Assimilation of satellite data for meteorology

John C. Derber

Environmental Modeling Center, National Centers for Environmental Prediction
National Weather Service, NOAA, DOC
John.Derber@noaa.gov

ABSTRACT
Over the last 20–30 years, the use of satellite data has evolved tremendously as data assimilation techniques and the observing system have evolved. In this paper, an overview of the assimilation of satellite data for meteorology is given and the relationship between the more detailed presentations at this seminar described. The presentation gives a short history of the use of satellite data, the evolution of data assimilation techniques, and additional considerations for the use of satellite data. Finally, a few of the future challenges are discussed.

1 Introduction

The use of satellite data in meteorological data assimilation has evolved tremendously over the last 20–30 years. Those of us that have been involved in the field over this time have been privileged to see (and in many cases participate in) the rapid development of the techniques for both the use of satellite data in data assimilation and assimilation techniques themselves. This rapid development continues as new satellite data sets are used, current satellite data sets are used more effectively, the use of cloud, aerosol, and constituent gas information becomes a reality, coupled ocean, surface and atmosphere assimilation systems are developed, and assimilation techniques continue to evolve.

Since this was the first full presentation of the Seminar, it was designed to give an overall perspective on the problem and relate the presentations which follow to each other. First, a short historical summary was given. Further details on the history can be found in a presentation by Eyre (2008) at the last similar ECMWF seminar. Second, a short summary of atmospheric data assimilation was given and related to the use of satellite data as an introduction to the subject. A more detailed presentation on atmospheric data assimilation was given by Lorenc at this meeting and many of the other concepts were discussed in more detail in the other presentations. Third, some additional considerations (such as bias correction, quality control, thinning or super-obbing, and monitoring), usually not directly addressed in the theory, are then presented. Finally some of the major future challenges are presented.

2 History

The evolution of the use of satellite data and data assimilation parallels the improvement in forecast skill (Figure 1). Note that this figure (and others in this presentation) is from NCEP, but similar figures can be found from most operational centres. The 5-day forecast skill has evolved from marginally useful (~.60) in the early 1980s (on average) to very useful (> .85) currently. Note the greater increase in forecast skill in the Southern Hemisphere than in the Northern Hemisphere. In the Southern Hemisphere, there are much fewer non-satellite observations than in the Northern Hemisphere. Also note that much of the improvement in the Southern Hemisphere relative to the Northern Hemisphere occurred after the direct use of radiances was incorporated into the operational systems.

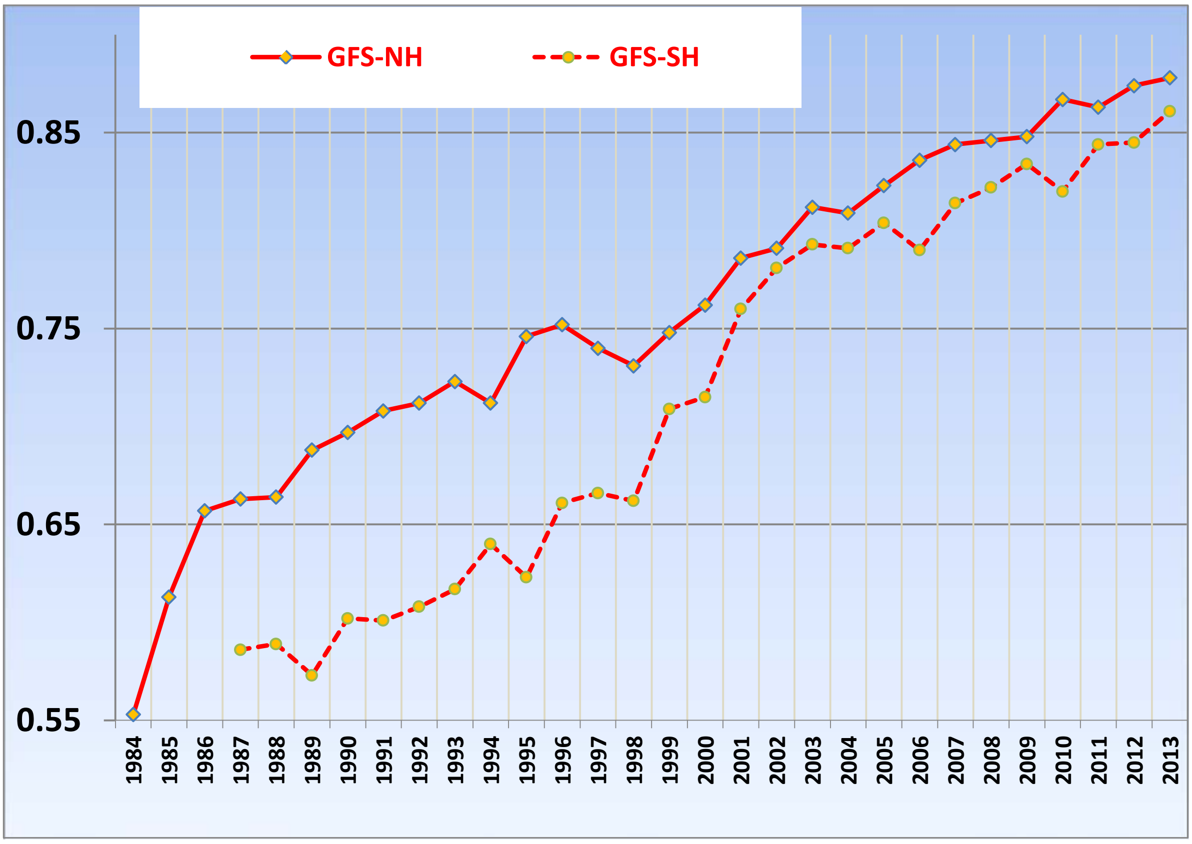

Figure 1: Annual mean forecast skill (anomaly correlation) for NCEP Global Forecast System since 1984. Northern Hemisphere (solid) - Southern Hemisphere (dashed).

In the early 1980s, experiments were performed using retrievals produced from the available satellites. These experiments showed a positive impact, and some operational centres began using them in their operations. By the late 1980s, impact experiments began to show much more mixed results, and considerable effort was expended trying to make them have a more positive impact.

What had changed from the early 1980s to the late 1980s?The quality of the retrievals probably had not decreased. However, the quality of the assimilation and forecast systems had improved, thus making the requirements stricter on the quality and information content of the retrievals. The retrieval process attempted to make temperature and moisture profiles so that the data could be used like the conventional radiosonde profiles. However, the temperature and moisture retrieval process is ill-posed with multiple temperature and moisture profiles potentially fitting the same observations. To get around this problem, a background (guess) field was used to define the unobserved degrees of freedom. Therefore, some of the information in the profile came, not from the observations but rather from the background, and the results were sensitive to the choice of the background. Also, by introducing a background and due to errors in the retrieval process, correlated errors were introduced into the retrievals. These correlated errors were difficult to model with errors both vertically and spatially between retrievals. In addition, errors in the detection of clouds and surface characterization were not always detected by the quality control. It became clear that the use of retrievals to extract the information in the observations was not the best strategy.

With the development of variational assimilation techniques in the early 1990s, the possibility of directly using radiances became real. (Note that with the Optimal Interpolation analysis scheme popular at the time, theoretically could use radiances directly, but this was not done in any operational system.) The variational schemes had the advantages of not requiring the analysis variables to be the same as the model variables or the observations, and all data was used at once.

At this point in time, the use of satellite data and the development of data assimilation techniques became very strongly coupled. Figure 2 shows the same progress of forecast skill as Figure 1, but also includes the lines showing two reanalyses done at NCEP. The first (CDAS) used an assimilation system close to the operational one in 1993 but at lower spatial resolutions. The second (CFSR) shows a reanalysis performed over the same time period using the assimilation system close to operations in 2006 (again at lower resolution than operations, but higher than CDAS). Between these two reanalyses, many enhancements to the forecast model and assimilation system were incorporated (including the direct use of radiances). Note many of the improvements were made possible by increased computational resources to allow algorithmic enhancements to the models and data assimilation systems.

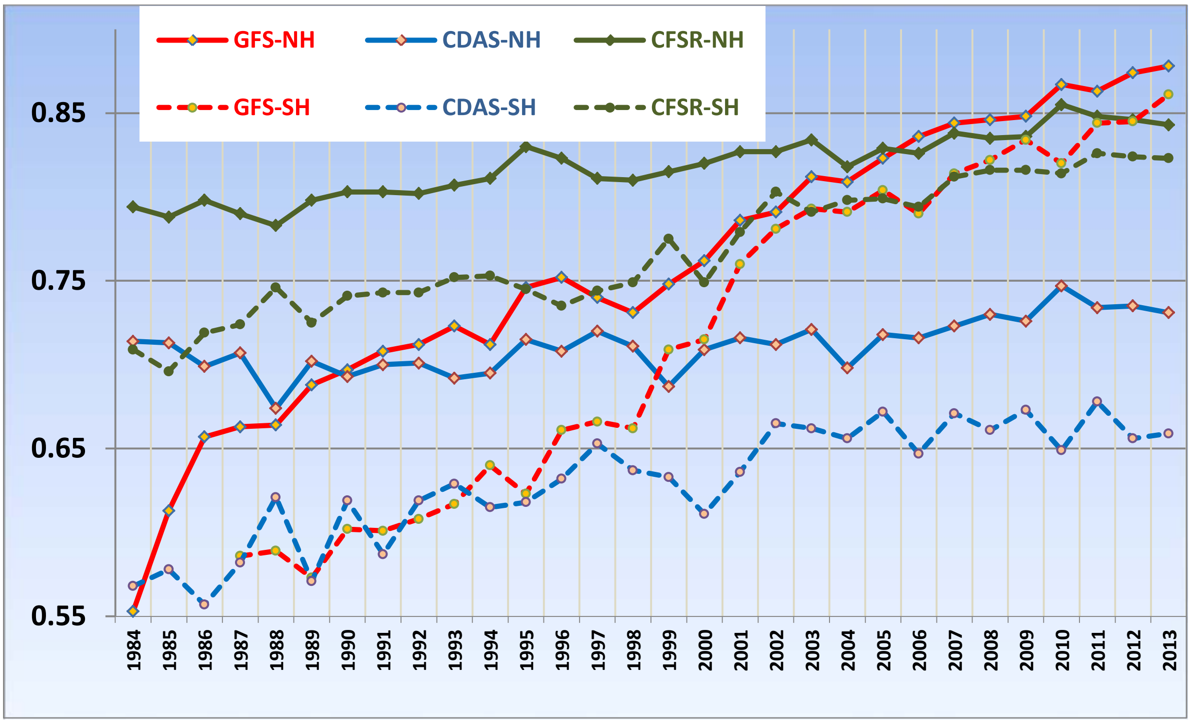

Figure 2: Same as Figure 1 except includes forecasts from CDAS (blue) and CFSR (green) reanalyses.

As shown in Figure 2, improvements in the system from the first reanalysis to the second reanalysis were dramatic. Note the continued increase in forecast skill after the second reanalysis. The evolution of the observing system over the period with a fixed assimilation/forecast system had a positive impact, but much less dramatic than that from the changes in the assimilation/forecast system. Thus, most of the improvement can be attributed to the changes in the assimilation/forecast system, not the observing system.

2 Data Assimilation

The development of data assimilation techniques has had a large impact on the use of satellite data. Data assimilation theory will be more completely reviewed by Lorenc later in this seminar. However, as an introduction, a few basic concepts are outlined here.

From a Bayesian perspective, we are looking for the probability of the atmospheric state given the observations and any prior information. To make a deterministic forecast, we want to find the most probable atmospheric state from this distribution and use it to initialize the forecast model. Then if the probability density functions (pdf's) are Gaussian, and there are no biases, it can be shown that the most probable atmospheric state can be found by minimizing the variational problem:

$$J[x] = ½(x-x_b)^T B^{-1} (x-x_b) + ½(y_o-H[x])^T (E+F)^{-1} (y_o-H[x]) \quad (1)$$

Where x is the analysis vector,
x_b is the background vector (usually a short-term forecast),
B is the background error covariance matrix,
y_o is the observation vector,
H is the forward operator transforming the analysis vector into the form of the observation vector, and
E and F are the instrument and representativeness error covariance matrices.

The two terms on the right hand side of the equation are often referred to as the background and observation terms, respectively. Additional terms can be included if more information is provided from other sources. For example, a constraint term can be included to penalize the lack of fit to certain known properties, e.g., moisture greater than zero, conservation of mass, etc. But for presentation simplicity, we will only consider the background and observation terms in this presentation.

Note that the same basic variational formulation is used both for atmospheric data assimilation and for physical retrievals from satellite radiances. However, there are some basic differences in the application. For atmospheric data assimilation, the problem is 3- or 4-dimensional, but retrievals are usually done in 1-D (along the path of the observations). Also, for atmospheric data assimilation, many different types of data are combined together to get the analysis (retrieval). However, because of the consistency in the formulation, if the background and background error are consistent, it can be shown that using retrievals and the retrieval error covariances in the atmospheric analysis would be identical to directly using the radiances in the analysis (if H is linear). This requires ensuring that all components of the two are consistent and passing to the analysis system both the n-component retrieval and the n x n retrieval error covariance matrix.

Focusing on the background term ($½(x-x_b)^T B^{-1} (x-x_b)$), several important details can be noted. The analysis variable (x) does not have to be the same as the model variables (or the observed variables). However, one must be able to transform the analysis variables into the model variables in order to start the forecast model. The background vector (x_b) contains all of the information extrapolated in time from previous analyses using the forecast model equations. This background vector contains a large amount of information (as much or more that the observations at any analysis time). With improvements in any previous analysis or in the forecast model, improvements in the background can be expected and will result in a feedback into an improved current analysis. See Mahfouf presentation for some additional details of the importance and future of the forecast modeling. The background error covariance (B) is extremely important for the quality of the analysis (Lorenc and Bormann presentations). B determines how the information in the observations is distributed spatially and between analysis variables. Considerable progress has been made in improving the specification of this matrix over the last 5 years, and most centers are now using techniques which make this matrix situational dependent (varying in time). Figures 3 and 4 show a comparison of the impact of a single temperature observation on the analysis given a static (unchanging in time) background error covariance (left panels) and a situation dependent background error (right panels).

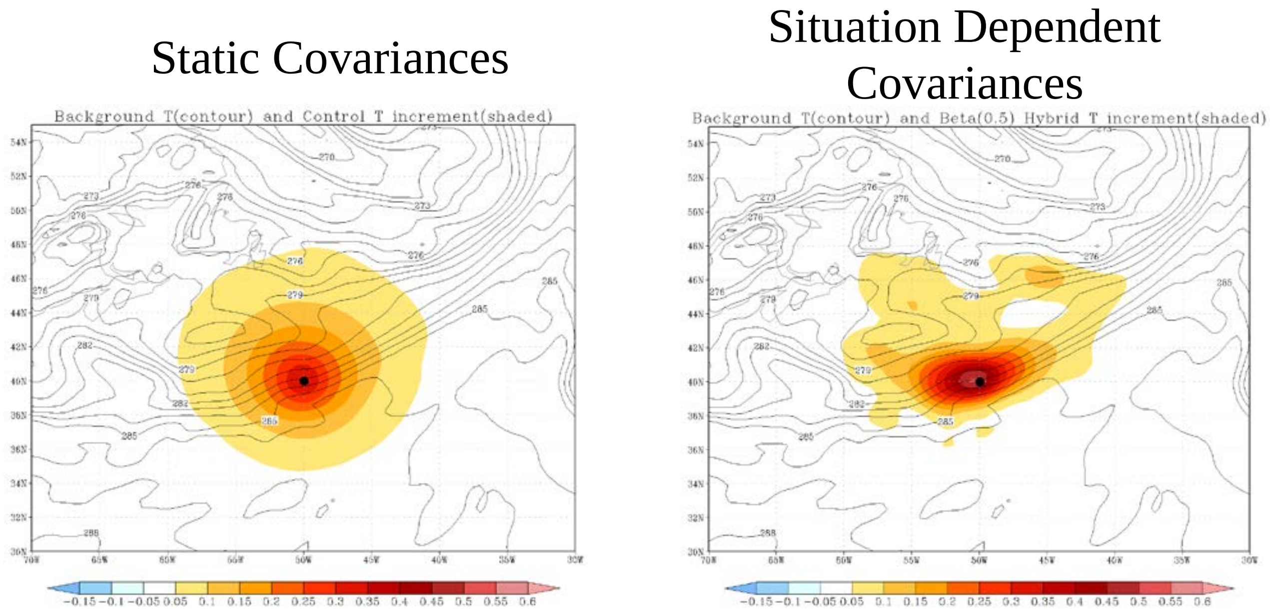

Figure 3: Analysis increment resulting from a single temperature observation (colored) superimposed on temperature field. Left panel -static covariances. Right panel - situation dependent covariances.

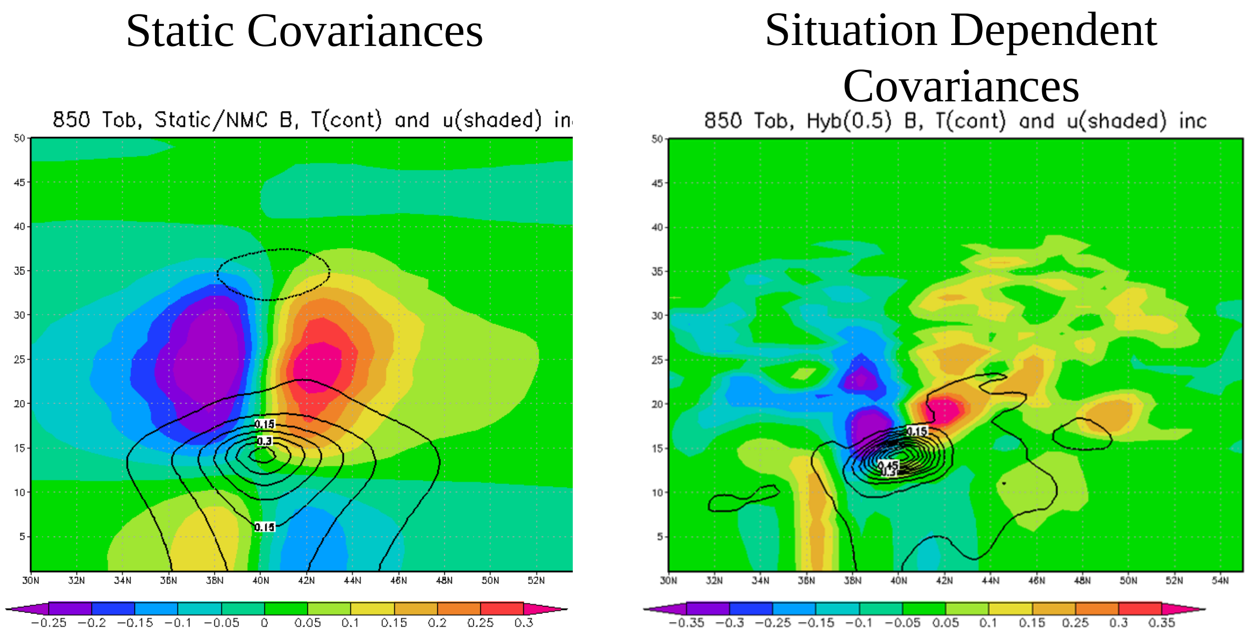

Figure 4: Same as Figure 3 except east-west vertical cross section through observation point. Shaded - Temperature increment. Contours - wind increment.

The second term of equation 1 is the observation term ($\frac{1}{2}(y_o\text{-}H[x])^T(E+F)^{-1}(y_o\text{-}H[x])$). The observation vector (y_o) contains all observations used in the analysis, and thus, all observations are used at the same time. The forward model (H) transforms the analysis variable into the same form as the observations. Since the observations contain the information to be used in the analysis, it is important to know the characteristics of the instruments and observations and to use this information in the forward model (and error characterization) to extract as much information out of the data as possible. There are several presentations at this seminar which specifically address instrument characteristics, forward models, and observational errors for various instruments. For satellite radiances, Collard, Bell, Geer, Kazumori, and Ruston made presentations on various satellite sounders with two additional presentations specific for the radiative transfer part of the forward model (Vidot and Karbou). In addition, Matricardi presented the use of the sounder radiances through principle components/reconstructed radiances. Non-sounder instruments had specific presentations including wind estimates (Forsythe), radar and lidar measurements (Janisková), GPS Radio Occultation (Healy), and Lidar winds (Rennie). Note that the H operator can also contain other components including a forecast model to extrapolate the information to the observation time. This is the basis of 4D-var systems.

The instrument error covariance (E) and the representativeness error covariance (F) determine how the observations are weighted within the analysis. See Bormann presentation in this seminar for details on the satellite radiances. The instrument error covariance specifies the actual errors in the observations. The representativeness error covariance specifies how well the observations are represented in the analysis. There is not a consensus into which term the forward model errors are included, but they must be included in one of the terms. An example of the representativeness error is during a 10 km global analysis, a radiosonde is launched into a 1 km thunderstorm. The 10 km analysis cannot represent the 1km thunderstorm, so there is a difference between the observation and the analysis resulting from neither an error in the observation or analysis themselves. Both instrument and representativeness error covariances can have correlated errors and can be situation dependent. In my opinion, improved specification of these errors is an under-developed part of the analysis problem.

3 Additional considerations for satellite observations

In addition to the components of the data assimilation discussed in the previous section which result directly from equation 1, there are further considerations which are quite important for the use of satellite data. These considerations are consistent with the theory but perhaps not immediately apparent from equation 1. Some components are necessary to satisfy the assumptions in the derivation of the variational problem (e.g., unbiased observations and background), and some are basic considerations of operational systems (e.g., choosing which observations will be used, computational resource restrictions, and monitoring of the data). All of these considerations can apply to all types of data, not just satellite data. However, because of the extreme volume and data characteristics, the considerations can become more important for satellite data than other types of data.

3.1 Bias correction

The basic assimilation theory assumes that the observations and the background are unbiased. When this is not true, the best solution is to remove the source of the bias whether it is from a model bias, forward model error, instrument characterization issue, or another source. Unfortunately, it is not trivial to remove or in fact determine the source of the bias since the truth is often not known. Thus, various bias correction schemes (some used as preprocessorsor as part of the variational problem system itself) that reduce the impact, but do not address the source of the error, have been developed.

Since the truth is not available, a proxy must be used. The observations are continually being compared to the background, and this difference will be an upper bound on the observational error unless the errors between the various components are strongly correlated. The sources of bias differences between observations and background can be from:

1. Inadequacies in the characterization of the instruments,
2. Deficiencies in the forward models,
3. Errors in processing data,
4. Errors in the background.

When there are bias differences between the background and observations, the biases resulting from the first three sources should be removed while the bias due to errors in the background should not be removed (otherwise the observations cannot correct the bias). Note that these biases can be quite complex and difficult to remove.

Figure 5 shows the real time mean biases between the satellite radiances for two legacy AMSU-A instruments as a function of scan position. Note two features of these figures. First, the biases are different for two realizations of the same instrument, indicating the bias correction is specific to the instrument, not the class of instruments. Second, there is a significant cross scan bias. Channels 7 and 8 should have very little impact from the surface, and there is no reason to suspect that the background would have a scan dependent bias. This bias is considerably larger than the signal in the data and must be accounted for in order to use the data.

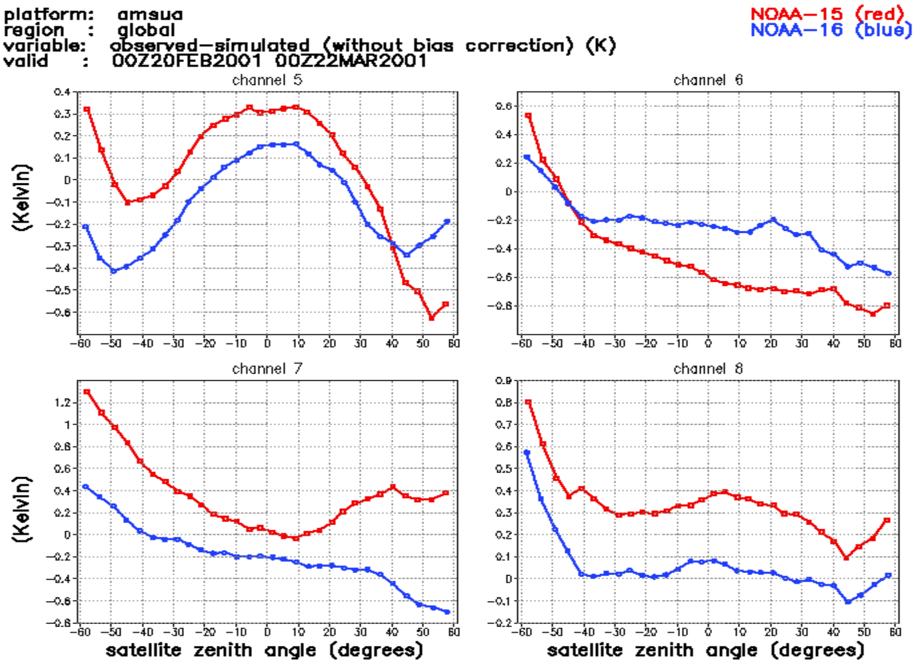

Figure 5: Scan dependent bias for AMSU-A on NOAA-15 and NOAA 16 for channels 5–8.

Figures 6 and 7 compare the standard deviation and bias before and after bias correction for channels 7 and 8 from AMSU-A on NOAA-18. Both the bias and standard deviation is significantly reduced. The standard deviation is reduced because the bias correction is situation dependent. Note that the bias does not go to zero. Unless the background is unbiased, the difference between the observations and background should not be zero. Comparing the bias in the un-bias corrected data to the standard deviation after bias correction demonstrates the importance of removing the bias before the information in the data can be used.

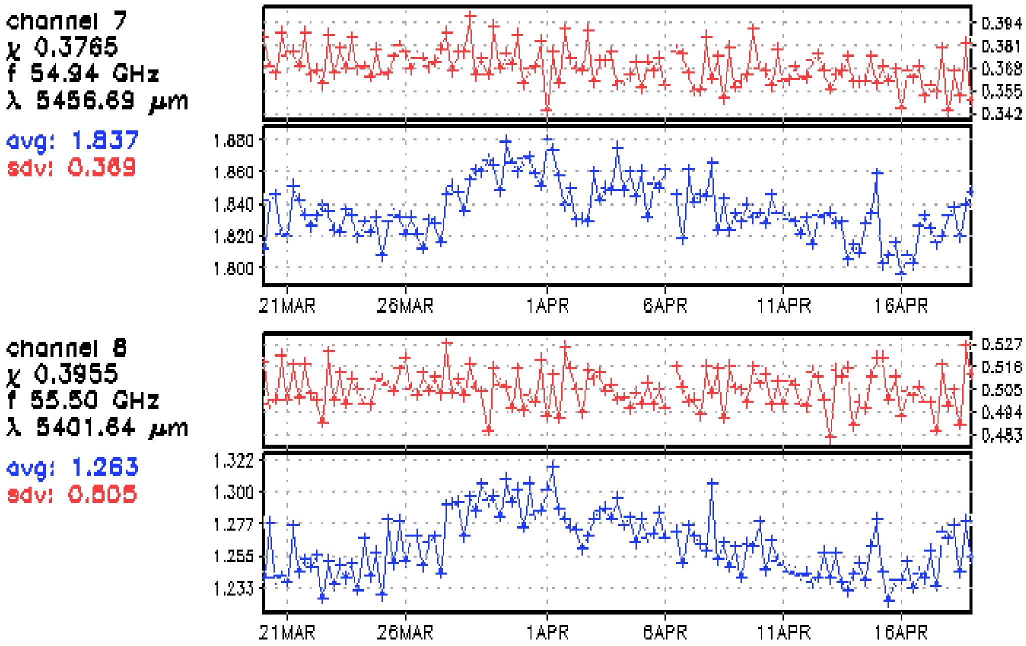

Figure 6: Standard Deviation and Bias between observations and simulated brightness temperatures from NOAA-18. No Bias correction. Upper panel - Channel 7. Lower panel - Channel 8.

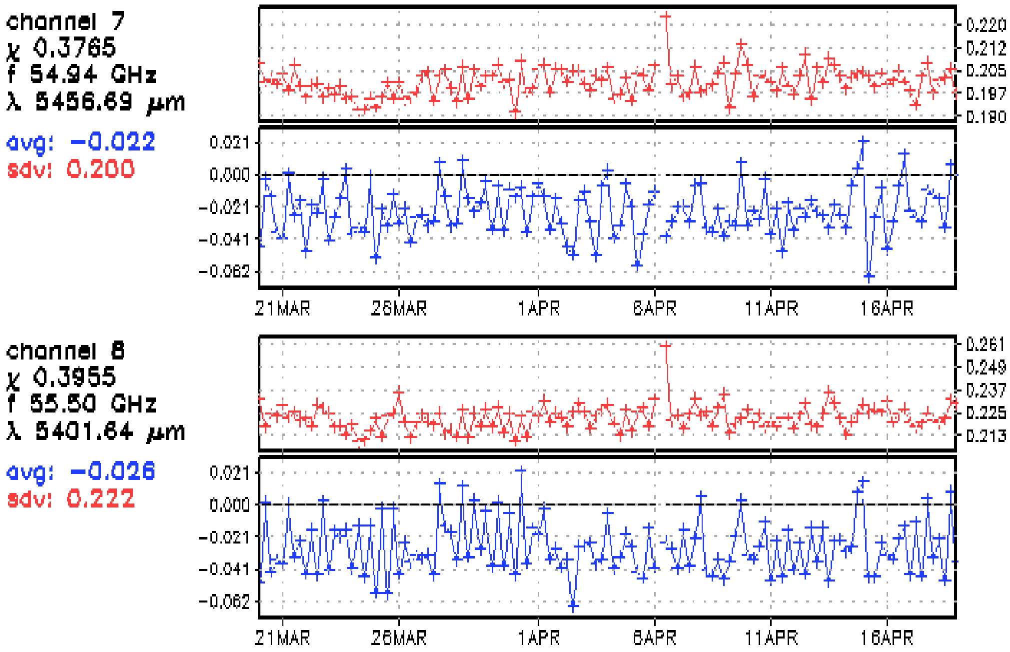

Figure 7: Same as Figure 6 except with Bias Correction.

3.2 Quality control

When performing atmospheric analyses, one does not necessarily want to use all observations. Observations having errors that make the error distribution non-Gaussian or observations not properly modelled can result in degraded analyses. The operational data assimilation systems tend to be conservative in their use of data because a few bad (or improperly modelled observations) can do more damage than many additional good observations. The effects are amplified since often these errors are strongly correlated (especially those due to forward model inadequacies). For example, if cloud effects are not properly incorporated into a radiative transfer model, the use of cloud impacted radiances are likely to result in highly correlated errors and a degraded analysis. Considerable effort is expended with each observation type developing appropriate quality control procedures.

3.3 Thinning and Super-obbing

Satellites can produce very large volumes of data. In theory, if observation errors are uncorrelated, using more observations should reduce the analysis error. However, observation errors are not always uncorrelated and computational considerations in data assimilation are important. To reduce computational costs, communication volumes and correlated errors, observations are often thinned or super-obbed. Note that the thinning or super-obbing of the observations can be done in several different ways (e.g., spatially, spectrally, temporally). Super-obbing is the combination of nearby observations to create a single more accurate observation. The use of principle components or reconstructed radiances (Matricardi) can be considered a type of super-obbing. To minimize the potential negative impact (or in the case of correlated error increase the positive impact) on the analysis of the thinning or super-obbing, the thinning and super-obbing algorithms are designed to retain as much independent information in the observations and reduce correlated error. To date many of the operational techniques are based on experience rather than a well-developed theory.

3.4 Data monitoring

When a new type of data is incorporated into an assimilation system, one of the first steps is to monitor the data by comparing it to what is produced by the forward model from the short term model forecasts. Generally, if the forward model is adequate, the differences are small, often only slightly larger than the expected observational error. Most operational centres continue generating these statistics after the new data is incorporated into the assimilation system. The statistics provide an excellent measure of the health of the observing system and the data assimilation system. When the statistics decrease, the changes can often be attributed to enhancements to the modelling/assimilation system. When the statistics increase, degradation in the instrument or in the modelling/assimilation system is usually the cause. For that reason, operational centres often maintain and exchange information from their monitoring systems. When an anomaly is noted, other operational systems are checked to determine if the anomaly is likely a result of instrument or data assimilations changes. Often the operational centres notice instrument problems prior to the data providers since the monitoring system results are often quite sensitive to small signals.

Figure 8 shows an example of the instrument issues that can be revealed by the monitoring. In this case, the problem was indicated beginning on July 2. After consulting other operational centres monitoring sites, it was decided that this was a significant instrument issue. NCEP stopped using the data on July 4. Other operational centres also removed the data close to this time. A filter wheel issue of reported by the provider, and the instrument was "fixed" on July 10. Note the large changes in the biases when the fix was applied. Due to the large change in bias, possibility of recurrence of the problem, age of the instrument, and redundancy of the observations, it was decided not to attempt to re-insert this data into our operational system.

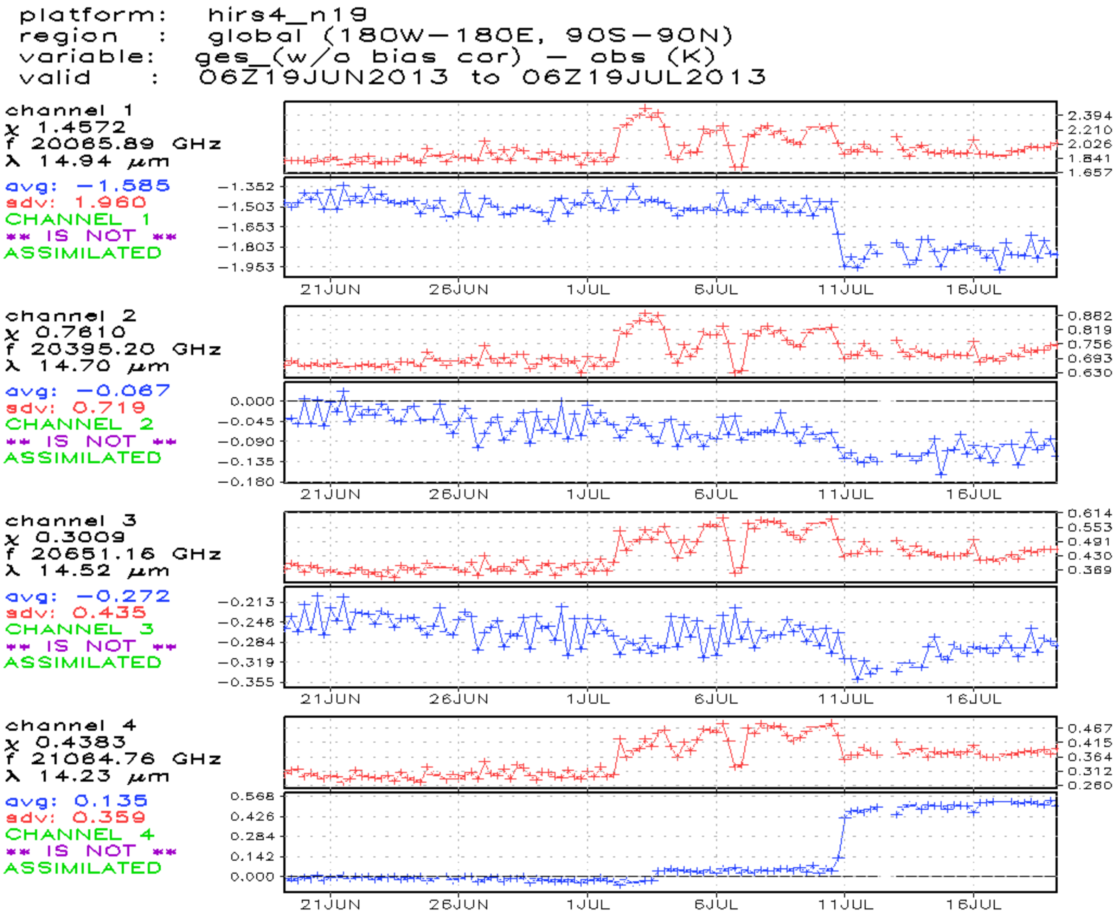

Figure 8: Time series of standard deviation and bias between observations and simulated observations for HIRS4 from NOAA-19.

4 Challenges and final comments

Despite the tremendous progress made in the last 20 years in data assimilation and the use of satellite data, many opportunities to improve the systems still exist. These opportunities include the areas of coupled assimilation, the improved use of the observational database, and improved techniques for assimilating the data. With all of these areas, it is important to note that "the devil is in the details". Small changes in parts of the system can have very large impacts (both positive and negative), and it is necessary to get everything "right".

One of the areas currently receiving more attention is coupled data assimilation, including coupling with aerosols, constituent gases, land, and oceans. Presentations on the subject of the assimilation of atmospheric composition (Elbern), land surface (Candy), and the ocean (Bertino) are part of the seminar. By coupling together different systems, it is anticipated that the observations which are sensitive to more than one system can be used better, and the interaction between the independent systems can be better modelled. For example, satellite radiance measurements are sensitive to aerosols, atmospheric constituents, land surface variables (e.g., temperature and moisture), and ocean surface variables (e.g., temperature, roughness, foam, etc.)If differences between the observation and background can be properly attributed to the component, the information should be more completely extracted from the observations. Without coupling, the signal from the uncoupled component can be improperly aliased into the solution.

Many challenges remain in the use of observations in the assimilation systems. More information can be extracted from observations currently being used by improving the forward model and by improving the specification of the observational error (correlated and uncorrelated components of both instrumental and representativeness). Very significant progress has been made in the last few years on the use of cloudy radiances, but much of the information in these data still remains unused. Observational data with useful information are not being used for a variety of reasons including real-time availability, insufficient forward models, instrument instability, and because sufficient resources have not been expended in developing the appropriate data assimilation components. With the many available sources of observations and the many planned future instruments (see presentations by Eyre, Klaes, and Goldberg), it has been difficult for operational centres to use all of the information that is available.

As the large scales become better defined, data assimilation development will shift to defining smaller scales. The future of data assimilation at these smaller scales is uncertain (see presentation by Auligné). As we get to smaller scales, the nonlinearity of the system becomes larger, and the relationship between the variables becomes less well defined. It is not clear that data assimilation techniques that work well at the large scale will work well at the smaller scales.

Atmospheric data assimilation systems were developed primarily for the purpose of initializing operational prediction systems. However, the use of the assimilation system for other purposes such as reanalysis (see presentation by Dee) and evaluating observational system component impact (see presentation by McNally) have increased. For reanalysis, the maintenance and recovery of information about the long-term observing system is extremely important. For evaluating observing system impacts, the evolution of the observing system and accounting for redundancies in the observing system present many challenges. For both reanalysis and observational impact, keeping up with the rapid advances in data assimilation and modelling remain a major challenge.

Despite the many recent improvements in the use of satellite data and data assimilation systems, the future promises many additional enhancements. In this presentation, I have attempted to give an overview of the problem and show the relationships to other presentations at this seminar. The assimilation of information in satellite data requires all components of the assimilation system to be designed and implemented in the best possible way. The final (and perhaps the most difficult) challenge of data assimilation and the use of satellite data is to prioritize among the many potential directions and only limited resources.

References

Eyre, J.R., 2007: Progress achieved on assimilation of satellite data in numerical weather prediction over the last 30 years. *ECMWF Seminar on Recent Developments in the Use of Satellite Observations in Numerical Weather Prediction,* 3–7 September 2007, ECMWF, 27 pp.