Atmospheric Observations and Experiments to Assess Their Usefulness in Data Assimilation

By Robert Atlas

Data Assimilation Office, Laboratory for Atmospheres, Code 910.4,
NASA Goddard Space Flight Center, Greenbelt, MD 20771

(Manuscript received 14 June 1995, in revised form 28 February 1996)

Abstract

Atmospheric observations consist of a mixture of in situ, visual, and remotely sensed observations. These provide an extensive database for research and numerical weather prediction. However, significant data deficiencies still exist, and new observing systems are continually being proposed. Observing system experiments (OSE's) are conducted to assess the usefulness of different types of existing atmospheric observations. Observing system simulation experiments (OSSE's) are conducted to evaluate the potential impact of proposed observing systems, as well as to determine tradeoffs in their design, and to evaluate data assimilation methodology.

This paper contains a review of the development of the global atmospheric observing system, a description of the principal types of data, an overview of OSE and OSSE methodology, and results from recent experiments to evaluate the relative utility of the principal atmospheric observing systems and the potential for new observing systems. These experiments show the critical contributions being made by both conventional and space-based observations, and indicate considerable potential for future satellite observing systems to improve data assimilation.

1. Introduction

There can be little doubt, even among the typically skeptical users of weather information, that the science of meteorology and the practice of weather forecasting have improved immensely over the last century. The improvements that have occurred are linked to very significant increases in our ability to observe the global atmosphere and in our computing capability over the last several decades. As a result of these advances, there has been a significant improvement in our understanding of atmospheric processes, and increasingly sophisticated models of the global atmosphere have been developed for both research and weather prediction.

At the present time, atmospheric observations include: direct (in situ) measurements of the primary atmospheric variables including wind velocity, pressure, temperature, and humidity; visual observations of clouds, visibility, and type of precipitation; and remotely sensed observations of temperature, moisture, clouds, and wind. This combination of observations provides an extensive data base for both initializing and verifying numerical weather prediction (NWP) models and for research to further our understanding of the atmosphere and its role in the Earth-Atmosphere-Ocean System. Nevertheless, significant data deficiencies still remain and new observing systems are continually proposed to improve the accuracy of analysis and forecast products.

Two types of experiments related to the effectiveness of atmospheric observations in data assimilation are described in this paper. These are: Observing System Experiments (OSE's), which are conducted in order to evaluate the impact of a given observing system in data assimilation; and Observing System Simulation Experiments (OSSE's), which are conducted to assess the potential impact of a new observing system. Following a brief history of meteorological observations and their utilization in synoptic analysis (in Section 2), a summary of the current World Weather Watch (WWW) conventional observing network and available satellite observations will be given (in Section 3). A description of OSE and OSSE methodology and recent experiments that have been conducted to assess the utility of different components of the WWW and the potential for new observing systems to improve data assimilation will be presented in Section 4. Finally, conclusions, and a brief look at some of the new
space-based observations that are expected to be available in the near future will be given in Section 5.

2. Historical background

The current observing system for the atmosphere developed over the last several hundred years. While visual observation of the atmosphere probably began with the earliest people on Earth, and early versions of simple meteorological instruments for qualitative observation of the atmosphere appeared more than two thousand years ago, systematic study of the atmosphere did not begin until instruments for measuring atmospheric properties were developed. (See Middleton, 1969, and Frisinger, 1977, for a detailed history of the invention of the standard meteorological instruments.)

The earliest meteorological instrument for quantitative observation, the rain gauge, was first used in Korea around 1440 (Godske et al., 1957). By the mid 1700's, all of the main surface meteorological instruments, the thermometer, barometer, hygrometer, and anemometer, were in use. Over the next hundred years these instruments were improved substantially and by the mid 1800's, with the invention of the telegraph, meaningful observing networks and national weather forecasting services were inaugurated. By the beginning of the 20th century a very limited global observing network was in place (Daley, 1991).

Most of the observations that were available at that time were poorly organized surface observations of pressure, temperature, wind, humidity, clouds, and precipitation. Nevertheless, significant advances in our ability to describe and predict the atmosphere were achieved. Prior to the invention of the telegraph, the only approach for studying and predicting the weather was the “local method” (Godske et al., 1957). Many important findings about the atmosphere were deduced from local observations only (Godske et al., 1957; Kutzbach, 1979). For example, H.W. Dove in 1840 was able to identify polar and equatorial currents of air, and formulate the hypothesis that the interaction of these currents would lead to the development of storms.

The “synoptic method” of using weather maps for the study of the atmosphere and for weather forecasting began with the transmission and plotting of weather observations. This led to the establishment of national weather forecasting services and to the development of increasingly sophisticated synoptic models (Bergeron, 1959). Fitzroy in 1863 developed a model of cyclones, which showed the interaction of the polar and equatorial currents that had been identified by Dove. Abercromby in 1887 and Lempfert and Shaw in 1906 developed useful models of cyclone characteristics. Further development of synoptic models occurred in 1918 with the establishment of a dense observational network in Norway by Vilhelm Bjerknes. This led to the discovery and classification of fronts and air masses and to the Polar Front Theory of Extratropical Cyclones by the “Bergen School” of meteorologists (Bergeron, 1959).

The application of the Bergen models resulted in significant advances in weather forecasting accuracy. Direct aerological (upper-air) measurements began in 1749 when a thermometer attached to a kite was sent aloft (Lally, 1985). The invention of the hot air balloon in 1783 and the box kite (which could carry meteorographs aloft) in 1893 greatly increased observational capabilities. However, upper-air observations in the early 1900’s were still extremely sporadic and were limited to indirect aerology (inferences from surface observations of clouds and their motion) and to the occasional ascents by kites, balloons, and instrumented aircraft.

The invention of the radiosonde in 1928 provided a vital new tool for observing the three-dimensional structure of the atmosphere. With the impetus of World War II, global upper-air observations were added to the surface observing network, and the routine analysis of upper-air weather maps began. These data provided the phenomenological basis for numerous advances in our understanding of atmospheric processes and led to the application of vorticity principles and jet stream dynamics to weather forecasting. These advances in conjunction with the development of the high speed electronic computer enabled the beginning of operational numerical weather prediction in the late 1950’s. However, the concentration of these “conventional” synoptic observations over land and primarily in the Northern Hemisphere resulted in large data void and data sparse regions of the globe. This limited both the accuracy and useful range of weather forecasts significantly.

The advent of meteorological satellites in the 1960’s provided an effective means to supplement the conventional observing network by providing observations, not only in the traditionally data-poor regions of the oceans, the Southern Hemisphere and stratosphere, but also at a higher horizontal resolution than the conventional surface-based observations. The first weather satellite, TIROS I, was launched in 1960. This and subsequent weather satellites provided images of cloud cover, which were used to improve analyses of pressure systems and fronts, and to make inferences of atmospheric stability, wind, moisture, and precipitation (Vaughan and Johnson, 1994). Quantitative data from satellites began in the late 1960’s with experimental temperature soundings derived from the radiances measured by polar orbiting satellites, and wind velocities determined from the motions of clouds seen in geostationary satellite images (Johnson, 1994). Since the 1970’s, satellite temperature soundings and cloud
drift winds have improved substantially in accuracy, and these data have been important components of the WWW Global Observing System (GOS). Additional types of satellite data have also become available in recent years and many new types of space-based observations are planned as part of the Earth Observing System (EOS)\(^1\).

### 3. The World Weather Watch

The WWW was adopted as a concept by the Fourth World Meteorological Congress in 1963. Its objectives, as stated in WMO Report No. 790 (1993) are:

1. To maintain an effective world-wide integrated system for the collection, processing, and rapid exchange of meteorological and related environmental data, analyses, and forecasts;

2. To make available, both in real time and historical archives, as appropriate, observational data, analyses, forecasts, and other products to meet the needs of all World Meteorological Organization (WMO) Members and Programmes, and of relevant programmes of other international organizations;

3. To arrange for the introduction of standard methods and technology which enable WMO Members to make best use of the WWW system and ensure an adequate level of services, and also the compatibility of systems for cooperation with agencies outside the WMO.

The WWW functions on three levels: global, regional, and national. It consists of three core elements:

- The Global Observing System (GOS), consisting of facilities and arrangements for making observations at stations on land and at sea, and from aircraft, environmental observation satellites and other platforms;

- The Global Telecommunication System (GTS), composed of an increasingly automated network of telecommunication facilities for the rapid, reliable collection and distribution of observational data and processed information;

- The Global Data-processing System (GDPS), consisting of World, Regional/Specialized and National Meteorological Centres to provide processed data, analyses, and forecast products.

\(^1\) EOS is the principal observational component of NASA’s “Mission to Planet Earth”, and will involve a series of satellites to provide coordinated measurements of the Earth’s surface, atmosphere and oceans (King et al., 1995).

In this paper, we will be concentrating on only the first of these elements. At the present time, the main components of the GOS are:

1. **SYNOP** (land) and **SHIP** (ocean) surface observations. These are composite observations that include temperature, pressure, humidity, wind velocity, clouds, and precipitation that are reported every 3 hours. Temperature is measured by hygrothermographs (accurate to \(+/- 1.0\) K) and to a lesser extent by mercury thermometers (accurate to \(+/- 0.1\) K). Surface pressure, measured by barographs, is generally accurate to 0.5 hPa, but reductions of pressure to sea level using local temperature and elevation result in errors of 1-3 hPa. Humidity is measured by hygrothermographs, wet and dry bulb thermometers, and hygrometers of varying design. There is typically a \(+/- 3\) % error for humidities of 20-80 %; 5 % or larger for the extremes of humidity. Wind velocity measured by wind vanes and anemometers is typically accurate to 10 degrees in direction and 3 m/sec in speed. In addition to instrument errors, there is an error of representativeness, that measures the ability of observations to represent particular scales of atmospheric phenomena. For example, surface wind observations over land are strongly influenced by local factors (small scale eddies from local obstructions, strong static stability, and local circulations). As a result, these observations may not be representative of synoptic scale motion. Wind measurements on moving ships are generally more representative, but less accurate due to ship motion and obstructions to the wind flow. In addition, these observations are concentrated along predetermined shipping routes and avoid significant storms whenever possible.

2. **TEMP** upper air observations of temperature (geopotential height), wind velocity, and humidity by radiosondes and rawinsondes. These observations are taken 2-4 times daily, mainly at 00 and 12 UTC, with a few at 06 and 18 UTC. These data are obtained mostly over continents and primarily in the Northern Hemisphere. Observations are recorded at preselected pressure levels referred to as “mandatory levels”. In addition, “significant level” data are recorded whenever the changes in the vertical exceed a prespecified limit. At the present time, there are approximately ten different manufacturers of radiosondes. For sensing temperature, European sondes previously used bimetallic elements, which require a radiation correction. The more recent approach uses thermocaps to relate temperature change to capacitance changes. In the United States, rod or flake thermistors are used. Errors for this instrument are about 1 K and tend to be random below 25 km. At higher levels, radiation errors of 1-2 K occur. Height errors increase with altitude due to both the accumulation of temperature errors during hydrostatic integration and to increasing er-
rors in placing the balloon as it rises. Height errors are 10–15 m at low levels, and 25–40 m in the upper troposphere. Wind is obtained by tracking balloon displacements; accuracies are about 3 m/sec in the lower troposphere, 5 m/sec in the upper troposphere (larger near strong jet streams). Humidity is measured by hygrometers in the U.S. and humicaps in Europe with errors of about 5 % under optimal conditions. (For a detailed discussion of the errors associated with radiosondes see Nash, 1993, and Schmidlin, 1993).

3. PILOT wind observations obtained by the manual tracking of pilot balloons. These are wind velocity profiles for the lower troposphere. The accuracy of these observations is slightly poorer than for rawinsondes due to errors in accounting for the ascent speeds of the balloons. In addition, the balloons cannot be tracked in clouds or fog.

4. AIREP manually reported and automatically reported aircraft observations. These are asynoptic observations of temperature and wind primarily near 200–300 hPa (cruising altitude). The accuracy is 0.5 K for temperature and 1–2 m/sec for wind.

5. SATEM temperature soundings that are retrieved from the radiances measured by polar orbiting satellites. Infrared and microwave radiometers aboard two NOAA polar orbiting satellites are used to observe radiances in 26 spectral channels. The radiance data are then converted to temperature profiles using a physically-based retrieval procedure at the present time. Global coverage is achieved by each satellite every twelve hours. The data produced are asynoptic layer mean temperatures at 100 km (or higher) horizontal resolution. The accuracy of mean temperatures between mandatory levels is about 2 K, but varies considerably with altitude, geographical location, and synoptic situation. Errors in the retrievals are horizontally correlated (Sullivan et al., 1993).

As an alternative to the use of satellite temperature retrievals, the radiances (either raw or cloud corrected) may be assimilated directly. The direct assimilation of cloud-corrected radiances using a one-dimensional variational approach was implemented at the European Centre for Medium-range Weather Forecasts (ECMWF) in 1993 (Eyre et al., 1993) and with a three-dimensional variational approach at the National Centers for Environmental Prediction (NCEP) in 1995 (J. Derber, personal communication, 1995).

6. SATOB single level wind velocity data inferred from the cloud motions detected by geostationary satellite imagery. The technology for generating these winds has evolved substantially. Originally, the movement of cloud elements was tracked manually. This was time consuming, difficult, and often inaccurate; with the major problems being the level assignment for the winds and errors due to cloud dynamics. More objective procedures for tracking, based on correlation techniques, are now being used (LeMarshall et al., 1994). Level assignment has improved somewhat, but is still a significant source of error. Coverage is nearly global between 60N and 60S; however sampling can still be a problem and result in a low wind speed bias for jet streams (Atlas, 1987; Kallberg and Delsol, 1985). Cloud drift winds are most accurate in the Tropics, with their accuracy ranging from 3 m/sec at low levels to 6–8 m/sec at upper levels.

Global coverage maps of each of the above observing systems for the 6-hour period centered at 00 UTC March 3, 1993 are shown in Figs. 1a–f. These figures illustrate the heterogeneity of the GOS as well as the typical coverage of each data type. Table 1 summarizes the amount of each data type and its percentage at this time. In addition to the above data, various other types of surface-based and space-based data are currently available for assimilation. These include: synthetic observations of sea level pressure, humidity or wind determined from geostationary satellite imagery (e.g., PAOB sea level pressure observations which cover most of the Southern Hemisphere oceans); sea level pressure and wind observations from moored and drifting buoys; mesoscale observations of precipitation and/or wind from ground-based radar and lidar; surface wind speed and vertically integrated water vapor from the Special Sensor Microwave Imager (SSM/I); and surface wind velocity data from the ERS-1 scatterometer. Both SSM/I and ERS-I provide very high resolution data over the global oceans, as shown in Figs. 1g and 1h.

4. Data impact studies related to the global observing system

Since the advent of meteorological satellites in the 1960’s, numerous experiments have been conducted in order to evaluate the impact of these and other data on atmospheric analysis and prediction. Such studies have included both OSE’s and OSSE’s. The OSE’s were conducted to evaluate the impact of specific observations or classes of observations on analyses and forecasts. Such experiments have been performed for selected types of conventional data and for various satellite data sets as they became available. (See for example the 1989 ECMWF/EUMETSAT workshop proceedings on “The use of satellite data in operational numerical weather prediction” and the references contained therein.) The OSSE’s were conducted to evaluate the potential for future observing systems to improve NWP and to plan for the Global Weather Experiment and more recently for EOS (Atlas et al., 1985a; Arnold and Dey, 1986; Hoffman et al., 1990). In addition, OSSE’s have been run to evaluate trade-offs in the design of observing systems and observ-
Fig. 1. Global coverage maps for 00 UTC March 3, 1993 +/- 3 hours.


4.1 Experimental Design

Although there are many possibilities for how an OSE may be conducted, the most typical procedure is as follows: First a "Control" data assimilation cycle is performed. This is followed by one or more experimental assimilations in which a particular type of data (or specific observations) are either withheld or added to the Control. Forecasts are then generated from both the Control and experimental assimilations every few days (to achieve relative independence of the forecast sample). The analyses and forecasts from each assimilation are then verified and compared in order to determine the impact of each data type being evaluated.

Experiments performed in this manner provide a quantitative assessment of the value of a selected type of data to the specific data assimilation system (DAS) that was used. In addition, the OSE also
provides useful information on the effectiveness of the DAS. This information can be used to improve the utilization of this and other data in the DAS, as well as to determine the value of the data.

The methodology currently used for OSSE’s is very similar to that described above for OSE’s. However, this methodology has evolved considerably since these experiments were first carried out in the 1950’s and 60’s (Arnold and DeY, 1986). The earliest simulation studies proceeded according to the following sequence of steps. First, an artificial history of the atmosphere was created by numerical integration of a model. Second, simulated “data” were created from the history by the addition of random variations to the history values for temperature, wind, and pressure. Third, the numerical integration that created the history was repeated, but with the meteorological variables in the model replaced by the simulated data at locations and times corresponding to the assumed pattern of observations.

OSSE’s of this type were conducted by Char-
ney et al. (1969), Jastrow and Halem (1970, 1973), Williamson and Kasahara (1971), Kasahara (1972), Gordon et al., (1972), and others in preparation for the Global Weather Experiment. These studies provided an analysis of the Global Atmospheric Research Program (GARP) data requirements, the "useful" range of predictability, and the need for reference level data. From the results, it was concluded that the assimilation of satellite-derived temperature profiles meeting the GARP data specifications would yield a substantial improvement to the accuracy of numerical weather forecasts. A later study by Cane et al. (1981) using a modified OSSE procedure indicated similar potential for satellite surface wind data.

An examination of the underlying rationale for the early simulation studies (Jastrow and Halem, 1973), as well as a comparison of the results of the above studies with the results of subsequent real data impact tests (e.g., Halem et al., 1982; Baker et al., 1984) indicated several important limitations. The most important weakness stems from the fact that the same numerical model was used both to generate the simulated observations and to test the effectiveness of these observations. (This is referred to as the "identical twin" problem.) Other weaknesses relate to the model-dependence of the studies and the specification of observational errors as random. These weaknesses can lead to overly optimistic results and incorrect conclusions from an OSSE.

The current methodology used for OSSE's was designed to increase the realism and usefulness of such experiments (Atlas et al., 1985b). In essence, the analysis/forecast simulation system (shown schematically in Fig. 2) consists of the following elements:

1. A long atmospheric model integration using a very high resolution "state of the art" numerical model to provide a complete record of the assumed "true" state of the atmosphere (referred to as the "nature run" or "reference atmosphere"). For the OSSE to be meaningful, it is essential that the nature run be realistic, i.e., possess a model climatology, average storm tracks, etc. that agrees with observations to within prespecified limits.

2. Simulated conventional and space-based observations from the nature run. All of the observations should be simulated with observed (or expected) coverages, resolution, and accuracy. In addition, bias and horizontal and vertical correlations of er-
Table 1.

<table>
<thead>
<tr>
<th>Principal components of the GOS at 0000 UTC March 3, 1993 ±3 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a. Conventional data</strong></td>
</tr>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>Surface Land Reports</td>
</tr>
<tr>
<td>Ship Reports</td>
</tr>
<tr>
<td>Buoys</td>
</tr>
<tr>
<td>Rawinsondes</td>
</tr>
<tr>
<td>Pilot Balloons</td>
</tr>
<tr>
<td>Aircraft</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
</tr>
<tr>
<td><strong>b. Satellite Temperature Soundings</strong></td>
</tr>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>Clear Retrievals</td>
</tr>
<tr>
<td>Partly Cloudy Retrievals</td>
</tr>
<tr>
<td>Cloudy Retrievals</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
</tr>
<tr>
<td><strong>c. Cloud-Drift Winds</strong></td>
</tr>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>High Level</td>
</tr>
<tr>
<td>Low Level</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
</tr>
<tr>
<td><strong>d. All</strong></td>
</tr>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>Conventional Observations</td>
</tr>
<tr>
<td>Satellite Temperature Soundings</td>
</tr>
<tr>
<td>Cloud-Drift Winds</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
</tr>
</tbody>
</table>

Fig. 2. Schematic of the simulation system.

Errors with each other and with the synoptic situation should be introduced appropriately. Two approaches have been used for this purpose (Atlas et al., 1985b; Hoffman et al., 1990). The simpler approach is to interpolate the nature run values to the observation locations and then add appropriate errors. The more complicated (and expensive) approach is to attempt to retrieve observations from the nature run in the same way as observations are retrieved from the real atmosphere.

3. Control and experimental data assimilation cycles. These are identical to the assimilation cycles in an OSE except that only simulated data are assimilated. In order to avoid the identical twin problem, a different model from that used to generate the nature run is used for assimilation and forecasting. Typically this model has less accuracy and resolution than the nature model. Ideally, the differences between the assimilation and nature models should approximate the differences between a "state of the
art" model and the real atmosphere.

4. Forecasts produced from the Control and Experimental assimilations. As with the OSE's, forecasts are generated every few days to develop an independent sample. The analyses and forecasts are then verified against the nature run to obtain a quantitative estimate of the impact of proposed observing systems and the expected accuracies of the analysis and forecast products that incorporate the new data.

An important component of the OSSE that improves the interpretation of results is validation against a corresponding OSE. In this regard, the accuracy of analyses and forecasts and the impact of already existing observing systems in simulation is compared with the corresponding accuracies and data impacts in the real world. Ideally, both the simulated and real results should be similar. Under these conditions, no calibration is necessary and the OSSE results may be interpreted directly. If this is not the case, then calibration of the OSSE results can be attempted by determining the constant of proportionality between the OSE and OSSE impact as described by Hoffman et al. (1990).

4.2 Results of recent observing system experiments

Most of the OSE's that have been conducted in recent years have been primarily concerned with the impact of space-based observations. In general, the satellite temperature soundings and cloud drift wind data have been found to be essential components of the GOS in the Southern Hemisphere, but have shown much smaller impact and mixed results in the Northern Hemisphere extratropics (Halem et al., 1982; Mo et al., 1995)\(^2\). Experiments with SSM/I wind speeds have demonstrated significant improvements to ocean surface wind analyses in the Tropics and Southern Hemisphere (Atlas et al., 1991,1993; Liu et al., 1993). Experiments to evaluate the usefulness of ERS-1 winds have shown mixed results. Hoffman (1993) and Stoffelen et al. (1994) reported on the impact of ERS-1 winds on the ECMWF Global Data Assimilation System (GDAS). These studies, found substantial modifications to surface wind analyses, but no consistent improvement in forecast accuracy. Bell (1994) using the UK Meteorological Office (UKMO) GDAS obtained substantial reductions in forecast errors in the Southern Hemisphere when ERS-1 wind vectors were assimilated, but a later study with this system showed less impact (R. Graham, personal communication, 1995). Finally, Atlas et al. (1995) using the Goddard EOS (GEOS) GDAS (Schubert et al., 1993) and in a later study using the NCEP GDAS demonstrated a significant improvement in Southern Hemisphere analyses and forecasts when ERS-1 data were included.

In addition to the testing of new space-based data types as they become available, it is important to evaluate the relative contributions of each of the components of the GOS to four-dimensional data assimilation. Table 2 summarizes an OSE that was conducted for July 1993 using the 4° latitude by 5° longitude, fourth-order version of the GEOS GDAS. The thirty-day Control assimilation used all conventional data, satellite temperature soundings, cloud drift winds, and PAOBS. Assimilation experiments were then run in which either ships and buoys, rawinsondes and pilot balloons, aircraft, SATEM, SATOB, or PAOB data were withheld from the assimilation. Six 5-day forecasts from each assimilation were generated and then verified against ECMWF analyses. It should be noted that this version of the GEOS GDAS has less forecast skill than higher resolution forecast systems in place at a number of operational centers\(^3\), and that it is not possi-

---

2 Recent results from the direct assimilation of radiances have shown improved impacts relative to the assimilation of TATEMS in both hemispheres (J. Derber, personal communication, 1995).

3 A documentation of the performance of the GEOS GDAS is contained in Schubert and Rood (1995). In addition, direct comparisons between the GEOS and NCEP models...
Fig. 3. (a) Relative accuracy of the average of 6 forecast cases from Control (heavy solid line), Control minus SATOBS (light solid line), Control minus SATEMS (light dashed line), Control minus PAOBS (heavy dashed line). (b) Relative accuracy of the average of 6 forecast cases from Control (heavy solid line), Control minus Aircraft (light solid line), Control minus Ships and Buoys (light dashed line), Control minus rawinsondes and pilot balloons (heavy dashed line).

able to infer what the results for a substantially different data assimilation system will be from this experiment. Nevertheless, documenting the impact of various data within a given DAS is extremely valuable, and comparison with similar studies for other DAS can eliminate the "system dependency" of the results.

Figures 3a and 3b summarize the forecast results of this experiment in terms of the average anomaly correlation over the six forecast cases for the Northern Hemisphere (NH) and Southern Hemisphere (SH) Extratropics. These figures show that in the GEOS GDAS: (1) SATEM's have a very significant positive impact (i.e. higher accuracy when these data are included) in the SH and a slight positive impact in the NH. (2) SATOBS have a positive impact (about half that of the SATEM's) in the SH and a small negative impact in the NH. (3) PAOBS have a small positive impact (about half that of SATOBS, except at day 5) in the SH and a negligible effect in the NH. (4) Aircraft data have a nearly negligible impact on GEOS forecasts in both hemispheres. (5) Ships and buoys have a small positive impact in both hemispheres (comparable to PAOBS in the SH and SATEM's in the NH). (6) Rawinsondes have a significant positive impact in both hemispheres. In the NH, removal of rawinsondes and pilot balloon observations results in a 36 hour reduction in forecast skill. In the SH, the combined effect of these observations is nearly three-fourths as large as the effect of the SATEM's.

A synoptic evaluation of each of the forecast cases was performed in order to determine the meteorological significance of the above data impacts. This included both subjective comparisons of a variety of prognostic fields and an objective cyclone tracking verification of all of the forecasts. For this pur-
pose, the error in cyclone position for each of the forecasts relative to the ECMWF analysis was computed. The average error for the Control forecast was then subtracted from the average error for each of the experiment forecasts. Figure 4 presents the average impact on cyclone position error over the six 5-day forecasts for each of the experiments, where positive values represent an increase associated with removal of the indicated type of data. This figure shows that in the Southern Hemisphere a very significant increase in position errors results from the removal of either SATEMS or rawinsondes. A much smaller, but still significant degradation results from the removal of each of the other data types. In the Northern Hemisphere, the impact of data removal on cyclones is much smaller. Only the removal of the rawinsondes results in greater than a 50 km increase in cyclone position error. There is a slight beneficial impact of SATEMS in the Northern Hemisphere, and all of the other data types have either a negligible or slightly negative impact.

As an illustration of the impact on cyclone prediction, Fig. 5 shows the 60 hour Control and Control-SATEMS sea level pressure forecasts from 00 UTC 21 July 1993 and the corresponding ECMWF analysis for verification. Very large differences between the two forecasts can be seen over much of the region shown. In particular, the removal of SATEMS results in an 18 hPa increase in pressure south of Australia where a cyclone had been excessively weakened, and a 30 hPa decrease in pressure associated with the development of a spurious cyclone near 58S, 45E. These prognostic impacts are traceable to specific initial state differences.

In general, the analysis differences are smaller than the forecast differences, but they are still substantial. Figure 6 presents 500 hPa height rms analysis differences between each of the experiments and the Control for the entire period of data assimilation, while Fig. 7 illustrates these differences at the end of the assimilation cycle. In agreement with the prognostic results, both of these figures show a very significant impact of SATEMS in the Southern Hemisphere and of rawinsondes in the Northern Hemisphere. The impact of the other data types investigated is smaller, with the aircraft data showing...
a nearly negligible impact. It should be noted that these results refer only to the use of these data in the GEOS DAS and may differ substantially for other data assimilation systems and for different seasons. For example, Graham (1994) has found a significant beneficial impact of aircraft data in selected cases in January 1994 using the UK Meteorological Office DAS, and Baede et al. (1987) demonstrated the usefulness of aircraft data in the absence of SATEMS.

4.3 Results of recent observing system simulation experiments

A collaborative effort between the European Centre for Medium Range Weather Forecasts (ECMWF), the National Meteorological Center (NMC) and the Goddard Laboratory for Atmospheres (GLA) was initiated in the mid 1980’s in order to provide a quantitative assessment of the potential impact of proposed observing systems on data assimilation and numerical weather prediction (Arnold and Dey, 1986). In the first set of experiments to be completed (Atlas et al., 1985b), the identical twin problem was avoided by using the 1.875° × 1.875° × 15 level ECMWF model to generate a 20-day nature run and the 4° × 5° × 9 level version of the GLA fourth-order model for assimilation and forecasting. NMC simulated observations at their expected locations with estimated errors.

Both real data and simulated data assimilation cycles were performed for the period from 00 UTC 10 November to 00 UTC 25 November 1979. The real data experiments included a Control cycle in which only conventional data were assimilated, and a FGGE cycle in which conventional and special FGGE data sets including TIROS-N temperature soundings and geostationary satellite cloud-track winds were assimilated. The main simulated data assimilation experiments consisted of corresponding Control and FGGE cycles, as well as a Control plus TIROS temperature profiles experiment, a control plus cloud-track winds, and a Control plus wind profiles. The wind profiles in this initial experiment were idealized. They were assumed to have the same coverage as the TIROS temperature soundings and only 1–3 m/s random errors.

Initial conditions for the simulated data assimilations were provided by a real data control assimilation from 0000 GMT 4 November to 0000 GMT 10 November 1979. Eight five-day forecasts were generated from each assimilation at 48 h intervals beginning on 11 November. In addition, a GLA model integration from the ECMWF analysis on 00 UTC 10 November was generated and then compared to the nature run as a measure of the differences between the two models. This comparison showed that the initial error growth rate between the two model forecasts was nearly identical to the differences between the ECMWF forecast and its analysis. This indicated that the simulation system is significantly more realistic than the identical twin systems used in the early OSSE’s.

Figures 8 and 9 present two of the main results from this OSSE. Figure 8 illustrates the relative impact of satellite temperature soundings and wind profiles on 400 hPa wind analyses over 15 days of assimilation. In the Southern Hemisphere, the assimilation of satellite temperature soundings resulted in a large (approximately 50 %) reduction in analysis error relative to the Control for this experiment. This impact occurred over the first several days of the assimilation, after which the errors are close to 4 m/s. In contrast, the assimilation of wind profiles resulted in a larger and much more rapid reduction of the errors in the initial state. Most of this impact occurred in the first 24 hours of assimilation. In the Northern Hemisphere, the impact of both types of satellite data was much smaller, but the wind profile data was approximately twice as effective as the temperature data in reducing analysis errors.

Figure 9 illustrates the relative impact of the different satellite data sets on the average of eight sea level pressure and 500 hPa height forecasts. This figure, in agreement with the above analysis results, shows a much larger improvement in forecast accuracy resulting from the assimilation of simulated satellite wind profile data than from temperature sounding data in the Southern Hemisphere. In the Northern Hemisphere, the influence of the satellite
Fig. 5. Example of the effect of data removal on sea level pressure for a portion of the Southern Hemisphere. Top map shows the 60 h Control forecast from 00 UTC 21 July 1993; middle map shows the corresponding 60 h Control minus SATEMS forecast. The ECMWF analysis for 12 UTC 23 July 1993 is shown on the bottom.
data sets was much smaller, but on occasion the simulated wind profiles showed a significant positive impact.

More recent OSSE's have investigated tradeoffs in the design and orbital configuration for a Laser Atmospheric Wind Sounder (LAWS), methodology for direction assignment of SSM/I surface wind speed data, and studies of the potential impact of the NASA Scatterometer (NSCAT). These OSSE's followed the same experimental procedure as outlined above. However, a higher resolution, more accurate ECMWF nature run, and a higher resolution DAS were used in the later experiments. A few selected examples from these studies are presented below as further illustrations of the applications of OSSE's.

Figure 10 summarizes results from an experiment conducted with the 2° lat. × 2.5° long. version of the GLA model to evaluate the relative impact of LAWS with three different power levels: 20, 5, and .2 joules. These power levels relate to the instruments ability to retrieve winds from aerosol measurements in the atmosphere. It was assumed that the 20 joule lidar could retrieve winds with 1° × 1° horizontal resolution and 1 km vertical resolution. At 5 joules there is a substantial reduction in the lidar's ability to retrieve winds from aerosols in the middle and upper troposphere, while at .2 joules winds are retrieved only where there are clouds and in the planetary boundary layer. Thus, for the purposes of this experiment the accuracy of the retrieved winds does not change between the different power levels, but the wind data coverage, particularly in middle and upper troposphere decreases substantially as the power decreases.
Fig. 7. 500 hPa height analysis differences (in geopotential meters) between each of the assimilations and the Control at 00 UTC 31 July 1993: (a) Control-SATEMS, (b) Control-SATOBS, (c) Control-PAOBS, (d) Control-Rawinsondes and pilot balloons, (e) Control-Ships and buoys, (f) Control-Aircraft.

The control for this experiment included all conventional observations plus SATEMS and satellite surface winds. The different wind profile data sets that were added to the control were retrieved from the nature run following assumed coverages and accuracies for each power level using the LAWS Simulation Model of Emmitt and Wood (1991). This experiment provides a quantitative estimate of the reduction of impact that would result from reduced power for a space-based wind profiler. Figure 10 shows cross sections of analysis accuracy for the Tropics and for the Northern and Southern Hemispheres Extratropics. These results demonstrate that even with a significant power reduction, a substantial improvement in analysis accuracy relative to the Control in the Southern Hemisphere Extratropics should result from the assimilation of wind lidar data.

Figure 11 presents an application of the simulation system to evaluate different methods for assigning direction to satellite surface wind speed data. As shown in Fig. 1g, SSM/I wind speed observations have excellent coverage and resolution. These data have been available since 1987, and other satel-
Fig. 8. Relative accuracy of 400 hPa zonal wind analyses for simulated Control, Control plus TIROS temperature soundings, and Control plus wind profiles, shown over 15 days of assimilation.

Fig. 9. Relative accuracy (in terms of the S1 skill score) of the average of eight forecasts from the Control, Control plus TIROS temperature soundings, and Control plus Wind Profiles experiments.
Fig. 10. Cross sections of rms error showing the relative accuracy of geopotential height analyses for the Control (solid line) and Control plus 20 joule LAWS (dashed line), 5 joule LAWS (dotted line) and .2 joule LAWS (dash-dot line) experiments.

lite surface wind speed data sets extend back to 1978. The major limitation associated with these data is the lack of directional information. Six different methods were proposed for direction assignment (Atlas and Bloom, 1989). These included: interpolation of model first guess fields (method 1), interpolation of analysis fields (method 2), an Ekman balance approach (method 3), a hybrid approach of 2 and 3 (method 4), a general balance approach (method 5), and a variational analysis approach (method 6). Figure 11 summarizes the results of these tests and clearly shows increased accuracy associated with the variational analysis method.

Following this experiment, a seven and one-half year data set of SSM/I wind vectors was produced using the variational approach, and limited data sets were produced using the other methods. Colocation of the “real” SSM/I wind vectors with ships and buoys (Atlas et al., 1991) indicated that the simulation results were slightly optimistic in terms of accuracy, but confirmed the relative ranking of the methods.

As a final illustration of the application of OSSE’s, Fig. 12 presents results from a recent experiment to assess the potential impact of NSCAT, which is scheduled for launch on the Japanese satellite ADEOS in 1996. In this experiment, assimilation cycles were run with the 2° latitude by 2.5° longitude version of the GEOS GDAS with different types of satellite data included. Figure 12 shows the relative accuracy of surface wind analyses over the first five days of the assimilation that results from (1) the assimilation of conventional data only (CONV), (2) the addition of SATEMS to the CONV assimilation (TOVS), (3) the addition of ERS - 1 scatterometer winds to the TOVS assimilation, and (4) the addition of scatterometer winds meeting the coverage and accuracy requirements for NSCAT to the TOVS assimilation. As can be seen, the assimilation of SATEMS results in a very substantial increase in surface wind analysis accuracy, in agreement with real data impact results. The addition of ERS-1 winds increases the accuracy further, also in agreement with real data studies. The assimilation of the idealized NSCAT data shows approximately twice the impact of ERS-1, indicating considerable potential for this and other advanced active and passive instruments to improve ocean surface analysis.

5. Conclusions and outlook for the future

The current Global Observing System (GOS) for the atmosphere evolved over the last several hundred years. Currently, it consists of a mixture of in situ, visual, and remotely sensed observations. These provide an extensive data base for research and numerical weather prediction.

A large number of observing system experiments (OSE’s) have been conducted to assess the impact of different types of atmospheric observations. In general, these experiments have shown space-based observing systems (SATEMS, SATOBS, PAOBS, ERS-1, SSM/I) to be very important components of the GOS. Nevertheless, conventional surface-based observations are still of critical importance.
Fig. 11. Relative accuracy of different directional assignment methods for SSM/I.

Fig. 12. Relative accuracy of surface wind analyses for the Southern Hemisphere over 5 days of assimilation.

Observing system simulation experiments (OSSE’s) provide an effective means to evaluate the potential impact of a proposed observing system, as well as to determine tradeoffs in their design, and to evaluate data assimilation methodology. Great care must be taken to ensure realism of the OSSE’s, and in the interpretation of OSSE results. Recent OSSE’s suggest considerable potential for space-based lidar wind profiles and for improved space-based surface winds. But these OSSE results are only valid for instruments meeting the expected accuracies and coverages that were simulated. Further experiments are needed to define the impact for a wider range of instrument and data characteristics.

Over the next several years many new space-based observing systems will become available. Some examples of these are: AMSU, an advanced microwave sounder which should provide improved temperature retrievals under extensive cloud cover; NSCAT and other advanced scatterometers for determining wind velocity over the global oceans with greater accuracy and coverage; AIRS for much higher vertical resolution and more accurate temperature and moisture profiles; MODIS for measuring cloud cover,
water vapor, and surface properties; TRMM PR for measuring tropical rainfall; and possibly a space-based lidar for global wind profiling (Baker et al., 1995). Some of the observing systems will provide data types which have never before been assimilated by numerical models. These will give exciting challenges to four-dimensional data assimilation in the near future.

Acknowledgments

This paper is dedicated to the memory of Dr. James W. Pfendtner, who headed the Data Assimilation Group at the NASA Goddard Space Flight Center for six years and supervised the development of the GEOS DAS. The author gratefully acknowledges the support of NASA Head-Quarters and the NSCAT project. Several of the OSSE’s referred to in this paper were performed in collaboration with G.D. Emmitt and S. Wood. I would also like to thank S.C. Bloom, J.C. Jusem, E. Brin, J. Ardizzone, J. Terry, and D. Bungato for significant contributions to this work.

References


Bell R.S., 1994: The assimilation of ERS-1 scatterometer winds. Forecasting Research, UKMO.


