Satellite observations of convection and their implications for parameterizations

J. Quaas\(^1\) and P. Stier\(^2\)
1 Institute for Meteorology, Universität Leipzig, Germany
2 University of Oxford, U.K.

1 Introduction

Parameterization development and evaluation ideally takes a two-step approach (Lohmann et al., 2007). Insight into new processes, and initial parameterization formulation should be guided by theory, process-level observations (laboratory experiments or field studies) or, if these are unavailable, by high-resolution modelling. However, once implemented into large-scale atmospheric models, a thorough testing and evaluation is required in order to assure that the parameterization works satisfactorily for all weather situations and at the scales the model is applied to. Satellite observations are probably the most valuable source of information for this purpose, since they offer a large range of parameters over comparatively long time series and with a very large, to global, coverage. However, satellites usually retrieve parameters in a rather indirect way, and some quantities (e.g., vertical wind velocities) are unavailable. It is thus essential for model evaluation

1. to assure comparability; and,
2. to develop and apply metrics that circumvent the limitations of satellite observations and help to learn about parameterizations.

In terms of comparability, the implementation of so-called “satellite simulators” has emerged as the approach of choice, in which satellite retrievals are emulated, making use of model information about the subgrid-scale variability of clouds, and creating summary statistics (Bodas-Salcedo et al., 2011; Nam and Quaas, 2012; Nam et al., 2014). In terms of process-oriented metrics, a large range of approaches has been developed, e.g. investigating the life cycle of cirrus from convective detrainment (Gehlot and Quaas, 2012), or focusing on the details of microphysical processes (Suzuki et al., 2011). Besides such techniques focusing on individual parameterizations, the data assimilation technique might be exploited, by objectively adjusting convection parameters and learning about parameter choices and parameterizations in this way (Schirber et al., 2013).

In this chapter, we will first introduce the available satellite data, consider their limitations and the approaches to account for these, and then discuss observations-based process-oriented metrics that have been developed so far.

2 Satellite data

Satellite instruments measure radiances and obtain information about the Earth’s surface and the atmosphere by inverse radiative transport modelling (“retrieval”). Information is contained in the wavelength-dependency of scattering or absorption by the Earth’s surface or particles or gases within the atmosphere. Further information is also available from changes in polarisation. The radiation measured by satellites fundamentally can be of three types:

1. sunlight reflected by the Earth’s surface or by particles in the atmosphere (passive remote sensing in the solar spectrum);
2. radiation emitted by the Earth’s surface or by gases or particles in the atmosphere (passive remote sensing in the infra-red or micro-wave spectrum); or,
3. radiation emitted by the satellite instrument and scattered back by the Earth’s surface or by particles in the atmosphere (active remote sensing).

Passive methods usually have the advantages that they may scan the atmosphere covering a broad swath, and that radiation in multiple wavelength bands may be measured, while active methods have the advantage that by measuring the time elapsed between emission and reception of the signal, a precise vertical sounding is possible.

In terms of viewing angle, there is fundamentally a distinction between “nadir” viewing instruments that measure radiation from a column perpendicular to the orbit directly below the satellite, and “limb” viewing instruments that measure radiation perpendicular to the orbit tangential to the Earth’s surface. Between these two, all inclined angles are also possible and are in use for atmospheric-science applications. Angles close to limb are rarely applied to observe the troposphere due to the long paths sunlight takes through the atmosphere in such cases, which hampers localisation of scatterers or emitters. As such, angles close to nadir are usually used, and passive instruments often scan perpendicular to the orbit to cover a wide swath, with more oblique viewing angles at the outside of the swath, and viewing nadir at its centre. Several satellites in use or in development for atmospheric sciences use cameras that observe at different angles along the orbit, measuring radiation from each column at different angles and thus allowing for a stereo-view and subsequently retrieval of additional information such as of the altitude of cloud- and aerosol layers.

Three kinds of orbits are in use for atmospheric science applications:
1. geostationary orbits at about 36,000 km above the equator. Such a satellite always “sees” one hemisphere, albeit at oblique angles for the remote parts of it. The most important advantage is that a column above a point on the Earth may be observed with high temporal resolution;

2. sun-synchronous polar orbits at lower altitudes (a few hundreds of km), observing each point on the Earth at approximately the same local time of the day for each overpass. The most important advantage is that almost the full globe is covered regularly, depending on the scanning across the track (swath width); and,

3. inclined low-Earth orbits, where, for example, the tropics are covered with higher frequency but polar regions are never overpassed.

In the context of this book, the interest is mostly in remote sensing of clouds and their environment. In the following, we will thus focus on retrieval methods for information about clouds.

Firstly, a retrieval has to decide whether or not a cloud is present in the pixel. In the solar spectrum, this is usually based on the assumption that a cloud scatters more radiation than the Earth’s surface or a cloud-free atmosphere. The distinction is difficult over bright surfaces (desert, snow, sea-ice, sun-glint over ocean) or in the presence of thick, possibly deliquesced, aerosol layers. In the terrestrial spectrum, clouds usually emit less infrared radiation than the surface below a clear atmospheric column, since temperature in the troposphere usually linearly decreases with altitude. The exceptions occur in cases of clouds or within an inversion layer, and the distinction between these is generally difficult within the atmospheric boundary layer. In any event, the detection of clouds becomes difficult if clouds are thin (small liquid- or ice water paths), and the distinction from haze is also ambiguous (Koren et al., 2007; Stubenrauch et al., 2013).

Beyond the detection of clouds, or equivalently (when defined at a larger spatial or temporal scale as the ratio of pixels detected as cloudy to the total number of pixels) the retrieval of cloud fraction, several other characteristics of clouds are of interest. Since the effect of clouds on radiation is so important for weather and climate, and since at the same time radiation is what satellites measure, often the radiative properties are characterised. The radiative effect in the solar spectrum is determined by the cloud albedo, which in turn is a function of cloud optical depth, $\tau_c$. For a layer in which the cloud particle size spectrum remains constant, the cloud optical depth can be written in terms of the liquid water path (mass-weighted vertical integral of the liquid water mixing ratio; similar relationships can be established for ice clouds), LWP, and the effective radius (third divided by second moment of the cloud particle size distribution, i.e. integral of the cube of the radius over the size distribution divided by the integral of the square of the radius over the size distribution), $r_e$, as $\tau_c = 3\text{LWP}/2r_e$. The radiative effect in the terrestrial spectrum is determined by the emission,
or brightness, temperature, $T_B$. Since usually temperature is monotonically
decreasing with altitude (increasing with pressure), cloud-top height, $z_{\text{top}}$, or
cloud-top pressure, $p_{\text{top}}$, can be used to characterise $T_B$.

Cloud optical depth can be retrieved by the reflectance measured in a pixel, with
a priori information about the reflectance of the Earth’s surface. If radiation
is measured in narrow wavelength bands in the visible-near-infrared spectrum,
then by choosing one wavelength where water has little effect and another where
it is more absorbing, $\tau_c$ and $r_e$ can jointly be retrieved if they are within cer-
tain ranges (Nakajima and King, 1990). Brightness temperature can directly
be inferred from measured terrestrial radiation, and can then be converted into
cloud-top altitude or pressure using ancillary information about temperature
profiles. A classic example where these three quantities — cloud fraction, cloud
optical depth and cloud top pressure — are used to characterise clouds is the
International Satellite Cloud Climatology Project (ISCCP). ISCCP produces
joint histograms of the cloud-top pressure and cloud optical depth (Rossow and
Schiffer, 1991, 1999, see also Fig. 1).

In the visible or infra-red spectrum, radiation is scattered or emitted from the
cloud top, or at least from shortly beneath it. In contrast, micro-wave radiation
can penetrate clouds and thus yield information about the full cloud depth (e.g.,
a liquid water path retrieval) and also information about precipitation. This
is particularly used in active remote sensing with radar instruments (Stephens
et al., 2002). When a large range of wavelengths is measured, such as by infrared
sounders, height information can be retrieved by exploiting the exponential de-
crease of density with height in the atmosphere (Susskind et al., 2003, 2006).

3 Sampling of clouds by satellite observations

While sampling biases are a general issue in observational studies, the near
global coverage of most satellite-based remote sensing instruments provides the
advantage of large sample sizes, reducing random sampling errors.

However, satellite resolution, overpass times, viewing geometry and coloca-
tion of instruments all contribute to sampling biases that often remain unquan-
tified.

3.1 Orbital considerations

The majority of satellite based cloud retrievals from space are based on instru-
ments aboard sun-synchronous polar orbiting satellites. While this has retrieval
advantages, such as consistent solar illumination angles and comparability due
to near identical local sampling times, such observations generally provide only
one snapshot of the cloud or convective life cycle. In regions with distinct diur-
nal variability this can introduce significant sampling biases. For example, cloud
top heights retrieved from the MODe rate resolution Imaging Spectroradiometer
Figure 1: ISCCP joint-histogram of cloud optical depth and cloud-top pressure defining 42 cloud classes that may be lumped into nine cloud classes loosely referring to the WMO-defined cloud types. If diagnosed by models, often a sub-visible cloud class (optical depths below 0.3) is additionally defined leading to 49 cloud classes in total. Note that the cloud optical depth is a measure for the solar-spectrum cloud radiative effect, while the cloud top pressure is a measure for the terrestrial-spectrum cloud radiative effect. When analysing satellite or model data, the histogram is normalised such that the integral over it yields the total cloud fraction in the region and time period considered (after Rossow and Schiffer, 1991).
(MODIS) onboard the Aqua satellite (early afternoon equatorial crossing time) are about 100 hPa higher over land as compared to MODIS aboard the Terra satellite (morning overpass time) (King et al., 2013).

Tilted orbits with lower inclination, such as the 35° inclination orbit of the Tropical Rainfall Measuring Mission (TRMM), are designed to statistically sample the diurnal cycle of convection. However, it should be noted that such orbits provide a spatially inhomogeneous sampling. Depending on the swath width of the instrument, this can result in an insufficient coverage to determine, for example, the diurnal cycle of precipitation from even 3 years of TRMM Precipitation Radar (footprint of approximately 5 km) data on spatial scales smaller than 12° and 4 hour temporal averaging (Negri et al., 2002). Such sampling issues are common to all active space-born instruments, characterised by the gain of vertical resolution at the price of a narrow horizontal footprint.

The high temporal resolution of state of the art instruments in geostationary orbits, such as the Spinning Enhanced Visible and Infrared Imager (SEVIRI) and Geostationary Earth Radiation Budget experiment (GERB) instruments with 15 minute resolution aboard the Meteosat Second Generation satellite centred over Africa, provide novel constraints on the observation of convection from space. In a Eulerian framework, the high time resolution allows detailed investigation of the diurnal cycle of convection. For example, Comer et al. (2007) use GERB to quantify the diurnal cycle of outgoing longwave radiation, which they decompose into surface radiation and convective development. However, the high time resolution of geostationary instruments also allows one to investigate convection in a Lagrangian framework. Williams and Houze (1987) developed algorithms to identify and track the evolution of convective systems from brightness temperature observations from geostationary satellites. Such tools have been extended to track and analyse stages of storm development (Zinner et al., 2008) and applied to derive long-term statistics of deep convection (Schröder et al., 2009). However, the superior temporal sampling of geostationary orbits comes at the price of inferior horizontal resolution due to the large distance from the Earth’s surface (36,000 km as compared to hundreds of km for a typical sun-synchronous polar orbit). SEVIRI has a resolution of 3 km for the standard channels at the sub-satellite point and one high-resolution 1 km channel in the visible. In comparison, the widely used polar orbiting MODIS instrument has two bands imaged at a 250 m resolution at the sub-satellite point with five bands at 500 m. Resolution has direct implications for the detection of convection: at km resolutions individual convective cells frequently do not completely fill the instrument’s pixels, which causes the retrieval to be either rejected, introducing sampling biases, or to be erroneous due to the contribution of surface radiances.

This effect is enhanced for slanted paths at higher solar and viewing zenith angles, introducing further sampling biases depending on the viewing geometry. Even for scenes with 100% cloud cover and fully filled instrument pixels the effects of 3D radiative transfer and preferred cloud field orientations can affect the retrieval of cloud properties (Kato and Marshak, 2009).

Another important source of sampling errors is introduced in studies relat-
ing multiple satellite-retrieved cloud- or aerosol properties. It is generally not possible to retrieve aerosol and cloud properties at the same location and time. Aerosol retrievals are strongly affected by cloud contamination while cloud retrievals are strongly affected by partially cloud filled pixels. As a consequence, both aerosol and cloud retrievals apply very conservative masks for their specific purpose, leaving an un-sampled zone that is rejected as potential cloud by the aerosol retrieval and as potential clear-sky by the cloud retrievals. This “twilight” zone is estimated to cover about 20% of the globe, introducing significant sampling biases in both cloud and aerosol retrievals.

3.2 Satellite simulators

Satellite instruments measure radiances, which are not considered directly as model variables. Based on the measured radiances, and applying inverse radiative transfer modelling with ancillary information, cloud quantities are retrieved to enable comparison with models (see above, 2). However, as discussed above, assumptions, resolutions and definitions are different between models and retrieval algorithms. Thus, the quantities as simulated in models, and as retrieved from satellite observations, differ not only due to model errors or measurement errors, but also due to inconsistencies. The aim of satellite simulators is to minimise these inconsistencies.

The pioneering ISCCP simulator (Klein and Jakob, 1999; Webb et al., 2001) and later on the comprehensive Cloud Feedback Model Intercomparison Project (CFMIP) Observational Simulator Package (COSP, Bodas-Salcedo et al., 2011) in three steps yield diagnostics comparable between retrievals and models.

1. A subcolumn sampler to bridge the gap in horizontal resolution. Since general circulation models (GCMs) usually work at resolutions of the order of 100 km, much coarser than satellite retrievals (1 to 10 km resolution), subgrid-scale variability needs to be taken into account for the cloud information. In most GCMs, the only subgrid scale information available stems from the fractional cloud cover and its vertical overlap; although more recent developments also take into account more information about cloud variability (Räisänen et al., 2004). Based this, subcolumns are generated that at each level are either cloud-free or overcast, with homogeneous clouds (Klein and Jakob, 1999).

2. A mimicking of the satellite retrieval. This takes into account instrument limitations and sensitivities such as lower bounds for cloud detectability or the indistinguishability of multi-level clouds by passive instruments that retrieve a mid-tropospheric cloud top pressure (e.g. for a thin cirrus overlying a boundary-layer cloud). It also generates, based on model assumptions, and possibly further ancillary assumptions, observables such as the lidar scattering ratio or radar reflectivity.

3. The diagnostics of summary statistics such as the ISCCP joint histogram
that contain relevant information at arbitrary scales. Other examples are the Contoured Frequency by Altitude Diagrams (CFADs) for lidar scattering ratio or radar reflectivity as a function of altitude.

Often, the retrievals have to be prepared in such a way that they are consistent with what the satellite simulator produces, e.g. for the lidar scattering ratio a vertical rather than horizontal averaging is used (Chepfer et al., 2010). Depending on the application, simplified simulators may be sufficient that just sample cloud-top quantities using the overlap assumption and take into account detection limits by satellite instruments (Quaas et al., 2004). The problem of vertical overlap is introduced along with the computation of a fractional cloud cover since for partly-cloudy skies in two vertical layers, a cloud might overlap either a cloud or clear sky, which affects radiation and precipitation processes. The standard assumptions are “maximum” overlap in which the smaller cloud always is completely covered by a larger cloud, and “random” overlap in which the probability of overlap is just the cloud fraction percentage itself; combinations of the two possibly taking into account the separation of cloudy layers are also in use.

4 Observational constraints on models

4.1 Convection and cloud-type identification

It has become clear that for a thorough observational constraint on cloud parameterizations, a separation of weather situations, or cloud regimes, is essential (Stephens, 2005). The regional separation of predominantly convective and stratiform regimes has often been used to investigate convection. In these studies, tropical clouds are frequently used synonymously with convection and no specific convective identification is applied (e.g. Yang and Slingo, 2001). However, convection also plays an important role in the extra-tropics and even in the tropics it is desirable to gain a better understanding of different convective elements, such as cores and anvil areas.

Jakob and Tselioudis (2003) introduced an objective identification of cloud regimes based on clustering, making use of ISCCP histograms (Tselioudis et al., 2000). The k-means clustering algorithm provides objective mean histograms for each cluster (cluster centroids), which correspond (subjectively) to cloud regimes in the conventional sense, including shallow cumulus, anvil cirrus and deep convective systems.

Such objective clustering of cloud regimes has been used to investigate the performance of climate models (K. Williams and Webb, 2009), highlighting limitations of current climate models in simulating specific cloud regimes such as a significant underestimation of the solar cloud radiative effect of deep convective clouds. It has also been used to assess aerosol-cloud interactions in a regime-based context (Gryspeerdt and Stier, 2012; Gryspeerdt et al., 2014).
4.2 Convective life cycles

Geostationary data, such as that available in the ISCCP dataset, are valuable for tracking clouds along their life cycle. Luo and Rossow (2004) used this data for tracking cloud systems, starting from each occurrence of a cloud characterised as “deep convective” by the ISCCP classification and following that through to convective outflow and anvil cirrus, which decays over time. This approach has proven useful in model analysis (Gehlot and Quaas, 2012) to identify deficiencies such as too little mid-tropospheric detrainment and too long lifetimes of cirrus from convective origins, possibly related to the ice crystal sedimentation parameterization.

4.3 Convective environment: relative humidity variability parameterization

Clouds and convection are embedded in a thermodynamic environment, in which the distribution of relative humidity is one of the most important characterisations. From the perspective of a GCM, important fluctuations of relative humidity take place at subgrid-scales. This subgrid-scale variability of relative humidity, indeed, is the basis of cloud parameterizations in GCMs (Sommeria and Deardorff, 1977; Sundqvist et al., 1989; Smith, 1990) and may be coupled to the convection parameterization (Tompkins, 2002; Klein et al., 2005). Usually the temperature variability is neglected, and in clouds, saturation is assumed, so that the variability in relative humidity translates into variability in the total water specific humidity. Unfortunately vertically resolved satellite retrievals of humidity at high resolutions do not exist. However, from proxies such as the spatial distribution of the vertically-integrated total water path (Weber et al., 2011) or the “critical relative humidity” (Quaas, 2012), evaluations of available parameterizations of the subgrid-scale variability indicate shortcomings of current schemes. This is particularly relevant since processes such as precipitation formation strongly depend on subgrid-scale variability (Weber and Quaas, 2012; Boutle et al., 2014).

4.4 Convective microphysics

In convective parameterizations, microphysical processes are often very crudely parameterized (Tiedtke, 1989). However, ample observations exist that may lead to more realistic parameterizations. This is particularly relevant in light of the multiple effects aerosols might have on convection, implying possible radiative forcings (Khain, 2008; Rosenfeld et al., 2008). The representation of such effects requires, as a necessary but probably not sufficient condition, that precipitation formation is formulated as dependent on droplet number concentrations, and that ice microphysics is included. In convective clouds, droplets first grow by vapour deposition onto them, leading to increased liquid water content but constant
droplet number concentrations. Once precipitation formation sets in, collision-coalescence processes lead to decreasing droplet concentrations at constant cloud water content. Suzuki et al. (2011) have demonstrated from satellite data that this transition between microphysical processes may be identified by investigating the radar reflectivity as a function of cloud droplet effective radius, which follows a power law to the sixth power for condensational growth of particles, and to the third power for collision-coalescence growth. The transition between the processes is not well captured by many models. Aircraft observations imply that accounting for the non-linearity and droplet-size dependency of the precipitation formation process may be as simple as introducing a threshold in the cloud droplet effective radius, by a “critical effective radius” or $r_e \approx 14 \, \mu m$ (Freud and Rosenfeld, 2012). Conditioning the frequency of occurrence of precipitation as retrieved by cloud radar on the cloud liquid water path (Suzuki et al., 2010), or comparing the frequency of occurrence of large radar reflectivities at low altitudes between models and observations (Nam et al., 2014), it is evident that models tend to produce light warm rain far too often. One reason for this is that models often do not account for cloud variability. Once it is taken into account, the problem is much less severe (see above).

5 Outstanding issues and Future Perspectives

The purpose of the present chapter has been to outline the use of satellite data for convection studies, and to suggest its huge potential. Given the vast amount of data available from satellite measurements, its potentials can hardly be overemphasised. There are already scientific achievements in process understanding and advances in the modelling of clouds and convection based on satellite observations. However, there are also several limitations that we still need to overcome.

Although satellite data is available at high resolution, it is usually only high in some particular sense: either temporally or spatially, and only horizontally or vertically. High temporal resolution from geostationary satellites comes at the expense of relatively coarse horizontal and virtually no vertical resolution. Active remote sensing allowing for vertical information comes at the expense of small horizontal and temporal coverage. No detailed, vertically resolved, information at the horizontal resolution of $O(100 \, m)$ is yet available that would be ideal for cloud and convection observations. Also, no vertically resolved information about relative humidity is available at horizontal resolutions better than $O(20 \, km)$.

We should also realise that satellites do not measure everything of interest. Mass flux, or vertical wind, information is only very indirectly available from satellites, and even future Doppler radar instruments as on the Earth Clouds, Aerosols and Radiation Explorer (EarthCARE) will only provide very coarse information (Kollias et al., 2014). In this respect, we should not naively expect that the satellite technology will replace the conventional sounding network in the near future. The important role of conventional soundings has already been
emphasised at several places in the book.

As such, it remains a science topic of its own right to develop methods and metrics that permit insights into processes, and allow for model evaluation and improvement, in order to exploit the wealth of satellite data that is available.

6 Further Readings

For further reading on the recent development of satellite technologies, Stephens et al. (2002); Rosenfeld et al. (2014) provide good starting points on cloud structures and microphysics, respectively.

References


