% underestimate. A way forward is to apply parameterisations in terms of retrievals of relative humidity to account for the aerosol swelling. These need information about aerosol hygroscopicity and relative humidity at the appropriate scale. Hygroscopicity information could rely on the kappa-Köhler parameterisation approach (Petters and Kreidenweis, 2007; Pringle et al., 2010), and a parameterisation of small-scale to mesoscale humidity variability could make use of approaches exploited in general circulation models (GCMs) (Quaas, 2012; Petersik et al., 2018). Another alternative would be to retrieve the amount of aerosol water, making use of the real part of the refractive index (Schuster et al., 2009). This would allow the translation of the size distribution of humidified aerosol particles to the corresponding dry size distribution. In the near future, accurate refractive index retrievals are expected from polarimeters such as the SPEXone instrument on the NASA PACE mission (Hasekamp et al., 2019b; Werdell et al., 2019), to be launched in 2022.

Summarising, using AOD as a proxy for CCN results in low-biased estimates of $\beta$ due to aerosol swelling. Approaches to parameterise the dry aerosol properties on the basis of the humidified one can help alleviate the problem.

2.4 Approaches using aerosol index, column-CCN, reanalysis or cloud base updraught

The aerosol index (AI$^1$) is defined as the product of AOD and the Ångström exponent (Deuzé et al., 2001). This latter quantity is the slope of the spectral variation in AOD and is typically larger for smaller particles (Ångström, 1929). AI is more weighted towards smaller particles, which makes it better suited as a proxy for CCN concentration at typical supersaturations than AOD. For log-normal size distributions, AI is approximately proportional to the column aerosol number concentration (Nakajima et al., 2001). Studies using models concluded that AI is a better predictor for CCN (Stier, 2016) and that AI–$N_d$ relationships are better suited to predict $\Delta N_{d,\text{ant}}$ than AOD–$N_d$ relationships (Penner et al., 2011; Grysspeerdt et al., 2017). However, retrievals of the Ångström exponent, and thus of AI, over land are not reported in operational products such as the MODIS dark target algorithm and are in general not as reliable as they are over ocean (Lee and Chung, 2013; Sayer et al., 2013).

Further refining this idea, Hasekamp et al. (2019a) aimed to retrieve the column CCN concentrations over oceans. The analysis of polarimetric observations allowed us to account for some aspects of the aerosol particle size distribution, and for particle sphericity, which is related to particle hygroscopicity. This column-CCN retrieval implied larger $\beta$, increasing the resulting RF$_{\text{aci}}$ by almost 50%. It is an example of how additional information from polarimetry is useful for studying the CCN-to-$N_d$ relationship.

However, neither the approach of Hasekamp et al. (2019a) nor the use of AI overcomes the problem of lack of horizontal and vertical coincidence of CCN and $N_d$ retrievals. An option to overcome this problem is to make use of additional model information. Satellite-retrieved AOD is assimilated into aerosol models, e.g. in the Copernicus Atmosphere Monitoring Service (CAMS, Benedetti et al., 2009; Inness et al., 2019) or the Modern-Era Retrospective Analysis for Research and Applications (version 2; MERRA-2 Gelaro et al., 2017). The model predictions are applied to obtain aerosol information beneath clouds. Such aerosol reanalysis information has been used for assessing RF$_{\text{aci}}$ in several studies (Bellouin et al., 2013; McCoy et al., 2017; Bellouin et al., 2020a). However, assessing the validity of model results requires extensive and rigorous evaluation, especially for coarsely resolved models with regard to aerosol scavenging below clouds. For this, independent data are required such as from ground-based observations or satellite observations from sensors other than those that are assimilated.

Yet another solution initially proposed by Feingold et al. (1998) and applied to satellite retrievals by Rosenfeld et al. (2016) is to parameterise the cloud base updraught, $w$, on the basis of cloud retrievals, rather than to retrieve the aerosol. For convective clouds, Zheng et al. (2015) suggested that $w$ scales with cloud base altitude, which can be retrieved from satellites. For stratocumulus clouds, Zheng et al. (2016) proposed that updraught is a function of cloud-top radiative cooling, and that this can be computed by radiative transfer modelling on the basis of cloud quantities retrieved from passive sensors and thermodynamic profiles from meteorological re-analyses. The retrieved profiles of $w_c$ together with derivations of supersaturation as a function of $w$ and $N_d$ (Rosenfeld et al., 2016) then allow the parameterisation of the CCN concentration at any given supersaturation. This approach does not suffer from the problem of a lower detection limit. However, it has not yet been used to quantify the Twomey effect.
Concluding, all four approaches alleviate many problems encountered when using AOD. An ideal solution may be the combination of several of these by also assimilating, in addition to AOD, polarimetric satellite observations, as well as lidar measurements, into the analysis of the atmospheric state in high-resolution models.

3 Remote sensing of cloud droplet concentrations

The problem of the remotely sensed $N_d$ as used to estimate $\beta$ has three different facets to it, which will be discussed in this section, namely the following.

- Consideration of $r_e$ rather than $N_d$ in aerosol–cloud interaction studies. In many studies, the droplet effective radius, $r_e$, is used, and the datasets are stratified with respect to $L$ in order to estimate $\beta$. This is very difficult to perform adequately and leads to biases.

- Biases in the retrieved $N_d$. For the assessment of sensitivity, systematic (rather than random) errors in retrieved $N_d$ are relevant. Also, $N_d$ is not retrieved in standard operational procedures, so that inconsistencies between the retrieval of standard components and in the computation of $N_d$ on the basis of retrievals can lead to additional errors.

- Relationship of $N_d$ formed at activation with retrieved and radiation-relevant $N_{d,\text{top}}$. Retrieved $N_{d,\text{top}}$ refers to the drop concentration within the top one to two optical depths of the clouds, and it is $N_{d,\text{top}}$ that is relevant for determining the cloud radiative effect. $N_d$ sink processes such as coagulation imply that $N_{d,\text{top}}$ is smaller than the one resulting from activation above cloud base, $N_{d,\text{base}}$.

$N_d$ is vertically constant for single-layer, purely liquid-water clouds with (i) a vertically homogeneous droplet size spectrum, (ii) for adiabatically stratified clouds or (iii) for sub-adiabatic clouds in which mixing is homogeneous. However, in many situations, precipitation formation or entrainment can lead to reduction of $N_d$ above cloud base. In such situations, it is $N_{d,\text{top}}$ that is relevant to determine the cloud radiative effect (cloud albedo in Eq. 2). Building on Eq. (4) thus gives

$$\Delta N_{d,\text{top}} = \frac{dN_{d,\text{top}}}{dN_{d,\text{base}}} \int_{\eta=-\infty}^{\infty} \frac{\partial N_{d,\text{base}}}{\partial \ln \sigma_w} \mathcal{P}(w) \mathcal{P}(a) dw$$

$$\cdot \Delta \ln \sigma_w = \beta \cdot \Delta \ln \sigma_{\text{e,ant}}.$$

When estimating $\beta$ as a regression coefficient from, for example, satellite-retrieved $N_d$ and a proxy for CCN such as AOD, it is thus this $\beta$ that is inferred.

3.1 Considering $r_e$ rather than $N_d$

Many past studies have used operationally retrieved $r_e$ rather than $N_d$ in aerosol–cloud interaction studies. However, $r_e$ is a function of both $N_d$ and $L$. This introduces the requirement for stratifying the data with respect to $L$ in order to estimate $\hat{\beta}$. To further complicate matters, $N_d$ and $L$ have been found to be correlated (e.g. Michibata et al., 2016; Gryspeerdt et al., 2019). A precise estimation of $\hat{\beta}$ is thus only possible for a large amount of data combined with suitable binning by $L$. Errors in this approach that are related to a lack of data increase at aggregated scales (McCormiskey and Feingold, 2012). Using derived $N_d$ is therefore preferable to avoid unnecessary complications.

3.2 Biases in the $N_d$ retrieval

Satellite retrievals of $N_d$ were extensively reviewed by Grosvenor et al. (2018). Since $N_d$ currently is not retrieved by operational algorithms and new developments to retrieve $N_d$ (e.g. from polarimetry) are still in their infancy, the most frequently used method is to infer $N_d$ from retrieved $r_e$ and cloud optical depth, $\tau_c$, using the relationship

$$N_d = \gamma \cdot \tau_c^{\frac{1}{2}} \cdot r_e^{-\frac{2}{3}},$$

where $\gamma \approx 1.37 \times 10^{-5} \text{m}^{-0.5}$ is a parameter provided as a constant here but more realistically depending on cloud base temperature and pressure, the adiabatic fraction, and the drop size distribution breadth (Boers et al., 2006; Quaas et al., 2006; Grosvenor et al., 2018). The relationship in Eq. (6) assumes that clouds are adiabatic or nearly adiabatic (i.e. adiabatic clouds or sub-adiabatic clouds with homogeneous mixing only; Brenguier et al., 2000). The most common method uses a bispectral approach to retrieve $r_e$ and $\tau_c$ (Nakajima and King, 1990). Various error sources lead to an overall retrieval error for $N_d$ (Grosvenor et al., 2018; Wolf et al., 2019). As can be deduced form Eq. (6), the most important contributions are from retrieval errors in $r_e$. Other error sources are the uncertainty in sub-adiabatic factor, the cloud model used in the retrieval, and the droplet size distribution width. Satellite retrievals of the vertical profile of cloud droplet size may help to improve the retrieval (Chang and Li, 2002; Chen et al., 2008). Grosvenor et al. (2018) identified biases of retrieved $N_d$ in particular for broken cloud regimes and at large solar zenith angles. In stratocumulus, it was suggested that the retrieval yields the most trustworthy results when considering only the brightest pixels (Zhu et al., 2018). For the ideal case of homogeneous, low-latitude stratiform clouds, relative errors in the $N_d$ retrieval at pixel scale are quantified as 78% (Grosvenor et al., 2018). In such cases, the error was assumed to be random. However, systematic errors occur in particular in broken cloud regimes and for large solar zenith angles, leading to an underestimation (broken cloudiness) and overestimation (large solar zenith angles), respectively, of $N_d$. Painemal et al. (2020) addressed the $N_d$ bias for
broken clouds by only sampling $N_d$ retrieved for large clouds (larger than 5 km $\times$ 5 km) to find that the relation between $N_d$ and aerosols is substantially enhanced.

For improvements in estimates of $N_d$, it would be beneficial to formulate a retrieval in terms of $N_d$ directly rather than in terms of $r_e$ and $\tau_c$. It is also possible to reduce uncertainties in retrievals of $r_e$ and $\tau_c$, or to reduce uncertainties related to assumptions of the vertical structure of the cloud and particle size distribution shape. Approaches to quantify and partly correct for retrieval biases as discussed in Grosvenor et al. (2018) include accounting for cloud heterogeneity by using those channels in passive imagers that provide spatial resolution that exceeds the one at which the standard retrieval products are provided (Zhang et al., 2016). The combination of passive observations with radar may further improve the retrieval (Posselt et al., 2017). Substantially more accurate retrievals of $r_e$ and additional relevant information about droplet size distributions may also come from multi-angular polarimetric measurements (Alexandrov et al., 2012a, b; Shang et al., 2019), which will be possible from orbit at pixel level from the Hyper-Angular Rainbow Polarimeter-2 (HARP-2) on the NASA PACE mission (Martins et al., 2018; McBride et al., 2019). Polarimetric retrievals allow the inference of the spectral width or general shape of the droplet size distribution at cloud top (Hu et al., 2007). This approach is not substantially sensitive to sub-pixel cloudiness, mixed-phase conditions and 3D radiative effects (Alexandrov et al., 2012b). The sensitivity of derived $N_d$ to uncertainties in $r_e$ from polarimetric retrievals may further be reduced by additionally inferring cloud physical thickness. In this case, $N_d$ can be inferred to be linear in $\tau_c$ and inversely linear in geometrical thickness and mean droplet extinction cross section at cloud top (Sinclair et al., 2019). The geometrical thickness may also be inferred from total and/or polarised reflectances measured in oxygen or water vapour absorption bands (Desmons et al., 2013; Sanghavi et al., 2015; Richardson et al., 2019; Sinclair et al., 2019) or by retrieving cloud base using lidar (Müllmenstädt et al., 2018) or using multi-angle observations (Böhnn et al., 2019). When exploiting passive observations together with lidar, $N_d$ at cloud top can be robustly inferred as the ratio of in-cloud extinction (lidar) and extinction cross section (passive). A slightly less direct approach using depolarisation to estimate extinction and effective radius to estimate extinction cross section has been presented by Hu et al. (2007).

### 3.3 Relationship between $N_d$ formed at CCN activation and retrieved radiation-relevant $N_d$

In stratiform clouds, droplets form in updraughts near cloud base which is where $N_d$ most closely relates to CCN. In convective clouds, updraught in some cases increases with height above cloud base. Hence, additional CCN may activate above cloud base and lead to vertically increasing $N_d$ in the lower third of the cloud with a decrease further up (Endo et al., 2015). However, in most cumulus clouds, and in stratiform clouds, $N_d$ is found to be largest at cloud base and to slightly decrease above it (Jiang et al., 2008; Small et al., 2009; vanZanten et al., 2011). In the approach discussed by Grosvenor et al. (2018), the retrieved $N_d$ is representative of the cloud-top reflectance, and thus the relevant proxy for the $N_d$ that matters for cloud albedo and RF$_{\text{act}}$ (Platnick, 2000). To which extent the microphysical structure of lower parts of a cloud exactly impacts radiation (weighting function) depends on the multiple scattering and thus on the vertical structure of $N_d$ itself (Platnick, 2000; Krishma et al., 2018). For vertically constant $N_d$, the retrieved $N_d$ represents the droplet concentration formed by CCN activation. However, there are $N_d$ sinks, in particular due to collision and coalescence (in liquid clouds, the autoconversion and accretion, or “warm rain” processes) that lead to droplet depletion. Wood (2006) demonstrated that the depletion is exponential in precipitation rate and estimated a loss in $N_d$ of 100 cm$^{-3}$ d$^{-1}$ for precipitation rates of 1 mm d$^{-1}$. There may also be lateral and vertical mixing (of heterogeneous type; Lehmann et al., 2009) of cloud air with environmental cloud-free air that can lead to the full evaporation of droplets. In both sinks for $N_d$, the one due to precipitation formation and the one due to mixing, the retrieved $N_d$ is expected to be smaller than the $N_d$ formed at activation of CCN. In an aged cloud, however, updraughts may have decayed such that no additional droplets are formed, while existing droplets persist, or may be advected from elsewhere. Also, in case they are very large, raindrops may break up into droplets, in which case $N_d$ is increased. Arguably, it is the right choice to relate the retrieved $N_d$ as the radiation-relevant one, to CCN, i.e. to use $\beta$, when computing the $N_d$-to-CCN sensitivity with the aim to constrain RF$_{\text{act}}$.

Cloud-resolving models are a good tool to investigate these interpretations (McComiskey and Feingold, 2012). Figure 1 shows an analysis of a large-domain large-eddy simulation with the ICON-LEM model (Heinze et al., 2017; Costa-Surős et al., 2020). CCN concentrations in these simulations are relaxed towards pre-computed spatially and temporally varying fields and are consumed at activation. In the 22 million grid columns, the droplet concentration at cloud top (what is retrieved from satellites) is compared to the maximum droplet concentration (approximately the concentration of activated CCN divided by formed droplets). This demonstrates that there is a link between the droplet concentration formed at activation and $N_d$ determining the cloud radiative effect at its top. These two quantities correlate rather well in the joint histogram, though that link is far from one to one. The second plot (Fig. 1b) assesses the possibility to infer cloud-top $N_d$ from cloud-top $r_e$ and $\tau_c$ (Grosvenor et al., 2018). For this, the MODIS simulator (Pincus et al., 2012) that is part of the Cloud Feedback Model Intercomparison Project (CFMIP) Observational Simulator Package (COSP; Boddas-Salcledo et al., 2011) is applied to the model output to compute cloud-top $r_e$ and $\tau_c$. From these, $N_d$ is computed.
as in Eq. (6). This approach mimics the satellite retrieval but assumes no retrieval errors; i.e. the comparison is a lower bound on the accuracy of the retrieved $N_d$ in representing the actual $N_d$ at cloud top. There is a meaningful co-variation in the two quantities, but it is far from perfect. In particular, there is a systematic overestimation of $N_d$ in the retrieval approach, especially at low $N_d$. The relative error even is a function of $N_d$, with larger relative errors at low $N_d$.

In conclusion, the fact that cloud-top $N_d$ is in general lower than $N_d$ at activation height implies that $\hat{\beta}$ is indeed somewhat smaller than unity. This is not a problem but rather a desired analysis result when studying the Twomey effect. However, $N_d$ obtained from retrieval products is biased high for low values of $N_d$ at top. This relative error, which is a function of $N_d$, implies that the regression between satellite-derived $N_d$ and CCN yields a sensitivity that is too weak.

4 Cloud-regime dependence

Aerosol–cloud interactions depend on cloud regime (Stevens and Feingold, 2009; Mülmenstädt and Feingold, 2018). When it comes to RF$_{aci}$, there are three reasons for this: (i) the radiative sensitivity (Oreopoulos and Platnick, 2008; Alterskjær et al., 2012), i.e. the first two terms on the right-hand-side of Eq. (2) (in particular the sensitivity expressed in Eq. 1); (ii) the updraught dependence of $\hat{\beta}$; and (iii) the dependence of the relation of cloud-top to cloud base $N_d$ to characteristics of turbulence and rain. The latter two are of interest here. “Cloud regime” thus here means, a cluster of clouds with similar $P(w)$ and similar $\frac{dN_d}{dN_d}$ in Eq. (5). When considering CCN at a certain supersaturation level, $\hat{\beta}$ is larger at larger updraught, $w$ (MacDonald et al., 2020). Broadly, cumulus clouds have larger $w$ than stratiform clouds. In addition, clouds over land usually have larger $w$ than clouds over ocean. Building on Eq. (5), this suggests a regime-based analysis expressed as

$$\Delta N_{d,\text{top,ant}} = \frac{dN_{d,\text{top}}}{d\ln a}_\text{regime} \cdot \Delta \ln a_{\text{ant}}. \tag{7}$$

Figure 2 shows the spatial distribution of the $N_d$–AI regression coefficient from its temporal variability within $1^\circ \times 1^\circ$ grid boxes. The large spatial heterogeneity is not straightforward to interpret. Some problems may be due to the lack of aerosol retrieval sensitivity (e.g. in regions with low CCN concentrations such as the southern oceans) or lack of vertical or horizontal co-incidence (e.g. in regions with heterogeneous aerosol and large cloud coverage such as mid-latitude storm tracks). However, aspects of the geographical heterogeneity may indeed be attributable to physical and relevant reasons. However, it is difficult to determine any attributable factors in the spatial and cloud-regime variations in $\hat{\beta}$ (Gryspeerdt and Stier, 2012) before retrieval errors are remedied.

In precipitating situations, the two-way interactions can lead to large challenges in determining the $\hat{\beta}$ term (Ekman et al., 2011). Precipitation scavenges aerosol and, in certain situations, the interplay between aerosol, droplet concentrations and precipitation determines both aerosol and droplet concentrations. This may yield bifurcations between situations with large $N_d$ in which no drizzle forms and very low $N_d$ and cloud dissolution when precipitation forms (e.g. Yamaguchi et al., 2017). In such situations, it is particularly challenging to identify the $N_d$–CCN concentration sensitivity.

5 Aggregation scale

The impact of aggregation scale on estimates of $\beta$ has been discussed in detail by McComiskey and Feingold (2012). Their key conclusion is that at scales larger than the cloud variability scale of about 1 to 10 km, aerosol and cloud data become de-correlated so that the diagnosed $\beta$ becomes less and less representative for individual cloud parcels. In turn, Sekiguchi et al. (2003) computed $\hat{\beta}$ for different aggregation scales and demonstrated that it actually increases with larger scales. An analysis of spatio-temporal vs. temporal-only co-variability of $N_d$ and AOD by Grandey and Stier (2010) found that $\hat{\beta}$ is larger when considering spatio-temporal variability over entire regions compared to only temporal variability at individual $1^\circ \times 1^\circ$ grid boxes. These results are opposite to those expected from the process-based conclusions of McComiskey and Feingold (2012). A possible problem in the Sekiguchi et al. (2003) study is their use of $r_e$ rather than $N_d$, and the subsequent need to stratify by $L$. McComiskey and Feingold (2012) demonstrated that this approach becomes more problematic with increasing aggregation scale. However, their analysis suggested a low-bias in $\beta$ at coarser scales due to stratification by $L$. Reduced $\hat{\beta}$ at small scales could occur if aerosol conditions become too homogeneous to diagnose the full range of co-variability due to smaller sample sizes at smaller scales.

Concluding, from a process point of view, aggregation over larger scales is expected to lead to a decrease in estimated $\hat{\beta}$. In turn, to study the large-scale Twomey effect, an aggregate $N_d$–CCN relationship is desired as it is the large-scale $\Delta N_{d,\text{ant}}$ that matters for the radiation perturbation and because the anthropogenic aerosol perturbation can only be inferred at a large scale. The often adopted choice of a $1^\circ \times 1^\circ$ gridding is somewhat motivated by the suggestion that this is a scale at which aerosol concentrations are considered homogeneous (Anderson et al., 2003) and loosely (to within a factor of about 2 in each horizontal direction; re-analyses are to closer 50 km scales, and many general circulation models still are as coarse as 200 km) related to the scale at which models infer the anthropogenic perturbation of CCN. A rigorous study on the scale dependency of $\hat{\beta}$ and the consequences thereof for RF$_{aci}$ would be desirable.
Figure 1. Analysis of $N_d$ in the “virtual reality” of a cloud-resolving simulation: droplet number concentration ($\text{cm}^{-3}$) from the ICON large-eddy simulation (156 m horizontal resolution) over the domain of Germany for 2 May 2013 (Heinze et al., 2017), for the overpass times of the Terra and Aqua satellites for which the swath of the MODIS instrument covered the domain (twice around 10:30 local solar time for Terra, twice around 13:30 for Aqua) even if no actual data are used in this analysis (Costa-Surós et al., 2020). Joint histograms, normalised along the $y$ axis as in Gryspeerdt et al. (2016) for (a) column-maximum (proxy for activated CCN) vs. cloud-top $N_d$ (taken at $\tau_c = 1$ integrated from cloud top) and (b) $N_d$ derived from $r_e$ and $\tau_c$ as in Grosvenor et al. (2018) vs. cloud-top $N_d$, where both quantities are computed as seen from a satellite using the Cloud Feedback Model Intercomparison project (CFMIP) Observational Simulator Package COSP (Bodas-Salcedo et al., 2011). The blue line is the mean in each bin for cloud-top $N_d$.

Figure 2. Regression coefficients of $N_d$ computed on the basis of retrievals of the MODerate Resolution Imaging Spectroradiometer (MODIS; Platnick et al., 2017) as in Grosvenor et al. (2018) and AI from MODIS (Levy et al., 2013) from the daily temporal variability in grid boxes of $1^\circ \times 1^\circ$.

6 Quantification for the regression coefficient

When sensitivities are approximated by linear regression coefficients from an ordinary least-squares (OLS) line fitting method, rather than derived in the form of joint histograms, the problem of regression dilution arises to the extent that the aerosol quantity shows errors: the regression coefficient becomes gradually smaller as the stochastic error increases (Cantrell, 2008; Pitkänen et al., 2016; Wu and Yu, 2018). Regression dilution, also known as regression attenuation, is a problem if the independent variable ($x$ axis) in the regression is subject to a statistical error. If the regression method does not take the statistical error into account, which is often the case (for example in OLS), the regression coefficient is always systematically biased low. In turn, statistical error on the dependent variable ($y$ axis) only causes uncertainty in the regression coefficient but no systematic bias. This is quantified for the column-CCN vs. $N_d$ sensitivity evaluated as a regression coefficient in Fig. 3. Due to the regression dilution, the sensitivity decreases by factors of 2 to 3 as the error in column CCN increases when considering relative errors of 50 %. This can to a large extent be remedied by ignoring data points at low CCN concentrations from the regression (Fig. 3b). However, this solution is limited to regions not dominated by low aerosol concentrations. Figure 3 also illustrates that an absolute bias in the data translates to relative bias in logarithmic scale. Therefore, if no bias correction is applied, an absolute bias in the data will cause a bias in the sensitivity estimates. As shown by Pitkänen et al. (2016), the regression dilution in turn becomes weaker at coarser ag-
The sensitivity of droplet concentration, $N_d$, to perturbations in the cloud-active aerosol, i.e. the cloud condensation nuclei (CCN) concentration at or above cloud base. The most widely used proxy of the cloud base CCN concentration is the aerosol optical depth (AOD), or alternatively the aerosol index (AI), taken from cloud-free pixels in the vicinity of the locations of the cloud retrievals. The four main caveats with AOD are the lack of vertical resolution, the additional influence of hygroscopic swelling, the fact that the detected aerosol might be not active as CCN nd the impossibility to retrieve it below clouds. In terms of the vertical resolution, satellite-based lidar offers help. However, current lidar retrievals are even more constrained to large aerosol concentrations than passive AOD retrievals. EarthCARE’s ATLID lidar will allow direct inference of the ratio of backscatter to extinction profile. Adding a second wavelength with ATLID capabilities and combining it with polarimetric measurements would substantially extend vertically resolved aerosol information content. In terms of horizontal co-location, trajectory computations may help to identify the aerosol representative of that affecting specific clouds. However, this requires extra effort and reliable information about trajectories. The hygroscopic swelling can be addressed by parameterisations that use retrievals and ancillary data to compute the swelling. Further relevant information is possible from polarimetric measurements.

Cloud droplet number concentration, $N_d$, is only indirectly available from current operational satellite retrievals. It is generally computed from retrieved cloud-top droplet effective radius, $r_e$, and cloud optical thickness, $\tau_c$, leading to substantial biases in comparison to the cloud-top droplet number concentration, especially in inhomogeneous, broken and/or precipitating cloud regimes. Sink processes for $N_d$ and variability due to atmospheric dynamics, including turbulent mixing, imply that the radiatively relevant cloud-top $N_d$ relates imperfectly to the $N_d$ formed by CCN activation. In addition, at a given CCN concentration, the updraught variability also leads to sensitivities of $N_d$ to CCN that are much less than 1. These latter two facts are not problematic when assessing the $N_d$ to aerosol sensitivity from data for the estimation of the Twomey effect. In fact, it is desirable to quantify at a large scale the net impact of aerosol perturbations of the (radiatively relevant) cloud-top $N_d$ that accounts for updraught and $N_d$ sink variability. However, it is necessary to operationally retrieve $N_d$, rather than to indirectly compute it from $r_e$ and $\tau_c$ retrievals. It is also necessary to improve these retrievals in particular for low droplet concentrations and broken cloud conditions. In addition, these retrievals should take into account additional information, e.g. about the onset of drizzle.

Regression dilution influences the statistically inferred sensitivity as a result of stochastic retrieval errors in CCN concentration. On the one hand, at aggregate scales, this...
problem becomes less relevant due to the autocorrelation of the aerosol concentrations. The relationship between $N_d$, which varies at cloud-dynamics scales, and CCN proxies becomes weaker at aggregate scales. Relative retrieval errors in $N_d$ that depend on actual $N_d$ (with larger high-biases at low true $N_d$) lead to a further reduction in the estimated sensitivity. It is thus necessary to account for the impact of CCN errors in the statistics and to optimise the resolution of $N_d$ and CCN retrievals towards cloud-scale resolutions.

The recent study by Hasekamp et al. (2019a) made use of polarimetric satellite measurements to suggest a global-ocean average $N_d$-to-CCN sensitivity of 0.66. This, combined with anthropogenic column-CCN concentrations and radiative sensitivities, translates into a global Twomey effect of $-1.1 \text{ W m}^{-2}$. The net effect of the remaining problems laid out above suggests that this likely is still too low an estimate for the $N_o$–CCN sensitivity, implying a stronger Twomey effect. However, the estimate is in line with an independent observation-based estimate of McCoy et al. (2020) that used differences in $N_d$ between pristine and polluted regions in combination with GCM results as an emergent constraint. In any case, it is desirable to add the extra steps to improve the quantification supported by data for process understanding as well as for evaluating and improving climate models.

In situ and ground-based observations, as well as analysis of cloud-resolving dynamical models, may be a path forward for the evaluation of critical aspects in the satellite-based analysis. Important steps would be the quantification of updraught PDFs for different cloud regimes and the assessment of horizontal homogeneity of aerosol concentrations.

Data availability. The data used in Fig. 1 are the simulation data as described in Costa-Surós et al. (2020, https://doi.org/10.5194/acp-20-5657-2020) and available upon request due to the large amount of data; it is securely saved in tape archives at the Deutsches Klimarechenzentrum (DKRZ), which will be accessible for 10 years. The MODIS data used in Fig. 2 was downloaded from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (https://ladsweb.nascom.nasa.gov/, last access: 19 November 2020, LAADS DAAC, 2020). The data used for Fig. 3 are archived and available at https://cera-www.dkrz.de/WDCU/ucearchseach?acronym=DKRZ_LTA_1002_ds00001 (last access: 19 November 2020, Hasekamp and Quaas, 2020).

Author contributions. JQ led the writing of the manuscript with significant contributions from all authors.

Competing interests. The authors declare there are no competing interests.

Acknowledgements. The work of Johannes Quaas, Annica M. L. Ekman, Athanasios Nenes and Philip Stier was supported by the European Union via its Horizon 2020 project FORCeS. The work of Athanasios Nenes was further supported by the European Research Council via the project PyroTRACH. Philip Stier was also supported by the European Research Council (ERC) project constRaining the EffecTs of Aerosols on Precipitation (RECAP). This review originated from discussions at the 2019 Nanjing workshop of the Aerosols-clouds-precipitation and climate (ACPC) initiative (http://acpcinitiative.org/, last access: 13 November 2020) and benefited from discussions within the group “Study of aerosol–cloud interactions based on satellite observations of the terrestrial underlying surface–atmosphere system: a new frontier of atmospheric science”, hosted by the International Space Science Institute (ISSI). We thank the German Climate Computing Centre (Deutsches Klimarechenzentrum, DKRZ) and the German Federal Ministry of Education and Research (BMBF) within the framework programme “Research for Sustainable Development (FONA)”, https://www.fona.de/ (last access: 13 November 2020) for making the ICON-LEM simulations available. Po-Lun Ma and Johannes Mülmenstädt were supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Earth System Modeling program. Johannes Quaas is grateful to the NASA Goddard Institute for Space Studies, New York, for hospitality during a research stay. We thank Andrew Ackerman and Patrick Chuang for constructive discussions. We thank the three anonymous reviewers and David Painemal for helpful comments on the earlier version of the manuscript. We acknowledge support from Leipzig University for Open Access Publishing.

Financial support. This research has been supported by the European Union’s Horizon 2020 research and innovation programme via H2020-EU.3.5.1 – Fighting and adapting to climate change (FORCeS (grant agreement no. 821205)), the European Research Council (ERC) via H2020-EU.1.1. – Excellent Science program (PyroTRACH (grant agreement no. 726165) and RECAP (grant agreement no. 724602)), the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Earth System Modeling program by the “Enabling Aerosol-cloud interactions at Global convection-permitting scales” (EAGLES) (project no. 74358), and the Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute (contract no. DE-AC05-76RL01830).

Review statement. This paper was edited by Frank Dentener and reviewed by three anonymous referees.

References


Hollstein, A. and Fischer, J.: Retrieving aerosol height from the oxygen A band: a fast forward operator and sensitivity study concerning spectral resolution, instrumental noise, and
Level-1 and Atmosphere Archive & Distribution System (LAADS), Distributed Active Archive Center (DAAC), Goddard Space Flight Center, Greenbelt, Maryland, available at: https://ladsweb.nascom.nasa.gov/, last access: 19 November 2020.


