# Diagnosing Model Systematic Error for Clouds and Precipitation

**Richard M. Forbes**

*European Centre for Medium-Range Weather Forecasts*
*Email: richard.forbes@ecmwf.int*

## 1. Introduction

Clouds are important for weather forecasting and climate prediction, not only for their direct role in the hydrological cycle, but for their interaction with radiation and the dynamics of the atmosphere. In order to improve our representation of clouds in atmospheric general circulation models (GCMs), we require observations that can describe the cloud field to validate and constrain the model parametrization schemes. Fortunately, the quantity, quality and variety of observations of cloud has increased over recent years due primarily to new passive (radiometers) and active (radar, lidar) instruments both from long-term ground based monitoring sites (Ackerman and Stokes, 2003; Illingworth et al., 2007) and on-board satellites (e.g. Stephens et al., 2002). With a model simulation and a set of cloud observations we can diagnose systematic errors in the cloud field. This may suggest that improvements in the model cloud parametrization scheme are required, but cloud is the result of many processes acting together and the error may actually be due to deficiencies in the representation of the radiation, convection, turbulence, cloud physics or the dynamics, all of which interact with each other (Figure 1). The process of model improvement is therefore not straightforward and can be complicated by uncertainties in the observations, differences in observed and model quantities, and compensating errors in the model. The aim of this paper is to highlight some of the main issues when diagnosing model systematic errors for clouds and precipitation in Numerical Weather Prediction (NWP) and climate models, with a focus on methodologies for diagnosing errors, comparing model cloud variables with remote sensing observations, and understanding and validating the representation of physical processes in the model. It is by no means a comprehensive description of the field, but will hopefully provide a few insights into the process of model parametrization improvement.

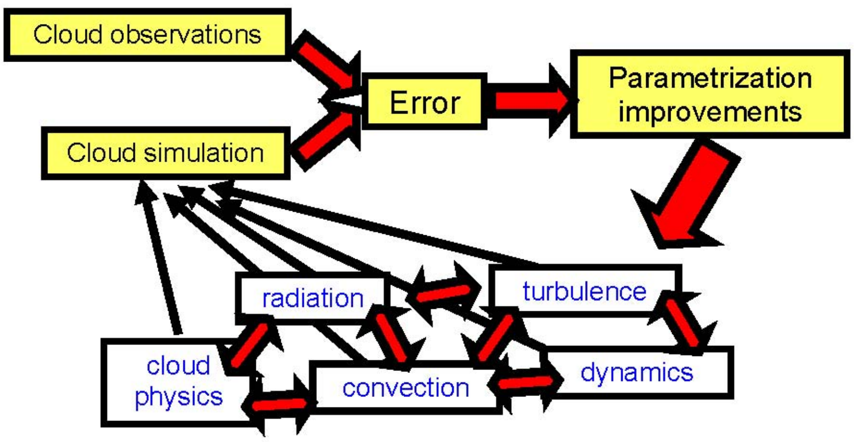


*Figure 1: Schematic of the process of diagnosing cloud-related model systematic errors and improving model parametrization schemes. (from A. Tompkins).*

## 2. Methodology for diagnosing model errors and improving parametrizations

Atmospheric general circulation models (GCMs) are used across a range of resolutions for Numerical Weather Prediction (NWP), monthly and seasonal forecasting, decadal and long-term climate prediction. One aspect relevant to the evaluation of all global modelling systems is the assessment of systematic errors in the "climate" of the model with the observations that are available (observational issues are discussed in the next section). Model climate evaluation in a long (year or multi-year) model integration can provide information on the main problem areas. However, as mentioned earlier, there is the possibility for feedbacks between different model errors, so an apparently poor representation of clouds may not be a problem with the cloud scheme, but be due to problems in the dynamical forcing or complex interactions between other processes in the model. For NWP, in addition to assessing the climate of the model it is possible to routinely compare models with observations in the data assimilation and short-range forecast environment, which allows the separation of model errors directly associated with physical processes from the overall drift and feedbacks of longer climate integrations (Rodwell and Palmer, 2007; Williams and Brooks, 2008). However, even when a general problem has been identified, it can be difficult to attribute the error to a specific deficiency and it is necessary to focus even more closely on particular physical processes in the model.

Jakob (2003) describes a strategy for the evaluation of cloud parametrizations in GCMs, which consists of identifying major problem areas through simulations of the model climate, focussing in on specific meteorological regimes that are likely to be contributing to the problem, identifying representative case studies, and investigating the problem in detail in these case studies to improve the parametrization. The whole process is then repeated with the improved parametrization (Figure 2). The focus on regimes provides the compromise between individual case studies (which may not be representative or conclusive), and general statistics, which may obscure the link to particular physical processes. Different approaches are required to isolate different types of cloud regimes. For certain regimes that dominate a particular geographic area, such as the maritime stratocumulus decks in the eastern oceanic basins or the trade-cumulus regions, it is relatively straight-forward to limit the diagnostics to these areas with certain criteria to distinguish the cloud type (e.g. high cloud fractions with tops below a certain height). Ahlgrimm et al. (2010) provides an example of an evaluation of the stratocumulus regime. For other transitory cloud systems it is necessary to devise more complex compositing techniques so that key dynamical and hence cloud structures remain intact. For example Klein and Jakob (1999) describe a technique to composite extra-tropical cyclones to assess cloud errors in relation to the cyclone centre. Other examples of compositing by dynamical regime include Webb et al. (2001), and Tselioudis and Jakob (2002).

Determining the regime where the systematic error is dominant brings us one step closer to solving the problem, but it often does not provide enough information to isolate the source of error. Choosing or designing a particular case study that is representative of the regime and highlights the specific problem can then be used to focus on the relevant physical processes through comparison with more detailed models and sensitivity studies.

The Global Energy and Water Cycle Experiment (GEWEX) Cloud System Study (GCSS) (Browning et al., 1993; Moncrieff et al., 1997; Randall et al., 2003) has provided a framework for cloud parametrization development and co-ordination of case studies to investigate particular model problems for different cloud types using Cloud System Resolving Models (CSRMs) and Single Column Models (SCMs – a single column model of a GCM). The GCSS paradigm is to use observations to evaluate parametrizations of subgrid-scale processes in CSRMs, to evaluate CSRM results against observational datasets, then to use the validated CSRMs to simulate cloud systems forced by large-scale observations, and finally to evaluate and improve SCMs by comparing to the observations and CSRM diagnostics. In principle, the improvements in the SCM can then feed directly into the full GCM, although feedbacks in the global system and compensating errors can make the implementation of these physically-based improvements more difficult. An example of the GCSS process for the simulation of a tropical squall line is provided by Redelsperger et al. (2000) and Bechtold et al. (2000). CSRMs are validated directly with observations. A reference CSRM then provides a baseline for the assessment of a number of SCMs with sensitivity results providing information for the further development of GCM convection schemes.

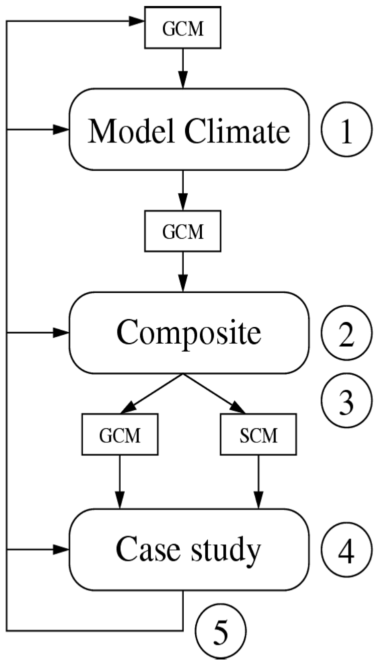


*Figure 2: Schematic of the process of diagnosing cloud-related model systematic errors and improving model parametrization schemes with steps:*

*(1) identify major problem areas,*
*(2) identify major problem regimes,*
*(3) identify typical cases,*
*(4) identify detailed problems*
*(5) improve parametrization.*

*(From C. Jakob)*

## 3. Comparing model and observations: Uncertainties and limitations

The question "What is a cloud ?" may at first glance have an obvious answer, but how do we define cloud boundaries, at what point do falling ice particles become precipitation rather than cloud, how do we class sub-visible cirrus, how do we compare a model representation of a cloud in a partially cloudy

1°x1° grid box with 1 km resolution remote sensing observations…..? A prerequisite for a model-observation comparison is an appropriate transform, either from model to observation space, or from observation to model space, in order to compare "like-with-like". These transforms need to take account of not only the different parameters observed and prognosed by the model, but also the different spatial scales in the model and observations, and the uncertainties and error characteristics of the observations. If these differences are not understood or dealt with appropriately, then the model evaluation can lead to misleading results. Here I focus on three aspects relevant to comparing models and observations, using an example of ECMWF model cloud and precipitation field evaluation with vertical profile observations from the 94GHz radar on-board the CloudSat satellite (Stephens et al., 2002). Although the example is specific to one observation type, the issues are generic and relevant for a wide range of observations.

### 3.1. Comparing model and observed parameters

One problem arises from the different quantities that are provided by the model and observed by the instrument. This particularly applies to remote sensing observations such as radiances, radar reflectivity or lidar backscatter, to be compared with model parameters such as ice/liquid water contents and cloud fraction. In order to transform to a common physical quantity, one option is to use a forward operator to process model output in terms of the observed parameters. For example, using a radar reflectivity forward model (e.g. Haynes et al. 2007) to calculate the attenuated radar reflectivity for the 94GHz CloudSat radar from the model stratiform and convective cloud and precipitation fields (an example is shown in Figure 3). In this case there is additional information required (which may be predicted or diagnosed in the model) such as hydrometeor particle size distributions and particle characteristics. There is, however, some ambiguity whether the source of reflectivity differences is

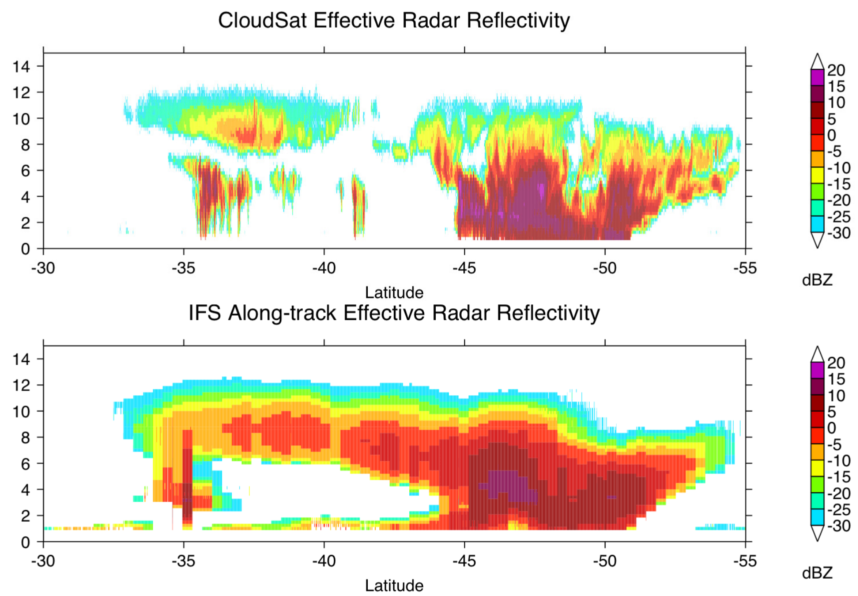


*Figure 3: Example cross-section of 94GHz attenuated radar reflectivity from CloudSat (upper panel) and from a radar forward operator applied to the ECMWF IFS model cloud and precipitation fields (lower panel) along the satellite track through a southern hemisphere mid-latitude front.*

from the amount of condensate prognosed by the model or from the microphysical assumptions that are used for the forward model. It is therefore also of interest to follow the reverse transform where this is feasible, for example using derived observation products (such as the CloudSat Level-2 products providing cloud mask and estimates of cloud phase and water contents). Delanoë and Hogan (2010) describe a method of deriving ice water content from the radar, lidar and radiometer onboard the A-Train satellites, combining data from different instruments to compare directly with the ice/snow prognostic fields in the model. The synergy of different instruments providing additional constraints is an important part of this methodology. There are still a number of assumptions that need to be made in estimating the model quantity from the observations, but approaching the validation problem from different angles with different sets of assumptions reduces the chance of mis-interpreting any results and provides increased confidence in the model evaluation.

### 3.2. Appropriate spatial and temporal matching

There is often a spatial and temporal mismatch between models and observations. The model quantities will be representative of the resolution of the model although sub-grid information such as cloud fraction may be available, whereas the observations may be point locations or small footprint satellite tracks. Again, returning to the CloudSat example, for a fair comparison it is necessary to extract the model data along the satellite track at the appropriate time (the model data in Figure 3 is matched with the CloudSat track and observation time) and then address the mismatch in spatial scales in the model (e.g. 50 km) and observations (~1 km). Sub-grid variability may be predicted by a model, for example in terms of a cloud fraction and an assumption of the vertical overlap of cloud. In which case there are two approaches to overcome the spatial mismatch problem; (1) average observations to a model representative spatial scale (e.g. the grid scale), and (2) statistically represent model sub-gridscale variability to compare directly with the observations. These two options are shown schematically in Figure 4. In option (2) above, the sub-grid scale distribution of cloud elements and vertical overlap is represented by a cloud generator (e.g. Räisänen et al. 2004). For the CloudSat example, the sub-grid cloud generator approach has the advantage that it is able to better represent the impact of attenuation in the model calculated vertical profiles of radar reflectivity. Representativity errors relating to statistical sampling and the fact that CloudSat samples a narrow (1D) track compared to the model 2D grid boxes are other issues that are not explored further here.

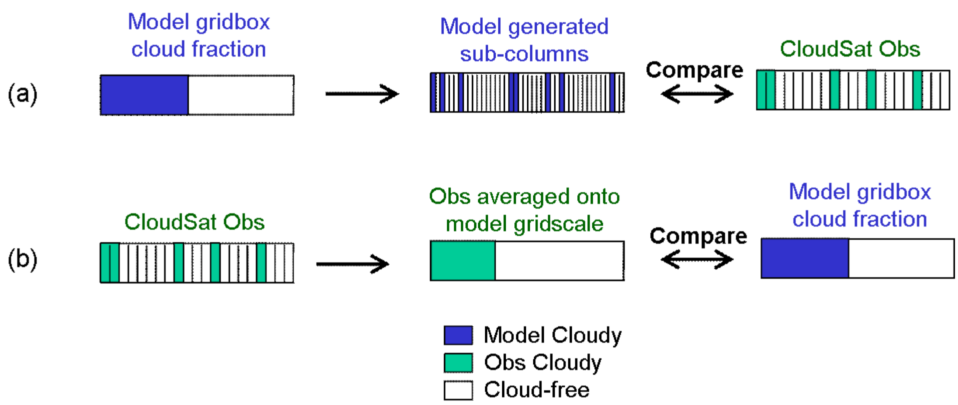


*Figure 4: Schematic of two approaches to addressing the spatial mismatch between the high resolution CloudSat observations and lower resolution model; (a) using a cloud generator on the model cloud fraction field to produce sub-columns representing the high resolution of the observations, (b) averaging the high resolution observations to the lower resolution model grid.*

### 3.3. Observation error characteristics and limitations

All observations have associated error characteristics and limitations in what they can observe. Different observational instruments will detect different characteristics of clouds and understanding these limitations is an important part of the process of model evaluation. For example, Li et al. (2008) and Waliser et al. (2009) compare global liquid and ice water paths, respectively, from a number of different satellite products with widely varying estimates. Some of the differences will be due to deficiencies in algorithms, but a substantial part will be due to the characteristics of the different instruments, either passive or active, observing different components of the cloud and precipitation field. This is the motivation for using a forward operator on the model data to compare against the observations, using all the relevant data available from the model to capture the limitations of the observations and thus perform a fairer comparison. However, for the alternative approach of deriving model variables from the observations, there is benefit in combining observational information from a variety of different sources to provide a more complete picture of the cloud field.

The example of the CloudSat radar and CALIPSO lidar can be used to illustrate the above. The CALIPSO lidar is most sensitive to small particles (backscatter is proportional to particle diameter squared, $D^2$), so the lidar is dominated by a return from cloud droplets and small ice particles, making it particularly useful for observing water clouds and thin cirrus. However, the backscatter signal is quickly attenuated by small water droplets and often only sees the top of water cloud, obscuring precipitation and cloud layers lower in the atmosphere. In contrast the CloudSat radar is more sensitive to large particles (reflectivity $\propto D^6$ ) and therefore is dominated by precipitation as well as cloud containing larger particles. However, the radar can miss cloud with small particles such as thin cirrus and many liquid water clouds. In the forward modelling approach, the radar or lidar forward operator has assumptions built in to emulate the observation limitations (i.e. radar sensitive to larger particles, attenuation, sensitivity threshold). However, it is apparent that the radar and lidar data sources are to a large extent complementary in what they observe, so a combined product has the potential to give a more complete picture of the occurrence of hydrometeors as well as water contents and other properties of cloud and precipitation (e.g. CloudSat GEOPROF-LIDAR product or the combined radar/lidar product developed at the University of Reading - Delanoë and Hogan, 2010).

### 3.4. Summary

The main points from the above discussion can be summarised by the following:

- Limitations and uncertainty:
    - All cloud observations have limitations; they provide a partial picture. We need to know the error characteristics of the observations (which are not always known!).
- Synergy of instruments:
    - Different observation sources have different strengths and weaknesses. We need to make the most of complementary information from different instruments.
- Comparing like-with-like:
    - Model and observed quantities need to be compared in “model-space”, or “observation-space”, taking account of spatial and temporal differences.

- Diagnosing model problems from different angles
  - Using different approaches and different instruments to help to diagnose model systematic errors can provide increased confidence in the robustness of results.

# 4. Some Concluding Remarks

This short paper has briefly explored a few of the questions arising when diagnosing model systematic error, with a focus on cloud and precipitation parametrization. It is by no means a comprehensive discussion, but touches on some of the issues a model developer is faced when evaluating an atmospheric model. Validating the model climate, determining systematic errors for composite/regimes and exploring individual case studies are all necessary steps, with the problems of different parameters, different spatial scales and limitations/error characteristics of the observations. It is crucial to obtain a fair comparison between model and observations in order to identify real model performance and deficiencies, rather than artefacts of the representativity problem. The evaluation process is not always straight forward and there is inevitably some uncertainty in the robustness of the model-observation comparisons. This is why it is important to approach the model evaluation from multiple angles, explore the sensitivity to uncertainties in the assumptions, understand the characteristics and limitations of both observation and model and focus on improving the representation of physical processes.

Model developments in GCMs over the next few years will lead to increasingly complex representations of cloud and precipitation with the possibility of additional predicted variables representing different hydrometeor categories, aerosol and information on sub-grid variability. This will provide further challenges for model evaluation in order to constrain the additional degrees of freedom introduced into the model, but observational data and innovative evaluation techniques will always remain a central activity in the process of model parametrization improvement and continued reduction of model systematic errors.

**REFERENCES**

Ackerman, T. P., and G. M. Stokes, 2003: The Atmospheric Radiation Measurement program. *Physics Today*, January 2003, 38-44.

Ahlgrimm, M., D. Randall and M. Köhler, 2009: Evaluating cloud frequency of occurrence and top height using space-borne lidar observations. *Mon. Wea. Rev.*, **137**, 4225-4237.

Bechtold, P. and co-authors, 2000: A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE. II: Intercomparison of single-column models and a cloud resolving model. *Quart. J. Roy. Meteor. Soc.*, **126**, 865-888.

Browning, K. A. and co-authors, 1993: The GEWEX Cloud System Study (GCSS) . *Bull. Amer. Meteor. Soc.*, **74**, 387-399.

Delanoë, J. and R. Hogan, 2010:Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *Submitted to J. Geophys. Res.*

Haynes, J. M., R. T. Marchand, Z. Luo, A. Bodas-Salcedo and G. L. Stephens, 2007: A multi-purpose radar simulation package: Quickbeam. *Bull. Amer. Meteor. Soc.*, **848**, 1723-1727.

Jakob, C., 2003: An improved strategy for the evaluation of cloud parametrizations in GCMs. *Bull. Amer. Meteor. Soc.*, **84**, 1387-1401.

Klein, S. A., and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.*, **127**, 2514-2531.

Li, J.-L., D. Waliser, C. Woods, J. Teixeira, J. BAcmeister, J. Chern, B.-W. Shen, A. Tompkins, W.-K. Tao and M. Köhler, 2008: Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses, 20th century IPCC AR4 climate simulations, and GCM simulations. *Geophys. Res. Lett.*, **35**, L19710, doi:10.1029/2008GL035427.

Illingworth, A. J., R. J. Hogan and co-authors, 2007: Cloudnet – Continuouse evaluation of cloud profiles in seven operational models using ground-based observations. *Bull. Amer. Meteor. Soc.*, **88**, 883-898.

Moncrieff, M. W., S. K. Krueger, D. Gregory, J.-L. Redelsperger and W.-K. Tao, 1997: GEWEX Cloud System Study (GCSS) Working Group 4: Precipitating convective cloud systems., *Bull. Amer. Meteor. Soc.*, **78**, 831-845.

Stephens, G. L., D. G. Vane and co-authors, 2002: The CloudSat mission and the A-Train. *Bull. Amer. Meteor. Soc.*, **83**, 1771-1790.

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.*, **121**, 3040-3061.

Tselioudis, G. and C. Jakob, 2002. Evaluation of mid-latitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *J. Geophys. Res*., **107**, 4781, doi:10.1029/2002JD002259.

Räisänen, P. H. Barker, M. Khairoutdinov, J. Li, D. Randall, 2004: Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Quart. J. Roy. Meteor. Soc*., **130**, 2047-2067.

Randall, D. A., and co-authors, 2003: Confronting models with data – The GEWEX Cloud Systems Study. *Bull. Amer. Meteor. Soc.*, **84**, 455-469.

Redelsperger, J.-L. and co-authors, 2000: A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE. I: Cloud-resolving models. *Quart. J. Roy. Meteor. Soc*., **126**, 823-863.

Rodwell, M. J, and T. N. Palmer, 2007: Using Numerical Weather Prediction to assess climate models. *Quart. J. Roy. Meteor. Soc*., **133**, 129-146.

Waliser, D., J. Li, C. Woods, R. Austin, J. Bacmeister, J. Chern, A. Del Genio, J. Jiang, Z. Kuang, H. Meng, et al., 2009: Cloud ice: A climate model challenge with signs and expectations of progress. *J. Geophys. Res*., **114**(D00A21, doi:10.1029/2008JD010015).

Webb, M., C. Senior, S. Bony, J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in three climate models. *Climate Dyn*., **17**, 905-922.

Williams, K. D., and M. E. Brooks, 2008: Initial tendencies of cloud regimes in the Met Office Unified Model. *J. Climate*, **21**, 833-840.