

MESO-NH cloud forecast verification with satellite observations

Jean-Pierre Chaboureau
University of Toulouse and CNRS, France

Mesoscale models offer an ideal framework for performing detailed and explicit simulations of cloud and aerosols. First, these models are able to follow the time evolution of different moments of the cloud and aerosol distributions in the context of real meteorological conditions. Second, the grid mesh of a mesoscale model is of the same size as a satellite pixel, facilitating a comparison without it being necessary to resort to additional assumptions on size. Here we adopt a model-to-satellite approach, in which satellite brightness temperature (BT) images are directly compared to BTs computed from predicted model fields. The approach is especially powerful in identifying discrepancies of cloud cover forecasts with BTs at 10.8 μm . The model-to-satellite approach associated with the BT difference (BTD) technique can also verify specific forecasts such as cirrus cover, dust occurrence, convective activity and overshoots. Applications of the approach to the French mesoscale model MESO-NH will be shown using MSG and AMSU observations. In particular, the BTD technique leads us to improve the cloud scheme of the MESO-NH model by tuning a critical parameter in a cirrus parameterization. The model-to-satellite approach is further combined with the calculation of meteorological scores for an objective evaluation of the model forecasts. The applications of the approach will be shown in the context of case studies and meteorological campaigns over the midlatitudes and the tropics.

Simulating satellite imagery from high-resolution WRF output

Dan Lindsey
Cooperative Institute for Research in the Atmosphere
Colorado State University, Fort Collins, CO

As part of the 2010 Spring Experiment at NOAA's Hazardous Weather Testbed in Norman, OK, output from the National Severe Storms Laboratory's (NSSL) 4-km WRF-ARW was used to generate synthetic satellite imagery for each infrared band on the GOES-R Advanced Baseline Imager (ABI). This novel approach served two primary purposes: 1) to familiarize forecasters with the ABI's new spectral channels, and 2) to allow forecasters to compare the model-simulated and observed satellite imagery as a means of evaluating the accuracy of the WRF model forecast.

Although formal cloud verification has not yet been attempted using the synthetic satellite data, it seems like an ideal model diagnostic with which to compare with observed satellite imagery. In this talk, we will explain how the synthetic imagery is created and discuss its potential use for model cloud verification.

Forecast jumpiness in the ECMWF, UK Met Office and NCEP ensemble systems

Dr. Ervin Zsoter

European Centre for Medium-Range Weather Forecasts
United Kingdom, and Hungarian Meteorological Service, Budapest, Hungary

The inconsistency between consecutive forecasts is an important aspect of the behaviour of any numerical forecast system. For example, one of the key advantages that should be expected from using an ensemble prediction system (EPS) instead of a single forecast, i.e. any control forecast, is a better consistency between consecutive runs.

We have investigated the inconsistency properties of the ECMWF, the UK Met Office and the NCEP ensembles with special emphasis on zigzagging or '*jumpiness*', when the EPS (represented by the ensemble mean), the single forecast, or sometimes both jump in a zigzagging fashion for a few consecutive runs.

The main conclusions are that the ensemble-mean forecast jumps less than the control forecast, which is particularly a reflection of the less frequent consecutive, zigzagging jumps. Results also showed that there is no link between forecast inconsistency and forecast error, i.e. forecasts with lower inconsistency do not necessarily have, on average, lower error. By contrast, results have confirmed an existence of a solid relationship between ensemble spread and forecast error.

Measuring Forecast Continuity

David P. Ruth

Meteorological Development Laboratory, Silver Spring, MD

The Meteorological Development Laboratory (MDL) produces model forecast guidance for use at NWS forecast offices. NWS forecasters almost always take into consideration previously issued forecasts. Most model forecasts, however, do not consider previously issued forecasts, and have been known to "flip-flop" as a result. In order to quantify this, MDL developed an index that measures the number of significant swings made over a series of forecast cycles for forecasts valid at the same time. This index is known as the Ruth-Glahn Forecast Convergence Score (FCS).

This presentation will show how the FCS is calculated. It will provide monthly scores for forecasts made by Gridded Model Output Statistics and the National Digital Forecast Database over matched 7-day periods.

Cloud Layer Verification at the Air Force Weather Agency

Matthew C. Sittel, UCAR Associate Scientist

Air Force Weather Agency, Offutt AFB, NE

The Diagnostic Cloud Forecast (DCF) model, run operationally at the Air Force Weather Agency (AFWA), produces forecasts of up to 5 layers of low, middle and high-based clouds. Traditional measures of cloud base and cloud top, while useful, are not practical for identifying multiple cloud layers, as visually these are not always apparent, nor are they 'seen' by most satellites. The NASA Earth System Science Pathfinder satellite, known as CloudSat, includes a cloud profiling radar (CPR) that can detect multiple cloud layers. The CPR examines columns of 125 vertical layers, each a height of approximately 240 meters, and reports the presence or absence of clouds in each layer. Multiple, vertically-adjacent cloud layers, when combined, form an approximation of a cloud layer suitable for comparison to the DCF forecast. This simple methodology is applied to a series of DCF cloud forecasts to produce 2x2 contingency tables of the presence or absence of clouds in the lowest 15 km of the atmosphere.

Using lidar and radar to evaluate cloud forecasts in operational models.

Anthony Illingworth, University of Reading, UK

For several years vertically pointing cloud radars and lidars have been providing vertical profiles of cloud properties which can be interpreted in terms of the variables held in operational models such as cloud liquid and ice water content and cloud fraction. This talk will describe the algorithms developed as part of the 'Cloudnet' program at several European sites for deriving these cloud quantities, mapping them on to the grids of six different operational models, and then deriving statistics to gauge the performance of the models in representing clouds. The advantage of this approach is that feedback on the performance of improved cloud parameterisation schemes in models can be provided within a month or so of their implementation. The 'Cloudnet' approach is now being extended to the ARM sites and to the spaceborne observations from the cloud radar on CloudSat and the lidar on Calipso.

Some initial results from Markov-chain modeling of sequences of lagged probability forecasts

Justin McLay
Naval Research Laboratory, Monterey, CA

Sequences of lagged ensemble-derived probability forecasts are treated as being realizations of a discrete, finite-step Markov chain. Technical aspects of the Markov chain are explored with the aid of a reforecast-ensemble dataset, including the order, homogeneity, and communication of the chain. Cluster analysis and conditional most likely path calculations are used to reveal the key patterns traced by the Markov chain. Monte-Carlo simulation of the lagged probability forecast sequences is undertaken using the transition law estimated from the reforecast ensemble. A simple dynamic decision model is applied to the Monte-Carlo sequence realizations using a basket of cost functions, and the resulting expense incurred by the decision model is conditioned upon the structure of the sequence realizations. Findings show that the greatest average expense is incurred by so-called "sneak" and "volatile" sequence structures, which are structures characterized by large and rapid increases in event probability at short lag times. The experiments also demonstrate how even small improvements in the amount of advance warning of an event can translate into a substantial reduction in decision expense. In general, the conditioned decision expense is found to be sensitive to sequence structure for a given cost function, to the parameters of a given cost function, and to the choice of cost function itself.

Evaluating temporal wind changes in mesoscale numerical forecasts using an object-based method

Daran L. Rife
National Center for Atmospheric Research, Boulder, CO

This talk documents a novel method for evaluating numerical weather prediction models by comparing the characteristics of temporal changes in simulated and observed 10-m (AGL) wind. Temporal objects, or wind events, are defined at the observation locations and at each grid point in the model domain as vector wind change over 2 h. Changes above the uppermost quartile of the distributions in the observations and simulations are empirically classified as significant, though this threshold can easily be adjusted to suit the needs of a particular application. Some analysis and interpretation of the attributes of significant changes will be shown for a 1-year collection of 1-day simulations by the MM5 over southern New Mexico.

The presenter will demonstrate how the simulated temporal shifts at each location tend to organize into spatially coherent structures, even though the events at individual locations are identified independently from one another. Such continuity suggests that displaying them in the same manner as traditional forecast fields are displayed could be useful to forecasters and allow them to monitor real-time observations for indications of such features and to mentally adjust the forecasts based on documented systematic errors in the amplitude, direction, and timing of wind events. Several examples will be shown to highlight the general applicability of this technique to any NWP system and any prognostic quantity.

When can high-resolution NWP model forecasts be trusted? Using error characteristics of the initial conditions to evaluate the likelihood of accurate forecasts

Craig Schwartz

National Center for Atmospheric Research, Boulder, CO

Rigorous evaluation of limited-area high-resolution, convection-allowing (no cumulus parameterization) numerical weather prediction model (CAM) output (e.g., the 2005-2010 NOAA Hazardous Weather Testbed Spring Experiments) has revealed that on some days CAM forecasts are remarkably accurate and provide valuable guidance for weather forecasters, while on other days, CAM solutions are wildly incorrect. It has also been demonstrated qualitatively that CAM forecasts are very sensitive to their initial and lateral boundary conditions that are typically provided by a coarser-resolution, larger-domain “parent” model. Thus, it is likely that the high day-to-day variability of CAM forecast accuracy is modulated by the quality of the initial conditions (ICs) provided by the parent.

Inspired by the “hit or miss” nature of CAM forecasts, this study introduces simple quantitative methods to explore whether error characteristics of the ICs provided by the parent model can be utilized as a predictor of 12- to 36-hr CAM forecast accuracy. Convection-allowing 3-km WRF-ARW forecasts covering most of the United States produced at NCAR during the 2007 and 2010 spring seasons were used to investigate the relationship between the ICs and CAM output. WRF forecasts in 2007 were initialized from 40-km North American Mesoscale (NAM) model grids, while model fields from the 13-km Rapid Update Cycle (RUC) model provided ICs in 2010. For each initialization, the parent’s fields were compared to radiosonde sites considered as “truth.” Then, the corresponding WRF forecast was also verified against soundings. Examining the relationship between the parent’s errors at the initial time and WRF *forecast* errors provided insight regarding the predictability of the WRF forecasts. WRF precipitation forecasts were also compared to the parent’s errors.

In the 2007 dataset, error characteristics of the ICs provided no indication regarding the likely accuracy of next-day WRF forecasts (i.e., “good” ICs did not necessarily mean good subsequent forecasts). However, results from 2010 indicated that the quality of the ICs might be a useful predictor of next-day WRF forecast accuracy. Although there were many configuration differences between the 2007 and 2010 datasets, conflicting results between the two years may be largely attributable to changes in the parent’s resolution and improved ICs in 2010 due to the RUC’s hourly cyclic data assimilation.

Development and Implementation of MODE Time Domain Object-Based

Randy Bullock

National Center for Atmospheric Research, Boulder, CO

Verification MODE Time Domain is an extension of the Method for Object-based Diagnostic Evaluation (MODE) that incorporates the time dimension in verification of gridded scalar fields such as precipitation. MODE is a spatial verification technique that is part of the Model Evaluation Tools (MET) verification software package. Many users have expressed a desire for an object-based verification tool similar to MODE that could handle time variability in forecasts and observations. Like 2D MODE, the Time Domain version uses fuzzy logic to perform object matching and merging, although since one of the dimensions is non-spatial, some object attributes carry new information. Additionally, tracking of spatial objects over time falls out of this approach with no additional effort. This extension of MODE is planned to be available in next year's release of MET.

Using MODE with Cloudsat Data

Paul Kucera

National Center for Atmospheric Research, Boulder, CO

This is a look at a preliminary version of a tool that is planned for a future release of the MET verification software package. It is an extension of the Method for Object-based Diagnostic Evaluation (MODE) that uses Cloudsat satellite data and model data to look at spatial objects in the vertical "curtain" swept out by the Cloudsat observation track. In addition to looking at general storm cases we plan to look at several cases where the Cloudsat track has passed very close to the eye of a hurricane, and compare the spatial objects found with those forecast by the model. Forecast and observed object attributes will be compared, and fuzzy logic matching and merging of objects will be done in a manner similar to that of traditional MODE.