DAO Office Note 96-13

Office Note Series on Global Modeling and Data Assimilation

Richard B. Rood, Head
Data Assimilation Office
Goddard Space Flight Center
Greenbelt, Maryland

New Data Types Working Document

Joanna Joiner
Robert Atlas
Steve Bloom
Genia Brin
Stephen Cohn
Arthur Hou
Dave Lamich
Dave Ledvina
Richard Ménard
Lars-Peter Riishøjgarrd
Richard Rood
Arlindo da Silva
Meta Sienkiewicz
Jim Stobie
Runhua Yang

Data Assimilation Office, Goddard Laboratory for Atmospheres
Goddard Space Flight Center, Greenbelt, Maryland

This paper has not been published and should be regarded as an Internal Report from DAO.

Permission to quote from it should be obtained from the DAO.

Goddard Space Flight Center
Greenbelt, Maryland 20771
June, 1996

Abstract

This working document describes activities related to the New Data Types Group of the Goddard Data Assimilation Office (DAO) and how these activities are prioritized. Included in this document are science drivers for new data types, brief descriptions of assimilation methodologies including advanced methodologies being explored at the DAO, a list of new data types being considered for assimilation and validation at the DAO, and how these data types are roughly prioritized. This document will be updated periodically.

On-line versions of this document are available from

`ftp://dao.gsfc.nasa.gov/pub/office_notes/on9613.ps.Z  (postscript)`

Contents

List of Tables

1 Introduction

The mandate of the Data Assimilation Office (DAO) at NASA/Goddard is to provide high-quality data sets to be used by the Earth science community to study a number of Earth Systems problems. As part of the DAO mandate, a major effort is being undertaken to assimilate new data types from current and future instruments. Emphasis is given to instruments that will fly as part of NASA's Mission to Planet Earth (MTPE) program, particularly those on board the Advanced Earth Observing System (ADEOS), the Tropical Rainfall Measuring Mission (TRMM), and the Earth Observing System AM-1 (EOS-AM1) scheduled for launches in 1996, 1997, and 1998, respectively (currently, this document only addresses satellites through EOS-AM1). The instruments aboard these satellites are designed to measure quantities related to atmospheric and surface parameters of particular relevance to Earth Systems study. More detailed information about components of the DAO's GEOS-1 system can be found in Takacs *et al.* (1994), Pfaendtner *et al.* (1995), and Schubert and Rood (1995). Documentation on the DAO's Physical-Space Statistical Analysis System (PSAS) can be found in da Silva and Guo (1996).

The effort required to assimilate new data types is non-trivial. For example, at other NWP centers, the effort to incorporate new data types has required approximately 5-10 person-years per new data type (and is an ongoing effort). New data types currently being considered for assimilation at the DAO may be classified into two categories: (1) Data types from instruments with an assimilation heritage at the DAO and other centers (*e.g.,* TOVS radiances, Scatterometer measurements, etc.) (2) Data types that do not have a long-term operational heritage (*e.g.,* precipitation, surface wetness, etc.). Data types in (2) are considered to be a higher risk than those in (1). The DAO effort to assimilate new data types includes developing advanced assimilation methodologies that will be required in order to handle the large volume of data from future instruments. Monitoring new data types before, during, and after assimilation will be an important component of the new data types effort. The effort will also involve interacting with instrument teams as well as interfacing with other DAO groups, such as the those concerned with covariance tuning, quality control, operations, and development.

This document attempts to define and prioritize the activities of the DAO's *New Data Types* group. This is a working document, meaning that it will evolve as experience is gained with new data types. Parts of the document are incomplete at this time and will be revised in the future. This document will be used by DAO staff to guide decisions in how resources related to *New Data Types* will be allocated in the present and future. It will also provide instrument teams with information about what will be required from them in order to effectively utilize observations in the Goddard Earth Observing System-Data Assimilation System (GEOS-DAS). Details on related topics such as the Kalman filtering effort, Observation System Simulation Experiments (OSSE), Observation System Experiments (OSE), and bias correction will appear elsewhere or will be incorporated into the document at a later time. Documentation on related topics can be found elsewhere; monitoring can be found in da Silva *et al.* (1996b), quality control in Dee and Trenholme (1996), quality assurance in Schubert (1996), and operations in Stobie (1996).

The outline of the document is as follows: Section 2 describes the scientific goals that drive the effort to incorporate new data types in the DAO's Data Assimilation System. This section begins with an overview and follows with subsections devoted the hydrological cycle, land-surface/atmosphere interaction, ocean-surface/atmosphere interaction, radiation (clouds, aerosols, greenhouse gases), atmospheric circulation, and constituents. For each scientific driver, examples of relevant data types are given. Section 3 discusses traditional and advanced methodologies for assimilating data and provides examples of relevant data types for each method. Section 4 describes several topics related to implementation: computational issues, data flow, and instrument team interaction. Finally, new and existing data types are prioritized in section 5. Cross-references of instrument to data type, scientific driver, use in GEOS, etc. are provided.

2 Scientific drivers

2.1 Overview

The NASA Mission to Planet Earth (MTPE) is a program designed to make use of ground, aircraft, and satellite-based measurements in order to better understand the systems that govern the Earth's climate, their interactions, and their variations. NASA's Earth Observing System (EOS) and several other satellite systems are key components of this program. Driving the selection of new data types in the GEOS-DAS is a subset of scientific objectives outlined by the MTPE program, the United States Global Change Research Program (USGCRP), and the Intergovernmental Panel on Climate Change (IPCC). These objectives include improving our understanding of (1) the hydrological cycle (interaction between the land and ocean surfaces with the atmosphere and transport of heat, moisture, and momentum) (2) the interaction of clouds and aerosols with radiation and its impact on climate (3) chemistry and transport of atmospheric constituents, such as ozone, in both the troposphere and stratosphere. These and other Earth systems topics are briefly discussed below in the context of data assimilation. Examples of relevant data types are provided.

2.2 Hydrological Cycle

Water substance plays an integral role in many of the processes operating within the Earth System. In the atmosphere, latent heat release through condensation of water vapor, producing clouds, is a major source of energy that regulates atmospheric circulation. A large fraction of the energy transferred from the surface to the atmosphere is in the form of latent heat due to evaporation which itself depends on the moisture gradient between the surface and atmosphere. In addition, the presence of clouds and precipitation determines the availability of solar radiation and water at the surface. Water vapor is also the predominant greenhouse gas and plays a crucial radiative role in the global climate system.

Clearly, an accurate depiction of the global atmospheric moisture field, its vertical and horizontal transport, and the transfer of water across its boundaries, is critical to understanding the hydrologic cycle and its impact on climate. In the past, moisture information used in the GEOS-DAS has come primarily from rawinsonde ascents. Satellite moisture estimates offer a significant improvement over the spatial and temporal sampling problems of the existing rawinsonde network, but they also have limitations. Polar orbiting satellites such as those in the Defense Meteorological Satellite Program (DMSP) carry microwave instruments which are sensitive to atmospheric moisture. The Special Sensor Microwave/Imager (SSM/I) for instance, gives an accurate estimate of the integrated water vapor in an atmospheric column (referred to as Total Precipitable Water, TPW) but contains little information about the vertical distribution of moisture which is critical to understanding many of the physical processes mentioned above. The SSM/I observations are also limited by the fact that they are only available over the oceans in precipitation/sea ice free regions. SSMT/2 provides similar coverage, but contains additional channels that provide information about the vertical structure of moisture. Other satellites, such as TIROS Operational Vertical Sounder (TOVS) which carries the High-resolution Infrared Sounder (HIRS) and the Microwave Sounding Unit (MSU), have better geographical coverage (over land as well as ocean) and can resolve the integrated water in approximately two thick slabs. Future microwave and IR instruments, such as MHS (AMSU-B), AIRS, and IASI, will have improved vertical resolution for moisture sounding. Although still fairly crude, the moisture estimates that are/will be available from these satellites are valuable both for assimilation, validation, and bias estimation. For example, information about global moisture derived from SSM/I and TOVS has been assimilated at DAO and other centers (*e.g.,* Ledvina and Pfaendtner, 1995; Derber *private communication*; Andersson *et al.*, 1994) and has significantly impacted global analyses.

Significant discrepancies have been found between precipitation estimates from analysis systems and those derived from satellite measurements. Systematic differences in both the intensity and spatial distribution of monthly mean precipitation are especially large in the tropics. Current research at the DAO is directed towards assimilating precipitation data. Improving the hydrological cycle in analyses in both the tropics and extra-tropics will ultimately aid in the understanding of climate sensitivity, climate variability, the role of dynamical feedback, and other physical processes. Improving the hydrological cycle in data assimilation systems will also benefit studies in which precipitation estimates from a DAS are used to drive surface hydrological models.

2.3 Land-Surface/Atmosphere Interaction

The land-surface is an important component in the Earth System affecting energy balance, the hydrological cycle, and chemical cycles. Latent heat, sensible heat, and momentum fluxes at the land-surface modulate near-surface turbulence and boundary layer convection on diurnal and longer time-scales. The land-surface plays an important role in the hydrological cycle receiving water from the atmosphere in the form of rain or snow with the soil and vegetation acting as reservoirs. Precipitation that does not infiltrate the soil forms surface runoff. The biosphere over the land surface also affects the carbon cycle through photosynthesis to produce the greenhouse gas CO_2.

Although still in its infancy, land surface data assimilation has generated interest in both environmental and meteorological communities. Over the past decade, state-of-art land surface models (LSM) have been coupled with general circulation models. Remote sounding instruments that infer land surface quantities will play an integral role in land-surface modeling and data assimilation in the future. Many new data types can be inferred remotely from satellite instruments that provide superior geographic coverage of the land-surface as compared with conventional surface station observations. These satellite-derived data types include surface (skin) temperature from infrared instruments such as TOVS, AVHRR (MODIS predecessor), and MODIS, as well as snow water equivalent content and soil moisture information from passive microwave instruments such as SSM/I and in the future TMI. In addition to measurements that can be assimilated, satellite-derived quantities such as vegetation indices from AVHRR and MODIS can also be used to specify or estimate model parameters in some LSMs. Use of these data should improve both short-term forecasts and analyses of climate events such as El Nino.

2.4 Ocean-Surface/Atmosphere Interaction

The ocean surface (including regions immediately above and below) is an interface between two of the great subsystems involved in the Earth System: the ocean and the atmosphere. Considerable Earth Systems Science research will be devoted to the interplay of dynamics and features between these two subsystems. There is a considerable disparity in the nature of our understanding of these subsystems and their interaction. A better understanding of the workings of the Earth System, especially for long time scales, will entail a detailed knowledge of the behavior of the coupled atmosphere-ocean system. El Nino and its global effects is certainly an example of this. Thus it is important to have improved ocean surface data to better delineate the mechanisms that govern the atmosphere-ocean interaction.

Past observations of ocean surface variables have largely been obtained from in situ platforms: ships, buoys and island data. These observing systems provide information on sea level pressure, winds, humidity and sea surface temperature. Such observations have suffered from a variety of problems, ranging from unrepresentative sampling to inherently poor quality. The most serious problem with these historical data sources lies in their poor sampling of the Tropics and the southern oceans.

There are currently in place satellite-based sensing systems which provide information related to sea surface temperature and sea level winds. For example, one source of information concerning winds in the lower atmosphere, cloud-tracked winds (CTW), has been available for a number of years. Research using data from the current space-based ocean surface wind sensors, SSM/I and Scatterometer, has found potential benefits and difficulties in using these data. The benefits are clear; the satellite sensors provide a tremendously improved sampling, in both space and time, of the ocean surface wind. The difficulties with using data from these sensors lie in drawing the proper inferences from the information provided by the sensors. For example, the SSM/I provides information related to the wind speed near the ocean surface - provided a number of assumptions pertaining to the wind-sea state relation and the nature of the boundary layer near the sea surface are valid. Similar problems arise in the consideration of Scatterometer data.

Improvements in the sensing of winds near the ocean surface will benefit a number of Earth Science research enterprises. More accurate surface winds will lead to improved estimates of momentum fluxes, and have a potentially positive impact on the fluxes of sensible heat and moisture for use in atmospheric modeling. Similarly improved estimates of momentum, heat and fresh water fluxes will be used for ocean models. Such wind data will provide important forcing, or boundary conditions, for research in coupled atmosphere-ocean modeling. Ultimately, the assimilation of surface wind data into a coupled atmosphere-ocean system will be used to correct the state of such a system. Another area of interest for these wind data is in ocean wave modeling; accurate high-resolution winds will improve the process of data assimilation of surface winds into ocean wave models. Finally, improved surface wind datasets, obtained from the global assimilation of these new surface wind data, will provide more useful forcing for offline model studies, such as sea-ice transport studies.

2.5 Radiation (Clouds, Aerosols, and Greenhouse Gases)

The transfer of electromagnetic radiation in the atmosphere, *i.e.,* absorption, emission, scattering, and refraction of energy by particles (clouds and aerosols) and molecules, is the most important process affecting energy transfer. Radiative transfer in the atmosphere is generally separated into two regions: (1) Solar or shortwave radiation that peaks near 0.5 μm in the visible part of the electromagnetic spectrum (2) Terrestrial or longwave radiation that peaks in the infrared part of the spectrum near 10μm. One consequence of molecular absorption is the so-called greenhouse effect in which a gas is virtually transparent to incoming shortwave radiation while absorbing and re-emitting longwave radiation. The major greenhouse gases are CO_2, H_2O, and O_3 with CO, CH_4, N_2O and others playing a smaller roll. The CO_2 concentration remains relatively homogeneous in both space and time, although observations show a seasonal variation and long-term increase. In contrast, concentrations of H_2O, O_3, and some of the minor greenhouse gases exhibit relatively large spatial and temporal variations. Climate changes resulting from variations in greenhouse gas abundances are difficult to predict because of complex radiative-convective feedbacks and remain a controversial issue.

Cloud properties have a large impact on energy balance at the top of the atmosphere and the surface. In the shortwave, suspended liquid water in clouds reflects incoming solar radiation back to space. In the longwave, liquid water drops absorb and re-emit radiation. Molod *et al.* (1995) have shown that systematic errors in the GEOS-1 general circulation model (GEOS1-GCM) linked to physical parameterizations of clouds, convection, and rainfall result in latitudinally-dependent systematic differences between observed and GEOS1-derived long and shortwave flux at both the upper and lower boundaries of the atmosphere. Currently, satellite observations (*e.g.,* from the Earth Radiation Budget Experiment or ERBE) are used to diagnose problems with model parameterizations.

Aerosols, like cloud particles, can significantly affect the heat balance of the Earth-atmosphere system by scattering and absorbing incoming solar radiation. Aerosols, or suspended particles in the atmosphere, include volcanic dust, sea spray, dust generated from wind, smoke from forest fires and biomass burning, particles produced during combustion, chemical reactions involving naturally occurring gases or gases formed during combustion, and cataclysmic impacts between the Earth and other solar system bodies. Temporal and spatial variations in aerosol distributions have not been given much attention in current data assimilation systems.

There currently exists an abundance of radiation-related data inferred from satellite and ground-based instruments, such as cloud-top temperature and cloud fraction in the ISCCP data base (derived from combined infrared and visible observations) and cloud-liquid water derived from passive microwave instruments such as SSM/I. Future instruments such as MODIS and the TRMM microwave imager (TMI) will provide additional estimates of these quantities. Ground-based aerosol observations date back to the early 1900's. Aerosol parameters have also been measured with satellite instruments such as SAGE, HALOE, and CLAES, primarily in the stratosphere. Advanced instruments such as MISR have been designed for improved observations of aerosol properties. Ground-based monitoring of greenhouse gas concentration has been taking place over the last several decades. In the future, more complete global monitoring of greenhouse gases will be accomplished with instruments such as MOPITT. These data types related to radiation have not yet been used in operational data assimilation systems. Research at the DAO is currently ongoing to more fully exploit these new data types in a DAS. This should lead to a better understanding of the role of radiation and dynamical feedback processes in regulating the climate system.

2.6 Atmospheric Circulation

Accurate, consistent, long-period measurements of global mass (height and temperature) and momentum (wind) fields are important for Earth System studies. Atmospheric circulations are the method by which sensible and latent heat, mass, and momentum fluxes are transported. The study of climate variability focuses in part on inter-annual differences and on anomaly fields (differences from climatological means), including changes in large-scale flows such as the Hadley circulation, which transports mass, energy and moisture from the tropics to the mid-latitudes. Another important long time-scale circulation is the El Nino-Southern Oscillation (ENSO), which has been linked to changes in the extra-tropical circulation resulting in floods and droughts. The stratospheric quasi-biennial oscillation (QBO) is a major component of the inter-annual variability in the stratosphere. Several documented deficiencies in GEOS-1, such as a weak Hadley cell and warm-biased tropopause temperatures in the tropics (see references in Schubert and Rood, 1995), can potentially be addressed with new data types (including improved use of current observations).

2.6.1 Tropospheric circulation and temperature

In the troposphere (and lower stratosphere), both winds and temperatures are observed with a network of conventional rawinsondes. This network is dense in the Northern Hemisphere over land. However, observation density and quality in the tropics, Southern Hemisphere, and over oceans, is much poorer. In these areas satellite observations provide excellent geographic coverage. Advances in global weather models and data assimilation have led to better global analyses and forecasts of these fields, but there is room for improvement, both in the observations themselves as well as the way in which they are used in assimilation systems. Advanced instruments for temperature sounding, such as AIRS and IASI, will provide measurements with higher information content about the mass field than the current NOAA polar-orbiting satellites. Yet, even with improved sounding capabilities, there are still both technical and scientific issues to be addressed when considering how best to assimilate remote satellite observations. These issues will be addressed in more detail in section 3.

2.6.2 Stratospheric circulation and temperature

The need for improved knowledge of stratospheric winds and temperatures is most strongly driven by efforts to understand stratospheric ozone variability and predict anthropogenic changes in the ozone distribution. Recently, there has been increased interest in role of the stratosphere in climate change, especially changes associated with upper tropospheric and lower stratospheric water vapor and ozone. In addition, new initiatives in tropospheric chemistry require better quantification of stratosphere-troposphere exchange.

Winds and temperature from data assimilation systems have been central in increasing the quanititive level of stratospheric chemistry. In the lower stratosphere, middle latitudes the winds are already of sufficient quality that transport calculations can be used to unify constituent measurements from different observing platforms. The chemical studies have also revealed places where the winds and temperatures remain of low enough quality that uncertainties in chemical assessments are still strongly tied to meteorological conditions.

Within the current GEOS system the most notable problems with the winds lie in the tropics and the subtropical middle latitude boundary. The transport between the subtropics and middle latitudes is an especially important topic because mixing of middle latitude pollutants into the tropics can strongly perturb the ozone sources. With regard to temperature, field scientists in the aircraft missions have noted a warm bias of GEOS temperatures in the lower stratosphere when the extreme cold events occur. Improvements in the temperature representation are needed because heterogeneous chemical processes become more important at extremely low tempertures. Long-term transport experiments show that recent versions of GEOS have extensive improvements in the representation of the seasonal, zonal-mean meridional circulation. However, misrepresentations of the equatorial upwelling and polar downwelling remain large enough that it is difficult to represent interannual variability in long-lived tracers.

Much of the projected improvement in stratospheric meteorological fields is expected to come from advances in modeling, analysis, and specification of improved error statistics. New data sources are limited primarily to improved temperature sounders, such as MLS and GPS, that have increased vertical resolution and accuracy. Indirect improvements to meteorological fields may also be derived from the assimilation of long-lived tracers. Stratospheric wind measurements from the Upper Atmosphere Research Satellite (UARS) have proven difficult to utilize and the DAO will follow the progress on impact studies being carried at at the UKMO.

2.7 Constituents

Building the capability of assimilating constituent observations is one of the long-term goals of the DAO. In such an assimilation scheme the constituent data are allowed to influence both the estimate of the tracer itself and - by the coupling of the fields through the linear transport equation - the estimate of the flow field (Daley, 1995; Riishøjgaard, 1996). Constituent assimilation in the present context means assimilation of observations of the minor constituents of the atmosphere, such as water vapor, N_2O, CO, CH_4, etc. However, one of the main motivations for recent development in methodologies for constituent assimilation is the availability of ozone derived from satellite-based instruments. Assimilation of these data would be useful for the atmospheric chemistry community since it would allow us to build a multi-year sequence of three-dimensional ozone fields consistent with atmospheric dynamics, available for each analysis time.

From a meteorological point of view the ozone data is of interest because the flow patterns around the tropopause level generates a very strong signal in total ozone observations. Conversely, this implies that the total ozone measurements carry information about the winds at these level, information that over large parts of the globe is not readily available from alternative sources. There are a number of different strategies that one can follow in order to retrieve this information, depending on the model underlying the data assimilation procedure. The basic ideas involved in constituent assimilation are intuitively appealing and conceptually simple. However, the actual design and implementation of a multivariate scheme in a general circulation model is technically a major effort, and there remain a number of scientific issues to be resolved. Thus it is far from clear at the present time which assimilation method (EKF, 4D-VAR, Fixed-interval smoothing, assimilation through a transport model or through a PV model, etc.) will be best suited for the purpose. Also, even though some of these methods theoretically should be insensitive to the initial specification of the forecast error covariances, this may not be true for a particular numerical implementation. Since the partitioning of the information present in the data will be determined by the error specified, work towards improving our understanding of the errors both of the observations and of the background model estimate is clearly needed and will be an integral part of constituent assimilation at the DAO.

3 Assimilation Methodology

Several approaches have been proposed and used to incorporate both conventional and satellite observations in data assimilation systems. These methods will be reviewed in subsequent sections following a brief review of statistical analysis and description of the DAO's Physical-space Statistical Analysis System (PSAS). We will focus here on assimilation of remotely sensed data from satellites in PSAS, although the concepts also apply to conventional data as well as other assimilation methods such as the Kalman filter.

3.1 Statistical Analysis

The objective of statistical interpolation is to produce an optimal estimate of the atmospheric state, given a set of observations and a first guess usually in the form of a short-term forecast. In the variational framework (*i.e.,* 3D-VAR), this can be accomplished by minimizing the likelihood functional

$$J(w) = (w - w^f)^T (P^f)^{-1} (w - w^f) + (w^o - h(w))^T (R^o)^{-1} (w^o - h(w)), \tag{1}$$

where $w \in I\!R^n$ is a vector representing the 3D state of the atmosphere, $w^f \in I\!R^n$ is the forecast, $w^o \in I\!R^p$ is the observation vector, and $h(w)$ is an observation operator that maps the 3D atmospheric state into observables. The first term on the RHS of (1) is weighted by the inverse of the forecast error covariance matrix $P^f \in I\!R^n \times I\!R^n$, while the second term is weighted by the inverse of the observation error covariance matrix $R^o \in I\!R^p \times I\!R^p$. Provided these covariances are specified correctly, the analysis state obtained by minimizing $J(w)$ is the mode of the conditional probability density function $p(w|w^f \bigcup w^o)$ and is derived from a maximum-likelihood principle assuming that forecast and observation errors are unbiased, normally distributed, and uncorrelated with each other.

Because the observation operator $h(w)$ is in general nonlinear, the minimum of $J(w)$ can be obtained by a quasi-Newton iteration of the form

$$w_{i+1} = w^f + P^f H_i^T \left(H_i P^f H_i^T + R^o \right)^{-1} \left[w^o - h(w_i) + H_i (w_i - w^f) \right], \tag{2}$$

where

$$H_i = \left. \frac{\partial h(w)}{\partial w} \right|_{w=w_i}. \tag{3}$$

The analysis vector, w^a, that minimizes $J(w)$ is given by

$$w^a = \lim_{i \to \infty} w_i. \tag{4}$$

In PSAS, a $p \times p$ system of equations is first solved at observation locations

$$\left(H_i P^f H_i^T + R^o \right) x_i = w^o - h(w_i) + H_i \left(w_i - w^f \right) \tag{5}$$

for the vector $x_i \in I\!R^p$ using a conjugate gradient algorithm (da Silva *et al.*, 1995; Guo and da Silva, 1995). The first term on the LHS of (5) is called the innovation covariance. The state at iteration $(i+1)$ is updated by an additional matrix-vector multiply, viz.

$$w_{i+1} = w^f + P^f H_i^T x_i. \tag{6}$$

The computation required for the solution of the linear system (5) is approximately $\mathcal{O}(N_{cg} p^2)$, where N_{cg} is the number of iterations of the conjugate gradient algorithm. N_{cg} depends on the conditioning of the innovation covariance matrix. The matrix-vector multiply in (6) requires $\mathcal{O}(np)$ floating point operations. The total operation count to solve (5)-(6) is approximately $\mathcal{O}[N_o(N_{cg}p^2 + np)]$, where N_o is the number of outer (quasi-Newton) iterations performed. It is evident any type of data compression that reduces the number of observables will significantly reduce computation in PSAS. At the DAO, development of improved methodology to assimilate new data types with very high spatial- or spectral-resolution has been driven largely by consideration of computational costs.

3.2 Direct radiance assimilation

In this document, radiance is a general term meaning a directly measured quantity (as opposed to a retrieved quantity). For example, radiance could refer to refractivity measured by Global Positioning System (GPS) receivers, backscatter from a scatterometer, or thermal/reflected radiation measured by a passive infrared or microwave sounder. The term radiance also applies to pre-processed raw radiance (*e.g.,* cloud-cleared radiance). Assimilation of radiances involves utilizing radiance measurements from a remote sounding instrument as the observable and specifying the radiance error covariance as the observation error covariance in (1). In addition the observation operator h in (1) that maps the state variables to the radiances must be specified.

For remote measurements, h is an approximate radiative transfer or empirical model relating the atmospheric state in grid space to the radiance at the observation location using a set of parameters such as spectral line data and/or calibration parameters. To assimilate radiances correctly, an appropriate radiance error covariance must be specified that incorporates both detector noise and observation operator error. Practical implementation at the DAO will use statistical modeling from innovation sequences or some form of online parameter estimation such as discussed in Dee (1995) to estimate radiance errors. Observation operator parameter errors are usually systematic and may be estimated and corrected for in part by utilizing independent observations and/or forecasts (*e.g.,* Eyre, 1992; Susskind and Pfaendtner, 1989).

Radiance assimilation is computationally feasible with current instruments and analysis schemes. However the cost of this approach may be prohibitive for future high-spectral resolution sounding instruments, such as AIRS and IASI, with large numbers of channels. More efficient approaches are currently being examined as an alternative to radiance assimilation for such instruments.

3.3 Traditional retrieval assimilation

Remotely sensed data have been traditionally assimilated in the form of physical-space retrievals. In this approach, radiances are processed off-line by data producers, and *familiar* data types such as temperature/moisture profiles or wind vectors are used in the assimilation system. Specifying the retrieval z as the observable w^o in (1), the observation operator h is a linear interpolation operator so that the iterated form of (2) can be reduced to

$$w^a = w^f + \left(P^f \mathcal{I}^T\right)\left(\mathcal{I} P^f \mathcal{I}^T + R^z\right)^{-1} (z - \mathcal{I} w^f), \tag{7}$$

where $z \in I\!\!R^p$ denotes the retrieved data, $R^z = \left\langle \epsilon^z (\epsilon^z)^T \right\rangle$ is the retrieval error covariance, and $\mathcal{I} \in I\!\!R^n \times I\!\!R^p$ is the interpolation operator used above. A more general form of (7) that includes the retrieval-forecast error cross-covariance, denoted $X = \left\langle \epsilon^z (\epsilon^f)^T \right\rangle$, is given by

$$w^a = w^f + \left(P^f \mathcal{I}^T - X^T \right) \left(\mathcal{I} P^f \mathcal{I}^T + R^z - \mathcal{I} X^T - X \mathcal{I}^T \right)^{-1} (z - \mathcal{I} w^f). \tag{8}$$

The assimilation of retrievals requires the specification of the retrieval error covariance matrix R^z and the retrieval-forecast error cross-covariance X. Operational implementations of retrieval assimilation are often based on statistically modeled retrieval error covariances under the assumption of stationarity, and horizontal homogeneity and isotropy. In addition, the retrieval-forecast error cross-covariance matrix X is often neglected as a result of the difficulty associated with modeling it and accounting for it in a DAS. At the DAO, TOVS retrieval errors have been estimated using a tuning algorithm that separates horizontally-correlated components of the error from uncorrelated components as described in da Silva *et al.* (1996a).

3.4 Consistent Assimilation of Retrieved Data (CARD)

Ménard (1995) proposed a potentially less expensive alternative to radiance assimilation that has a more theoretically sound basis than traditional retrieval assimilation. This method combines Rodgers' (1990) characterization of retrieval errors with Kalman filter theory leading to the consistent assimilation of retrieved data (CARD). The implementation of CARD may be considerably less expensive than radiance assimilation for advanced instruments where the number of radiance measurements is much greater than the number of retrieved products (*i.e.,* the retrieval is a form of data compression). The retrieval of geophysical parameters is a nonlinear estimation process. Often the problem is ill-posed and in this case requires the use of prior information. For example, nadir-viewing infrared and microwave profiling instruments use forecasts as prior information for interactive physical retrievals. Prior information could also come from climatology or a representative ensemble of profiles used to create a regression, pattern recognition, or neural net retrieval algorithm. The general CARD approach requires the specification of error characteristics for the radiances as well as the prior information. Because the statistical characteristics of the prior information are often not known and/or difficult to model and account for in a DAS, some modification to the retrieval may be necessary in order to eliminate the effects of prior information. In the following two subsections, examples of different CARD implementations are given.

3.4.1 Physical Space

If little or no prior information is used in the retrieval, the retrieval errors may be computed by propagation of instrument error (including detector noise and transfer modeling errors) as described in Rodgers (1990). In this way, the state dependence of the retrieval error is accounted for. If the state dependence is to be properly accounted for, instrument error must be specified (as in the case of radiance assimilation). Once the retrieval error is estimated, the retrievals can then be assimilated in a consistent manner. This approach has been used by Ménard *et al.* (1995) to assimilate constituent data from the Cryogenic Limb Array Etalon Spectrometer (CLAES) using a Kalman filter algorithm where retrieval errors were reported by the instrument team.

3.4.2 Phase Space

If a significant amount of prior information is incorporated into the retrieval, the implementation becomes more difficult as a result of the cross-correlation between the prior information and forecast errors X in (8). Joiner and da Silva (1996) describe several ways to modify retrievals in order to remove the effect of the prior information thereby effectively removing X. These approaches involve (1) filtering the portion of the retrieval affected by the use of prior information (*i.e.,* null-space filtering or NSF) or (2) performing a partial eigendecomposition (PED) retrieval in which prior information is not incorporated into the retrieval. These approaches compress radiance observations in a single sounding to a small number of orthogonal functions that will impact the data assimilation system. For example, Joiner and da Silva (1996) showed that over 500 radiances observations from the AIRS instrument could be compressed into approximately 10 pieces of relevant information about the temperature profile. The observations may then be compressed horizontally (*i.e.,* combined into a super-observation or super-ob) provided the retrievals are defined in terms of a consistent set of basis functions. The approaches described here have been demonstrated with simulated data in 1-D thus far. Full implementation of both direct radiance assimilation and the CARD/PED approach using TOVS data in PSAS is underway at the DAO so that these methods can be compared both in terms of cost and analysis quality.

4 Implementation

4.1 Data flow and Computational Issues

One important consideration for incorporating new data types an appropriate data assimilation method is the bandwidth required to accommodate data flow between the site where the raw data is pre-processed and the site where the data is assimilated. For future high-spatial and high-spectral resolution instruments, such as MODIS and AIRS, the required bandwidths (assuming no data compression) can be quite large. Using level 3 products when possible, super-obbing, and/or the CARD methodologies described in section 3.4 that allow for data compression, a significant reduction in data flow as well as computation in PSAS can be achieved. To achieve this reduction in data flow, the preprocessing required to produce appropriate retrievals and to perform super-obbing must take place at the site where the data resides. Furthermore, quality control (and in some cases covariance tuning) will have to be performed at the data site in conjunction with data compression. Interaction with the EOSDIS project will be required to ensure that adequate facilities are in place to accommodate data flow and data compression at the data site.

4.2 Instrument Team Interaction

Members of instrument teams are familiar with the intricate and unique problems associated with a particular instrument (*e.g.,* calibration and other sources of systematic error in the observations and operators). In many cases, instrument teams have developed the tools that are necessary to effectively use the data in a DAS (*e.g.,* observation operators). In the past, interaction between instrument teams and data assimilation teams has been weak. The DAO and New Data Types group would like to foster stronger interaction with instrument teams. For this interaction to succeed, strong commitment by both the DAO and instrument teams is required. A prototype for this interaction is planned with the MLS team.

The current plan for instrument team interaction is that at least one representative from an instrument team will work closely with the DAO for an extended period of time (of the order of a few months to a year). The time required to integrate the new data type into the DAS will depend on several factors including the amount of previous experience with a similar data type, the readiness of tools to assimilate the new data type (*i.e.,* observation operators), the quality of the data type (including the ability to remove bias), and the amount of pre-processing needed prior to assimilation. In most cases, instrument teams will be expected to provide the DAO with the tools needed to properly assimilate the data, such as observation operators and their derivatives. The instrument team member will interact

Table 1: Seminar Speakers for New Data Types Group

Speaker	affiliation	Topic
Steve Bloom	NASA	Ocean-surface Wind Retrievals
Al Chang	NASA	Snow Water Equivalent Retrieval
John Derber	NMC/NCEP	Radiance Assimilation at NMC
Anne Douglass	NASA	Ozone 3D Chemistry-Transport Model
Gregory Gurevich	USRA/UMd	Satellite Tomography
Paul Houser	U Arizona	Remote Sensing of Soil Moisture
Randy Koster	NASA	Mosaic Land Surface Model
Chris Kummerow	NASA	Precipitation Retrievals (SSM/I and TRMM)
Venkataraman Lakshmi	NASA	Soil Moisture from SSM/I
Andrea Molod	NASA	Coupling of Land Surface Model at DAO
Howard Motteler	UMBC	Neural Net Retrievals (SSMT/2 and AIRS)
Bill Seegar	Aberdeen	Bird-based meteorological measurements
Larrabee Strow	UMBC	Fast and hyperfast IR radiative transfer modeling
David Tobin	UMBC	Infrared spectral lineshapes

with DAO staff specializing in covariance tuning, quality control, and the PSAS interface. The instrument team member is expected to provide baseline models for systematic error correction and covariance modeling and will assist with quality control and monitoring.

Another aspect of instrument team interaction involves teams utilizing the data assimilation system to diagnose and correct problems with the instrument (*i.e.,* calibration error) and observation operators (*e.g.,* tuning the forward model). An example of this is the work done with the ERS-1 scatterometer by Stoffelen and Anderson (1995). This interaction will be part of the monitoring activity for any new data type to be used in the GEOS system.

As part of the DAO commitment to bring together instrument teams and data assimilation teams, the DAO has agreed to host and help organize a satellite assimilation workshop headed by Ron Errico and George Ohring in March 1997. The New Data Types group has also invited several instrument team members and scientists working on related projects to give seminars. Table 1 lists speakers to date that have given seminars at the DAO. Table 2 lists other visitors, collaborators, and consultants in contact with the DAO New Data Types group.

5 Priorities

During several meetings of the New Data Types group, instruments on past, present, and future platforms were reviewed in order to prioritize their usage in the GEOS/DAS. Data types were selected to be used either for assimilation, analysis (non-cycled), or validation. When two or more instruments provided redundant information, generally one was selected for assimilation, allowing for the other to be used as independent validation (or perhaps bias estimation). An initial selection of a data type for validation may change to assimilation after monitoring if (1) the data appear to be of higher quality than data currently used for assimilation (2) a commitment is made by an instrument team (3) no commitment by an instrument team is made, but resources within the DAO permit. Da Silva *et al.* 1996b describe the planned DAO On-Line Monitoring System (DOLMS).

Each data type was initially assigned a score on a four point scale based on an estimate of the cost-to-benefit ratio, where benefit is defined primarily in terms of meeting the scientific

Table 2: Visitors, Consultants, and Collaborators

Contact	Affiliation	Subject
P.K. Bhartia	NASA	TOMS/SBUV retrievals
Moustafa Chahine	JPL	AIRS retrievals
Jean Dickey	JPL	GPS ground-based H_2O retrievals
Steve Engman	NASA	Soil moisture from passive microwave
Evan Fishbein	JPL	MLS retrievals
Larry Gordley	GATS	CLAES retrievals, limb sounding radiative transfer
Christian Keppenne	JPL	GPS ground-based H_2O retrievals
Arlin Krueger	NASA	TOMS retrievals
Steven Marcus	JPL	GPS ground-based H_2O retrievals
Piers Sellers	NASA	Sib2 Land-Surface Model
Max Suarez	NASA	Mosaic Land Surface Model, etc.
Joel Susskind	NASA	TOVS/AIRS retrievals
Joe Waters	JPL	MLS retrievals

goals described in section 2 and cost is estimated primarily in terms of person labor and to a lesser extent computational cost (off-line and on-line computation, data storage and transfer). Factored into the *benefit* are the spatial and temporal coverage (*e.g.,* time-period and coverage for which data type is available). The *cost* estimate includes factors such as instrument team commitment, previous experience with a similar data type, an estimate of how much work is needed to produce the observation operator, known difficulties with calibration and systematic error, and an estimate of the amount of preprocessing needed prior to assimilation. The initial scores were translated into either a ranking of either high or low priority. These rankings are listed in following subsections along with relevant information about each data type considered. The prioritization is to be used as a guide for allocating resources and does not necessarily indicate a commitment by the DAO to assimilate or use a particular data type. The actual use of a data type in the GEOS system will depend on resources available (including instrument team commitments).

5.1 Priorities grouped by science topic

New data types are first grouped roughly by science topic. The groupings below are not listed in order of priority. For a given topic, all data types considered are listed. For each data type, the expected use in the GEOS system is also listed, *i.e.,* whether the data type will be used in the first look analysis, final platform, reanalysis only, pocket analysis, etc. (see Stobie, 1996, or da Silva *et al.*, 1996b, for a summary of operations). Although we found some data types to be strong candidates for assimilation, they may have been selected for validation on the basis of personnel considerations.

If available, a specific product name is listed. Data volumes are also listed. These were estimated from B. Bass (*private communication*, most recent estimate) when available or from values reported in NASA (1994). These are given in Gigabyte/day for level 2 products unless otherwise noted. Caution should be exercised in extrapolating from the figures listed here for data flow calculations. In some cases super-obbing will be performed to reduce data flow. For example, the data volume from super-obbed or level 3 MODIS data types will be more than a factor of 100 less than the data volume of the level 2 products. In some cases, data types are produced less than once per assimilation period (*e.g.,* land-surface products from MODIS to be used as boundary conditions produced once every 6, 15, or 30 days) and need not be re-transmitted every assimilation cycle. Estimates for data volume for level 3 products are not yet available.

The notations used in the tables are defined as follows:

1L: First Look Analysis (performed 12-24 hours after data time, used by EOS instrument teams for retrieval algorithms, etc.)

FP: Final Platform including reanalysis of final platform (runs several months after data time, includes data from first look as well as data from the EOS platform)

RA: Reanalysis (multi-year reprocessing of pre-EOS era and EOS-era using frozen system)

V: Validation

OA: Offline Analysis

B: Boundary Condition

PA: Pocket Analysis (similar to reanalysis, but for a limited time-period)

H: High priority

L: Low priority

S: Candidate for super-obbing/pre-processing at data site

E: Estimate

*past

‡future

†past, present, future

5.1.1 Temperature

There is an abundance of temperature data available for assimilation and validation. For both the first look system and final platform, TOVS was given the highest priority. The other measurements, with the exception of MLS, were selected for validation. Currently, in addition to conventional data, retrieved temperature (height) profiles from TOVS (from

NESDIS and GLA) are assimilated in the traditional manner. This approach has been implemented with covariance tuning as described in da Silva *et al.*(1996a) in GEOS-2.

The advanced methodologies (radiance assimilation and phase-space retrieval assimilation) described in section 3 are currently being implemented and evaluated with TOVS Pathfinder data in anticipation of future sounders such as AIRS and IASI. For the first look system as well as final platform and reanalysis with GEOS-3 and beyond, the most appropriate method will be selected on the basis of cost and quality of the analysis. The selected method will be implemented with NESDIS radiances/retrievals for the first look system. If a reliable alternate data source for TOVS cloud-cleared radiances and retrievals is available in near real-time for the first look system, a selection will be made on the basis of product quality after monitoring has been performed.

The strategy for implementing new methodologies with TOVS is as follows:

(1) A 1-D simulation of the PED approach using TOVS and AIRS radiances was completed in early 1996. The results showed that the data compression for both AIRS and TOVS should significantly reduce computation in PSAS. Based on the 1-D results, the theoretical foundation for this approach appears sound.

(2) A fast forward radiative transfer model was extracted from the TOVS Pathfinder code (Susskind *et al.*, 1983) and an analytic Jacobian (can be used as TLM and adjoint) was added. This model, called GLATOVS, is described in Sienkiewicz (1996). This model is currently being compared (in several respects) with RTTOVS.

(3) Systematic error correction to account for errors in the forward model and instrument calibration is an integral part of assimilating data from any microwave/IR sounder. In early 1996 a simulation of forward model error was completed. The results showed that, as expected, systematic error can be as large or larger detector noise. In the first part of 1996, several systematic error correction schemes have been compared in a simulation environment. Currently, a physically-based correction scheme is being implemented with TOVS Pathfinder data. After the algorithm has been tested, validated, and frozen, the results will be documented. The documentation should be completed in summer of 1996.

(3) After the systematic error correction scheme has been implemented, the radiance error covariance will be estimated using the approach described in da Silva *et al.*(1996a). This activity will take place in summer of 1996.

(4) Beginning in fall 1996, when observation operators have been implemented in PSAS, a comparison of the PED and the direct radiance assimilation approaches will be completed using TOVS Pathfinder data for one season. The results will also be compared with the current method of tuned retrieval assimilation. One approach will be selected on the basis of cost and analysis quality for use in future systems.

(5) The selected method will be implemented with NESDIS radiances in winter 1996-1997. The quality of the analysis with radiances from NESDIS and Pathfinder will be compared. If there is a significant difference in data quality, the Pathfinder data will be used for reanalysis.

Stratospheric temperature measurements from MLS were also given a high priority for assimilation as a result of a commitment from the MLS team. The availability of MLS temporal coverage makes it a candidate for pocket analysis. The first meeting with the MLS team will take place in late June of 1996. The project design will begin at that time. The project is expected to begin in late summer 1996 and should last approximately one year.

5.1.2 Moisture Assimilation

Conventional data has been the primary source of moisture data in the GEOS/DAS to date. Experiments with assimilating total precipitable water (TPW) from SSM/I at the DAO (Ledvina and Pfaendtner, 1995) have shown positive impact especially in the tropics where the GCM is known to have a dry bias. Both TOVS and SSM/I (and TMI) provide information about TPW. In addition, TOVS provides some additional information about

Table 3: Priorities for assimilating temperature data

Instrument	Satellite	Observable	Use/Priority	Vol (GB/day)
†Conventional	N/A	ANC_NOAA_RDSONDS_OBS	1L **H**	0.010
†TOVS, ATOVS	POES (NESDIS)	ANC_NESDIS_TOVS_THKN_RETS	1L **H**	0.030
†TOVS, ATOVS	POES (GLA)	Trop/Strat	RA **H**	0.2
†MLS	UARS	Strat	RA **H**	0.02
†GPS	GPSMET	Trop/Strat	V **H**	0.0E
†SSM/T	DMSP	Trop/Strat	V **H**	?
†GOES sounder	GOES	Trop	V **L**	?S
*LIMS	Nimbus 7	Strat	PA **H**	?
*CLAES	UARS	Strat	V **H**	0.02
‡MODIS	EOS AM	Trop (MOD30)	V **H**	11.2S
‡IMG	ADEOS	Trop/Strat	V **L**	?S

the profile and has coverage over both land and water. The capability to assimilate moisture information from both TOVS (as described above) and SSM/I will be in place for GEOS 3 and beyond. The final decision on which data type or combination thereof will be made on the basis of computational cost and analysis quality.

A collaborative effort is underway with scientists at JPL and DAO to use ground-based GPS data in the GEOS system. The actual assimilation of GPS data into the GEOS system will begin in approximately two years after initial studies have been completed. The number of GPS ground-based measurement systems is expected to increase dramatically in the future and initial studies show the TPW from these systems to be of high quality.

Several other instruments received high priority for validation. Although several of these data types are strong candidates for assimilation in the GEOS system (*e.g.,* MLS and SSMT/2), resources in addition to those currently in place would be required for this as these data types have little heritage.

5.1.3 Convective/Precip. Retrieval Assimilation

Developing the capability to assimilate convective data, including precipitation, is an ongoing research project at DAO. The goal is to have this capability ready in fall of 1996, and to compare this approach with the physical initialization approach used by Krishnamurti *et al.* (1991) in winter of 1996. The basic approach is to use a physical model (RAS) as an observation operator to perform a 1-D retrieval of water vapor, which is then assimilated in PSAS in the context of the CARD/PED approach described above. Currently, a simplified version of this algorithm has been coded and is undergoing tests. This approach does not assume that either the data or the model are perfect. The first implementation will use data inferred from SSM/I. TRMM data will be used when it becomes available. In the future, other data types including inferred cloud top temperature and OLR, will be assessed for assimilation in a similar manner. A summary of convective data types is given in table 5.

5.1.4 Land Surface

The DAO is currently developing the capability to assimilate land-surface data. At the present time, the Mosaic (Koster and Suarez, 1992) land-surface model (LSM) has just been coupled with the GEOS GCM and is undergoing tests. Currently, off-line tests with the Mosaic LSM are also being performed. The Mosaic LSM does not have the ability to accept satellite data in order to specify boundary conditions. It is planned that part of the Sib2 (*e.g.,* Sellers, *et al.*, 1986) LSM that accepts satellite data will be integrated

Table 4: Priorities for assimilating water vapor data

Instr.	Satellite	Observable	Use/Priority	Vol (GB/day)
†Conven.	N/A	ANC_NOAA_RDSONDS_OBS	1L **H**	0.01
†TOVS, ATOVS	POES (NESDIS)	ANC_NESDIS_TOVS_RETS	1L **H**	0.030
†TOVS	POES (GLA)	Trop	RA **H**	0.2
†SSM/I	DMSP	ANC_MSFC_SSMI_PRCP_WATER	1L **H**	0.050
†GPS	Ground	TPW	RA **H**	0.0
†SAGE	ERBS, etc.	Strat	V **H**	0.0
†MLS	UARS	Strat	V **H**	0.02
†SSM/T2	DMSP	Trop	V **H**	?
†GOES	GOES	Trop (prof)	V **L**	?S
*HALOE	UARS	Strat	V **H**	0.0
‡TMI	TRMM	TPW	FP **H**	?
‡MODIS	EOS AM	Trop (MOD30_L2)	V **H**	7.192S
‡IMG	ADEOS	Trop	V **L**	?S

Note: Volumes in GB/day

Table 5: Priorities for assimilating convective retrievals

Instrument	Satellite	Observable	Use/Priority	Vol (GB/day)
†SSM/I	DMSP	Prec rate, surf	FP **H**	0.04
†SSM/I	DMSP	Cloud liq H_2O	V **H**	0.04
†ISCCP	(POES)	Cloud Top, frac	V **H**	0.01
†Conven.	N/A	Cloud base	V **H**	0.0
†TOVS	POES (GLA)	Cloud top	V **H**	0.2
†Conven.	N/A	Radar Precip	V **H**	0.03
†MSU	POES (Path)	Precip ocean	V **L**	0.01
*ERBE	ERBS	TOA flux	V **H**	0.0
‡TMI	TRMM	(M69) TMI_PROF_L2A-12	FP **H**	1.314S
‡TMI,PR	TRMM	TRMM_COMB_L2B-31(M73)	V **H**	0.848S
‡PR	TRMM	PR Prof (M72)	V **H**	1.8S
‡MODIS	EOS AM	Cloud (MOD06)	V **H**	1.13S
‡CERES	TRMM	TOA, cloud (CER07)	V **H**	0.26S
‡CERES	EOS AM	TOA, cloud (CER07)	V **H**	0.26S
‡AMSU	POES	Cloud liq H_2O	V **L**	TBD

Table 6: Priorities for assimilating land-surface data

Instrum.	Satellite	Observable	Use/Priority	Vol (GB/day)
†Conven.	N/A	T_a, q_a ANC_NOAA_SFC_OBS	FP, RA **H**	0.05
†TOVS	POES (GLA)	T_s	FP, RA **H**	0.2
†SSM/I	DMSP	Snow Wat.Eq. (SWE)	FP, RA **H**	0.01
†SSM/I	DMSP	Surface Wetness	FP, RA **H**	0.01
†AVHRR	POES	ANC_NESDIS_NDVI	B, RA **H**	0.069S
*SMMR	Nimbus7	SWE	RA **H**	0.00
‡MODIS	EOS AM	T_s (MOD11_L2)	V (FP) **H**	6.376
‡MODIS	EOS AM	NDVI (MOD34_L3_10DY)	B, FP **H**	93.558GB/10day-S
‡MODIS	EOS AM	LAI, FPAR (MOD15_L4_10DY)	B, FP **H**	1.502GB/10day-S
‡MODIS	EOS AM	Evaptrans (MOD16_L3_10DY)	B, FP **H**	22.537GB/10day-S
‡MODIS	EOS AM	(MOD09_BRDF_L3_16DY)	B, FP **H**	108.16GB/16day-S
‡MODIS	EOS AM	Land type (MOD12_L3_32DY)	B, FP **H**	0.417/32day-S
‡MODIS	EOS AM	LS res (MOD41_L2)	B, FP **H**	23.652/day???
‡ASTER	EOS AM	T_s (on demand only)	V **L**	?S

with the Mosaic model in GEOS. When this is accomplished, satellite derived products such as the Normalized Differential Vegetation Index (NDVI) and others listed in Table 6 from AVHRR and MODIS will be ingested. In addition to surface station observations (temperature T_a and humidity q_a), we are investigating the feasibility of assimilating surface (skin) temperature (T_s) derived from IR instruments, as well as snow water equivalent (SWE) and surface wetness inferred from passive microwave instruments. It is expected that some combination of these data types will be used in the final platform system.

5.1.5 Ocean Surface

Ocean wind data from both passive microwave (SSM/I and SMMR) and scatterometers (ERS-1 and NSCAT) were identified as high priority data types for the 1998 system (both first look and final platform). Table 7 lists the details of ocean surface data types considered. For the 1998 system, the capability of assimilating data from both passive microwave and scatterometers will be in place. However, before one or more of these data types are used, the data will have been monitored and corrected if necessary for systematic errors (including observation operator error). The specific plans for use of NSCAT data are outlined below.

Starting in December 1996, the Synoptic Evaluation Group will start receiving NSCAT science products, which includes backscatter measurements, (Level 1.7 data), retrieved ranked wind ambiguities (Level 2.0 data), as well as gridded wind fields (Level 3.0 data). The operational release of the data is scheduled to start about March 1997, although the data format will change. The initial plan is to assimilate NSCAT retrieved wind velocity and possibly employ an ambiguity removal algorithm.

The process of assimilating surface wind data includes the following steps:

(1) A process, known as MOVEZ, is used to create a suitable first guess wind field by a reduction process of the lowest model level wind field to 10m, consistent with the GEOS PBL. The multivariate wind and sea level pressure analysis is performed. Subsequently, the analyzed wind field is extended up to the lowest vertical model level and the analysis increments are computed. The analysis increments of the upper air wind field at 850 hPa and the lowest model level are then interpolated linearly in log (P) to the vertical model levels in between. This work will begin approximately 4 weeks after the GEOS-2 (with PSAS) is frozen.

Table 7: Priorities for assimilating ocean-surface data

Instrument	Satellite	Observable	Use/Priority	Vol (GB/day)
†Conventional	N/A	u, v, T_{air}, q_s	1L **H**	0.0
†SSM/I	DMSP	Wind speed	1L, RA **H**	0.01
†ERS-1,2	ERS-1,2	speed, dir, σ^o	1L, FP **H**	0.20E
*SMMR	Nimbus-7	speed	V **L**	0.01
‡NSCAT	ADEOS	ANC_NESDIS_NSCAT_WNDPDT	1L **H**	0.046
‡NSCAT	ADEOS	ANC_JPL_NSCAT_WNDPDT	FP **H**	0.046

(2) In order to maximize the influence of the surface wind field in the GEOS DAS, the NSCAT wind vectors are also extended by a process analogous to MOVEZ to the model lowest vertical level. The observation innovations are computed and used by the multivariate mass - wind analysis. A check is made of the atmospheric stability, and only the observations under the unstable or neutral conditions are used. The 3-D global analysis that makes use of the data and propagates the information in the vertical according to the background error vertical correlation function. This particular process has experimentally been applied to ERS-1 scatterometer data within the GEOS-1 DAS environment and presented in Tokyo in 1995. It resulted in modest improvement of the resultant forecast experiments. Within the framework of PSAS it is expected to result in a more significant impact of scatterometer data due to a truly global three dimentional design of PSAS. This step will begin approximately 4 weeks after step (1).

(3) A simple surface wind ambiguity removal algorithm is in place in GEOS-1 DAS. It will be tested with the new GEOS DAS. The algorithm compares the directions of available ambiguous wind vectors with the background field and chooses the one closest to it. In the future an interactive procedure or a 2D Variational ambiguity removal algorithm might replace the current one. If done, this work would begin in Jan or Feb of 1997.

5.1.6 Constituents

The plans for constituent assimilation will be given at a later time. The data types considered for ozone assimilation are given in table 8. The data types considered for CO assimilation are given in table 9.

5.1.7 Wind profile

Currently, wind profile data used in GEOS are obtained from radiosondes and derived cloud track winds. Several of the data types listed in table 10 appear to be strong candidates for assimilation (*e.g.,* water vapor winds similar to cloud track winds). These new data types can enhance the coverage of current wind profile data. However, because there is little heritage for assimilating these data types, they are presently selected for validation. As resources permit, some of these data types may eventually be assimilated.

5.1.8 Aerosols

Although consideration has been given to aerosol assimilation (Ménard, personal notes), this topic has been given an overall low priority at present. Therefore, all data types listed in table 11 were selected for validation.

Table 8: Priorities for assimilating ozone data

Instrument	Satellite	Observable	Use/Priority	Vol (GB/day)
†TOMS	N7/Meteor/ADEOS	total	OA **H**	0.01
†SBUV	N7/POES	prof	V (OA) **H**	0.001
†Conven.	N/A	prof	V **H**	0.0
†SAGE	ERBS, etc.	strat prof	V **H**	0.0
†MLS	UARS	upper strat prof	V **H**	0.02
†GOME	ERS-2	total, prof	V **L**	?
†TOVS	POES (GLA)	lower strat	V **L**	0.2
*CLAES	UARS	upper strat prof	V **L**	0.02
*HALOE	UARS	upper strat prof	V **L**	0.02
‡ILAS	ADEOS	upper strat prof	V **L**	?
‡IMG	ADEOS	lower strat prof	V **L**	?

Table 9: Priorities for assimilating CO data

Instrument	Satellite	Observable	Use/Priority	Vol (GB/day)
*MAPS	Shuttle	prof	OA **H**	0.0
‡MOPITT	EOS AM	prof	OA **H**	TBD
†Conven.	N/A	prof	V **H**	0.0

Table 10: Priorities for assimilating wind profiles

Instrument	Satellite	Observable	Use/Priority	Vol (GB/day)
†Conventional	N/A	ANC_NOAA_RDSONDS_OBS	1L **H**	0.01
†Conventional	N/A	ACARS_ANC_NOAA_ARCFT_OBS	1L **H**	0.01
†cloud track	GOES	ANC_NESDIS_GOES_WIND_MOTION	1L **H**	0.008
†RADAR	N/A	trop/strat/mes	V**H**	?
water vapor	GOES/other	trop	V **H**	?
*HRDI	UARS	strat/mes	V **L**	0.003
*WINDII	UARS	mes	V **L**	0.03
*MODE	Shuttle		V **L**	?
‡SWIPE	TBD			

Table 11: Priorities for use of aerosol data

Instrument	Satellite	Observable	Use/Priority	Vol (GB/day)
†SAGE	ERBS, etc.	Strat	V **L**	0.0
†TOMS	Nimbus 7	Strat/Trop	V **L**	0.01
*CLAES	UARS	Strat	V **L**	0.02
*HALOE	UARS	Strat	V **L**	0.02
‡MISR	EOS-AM	properties	V **L**	?
‡ILAS	ADEOS	Strat	V **L**	?

Table 12: High-priority data types from POES satellite

Instrument	Observable	Use
†TOVS	Temp. Prof/rad	1L
†TOVS	H_2O Prof/rad	1L
†TOVS	T_s	FP, RA
†AVHRR	ANC_NESDIS_NDVI	B, RA

Table 13: High-priority data types from UARS satellite

Instrument	Observable	Use
†MLS	Strat T	RA
†MLS	Upper Trop/Strat H_2O	V
†MLS	upper strat O_3	V
*CLAES	Strat T	V
*HALOE	Strat H_2O	V
*HALOE	upper strat prof O_3	V

5.2 Priorities grouped by satellite

This section categorizes high priority items by satellite. The priorities for the POES, UARS, TRMM, ADEOS, and EOS-AM1 satellites are listed in tables 12-16, respectively.

5.3 Priorities grouped by use in GEOS

This section categorizes high priority items according to use in GEOS. Data types are grouped as either used for first look, final platform (data types in addition to those used for the first look), or reanalysis/pocket analysis (may include data types used in the first look and final platform) in tables 17-19, respectively. Some of the high-risk data types (such as land-surface data) have not been included here. Data types used for validation are not listed here.

Table 14: High-priority data types from TRMM satellite

Instrument	Observable	Use
‡TMI	PR Prof (M69)	FP
‡TMI,PR	PR Prof (M73)	V
‡PR	PR Prof (M72)	V
‡CERES	cloud (CER07)	V

Table 15: High-priority data types from the ADEOS satellite

Instrument	Observable	Use
‡NSCAT	wind speed, dir, σ^o	FP
†TOMS	total O_3	OA

Table 16: High-priority data types from EOS AM1 satellite

Instrument	Observable	Use
‡MODIS	NDVI	B, FP
‡MODIS	LAI, FPAR	B, FP
‡MODIS	Evaptrans	B, FP
‡MODIS	BidirRefl	B, FP
‡MODIS	Land type	B, FP
‡MODIS	LS res	B, FP
‡MODIS	Cloud	V
‡MODIS	T_s	V
‡MODIS	T,q Prof (MOD30)	V
‡MOPITT	CO Prof	OA
‡CERES	TOA, cloud (CER07)	V

Table 17: High-priority data types for first look system

Instrument	Platform	Observable	Volume
†Conventional	N/A	ANC_NOAA_RDSONDS_OBS	0.010
†TOVS, ATOVS	POES (NESDIS)	ANC_NESDIS_TOVS_THKN_RETS (T, q, radiance)	0.030
†Conventional	ship, buoy	u, v, T_{air}, q_s	0.0
†SSM/I	DMSP	ANC_MSFC_SSMI_PRCP_WATER	0.050
†SSM/I	DMSP	Wind speed	0.01
†ERS-1	ERS-1	speed, dir	0.20E
‡NSCAT	ADEOS	ANC_NESDIS_NSCAT_WNDPDT	0.046
†Conventional	N/A	ACARS_ANC_NOAA_ARCFT_OBS	0.01
†cloud track	GOES	ANC_NESDIS_GOES_WIND_MOTION	0.008

Table 18: High-priority data types for final platform.

Instrument	Platform	Observable	Volume
†SSM/I	DMSP	Prec rate, surf	0.04
‡TMI	TRMM	(M69) TMI_PROF_L2A-12	1.314S
‡NSCAT	ADEOS	ANC_JPL_NSCAT_WNDPDT	0.046
‡MODIS	EOS AM	NDVI (MOD34_L3_10DY)	93.558GB/10day-S
‡MODIS	EOS AM	LAI, FPAR (MOD15_L4_10DY)	1.502GB/10day-S
‡MODIS	EOS AM	Evaptrans (MOD16_L3_10DY)	22.537GB/10day-S
‡MODIS	EOS AM	(MOD09_BRDF_L3_16DY)	108.16GB/16day-S
‡MODIS	EOS AM	Land type (MOD12_L3_32DY)	0.417/32day-S
‡MODIS	EOS AM	LS res (MOD41_L2)	23.652/day???

Table 19: High-priority data types for reanalysis and/or pocket analysis.

Instrument	Platform	Observable	Volume
†MLS	UARS	Strat T	0.02
†TOVS	POES (GLA)	T, q	0.2

6 References

Andersson, E., Pailleux, J., Thépaut, J. N., Eyre, J. R., McNally, A. P., Kelly, G. A., and P. Courtier, 1994: Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **120**, 627-653.

Daley, R., 1995: Estimating the wind field from chemical constituent observations: Experiments with a one-dimensional extended Kalman filter. *Mon. Wea. Rev.*, **123**, 181-198.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123**, 1128-1145.

Dee, D. P., and A. Trenholme, 1996: DAO QC Strategy Document, *DAO Office Note* 96-15, Data Assimilation Office, Goddard Space Flight Center, Greenbelt, MD 20771 .

Eyre, J. R., 1992: A bias correction scheme for simulated TOVS brightness temperatures. ECMWF Tech. Memo 176.

Guo, J. and A. M. da Silva, 1995: Computational Aspects of Goddard's Physical-space Statistical Analysis System (PSAS). *Preprints, Second UNAM-CRAY Supercomputing Conf. on Numerical Simulations in the Environmental and Earth Sciences.* Mexico City, Mexico, 6pp.

Joiner, J. and A. M. da Silva, 1996: Efficient methods to Assimilate Satellite Retrievals Based on Information Content, submitted to *Mon. Weath. Rev.*

Koster, R. D., and M. J. Suarez, 1992: Modeling the land surface boundary in climate models as a composite of independent vegetation stands, *J. Geophys. Res.* **97**, 2697-2715.

Krishnamurti, T. N., Xue, J., Bedi, H.S., K. Ingles, and D. Oosterhof, 1991: Physical initialization for numerical weather prediction over the tropics. *Tellus*, **43AB**, 53-81.

Ledvina, D. V., and J. Pfaendtner, 1995: Inclusion of Special Sensor Microwave/Imager (SSM/I) total precipitable water estimates into the GEOS-1 Data Assimilation System, *Mon. Weath. Rev.,* **123**, 3003-3015.

Ménard, R., 1995: Middle atmosphere assimilation of UARS constituent data using Kalman filtering: Preliminary results. *International symposium on assimilation of observations in meteorology and oceanography*, Tokyo, Japan, 273-278.

Ménard, R., 1996: Consistent assimilation of retrieved data (CARD), private notes and manuscript in preparation.

Molod, A., H. M. Helfand, and L. L. Takacs, 1995: The climatology of parameterized physical processes in the GEOS-! GCM and their impact on the GEOS-1 Data Assimilation System, *J. Climate.*

NASA/Goddard Space Flight Center, 1994; *Science Data Plan for the EOS Data and Information System covering EOSDIS Version 0 and Beyond*, Document version 4, Ed. M. Schwaller and B. Krupp.

Riishøjgaard, L. P., 1996: On four-dimensional variational assimilation of ozone data in weather prediction models. *(accepted for publication in the Q. J. R. Meteorol. Soc.)*

Rodgers, C. D., 1990: Characterization and error analysis of profiles retrieved from remote sounding measurements. *J. Geophys. Res.* **95**, 5587-5595.

Pfaendtner, J., S. Bloom, D. Lamich, M. Seablom, M. Sienkiewicz, J. Stobie, A. da Silva, 1995: Documentation of the Goddard Earth Observing System (GEOS) Data Assimilation System–Version 1. *NASA Tech. Memo. No. 104606*, Vol. **4**, Goddard Space Flight Center, Greenbelt, MD 20771.

Schubert, S.D., R. B. Rood, and J. Pfaendtner, 1993: An assimilated data set for earth science applications. Bul. Amer. Meteor. Soc., 74, 2331-2342.

Schubert, S., C.-K. Park, Chung-Yu Wu, W. Higgins, Y. Kondratyeva, A. Molod, L. Tkacs, M. Seablom, and R. Rood, 1995: A Multiyear Assimilation with the GEOS-1 System: Overview and Results. NASA Tech. Memo. 104606, Vol. 6. Goddard Space Flight Center, Greenbelt, MD 20771.

Schubert, S. D., 1996: DAO's Quality Assessment Team: Definition and Planing Document. *DAO Office Note 96-??.*. Data Assimilation Office, Goddard Space Flight Center, Greenbelt, MD 20771.

Schubert, S. D. and R. Rood, 1995: Proceedings of the Workshop on the GEOS-1 Five-Year Assimilation. NASA Tech. Memo. 104606, Vol. 7, Goddard Space Flight Center, Greenbelt, MD 20771.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher, 1986: A simple boisphere model (SiB) for use within general circulation models, *J. Atmos. Sci.*, **43**, 505-531.

Sienkiewicz, M., 1996: The GLA TOVS rapid algorithm forward radiance modules and Jacobian version 1.0, *DAO Office Note* 96-08, Data Assimilation Office, Goddard Space Flight Center, Greenbelt, MD 20771 .

da Silva, A. M., Pfaendtner, J., Guo, J., Sienkiewicz, M., and S. E. Cohn, 1995: Assessing the effects of data selection with DAO's Physical-space Statistical Analysis System. *International symposium on assimilation of observations in meteorology and oceanography*, Tokyo, Japan, 273-278.

da Silva, A., and J. Guo, 1996: Documentation of the Physical-Space Statistical Analysis System (PSAS) Part I: The Conjugate Gradiant Solver version PSAS-1.00 *DAO Office Note* 96-02, Data Assimilation Office, Goddard Space Flight Center, Greenbelt, MD 20771 .

da Silva, A. M., Redder, C., and D. P. Dee, 1996a: Modeling retrieval error covariances for global data assimilation. *Proc. 8^{th} Conf. Satell. Meteor.*, Atlanta, GA., in press.

da Silva, A., Ekers, K., and A. Conaty, 1996b: Requirements for DAO's On-Line Monitoring System (DOLMS) version 1.00, *DAO Office Note* 96-14, Data Assimilation Office, Goddard Space Flight Center, Greenbelt, MD 20771 .

Stobie, J, 1996: Data Assimilation Computing and Mass Storage Requirements for 1998. *DAO Office Note 96-16.*. Data Assimilation Office, Goddard Space Flight Center, Greenbelt, MD 20771.

Stoffelen A., and D. Anderson, 1995: The ECMWF contribution to the characterisation, interpretation, calibration, and validation of ERS-1 scatterometer backscatter measurements and winds, and their use in numerical weather prediction models, *European Space Agency Contract Report.*

Susskind, J., Rosenfield, J., and D. Reuter, 1983: An accurate radiative transfer model for use in the direct physical inversion of HIRS and MSU temperature sounding data. *J. Geophys. Res.* **88**, 8550-8568.

Susskind, J., and J. Pfaendtner, 1989: Impact of iterative physical retrievals on NWP. Report on the Joint ECMWF/EUETSAT Workshop on the Use of Satellite Data in operational weather prediction: 1989-1993. Vol. 1, 245-270. T. Hollingsworth, Ed.

Takacs, L.L., A. Molod, and T. Wang, 1994: Documentation of the Goddard Earth Observing System (GEOS) General Circulation Model–Version 1. NASA Tech. Memo. No. 104606, volume 1, Goddard Space Flight Center, Greenbelt, MD 20771.

7 Appendix

7.1 Appendix A: Acronyms

CARD: Consistent Assimilation of Retrieved Data

CTW: Cloud track winds

DAO: Data Assimilation Office (NASA Goddard Space Flight Center Laboratory for Atmospheres)

DAS: Data Assimilation System

EOSDIS: Earth Observing System Data and Information System

GCM: General Circulation Model

GEOS: Goddard Earth Observing System

EKF: Extended Kalman Filter

LSM: Land-Surface Model

MTPE: Mission to Planet Earth

NDVI: Normalized Differential Vegetation Index

NWP: Numerical Weather Prediction

PSAS: Physical-space Statistical Analysis System

PV: Potential Vorticity

SWE: Snow Water Equivalent

TPW: Total precipitable water

7.2 Appendix B: Instruments and Satellites

ADEOS: Advanced Earth Observting Satellite (Mid-late 1996)

AIRS: Atmospheric Infrared Sounder (EOS PM)

AMSU A-B: Advanced Microwave Sounding Unit (POES, EOS PM)

AVHRR: Advanced Very High-Resolution Radiometer

ASTER: Advanced Spaceborne Thermal Emission and Reflection Radiometer (EOS AM)

ATOVS: Advanced TOVS; HIRS3/AMSU (POES)

CERES: Clouds and Earth's Radiation Energy System (TRMM, EOS AM)

CLAES: Cryogenic Limb Array Etalon Spectrometer (UARS)

DMSP: Defense Meteorological Satellite Program (currently operational)

EOS AM1: Earth Observting Satellite AM (June 98 launch)

EPS: EUMETSAT (European Meteorology Satellite) Polar System

ERBE: Earth Radiation Budget Experiment (ERBS)

ERBS: Earth Radiation Budget Satellite

ERS-1,2: European Remote Sensing Satellite (Scatterometer, 6 channel IR-Visible radiometer)

GOES: Geostationary Observational Environmental Satellite (Imager and 18 channel visible and infrared sounder, currently operational)

GOME: Global Ozone Monitoring Experiment (ERS-2)

GPS: Global Positioning System

HALOE: Halogen Occultation Experiment (UARS)

HIRS2/3: High-Resolution InfraRed Sounder (POES)

HRDI: High Resolution Doppler Imager

IASI: Infrared Atmospheric Sounding Interferometer (EPS)

ILAS: Improved Limb Atmospheric Spectrometer (ADEOS)

IMG: Interferometric Monitor for Greenhouse Gases (ADEOS)

ISCCP: International Satellite Cloud Climatology Project (several IR and visible instruments aboard different satellite)

LIMS: Limb Infrared Monitor of the Stratosphere (Nimbus 7)

MAPS: Measurement of Atmospheric Pollution from Satellites (?)

MHS: Microwave Humidity Sounder (EOS-PM)

MLS: Microwave Limb Sounder (UARS)

MODIS: Moderate-Resolution Imaging Spectrometer (EOS AM)

MSU: Microwave Sounding Unit

MOPITT: Measurement of Pollution in the Troposphere (EOS AM)

NSCAT: NASA Scatterometer (ADEOS)

POES: Polar Orbiting Environmental Satellite (Currently Operational)

PR: Precipitation Radar (TRMM)

SBUV: Satellite Backscatter Ultraviolet radiometer (Nimbus 7, POES)

SAGE: Stratospheric Aerosol and Gas Experiment (ERBS)

SMMR: Scanning Multispectral Microwave Radiometer (?)

SSM/I: Special Sensor Microwave/Imager (DMSP)

SSM/T: Special Sensor Microwave (Temperature sounder) (DMSP)

SSM/T2: Special Sensor Microwave (Water vapor sounder) (DMSP)

SSU: Stratospheric Sounding Unit (POES)

TMI: TRMM Microwave Imager (TRMM)

TOMS: Total Ozone Mapping Spectrometer (ADEOS, Meteor, Earth Probe, Nimbus 7)

TOVS: TIROS Operational Vertical Sounder; HIRS2/MSU/SSU (POES)

TRMM: Tropical Rainfall Measuring Mission (summer '97 launch)

UARS: Upper Atmospheric Research Satellite (some instruments in operation)

WINDII: Wind Imaging Interferometer (UARS)