Introduction

A workshop on the "Parametrization of clouds and precipitation across model resolutions" was held at ECMWF between 5 and 8 November 2012. The ECMWF strategy for the development of physical parametrizations over the next decade places particular emphasis on moist physics and it was timely for a workshop to discuss the latest advances in research and development on this topic. There are still many questions of how best to parametrize microphysical processes and represent the hydrological, radiative and dynamical impacts of cloud and precipitation across an increasing range of model resolutions, from the global to the convective scale.

The workshop brought together more than 40 international scientists to discuss progress, exchange ideas and provide recommendations for the direction of future cloud parametrization development at ECMWF and in the wider research community. The workshop focussed on three themes; (i) the appropriate level of complexity and numerical formulation of cloud and precipitation microphysics parametrizations, (ii) how to represent the impacts of sub-grid heterogeneity efficiently and consistently across a range of model resolutions, (iii) how to get the most benefit from observations with an emphasis on evaluating and constraining cloud and precipitation processes. Two days of presentations were followed by a day of working group discussions on the three workshop themes and a few points from the working groups are highlighted below.

On microphysics complexity, a single moment bulk microphysics scheme with prognostic variables for cloud liquid, cloud ice, rain and snow was considered appropriate for global-scale NWP at present, although alternatives to the "artificial" split between ice/snow categories should be considered given the uncertainties in nucleation and autoconversion. Additional prognostic variables for number concentration should be investigated in the future as well as a move towards consistency of microphysical assumptions across model parametrizations. For NWP the role of aerosol-cloud interactions in affecting the skill of precipitation forecasts has not yet reached a consensus and further research is required.

For sub-gridscale heterogeneity, the prognostic PDF approach has many attractions, but there are still uncertainties in how to represent the ice phase, mixed phase and precipitation processes. More research is needed in this area, but a hybrid approach targeting the benefits of both the current prognostic cloud fraction and PDF schemes could be the best way forward. High resolution cloud resolving models combined with observations should be used more rigorously in order to provide information on sub-grid heterogeneity and help to formulate source and sink terms for cloud parametrization.

On observations and evaluation, there should be emphasis on using a wide range of observation types to investigate statistical relationships between different variables and to provide a more holistic evaluation of the model to reduce compensating errors. Skill metrics for cloud and radiation should be routinely produced and may provide a more sensitive measure of improvements in skill than the traditional large-scale measures.

ECMWF would like to thank all participants for their contributions to a successful and stimulating workshop.

WG1: Parametrization of Cloud Microphysics

Participants: Paul Field (Chair), Alessio Bozzo (Rapporteur), Hanna Joos, Rita Nogherotto, Lee Hawkness-Smith, Irena Ivanova, Yann Seity, Christine Lac, Ulrike Lohmann, Peter Yau, Greg Thompson, Richard Forbes, Noureddine Semane, Sylvie Malardel

The discussion covered some broad topics such as the required complexity of cloud microphysics, consistency between different parametrizations linked to clouds and the importance of non-regionality in the representation of clouds. Specific questions addressed the matter of dealing with aerosol-cloud interactions, investigation of numerical issues and whether there were outstanding cloud microphysical processes that were not currently well represented.

1. Complexity

Should the computational cost of the cloud microphysics be increased to accommodate additional numbers of moments and hydrometeor species? The table overleaf summarises the discussion concerning complexity of microphysics for three different model configurations that were considered: a convection permitting model for NWP (Hi), a parametrized convection model for NWP (Lo) and a parametrized convection model for climate (Cl).

It was felt that the uncertainties related to ice nucleation and ice autoconversion (through the aggregation and depositional growth) meant that the separation of the ice phase into separate cloud ice and snow categories was not warranted.

Prognostic rain is usually included in the Hi configuration, but including it in the Lo configuration was considered advantageous. It has been shown to improve the character of light rain in other global models (e.g Walters et al. 2011). In addition, increasing the complexity of the rain representation to allow prediction of both mixing ratio and number would improve the treatment of rain that can be produced through either warm rain processes or from the melting of snow.

For the Hi configuration it would be advisable to include graupel, with up to three moments required to capture the strong size sorting associated with this species. The three moments would be mixing ratio, concentration and density or size distribution shape. An alternative way to represent graupel may be through the introduction of a rimed fraction prognostic associated with the snow/ice species.

For the cloud water and ice/snow category it was recognised that number concentration would need to be predicted if aerosol-cloud interactions were to be represented. For NWP the role of these interactions in affecting the skill of forecasts has not reached a consensus. Therefore, the addition of these extra moments would be to allow research into this area.

Species	Prognostic number concentration?	Prognostic mass mixing ratio?	Additional moments?	Notes
Cloud water	Hi, Lo, Cl	Hi, Lo, Cl		For exploring aerosol effects, concentration needed.
Rain	Hi	Hi, Lo		Prognostic rain beneficial. Number concentration also potentially better for representation of warm rain formation, sedimentation and evaporation.
Drizzle	?	?		Is a drizzle category needed? It depends whether rain is diagnostic. Adds additional complexity.
Ice	Hi, Lo, Cl	Hi, Lo, Cl		Snow and ice can be combined in one category. Number concentration helps explicit representation of aggregation. Number concentration allows capture of secondary multiplication (only Hi-res). Also, needed for exploring aerosol-cloud effects.
Snow	Hi, Lo, Cl	Hi, Lo, Cl	Riming(Hi)	As above, snow and ice can be combined in one category. Number concentration helps explicit representation of aggregation. With a riming factor graupel can also be represented in the ice/snow category.
Graupel	Hi	Hi	Density(Hi)	Number concentration for size sorting. Density to represent hail.
Aerosol	Cl	Cl		For aerosol effects in climate models, 1 or 2 (small and large) soluble and insoluble species may be enough?

Table 1: Summary of microphysics representation and complexity for a convection permitting model for NWP (Hi), a parametrized convection model for NWP (Lo) and a parametrized convection model for climate (Cl).

2. Consistency

There was a consensus that efforts should be made to harmonize parametrization settings that involve cloud representations; for instance, cloud schemes, forward model operators for simulating satellite observations and assimilation of observations relating to cloud. Such consistency in principle allows model developers to capitalise on the diverse range of observations available, e.g. correcting errors highlighted by radar observations may lead to improvements also for microwave brightness temperatures. Notwithstanding the conclusions concerning complexity it was recognised that true consistency may require an increase in the complexity of particle size distribution representations to forms with greater degrees of freedom than currently employed (e.g. 3 parameter rain distribution).

As an adjoint to consistency it is recommended that all geographically fixed structural controls on parametrizations should be avoided (e.g. latitude dependence). Besides not being physically justifiable it could potentially hide effects from climate change in climate configurations of a model. It should potentially be possible to capture differences between regions (e.g. fronts versus deep convection) from differences in the characteristics of cloud regimes (e.g. diagnosis of ice particle size distribution from temperature and ice water content).

3. Subgrid representation

The representation of sub-grid heterogeneity of cloud condensate species is important for non-linear microphysical processes such as autoconversion/accretion. The spatial correlation between cloud and precipitation is also important sub-grid information that will affect the microphysical process rates. The effect of the subgrid variability on microphysical process rates can be represented by enhancement factors. However, enhancement factors of order 1-2 are typically smaller than differences between autoconversion/accretion schemes and tuning factors applied. Are enhancement factors for non-linear processes a good enough approach? One reason why some models have large tuning factors for autoconversion is because accretion is too weak and overcompensation of the autoconversion is used to correct the problem with the accretion process that should often dominate. One way to introduce subgrid information is through a diagnostic parametrization approach, based on information from a range of observations, linked to the prognostic variables in the model. An alternative is to use resolved variance at some large scale to infer sub-grid variance of humidity and cloud properties.

Other subgrid considerations include the representation of sub-grid precipitation enhancement over orography due to the seeder-feeder effect.

The effect of turbulence enhancement on collision-coalescence rates is not yet known for atmospheric relevant Reynolds numbers. Direct Numerical Simulation (DNS) needs to be performed at higher Reynolds numbers to build robust parametrizations of turbulent enhancement.

If aerosol-cloud interactions are to be investigated, then it is clear that there needs to be a representation of the sub-grid vertical velocity or relative humidity PDF to provide the correct supersaturation for droplet activation.

4. Cloud-aerosol interaction

The incorporation of aerosol-cloud interactions into NWP models is still very basic. Does including aerosol-cloud interactions improve the forecast? So far, there is no evidence for that in terms of precipitation, but maybe in terms of surface radiation and surface temperature (as seen in convection permitting simulations, Seifert et al., 2012). The current lack of convincing evidence of aerosol-cloud effects leading to improved weather forecasts suggests that, operationally, it may be better to concentrate resources on other aspects such as resolution improvement rather than aerosol complexity. However, it is noted that other groups are using convection permitting models to explore this aspect. There is conflicting evidence concerning precipitation effects (pro: Bell et al., 2009; Rosenfeld et al. 2007, contra: Alpert et al., 2008; Barmet et al., 2009) and therefore a better understanding of aerosol effects on precipitation is still needed at this point in time.

In the near future, if aerosol-cloud interactions are better understood the degree of sophistication of the aerosol representations needs to be explored. For instance, the Met Office has a single moment aerosol prognostic variable currently in the operational high resolution model.

For climate applications capturing aerosol-cloud interactions (first indirect effect, effect on effective radius and hence radiation) is necessary even if they are not well understood yet. Then two-moment cloud microphysics schemes in cloud water and cloud ice are needed and probably more than a single aerosol category.

Finally, for aerosol-cloud interactions it is suggested that a sanity check is performed on current parameterizations to assess the validity of any implicit aerosol effects on clouds (e.g. cloud number concentration dependence on wind speed for the calculation of cloud effective radius for radiation).

5. Missing processes

There was a consensus that the heterogeneous nucleation of ice in most GCMs is crudely represented which can lead to errors in the distribution of the phase of water. This provided some motivation for considering the possibility of introducing prognostic ice nuclei, although uncertainties in the processes are still considerable.

Ice production through secondary multiplication is typically not represented - it requires a prognostic number concentration for the ice species. This effect tends to occur in optically thick clouds and so may not be important from a radiative point of view. However, this process is very efficient at glaciating convective clouds at warm temperatures ($T > -10^{\circ}C$) with impacts on radiation through cloud phase and cloud lifetime changes.

6. Numerical issues

A number of issues were identified that could be usefully explored using simplified kinematic frameworks. This simpler framework should be used to investigate dependency of microphysics on timestep and resolution. It was noted that iterating cloud microphysical processes can lead to some improvement in the light rain representation through an improved representation of autoconversion, sedimentation and evaporation.

It was decided that improved vertical resolution ($dz < 200m$, ensuring consistency with horizontal resolution) would lead to a better representation of the melting layer, boundary layer clouds and mixed phase clouds.

7. Recommendations

The following are the recommendations from this Working Group to consider for microphysics parametrization development:

1. **Microphysics Complexity** (in approximate order of priority)
   - Combine ice and snow species into one ice phase species.

- Have a prognostic representation of rain (even for low resolution) and include substepping to improve representation of autoconversion, sedimentation and evaporation.
- Add an extra moment to rain (number concentration) to improve sedimentation and evaporation.
- Add number concentration for cloud liquid (Nc) for future aerosol-cloud interactions.
- Add number concentration for ice/snow (Nis) to represent aggregation, capture secondary ice multiplication and provide an aerosol-nucleation link.
- For convection permitting models, include a triple-moment representation of graupel.

2. Consistency

- Harmonize microphysical assumptions in parametrizations, forward model operators, and assimilation schemes to achieve consistency where possible.
- Avoid all geographically fixed structural controls on parametrizations (e.g. cloud properties as a function of latitude).

3. Sub-grid representation

- Explore the introduction of enhancement factors to correct for the impact of sub-grid heterogeneity of cloud/precipitation on non-linear microphysical processes.

4. Aerosol-cloud interactions

- At the moment, due to lack of evidence of an impact of aerosol-cloud interactions on numerical weather prediction, resources would be better spent on other aspects of the model rather than increasing the complexity of aerosol representation.
- However, continue to explore the effects of aerosol-cloud interactions on quantitative precipitation forecasting and cloud lifetime. We need better observational tests to disentangle aerosol and meteorological impacts.
- Climate applications do require aerosol-cloud interactions to be represented.

5. Other processes and numerical issues

- Representation of heterogeneous ice nucleation in GCMs is an area that needs improvement.
- Use kinematic models to investigate time and vertical resolution dependencies of microphysics.
- Introduce improved vertical resolution ($dz < 200m$) to better represent thin mid-tropospheric mixed-phase clouds, boundary layer cloud and melting layer processes. Don’t concentrate all resources on horizontal resolution alone – consider consistent increase in vertical resolution.

WG2: Representing sub-grid cloud variability across model resolution

Participants: Cyril Morcrette (Chair), Irina Sandu (Rapporteur), Peter Bechtold, Anton Beljaars, Ian Boutle, Peter Caldwell, Jean-François Geleyn, Luc Gerard, William Gustafson, Tabito Hara, Peter Hill, Berndt Kärcher, Marat Khairoutdinov, Daniel Klocke, Martin Köhler, Vincent Larson, Roel Neggers, Robert Pincus, Thorsten Reinhardt, Axel Seifert.

1. Representation of subgrid variability

Which variables and formulation should we use to represent subgrid variability (e.g. cloud fraction/condensate/humidity or mean/variance/skewness)? What are the appropriate degrees of freedom?

There are two distinct approaches to representing sub-grid cloud variability. The first consists of schemes built around macroscopic variables such as cloud water and cloud fraction (CF schemes), and the second explicitly considers the PDF of sub-grid variability via their means, variances and skewness (PDF schemes). There is no clear argument for which representation is better as each has advantages and difficulties. **Hybrid approaches targeting the benefits of both the CF and PDF approach could be the best method**.

CF schemes will advect cloud fraction as a conserved variable, however the cloud fraction is not a physically conserved variable in the same way as energy, moisture or momentum. As a result, this can provide room for development as there is some freedom in how some terms are formulated. CF schemes also lend themselves more easily to pragmatic development work as it is easier to foresee how changing a source or sink term for condensate or cloud fraction will change the cloud simulated by the model. This is harder for PDF-schemes, where the variables are moments of the PDF, but since these are derived from first-principles, there is also less room for tuning.

The formulation of PDF schemes is more rigorous, self-consistent and elegant. However, it appears that some clouds fields are associated with bi-modal total-water PDFs. These would be difficult to model in a PDF scheme that does not explicitly allow for bi-modality. An elegant way to include the ice-phase in PDF schemes is not available yet (Kärcher and Burkhardt, 2008). Also the coupling of mass-flux convection schemes with PDF schemes is very complicated and we lack the information of the relevant terms (Klein et al., 2005). The DUALM scheme for example uses a combination of mass-flux and bimodel PDFs (Neggers et al., 2009).

CF-schemes have been used for operational NWP for a number of years at both ECMWF and the Met Office. As a result, output from global models using these schemes is looked at every day and gets evaluated using routine verification metrics. This provides a wealth of information about the performance of the scheme beyond what a parametrization scheme developer could hope to obtain on their own. This has helped in the ongoing development and improvement of CF schemes. In comparison, PDF schemes have not been subjected to daily validation within an NWP system. So, although there is vastly more experience of developing and tuning CF schemes than PDF schemes, it was felt that the **continued development and evaluation of PDF schemes is definitely worthwhile and it is certainly too soon to rule them out.**

It was also noted that neither CF or PDF schemes provide any information about vertical overlap. In order to correctly represent the impact of the clouds on radiative and microphysical processes, some information about vertical correlation is required. **Information about the vertical correlation of cloud and precipitation could be obtained from LES or radar/lidar studies before being incorporated into the scheme.**

Future development of either type of scheme should include an analysis of LES data. This would help formulate additional degrees of freedom within CF-schemes for example by adding complexity to represent convective cores and convective or stratiform cirrus. Analysis of LES data could also help validate the formulation of the source and sink terms used within CF schemes. In the case of PDF schemes, analysis of LES data would help to define the source and sink terms for the different moments of the PDF and help to determine optimal PDF shapes.

2. How can we best represent the ice phase, ice supersaturation and mixed-phase microphysics in a statistical pdf scheme?

2.1. PDF/Thermodynamics

It is necessary to include an additional PDF for the ice phase or to include an additional prognostic variable to provide a memory of ice supersaturation and to track ice nucleation history. For a next step, we need to allow for finite diffusion times (in-cloud sub- and super-saturations). Although saturation-adjustment is suited to liquid-phase clouds, it is necessary to **abandon saturation adjustment in ice-containing clouds. This is important for predicting the vapour/ice partitioning in thin cirrus**, for modelling the tropical tropopause layer and representing water vapour transport into the stratosphere.

It is unclear how to best describe mixed-phase clouds, although a temperature-dependent approach (to determine phase partitioning in mixed-phase cloud regime) is not recommended. **Instead using a separate ice fraction variable in addition to a CF or PDF method for liquid cloud may be the best approach.**

2.2. Variability

Since convective ice is different from large-scale cirrus, **it may be better to treat convective ice cloud separately from large-scale ice cloud for the purposes of microphysics and radiation.**

In the longer term, **two-moment microphysics would be needed to differentiate cirrus properties arising from different forcing histories**. This would introduce a regime dependence and have implications for microphysics, radiation and the upper-troposphere water budget.

More work is required to reconcile the high number of ice crystals (>100 per litre, expected from theory in conditions of high small-scale cooling rates) with the high number of ice crystals which are observed from aircraft but which are an artefact of larger crystals shattering as they enter the probe.

Heterogeneity in vertical velocity (w) is important as it drives nucleation. In the boundary layer we use turbulent kinetic energy (TKE) to assess the variability in subgrid vertical velocity, but this is not a good predictor of vertical air motion variability in the upper-troposphere. **More research is required into how to represent subgrid variability of vertical velocity arising from gravity waves**, for example. This would require more targeted observations (along with microphysical data) and/or very high resolution modelling.

2.3. Nucleation

Microphysical information about primary ice initiation should be used to try to include ice physics and ice cloud fraction into the PDF framework.

We are **missing fundamental knowledge about heterogeneous ice nucleation**, both on the level of a microscopic theory and with regard to observations of the ice-nucleated fraction of aerosol populations. However, it is unclear whether in-situ measurements will help settle this problem quickly, while progress on the theoretical aspects is expected to be difficult.

3. How do we treat sub-grid precipitation, overlap with cloud and evaporation of rain in a PDF scheme?

Need to consider the timescales of processes in relation to the time-step of model. A fast process such as condensation can be represented by a PDF approach, but a slow process such as microphysics is more difficult. To the extent that precipitation formation is far away from equilibrium and the process timescales are relevant, then this becomes challenging in PDF-based schemes. If the rain formation is approximated as a fast and local process, like in a Kessler-type autoconversion scheme, then the PDF approach is an appropriate and efficient parameterization. If spatial and temporal correlations become important the standard PDF schemes are no longer appropriate (similar to the problems with non-equilibrium ice or mixed-phase). This issue of non-local precipitation has not been worked through in detail for PDF schemes, i.e., it is not yet possible to say whether this is important or not and on which scales it might become important.

In a PDF scheme, overlap (and maybe evaporation) can be treated by rank correlation or by sampling sub-columns (like in McICA), but numerous simplifications are necessary. The latter can become expensive and might make it necessary to use diagnostic precipitation in the subcolumns. On the other hand, diagnostic precipitation assumptions would deteriorate the representation of precipitation because of the lack of memory. The alternative of keeping the subcolumns over several time steps and maybe even advecting them basically leads towards a stochastic parameterization with sub-grid memory. This might be attractive, but is unexplored territory and computationally expensive. The next level of detail would be to explicitly represent individual clouds (or super-clouds in analogy to the super-droplets method) by stochastic elements with memory and a life cycle (which is currently being explored, e.g. by the Hans-Ertel Centre Group at Hamburg as an extension of the Plant-Craig approach).

Alternatively one could use the PDF-approach only for non-precipitating clouds and treat precipitation elsewhere, but then consistency is immediately lost. A more pragmatic approach would be to introduce an empirical rain overlap scheme based on observations or high-resolution modelling from LES/CRMs.

4. How do we consistently represent sub-grid heterogeneity with other sub-grid parameterizations?

Using a unique PDF for all processes would be the most elegant and simplest to implement. However, the impacts of the choice of shape on the performance of the scheme are not obvious. The ideal would be to handle all sub-grid variability via a PDF which in principle should be of the same form for all the processes (microphysics, radiation, cloud, etc). Since all possible observed PDF shapes cannot be captured by a single PDF form, one needs to find one that performs well in the most types of cases.

PDFs which are defined as a superposition of multiple PDF would allow more flexibility, but would also introduce more complexity. So one can consider a superposition of two Gaussians or two Beta functions for example. Schemes that have more complexity and hence can represent the shape of the PDFs better may be expected to lead to better performance. However more complex schemes are also harder to develop and implement. It is still not clear what the optimal balance is in terms of complexity and performance. In part this is because it depends on what is meant by performing well. Do we want a PDF scheme to (i) accurately represent the shape of the humidity and condensate PDFs in a range of regimes, or do we want the scheme to (ii) accurately represent the impact of the sub-grid nature of the clouds on the microphysical process rates and radiative transfer. Due to compensating errors, one can imagine doing well according to the second measure despite not doing so well according to the first. This is in part because the cloud condensate is in the tail of the total water PDF and it is this that is important for the microphysics and radiation.

When choosing the form of the PDF, one also needs to consider the numerical aspects of its implementation such as whether it is bounded or unbounded.

5. Is the way forward for sub-grid heterogeneity and microphysics to use a statistical sampling (or Monte Carlo) approach or to use analytical upscaling?

The effects of sub-grid heterogeneity on microphysics can be represented by splitting the grid-column into a number of sub-columns and statistically representing the heterogeneity in humidity/cloud/precipitation (statistical sampling, SS), or by integrating over assumed PDFs representing the heterogeneity to give enhancement factors (analytical upscaling, AU). Statistical sampling is able to represent non-local processes, which is more difficult with analytical upscaling. However, computational cost and statistical noise are the two main issues with implementing statistical sampling in an NWP model. **It is only by both SS and AU both being implemented in a model that their relative performance can be compared.**

When time-steps are short and an accurate calculation of microphysical rates requires the microphysics to have memory (such that cloud and rain formation interact with the sub-grid scale structure and vertical overlap of cloud and rain) it is not clear how to proceed with either the statistical sampling or analytical upscaling methods.

The motivation for using the statistical sampling method is to be able to represent the effects of heterogeneity on non-local processes like, for example, rain evaporation, accretion or cloud overlap.

Analytical upscaling can treat only certain local cloud processes like autoconversion, but it is difficult to extend the AU to non-local processes as generally the PDFs do not have explicit spatial and temporal information. One of the disadvantages of the statistical sampling approach is the relatively high computational cost, although that concern could be somewhat alleviated when massively parallel computers are used as the statistical sampling method parallellizes nicely when the samples are independent. The difficulty with the statistical sampling approach is also the choice of sample size. A small number of samples can generate noise on the dynamical grid, which may or may not be beneficial.

6. What aspects of sub-grid schemes need to vary with resolution? At what resolution can we ignore all sub-grid heterogeneity?

One cannot define a clear scale, in absolute terms, at which heterogeneity can be ignored because nature is inherently multi-scale, with additional details added at ever smaller scales as one seeks more detail. However, there are practical scales beyond which smaller scales do not heavily impact the solution.

Fortunately, **detailed analysis of observations and very high-resolution direct simulations can help define the relevant scales that should be preferentially addressed by the parameterization schemes**. A targeted study for this purpose is encouraged, for example the forthcoming HP(CP)2 LES study.

The specific aspects of the schemes that need to vary with resolution depend upon the particular parameterization component and processes within each component. As an example, heterogeneity differs vertically. Near the surface turbulence mixes the atmosphere leading to homogenization, while in the upper troposphere gravity waves contribute to gradients and heterogeneity.

Interactions between schemes is an important aspect of making the entire physics suite accurate at different resolutions; for example, interactions between the boundary-layer scheme, the shallow convection scheme, the cloud macrophysics and the cloud microphysics in the case of stratocumulus clouds transitioning to shallow cumulus convection. These interactions need to be directly addressed in the design of the physics suite so that the different schemes are working together and contributing to the solution in a consistent way. This correct balance should also be maintained across the range of grid resolutions being considered. This could be done by explicitly considering the ratio of the scale of the processes involved compared to the size of the grid. Unfortunately, the different physical processes exhibit variability on a range of different scales and there is no single scale for a given process.

As the grid-size and timestep decreases some processes start to become possible to represent explicitly, and this had led to NWP being run without a parametrization of deep convection. However, for the “grey zones” between grid-lengths of the order of tens of kilometres to grid lengths of the order of kilometres, it is not clear what the best way of representing a process is. Also, the range of resolutions over which this uncertainty exists is different for the different physical processes (e.g. deep convection, shallow convection, boundary-layer mixing, turbulence, microphysics and radiation).

This uncertainty in how we parametrize each physical process and our uncertainty in how the parametrized processes are interacting with each other should be reflected by more stochasticity in the formulation of the parameterizations and/or larger ensemble sizes.

Time scales are an issue in addition to spatial scales. The time scale of microphysics is typically seconds to minutes while for cumulus it is closer to an hour. This leads to numerical issues in that as the physics timestep changes the solution is not robust. The extent of this problem can be investigated by running models with very short timesteps for comparison with operational-like configurations.

While much attention has focused on issues of horizontal resolution, vertical resolution is of equal importance, particularly when running models at smaller grid spacings. To handle gradients at cloud top and bottom boundaries, where the model resolution is often too coarse, there are multiple options available for a better treatment of vertical boundaries, such as higher order numerics or interpolation/extrapolation methods with assumptions about the properties of the vertical gradients.

7. At what grid spacing do microphysics uncertainties dominate over sub-grid heterogeneity uncertainties?

The cross-over grid spacing is problem-dependent (e.g. ocean, land, weak convection, strong). Certain meteorological regimes are more spatially homogeneous but still microphysically complex (e,g, mixed-phase stratus). For instance, LES using single-moment microphysics show that drizzle production in stratocumulus can depend on grid spacing down to 100 m or so. Similarly, CRM simulations show that subgrid vertical velocity is relevant for aerosol activation down to 500 m or so. On the other hand, for forecasts of rain using a simpler microphysics scheme, information about cloud fraction may not be needed by the microphysics scheme even up to 30 km.

The cross-over grid spacing also depends on the relative accuracy of the microphysics scheme versus the subgrid-variability scheme. When analyzing the relative errors, one should use a microphysics scheme that has been developed for the local (smallest) scale, not one that has been tuned to work at coarse resolution by compensating for a lack of subgrid variability.

A possible way forward is by evaluating data from LES or observations to separate uncertainty in subgrid variability from uncertainty in subgrid microphysics. Using an LES, one could average over the horizontal, and feed the averaged fields into a microphysics single-column calculation, and thereby find the errors incurred by ignoring subgrid variability. These could be compared to differences caused by the use of different microphysics schemes.

A related way forward is to use satellite observations (as presented by Matt Lebsock at this workshop) or to use satellite and ground-based retrievals and aircraft data (e.g. Boutle et al., 2013). They found enhancement factors for autoconversion and accretion that represent the errors due to ignoring subgrid variability. These can be compared to differences in process rates as computed by different autoconversion or accretion schemes of varying sophistication.

8. References

Boutle, I. A., Abel, S. J., Hill, P. G. and Morcrette, C. J. (2013), Spatial variability of liquid cloud and rain: observations and microphysical effects. Q.J.R. Meteorol. Soc.. doi: 10.1002/qj.2140

Kärcher, B. and Burkhardt, U. (2008). A cirrus cloud scheme for general circulation models. Q. J. Roy. Meteor. Soc., 134, 1439–1461. DOI: 10.1002/qj.301.

Klein, S. A., Pincus, R., Hannay, C., and Xu, K.-M. (2005). How might a statistical cloud scheme be coupled to a mass-flux convection scheme? J. Geophys. Res., 110(D15S06). DOI:10.1029/2004JD005017.

Neggers, R., Köhler, M., and Beljaars, A. C. M. (2009). A dual mass flux framework for boundary layer convection. Part I: Transport. J. Atmos. Sci., 66, 1465–1487. DOI:10.1175/2008JAS2635.1.

WG3. Constraining cloud and precipitation parametrizations with observations

Participants: Robin Hogan (Chair), Maike Ahlgrimm (Rapporteur), Hélène Chepfer, Andy Heymsfield, Edouard Martins, Sabatino DiMichele, Matt Lebsock, Marta Janisková, Robin Hogan, Frank Brenner, Keith Williams, Christoph Wittmann, Thomas Haiden, Elias Holm, Johannes Quaas.

1. Synergy of different observation types currently available

- Evaluation of clouds and precipitation in a model should go beyond comparing geographical distributions of individual parameters to individual observational datasets.

- For the observations to be used, this requires the highest possible spatiotemporal resolution, to be as close as possible to the process scales. Determining the s**ubgrid PDF of total water** requires combining observations sensitive to clouds and water vapour. From the ground this could be Raman lidar from the Earlinet lidar network plus cloud retrievals from Cloudnet. There have been aircraft observations of total water in the boundary layer (e.g. Wood & Field, JAS 2000) which could be extended to higher altitudes. From space we need to combine CloudSat and Calipso for clouds with microwave radiometers for vapour, but the latter may have too coarse a spatial scale.

- **We recommend comparing statistical relationships between different variables as a strategy to focus on individual parameterisations**, e.g. drizzle versus liquid water path, which can be done from the ground or from space, helping to inform autoconversion and accretion parametrizations (Suzuki et al., JAS 2011; Matt Lebsock's talk at this workshop), or reflectance versus cloud cover (Konsta, 2012).

- A complementary strategy is to investigate samples **conditioned on specific cloud regimes**, e.g. cumulus versus stratocumulus (Nam et al., GRL 2012; Ahlgrimm and Forbes, MWR 2012) to pinpoint deficiencies in specific aspects of the boundary-layer parametrization.

- **We recommend simultaneously evaluating cloud properties (using cloud radar and lidar) with evaluation of radiative fluxes and precipitation**. In this way errors in clouds can directly be related to errors in precipitation and radiation.

2. Fair model evaluation accounting for retrieval and forward-model uncertainties

- **Satellite simulators** for lidar, radar and passive cloud retrievals (e.g. the CALIPSO, CloudSat and ISCCP simulators within COSP) are useful tools and should be used for the evaluation of clouds in models. However, in drawing conclusions from such an evaluation, one should **perform sensitivity tests** to determine the spread in the forward-modelled observations due to input assumptions. Dialogue with observationalists is necessary to get a good estimate for the range of uncertainty. Moreover, even if the model makes an explicit assumption about these properties (e.g. the Met Office model explicitly assumes an exponential raindrop size distribution), the results may be sensitive to this assumption such that differences in forward modelled signals cannot be unambiguously attributed to errors in water content, which is usually what we want to get out of such a comparison.

- **Tools and data products from the CFMIP intercomparison** could prove useful for model evaluation and are available from http://climserv.ipsl.polytechnique.fr/cfmip-obs/. Since a version of a lidar and radar forward model (similar to Actsim/Quickbeam) is already used with the ECMWF IFS model, the Actsim and Quickbeam metrics and scores used in CFMIP could be used. Conclusions drawn from the CFMIP intercomparison about deficiencies in global models will apply to the IFS and should be taken into consideration.

- The FASTER and EUCLIPSE projects are comparing models to observation retrievals at many surface supersites. Moreover, they are running multiple single column models (SCMs) with perturbed physics over various sites as part of parametrization development. **Since the ECMWF SCM will shortly be freely available, we recommend submitting it to the FASTER/EUCLIPSE projects to be incorporated into their systems**.

- To understand the underlying physics and identify problematic processes, retrieved quantities remain essential. **It is recommended only to use (and only to produce) retrievals that report a reliable characterization of the associated errors**. Ideally this should include propagation of measurement errors and propagation of uncertainties in assumptions. Note that these errors should then be analogous to what is obtained from the sensitivity studies we recommended in the first bullet of this section for the case of simulators. In addition, reporting of averaging kernel information can be useful to elucidate the dependence of a retrieved quantity on *a priori* information, and should be encouraged. The spread of different retrievals of one quantity provides useful complementary information for an error bar, but where there is a big spread, discussion is needed with observationalists: some retrievals are known to perform poorly in some situations, and some retrievals have been deprecated by newer ones and should not be considered.

- Care should be taken to understand representativity issues in observational data. In particular, the spatial resolution of the observations relative to the model must be considered. Note that when a climatology (a long-term mean) of the variable is computed then the representativity errors average away, **but for data assimilation and for pointwise comparisons, as used in case studies, these errors can be important**. See Stiller (JGR 2010) and Lopez et al. (2011 ECMWF Tech Memo). Note also that higher order moments of PDFs from single-site observations (e.g. skewness of relative humidity from Raman lidar) may not be representative of the spatial distributions.

- An obstacle for model evaluation is that no observational dataset has both global coverage, vertical resolution and measures the full diurnal cycle. The A-train does have two points in the diurnal cycle, but this is not enough for evaluating convection: convective parameterizations tend to lead to a diurnal cycle shifted too early in the day, so timing errors could appear as biases when A-train comparisons at these two times are examined. Therefore, in addition to the A-Train, **there is a need to use diurnally resolved observations such as from SEVIRI and GOES geostationary satellites**.

3. Cloud and precipitation verification

- The skill of cloud forecasts in NWP has remained fairly flat, whereas traditional large-scale measures such as geopotential height and even precipitation have improved. However, there has been work to objectively evaluate the skill of cloud forecasts using radar and lidar from the ground (Illingworth et al., BAMS 2007; Hogan et al., QJRMS 2009) and from space (Hogan et al., in prep.). These measures demonstrate that cloud forecasts degrade more rapidly with lead time (within 4 days rather than >8 days for geopotential height) so **we recommend that cloud verification is used in an operational context as it could be more sensitive to measures of skill when comparing a new model cycle to the previous one**, especially if the main changes are to cloud processes.

- **We recommend including skill metrics based on surface and/or top-of-atmosphere radiation fluxes** (from BSRN and satellites, respectively) since they are readily available, sensitive to clouds and their optical properties, and more objective than human observations of cloud cover.

4. Process-based model evaluation

- *Excessive drizzle frequency*. Splitting frequency/intensity of precipitation in skill metrics was useful for pinpointing this issue. A metric proposed in Matt Lebsock's talk at this workshop is useful to evaluate the processes underlying precipitation formation. **The relationship between liquid water path of stratocumulus (and cumulus) and the drizzle beneath them is key, and can be derived observationally** both from the combination of CloudSat and MODIS, and from supersites that have microwave radiometers and radar-lidar drizzle retrievals. The former has global coverage and the latter enables compositing by aerosol content to evaluate the sensitivity to aerosols.

- *Boundary layer clouds*. Boundary layer clouds in models are intimately coupled to the boundary-layer scheme, yet a weakness of current practice is that evaluation and development of the boundary-layer scheme and the cloud/microphysics schemes are generally carried out separately. **Doppler lidar offers great potential for evaluating turbulence predicted by boundary-layer schemes**, including the source of the turbulence (e.g. using skewness to distinguish surface-driven turbulence from cloud-top driven turbulence), and have been used to diagnose boundary-layer type to evaluate models (Harvey et al., QJ in press). Since improving boundary-layer schemes is key for improving prediction of boundary-layer clouds, **Doppler lidar could be combined with surface-based cloud radar and lidar observations to evaluate the clouds and boundary-layer properties together**. A particular example of known model bias is wintertime continental stratus, and with more Doppler lidars being deployed every year, there is potential to use this approach to elucidate physical reasons for this error.

- *Aerosol-cloud interactions*. To what degree should we attempt to include aerosol-cloud interactions in the IFS? It was felt to be much too soon to include aerosol effects on ice or mixed-phase clouds, but it was noted that the Met Office forecast model now has a single prognostic aerosol variable affecting both drizzle rate and cloud reflectance in boundary-layer cloud (the effect is also in the WRF-CHEM, LMDZ-INCA and CHIMERE models). Work is ongoing to evaluate the Met Office scheme using simultaneous observations of liquid water path, drizzle rate and surface aerosol concentrations. While the observations are quite clear that increased aerosol reduces drizzle rate for a given liquid water path, effects on cloud lifetime are very difficult to verify. **We conclude that the IFS *may* benefit from use of a more realistic representation of aerosol amount, but further research is required.** A half-way-house between a land-sea dependence and prognostic aerosol would be to use an **aerosol climatology** (e.g. from MACC or Kinne et al., JAMES 2012 for ECHAM as example) for aerosol amount paired with a simple parameterization relating aerosol amount to CCN.

5. Mixed-phase clouds

- There is a growth in the amount of cloud phase data coming on line, particularly given the latest analysis of CALIPSO lidar depolarization observations. Supercooled liquid water in mixed-phase cloud is important for its radiative impacts amongst others, but **can we evaluate the global distribution of supercooled liquid cloud occurrence?** This can be done in both a retrieval comparison sense and using a CALIPSO lidar simulator. There is potential to evaluate the liquid water path through these clouds using AMSR-E subset by supercooled clouds identified with CALIPSO, although one would need to average to reduce the random error (documented systematic error in AMSR-E LWP is less than 10 g m^{-2}). There is a possibility also to evaluate ice water content falling from supercooled layers, but we need to be aware that the particle types are much more pristine (planar in particular) than the irregular aggregates elsewhere, so the CloudSat retrievals may have a bias.

- There have been **observations showing a higher occurrence of supercooled liquid in the southern hemisphere than the northern hemisphere** at mid-latitudes at the same temperature and latitude (e.g. Fig. 4 of Hogan et al. GRL 2004; Fig. 6c of Hu et al. JGR 2010). This could be due to a lower concentration of ice nuclei in the cleaner southern hemisphere. Can the model capture this difference without having explicit aerosol effects? If so, it would imply a dynamical mechanism. Otherwise, there may be a case for incorporating some kind of aerosol dependence, but it is likely at least initially that the model-observation differences are larger than the hemispheric differences.

6. How important is it to get observations of vertical velocity?

- A potentially exciting new measurement coming on line from around 2015 is the EarthCare satellite Doppler velocity. *In ice clouds*, due to the low vertical winds (which can also be horizontally averaged to reduce them to almost zero) this is largely a measure of the terminal fall-speed. The first important step is to forward-model this variable accounting for the radar weighting and the size distribution, using the most recent particle fall-speed relationship (e.g. Heymsfield and Westbrook). For this to be most useful the ECMWF model ought to have an explicit size distribution assumption. **Ice fall speeds are sometimes used for tuning GCMs to get the correct upper-tropospheric cloud-radiation impacts and this data will be able to constrain the model assumptions**. We have yet to see whether EarthCARE's Doppler velocity is accurate enough to provide a good constraint on the model.

- For the ECMWF global model and other GCMs, it is less clear how much we can profit from the Doppler velocity in *convective clouds* (where it will be the superposition of fall speeds and updrafts) because the resolution of the model necessitates the use of a convection parametrization. However, there will be lots of scope for studies that help us understand convective clouds more generally involving cloud-resolving models, and hopefully these will be of benefit for convection scheme development more indirectly.

7. Miscellaneous

- *Aircraft data*. The University of Wisconsin maintains a database containing aircraft observations in a common format, corrected for shattering effects and covering a range of observational campaigns from the Arctic, mid-latitudes and tropics (http://www.ssec.wisc.edu/ice_models/microphysical_data.html). The archived data contain radiative and optical properties and size distributions. **Recommendation: To explore published literature based on multiple aircraft campaigns for potential application in parametrization development.** Aircraft observations also provide a potential source of information for estimates of upscaling. Availability of aircraft data from the US, Europe and elsewhere in a common format and removal of known problems would facilitate their use by a wider community and should be encouraged.

- *Data assimilation*. The analysis of the initial tendencies of the model after data assimilation allows the full suite of data sources to be used to identify problems in the parameterised physics (Rodwell et al.) and could be extended. Moreover, **first-guess departures,** differences between short-range forecasts and observational data during the assimilation cycle provide valuable information for model evaluation. For the assimilation of cloud-affected observations, such as infra-red radiances, the assimilation is difficult but will become easier as the model representation and positioning of clouds becomes more accurate. We reiterate the need for near-realtime EarthCare observations for assimilation.