ECMWF Feature article

from Newsletter Number 143 – Spring 2015

METEOROLOGY

The skill of ECMWF cloudiness forecasts

tounka25/iStock/Thinkstock

www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/lee5bz2g

This article appeared in the Meteorology section of ECMWF Newsletter No. 143 – Spring 2015, pp. 14–19.

The skill of ECMWF cloudiness forecasts

Thomas Haiden, Richard Forbes, Maike Ahlgrimm, Alessio Bozzo

Correctly predicting cloudiness is an important part of a successful weather forecast. Cloud cover is not just of interest in its own right but also has a major impact on other parameters, such as temperature and solar radiation. It is, however, often highly variable in terms of time and location and can therefore be difficult to forecast. The skill of ECMWF cloudiness forecasts improves if time-averaged values and ensemble forecasts are used, but some regions of the world pose particular challenges.

We investigate cloud forecast skill in the ECMWF model by verifying total cloud cover and solar radiation forecasts against surface observations and satellite data, by analysing the scale-dependence of skill, and by evaluating both high-resolution and ensemble forecasts of cloudiness and solar radiation.

Skill in predicting cloudiness

The difference in downward solar radiation at the surface between an overcast and a clear sky can reach several 100 Wm^{-2} during daytime and thereby strongly impact the weather experienced at the surface as well as the development of the atmospheric boundary layer. At night, the reduced net outgoing longwave radiation in the presence of clouds, especially of low clouds, may alter the structure of the stable boundary layer and strongly affect minimum temperatures. Hence the skill in predicting cloudiness has implications for the skill in predicting other parameters, too.

However, cloudiness is one of the more difficult parameters to forecast. This is illustrated in Figure 1, which shows the skill relative to climatology of the high-resolution forecast (HRES) in ECMWF's Integrated Forecasting System (IFS) in predicting various surface parameters in Europe. As a verification measure, we use the standard deviation of the forecast error, which is independent of any constant biases in either the forecast or the verifying observations. For 2-metre temperature and 24-hour precipitation there is positive skill beyond forecast day 5, whereas for 10-metre wind speed and total cloud cover it becomes negative as early as on days 3 and 2, respectively. If skill is measured relative to persistence, as was done by *Köhler* (2005) in a similar diagram, then the ranking of precipitation and temperature is reversed but total cloud cover still ranks lowest.

Deterministic forecast skill for total cloud cover lags behind other parameters and has improved relatively little over the last ten years, despite some recent progress (*Haiden et al.*, 2015). One explanation could be that, unlike 24-hour precipitation, total cloud cover observed at surface stations (SYNOP observations) is a near-instantaneous quantity. A substantial part of the diurnal variability in cloud cover is due to sub-grid scale variations in vertical motion and humidity fields, which have low predictability.

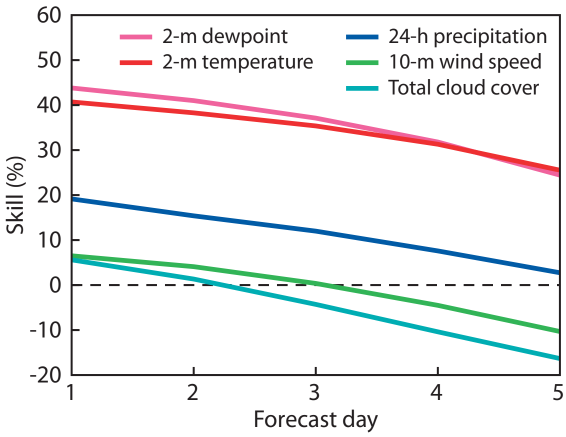

Figure 1 Forecast skill relative to climatology of various surface parameters for Europe in 2014. Skill is defined as the standard deviation of the forecast error divided by the standard deviation of the climatology error (average over the period 1980–2009), where the error is calculated using SYNOP observations. This verification measure is subtracted from 1 and multiplied by 100 to give a skill score that is 100% for a perfect forecast and 0% when equivalent to climatology.

 doi:10.21957/lee5bz2g

Time and space scale-dependence of cloud forecast skill

Total cloud cover

Cyclonic activity and convective processes create cloud systems with a high degree of spatial and temporal variability. Thus part of the low skill of cloudiness forecasts verified at a certain time at a given location may be due to the limited representativeness of the observations. If we verify not just instantaneous values of total cloud cover against SYNOP observations (at 12 UTC, say) but also averages over several reporting times (e.g. 06, 12, 18 UTC), the loss of predictability associated with intra-diurnal variations in cloud cover can be estimated.

There is also a spatial representativeness mismatch between forecasts and SYNOP observations of total cloud cover, whether they are made visually or by an instrument. The area covered by visual observation typically varies between 10 and 100 km around a station, depending on visibility and topography (*Mittermaier*, 2012). Automated observations are derived from ceilometers, which measure cloud cover overhead. Depending on the wind speed in the cloud layer, the scanned area may or may not be representative of the model grid-scale.

Verifying cloud cover averaged over a certain time period reduces this spatial mismatch to a certain extent, depending on how strongly local cloud cover changes are governed by advection. It does not work in situations such as convective clouds over mountainous terrain in the presence of weak flow, where quasi-stationary sub-grid scale variations may persist during the day.

A comparison of the skill, relative to climatology, in predicting instantaneous and averaged cloudiness in Europe is given in Figure 2. The difference in performance for the two quantities is equivalent to about two forecast days and is only weakly dependent on lead time. Comparison with Figure 1 shows that forecasts of averaged cloudiness have a skill comparable to that of 24-hour precipitation.

Solar radiation

Daily averages of downward solar radiation can be used as a proxy for average cloud cover. To assess the skill of such radiation forecasts, we use daily averages of downward solar radiation at the surface provided by the Climate Monitoring Satellite Application Facility (CM SAF). The data is based on observations from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the geostationary Meteosat satellites and has been verified extensively against Baseline Surface Radiation Network (BSRN) data.

Since the main interest here is to use the radiation totals as a proxy for cloudiness, the amount of incoming solar radiation at the surface is normalized by comparing it with typical clear-sky values for a given latitude and season. Day-to-day variations in the normalized values are primarily due to variations in cloudiness, but in some desert areas they also represent the effect of varying dust aerosol content.

Figure 2 shows how the skill of the daily averaged solar radiation forecast compares to that of the averaged cloud cover forecast. Note that the radiation verification has been made for all grid points within the European domain using the daily mean derived from satellite observations every 15 minutes, whereas the SYNOP cloudiness verification is only for about 700 station locations averaged over just 3 validity times. The resulting medium-range skill is slightly higher than for the averaged cloud skill, but both skill measures show the forecast outperforming climatology up to day 4.

The dependence of forecast skill on spatial scale is analysed by averaging the forecast and satellite-derived radiation fields from the original grid at 16 km resolution onto increasingly coarser grids. As a measure of correspondence between time-series of forecasts and satellite-derived values at each grid point, we use the correlation coefficient. Figure 3 shows that, in Europe, the correlation increases by about 0.07 when the scale increases from 20 to 200 km. In the tropics, where the correlation is much lower to begin with, the increase is larger, about 0.11. Over the Southern Ocean, the level of correlation is comparable to the tropics, but with an increase of only 0.05. This suggests that the low radiation forecast skill over the Southern Ocean is more strongly associated with processes on scales resolved by the model.

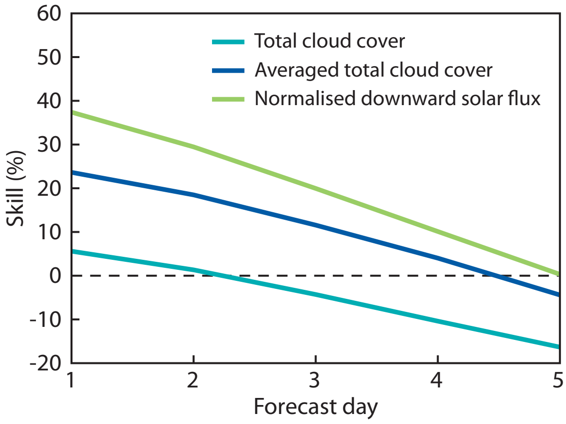

Figure 2 Forecast skill for Europe of instantaneous (00, 12 UTC) and averaged [(06+12+18 UTC)/3] total cloud cover in 2014 from verification against SYNOP, as well as forecasts of daily averages of the normalized downward solar flux at the surface, verified against satellite-derived values.

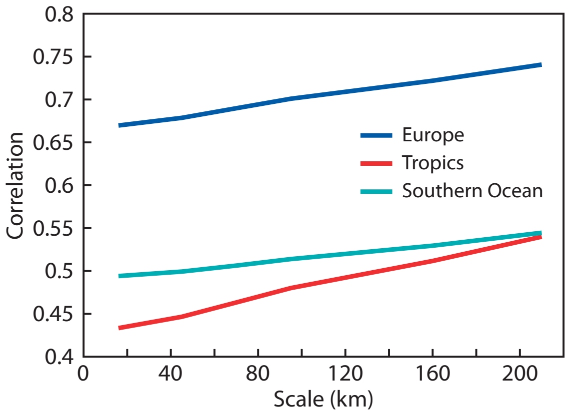

Figure 3 Correlation between forecasts and satellite-derived observations of daily averages of the normalized downward solar flux at the surface on forecast day 3 in 2014 as a function of horizontal scale for Europe (35°–70°N, 12°W–42°E), tropics (20°N–20°S) and Southern Ocean (50°–70°S).

Geographical variations in skill

Overall, the bias in ECMWF's solar radiation forecast is relatively small compared to the magnitude of non-systematic errors, even at short lead times. However, there are certain regions where a specific cloud type dominates. Here, systematic model deficiencies have a more marked effect. The region off the south-western African coast, where the forecast underestimates cloud forcing (Figure 4a), is dominated by marine stratocumulus. Forecasts predict too much solar radiation at the surface over the Southern Ocean, as well as in the mid-latitudes of the northern hemisphere. In terms of absolute values, these correspond to biases of 10–20 Wm^{-2}. There is also a positive bias over North Africa and a small negative bias over the western part of the Atlantic and Indian Oceans.

Note that an overestimation of solar radiation at the surface can be due to a negative bias in cloud fraction, an underestimation of cloud optical thickness, or other reasons, such as an underestimate of aerosol. Some of these systematic errors are discussed in more detail below.

The Southern Ocean and supercooled liquid water

The high latitudes of the Southern Ocean are an area where overestimation of solar radiation at the surface is a common problem in both numerical weather prediction and climate models. There may be a number of reasons for systematic radiation errors in this region, but *Forbes & Ahlgrimm* (2014) highlight the large areas of stratiform low-level cloud that are prevalent over the Southern Ocean as a cloud regime where cloud forcing can be underestimated. Clouds found in this region are often mixed-phase, containing both liquid and ice at sub-zero temperatures, and the correct cloud phase as well as cloud cover and amount of condensate is required to represent the correct impact on radiation.

In 2010, the cloud scheme in the IFS was changed from a diagnostic temperature-dependent liquid/ice phase split to include separate prognostic variables for cloud liquid and ice and associated sources and sinks. This enabled the model to better represent the vertical structure and variable proportions of liquid and ice found in these stratiform clouds (*Forbes & Ahlgrimm*, 2014). This was not only important for the Southern Ocean but also reduced radiation errors for winter stratiform boundary layer cloud over Northern Hemisphere land, with a direct impact on 2-metre temperatures. A significant contribution to the remaining radiation errors over the Southern Ocean and over the North Atlantic mid-to-high latitudes in Figure 4 is due to the lack of supercooled liquid water in convective clouds, which will be addressed in a future IFS cycle.

doi:10.21957/lee5bz2g

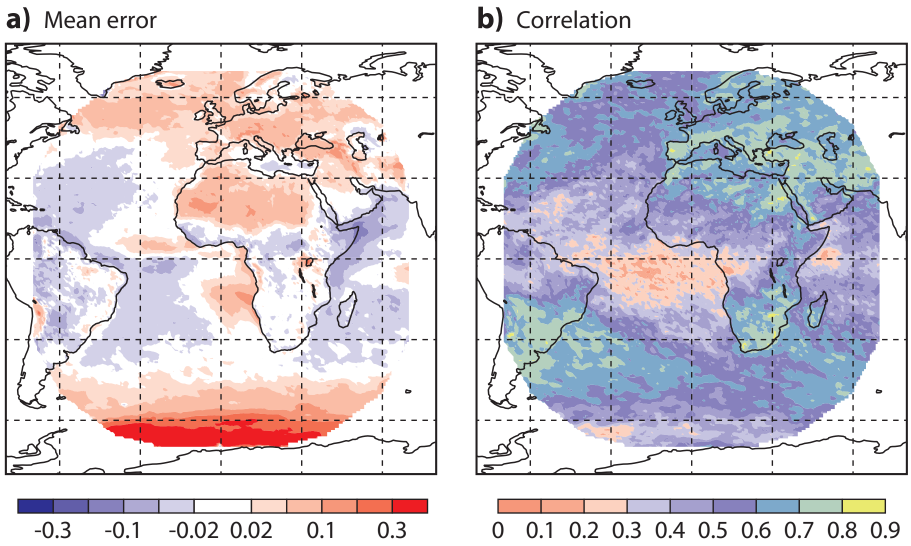

Figure 4 Verification of normalized downward solar radiation at the surface on forecast day 3 in terms of (a) the mean error and (b) correlation of daily averages for the year 2014.

Boundary layer cloud and cloud liquid water

The cloud types dominating the subtropical oceanic regions where surface downward solar radiation is overestimated are stratocumulus (including transitions to cumulus) and stratus. Marine stratocumulus are typically only a few hundred metres thick, and cloud-top radiative cooling plays an important part in their generation and maintenance. The boundary layer associated with stratocumulus is usually well-mixed and topped by an inversion. Thus there are similarities with continental stratus in central Europe, suggesting that model changes aimed at improving the prediction of the former may also improve forecasts of the latter (*Haiden & Trentmann*, 2015).

Work is under way to better integrate the boundary layer, shallow convection and cloud schemes with the aim of improving the stratocumulus-to-cumulus transition and stratocumulus cloud cover and radiative forcing over land and ocean, as well as improving the overly strong radiative forcing in the trade cumulus regions in the western half of the ocean basins.

As mentioned above, model biases in cloud radiative forcing may be due to errors in cloud fraction or optical properties, which for warm-phase cloud are dependent on the assumed effective radius of cloud particles and the liquid water path. Precipitation processes (formation and evaporation) modify the amount of liquid condensate in the cloud as well as the humidity in the sub-cloud layer, which in turn can affect the cloud cover. The next IFS cycle (41r1), to be operational this spring, includes changes to the parametrization of rain formation (autoconversion and accretion) and evaporation (*Ahlgrimm & Forbes*, 2014). This increases the liquid water path, particularly for stratiform cloud over land, which will help to reduce the bias in shortwave cloud forcing.

Africa and dust aerosol

A large negative bias can be seen over the Arabian Sea, and this is more prominent during the summer months. It is linked to the monthly mean climatological aerosol distribution used in the operational model. The distribution of mineral dust in particular dominates the total aerosol optical depth (AOD) in the region, and the negative bias in the solar radiation at the surface suggests an overestimation of the dust AOD. Also related to the climatological dust AOD is the positive bias over the Sahara, although part of it appears to be the result of an underestimation of the surface albedo in the CM SAF radiation product in the area. Tests with a new aerosol climatology based on prognostic aerosol species implemented in the MACC (Monitoring Atmospheric Composition and Climate) system at ECMWF show that the solar radiation at the surface will benefit from an improved AOD distribution and revised optical properties of the aerosol species.

Variation in skill

The skill in forecasting total cloud cover decreases substantially from the mid-latitudes towards the subtropics and tropics. At lower latitudes, a larger portion of the vertical motion field is due to convective processes, which have lower predictability. Figure 4b shows the temporal correlation between forecasts and satellite-derived values. It measures the ability of the model to forecast day-to-day variations of total cloud cover.

In Europe, correlation values of around 0.8 are found, implying that the forecast is able to capture about two thirds of the day-to-day variability. Within Europe, the highest skill is found in Mediterranean land areas. This is due to the less frequent occurrence of low stratus in these areas compared with central and northern Europe. As analysed by *Haiden & Trentmann* (2015), low stratus is underestimated in large parts of continental Europe. It is one of the primary causes of large 2-metre temperature forecast errors in the short to medium range during autumn and winter in Europe.

Compared to the mid-latitudes, skill is rather low in tropical and subtropical areas, with correlations dropping below 0.3, which means that less than 10% of the day-to-day variability is captured by the model. The large area of low skill in the southern tropical Atlantic extends from the Intertropical Convergence Zone (ITCZ) of frequent deep convection into areas more dominated by cumulus and stratocumulus.

Use of ensemble forecasts of cloudiness

We investigate the usefulness of ensemble forecasts (ENS) of cloudiness by first verifying the mean and the median of the ensemble members' total cloud cover forecast against SYNOP observations. We then verify the mean and median of the ensemble members' radiation forecast against satellite data. The error-spread relationship is also considered.

Verification against SYNOP observations

Figure 5a compares the skill of instantaneous total cloud cover relative to climatology of the HRES, the ensemble control (CTRL), ensemble mean (MEAN), and ensemble median (MEDIAN). While the CTRL performs slightly worse than the HRES, both MEDIAN and MEAN show considerably higher skill, and a smaller loss of skill with lead time. The fact that MEDIAN and MEAN outperform the single forecasts by such a large margin, particularly in the short range, is perhaps surprising and we need to ask to what extent it is due to the fact that cloud cover is a parameter that is physically limited to a finite interval, causing error measures to favour forecasts that avoid extremes.

Figure 5b compares the frequency distribution of forecasts at day 3 with observations. In both the HRES and the CTRL there is a much higher occurrence of clear-sky and overcast conditions than in the observations, while in the range from 1–7 okta the frequency is underestimated. The MEDIAN behaves similarly but its distribution is slightly smoothed, which brings it closer to the observations.

The MEAN is much smoother still, with the typical U-shape of the distribution almost lost by the averaging over ensemble members. Although the MEAN underestimates the occurrence of 0 and 8 oktas, it is closest to the observed distribution overall and therefore has highest skill. However, it is important to remember that there is a spatial representativeness mismatch between the modelled and observed cloud cover and uncertainty in the comparison. For example, in SYNOP observations of cloud cover a value of 0 okta is given only if the sky is completely cloudless, and 8 okta are given only if there is no gap visible through the cloud layer(s). To take account of these uncertainties, if the frequencies in Figure 5b for 7 and 8 okta are taken together, then HRES, CTRL and MEDIAN all match the observed combined frequency for the high cloud cover regime quite well, and better than MEAN. This is not the case though for the combined frequency of 0 and 1 okta, which is too high in all forecasts except the MEAN.

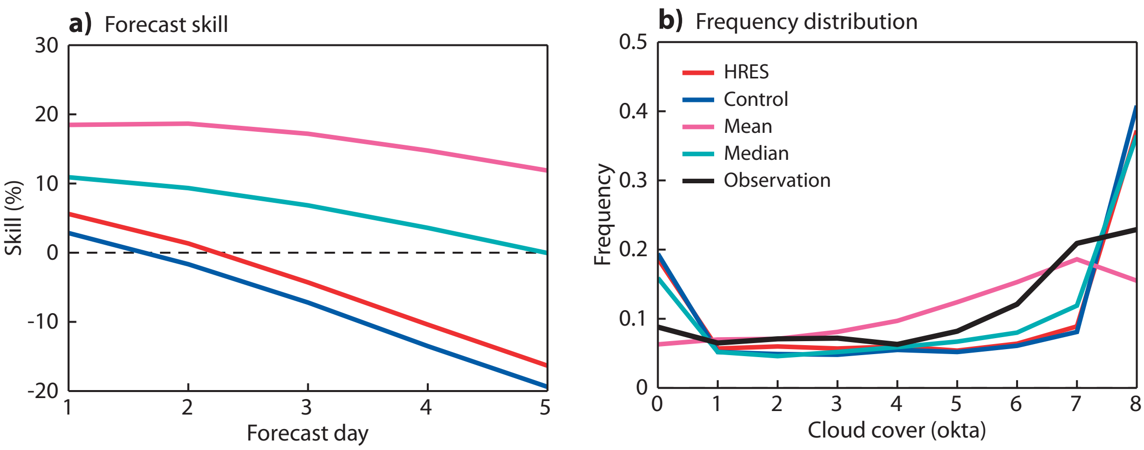

Figure 5 (a) Forecast skill of total cloud cover (00,12 UTC) in 2014 for Europe and (b) frequency distribution for forecast day 3 from the high-resolution forecast (HRES), the ensemble control (CTRL), ensemble mean (MEAN) and ensemble median (MEDIAN) and for observed (SYNOP) cloud cover (Observation).

doi:10.21957/lee5bz2g

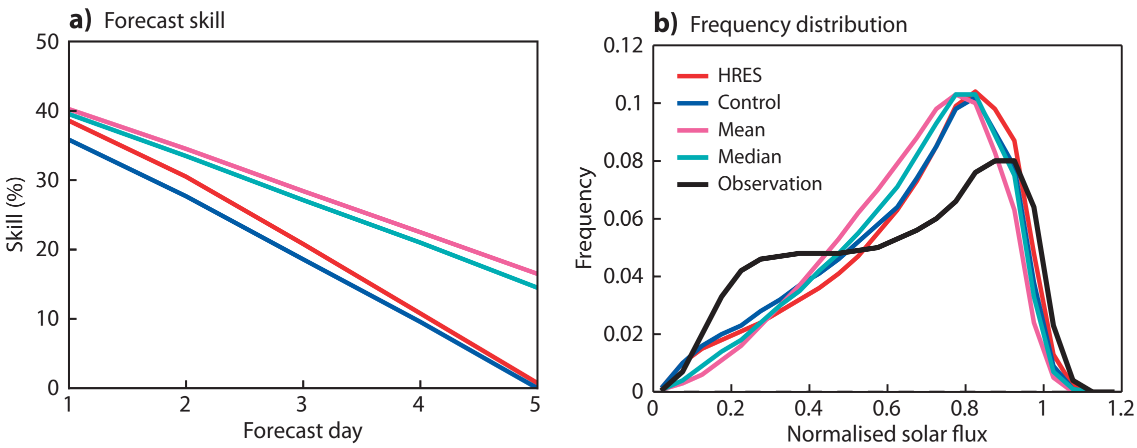

Figure 6 (a) Forecast skill of the normalised downward solar flux at the surface in 2014 for Europe and (b) frequency distribution for forecast day 3 from the high-resolution forecast (HRES), the ensemble control (CTRL), ensemble mean (MEAN) and ensemble median (MEDIAN) and for satellite-derived fluxes (Observation).

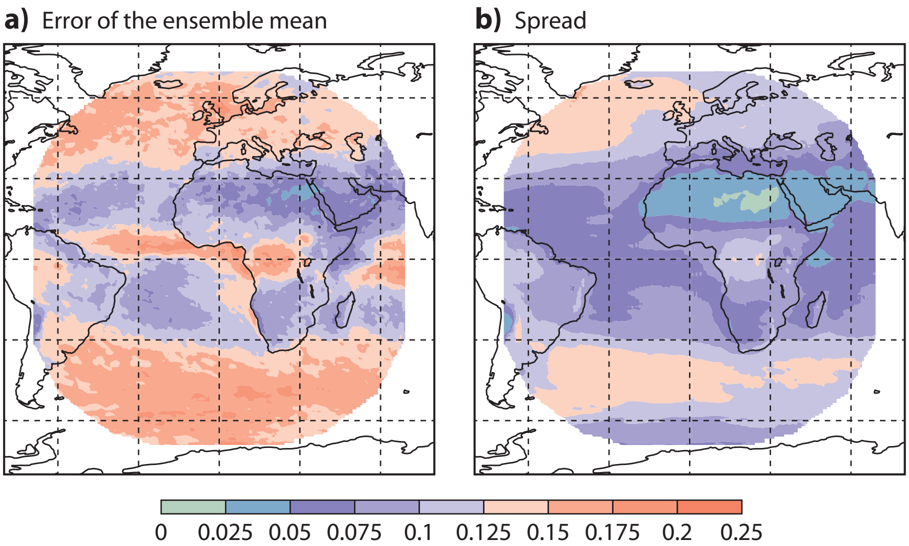

Figure 7 (a) Error of the ensemble mean (i.e. standard deviation of the ensemble mean), and (b) the spread (i.e. ensemble standard deviation around the mean) for daily averages of the normalised downward solar radiation at the surface in 2014 on forecast day 3.

Verfication against satellite data

To see whether the above results carry over to daily averaged radiation forecasts, we perform a similar analysis based on satellite-derived normalized downward solar radiation. Figure 6a shows that the skill of the HRES, CTRL, MEAN and MEDIAN follows the same ranking as in the verification against SYNOP (Figure 5a). However, the benefit of using ensemble information is smaller, especially at short lead times.

The frequency distribution of the normalized radiation values (Figure 6b) is rather different from that of instantaneous cloud cover. In the satellite-derived data it peaks near 0.9 (corresponding to 90% of the typical clear-sky flux at the given latitude and season) and has an almost bimodal shape with a plateau near 0.4. The forecast distributions, which are relatively similar to each other, put the peak closer to 0.8 and they appear to underestimate the frequency of cases in the part of the distribution that represents cloudy to overcast conditions. This is consistent with the results from an evaluation of radiation bias for different cloudiness regimes over central North America (*Ahlgrimm & Forbes*, 2012), where the IFS underestimates the cloud forcing of more overcast cloud (corresponding to normalised solar fluxes of 0.1 to 0.4) due to too little liquid water path and overestimates the cloud forcing in the broken cloud cover regime (corresponding to normalised solar fluxes closer to 1).

Error-spread relationship

An important property of ensemble forecasts is the agreement between error and spread. While this has been monitored for upper-air fields and surface parameters for some time, the solar radiation data makes it possible to obtain a more complete picture with respect to cloudiness. Figure 7a shows the standard deviation of the error of the ensemble mean. There are broad areas of increased error due to cyclonic activity in the mid-latitudes, and due to deep convection associated with the ITCZ. The geographical distribution of ensemble spread (Figure 7b) captures the increase of error from the sub-tropics to the mid-latitudes reasonably well. In the tropics, a comparable match between error and spread is found only over land, whereas over the tropical Atlantic and Indian Ocean the ensemble appears to be strongly underdispersive.

Summary and outlook

While positive forecast skill for instantaneous total cloud cover in the mid-latitudes extends only out to days 2–3, forecasts of daily averaged quantities, such as daily means of solar radiation at the surface, are found to be skilful out to day 5. By using ensemble-derived quantities such as ensemble mean or median total cloud cover, the skilful range can be extended further by 1–2 days. The decrease of forecast skill from the mid-latitudes towards the sub-tropics and tropics is considerable. For a forecast at day 3, the correlation between forecast and satellite-derived downward solar radiation decreases from 0.6–0.8 to 0.2–0.4. Similarly, in the mid-latitudes ensemble error and spread match reasonably well, whereas the ensemble forecast is more strongly underdispersive in the area of the ITCZ. Cloud systems generated by deep convection are inherently less predictable than those associated with extra-tropical cyclones, and the ensemble forecast is not yet able to fully capture the uncertainty in initial conditions and model physics for such systems.

Starting with the introduction of the 5-species cloud microphysics scheme in 2010, changes to microphysical parametrization have led to an increase in cloud forecast skill in recent years (*Haiden et al.*, 2015). At the time of writing, the new model cycle 41r1 has just been evaluated in the operational parallel suite, showing a reduction in total cloud cover root-mean-square error (RMSE) of about 1–2% compared to the operational forecast. With regard to evaluation, there are plans to include satellite-derived bulk cloud properties, such as liquid water path and cloud top temperature, in order to stratify cloud forecast errors and facilitate their attribution to specific cloud processes.

A number of further improvements to cloudiness forecasts, and reduced systematic errors in radiation forecasts, are on the horizon. In particular, work is under way on improved integration of the boundary layer, shallow convection and cloud schemes to improve the stratocumulus and stratocumulus-to-cumulus transition, and on an improved representation of turbulence driven by cloud top cooling for stable boundary layers, which will help to reduce deficiencies in night-time continental stratus. A better representation of mixed-phase cloud, particularly for convection in cold-air outbreaks, will reduce the high-latitude, oceanic, solar radiation bias, and an improved aerosol climatology will reduce errors over North Africa and the Arabian Sea.

Further reading

Ahlgrimm, M. & **R. Forbes,** 2014: Improving the representation of low clouds and drizzle in the ECMWF model based on ARM observations from the Azores. *Mon. Wea. Rev.*, **142,** 668–685.

Forbes, R.M. & **M. Ahlgrimm,** 2014: On the representation of high-latitude boundary layer mixed-phase cloud in the ECMWF global model. *Mon. Wea. Rev.*, **142,** 3425–3445.

Haiden, T., M. Janousek & **D. Richardson,** 2015: Forecast performance 2014. *ECMWF Newsletter No. 142,* **4**.

Haiden, T. & **J. Trentmann,** 2015: Verification of cloudiness and radiation forecasts in the greater Alpine region. *Meteorol. Z.,* **24**, (in press).

Köhler, M., 2005: Improved prediction of boundary-layer clouds. *ECMWF Newsletter No. 104*, 18–22.

Mittermaier, M., 2012: A critical assessment of surface cloud observations and their use for verifying cloud forecasts. *Q. J. R. Meteorol. Soc.*, **138**, 1794–1807.

doi:10.21957/lee5bz2g